

Integrated mass spectrometry imaging and single-cell transcriptome atlas strategies provide novel insights into taxoid biosynthesis and transport in *Taxus mairei* stems

Chunna Yu^{1,2}, Kailin Hou^{1,2}, Hongshan Zhang^{1,2,3}, Xueshuang Liang^{1,2}, Cheng Chen⁴, Zhijing Wang⁴, Qicong Wu^{1,2}, Ganlin Chen^{1,2}, Jiaxu He^{1,2}, Enhui Bai^{1,2}, Xinfen Li^{1,2}, Tingrui Du^{1,2}, Yifan Wang^{1,2}, Mingshuang Wang^{1,2}, Shangguo Feng^{1,2}, Huizhong Wang^{1,2,*} and Chenjia Shen^{1,2,3,*} 

¹College of Life and Environmental Sciences, Hangzhou Normal University, Hangzhou 311121, China,

²Zhejiang Provincial Key Laboratory for Genetic Improvement and Quality Control of Medicinal Plants, Hangzhou Normal University, Hangzhou 311121, China,

³Kharkiv Institute, Hangzhou Normal University, Hangzhou 311121, China, and

⁴College of Pharmacy, Hangzhou Normal University, Hangzhou 311121, China

Received 3 March 2023; revised 30 April 2023; accepted 18 May 2023.

*For correspondence (e-mail shencj@hznu.edu.cn; whz62@163.com)

SUMMARY

Taxol, which is a widely used important chemotherapeutic agent, was originally isolated from *Taxus* stem barks. However, little is known about the precise distribution of taxoids and the transcriptional regulation of taxoid biosynthesis across *Taxus* stems. Here, we used MALDI-IMS analysis to visualize the taxoid distribution across *Taxus mairei* stems and single-cell RNA sequencing to generate expression profiles. A single-cell *T. mairei* stem atlas was created, providing a spatial distribution pattern of *Taxus* stem cells. Cells were reordered using a main developmental pseudotime trajectory which provided temporal distribution patterns in *Taxus* stem cells. Most known taxol biosynthesis-related genes were primarily expressed in epidermal, endodermal, and xylem parenchyma cells, which caused an uneven taxoid distribution across *T. mairei* stems. We developed a single-cell strategy to screen novel transcription factors (TFs) involved in taxol biosynthesis regulation. Several TF genes, such as endodermal cell-specific *MYB47* and xylem parenchyma cell-specific *NAC2* and *bHLH68*, were implicated as potential regulators of taxol biosynthesis. Furthermore, an ATP-binding cassette family transporter gene, *ABCG2*, was proposed as a potential taxoid transporter candidate. In summary, we generated a single-cell *Taxus* stem metabolic atlas and identified molecular mechanisms underpinning the cell-specific transcriptional regulation of the taxol biosynthesis pathway.

Keywords: scRNA-seq, Matrix-assisted laser desorption/ionization imaging mass spectrometry analysis, taxol biosynthesis, *Taxus* stem, transcription regulation, transporter.

INTRODUCTION

Taxus wallichiana var. *mairei* (*T. mairei*) is a well-known gymnosperm widely distributed across South China with distinctive medicinal qualities (Hu et al., 2022). Several bioactive ingredients, including taxoids, have been extracted and isolated from the *T. mairei* stem bark (Chen et al., 2020). Taxol is a highly effective anticancer drug and is a complex and structural taxoid derived from *Taxus* trees (Hao et al., 2017). In recent decades, as market demands for taxol have increased, forest destruction has seriously impacted wild *Taxus* resources (Feng et al., 2023; Sanchez-Munoz et al., 2018).

While the taxol biosynthesis pathway and its regulatory mechanisms have been extensively studied, they require

more characterization. Dozens of enzymes with key roles in taxol biosynthesis have been identified, including taxadiene synthase (TS), taxadien-5 α -ol *O*-acetyltransferase (TAT), taxane 2 α -*O*-benzoyl-transferase (TBT), 10-deacetylbaicatin III 10-*O*-acetyltransferase (DBAT), and 3'-*N*-debenzoyl-2'-deoxytaxol-*N*-benzoyltransferase (DBTNBT) (Yu et al., 2021). Also, several cytochrome P450 (CYP) family members, including 2 α -, 5 α -, 7 β -, 9 α -, 10 β -, 13 α -, and 14 β -hydroxylases, have been shown to chemically modify the taxane skeleton (Kaspera & Croteau, 2006). Moreover, several transcription factors (TFs) are reportedly involved in the regulation of taxol biosynthesis (Kuang et al., 2019). In *Taxus media*, the MYB39-bHLH13 TF complex has key roles in the sexual dimorphic accumulation of taxol (Yu et al., 2022). In *Taxus*

chinensis, MYB29a is an activator of abscisic acid (ABA)-mediated taxol biosynthesis and WRKY33 is a repressor of salicylic acid (SA)-mediated taxol biosynthesis (Cao et al., 2022; Chen, Zhang, et al., 2021). While progress has been made in the TF-mediated regulation of taxol biosynthesis, the regulatory networks underpinning taxol biosynthesis require more investigation.

Taxol was first isolated from the stem bark of the Pacific yew (*Taxus brevifolia*) and subsequently detected in all *Taxus* species (Hu et al., 2022). In plants, the stem is the major accumulation site of medicinal compounds, e.g., several unique active constituents, such as vijayayosin, pterosupin, marsupsin, and pterostilbene, accumulate in the stem bark and heartwood of *Pterocarpus marsupium* (Ahmad et al., 2022). Ellagitannins and triterpenoids, which are used in traditional medicine, significantly accumulate in the stem bark of *Anogeissus leiocarpus* DC (Akanke et al., 2022). The inner stem bark decoction preparation from *Cariniana rubra* Gardner ex Miers is used to treat inflammatory diseases in folk medicine (Silva et al., 2021). Several novel amide alkaloids and carbazole alkaloids, such as claulansines A–J, were detected and isolated from *Clausena lansium* stems (Sun et al., 2021).

Although stems are the main taxol extraction sites, taxoids are unequally distributed across stem tissues (Fett Neto & DiCosmo, 1992). In 1-year-old twigs from the European yew (*Taxus baccata*), baccatin IV and taxine B primarily accumulate in the cortex, taxinine M is mainly distributed in the pith and cortex, and taxuspine W mainly accumulates in vascular tissue (Arendowski & Ruman, 2017). In *T. media*, taxol is primarily localized to outer stem layers, including the phloem and dead bark (Soliman & Raizada, 2020). We previously showed that 10-deacetylbaccatin III (10-DAB) was highly accumulated in the phloem, baccatin III (BAC) in the pith, and 10-deacetylpaclitaxel (10-DAP) in the cortex and phloem (Yu et al., 2020). Taxol concentrations and associated intermediates are believed to be correlated with the expression levels of associated biosynthesis genes (Bamnashin et al., 2022). However, little is known about the taxoid distribution and the expression patterns of taxol biosynthesis genes across distinct stem tissues.

The identification of chemicals in plant samples using matrix-assisted laser desorption/ionization-mass spectrometry (MALDI-IMS) is a newly developed MS-based metabolomics approach (Xiang et al., 2022). Single-cell RNA sequencing (scRNA-seq) is a high-resolution technology used to investigate cell heterogeneity and identify different cell types (Bawa et al., 2022) that has been used to elucidate several biological phenomena in plants. Time-series scRNA-seq approaches have provided insights into circadian clock mechanisms in *Arabidopsis* (Torii et al., 2022). Leaf cell-specific scRNA-seq identified essential palisade layer functions in UV-B protection mechanisms (Procko et al., 2022). The technology was instrumental in identifying

several organ development-related regulators, including epidermis development regulators, TFs regulating mesophyll cell development, regulators involved in the developmental trajectories of floret and inflorescence meristems, regulators of wood formation processes in poplar, and novel regulators required for the early development of vein patterns in cotyledons (Chen, Tong, et al., 2021; Liu, Guo, et al., 2022; Liu, Wang, et al., 2022; Tao et al., 2022; Zong et al., 2022). Similarly, single-cell transcriptome atlases have been used to identify novel catechin ester metabolic pathways in tea trees and provided novel scRNA-seq applications to study plant secondary metabolism (Wang et al., 2022).

To clarify taxol biosynthesis cell types, mass spectrometry (MS) imaging integrated with scRNA-seq was used to generate metabolic profiles of *T. mairei* stems at the single-cell level. Screening cell-specific TFs involved in taxol biosynthesis and potential secondary metabolite transporters will provide a better understanding of the regulatory networks underlying taxol biosynthesis and transport. Our study provides insights into the regulation of secondary metabolism in medicinal plants.

RESULTS

Overview of MS imaging data

To detect different primary and secondary metabolites in *T. mairei* stems, MALDI-IMS analysis was conducted (Figure 1a). To understand patterns in stems, segmentation analysis was performed, which generated differentially colored maps with an average of 6476 data points (three repetitions, Table S1). Maps clearly showed that most metabolites accumulated in tissue-specific patterns (Figure 1b).

Principal component analysis (PCA) separated all the detected metabolites into six principal components (PCs), referring to different tissues of the stem, including companion cell & phloem (PC1), pith (PC2), endodermis (PC3), xylem (PC4), epidermis (PC5), and cork (PC6) (Figure 1c). PCA loading data are shown in a heatmap (Figure 1d). To identify PC-specific metabolites, clustering analyses grouped all MS features into six clusters; Cluster I contained 2444 features which accumulated at very low levels in all PCs; Cluster VI contained 1345 features which accumulated at high levels in all PCs; Cluster III contained 985 features; PC5 contained specifically accumulated features; and Cluster II contained 1093 features, which highly accumulated in PC2 and PC5 (Figure 1e).

Visualization of taxoid accumulation in *T. mairei* stems

Taxoids are major bioactive components in *T. mairei* stems (Yu et al., 2020). Nine taxoids were detected by MALDI-IMS analysis, including four taxol biosynthesis-related taxoids and five taxol analogs. For taxol pathway-related metabolites, 10-deacetyl paclitaxel was evenly distributed across

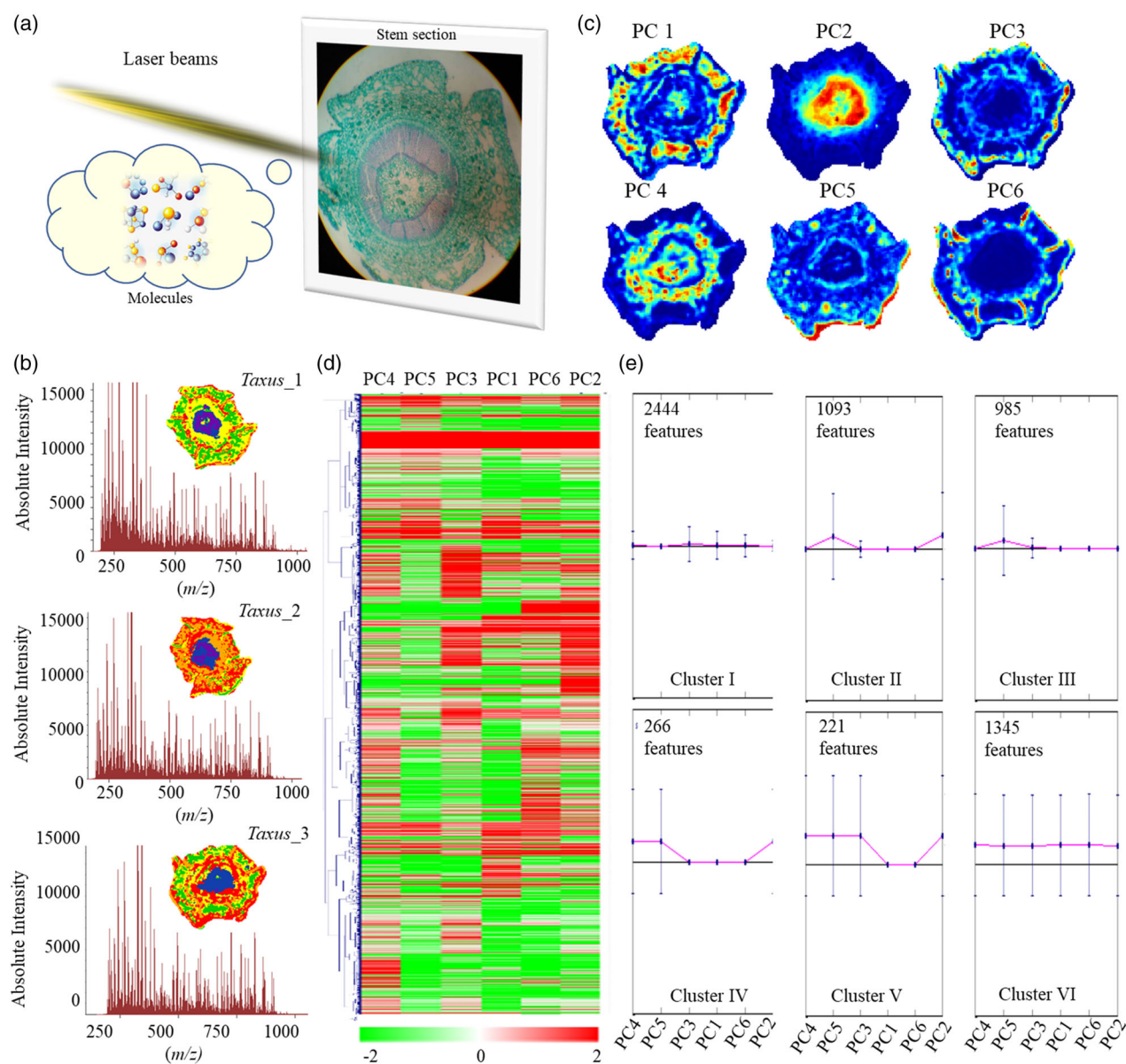


Figure 1. Overview of the MS imaging data. (a) Sampling and sectioning strategy for MS imaging. (b) Three independent segmentation analyses were performed, outputting differentially colored maps with on average 6476 data points. (c) PCA of all the detected metabolites in *T. mairei* stems. Six PCs indicated different accumulation patterns of the detected metabolites. (d) Heatmap showing the relative accumulation levels of each detected metabolite in different PCs. The heatmap scale ranges from -2 to $+2$ on a $\log_2$ scale. (e) Clustering analysis grouped all the detected MS features into six clusters.

