

# Unraveling Developmental Dynamics and Triterpene Saponin Biosynthesis in *Panax notoginseng* Stem Apices by Single-Cell Transcriptomics

**Mei Liu**

Kunming University of Science and Technology

**Lifang Yang**

Kunming University of Science and Technology

**Junda Guo**

Kunming University of Science and Technology

**Hanye Wang**

Kunming University of Science and Technology

**Saiying Yu**

Kunming University of Science and Technology

**Panpan Wang**

Kunming University of Science and Technology

**Ye Yang**

Kunming University of Science and Technology

**Luqi Huang**

China Academy of Chinese Medical Sciences

**Xiuming Cui**

China Academy of Chinese Medical Sciences

**Yuan Liu**

`liuguan513@kust.edu.cn`

Kunming University of Science and Technology

---

## Research Article

**Keywords:** *Panax notoginseng*, stem apices, single cell atlas, developmental trajectory, saponin synthesis

**Posted Date:** May 24th, 2024

**DOI:** <https://doi.org/10.21203/rs.3.rs-4421877/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

**Additional Declarations:** No competing interests reported.

---

# Abstract

## Background

*Panax notoginseng* is often regarded as the foremost blood-nourishing herb due to its rich content of triterpene saponins, which possess various pharmacological activities. The transcriptional regulatory mechanisms governing stem apex development and the expression of its major component, triterpene saponins, remain poorly understood.

## Results

In this study, we constructed a single-cell transcriptome atlas of stem apices, and the cells were divided into seven major cell types, comprising a total of 14 subclusters. Subsequently, we analyzed the developmental trajectory patterns and gene expression features of stem apices and inferred relevant transcription factors associated with stem apex development. We discovered potential key transcription factors, *PTI5* and *IDD12*, involved in the early development of stem apices. Additionally, we investigated the expression and regulation of the triterpene saponin synthesis pathway in stem apices.

## Conclusions

Therefore, the aim of this study is to elucidate the developmental trajectory of *P. notoginseng* stem apices and identify key transcription factors, opening a new avenue for future research of *Panax* genus.

## Background

*Panax notoginseng* (Burk.) F. H. Chen (PN) is a perennial herbaceous plant that belongs to the *Panax* genus and is widely used in traditional Chinese medicine (TCM) [1]. PN contains numerous compounds, such as saponins, flavonoids, polysaccharides, and amino acids, among which triterpenoid saponins are particularly valuable in the medical field [2]. PN is highly valued in TCM as a major raw material for various patent medicines, including Xuesaitong Capsules and Sanqi Tongshu Capsules, owing to its remarkable blood-activating effect. PN's triterpenoid saponins exhibit diverse therapeutic effects, including anti-inflammatory, antioxidant, hepatoprotective, and anticancer properties. Recent research highlights their potential in treating cardiovascular diseases, liver disorders, and cancers [3–5]. By 2020, the annual total output value of industries related to *P. notoginseng* had reached \$15 billion in China [6]. Based on statistical data from China Customs, the export volume of PN ranked second among TCM herbs in Yunnan Province in 2022. Despite primarily utilizing the root for medicinal purposes [7], previous studies have revealed that the stems, leaves, flowers and fruits of PN also possess pharmacological effects [8, 9].

Stem apices play a pivotal role in plant growth and development, particularly in the above-ground parts. Almost all aerial components are derived from the stem apical meristem (SAM) tissue [10]. SAM tissue sustains the presence of stem cells and drives the generation of differentiated cells, with alterations in its cellular state primarily regulating the overall growth dynamics of the tissue [11, 12]. Therefore, a remarkable feature of the SAM tissue is its capacity to continuously generate new organs, thereby providing support throughout the plant's entire lifespan. SAM tissue exhibits a hierarchical structure consisting of three well-defined cell layers. The outermost layer governs the epidermal development of plant stems, leaves, and flowers, ensuring the establishment of protective and structural tissues. The middle layer actively promotes the formation of leaf mesophyll tissue and facilitates germ cell development. The innermost layer orchestrates vascular tissues and pith development, facilitating the transport of water, nutrients, and providing structural support throughout the plant [13–15]. Within the SAM, pivotal and distinctive zones exist, encompassing the central zone as a paramount entity for sustaining the stem cell population, the organizing center is the regulatory hub governing the establishment and maintenance of the stem cell ecological niche, and the peripheral zone coordinates the initiation and advancement of lateral organs [16–19]. The coordinated interplay among these zones within the SAM ensures the growth and development of lateral organs and the maintenance of the stem cell reservoir. Modifications to the cellular organization hierarchy and cellular state at stem apices are critical for regulating the growth dynamics of plants.

In recent years, significant breakthroughs have been achieved in the field of plant research using single-cell transcriptomic (scRNA-seq) technology, which is a powerful tool to study the processes underlying plant growth and development, with *A. thaliana* emerging as a pivotal subject of investigation, with analyses focusing on various tissue types such as roots, stems, and leaves [13, 20–25]. Additionally, attention has been directed to other plants including the cabbage [26, 27], maize [28–31], rice [32–34]. Notably, the roots of *Lotus japonicus* and the stems of *Populus L.* have also garnered research interest [35–37]. Among cash crops, peanut leaves, tea leaves, tobacco corollas, strawberry leaves, cotton ovules and leaves, and *Bombax ceiba* have attracted extensive interest [38–44]. Furthermore, the medicinal plant has placed particular emphasis on the leaves of *Nepeta tenuifolia*, the foliage of *Catharanthus roseus*, and the stems and leaves of *Taxus mairei* [45–48]. These pioneering studies have provided new avenues for the depth exploration of fundamental processes in plant biology, such as organ development and cellular differentiation. However, several factors have hindered the progress of scRNA-seq studies in plants, including species diversity within plants, limitations in protoplast preparation techniques, and scarcity of cell-type marker genes in non-model organisms. To date, research on *Panax* has primarily focused on genomics, transcriptomics, metabolomics, and proteomics. However, no reports of significance scRNA-seq regarding *Panax* have been uncovered, which hinders the progress in expanding our knowledge of *Panax* genus. Therefore, a more comprehensive exploration of PN using scRNA-seq holds promise and presents opportunities.

The objective of this study was to construct a single-cell atlas for PN stem apices, and to explore the growth, development, and transcriptional regulatory of stem apices. Notably, we emphasize the significant role of transcription factors (TFs) in the developmental process of PN stem apices.

Additionally, we investigated the triterpene saponins regulatory mechanisms of biosynthesis and distribution across distinct cell types. These findings offer valuable references and guidance for prospective genetic enhancements and medicinal utilization of PN, establishing a foundation for further exploration and harnessing of the medicinal value of the *Panax* genus.

## Results

# Construction of the comprehensive single-cell atlas for the stem apices of PN

PN seeds harvested in November 2021, despite being morphologically mature but not physiologically mature, germinated successfully after about 40 days using a moist sand bed at room temperature to induce physiological maturation. Stem apices were sectioned for observation of their structure and morphology (Fig. 1A, 2A), followed by the collection of a large number of stem apices for protoplast preparation to increase the proportion of cells in the samples. The 10×Genomics Chromium single-cell platform was used for droplet-based scRNA-seq (Fig. 1A).

The Cell-Ranger software (10×Genomics, v4.0.0) was used for raw sequencing data analysis and PN genome alignment. We obtained a high-quality gene expression matrix comprising 7,987 cells and 29,611 genes (Additional file 1: Table S1). To validate the reliability of the scRNA-seq data, we conducted correlation analysis between the scRNA-seq and bulk RNA sequencing (bulk RNA-seq) datasets, which revealed a robust correlation between the two datasets (Additional file 1: Fig. S1A). After quality control, the top 2,000 variable genes (dims = 30) were selected based on their expression levels. A dimensionality reduction was achieved using uniform manifold approximation and projection (UMAP), coupled with t-distributed stochastic neighbor embedding (t-SNE), thereby facilitating the implementation of cellular clustering analysis (Fig. 1B). This analysis revealed 14 distinctive cell clusters, with each possessing a unique identity and functional relevance (Fig. 1B). Notably, cluster 0 presented the highest cell count at 1,887 cells (20.87% of the total cell population), whereas cluster 13 encompassed only 26 cells (0.33% of the total cell population) (Additional file 1: Fig. S1B). To reveal the identified intracellular cell types, Seurat's FindAllMarkers function was used to conduct differential gene expression analysis and identify differentially expressed genes (DEGs) within each cluster (Additional file 1: Table S2). Because of the limited exploration of cell-type marker genes at the single-cell level in PN, a set of marker genes for distinct cell types was compiled based on existing single-cell research findings from relevant plant species, such as *Arabidopsis*, poplar, and Chinese cabbage (Additional file 1: Table S3). To gain a deeper understanding of the biological processes associated with the distinct cell types, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses of the DEGs in each cell type were performed. We successfully classified the 14 cell clusters into seven distinct cell types: mesophyll cells (MC), epidermal cells (EC), shoot meristematic cells (SMC), proliferating cells (PC), vascular cells (VC), xylem cells (XC), and xylem parenchyma cells (XPC) (Fig. 1B, Additional file 1: Fig.S1G).