the whole stem, 10-DAB and BAC were mainly distributed in the endodermis and phloem, and taxol specifically accumulated in the epidermis. For other taxanes, 2'-desacetoxy austrospicatin highly accumulated in the pith and xylem, 7-xylotaxol B and cephalomannine accumulated in the epidermis, and 10,13-deacetyl-abeo-baccatin IV and 10-deacetyltaxol C accumulated in the endodermis and phloem (Figure 2a). Quantitative analyses indicated taxol pathway metabolite and taxol analog proportions in different PCs, confirming their tissue-specific accumulation in *T. mairei* stems (Figure 2b,c). To identify the genetic

mechanisms underlying this uneven distribution in *T. mairei* stems, single-cell transcriptomic analyses were performed.

scRNA-seq of *T. mairei* stems

Approximately 1.09×10^7 protoplasts mL^{-1} were isolated from independent stem samples. The basic protoplast isolation process, from stems to single cells, is shown in Figure 3a. Purified protoplasts were loaded onto a commercial 10 $\times$ Genomics platform for scRNA-seq (Figure 3b). For single-cell sequencing, cells were marked with

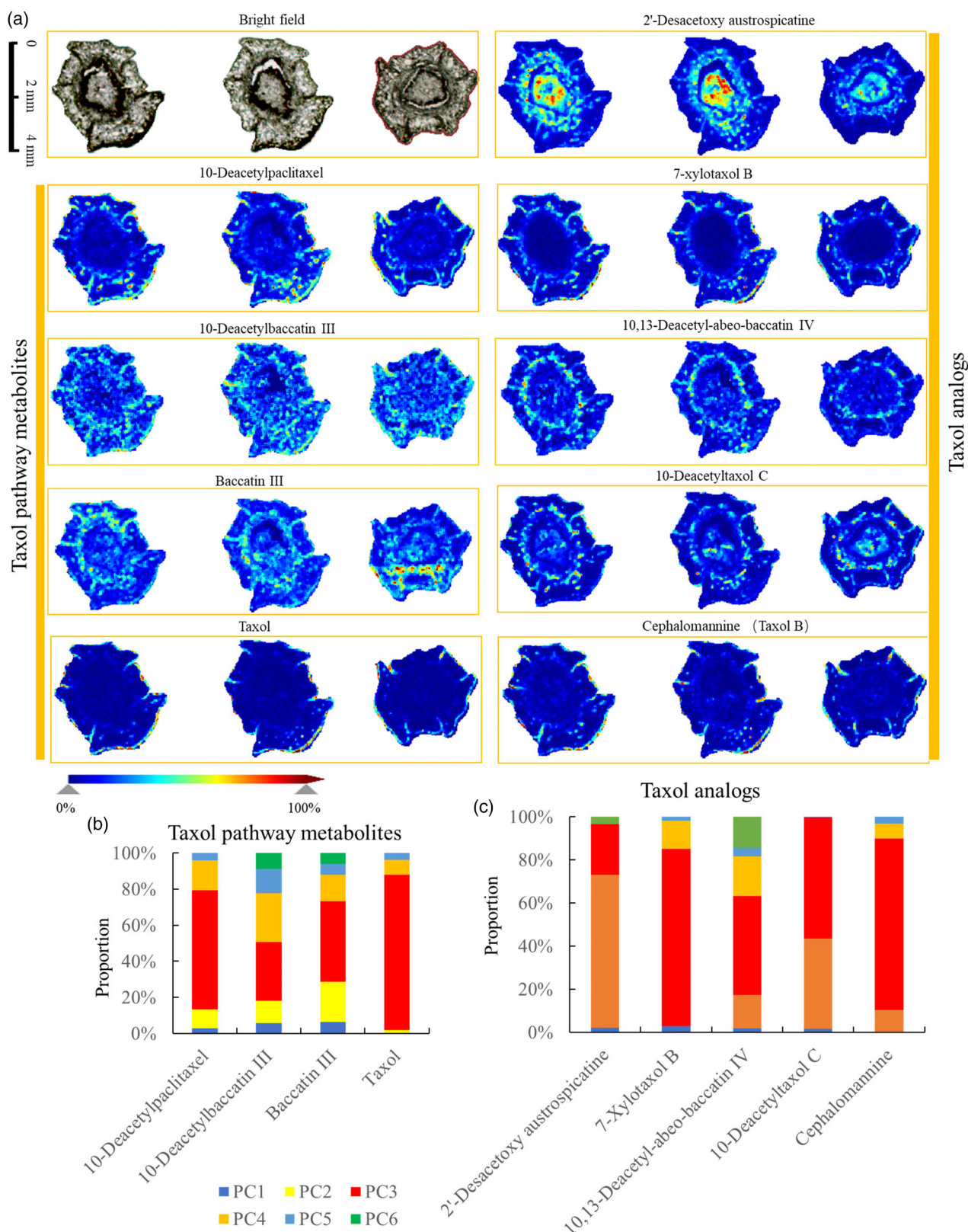


Figure 2. Visualization of taxoid accumulation in the *T. mairei* stems. (a) The tissue-specific distribution of nine taxoids, including four taxol biosynthesis-related taxoids and five taxol analogs, detected by MALDI-IMS analysis. The color scale ranges from 0 to 100%. (b) The proportions of taxol pathway metabolites in different PCs. (c) The proportions of taxol analogs in different PCs.

barcodes and transcripts were marked with unique molecular identifiers (UMIs), and cells were quantified using barcodes and UMI counts (Figure 3c). scRNA-seq produced a pool of 21 233 cells (7660 cells in *Taxus*-1, 8411 cells in *Taxus*-2, and 5162 cells in *Taxus*-3), with a median of 2996 genes per cell (Figure 3d). Over three biological repeats, sequencing saturation varied between 52.3 and 63.6% (Table S2). The *t*-distributed stochastic neighbor embedding (*t*-SNE) projection of cells, colored by different biological repeats, is shown in Figure 3e and the *t*-SNE projection of cells colored with UMI counts is shown in Figure 3f. Raw *t*-SNE data imported into Seurat loupe software are listed in Table S3.

Gene expression and cell clustering

In total, 14 788 genes with high expression levels were used for stem cell clustering (Table S4). The transcriptomic profiles of cells were projected in an unsupervised manner and classified into 18 clusters containing distinct cell numbers (Figure S1a). The two largest clusters were Cluster 0 (2120 cells) and Cluster 1 (1968 cells), and the smallest two were Cluster 16 (108 cells) and Cluster 17 (99 cells) (Figure S1b). We then reduced the large dataset using linear dimensional reduction analysis and identified 1530 cluster-enriched genes (Table S5). Among genes, three representative cluster-enriched genes for each cluster are shown in Figure 3g and Table S6.

Cell type marker genes and cluster annotation of stem cells

Several cell type marker genes in plant stems have been reported (Apelt et al., 2022; Chen, Tong, et al., 2021). Using sequence alignments, orthologous *Arabidopsis* cell type marker genes were identified in *T. mairei*, including 17 photosynthetic cell-, 10 cambium region-, 10 phloem parenchyma cell-, 40 xylem cell-, three xylem mother cell-, one phloem mother cell-, two endodermal cell-, nine xylem parenchyma cell-, four cork cell-, two epidermal cell-, two companion cell-, two cortex/endodermal cell-, five cortex/endodermis initial cell-, and seven sieve element-expressed genes (Table S7). Specific cell type-expressed *T. mairei* genes were compared with cluster-enriched genes to annotate cell clusters.

In our study, 13 cell type marker genes were used. *ABA REPRESSOR1* (*ABR1*, ctg510_gene.6) is a xylem parenchyma cell-specific gene and was enriched in Clusters 1 and 7 (Chen, Tong, et al., 2021). *APOPLASTIC AMINE OXIDASE 1* (*AO1*, ctg2853_gene.1) was highly expressed in xylem tissue and enriched in Cluster 2 (Ghuge et al., 2015). *GLABRA 2* (*GL2*, ctg8332_gene.2) is required for cell differentiation in the epidermis and was enriched in Clusters 4 and 17 (Lin & Aoyama, 2012). *PHOTOSYSTEM II REACTION CENTER W* (*PsbW*, ctg141_gene.24) is a photosystem II component-encoding gene and was enriched in Clusters 5 and 16 (Garcia-

Cerdan et al., 2011). *HB8* (ctg2829_gene.2) is an auxin-associated initiator of vascular cell differentiation and was enriched in Clusters 8 and 9 (Ohashi-Ito et al., 2013). *IRREGULAR XYLEM 1* (*IRX3*, ctg3909_gene.31) is a key regulator-encoding gene and was enriched in Cluster 9 (Schultz & Coleman, 2021). Another *HB8* gene (ctg38_gene.16) was enriched in Cluster 10. *BREVIS RADIX* (*BRX*, ctg19407_gene.1) is a molecular rheostat adjuster of auxin flux that promotes root phloem differentiation and was enriched in Cluster 11 (Marhava et al., 2018). *DA1-RELATED PROTEIN 2* (*DAR2*, ctg579_gene.1) is a marker gene of companion cells and was enriched in Cluster 11 (Zhang et al., 2019). *NRT1/PTR FAMILY 3.1* (*NPF3.1*, ctg2761_gene.7) is a gibberellic acid transporter-encoding gene expressed in the endodermis and was enriched in Cluster 12 (Tal et al., 2016). *CALLOSE SYNTHASE 7* (*CalS7*, ctg1793_gene.8) is a phloem-specific gene involved in callose deposition in developing sieve elements during phloem formation and was enriched in Cluster 15 (Xie et al., 2011). *SCARFACE* (*SFC*, ctg564_gene.13) is a cambium-specific gene and was also enriched in Cluster 15. *ALIPHATIC SUBERIN FERULOYL TRANSFERASE* (*ASFT*, ctg10348_gene.6) is a cork marker gene and was enriched in Cluster 16 (Chen, Tong, et al., 2021) (Figure 3h). Thus, a high degree of cell heterogeneity existed in *T. mairei* stems.

Most clusters were annotated according to *T. mairei* marker genes. No known cell type markers represented Clusters 0, 3, 6, and 14. Clusters 1 and 7 mainly contained xylem parenchyma cells. Cluster 2 represented xylem cells. Epidermal cells were enriched in Clusters 4 and 17. Photosynthetic (cork) cells were enriched in Clusters 5 and 16. Clusters 8 and 9 mainly contained vascular cells. Clusters 10 and 13 represented xylem mother cells. Cluster 11 contained companion and phloem cells. Cluster 12 mainly contained endodermal cells, while cambium region and sieve element cells were enriched in Cluster 15 (Figure 3i). The approximate distribution of cell populations in *T. mairei* stems is shown in Figure 3j.

Metabolic trajectories in *T. mairei* stems

To perform *T. mairei* metabolic profiling at the single-cell level, Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were performed to identify primary and secondary metabolism-related genes (Figure S2a). In primary metabolism analyses, four saccharide-related KEGG terms consisting of 791 genes, 17 amino acid-related KEGG terms consisting of 1651 genes, and five fatty acid-related KEGG terms consisting of 495 genes were identified. In secondary metabolism analyses, 438 phenylpropanoid-, 192 flavonoid-, 573 alkaloid-, 223 terpenoid-, 593 lipid-, 125 steroid-, 150 quinone-, and 84 oxidation-related genes were identified (Table S8).

In pseudotime trajectory analyses, certain cell types were set as the initial point, such as cambium region and phloem cells (Figure S2b). The pseudotime trajectory had six

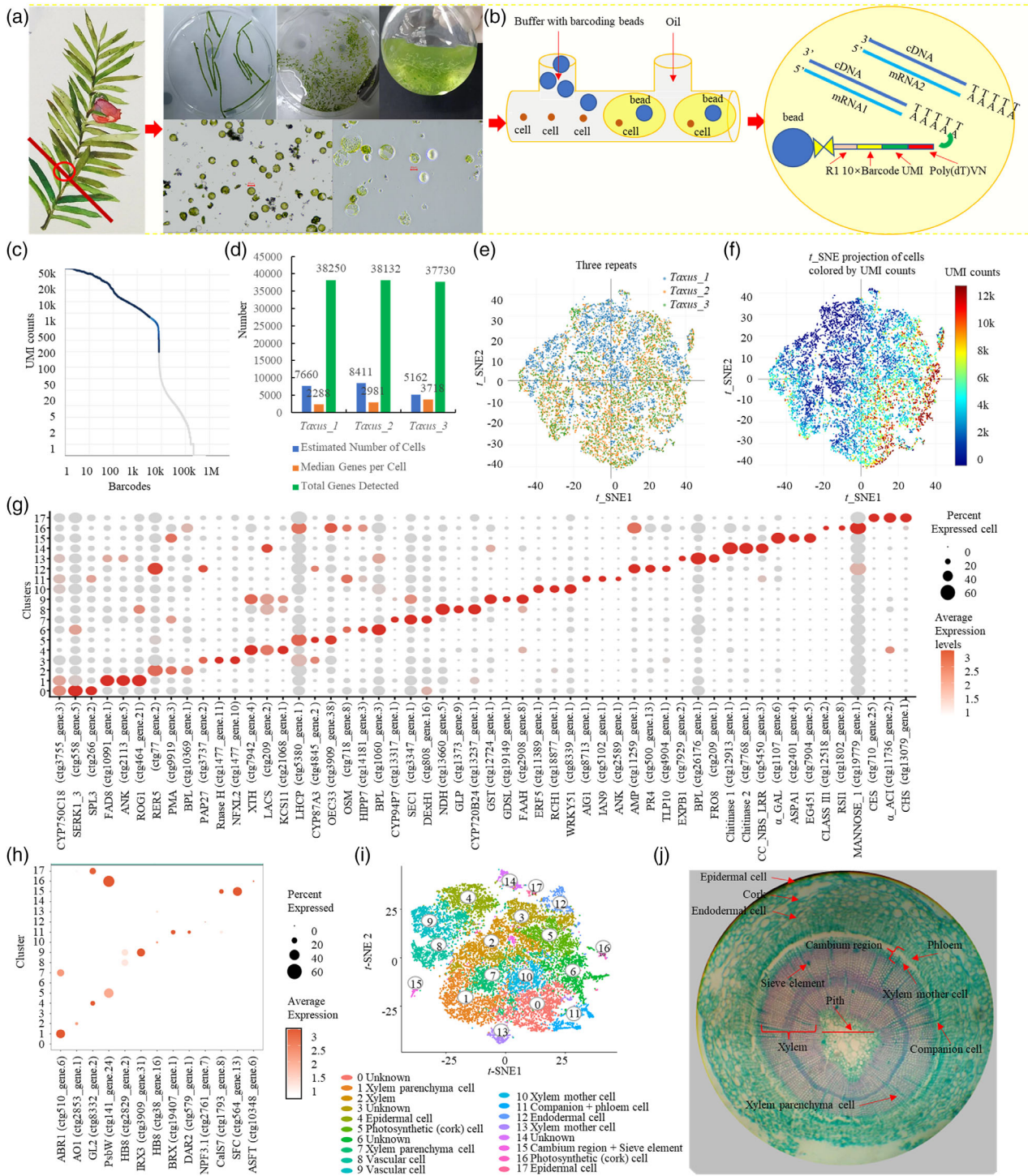


Figure 3. scRNA-seq and cell clustering of the *T. mairei* stem. (a) The basic process of protoplast isolation from stems to single cells. (b) Workflow of the scRNA-seq of *T. mairei* stems. Protoplasts isolated from the stems were loaded into a 10× Genomics Chromium Controller. (c) The number of effective cells was quantified by barcodes and UMI counts. (d) The numbers of cells and genes detected in each sample repeat (*Taxus_1*, *Taxus_2*, and *Taxus_3*). (e) t-SNE projection of cells colored by different biological repeats. (f) t-SNE projection of cells colored by UMI counts. (g) The expression pattern of three representative cluster-enriched genes for each cluster. Dot diameter indicates the proportion of cluster cells expressing a given gene. Color bar indicates the average expression level of each gene in different clusters. (h) The expression pattern of 13 tissue-specific expressed marker genes in 18 different clusters. (i) Visualization of 18 cell clusters using the t-SNE method. Dots indicate individual cells and colors indicate different cell clusters. (j) Schematic diagram of the stem segment tissue locations.

branch points, which divided cells into 13 states (Figure S2c). Xylem parenchyma cells predominantly occupied branches 1 and 5. Xylem, xylem mother, and vascular cells were evenly distributed across all branches. Photosynthetic (cork) cells were mainly distributed in branch 4, and companion, phloem, and cambium cells mainly occupied branch 3 (Figure S2d). The expression patterns of 22 flavonoid pathway- (Table S9), 36 lipid metabolism- (Table S10), 43 phenylpropanoid metabolism- (Table S11), 24 terpenoid metabolism- (Table S12), 11 steroid biosynthesis- (Table S13), and 35 glucosinolate metabolism-related genes (Table S14) during cell differentiation processes were analyzed (Figure S2e).

Cell type-based taxol biosynthesis in *T. mairei*

Our scRNA-seq analyses identified several genes encoding eight key enzymes involved in the taxol biosynthesis pathway, including *TS*, *T5OH*, *T13OH*, *TAT*, *T10OH*, *DBAT*, *DBTNBT*, and *T14OH* (Figure 4a). Some enzymes had more than one encoding gene, with differential cell coverage. For the three *TS* genes, *ctg6088_gene.1* expression had the largest cell coverage (1.24%). For the five *T5OH* genes, *ctg2768_gene.2*, *ctg7747_gene.2*, and *ctg5306_gene.3* were expressed in >1% of total cells. For the five *T13OH* genes, *ctg4364_gene.1* expression had the largest cell coverage (0.53%). For the six *TAT* genes, *ctg195_gene.25* expression had the largest cell coverage (0.57%). For the four *T10OH* genes, *ctg2120_gene.5* and *ctg_gene.32* were expressed in >2% of total cells. For the two *DBAT* genes, *ctg5026_gene.32* expression had the largest cell coverage (0.58%). For the two *DBTNBT* genes, *ctg887_gene.19* expression had the largest cell coverage (0.90%). The *T14OH* gene (*ctg8816_gene.2*) was expressed in approximately 1.46% of total cells (Figure 4b). Cell-specific expression analyses showed that most taxol-related genes were expressed in Clusters 1, 7, 12, and 17. According to cell type annotations, *TS*, *T5OH*, and *T14OH* genes were expressed in endodermal cells, *T10OH* and *DBTNBT* genes in xylem parenchyma, and *DBAT* in epidermal cells (Figure 4c). The UMAP visualization of the expression patterns of 11 major taxol-related genes is shown in Figure 4d.

The results of our successive differentiation trajectory analysis during bark development, from phloem to epidermis, are shown in Figure 4e. During bark development, *T10OH* (*ctg2120_gene.10*) and *T5OH* (*ctg7747_gene.2*) were highly expressed at early stages, while *T13OH* (*ctg4364_gene.1*) and *DBAT* (*ctg5026_gene.32*) were highly expressed at middle stages and *DBTNBT* (*ctg887_gene.19*) was highly expressed at later stages (Figure 4f). Successive differentiation trajectory analysis during xylem development, from xylem mother cells to xylem, is shown in Figure 4g. During xylem development, *T10OH* (*ctg2120_gene.5*), *T13OH* (*ctg4364_gene.1*), *T5OH* (*ctg5306_gene.3* and *ctg7747_gene.2*), and *T14OH* (*ctg8816_gene.2*) were highly expressed at initial stages, while

TAT (*ctg195_gene.25*) and *TS* (*ctg6088_gene.1*) were highly expressed at later stages (Figure 4h).

CYP superfamily expression patterns

Many taxol biosynthesis-related hydroxylases are key CYP superfamily members (Kaspera & Croteau, 2006). Many CYP family members (487 genes) were identified in a previous genomic analysis (Xiong et al., 2021) (Table S15). Traditional transcriptomic analyses investigated CYP family expression patterns at different tissue levels and identified 42 genes in bark, 114 in strobilus, 130 in leaf, and 90 in root (Figure 5a,b). scRNA-seq identified CYP gene expression patterns at the single-cell level. Interestingly, CYP725 subfamily members were enriched in Clusters 1 and 12 (Figure 5c), e.g., *ctg17930_gene.1* and *ctg2120_gene.6* were significantly expressed in Cluster 1, while *ctg7747_gene.2* and *ctg8379_gene.13* were highly expressed in Cluster 12 (Figure 5d).

Pseudotime trajectory analysis of cells in Cluster 1 identified three branch points dividing xylem parenchyma cells into six states (Figure 5e). Heatmaps show Cluster 1 CYP725 gene expression patterns during xylem parenchyma cell differentiation (Figure 5f). Pseudotime trajectory analysis of Cluster 12 showed three branch points dividing endodermal cells into six states (Figure 5g). Heatmaps show Cluster 12 CYP725 gene expression patterns during endodermal cell differentiation (Figure 5f). The expression of CYP725 family genes upon methyl jasmonate (MeJA) treatment was analyzed (Figure S3). Interestingly, the expression of several CYP725 family genes, such as *ctg593_gene.12*, *ctg2768_gene.2*, *ctg593_gene.6*, *ctg7168_gene.10*, *ctg2120_gene.6*, and *ctg13383_gene.3*, was more than 5-fold increased upon MeJA treatment.

The identification of cell-specific TFs in taxol biosynthesis

Currently, our understanding of the transcriptional regulation of taxol biosynthesis at the single-cell level is limited. scRNA-seq showed that most taxol biosynthesis-related genes in cells belonged to Clusters 1, 7, 12, and 17. Of cluster-enriched TFs, we identified 20 Cluster 1-enriched TFs, five Cluster 7-enriched TFs, five Cluster 12-enriched TFs, and six Cluster 17-enriched TFs (Figure 6a and Table S16).

To investigate potential TFs implicated in taxol biosynthesis, the promoter sequences of eight taxol biosynthesis-related genes, identified by scRNA-seq, were scanned for known TF-binding elements; two MBEs, one GT element, and one GCC-box were identified in the *TS* promoter, two MBEs and one GT element were identified in the *T5OH* promoter, three MBEs and one GCC-box were identified in the *T13OH* promoter, two MBEs and two GCC-boxes were identified in the *T14OH* promoter, two W-boxes, two ABREs, two MBEs, and one AuxRE were identified in the *TAT* promoter, two MBEs, three ABREs, and one