ECs were predominant within the cell population and accounted for 32.29% of the total cells (Table S4). This specific cell population was found in clusters 1, 4, 8, 11, and 12 (Fig. 1B and C). ECs express an array of distinct genes, including FIDDLEHEAD (*FDH*), DEFECTIVE IN CUTICULAR RIDGES (*DCR*), *LTP1*, and GLABRA 2 (*GL2*), all of which assumed pivotal functions within the EC differentiation process (Fig. 1C-D). Furthermore, in cluster 2, we observed specific high expression levels of marker genes associated with SMC, such as LIGHT SENSITIVE HYPOCOTYLS 3 (*LSH3*), KNOTTED 1-LIKE HOMEODOMAIN (*KN1*), and *KNAT6*. Notably, MC (cluster 7) exhibited high expression levels of genes linked to photosynthesis, such as Photosystem II subunit R (*PSBR*), Photosystem I subunit D (*PSAD*), Photosystem I subunit F (*PSAF*), and Photosystem II reaction center W (*PSBW*) (Fig. 1C-D). Genes highly expressed in cluster 9 included Walls Are Thin 1 (*WAT1*), which stimulates growth hormone signaling within microtubules, Wuschel-related homeobox 4 (*WOX4*), which promotes the proliferation and differentiation of vascular tissue cells, and Octopus (*OPS*), which actively participates in microtubule differentiation (Fig. 1C). In cluster 3, we discovered highly expressed genes associated with lignin synthesis and ABA signaling, such as Cinnamyl alcohol dehydrogenase 7 (*CAD7*) and Aba repressor1 (*ABR1*), therefore, this cluster is classified as XPCs (Fig. 1, C and D). Furthermore, we observed the expression of XC marker genes, including xylem-specific papain-like cysteine peptidase 1 (*XCP1*), Cellulose synthase 8 (*CESA8*), and Laccase 11 (*LAC11*), predominantly in cluster 13. PCs (clusters 5, 6, and 10) expressed marker genes linked to cell cycle regulation, including *KN*, High mobility group A (*HMGA*), High-mobility group box 6 (*HMGB6*), and Cyclin-dependent kinase B2 (*CDKB2-2*) (Fig. 1C-D). The majority of cells are in the G1 and S phases of the cell cycle (Additional file 1: Fig. S1 C-F). Stem apex cells exhibit increased metabolic activity, resulting in the rapid synthesis of RNA and proteins, leading to a significant increase in cell volume. Furthermore, we observed that cells in the G1 phase display strong division activity and constitute a specific proportion of the total cell population. We examined the cell cycle genes across different cell types, and our findings validate the distribution of our cell cycle genes. However, it is noteworthy that the PC cell type exhibits the highest count in the S phase (Additional file 1: Fig. S1F). To validate the identified cell types, we visualized the screened marker genes, compared the enriched marker genes in each cell cluster (Additional file 1: Fig. S1H), and performed RNA *in situ* hybridization to confirm the marker genes in MC, XC, and XPC (Fig. 2B). However, cluster 0 could not be classified based on a known cell type because its identity remained elusive according to known marker genes.

After GO and KEGG analyses, each cell type exhibited pronounced functional specificity. ECs are related to cell membranes and walls, amino and nucleotide sugar metabolism. MCs are closely associated with functions predominantly centered on photosynthesis and chloroplasts. VCs were linked to the cytoskeleton and cell motility, including cytoskeleton construction and the *MAPK* signaling pathway (Additional file 1: Fig. S2, S3, Table S5).

## Differentiation trajectory of PN stem apices

We re-clustered the PC cells, which were classified into six distinct subcellular clusters (Fig. 3A). We identified specific EC marker genes, namely protodermal factor1 (*PDF1*), *LTP1*, and *DCR*, within

subcellular clusters 2, 3, and 5 (Fig. 3B-C). In subcellular cluster 4, we observed predominant expression of photosynthesis-related genes, such as *PSBC*, *PSBR*, and *PSBA* (Fig. 3B), which are necessary to sustain normal photosynthesis and facilitate light energy conversion. Conversely, genes such as *Arabidopsis* homeobox 8 (*ATHB-8*), an early marker of development associated with VC, and *WAT1* exhibited relatively low expression within the PC subpopulation.

To explore the intricate path of differentiation in PN stem apex cells, we performed cell mimetic timing analyses to visualize the distribution of mimetic timing across all cell clusters (Fig. 3D- F, Additional file 1: Fig. S4A). The results showed the existence of two pivotal branching points and five distinct stages of stem apex cell differentiation (Fig. 3D). EC populations were constant, indicating their importance in the development of stem apices. Using a cell-mimetic time series, we identified high expression levels of microtubule organization-associated genes specific to VC, XC, and XPC in stage 4 compared to other stages. Furthermore, our analyses showed a predominance of UN in stages 1 and 3 (Fig. 3E, Additional file 1: Fig. S4A), indicating a potential role in stem apex development. We selected the top 20 DEGs in each cell type and categorized them into three distinct modules (Fig. 3G). In the early stages of the proposed cellular chronology, these genes predominantly participated in photosynthesis and protein complexes and facilitated vital biological processes, such as cell division, tissue growth, cell signaling, and metabolism. The intermediate module likely contributes to cell membrane formation during plant growth and development, and signaling events and lipid metabolism regulated by phytohormones. The final module centered on metabolic processes highlights the finely orchestrated regulation of gene expression at this stage, thus ensuring precise and appropriate growth and developmental processes. Additionally, by examining the expression of marker genes in each cell type along the cell differentiation trajectory (Additional file 1: Fig. S4B), we validated the identity and functional roles of these cell types.

Our results revealed the sequential developmental trajectory of PN stem apex cells, which are characterized by the successive differentiation of epidermal, chloroplastic, and vascular cell populations. We performed RNA velocity calculations for distinct cell types using Velocyto and incorporated RNA velocities into UMAP analysis (Fig. 3H). The results revealed increased RNA velocities in the EC, MC, SMC, PC, and VC populations, which is indicative of the rapid and dynamic differentiation processes occurring within these cells (Additional file 1: Fig. S4C-D).

## Transcription factor co-expression networks

We analyzed the expression patterns of identified transcription factors across various cell types by heatmap, highlighting the specific expression of these TFs in different cell types (Fig. 4A). Subsequently, we identified a significant set of TF families, including MYB, AP2, FAR1, bHLH, C2H2, NAC, HB, C2C2, C3H, and GRAS (Fig. 4B). Further investigations revealed the gene expression of these TF families, such as MYB, AP2, FAR1, bHLH, and C2H2. It is worth noting that these differences were particularly pronounced in cell types associated with microtubules, such as VC, XPC, and XC (Fig. 4C, Additional file 1: Fig. S5A). Subsequently, we conducted a KEGG analysis of DEGs and discovered a significant pathway that encompasses the ribosome, lipid biosynthesis proteins, and glycosyltransferases (Additional file 1:

Fig. S5B). We identified several key TF families, such as the ERF, C2H2, ARF, DOF, and the YABBY families (Fig. 4D).

Based on the analysis of the TF-DEG network, we found a close association between *PTI5* in the ERF transcription factor lineage and the indeterminate-domain 12 (*IDD12*) in the C2H2 transcription factor family, with the regulation of pivotal genes throughout the growth and developmental stages of PN stem apices. Furthermore, we observed that during seed germination under conditions of moisture and darkness, both *PTI5* and its closely linked counterpart *FAD2*, recognized for their role in enhancing plant cold hardiness, were induced in the emerging EC (Fig. 4E, Additional file 1: Fig. S5C). Pseudo-timing analysis revealed that there is a consistent developmental trajectory between *PTI5* and EC, while *IDD12* aligns with MC (Additional file 1: Fig. S5C). We used RNA *in situ* hybridization to determine the expression patterns of *PTI5*, *IDD12*, and *FAD2*, and the results showed predominant expression signals of *PTI5* and *FAD2* within EC (Fig. 2C, Additional file 1: Fig. S5C), whereas *IDD12* exhibited primary expression signals within MC (Fig. 2C, Additional file 1: Fig. S5C).

### Cytotype-specific gene expression in PN and *A. thaliana*

We investigated the analogies and discrepancies in the developmental processes of stem apices in two dicotyledonous plants, PN and *A. thaliana*. Initially, we obtained single-cell transcriptome information from *A. thaliana* stem apices. Our analysis validated the accuracy of our approach by confirming a cell type distribution consistent with previous studies (Fig. 5A and Fig. S6 A-C). We merged the transcriptome data of PN and *A. thaliana* for visualization using UMAP plots (Fig. 5B) and computed correlations between cell types (Fig. 5C). We observed significant correlations among the various cell types in the stem apices of both species. Notably, the MC (0.82), PC (0.86), and VC (0.73) clusters exhibited similar spatial distributions in the UMAP plot (Fig. 5B-C). We employed a Sankey diagram to illustrate the transitions between the shared cell types in PN and *A. thaliana* (Fig. 5D). The diagrams suggested highly aligned transcriptomic profiles. We observed a noteworthy correlation between the SMC of PN and the VC of *A. thaliana*, the Knotted1-like homeobox (KNOX) proteins play a central role in orchestrating the maintenance of the pluripotent cell population at the shoot apical meristem in the realm of plants. We also observed that KNOX homologues (KNAT6) were expressed in both cell types (Fig. 1D).

## Cell communication analysis of PN

In the analysis, we identified a total of 1640 ligand-receptor combinations, including 264 pairs of autocrine and 1403 pairs of paracrine ligand-receptor pairs. We constructed a cell-cell communication network between different cell types (Fig. 5E). The results revealed 58 significant ligand-receptor pairs within the entire cell communication network, with VC-VC pairs being prominent. Importantly, VC-XPC, VC-SMC, and VC-MC were identified as the major cell types expressing ligands (Fig. 5E-F). The top10 ligand-receptor pairs showed that in most cell pairs, there were ligand-receptor pairs like MPK3 – AT5G24010, UBQ11 – AT5G24010, and UBQ3 – AT5G24010. However, in SMC – XC and VC-XC cell-cell pairs, we observed the ligand-receptor pair EXL2 – XCP1 (Additional file 1: Fig. S6D). This suggests that

there may be differences in the regulation of ligand-receptor pairs among different cell-cell pairs. In conclusion, VC plays a significant role in intercellular interactions, which is important for our understanding of cell communication and signal transduction in PN stem apices.