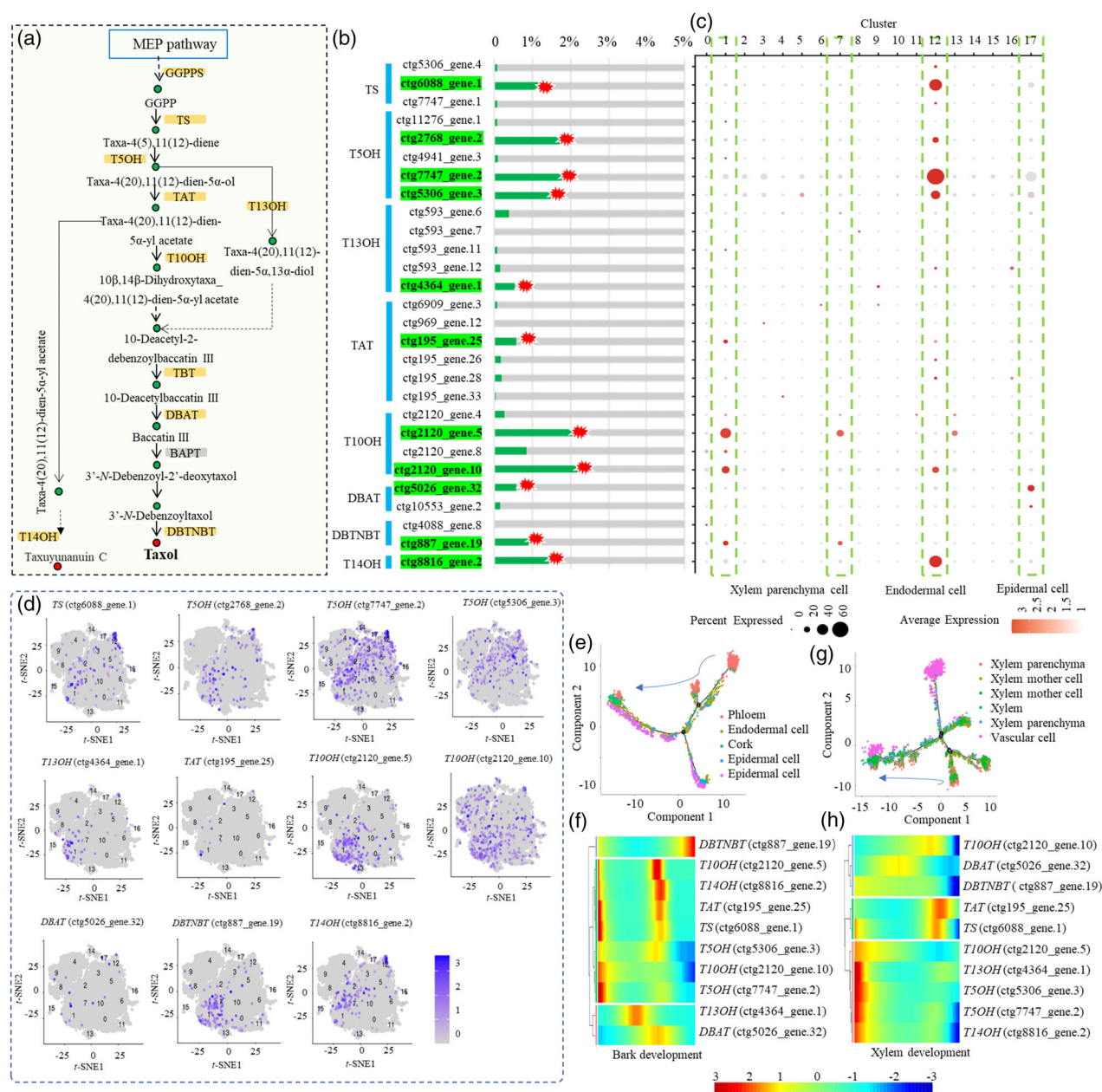


Figure 4. Cell type specificity of taxol biosynthesis in *T. mairei*. (a) Overview of the paclitaxel biosynthesis pathway, including enzymes and intermediates. (b) The cell coverage of taxol biosynthesis-related genes. Each key enzyme was encoded by at least one gene. Red stars indicate the encoding gene with the largest cell coverage. (c) Cell-specific expression analysis of taxol biosynthesis-related genes. Clusters 1 and 7 refer to xylem parenchyma cells, Cluster 12 refers to endodermal cells, and Cluster 17 refers to epidermal cells. (d) UMAP visualization of expression patterns of the 11 major taxol-related genes. (e) The successive differentiation trajectory during the bark development from phloem to epidermis. (f) Heatmap showing the expression levels of taxol-related genes during bark development. (g) The successive differentiation trajectory during xylem development from xylem mother cell to xylem. (h) Heatmap showing the expression levels of taxol-related genes during xylem development. The heatmap scale ranges from -3 to +3 on a log₂ scale.

GT element were identified in the *T10OH* promoter, two MBEs and two ABREs were identified in the *DBTNBT* promoter, and finally, three ACA elements, one ABRE, one GCC-box, and two MBEs were identified in the *DBAT* promoter (Figure 6b).

Co-localization analyses identified regulatory relationships between taxol biosynthesis-related genes and

upstream TFs. For Clusters 1 and 7, for *TAT*, *T10OH*, and *DBTNBT* genes, several TFs, such as MYB11/32/40/64, NAC1/2/22, bHLH68, WRKY31/51, and GT_2B, were identified as potential regulators. For Cluster 12 and *TS*, *T5OH*, *T13OH*, and *T14OH* genes, four TFs, including ERF16/115, GT_2, and MYB47, were identified as potential regulators. For Cluster 17 and the *DBAT* gene, five TFs, including

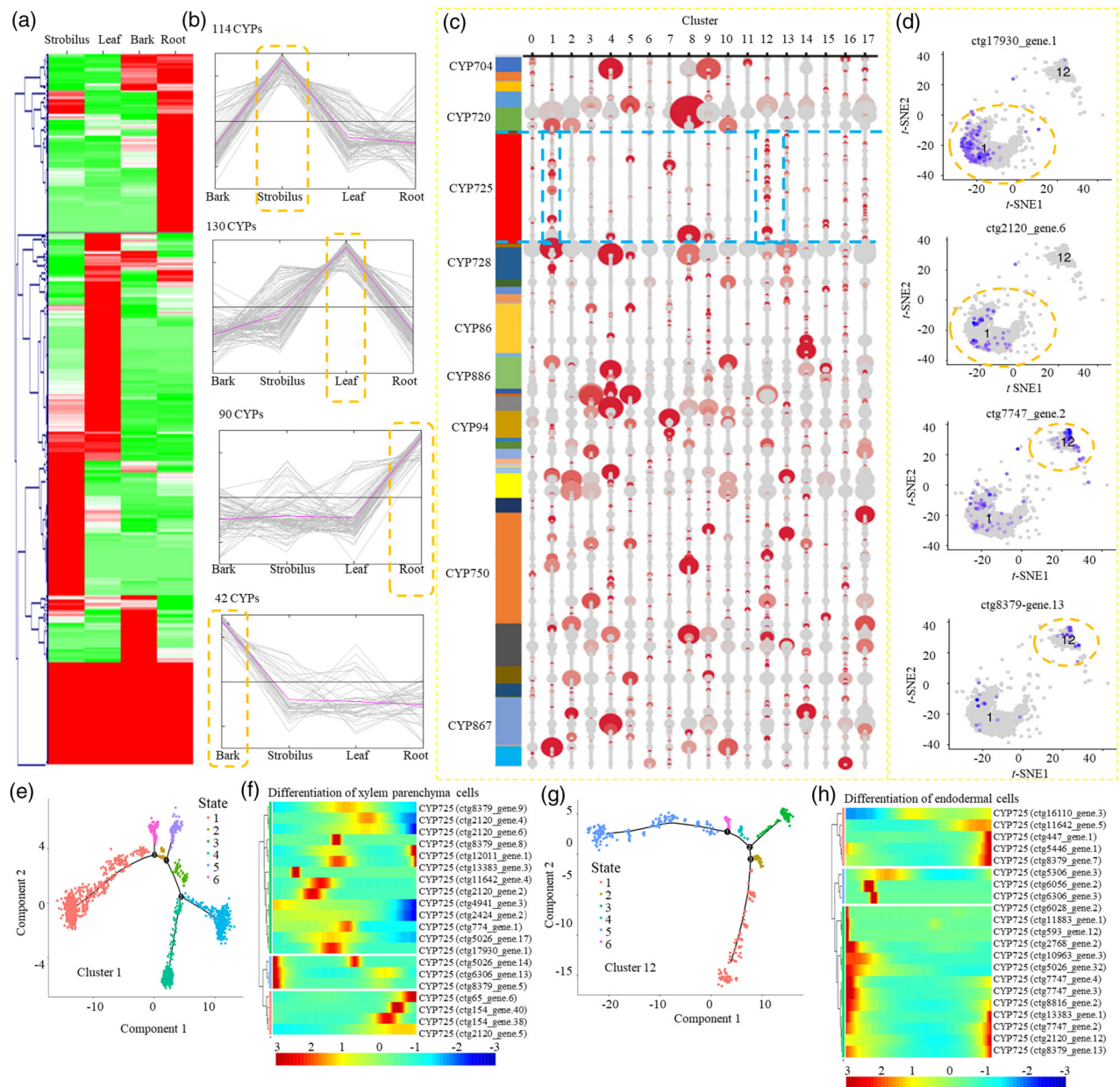


Figure 5. Expression patterns of CYP superfamily. (a) Expression patterns of CYP family members at the tissue level as determined by bulk RNA-seq. (b) MeV clustering analysis identified organ-specific CYP genes. (c) Cell-specific expression analysis of CYP family members at the single-cell level by scRNA-seq. Blue boxes indicate that the CYP725 subfamily members were enriched in the cells belonging to Clusters 1 and 12. (d) UMAP visualization of expression patterns of four selected CYP725 family members in Clusters 1 and 12. (e) Pseudotime trajectory analysis of the cells in Cluster 1. (f) Heatmaps showing the expression patterns of specific Cluster 1-expressed CYP725 genes during the differentiation of xylem parenchyma cells. (g) Pseudotime trajectory analysis of the cells in Cluster 12. (h) Heatmaps showing the expression patterns of specific Cluster 12-expressed CYP725 genes during the differentiation of endodermal cells. The heatmap scale ranges from -3 to $+3$ on a $\log_2$ scale.

MYBL, MYB69, TEM1, RAV1, and NAC3, were identified as potential regulators (Figure 6c).

Verifying cell-specific TF binding to targets

Four TFs, including NAC2 (ctg541_gene.6), WRKY31 (ctg8684_gene.5), MYB47 (ctg293_gene.5), and bHLH68 (ctg458_gene.1), were selected to verify cell-specific TF

binding to targets. NAC2, WRKY31, MYB47, and bHLH68 full-length sequences were successfully cloned by PCR (Figure S4a), followed by prokaryotic expression and affinity purification of GST-tagged TF proteins (Figure S4b–e). Electrophoretic mobility shift assay (EMSA) data showed that MYB47 directly bound to MBE sites in *TS* and *T5OH* promoters, while bHLH68 directly bound to ABRE sites in

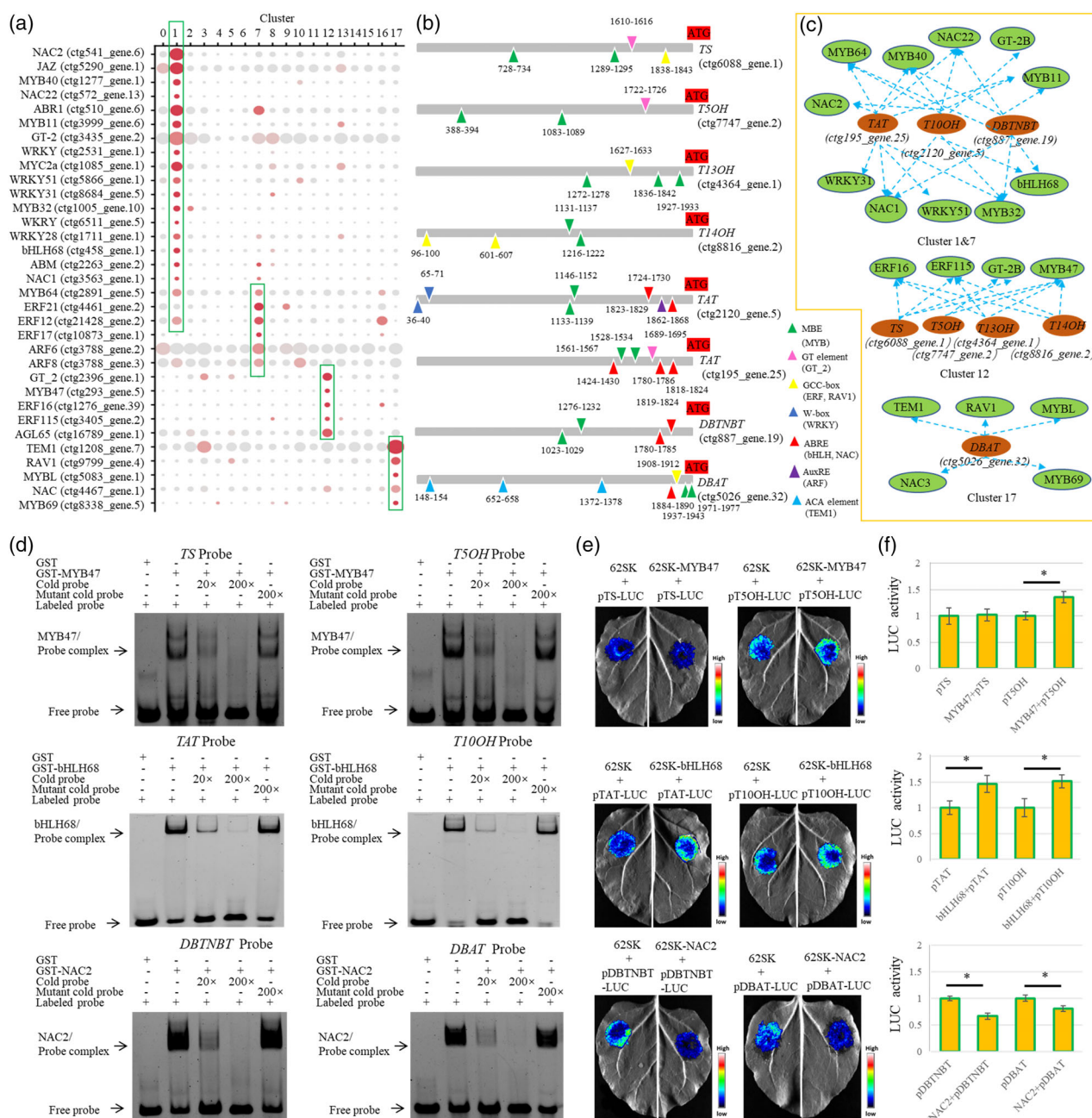


Figure 6. Identification of cell-specific TFs involved in taxol biosynthesis. (a) Cell-specific expression analysis of TFs at the single-cell level by scRNA-seq. Green boxes indicate the cell cluster-specific TFs. (b) Screening of TF-binding elements in the promoter regions of *TS*, *T5OH*, *T13OH*, *T14OH*, *TAT*, *T10OH*, *DBTNBT*, and *DBAT* genes. Triangles indicate different TF-binding elements, including MBEs (green), GT-elements (pink), GCC-boxes (yellow), W-boxes (dark green), ABREs (red), AuxREs (purple), and ACA elements (light blue). (c) Prediction of transcriptional regulation networks according to the cell type-specific expression pattern. (d) EMSA for the binding of cell-specific TFs to their targets. GST only or the TF-GST fusion protein was incubated with the probes containing the binding elements derived from the promoters. - and + represent absence and presence, respectively, and 20× and 200× show increasing amounts of probes for competition. (e) The dual-luciferase assays in tobacco leaves showed that co-transformation of TFs activates the promoters of taxol biosynthesis-related genes. (f) Effects of MYB47, bHLH68, and NAC2 on the luciferase activities of the *TS/T5OH*, *TAT/T10OH* and *DBTNBT/DBAT* reporters, respectively. Relative LUC activity represents the activity ratio of firefly luciferase to Renilla luciferase. Each value is the mean ± standard deviation of three biological repeats. **P* < 0.05.