## Transcriptional regulation of saponin biosynthesis

We aimed to determine the cellular specificity of triterpene saponin biosynthesis in the apices to better understand this metabolic pathway. Our study identified 19 crucial gene families involved in the upstream processes of the triterpene saponin synthesis pathway, which comprised genes in the mevalonate (MVA) pathway, such as ACAT, HMGS, HMGR, MVK, PMK, and MVD2; elements of the plastidial methylerythritol 4-phosphate (MEP) pathway, including DXS, DXR, CMS, CMK, MCS, HDS, and HDR; and participants in the 2,3-oxidosqualene biosynthetic pathway, such as IDI, GGPS, GGR, FPS, SS, and SQE (Additional file 1: Table S6). We observed their widespread expression across all cellular compartments within the stem apices and evaluated the expression levels of the genes associated with saponin biosynthesis in major cellular compartments within the stem apices (EC, MC, VC, and PC) (Fig. 6A). Notably, the relative expression of these genes was elevated in EC and MC. Thus, within the stem apex regions, the pivotal genes driving saponin biosynthesis primarily exerted their influence on epidermal and mesophyll cells. Furthermore, we observed the overall distribution of genes involved in the triterpene saponin biosynthetic pathway and the developmental trajectory of these genes in a pseudo-temporal manner (Additional file 1: Fig. S7). Therefore, while stem apices may not serve as the central hub for triterpene saponin synthesis, they nevertheless participate in certain aspects of this critical metabolic pathway, thereby providing a fundamental synthetic capacity for PN stem apices. Moreover, we constructed a co-expression network for triterpene saponins in the stem apices. Our results showed that the transcription factor families closely associated with the triterpene saponin biosynthesis pathway included the NAC, bHLH, MYB, ERF transcription factors (Fig. 6B).

## Discussion

### Generating the developmental cell atlas of stem apices in *Panax* genus

PN is a medicinal plant used in TCM that has received attention because of its therapeutic potential. However, current research on PN primarily focuses on its roots, stems, and leaves. Nonetheless, the developmental dynamics and gene regulation of apical meristematic tissue and aboveground tissue maturation in PN remain unknown. scRNA-seq technology, which encompasses high-resolution gene expression analysis and reveals cellular heterogeneity, is better than bulk RNA sequencing methods [49, 50]. In this study, we used scRNA-seq technology to curate a single-cell transcriptome atlas encompassing the early stages of PN stem development. We identified seven cell types pivotal to early stem growth and morphogenesis, and found their expression patterns. Interestingly, these expression patterns resembled those observed in analogous cell types within model plants such as *Arabidopsis* [13], *Brassica* [27], and *Populus* [36]. Encompassing SMC, MC, EC, PC, VC, and other cell types, this suggests a general similarity in cell type composition across aerial plant parts. Further investigation revealed similarities in structure and expression patterns of SMC, MC, EC, PC, VC, and other cell types between

PN and *Arabidopsis* at the single-cell transcriptome level, exhibiting a neighboring distribution in the cell spectrum. The expression of KNOX homologs in SMC and VC cell types is noteworthy, and previous studies have reported the expression of KNOX genes in these cell types in model plants such as poplar [51]. This observation is consistent with our findings, which revealed the expression of KNAT6 in SMC and VC cell types. The conservation of KNOX protein expression patterns across different plant species suggests a fundamental role in meristematic activities and tissue differentiation during plant development. This finding underscores the significance of KNOX proteins as conserved regulators of plant growth and development, with potential implications for understanding the molecular mechanisms underlying meristem function and secondary growth regulation. This similarity suggests the conservation of functions and regulatory mechanisms across different plant species. Also, we validated the results through RNA *in situ* hybridization experiments, thus providing a valuable reference for annotating the cells of *Panax* genus medicinal plants. Remarkably, most of the enriched GO and KEGG results were aligned with the expected physiological functions of the respective cell types, indicating that these cells execute specific and crucial functional roles in the dynamic orchestration of plant growth and development. Our results underscore the divergent levels of single-cell transcriptional regulation among different cell clusters, confirming that PN stem apices exhibit cell heterogeneity.

We elucidated the intricate developmental trajectories of pivotal cell types intimately associated with the growth and maturation of the PN stems. This indicates the importance of these cell types in the overall development of PN stems. Moreover, we observed a substantial increase in the expression of MC- and EC-related genes during the initial phases of stem apex development. In contrast, VC-related and growth hormone-associated genes manifested progressively in later stages, consistent with previous research findings [39], supporting the pivotal role of these genes during the initial developmental stages of PN stem apices. Using cell-mimetic time series we identified high expression levels of microtubule organization-associated genes specific to VC, XC, and XPC in stage 4 compared to other stages (Additional file 1: Fig. S4A). This observation suggests gradual and incremental formation of vascular tissues through the development of these cell types during the maturation of stem apices.

Considering the limitations of this study, stem apices from 40-day-old germinated plants were selected for sampling due to the current sequencing technology's limitation to this specific period. While samples from this period can provide valuable information, there are still limitations. Therefore, future research could consider advancing sequencing technologies and utilizing samples from more time points to gain a more comprehensive understanding.

## Unveiling the regulatory network of PN stem apices by analyzing the co-expression network of TF-DEGs

Transcription factors play an important role in fundamental biological processes [52]. Our study successfully elucidated the crucial functions of transcription factor families in plant growth and development. We constructed a network that interconnected transcription factors and DEGs (Fig. 4D).

Among these families are the ERF family, intricately associated with ethylene signaling [53]; the C2H2 family, actively involved in developmental processes and secondary metabolism[54]; the IAA and ARF families, responsible for the response to growth hormones [55]; the DOF family, engaged in plant growth, metabolism, and hormone signaling [56]; and the YABBY family, which actively participates in morphogenesis, development, and phytohormone regulation [57]. In the TF-DEG network, it was observed that the YABBY, HD, C2H2, and ARF families collectively regulated 20 differentially expressed genes (DEGs), while the DOF, IAA, MYB-related, bZIP, and G2-like families collectively regulated 8 DEGs. Specifically, major regulatory transcription factors such as PTI5, IDD12, AFR, and YAB4 played critical roles in the network. Among them, PTI5 and IDD2 respectively regulated 19 DEGs (Fig. 4D). These results highlight the importance of transcription factor families in regulating plant growth and development, providing a foundation for further research. The TF-DEG co-expression network construction revealed a significant regulatory impact of these TF families on PN seed stratification and germination, displaying close associations with phytohormone interactions, particularly the role of the ethylene response factor *PTI5* in primary regulation, in the ethylene signaling pathway [58, 59]. This suggests a pivotal role for this pathway in cell differentiation, environmental adaptation, and overall developmental processes. Previous studies have revealed the significant role of *PTI5*, a TF that intricately controls several defense-related genes through its response to salicylic acid regulation and ethylene signaling pathways. As a transcriptional activator, it controls the expression of genes associated with ripening through the ethylene signaling pathway [58]. These findings provide valuable insights into the function of PTI5 in the growth and development of PN stem apices.

Furthermore, we found that *PTI5* and *FAD2* exhibit co-expression patterns during PN seed stratification and germination. we found notable expression variations in these genes that are closely linked to the regulation of seed development. Specifically, we observed an increase in *FAD2* expression during the mid-seed maturation stage, indicating its involvement in developmental regulation. *FAD2* increases the dienophilic fatty acid content, which is responsible for cold tolerance during the propagation stage [60, 61]. This indicates that *FAD2* may be linked to mechanisms involved in fatty acid synthesis and environmental adaptation during stem apex development. We observed that during seed germination under conditions of moisture and darkness, both *PTI5* and its closely linked counterpart *FAD2* were induced within the emerging EC, and their expression distributions during the late stage of seed germination were also similar (Fig. 4E, Additional file 1: Fig. S5C). This suggests that *PTI5* may participate in regulating the processes of seed germination, subsequent growth, and development of PN by modulating the expression of the *FAD2* gene within EC. These results support further investigations aimed at understanding the role of *PTI5* and *FAD2* in the growth and development of PN stem apices.

Additionally, we determined the IDD TF subfamily is important for the control of phytohormones during plant development, influencing the spatial accumulation of growth hormones, morphogenesis of plant organs, and gravitropic response [62]. IDD12 was demonstrated an additive effect on rice MDPK activation while simultaneously promoting sheath blight (ShB) resistance, thereby highlighting the significance of IDD12 in primary regulation [63]. Furthermore, our findings demonstrate that the ribosomal protein (RP) gene family associated with the IDD12 transcription factor subfamily is

predominantly located in MC (Additional file 1: Fig. S5C). Through pseudo-timing analysis, we consistently observed a developmental trajectory in *PTI5* and EC, whereas *IDD12* aligned with MC (Fig. 4C, Additional file 1: Fig. S5C). This suggests that *PTI5* and *IDD12* may have cell-specific regulatory functions in the growth and development of PN stem apices. RNA *in situ* hybridization was conducted to examine the expression patterns of *PTI5*, *IDD12*, and *FAD2*. The results shed light on the potential roles of *PTI5* and *IDD12* as transcriptional regulators during the early stages of seed germination in PN growth and development. These discoveries not only enhance our understanding of the functions of *PTI5* and *IDD12* but also deepen our insights into the early growth dynamics and developmental aspects of PN stem apices.