TAT and *T10OH* promoters and NAC2 directly bound to ABRE sites in *DBTNBT* and *DBAT* promoters (Figure 6d). No binding was identified between WRKY31 and the *TAT* promoter.

Multiple sequence alignment analyses showed that MYB47 contained a classic R2R3 domain, bHLH68 contained a classic basic helix-loop-helix domain, and NAC2 contained five common motifs (A–E) (Figure S5). To

determine MYB47, bHLH68, and NAC2 transcription activities *in vivo*, dual-luciferase reporter assays were performed, showing that MYB47 significantly activated *T5OH* expression. However, MYB47 overexpression did not affect *TS* expression. bHLH68 significantly activated *TAT* and *T10OH* expression. NAC2 significantly repressed *DBTNBT* and *DBAT* expression (Figure 6e,f). Thus, MYB47 and bHLH68 were potential transcriptional activators and NAC2 was a potential transcriptional inhibitor in the taxol biosynthesis pathway.

Predicting a secondary metabolite transporter

Many ABC family proteins function as secondary metabolite transporters in different plants (Theodoulou & Kerr, 2015). In our study, 235 ABC superfamily members were identified, including seven ABCA, 66 ABCB, 25 ABCC, three ABCD, one ABCE, 10 ABCF, 73 ABCG, and three ABCI subfamily members, as well as 30 ABC-like proteins and 17 MDR-like ABC members (Table S17).

In the human ABC superfamily, ABCG2 is reportedly responsible for toxic xenobiotic and metabolite efflux (Moinul et al., 2022). Evolutionary analyses showed that *T. mairei* ABCG2 (ctg5370_gene.6) was homologous to human ABCG2 isoforms 1/2/3 (Figure 7a). No cell-specific *T. mairei* ABCG2 expression patterns were observed using feature plot analysis, suggesting extensive functions in different cell types (Figure 7b,c). Seven transmembrane domains were identified at the TcABCG2 C-terminus (Figure 7d). To confirm the subcellular localization of TcABCG2, a GFP-fused TcABCG2 protein was transiently expressed in *Nicotiana tabacum* epidermal cells, showing TcABCG2 localized to the cell membrane and was a potential secondary metabolite transporter (Figure 7e).

DISCUSSION

Hundreds of taxoids are reportedly enriched in *Taxus* stems, particularly the stem bark (Uniyal, 2013). It has been hypothesized that native *Taxus* plants transport taxol to avoid significant taxol cytotoxicity (Soliman et al., 2015; Soliman & Raizada, 2020). In *T. media*, BAC is highly accumulated in wood and taxol is primarily localized to outer tissues such as phloem and bark (Soliman & Raizada, 2020). Previous phytochemical analyses confirmed that taxol and 10-DAB mainly accumulated in the phloem and BAC primarily accumulated in the pith (Yu et al., 2020). However, the spatial distribution of taxol and other secondary metabolites in *Taxus* stems remains largely uncharacterized.

High-resolution MALDI-2-MS imaging can be used to explore and visualize metabolic heterogeneity in a specific organ (Bien et al., 2022). In our study, MS features were separated into six distinct clusters and provided six different visual accumulation patterns. PC3/5/6 indicated outer stem accumulation patterns, while PC1 indicated middle

stem accumulation patterns and PC2/4 indicated inner stem accumulation patterns. Our data provided detailed and precise localization information for different bioactive components in *Taxus* stems and may promote the comprehensive utilization of *Taxus* trees.

The taxol biosynthesis pathway involves several intermediates with different chemical structures and molecular weights (Yu et al., 2021). Some compounds, such as 10-DAB and BAC, are important raw materials for the semi-synthesis of taxol in the pharmaceutical industry (Wang, Zhang, et al., 2021). Interestingly, three key taxol biosynthesis intermediates, 10-DAP, 10-DAB, and BAC, were detected by our MALDI-2-MS imaging strategy, and displayed an even 10-DAP distribution across the whole stem, relative 10-DAB and BAC enrichment in outer stem tissues, and distinct taxol concentrations in epidermis cells. Distinctive accumulation patterns indicated different synthesis sites and potential taxoid transport mechanisms across *T. mairei* stems.

In addition to taxol, several taxol analogs, such as 10,13-deacetyl-abeo-baccatin IV, from the same *Taxus* tree sites, are expected to possess pharmacological properties (Yadava et al., 2011). Special taxol analogs, taxol B (cephalomannine) and taxol C, not directly involved in the taxol pathway, can be converted to taxol via chemical conversion procedures (Xue et al., 2020). Thus, considering its economic availability, taxol analogs have attracted much attention from chemists and pharmacists. Our MALDI-2-MS imaging showed that the accumulation patterns of three taxol analogs, 7-xyloxol B, 10,13-deacetyl-abeo-baccatin IV, and 10-deacetyl taxol C, were similar to the 10-DAB and BAC patterns, suggesting a possible *in situ* conversion step between taxol intermediates and taxol analogs. In *T. mairei*, both taxol and cephalomannine accumulated in the epidermis. It is no coincidence that cephalomannine is a major unwanted, ineluctable by-product of the taxol extraction process from plant cell cultures (McPartland et al., 2012).

scRNA-seq is used to identify cell types and investigate cell heterogeneity at high resolution (Islam et al., 2014). However, a lack of complete genome information and efficient protoplast preparation methods has limited scRNA-seq applications for woody plants, particularly in highly lignified tissues (Chen, Tong, et al., 2021). Recent *T. mairei* genome information and optimized protoplast isolation protocols have facilitated the generation of a comprehensive single-cell transcriptional *Taxus* stem landscape (Xiong et al., 2021). Our scRNA-seq analysis produced a pool of 21 233 cells (average 7078 cells in each repeat), which was larger than that of poplar scRNA-seq (3626 and 3170 cells from bark and wood tissue, respectively) and similar to that of tea tree scRNA-seq (9458 and 7519 cells from first and third leaves, respectively), but lower than that of rice (*Oryza sativa*) (237 431 cells from

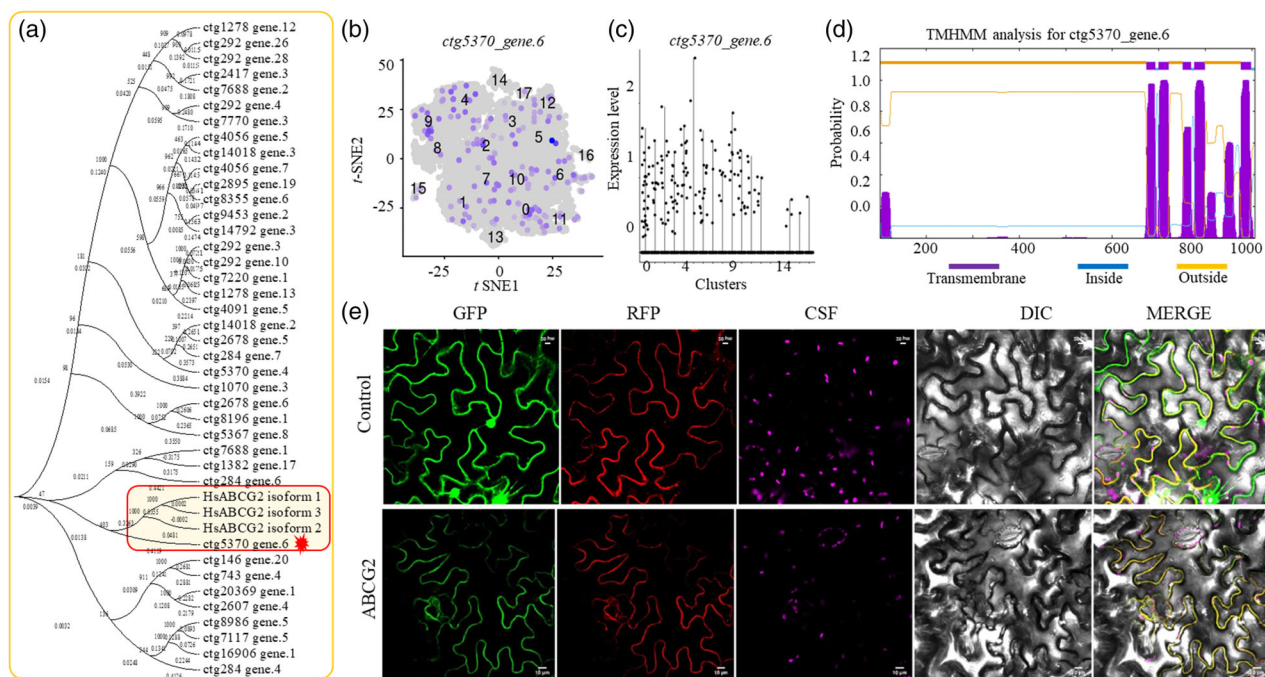


Figure 7. Prediction of a secondary metabolite transporter. (a) Evolutionary analysis identified the homologous protein of ABCG2 in *T. mairei* with its homologous genes in human. (b) UMAP visualization of expression patterns of the ABCG2 gene. (c) The expression level of ABCG2 in different cell clusters. (d) Prediction of the transmembrane domains of ABCG2. Purple lines indicate transmembrane domains, blue lines indicate intracellular domains, and yellow lines indicate extracellular domains. (e) Transient expression of the *T. mairei* ABCG2 protein in *Nicotiana tabacum* epidermal cells. Control, 35S::GFP vector; RFP, red fluorescent protein (RFP::OsMAC1); CSF, chloroplast spontaneous fluorescence; DIC, digital image correlation.

leaves and roots) and Arabidopsis (36 643 cells from the vegetative shoot apex) (Chen, Tong, et al., 2021; Wang et al., 2022; Wang, Huan, et al., 2021; Zhang et al., 2021). Sufficient cells ensured that our genome-scale expression profiling covered all stem cell types.

Several bulk RNA-seq studies have reported gene-metabolite coordination at different tissue levels in *Taxus* (He et al., 2018; Mubeen et al., 2018); however, correlations between gene expression profiles and metabolite accumulation in *Taxus* cell types have been rarely investigated. Cell annotations are required to define cell types in non-model plants. Some marker genes in model plants, e.g., Arabidopsis, express homologous genes which show similar expression patterns when compared with woody trees (Chen, Tong, et al., 2021; Wang et al., 2022). In our study, 13 cluster-specific markers were used to annotate cell types in *T. mairei* stems. Nevertheless, the lack of efficient genetic transformation systems and stable *in situ* hybridization methods has limited the precise identification of *Taxus* cell type markers (Shao et al., 2021).

scRNA-seq technology has been used to generate several plant cell development atlases. For example, a single-cell atlas defined the fate determination of fiber cell initiation in cotton (*Gossypium hirsutum*), leaf development in tea, early reproductive development in rice, vegetative shoot apex and root development in Arabidopsis, and

stem-differentiating xylem in poplar (Qin et al., 2022; Shao et al., 2021; Wang et al., 2022; Zhang et al., 2021; Zong et al., 2022). However, plant cell metabolic atlases have not yet been reported.

Based on tissue-specific markers, we created the first single-cell *Taxus* stem atlas which showed spatial distribution patterns in cells. Pseudotime trajectory analyses reordered cells along one main developmental trajectory, generating temporal distribution patterns in cells (Zheng et al., 2021). Cells from different clusters, representing different cell types, were separately distributed in the pseudotime trajectory. As *Taxus* is a well-known medicinal plant, we focused on the dynamic expression patterns of six representative secondary metabolic pathway-related genes during stem development. For example, geranylgeranyl diphosphate synthase (GGPPS), which is a key diterpene biosynthesis pathway enzyme and therefore a non-specific taxane enzyme, is encoded by five genes in *T. mairei*. Critically, one GGPPS-encoding gene (ctg4334_gene.2) was expressed in early stages and four genes (ctg14140_gene.5, ctg1897_gene.3, ctg3914_gene.4, and ctg12725_gene.16) were expressed in later stem developmental stages. Therefore, developmental stages are important factors regulating plant secondary metabolism.

Metabolic engineering relies on the selection of key rate-limiting enzymes; thus, multiple encoding genes can

provide several outcomes. TS is the rate-limiting enzyme in the taxol biosynthesis pathway and was encoded by three genes (Escrich et al., 2022). Among these genes, ctg5306_gene.4 and ctg7747_gene.1 showed extremely low cell coverage (<0.1%), suggesting silenced expression in most stem cells. DBTNBT was encoded by two genes in *T. mairei* and is a key flux-limiting enzyme of the taxane biosynthesis pathway (Kashani et al., 2018). Unlike active DBTNBT (ctg887_gene.19), ctg4088_gene.8 was completely suppressed in most stem cells. Interestingly, the artificial induction of these silenced genes could increase taxol levels in different cell lines.

Cell expression analysis showed that most known taxol biosynthesis pathway genes were expressed in epidermal, endodermal, and xylem parenchyma cells. Endodermal cells specifically expressed TS, T5OH, and T14OH, suggesting their downstream products probably accumulated in endodermal cells. Xylem parenchyma cells specifically expressed T10OH, suggesting these cells were likely the main biosynthesis sites of 10 β ,14 β -dihydroxytaxa-4(20) and 11(12)-dien-5 α -yl acetate. Furthermore, our single-cell atlas helped us analyze the spatiotemporal expression patterns of taxol biosynthesis pathway genes during bark and xylem development. The DBTNBT gene (ctg887_gene.19) was mainly expressed in later bark development stages, suggesting that senescent bark is an important taxol biosynthesis tissue. Most genes, such as T10OH (ctg2120_gene.5), T13OH (ctg4364_gene.1), and TS (ctg8816_gene.2), were expressed in early xylem developmental stages, suggesting that mature xylem is not the main taxol biosynthesis site. The cell-specific expression of taxol biosynthesis-related genes may cause uneven taxoid distribution across *T. mairei* stems. The phytohormone MeJA has been found to significantly elicit the production of paclitaxel in cultured *T. media*, *Taxus canadensis*, and *Taxus cuspidata* cells (Sun et al., 2013). In *T. chinensis*, the expression pattern of CYP genes during the long-time subculture and MeJA elicitation were analyzed, resulting in the identification of six novel CYP725 genes as candidates for taxol biosynthesis (Liao et al., 2017). Interestingly, several MeJA-induced CYP725 family genes showed tissue-specific expression in the *T. mairei* stem, suggesting a potential role in MeJA-induced taxol accumulation.

The taxol biosynthesis pathway is controlled by several environmental and genetic factors (Feng et al., 2023; Yu et al., 2020; Yu et al., 2022; Zhan et al., 2022). Recent studies on taxol biosynthesis regulators mainly focused on stimulus-responsive TFs (Cao et al., 2022; Chen, Zhang, et al., 2021; Zhang et al., 2015). We provided a single-cell perspective on novel TF screening, which was consistent with cell-specific expression patterns of taxol biosynthesis pathway genes. *cis*-Elements in promoter regions and cell type-specific expression patterns are important criteria for TF screening. Accordingly, several novel TFs, such as NAC, WRKY, bHLH, and ERF family TFs, were putatively involved

in the transcriptional regulation of the taxol biosynthesis pathway. The *in vitro* binding and *in vivo* regulation activities of several TFs were verified using EMSA and dual-luciferase reporter assays, respectively.

Previously, the roles of *T. media* MYB39/MYB3 and *T. chinensis* MYB29a in regulating taxol biosynthesis were identified (Cao et al., 2022; Yu et al., 2022). Our downstream target screening showed several different TF–target pairs, such as TmMYB3–TS, TmMYB3–TBT, TmMYB39–GGPPS, TmMYB39–T10OH, and TcMYB29a–T5OH. Also, a new MYB TF gene, *T. mairei* MYB47, and another TF–target pair, MYB47–T5OH, were identified. In woody trees, NAC family TFs are reportedly involved in secondary cell wall formation (Chen et al., 2022). NAC154 overexpression modulated phenolic glycoside levels in poplar stem wood (Jervis et al., 2015). In *T. mairei*, DBTNBT and DBAT were identified as potential targets of NAC2, suggesting an important role for NAC2 in xylem parenchyma cell-specific taxoid metabolism. Also, epidermal cells specifically expressed TEM1 and RAV1, endodermal cells specifically expressed GT_2 and ERF16/115, and xylem parenchyma cells specifically expressed JAZ, ABR11, and WRKY31/51, all of which encode potential TFs involved in the cell-specific transcriptional regulation of taxol biosynthesis.

While many intermediate metabolites were predicted as being synthesized in xylem parenchyma and endodermal cells, the end-product taxol is enriched in epidermal cells (Soliman & Raizada, 2020). Precursor availability affects each step of the taxol pathway (Yu et al., 2018). Taxol accumulation sites are not completely coincident with precursor and intermediate sites, suggesting taxoid movement across stems (Soliman & Raizada, 2020). We hypothesized that intermediate and/or end-product transport is required for taxol redistribution in stems. In humans, several ABC family members were reported as having taxol binding activities, e.g., human ABCG2 has indispensable multidrug resistance roles in cancer cells (Moinul et al., 2022). But, the taxoid transporter in *Taxus* trees is unknown. In our study, an ABCG2 family transporter-encoding gene was cloned and was suggested as a potential taxol transporter candidate.

Although taxol synthesis is dispersed across different type cells, the taxol end-product is finally transported to and accumulates in the outer epidermis of *T. mairei*. A previous study reported that taxol localization to hydrophobic bodies had evolved to protect cells from the cytotoxic effects of taxol (Soliman & Raizada, 2020). Taxoid transportation and redistribution may be effective strategies for protecting young tissue against taxol cytotoxicity. Also, taxol accumulation in outer tissue protects against front-line pathogen invasion, at the epidermis in most cases (Schafer & Wink, 2009).

In conclusion, MS imaging technology was used to characterize *T. mairei* stems. MALDI-2 MS imaging analysis

identified four taxol biosynthesis-related taxoids and five taxol analogs, thereby confirming tissue-specific taxoid accumulation in *T. mairei* stems. To identify the genetic mechanisms underpinning taxoid distribution in specific tissues, single-cell transcriptomic analysis was performed. We identified 18 cell clusters with several cell type marker genes, suggesting a high degree of cell heterogeneity in *T. mairei* stems. We created a single-cell *Taxus* stem metabolic atlas and identified spatial and temporal expression patterns of several taxol biosynthesis pathway genes. Furthermore, novel and cell-specific TFs in taxol biosynthesis were identified, including MYB47, bHLH, and NAC2. An ABCG2 family transporter-encoding gene was cloned, and was suggested as a putative taxol transporter candidate. We established a comprehensive, transcriptional, single-cell resolution landscape of the major cell types in *T. mairei* stems and generated a valuable resource outlining the specific cell type regulation of taxol biosynthesis (Figure S6).

METHODS

Plant materials

Three-year-old *T. mairei* trees were planted in an experimental field at the 'Cangqian' campus of Hangzhou Normal University, Hangzhou, China. Three independent young stems were collected and prepared for cryosection and spatial metabolomic analysis. Three independent stems were used for protoplasting and single-cell transcriptomic analysis.

Sampling and sectioning for spatial metabolomics

The young stem samples were cryosectioned using a Leica CM1950 freezing microtome. The slice thickness was set to 20 μm . The stem samples were transferred to a pre-cooled microtome for cooling for 1 h. After fixing and adjusting to a suitable angle, the stems were sliced according to the operating instructions. The resulting sections were transferred to a MALDI-2 special pre-cooled conductive glass slide. Then, the slides were transferred to a vacuum dryer for vacuum sealing for 30 min.

Matrix spraying and mass spectrometry imaging

The MALDI-2 special matrix was purchased from Bruker company. The matrix solution was evenly sprayed on the special conductive slide containing stem sections by a TM-Sprayer instrument (Soltwisch et al., 2015). The conductive slide coated with the substrate was placed on the target disk of a timsTOF fleX MALDI-2 (Bruker Daltonics GmbH, Bremen, Germany). The slice detection area was ticked out using Bruker Data Imaging software and divided into several 2D points according to the slice size. Under laser irradiation, the stem sections were scanned, and the chemical molecules released by ionization at the target sites were detected by MS to produce raw MS imaging files.

Analysis procedures of MS imaging data

The raw imaging files were imported into SCiLSTMLab 2021 software for data processing, including baseline subtraction, peak alignment, and data smoothing and normalization. Using the non-supervised spatial clustering application, the imaging pixels of the target area were clustered, and the areas with similar metabolite accumulation patterns were marked with the same color.

Feature extraction and qualitative analysis were carried out using a Bruker MetaboscapeTM workstation. The results were compared with the theoretical mass (molecular mass error within 10 ppm) in Bruker MetaboBASE Library metabobase 3.0 to identify the metabolites. PCA was conducted to analyze the distribution of the metabolites on the slices, and several PCs obtained could represent different imaging modes. Each principal component was a linear combination of all features.

Preparation of the *Taxus* stem protoplasts and single-cell suspensions

A total of 15–20 young stems of *T. mairei* were cut into dishes with enzymolysis solution containing cellulase R-10, isolation enzyme P, mannitol, KCl, and MES. The stems were cut into 0.1–0.2-mm fragments using a sharp scalpel and quickly transferred to the enzymolysis solution of the triangular flask. Then, the samples were treated with a vacuum pump for 30 min to completely immerse the blade filaments for enzymolysis. The samples were shaken at 28°C in the dark for 1 h. After the enzymatic hydrolysis, the supernatant was pipetted, and 10 mL of W5 solution (0.9 g NaCl, 1.84 g CaCl_2 , 0.0375 g KCl, and 0.0425 g MES, pH = 5.7) was added to gently resuspend the cells. The supernatant was discarded and the protoplasts were gently resuspended with 8% mannitol solution. The resulting cell suspension was filtered by passing through a 30–70- μm stacked cell strainer and centrifuged at $300 \times g$ for 5 min at 4°C. The suspension was resuspended in 100 μL Dead Cell Removal MicroBeads and dead cells were removed using the Miltenyi[®] Dead Cell Removal Kit (MACS 130-090-101). The overall cell viability was confirmed by trypan blue exclusion, which needed to be above 85%. Single cells in the single-cell suspension were counted and the cell count was adjusted to 700–1200 cells μL^{-1} .

scRNA-seq library construction and sequencing

Single-cell suspensions were loaded onto a 10 $\times$ Chromium platform to capture 5000 single cells according to the manufacturer's instructions of the 10 $\times$ Genomics Chromium Single-Cell 3 kit (V3). Libraries were sequenced on an Illumina NovaSeq 6000 sequencing system (paired-end multiplexing run, 150 bp) by LC-Bio Technology Co., Ltd. (Hangzhou, China) at a minimum depth of 20 000 reads per cell.

Raw data processing and visualization

Sequencing results were demultiplexed and converted to FASTQ format using Illumina bcl2fastq software (version 2.20). Sample demultiplexing, barcode processing, and single-cell 3' gene counting were performed using the Cell Ranger pipeline (version 6.1.1) and scRNA-seq data were aligned to the *T. mairei* reference genome. A total of 21 233 single cells captured from three independent stems of *T. mairei* were processed using 10× Genomics Chromium Single Cell 3' Solution. All cells were checked by quality control with the following thresholds: all genes expressed in less than three cells were removed, $500 < \text{number of genes expressed per cell} < 5000$, UMI counts < 500 , and the percentage of mitochondrial DNA-derived genes $< 25\%$. The Cell Ranger output was loaded into Seurat (version 3.1.1) for dimensional reduction, clustering, and analysis of scRNA-seq data. To visualize the data, the dimensionality of all cells was reduced into 2D space using the *t*-SNE method.

The 'Normalization' function of Seurat software was used to calculate the expression values of genes. PCA was performed using the normalized expression value. The top 10 PCs were used to perform clustering and *t*-SNE analysis.

Marker gene and cell type identification

To find clusters, the weighted shared nearest neighbor graph-based clustering method was applied. Marker genes for each cluster were identified with the Wilcoxon rank-sum test (default parameter is 'bimod': Likelihood-ratio test) with default parameters via the FindAllMarkers function in Seurat. This selected marker genes which are expressed in more than 10% of the cells in a cluster and with average log(fold change) > 0.25 .

To identify cell types, orthologous gene alignments of the known marker genes in the model plant Arabidopsis were performed. The Arabidopsis markers for different stem cell types were selected from the Plant Cell Marker Database (PCMDB, <http://www.tobaccodb.org/pcmdb/>) and their protein sequences were downloaded from the TAIR database (Jin et al., 2022). Using the Arabidopsis marker proteins as the query sequences, the homologous sequences of the *T. mairei* stem were searched by a local BLASP program. The proteins with top hit scores were selected and annotated as corresponding cell type markers of the *T. mairei* stem. Next, we isolated four different tissue types from *T. mairei* stems, namely, epidermal cells, endodermal cells, xylem cells, and pith cells. The expression levels of several *T. mairei* marker genes in four stem tissues were quantified by quantitative real-time reverse transcription PCR (qRT-PCR). qRT-PCR was performed according to our previous paper (Yu et al., 2017).

Pseudotime trajectory analysis

Single-cell trajectories were analyzed using Monocle software (version 2.0). The expression matrix of all cells was used to perform pseudotime analysis. Expression levels of key genes related to various primary and secondary metabolic processes were normalized by Monocle 2. Based on the pseudotime analysis, branched expression analysis modeling was applied to identify the genes at branch points.