### **Expression of genes related to secondary metabolite biosynthesis at the single-cell level during the developmental process**

Our research explored the saponin biosynthesis process in PN's stem apices. The results revealed enhanced saponin-related gene activity in the stem apex cells, mostly in MC and EC types, implying the crucial role of triterpene saponins in early plant development. Additionally, we discovered transcription factors regulating triterpenoid saponin biosynthesis, corroborating prior findings [64]. These findings serve as the foundation for a comprehensive analysis of the pathway governing the synthesis of newly formed saponins in PN. Although stem apices are the site of plant growth, the extent of research dedicated to the stem apices of medicinal plant remains relatively limited. Nonetheless, triterpene saponins, the primary active component in PN, are abundant throughout various growth stages and tissue locations [65].

We discovered that TF families closely associated with the triterpene saponin biosynthesis pathway included the extensively studied NAC TFs (which are involved in governing secondary metabolism in plants) [66]. Simultaneously, the bHLH transcription factor family is a pivotal regulatory position in medicinal plant biosynthesis [67], while the MYB transcription factor family is instrumental in orchestrating gene expression regulations within secondary metabolic pathways of plants [68]. Furthermore, it is plausible that ERF family transcription factors also hold substantial functional significance regarding triterpenoid saponin synthesis in PN [69]. Triterpene saponins perform crucial regulatory functions throughout plant secondary metabolic processes, and are involved in overseeing diverse biological processes or mediating signaling pathways. Interconnection analysis between transcription factors and triterpene saponin biosynthesis pathways specifically present within PN stem apices revealed a network of stem apices-derived triterpene saponins. This result gives us insights for a deeper understanding of the intricate mechanisms underlying PN triterpene saponin synthesis. Future research will require that the functionalities and regulatory mechanisms surrounding triterpene saponins optimize the triterpene saponin content within PN plants using molecular design breeding, enhancing the medicinal value of PN.

## **Conclusions**

In this study, we used scRNA-seq technology to analyze the stem apices of the medicinal plant PN. This allowed us to construct a comprehensive single-cell atlas of PN stem apices and develop cellular developmental trajectories. Importantly, our findings demonstrate the critical roles of the transcription factors PTI5 and IDD12 in the development of PN stem apices. Additionally, we identified the families of transcription factors involved in triterpene saponin biosynthesis and regulation. These findings provide valuable insights for future investigations into the biosynthesis pathways of saponins in PN stem apices and deepen our understanding of plant growth and secondary metabolism mechanisms. Overall, our research contributes to the elucidation of biological processes in plants and offers useful references for genetic enhancement and exploration of the medicinal value of PN. Further research can expand on the functionalities and applications of these findings, propelling further development and applications in the field of botany.

## Methods

### PN shoot tip growth conditions and protoplast isolation

PN seeds were harvested in Wenshan, Yunnan Province, People's Republic of China (latitude N23°23'20.26", longitude E104°13'49.49", elevation: 1258.89 m). Subsequently, the crimson pericarp enveloping the seeds was removed, and then the seeds were placed in a bed of moist sand to facilitate the stratified process of germination. Upon germination, nascent stems situated between the cotyledons and embryonic roots were collected carefully. Initially, we performed paraffin sectioning of the stem apices and toluidine blue staining to facilitate a more accurate analysis of the characteristics and tissue structure. Subsequently, the tender stems were placed carefully in test tubes containing an enzymatic solution consisting of 2% cellulase R10, 1% macerozyme R10, 0.4 M mannitol, 20 mM KCl, 20 mM MES, 10 mM CaCl<sub>2</sub>, 5 mM-Mercaptoethanol, and 0.1% BSA. Enzymatic digestion proceeded under dark conditions for 4 h. Following digestion, the enzyme solution was diluted with W5 solution (containing 154 mM NaCl, 125 mM CaCl<sub>2</sub>, 5 mM KCl, and 2 mM MES) to facilitate filtration-based elimination of protoplasts. After removing the supernatant, the protoplasts were rinsed three times and resuspended. The protoplast concentration and viability were subsequently evaluated using trypan blue staining to ensure the selection of superior-quality protoplasts for the construction of the single-cell RNA-seq library.

### Bulk and scRNA-seq of PN stem apices

This study employed the 10× Chromium Single Cell 3' platform (10× Genomics, Pleasanton, CA, USA) to construct scRNA-seq libraries according to the manufacturer's protocol. Specifically, protoplasts isolated from stem apices were loaded onto Chromium microfluidic chips using 3'v2 chemistry and barcoded using a 10× Chromium Controller to generate concentrated single-cell GEMs, which were enclosed within latex beads. These barcoded cells were reverse transcribed and scRNA-seq libraries were generated using the Chromium Single Cell 3' Reagent Kit v2 (10× Genomics). Subsequently, Illumina sequencing was conducted using a HiSeq 2000 platform (San Diego, CA 92121, USA) in strict accordance with the manufacturer's instructions.

For bulk RNA-seq, a substantial number of stem apices located between the cotyledons and embryonic roots were collected using the previously described procedure (Fig. 1A). Total RNA was extracted from these samples, the NEBNext Ultra RNA Library Prep Kit for Illumina (NEB, USA) was used to construct sequencing libraries, and sequencing was performed using the HiSeq 2000 platform (San Diego, CA 92121, USA). The sequencing data were quality-controlled and cleaned using Fastp software (v0.23.2) [70] to ensure high-quality reads. Subsequently, STAR software (2.1.0) [71] was used to align the sequencing reads to the PN genome [64], thus enabling the generation of precise expression matrices.

Finally, R (<http://www.r-project.org/>) software was employed to calculate Spearman's correlation coefficients to evaluate the correlations between the RNA-seq data from the stem apex samples and the scRNA-seq datasets.

### **RNA In situ hybridization**

The same scRNA-seq samples were collected and fixed in a formaldehyde-acetic acid-ethanol fixative (formalin, glacial acetic acid, and 50% ethanol) as previously described. Following this, we designed probe sequences that included all genes outlined in Additional file 1: Table S7. The samples were then embedded, sectioned, and digested with proteinase K at 37°C. The samples were rinsed with pure water and washed with PBS, and hybridization solution containing the probes was added. After treatment with blocking solution, anti-DIG-AP was added, incubated, and washed several times with TBS. We added the BCIP/NBT color development solution to visually observe the positive expression under a microscope (NIKON ECLIPSE CI, Japan). Once the color developed, we rinsed the samples with pure water again before encapsulating them in glycerol gelatin. We then observed these samples under the microscope and captured images for further analysis and interpretation.

## **scRNA-seq data pre-processing**

Single-cell RNA sequencing reads were demultiplexed using Cell Ranger (v4.0.0, <https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/what-is-cell-ranger>) and aligned to the PN genome [64]. Gene expression quantification and expression matrices were generated using the "cellranger mkref" and "cellranger count" commands, respectively. Subsequently, the scRNA-seq data were preprocessed using the Seurat package (v4.3.0, <https://satijalab.org/seurat/>) as described previously. DoubletFinder (v2.0.3) [72] was employed to identify potential doublets. The default artificial doublet rate (pN) was set at 0.25, and the optimal neighborhood size (pK) was determined using the "paramSweep\_v3" function with "PC = 1:30" following the guidelines provided by the authors. Among the tested parameters, we selected the largest pK value because it yielded the most favorable results. To account for potential doublet rates, we adjusted the nExp value based on the proportion of homozygous doublets and observed doublet rate. Following the determination of nExp, pN, and pK, we applied the "doubletFinder\_V3" function to the preprocessed Seurat data to predict and remove doublets from further analysis. To assess the reliability of the remaining single cells, we computed the artificial neighbor ratio (pANN) for each cell. The pANN

threshold was determined based on the expected number of doublets (nExp), thereby ensuring that only bona fide single cells were retained for subsequent analyses.

## Cell clustering and identification of cluster-enriched genes

Based on our previous data processing steps, we employed a series of filtering criteria to remove low-quality cells and genes, ultimately selecting only those genes expressed in at least three cells, and limiting the number of expressed genes per cell to within the 200–9500 range. Additionally, we eliminated cells with UMI counts < 500 and mitochondrial read percentages  $\leq 5\%$ . To standardize gene expression measurements, we used the "NormalizeData" function to scale the total expression of each gene in the corresponding cell by a factor of 10000. We then utilized the "FindVariableFeatures" function to identify 2000 variable features through the "vst" method and subsequently reduced data dimensionality via principal component analysis (PCA), where the top 30 principal components were determined using the "ElbowPplot" function. Following the clustering of cells using the "FindNeighbors" and "FindClusters" functions, we used non-linear dimensionality reduction algorithms ("RunTSNE" and "RunUMAP") to visualize the data. We used the plotly package (v. 4.10.2) to create a 3D UMAP embedding. By using the "RunUMAP" function with the parameter (n.components = 3L), we generated a 3D UMAP scatter plot with three UMAP dimensions: 'UMAP\_1', 'UMAP\_2', and 'UMAP\_3'. Utilizing the "FindAllMarkers" function in the Seurat R package, we calculated the spatial distribution of DEGs and filtered out significantly enriched marker genes in each cluster ( $PCT.1 \geq 0.25$ ,  $PCT.2 \leq 0.1$ ). Finally, we conducted GO and KEGG pathway enrichment analyses of the DEGs using R based on hypergeometric distribution.