Promoter isolation and cis-element scanning

Genomic DNA was isolated from the *T. mairei* stem using the classical cetyl trimethyl ammonium bromide method. The 2000-bp promoter sequences of the *TS*, *T5OH*, *T10OH*, *T13OH*, *T14OH*, *TAT*, *DBAT*, and *DBTNBT* genes were extracted from the *T. mairei* genomic sequence (Xiong et al., 2021). The promoter sequences were scanned using the PlantCARE program (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>). Several classic binding elements, including ERF binding sites (GCCGCC and CCGAC), GT₂ binding sites (GGTAAT and GGTAAT), bHLH binding sites (CANNTG, CACG(A/C)G, and CACGTG), TEM1 binding sites (CAACAC and ACACAG), and MYB binding sites (CAACTG and CAACCA), were screened.

Prokaryotic expression and electrophoretic mobility shift assay

The full-length coding sequences (CDSs) of target TF genes were inserted into pET30a vector and expressed in *Escherichia coli* cells to produce recombinant His-TF proteins. The expression of recombinant proteins was induced by 1 mM isopropyl β-D-1-thiogalactopyranoside and proteins were purified using Clontech His60 Ni Superflow Resin according to the manufacturer's procedures. EMSA was performed using the Light Shift Chemiluminescent EMSA Kit (GS009, Beyotime, China) according to our previous works (Yu et al., 2020; Yu et al., 2022). All primers and probe sequences are listed in Table S18.

Dual-luciferase reporter assay

The pGreenII0800-LUC and pGreenII62-SK vectors were used to produce target constructs. The full-length CDSs of *bHLH68*, *MYB47*, and *NAC2* were cloned and inserted into the pGreenII62-SK vector as effectors. The promoters of *TAT*, *TS*, *T5OH*, *DBTNBT*, and *DBAT* were cloned and inserted into the pGreenII0800-LUC vector as reporters. The GAL4BD and VP16 were used as positive and negative controls, respectively. All completed constructs were co-transformed into tobacco (*Nicotiana tabacum*) leaves by *Agrobacterium tumefaciens* GV3101-mediated transient expression. The firefly luciferase (LUC) and Renilla luciferase (REN) activities were determined using a dual-luciferase assay kit (Promega, Madison, USA). The primer sequences are shown in Table S18.

Subcellular localization

The full-length CDS of the *T. mairei* ABCG2 (ctg5370_gene.6) gene was cloned into the pBWA(V)HS vector with GFP. OsMCA1-mKATE was used as cell membrane marker. The primer sequences are shown in Table S18. All constructs were transformed into *Nicotiana tabacum* epidermal cells using a GV3101-mediated transient expression system.

Bioinformatics and statistical analysis

Unrooted phylogenetic trees were built using MEGA 7.1 software and the CLUSTALW program employing the neighbor-joining method. The physical and chemical properties of the transporter were analyzed online using ExPASy ProtParam (<https://web.expasy.org/protparam/>). The transmembrane domain of the transporter was analyzed using the TMHMM online tool (<https://services.healthtech.dtu.dk/service.php?TMHMM-2.0>).

Statistical analysis was performed using SPSS software ver.19.0 (SPSS Inc., Chicago, IL). Significance was determined at $P < 0.05$ by analysis of variance followed by Duncan's least significant difference test. For hierarchical clustering and visualization, all data were Z-score-normalized using MeV software by Euclidean distance.

AUTHOR CONTRIBUTIONS

CY, HW, and CS conceptualized the initial study; CY, KH, HZ., XL, CC, ZW, GC, JH, EB, XL, TD, YW, and QW were involved in the design of experiments; CY, KH, HZ, XL, CC, and MW performed the lab experiments; HW and CS drafted the initial article. All authors discussed the results, reviewed the article, and approved the final article.

ACKNOWLEDGMENTS

This work was funded by the National Natural Science Foundation of China (32270382 and 32271905), the Zhejiang Provincial Natural Science Foundation of China (Grant Nos. LY23C160001, LY18C050005, LY19C150005, and LY19C160001), and the Opening Project of the Zhejiang Provincial Key Laboratory of Forest Aromatic Plants-based Healthcare Functions (2022E10008). We are grateful to LC Sciences company (Hangzhou, China) and Shanghai Applied Protein Technology Co., Ltd (Shanghai, China) for transcriptomic and metabolomic analysis, respectively. We are also grateful to Professor Jianbin Yan (Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen, China) for genomic analysis.

CONFLICT OF INTEREST STATEMENT

The authors have no relevant financial or non-financial interests to disclose.

DATA AVAILABILITY STATEMENT

The *T. mairei* reference genome was downloaded from the NCBI database (ID: PRJNA730337). scRNA-seq data are available on NCBI with ID 'PRJNA730337'. The metabolite

measurement data generated and analyzed during the current study are available in the 'Baidu Netdisk' (<https://pan.baidu.com/s/1hXL8X9bCF5ZLi4UGKXRIHQ>; password: tjhg). All sequences of taxol biosynthesis-related genes and CYP450 genes are provided as Dataset S1.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Figure S1. The transcriptomic profiles of the genes with high variation in expression level.

Figure S2. Metabolism trajectories of *T. mairei* stems.

Figure S3. The expression of CYP725 family genes upon MeJA treatment.

Figure S4. Gene cloning and prokaryotic expression.

Figure S5. Multiple sequence alignment analysis of MYB47, bHLH68, and MAC2 proteins.

Figure S6. A model of cell-specific regulation and transport of taxol.

Table S1. List of qualitative and quantitative results of MS imaging.

Table S2. Sequencing statistics and quality control.

Table S3. Seurat loupe cell t-SNE data.

Table S4. The detailed information of the genes with high variation in expression.

Table S5. The detailed information of 1530 cluster-enriched genes.

Table S6. The detailed information of the representative cluster-enriched genes for each cluster.

Table S7. Gene list used for cell type identification in this study.

Table S8. KEGG analysis of the genes identified in *T. mairei*.

Table S9. The detailed information of the genes involved in flavonoid metabolism.

Table S10. The detailed information of the genes involved in lipid metabolism.

Table S11. The detailed information of the genes involved in phenylpropanoid metabolism.

Table S12. The detailed information of the genes involved in terpenoid metabolism.

Table S13. The detailed information of the genes involved in steroid metabolism.

Table S14. The detailed information of the genes involved in glucosinolate metabolism.

Table S15. The detailed expression levels of 487 CYP family members in different cell clusters.

Table S16. The detailed information of the four cluster-enriched TFs.

Table S17. The detailed information of the ABC superfamily in *T. mairei*.

Table S18. The primer sequences used in the present study.

Dataset S1. All the CDSs of taxol biosynthesis-related genes and CYP725 genes.

REFERENCES

- Ahmad, A., Ahmad, N., Anis, M., Faisal, M., Alatar, A.A., Abdel-Salam, E.M. et al. (2022) Biotechnological advances in pharmacognosy and *In vitro* manipulation of *Pterocarpus marsupium* Roxb. *Plants*, **11**, 247.
- Akande, T., Khatib, M., Ola Salawu, S., Afolabi Akindahunsi, A., Di Cesare Mannelli, L., Ghelardini, C. et al. (2022) (1)H NMR and HPLC-DAD-MS for

- the characterization of ellagitannins and triterpenoids of less investigated *Anogeissus leiocarpus* DC (Combretaceae) stem bark. *Food Chemistry*, **375**, 131813.
- Apelt, F., Mavrothalassiti, E., Gupta, S., Machin, F., Olas, J.J., Annunziata, M.G. *et al.* (2022) Shoot and root single cell sequencing reveals tissue- and daytime-specific transcriptome profiles. *Plant Physiology*, **188**, 861–878.
- Arendowski, A. & Ruman, T. (2017) Laser desorption/ionisation mass spectrometry imaging of European yew (*Taxus baccata*) on gold nanoparticle-enhanced target. *Phytochemical Analysis*, **28**, 448–453.
- Bamnashin, M., Mirjalili, M.H., Naghavi, M.R., Cusido, R.M. & Palazon, J. (2022) Gene expression pattern and taxane biosynthesis in a cell suspension culture of *Taxus baccata* L. subjected to light and a phenylalanine ammonia lyase (PAL) inhibitor. *Journal of Photochemistry and Photobiology, B*, **234**, 112532.
- Bawa, G., Liu, Z., Yu, X., Qin, A. & Sun, X. (2022) Single-cell RNA sequencing for plant research: insights and possible benefits. *International Journal of Molecular Sciences*, **23**, 4497.
- Bien, T., Koerfer, K., Schwenzfeier, J., Dreisewerd, K. & Soltwisch, J. (2022) Mass spectrometry imaging to explore molecular heterogeneity in cell culture. *Proceedings of the National Academy of Sciences of the United States of America*, **119**, e2114365119.
- Cao, X., Xu, L., Li, L., Wan, W. & Jiang, J. (2022) TcMYB29a, an ABA-responsive R2R3-MYB transcriptional factor, upregulates Taxol biosynthesis in *Taxus chinensis*. *Frontiers in Plant Science*, **13**, 804593.
- Chen, K., Liu, X.Q., Wang, W.L., Luo, J.G. & Kong, L.Y. (2020) Taxumarienes A–G, seven new alpha-glucosidase inhibitory taxane-diterpenoids from the leaves of *Taxus mairei*. *Bioorganic Chemistry*, **94**, 103400.
- Chen, Y., Tong, S., Jiang, Y., Ai, F., Feng, Y., Zhang, J. *et al.* (2021) Transcriptional landscape of highly lignified poplar stems at single-cell resolution. *Genome Biology*, **22**, 319.
- Chen, Y., Zhang, H., Zhang, M., Zhang, W., Ou, Z., Peng, Z. *et al.* (2021) Salicylic acid-responsive factor TcWRKY33 positively regulates Taxol biosynthesis in *Taxus chinensis* in direct and indirect ways. *Frontiers in Plant Science*, **12**, 697476.
- Chen, Z., Peng, Z., Liu, S., Leng, H., Luo, J., Wang, F. *et al.* (2022) Overexpression of PeNAC122 gene promotes wood formation and tolerance to osmotic stress in poplars. *Physiologia Plantarum*, **174**, e13751.
- Eschrich, A., Cusido, R.M., Bonfill, M., Palazon, J., Sanchez-Munoz, R. & Moyano, E. (2022) The epigenetic regulation in plant specialized metabolism: DNA methylation limits paclitaxel in vitro biotechnological production. *Frontiers in Plant Science*, **13**, 899444.
- Feng, S., Kailin, H., Zhang, H., Chen, C., Huang, J., Wu, Q. *et al.* (2023) Investigation of the role of TmMYB16/123 and their targets (TmMTP1/11) in the tolerance of *Taxus media* to cadmium. *Tree Physiology*, tpad019. Available from: <https://doi.org/10.1093/treephys/tpad019>
- Fett Neto, A.G. & DiCosmo, F. (1992) Distribution and amounts of taxol in different shoot parts of *Taxus cuspidata*. *Planta Medica*, **58**, 464–466.
- Garcia-Cerdan, J.G., Kovacs, L., Toth, T., Kereiche, S., Aseeva, E., Boekema, E.J. *et al.* (2011) The PsbW protein stabilizes the supramolecular organization of photosystem II in higher plants. *The Plant Journal*, **65**, 368–381.
- Ghuge, S.A., Carucci, A., Rodrigues-Pousada, R.A., Tisi, A., Franchi, S., Tavaladoraki, P. *et al.* (2015) The MeJA-inducible copper amine oxidase AtAO1 is expressed in xylem tissue and guard cells. *Plant Signaling & Behavior*, **10**, e1073872.
- Hao, J., Guo, H., Shi, X., Wang, Y., Wan, Q., Song, Y.B. *et al.* (2017) Comparative proteomic analyses of two *Taxus* species (*Taxus × media* and *Taxus mairei*) reveals variations in the metabolisms associated with paclitaxel and other metabolites. *Plant & Cell Physiology*, **58**, 1878–1890.
- He, C.T., Li, Z.L., Zhou, Q., Shen, C., Huang, Y.Y., Mubeen, S. *et al.* (2018) Transcriptome profiling reveals specific patterns of paclitaxel synthesis in a new *Taxus yunnanensis* cultivar. *Plant Physiology & Biochemistry*, **122**, 10–18.
- Hu, W.J., Liu, T.W., Zhu, C.Q., Wu, Q., Chen, L., Lu, H.L. *et al.* (2022) Physiological, proteomic analysis, and calcium-related gene expression reveal *Taxus wallichiana* var. *mairei* adaptability to acid rain stress under various calcium levels. *Frontiers in Plant Science*, **13**, 845107.
- Islam, S., Zeisel, A., Joost, S., La Manno, G., Zajac, P., Kasper, M. *et al.* (2014) Quantitative single-cell RNA-seq with unique molecular identifiers. *Nature Methods*, **11**, 163–166.
- Jervis, J., Hildreth, S.B., Sheng, X., Beers, E.P., Brunner, A.M. & Helm, R.F. (2015) A metabolomic assessment of NAC154 transcription factor overexpression in field grown poplar stem wood. *Phytochemistry*, **115**, 112–120.
- Jin, J., Lu, P., Xu, Y., Tao, J., Li, Z., Wang, S. *et al.* (2022) PCMDB: a curated and comprehensive resource of plant cell markers. *Nucleic Acids Research*, **50**, D1448–D1455.
- Kashani, K., Jalali Javaran, M., Sabet, M.S. & Moieni, A. (2018) Identification of rate-limiting enzymes involved in paclitaxel biosynthesis pathway affected by coronatine and methyl-beta-cyclodextrin in *Taxus baccata* L. cell suspension cultures. *Daru*, **26**, 129–142.
- Kaspera, R. & Croteau, R. (2006) Cytochrome P450 oxygenases of Taxol biosynthesis. *Phytochemistry Reviews*, **5**, 433–444.
- Kuang, X., Sun, S., Wei, J., Li, Y. & Sun, C. (2019) Iso-seq analysis of the *Taxus cuspidata* transcriptome reveals the complexity of Taxol biosynthesis. *BMC Plant Biology*, **19**, 210.
- Liao, W., Zhao, S., Zhang, M., Dong, K., Chen, Y., Fu, C. *et al.* (2017) Transcriptome assembly and systematic identification of novel cytochrome P450s in *Taxus chinensis*. *Frontiers in Plant Science*, **8**, 1468.
- Lin, Q. & Aoyama, T. (2012) Pathways for epidermal cell differentiation via the homeobox gene GLABRA2: update on the roles of the classic regulator. *Journal of Integrative Plant Biology*, **54**, 729–737.
- Liu, Z., Guo, C., Wu, R., Wang, J., Zhou, Y., Yu, X. *et al.* (2022) Identification of the regulators of epidermis development under drought- and salt-stressed conditions by single-cell RNA-seq. *International Journal of Molecular Sciences*, **23**, 2759.
- Liu, Z., Wang, J., Zhou, Y., Zhang, Y., Qin, A., Yu, X. *et al.* (2022) Identification of novel regulators required for early development of vein pattern in the cotyledons by single-cell RNA-sequencing. *The Plant Journal*, **110**, 7–22.
- Marhava, P., Bassukas, A.E.L., Zourelidou, M., Kolb, M., Moret, B., Fastner, A. *et al.* (2018) A molecular rheostat adjusts auxin flux to promote root protophloem differentiation. *Nature*, **558**, 297–300.
- McPartland, T.J., Patil, R.A., Malone, M.F. & Roberts, S.C. (2012) Liquid-liquid extraction for recovery of paclitaxel from plant cell culture: solvent evaluation and use of extracts for partitioning and selectivity. *Biotechnology Progress*, **28**, 990–997.
- Moinul, M., Amin, S.A., Jha, T. & Gayen, S. (2022) Updated chemical scaffolds of ABCG2 inhibitors and their structure-inhibition relationships for future development. *European Journal of Medicinal Chemistry*, **241**, 114628.
- Mubeen, S., Li, Z.L., Huang, Q.M., He, C.T. & Yang, Z.Y. (2018) Comparative transcriptome analysis revealed the tissue-specific accumulations of Taxanes among three experimental lines of *Taxus yunnanensis*. *Journal of Agricultural and Food Chemistry*, **66**, 10410–10420.
- Ohashi-Ito, K., Oguchi, M., Kojima, M., Sakakibara, H. & Fukuda, H. (2013) Auxin-associated initiation of vascular cell differentiation by LONESOME HIGHWAY. *Development*, **140**, 765–769.
- Procko, C., Lee, T., Borsuk, A., Bargmann, B.O.R., Dabi, T., Nery, J.R. *et al.* (2022) Leaf cell-specific and single-cell transcriptional profiling reveals a role for the palisade layer in UV light protection. *Plant Cell*, **34**, 3261–3279.
- Qin, Y., Sun, M., Li, W., Xu, M., Shao, L., Liu, Y. *et al.* (2022) Single-cell RNA-seq reveals fate determination control of an individual fiber cell initiation in cotton (*Gossypium hirsutum*). *Plant Biotechnology Journal*, **20**, 2372–2388.
- Sanchez-Munoz, R., Bonfill, M., Cusido, R.M., Palazon, J. & Moyano, E. (2018) Advances in the regulation of in vitro paclitaxel production: methylation of a Y-patch promoter region alters BAPT gene expression in *Taxus* cell cultures. *Plant & Cell Physiology*, **59**, 255–2267.
- Schafer, H. & Wink, M. (2009) Medicinally important secondary metabolites in recombinant microorganisms or plants: progress in alkaloid biosynthesis. *Biotechnology Journal*, **4**, 1684–1703.
- Schultz, J.A. & Coleman, H.D. (2021) Pectin and Xylan biosynthesis in poplar: implications and opportunities for biofuels production. *Frontiers in Plant Science*, **12**, 712083.
- Shao, F., Wilson, I.W. & Qiu, D. (2021) The research progress of taxol in *Taxus*. *Current Pharmaceutical Biotechnology*, **22**, 360–366.
- Silva, D., Adriana, F., Martins, D.T.O., Borges, Q.I., Lindote, M.V.N., Zoratti, M.T.R. *et al.* (2021) Methanolic extract of *Cariniana rubra* Gardner ex Miers stem bark negatively regulate the leukocyte migration and TNF-alpha and up-regulate the annexin-A1 expression. *Journal of Ethnopharmacology*, **270**, 113778.

- Soliman, S.S., Greenwood, J.S., Bombarely, A., Mueller, L.A., Tsao, R., Mosser, D.D. *et al.* (2015) An endophyte constructs fungicide-containing extracellular barriers for its host plant. *Current Biology*, **25**, 2570–2576.
- Soliman, S.S.M. & Raizada, M.N. (2020) Sites of biosynthesis and storage of Taxol in *Taxus media* (Rehder) plants: mechanism of accumulation. *Phytochemistry*, **175**, 112369.
- Soltwisch, J., Kettling, H., Vens-Cappell, S., Wiegmann, M., Muthing, J. & Dreisewerd, K. (2015) Mass spectrometry imaging with laser-induced postionization. *Science*, **348**, 211–215.
- Sun, G., Yang, Y., Xie, F., Wen, J.F., Wu, J., Wilson, I.W. *et al.* (2013) Deep sequencing reveals transcriptome re-programming of *Taxus × media* cells to the elicitation with methyl jasmonate. *PLoS One*, **8**, e62865.
- Sun, X., Li, C., Ma, J., Zang, Y., Huang, J., Chen, N. *et al.* (2021) New amide alkaloids and carbazole alkaloid from the stems of *Clausena lansium*. *Fitoterapia*, **154**, 104999.
- Tal, I., Zhang, Y., Jorgensen, M.E., Pisanty, O., Barbosa, I.C., Zourelidou, M. *et al.* (2016) The Arabidopsis NPF3 protein is a GA transporter. *Nature Communications*, **7**, 11486.
- Tao, S., Liu, P., Shi, Y., Feng, Y., Gao, J., Chen, L. *et al.* (2022) Single-cell transcriptome and network analyses unveil key transcription factors regulating mesophyll cell development in maize. *Genes*, **13**, 374.
- Theodoulou, F.L. & Kerr, I.D. (2015) ABC transporter research: going strong 40 years on. *Biochemical Society Transactions*, **43**, 1033–1040.
- Torii, K., Inoue, K., Bekki, K., Haraguchi, K., Kubo, M., Kondo, Y. *et al.* (2022) A guiding role of the *Arabidopsis* circadian clock in cell differentiation revealed by time-series single-cell RNA sequencing. *Cell Reports*, **40**, 111059.
- Uniyal, S.K. (2013) Bark removal and population structure of *Taxus wallichiana* Zucc. in a temperate mixed conifer forest of western Himalaya. *Environmental Monitoring and Assessment*, **185**, 2921–2928.
- Wang, H., Zhang, B.Y., Gong, T., Chen, T.J., Chen, J.J., Yang, J.L. *et al.* (2021) Construction of acetyl-CoA and DBAT hybrid metabolic pathway for acetylation of 10-deacetylbaicatin III to baicatin III. *Acta Pharmaceutica Sinica B*, **11**, 3322–3334.
- Wang, Q., Wu, Y., Peng, A., Cui, J., Zhao, M., Pan, Y. *et al.* (2022) Single-cell transcriptome atlas reveals developmental trajectories and a novel metabolic pathway of catechin esters in tea leaves. *Plant Biotechnology Journal*, **20**, 2089–2106.
- Wang, Y., Huan, Q., Li, K. & Qian, W. (2021) Single-cell transcriptome atlas of the leaf and root of rice seedlings. *Journal of Genetics and Genomics*, **48**, 881–898.
- Xiang, L., Wang, F., Bian, Y., Harindintwali, J.D., Wang, Z., Wang, Y. *et al.* (2022) Visualizing the distribution of phthalate esters and plant metabolites in carrot by matrix-assisted laser desorption/ionization imaging mass spectrometry. *Journal of Agricultural and Food Chemistry*, **70**, 15311–15320.
- Xie, B., Wang, X., Zhu, M., Zhang, Z. & Hong, Z. (2011) CalS7 encodes a callose synthase responsible for callose deposition in the phloem. *The Plant Journal*, **65**, 1–14.
- Xiong, X., Gou, J., Liao, Q., Li, Y., Zhou, Q., Bi, G. *et al.* (2021) The *Taxus* genome provides insights into paclitaxel biosynthesis. *Nature Plants*, **7**, 1026–1036.
- Xue, B., Zhao, J., Fan, Y., Chen, S., Li, W., Chen, J. *et al.* (2020) Synthesis of Taxol and docetaxel by using 10-Deacetyl-7-xylosyltaxanes. *Chemistry & Biodiversity*, **17**, e1900631.
- Yadava, U., Gupta, H. & Roychoudhury, M. (2011) Docking studies of two Taxane diterpenoids (10,13-deacetyl-Abeo-Baccatin-iv and 5-acetyl-2-Deacetoxydecinnamoyl-Taxinine-0.29hydrate) with microtubule. *Japanese Journal of Medical Science & Biology*, **2**, 351.
- Yu, C., Guo, H., Zhang, Y., Song, Y., Pi, E., Yu, C. *et al.* (2017) Identification of potential genes that contributed to the variation in the taxoid contents between two *Taxus* species (*Taxus media* and *Taxus mairei*). *Tree Physiology*, **37**, 1659–1671.
- Yu, C., Huang, J., Wu, Q., Zhang, C., Li, X.L., Xu, X. *et al.* (2022) Role of female-predominant MYB39-bHLH13 complex in sexually dimorphic accumulation of taxol in *Taxus media*. *Horticulture Research*, **9**, uhac062.
- Yu, C., Luo, X., Zhan, X., Hao, J., Zhang, L., YB, L.S. *et al.* (2018) Comparative metabolomics reveals the metabolic variations between two endangered *Taxus* species (*T. fuana* and *T. yunnanensis*) in the Himalayas. *BMC Plant Biology*, **18**, 197.
- Yu, C., Luo, X., Zhang, C., Xu, X., Huang, J., Chen, Y. *et al.* (2020) Tissue-specific study across the stem of *Taxus media* identifies a phloem-specific TmMYB3 involved in the transcriptional regulation of paclitaxel biosynthesis. *The Plant Journal*, **103**, 95–110.
- Yu, C., Zhang, C., Xu, X., Huang, J., Chen, Y., Luo, X. *et al.* (2021) Omic analysis of the endangered Taxaceae species *Pseudotaxus chienii* revealed the differences in taxol biosynthesis pathway between *Pseudotaxus* and *Taxus yunnanensis* trees. *BMC Plant Biology*, **21**, 104.
- Zhan, X., Chen, Z., Chen, R. & Shen, C. (2022) Environmental and genetic factors involved in plant protection-associated secondary metabolite biosynthesis pathways. *Frontiers in Plant Science*, **13**, 877304.
- Zhang, M., Li, S., Nie, L., Chen, Q., Xu, X., Yu, L. *et al.* (2015) Two jasmonate-responsive factors, TcERF12 and TcERF15, respectively act as repressor and activator of tasy gene of taxol biosynthesis in *Taxus chinensis*. *Plant Molecular Biology*, **89**, 463–473.
- Zhang, T.Q., Chen, Y. & Wang, J.W. (2021) A single-cell analysis of the Arabidopsis vegetative shoot apex. *Developmental Cell*, **56**(1056–1074), e1058.
- Zhang, T.Q., Xu, Z.G., Shang, G.D. & Wang, J.W. (2019) A single-cell RNA sequencing profiles the developmental landscape of Arabidopsis root. *Molecular Plant*, **12**, 648–660.
- Zheng, Z., Qiu, X., Wu, H., Chang, L., Tang, X., Zou, L. *et al.* (2021) TIPS: trajectory inference of pathway significance through pseudotime comparison for functional assessment of single-cell RNAseq data. *Briefings in Bioinformatics*, **22**, bbab124.
- Zong, J., Wang, L., Zhu, L., Bian, L., Zhang, B., Chen, X. *et al.* (2022) A rice single cell transcriptomic atlas defines the developmental trajectories of rice floret and inflorescence meristems. *The New Phytologist*, **234**, 494–512.