## Pseudo-time analysis

We conducted a comprehensive analysis of the cellular developmental trajectories using the Monocle package (v2.28.0) [73] to explore cellular maturation and progression. Initially, we transformed the original integrated Seurat object into a CellDataSet object, which ensured the seamless integration of the data with Monocle's analytical framework. Subsequently, we employed the DDRTree method, a robust dimensionality-reduction technique, to capture the underlying structure of the dataset. The "orderCells" function was then used to accurately deduce the developmental trajectories of individual cells and reveal the sequential progression and lineage relationships. To visually represent these trajectories, we employed the "plot\_cell\_trajectory" function, which yielded visually insightful plots. Additionally, we conducted a GO enrichment analysis to elucidate the functional implications of specific genes within the identified trajectories, thereby revealing the biological processes and pathways that drive cellular differentiation and maturation.

## RNA velocity analysis

RNA velocity analysis was performed using the Velocyto tool (v 0.6) to process BAM files acquired from CellRanger [74]. This advanced software facilitated the quantification of spliced and un-spliced RNA read matrices for each gene in individual cells and provided the data required to calculate RNA velocity values using default parameters. Additionally, we projected the RNA velocity vectors onto UMAP space, thereby

enabling the visualization and interpretation of the data. The application of RNA velocity analysis provided important insights into the intricate dynamics and directionality of gene expression changes and valuable information regarding the developmental trajectories and regulatory mechanisms within the investigated cells.

### **Integrated analysis of scRNA-seq data from PN and *A. thaliana* stem apices**

Single-cell transcriptome data of *A. thaliana* downloaded from *A. thaliana* stem apices databases [13] and subjected to assembly quantification using CellRanger were compared with those of PN stem apices from our experiment. To combine both datasets, we used the FindIntegrationAnchors function, a tool that defines anchors between the *A. thaliana* and PN datasets. Subsequently, the IntegrateData function was applied to these anchor points with additional default parameters to merge the datasets. We then preprocessed and downscaled the integrated data using ScaleData and RunPCA (dim = 1:20) operations. To gain insights into the cell types of *A. thaliana* and PN stem apices, we implemented the UMAP downscaling technique along with Shared Nearest Neighbor (SNN) plot construction, utilizing UMAP outcomes and default PCA approximation. Cluster analysis was performed on our integrated assay at a clustering resolution of 0.7, which helped to classify the cells into distinct clusters. Finally, to assess cell type similarities between PN and *A. thaliana*, we used the MetaNeighbor-predicted area under the curve scores as an indicator, with higher scores signifying greater similarity [75].

## **Cell Communication Analysis**

Utilizing the *A. thaliana* data in PlantPhoneDB (<https://jasonxu.shinyapps.io/PlantPhoneDB/>)[76], we constructed ligand-receptor pairs for *P. notoginseng* to explore how PN stem apices identify signaling molecules at single-cell resolution. Cell-cell interactions were deduced based on the expression of ligand-receptor pairs in different cell types, and we selected ligand-receptor pairs with significant cell communication (p-values < 0.05) through significance calculations.

## **scRNA-seq co-expression analysis**

To discern TFs within scRNA-seq data, we conducted a comparative analysis using the iTAK database, an invaluable resource for TF identification and classification [77]. Within the R environment, we computed the Pearson correlation coefficients between each gene and TF using the "cor.test" function, which identified genes exhibiting similar expression patterns. Subsequently, we imported TFs and DEGs into Cytoscape (v3.9.0)[78] and constructed a TF-DEG co-expression network via scRNA-seq co-expression analysis. This intricate network provides important insights into the potential regulatory relationships between TFs and DEGs within the context of cellular development. Moreover, we extended our analysis to include a network of genes and TFs associated with important enzymes in the saponin biosynthesis pathway, thereby facilitating a comprehensive understanding of the regulatory network governing saponin biosynthesis. This integrated approach enabled us to reveal the potential TF-mediated regulatory mechanisms underlying saponin biosynthesis in PN.

## **Declarations**

## Acknowledgments

We thank the editor and anonymous reviewers for providing valuable comments on the manuscript. YL acknowledges financial support from the Major Science and Technology Special Project of Yunnan Province (Grant No. 202202AG050021) and National Science Foundation of China (Grant No. 32360151). The Ability Establishment of Sustainable Use for Valuable Chinese Medicine Resources (Grant No. 2060302).

## Authors' contributions

M.L. and Y.L. wrote and reviewed the manuscript, M.L., L.F.Y., J.D.G., H.Y.W., S.Y.Y., and P.P.W. collected the samples, Y.L., and Y.Y. provided analysis ideas and reviewed the manuscript. M.L. and Y.L. reviewed the manuscript and prepared the illustrations, L.Q.H, X.M.C. and Y.L. supervised the work. All authors have read and approved the final paper.

## Funding

YL acknowledges the financial support from the Major Science and Technology Special Project of Yunnan Province (Grant No. 202202AG050021) and National Science Foundation of China (Grant No. 32360151, 31960134), and National Key R&D Program (Grant No. 2023YFC3503802)

## Availability of data and materials

The scRNA-seq and RNA-seq data have been deposited at the China National GeneBank (CNGBdb, <https://db.cngb.org>, Number: CNP0004665) and are publicly available as of the date of publication.

## Ethics approval and consent to participate

Not applicable.

## Consent for publication

Not applicable.

## Competing interests

The authors declare that they have no competing interests.

## References

1. Lee S, Reddy CK, Ryu JJ, Kyung S, Lim Y, Park MS, Kang S, Lee CH. Solid-State Fermentation With *Aspergillus cristatus* Enhances the Protopanaxadiol- and Protopanaxatriol-Associated Skin Anti-aging Activity of *Panax notoginseng*. *Front Microbiol.* 2021, 12:602135.<http://doi.org/10.3389/fmicb.2021.602135>.

2. Gao M, Cao X, Wei S, Huang X, Ouyang H, Chang Y, Shi R, He J. Quantitative Comparison and Chemical Profile of Different Botanical Parts of *Panax notoginseng* From Different Regions. *Front Nutr.* 2022, 9:841541.<http://doi.org/10.3389/fnut.2022.841541>.
3. Zhang E-Y, Gao B, Shi H-L, Huang L-F, Yang L, Wu X-J, Wang Z-T. 20(S)-Protopanaxadiol enhances angiogenesis via HIF-1 $\alpha$ -mediated VEGF secretion by activating p70S6 kinase and benefits wound healing in genetically diabetic mice. *Experimental & Molecular Medicine.* 2017, 49:e387-e87.<http://doi.org/10.1038/emm.2017.151>.
4. Luo F-C, Wang S-D, Qi L, Song J-Y, Lv T, Bai J. Protective effect of panaxatriol saponins extracted from *Panax notoginseng* against MPTP-induced neurotoxicity in vivo. *Journal of Ethnopharmacology.* 2011, 133:448 – 53.<http://doi.org/https://doi.org/10.1016/j.jep.2010.10.017>.
5. Xia L, Liu X, Mao W, Guo Y, Huang J, Hu Y, Jin L, Liu X, Fu H, Du Y, Shou Q. *Panax notoginseng* saponins normalises tumour blood vessels by inhibiting EphA2 gene expression to modulate the tumour microenvironment of breast cancer. *Phytomedicine.* 2023, 114:154787.<http://doi.org/https://doi.org/10.1016/j.phymed.2023.154787>.
6. Yang Y, Zheng K, Guo LP, Wang CX, Zhong DB, Shang L, Nian HJ, Cui XM, Huang SJ. Rapid determination and dietary intake risk assessment of 249 pesticide residues in *Panax notoginseng*. *Ecotoxicol Environ Saf.* 2022, 233:113348.<http://doi.org/10.1016/j.ecoenv.2022.113348>.
7. Bai Y, Liu H, Pan J, Zhang S, Guo Y, Xian Y, Sun Z, Zhang Z. Transcriptomics and Metabolomics Changes Triggered by Inflorescence Removal in *Panax notoginseng* (Burk.). *Front Plant Sci.* 2021, 12:761821.<http://doi.org/10.3389/fpls.2021.761821>.
8. Xiong Y, Chen L, Man J, Hu Y, Cui X. Chemical and bioactive comparison of *Panax notoginseng* root and rhizome in raw and steamed forms. *J Ginseng Res.* 2019, 43:385 – 93.<http://doi.org/10.1016/j.jgr.2017.11.004>.
9. Shi R, Xiong B, He S, Liu C, Ben-Asher J, Horowitz AR, Wang S, He X. Comparative metabolic profiling of root, leaf, fruit, and stem tissues of *Panax notoginseng*. *International Journal of Food Properties.* 2022, 25:1132–45.<http://doi.org/10.1080/10942912.2022.2071294>.
10. Xue Z, Liu L, Zhang C. Regulation of Shoot Apical Meristem and Axillary Meristem Development in Plants. *Int J Mol Sci.* 2020, 21.<http://doi.org/10.3390/ijms21082917>.
11. Shi B, Vernoux T. Hormonal control of cell identity and growth in the shoot apical meristem. *Current Opinion in Plant Biology.* 2022, 65:102111.<http://doi.org/https://doi.org/10.1016/j.pbi.2021.102111>.
12. Vernoux T, Besnard F, Godin C. What shoots can teach about theories of plant form. *Nature Plants.* 2021, 7:716 – 24.<http://doi.org/10.1038/s41477-021-00930-0>.
13. Zhang TQ, Chen Y, Wang JW. A single-cell analysis of the *Arabidopsis* vegetative shoot apex. *Dev Cell.* 2021, 56:1056-74 e8.<http://doi.org/10.1016/j.devcel.2021.02.021>.
14. Charuvi D, Kiss V, Nevo R, Shimoni E, Adam Z, Reich Z. Gain and loss of photosynthetic membranes during plastid differentiation in the shoot apex of *Arabidopsis*. *Plant Cell.* 2012, 24:1143–57.<http://doi.org/10.1105/tpc.111.094458>.

15. Dayan J, Voronin N, Gong F, Sun TP, Hedden P, Fromm H, Aloni R. Leaf-induced gibberellin signaling is essential for internode elongation, cambial activity, and fiber differentiation in tobacco stems. *Plant Cell*. 2012, 24:66–79.<http://doi.org/10.1105/tpc.111.093096>.
16. Nimchuk ZL, Tarr PT, Ohno C, Qu X, Meyerowitz EM. Plant stem cell signaling involves ligand-dependent trafficking of the CLAVATA1 receptor kinase. *Curr Biol*. 2011, 21:345 – 52.<http://doi.org/10.1016/j.cub.2011.01.039>.
17. Betsuyaku S, Takahashi F, Kinoshita A, Miwa H, Shinozaki K, Fukuda H, Sawa S. Mitogen-activated protein kinase regulated by the CLAVATA receptors contributes to shoot apical meristem homeostasis. *Plant Cell Physiol*. 2011, 52:14–29.<http://doi.org/10.1093/pcp/pcq157>.
18. Hendelman A, Kravchik M, Stav R, Frank W, Arazi T. Tomato HAIRY MERISTEM genes are involved in meristem maintenance and compound leaf morphogenesis. *J Exp Bot*. 2016, 67:6187 – 200.<http://doi.org/10.1093/jxb/erw388>.
19. Tian C, Wang Y, Yu H, He J, Wang J, Shi B, Du Q, Provart NJ, Meyerowitz EM, Jiao Y. A gene expression map of shoot domains reveals regulatory mechanisms. *Nat Commun*. 2019, 10:141.<http://doi.org/10.1038/s41467-018-08083-z>.
20. Zhang TQ, Xu ZG, Shang GD, Wang JW. A Single-Cell RNA Sequencing Profiles the Developmental Landscape of *Arabidopsis* Root. *Mol Plant*. 2019, 12:648 – 60.<http://doi.org/10.1016/j.molp.2019.04.004>.
21. Denyer T, Ma X, Klesen S, Scacchi E, Nieselt K, Timmermans MCP. Spatiotemporal Developmental Trajectories in the *Arabidopsis* Root Revealed Using High-Throughput Single-Cell RNA Sequencing. *Developmental Cell*. 2019, 48:840 – 52.e5.<http://doi.org/https://doi.org/10.1016/j.devcel.2019.02.022>.
22. Kim J-Y, Symeonidi E, Pang TY, Denyer T, Weidauer D, Bezruczyk M, Miras M, Zöllner N, Hartwig T, Wudick MM, et al. Distinct identities of leaf phloem cells revealed by single cell transcriptomics. *The Plant Cell*. 2021, 33:511 – 30.<http://doi.org/10.1093/plcell/koaa060>.
23. Tenorio Berrío R, Verstaen K, Vandamme N, Pevernagie J, Achon I, Van Duyse J, Van Isterdael G, Saeys Y, De Veylder L, Inzé D, Dubois M. Single-cell transcriptomics sheds light on the identity and metabolism of developing leaf cells. *Plant Physiology*. 2021, 188:898– 918.<http://doi.org/10.1093/plphys/kiab489>.
24. Liu Z, Zhou Y, Guo J, Li J, Tian Z, Zhu Z, Wang J, Wu R, Zhang B, Hu Y, et al. Global Dynamic Molecular Profiling of Stomatal Lineage Cell Development by Single-Cell RNA Sequencing. *Molecular Plant*. 2020, 13:1178-93.<http://doi.org/https://doi.org/10.1016/j.molp.2020.06.010>.
25. Xia K, Sun H-X, Li J, Li J, Zhao Y, Chen L, Qin C, Chen R, Chen Z, Liu G, et al. The single-cell stereo-seq reveals region-specific cell subtypes and transcriptome profiling in *Arabidopsis* leaves. *Developmental Cell*. 2022, 57:1299 – 310.e4.<http://doi.org/https://doi.org/10.1016/j.devcel.2022.04.011>.
26. Guo X, Liang J, Lin R, Zhang L, Zhang Z, Wu J, Wang X. Single-cell transcriptome reveals differentiation between adaxial and abaxial mesophyll cells in *Brassica rapa*. *Plant Biotechnology*

- Journal. 2022, 20:2233-35.<http://doi.org/https://doi.org/10.1111/pbi.13919>.
27. Sun X, Feng D, Liu M, Qin R, Li Y, Lu Y, Zhang X, Wang Y, Shen S, Ma W, Zhao J. Single-cell transcriptome reveals dominant subgenome expression and transcriptional response to heat stress in Chinese cabbage. *Genome Biol.* 2022, 23:262.<http://doi.org/10.1186/s13059-022-02834-4>.
  28. Tao S, Liu P, Shi Y, Feng Y, Gao J, Chen L, Zhang A, Cheng X, Wei H, Zhang T, Zhang W. Single-Cell Transcriptome and Network Analyses Unveil Key Transcription Factors Regulating Mesophyll Cell Development in Maize. *Genes (Basel).* 2022, 13.<http://doi.org/10.3390/genes13020374>.
  29. Satterlee JW, Strable J, Scanlon MJ. Plant stem-cell organization and differentiation at single-cell resolution. *Proceedings of the National Academy of Sciences.* 2020, 117:33689-99.<http://doi.org/doi:10.1073/pnas.2018788117>.
  30. Cao Y, Ma J, Han S, Hou M, Wei X, Zhang X, Zhang ZJ, Sun S, Ku L, Tang J, et al. Single-cell RNA sequencing profiles reveal cell type-specific transcriptional regulation networks conditioning fungal invasion in maize roots. *Plant Biotechnology Journal.* 2023, 21:1839–59.<http://doi.org/https://doi.org/10.1111/pbi.14097>.
  31. Xu X, Crow M, Rice BR, Li F, Harris B, Liu L, Demesa-Arevalo E, Lu Z, Wang L, Fox N, et al. Single-cell RNA sequencing of developing maize ears facilitates functional analysis and trait candidate gene discovery. *Dev Cell.* 2021, 56:557–68 e6.<http://doi.org/10.1016/j.devcel.2020.12.015>.
  32. Zhang T-Q, Chen Y, Liu Y, Lin W-H, Wang J-W. Single-cell transcriptome atlas and chromatin accessibility landscape reveal differentiation trajectories in the rice root. *Nature Communications.* 2021, 12:2053.<http://doi.org/10.1038/s41467-021-22352-4>.
  33. Wang Y, Huan Q, Li K, Qian W. Single-cell transcriptome atlas of the leaf and root of rice seedlings. *Journal of Genetics and Genomics.* 2021, 48:881 – 98.<http://doi.org/https://doi.org/10.1016/j.jgg.2021.06.001>.
  34. Zong J, Wang L, Zhu L, Bian L, Zhang B, Chen X, Huang G, Zhang X, Fan J, Cao L, et al. A rice single cell transcriptomic atlas defines the developmental trajectories of rice floret and inflorescence meristems. *New Phytologist.* 2022, 234:494–512.<http://doi.org/https://doi.org/10.1111/nph.18008>.
  35. Sun Z, Jiang S, Wang D, Li L, Liu B, Ran Q, Hu L, Xiong J, Tang Y, Gu X, et al. Single-cell RNA-seq of *Lotus japonicus* provide insights into identification and function of root cell types of legume. *Journal of Integrative Plant Biology.* 2023, 65:1147–52.<http://doi.org/https://doi.org/10.1111/jipb.13435>.
  36. Chen Y, Tong S, Jiang Y, Ai F, Feng Y, Zhang J, Gong J, Qin J, Zhang Y, Zhu Y, et al. Transcriptional landscape of highly lignified poplar stems at single-cell resolution. *Genome Biology.* 2021, 22:319.<http://doi.org/10.1186/s13059-021-02537-2>.
  37. Xie J, Li M, Zeng J, Li X, Zhang D. Single-cell RNA sequencing profiles of stem-differentiating xylem in poplar. *Plant Biotechnology Journal.* 2022, 20:417 – 19.<http://doi.org/https://doi.org/10.1111/pbi.13763>.
  38. Liu H, Hu D, Du P, Wang L, Liang X, Li H, Lu Q, Li S, Liu H, Chen X, 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, 19:2261 – 76.<http://doi.org/10.1111/pbi.13656>.
39. Wang Q, Wu Y, Peng A, Cui J, Zhao M, Pan Y, Zhang M, Tian K, Schwab W, Song C. Single-cell transcriptome atlas reveals developmental trajectories and a novel metabolic pathway of catechin esters in tea leaves. *Plant Biotechnol J.* 2022, 20:2089 – 106.<http://doi.org/10.1111/pbi.13891>.
  40. Kang M, Choi Y, Kim H, Kim SG. Single-cell RNA-sequencing of *Nicotiana attenuata* corolla cells reveals the biosynthetic pathway of a floral scent. *New Phytol.* 2022, 234:527 – 44.<http://doi.org/10.1111/nph.17992>.
  41. Bai Y, Liu H, Lyu H, Su L, Xiong J, Cheng ZM. Development of a single-cell atlas for woodland strawberry (*Fragaria vesca*) leaves during early *Botrytis cinerea* infection using single cell RNA-seq. *Hortic Res.* 2022, 9.<http://doi.org/10.1093/hr/uhab055>.
  42. Long L, Xu F-C, Wang C-H, Zhao X-T, Yuan M, Song C-P, Gao W. Single-cell transcriptome atlas identified novel regulators for pigment gland morphogenesis in cotton. *Plant Biotechnology Journal.* 2023, 21:1100–02.<http://doi.org/https://doi.org/10.1111/pbi.14035>.
  43. Sun Y, Han Y, Sheng K, Yang P, Cao Y, Li H, Zhu Q-H, Chen J, Zhu S, Zhao T. Single-cell transcriptomic analysis reveals the developmental trajectory and transcriptional regulatory networks of pigment glands in *Gossypium bickii*. *Molecular Plant.* 2023, 16:694–708.<http://doi.org/https://doi.org/10.1016/j.molp.2023.02.005>.
  44. Ding Y, Gao W, Qin Y, Li X, Zhang Z, Lai W, Yang Y, Guo K, Li P, Zhou S, Hu H. Single-cell RNA landscape of the special fiber initiation process in *Bombax ceiba*. *Plant Communications.* 2023:100554.<http://doi.org/https://doi.org/10.1016/j.xplc.2023.100554>.
  45. Zhou P, Chen H, Dang J, Shi Z, Shao Y, Liu C, Fan L, Wu Q. Single-cell transcriptome of *Nepeta tenuifolia* leaves reveal differentiation trajectories in glandular trichomes. *Frontiers in Plant Science.* 2022, 13.<http://doi.org/10.3389/fpls.2022.988594>.
  46. Sun S, Shen X, Li Y, Li Y, Wang S, Li R, Zhang H, Shen G, Guo B, Wei J, et al. Single-cell RNA sequencing provides a high-resolution roadmap for understanding the multicellular compartmentation of specialized metabolism. *Nature Plants.* 2023, 9:179 – 90.<http://doi.org/10.1038/s41477-022-01291-y>.
  47. Yu C, Hou K, Zhang H, Liang X, Chen C, Wang Z, Wu Q, Chen G, He J, Bai E, et al. Integrated mass spectrometry imaging and single-cell transcriptome atlas strategies provide novel insights into taxoid biosynthesis and transport in *Taxus mairei* stems. *The Plant Journal.* 2023, 115:1243–60.<http://doi.org/https://doi.org/10.1111/tpj.16315>.
  48. Zhan X, Qiu T, Zhang H, Kailin H, Liang X, Chen C, Wang Z, Wu Q, Wang X, Li X-I, et al. Mass spectrometry imaging and single-cell transcriptional profiling reveal the tissue-specific regulation of bioactive ingredient biosynthesis in *Taxus* leaves. *Plant Communications.* 2023:100630.<http://doi.org/https://doi.org/10.1016/j.xplc.2023.100630>.
  49. Li GW, Nan F, Yuan GH, Liu CX, Liu X, Chen LL, Tian B, Yang L. SCAPTURE: a deep learning-embedded pipeline that captures polyadenylation information from 3' tag-based RNA-seq of single cells.

- Genome Biol. 2021, 22:221.<http://doi.org/10.1186/s13059-021-02437-5>.
50. Eling N, Richard AC, Richardson S, Marioni JC, Vallejos CA. Correcting the Mean-Variance Dependency for Differential Variability Testing Using Single-Cell RNA Sequencing Data. *Cell Syst*. 2018, 7:284 – 94.e12.<http://doi.org/10.1016/j.cels.2018.06.011>.
51. Groover AT, Mansfield SD, DiFazio SP, Dupper G, Fontana JR, Millar R, Wang Y. The *Populus* homeobox gene *ARBORKNOX1* reveals overlapping mechanisms regulating the shoot apical meristem and the vascular cambium. *Plant Molecular Biology*. 2006, 61:917 – 32.<http://doi.org/10.1007/s11103-006-0059-y>.
52. Zhang Y, Zhou J, Wei F, Song T, Yu Y, Yu M, Fan Q, Yang Y, Xue G, Zhang X. Nucleoredoxin Gene TaNRX1 Positively Regulates Drought Tolerance in Transgenic Wheat (*Triticum aestivum* L.). *Front Plant Sci*. 2021, 12:756338.<http://doi.org/10.3389/fpls.2021.756338>.
53. Yang R, Liu J, Lin Z, Sun W, Wu Z, Hu H, Zhang Y. ERF transcription factors involved in salt response in tomato. *Plant Growth Regulation*. 2018, 84:573 – 82.<http://doi.org/10.1007/s10725-017-0362-4>.
54. Won DC, Kim YJ, Kim DH, Park H-M, Maeng PJ. The putative C2H2 transcription factor RocA is a novel regulator of development and secondary metabolism in *Aspergillus nidulans*. *Journal of Microbiology*. 2020, 58:574 – 87.<http://doi.org/10.1007/s12275-020-0083-7>.
55. Luo J, Zhou JJ, Zhang JZ. Aux/IAA Gene Family in Plants: Molecular Structure, Regulation, and Function. *Int J Mol Sci*. 2018, 19.<http://doi.org/10.3390/ijms19010259>.
56. Zou X, Sun H. DOF transcription factors: Specific regulators of plant biological processes. *Front Plant Sci*. 2023, 14:1044918.<http://doi.org/10.3389/fpls.2023.1044918>.
57. Zhang T, Li C, Li D, Liu Y, Yang X. Roles of YABBY transcription factors in the modulation of morphogenesis, development, and phytohormone and stress responses in plants. *J Plant Res*. 2020, 133:751 – 63.<http://doi.org/10.1007/s10265-020-01227-7>.
58. Sun M, Qiu L, Liu Y, Zhang H, Zhang Y, Qin Y, Mao Y, Zhou M, Du X, Qin Z, Dai S. Pto Interaction Proteins: Critical Regulators in Plant Development and Stress Response. *Front Plant Sci*. 2022, 13:774229.<http://doi.org/10.3389/fpls.2022.774229>.
59. Wu C, Avila CA, Goggin FL. The ethylene response factor Pti5 contributes to potato aphid resistance in tomato independent of ethylene signalling. *J Exp Bot*. 2015, 66:559 – 70.<http://doi.org/10.1093/jxb/eru472>.
60. Dar AA, Choudhury AR, Kancharla PK, Arumugam N. The FAD2 Gene in Plants: Occurrence, Regulation, and Role. *Front Plant Sci*. 2017, 8:1789.<http://doi.org/10.3389/fpls.2017.01789>.
61. Shi J, Cao Y, Fan X, Li M, Wang Y, Ming F. A rice microsomal delta-12 fatty acid desaturase can enhance resistance to cold stress in yeast and *Oryza sativa*. *Molecular Breeding*. 2012, 29:743 – 57.<http://doi.org/10.1007/s11032-011-9587-5>.
62. Cui D, Zhao J, Jing Y, Fan M, Liu J, Wang Z, Xin W, Hu Y. The Arabidopsis IDD14, IDD15, and IDD16 Cooperatively Regulate Lateral Organ Morphogenesis and Gravitropism by Promoting Auxin Biosynthesis and Transport. *PLOS Genetics*. 2013, 9:e1003759.<http://doi.org/10.1371/journal.pgen.1003759>.

63. Cui Z, Xue C, Mei Q, Xuan Y. Malectin Domain Protein Kinase (MDPK) Promotes Rice Resistance to Sheath Blight via IDD12, IDD13, and IDD14. *International Journal of Molecular Sciences*. 2022, 23:8214
64. Jiang Z, Tu L, Yang W, Zhang Y, Hu T, Ma B, Lu Y, Cui X, Gao J, Wu X, et al. The chromosome-level reference genome assembly for *Panax notoginseng* and insights into ginsenoside biosynthesis. *Plant Commun*. 2021, 2:100113.<http://doi.org/10.1016/j.xplc.2020.100113>.
65. Wei G, Dong L, Yang J, Zhang L, Xu J, Yang F, Cheng R, Xu R, Chen S. Integrated metabolomic and transcriptomic analyses revealed the distribution of saponins in *Panax notoginseng*. *Acta Pharm Sin B*. 2018, 8:458 – 65.<http://doi.org/10.1016/j.apsb.2017.12.010>.
66. Kumar R, Das S, Mishra M, Choudhury DR, Sharma K, Kumari A, Singh R. Emerging roles of NAC transcription factor in medicinal plants: progress and prospects. *3 Biotech*. 2021, 11:425.<http://doi.org/10.1007/s13205-021-02970-x>.
67. Ji Y, Xiao J, Shen Y, Ma D, Li Z, Pu G, Li X, Huang L, Liu B, Ye H, Wang H. Cloning and Characterization of AabHLH1, a bHLH Transcription Factor that Positively Regulates Artemisinin Biosynthesis in *Artemisia annua*. *Plant and Cell Physiology*. 2014, 55:1592 – 604.<http://doi.org/10.1093/pcp/pcu090>.
68. Thakur S, Vasudev PG. MYB transcription factors and their role in Medicinal plants. *Molecular Biology Reports*. 2022, 49:10995 – 1008.<http://doi.org/10.1007/s11033-022-07825-z>.
69. Deng B, Huang Z, Ge F, Liu D, Lu R, Chen C. An AP2/ERF Family Transcription Factor PnERF1 Raised the Biosynthesis of Saponins in *Panax notoginseng*. *Journal of Plant Growth Regulation*. 2017, 36:691–701.<http://doi.org/10.1007/s00344-017-9672-z>.
70. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*. 2018, 34:i884-i90.<http://doi.org/10.1093/bioinformatics/bty560>.
71. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2012, 29:15–21.<http://doi.org/10.1093/bioinformatics/bts635>.
72. McGinnis CS, Murrow LM, Gartner ZJ. DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Syst*. 2019, 8:329 – 37.e4.<http://doi.org/10.1016/j.cels.2019.03.003>.
73. Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, Trapnell C. Reversed graph embedding resolves complex single-cell trajectories. *Nature Methods*. 2017, 14:979 – 82.<http://doi.org/10.1038/nmeth.4402>.
74. La Manno G, Soldatov R, Zeisel A, Braun E, Hochgerner H, Petukhov V, Lidschreiber K, Kastrioti ME, Lönnerberg P, Furlan A, et al. RNA velocity of single cells. *Nature*. 2018, 560:494 – 98.<http://doi.org/10.1038/s41586-018-0414-6>.
75. Crow M, Paul A, Ballouz S, Huang ZJ, Gillis J. Characterizing the replicability of cell types defined by single cell RNA-sequencing data using MetaNeighbor. *Nature Communications*. 2018, 9:884.<http://doi.org/10.1038/s41467-018-03282-0>.

76. Xu C, Ma D, Ding Q, Zhou Y, Zheng H-L. PlantPhoneDB: A manually curated pan-plant database of ligand-receptor pairs infers cell–cell communication. *Plant Biotechnology Journal*. 2022, 20:2123–34.<http://doi.org/https://doi.org/10.1111/pbi.13893>.
77. Zheng Y, Jiao C, Sun H, Rosli Hernan G, Pombo Marina A, Zhang P, Banf M, Dai X, Martin Gregory B, Giovannoni James J, et al. iTAK: A Program for Genome-wide Prediction and Classification of Plant Transcription Factors, Transcriptional Regulators, and Protein Kinases. *Molecular Plant*. 2016, 9:1667–70.<http://doi.org/https://doi.org/10.1016/j.molp.2016.09.014>.
78. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003, 13:2498 – 504.<http://doi.org/10.1101/gr.1239303>.

## Figures

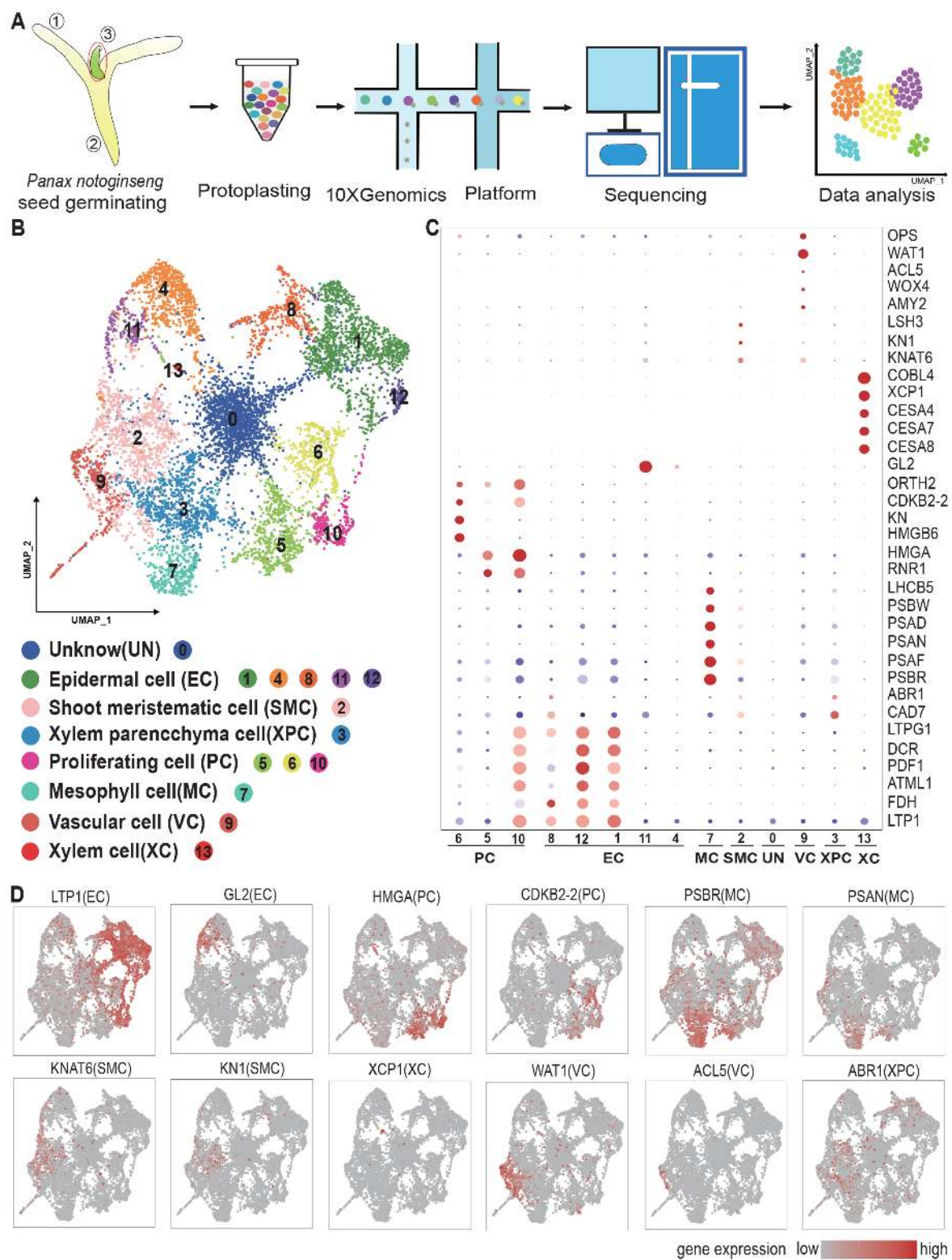


**Figure 1**

Legend not included with this version.

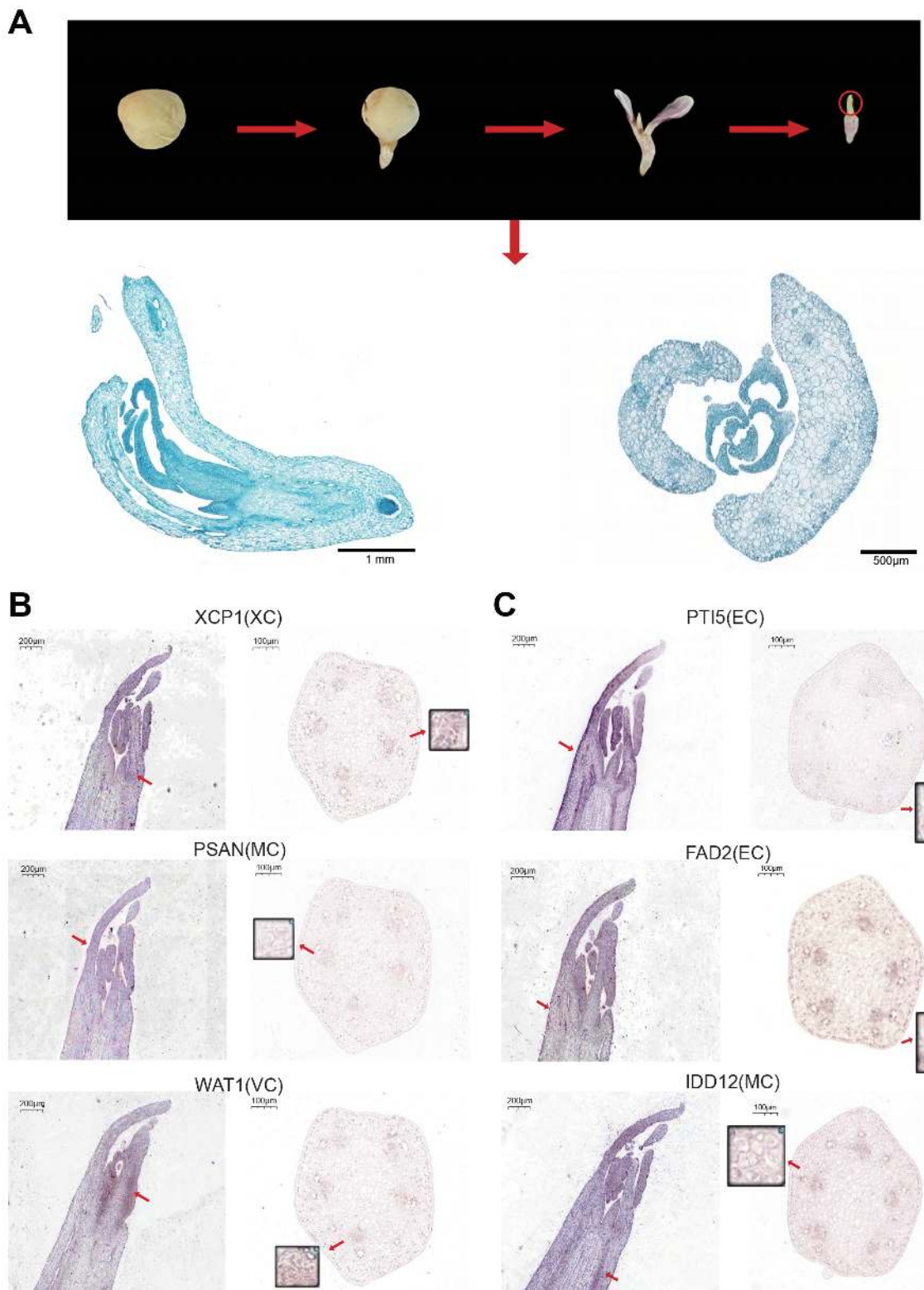


**Figure 2**

Legend not included with this version.

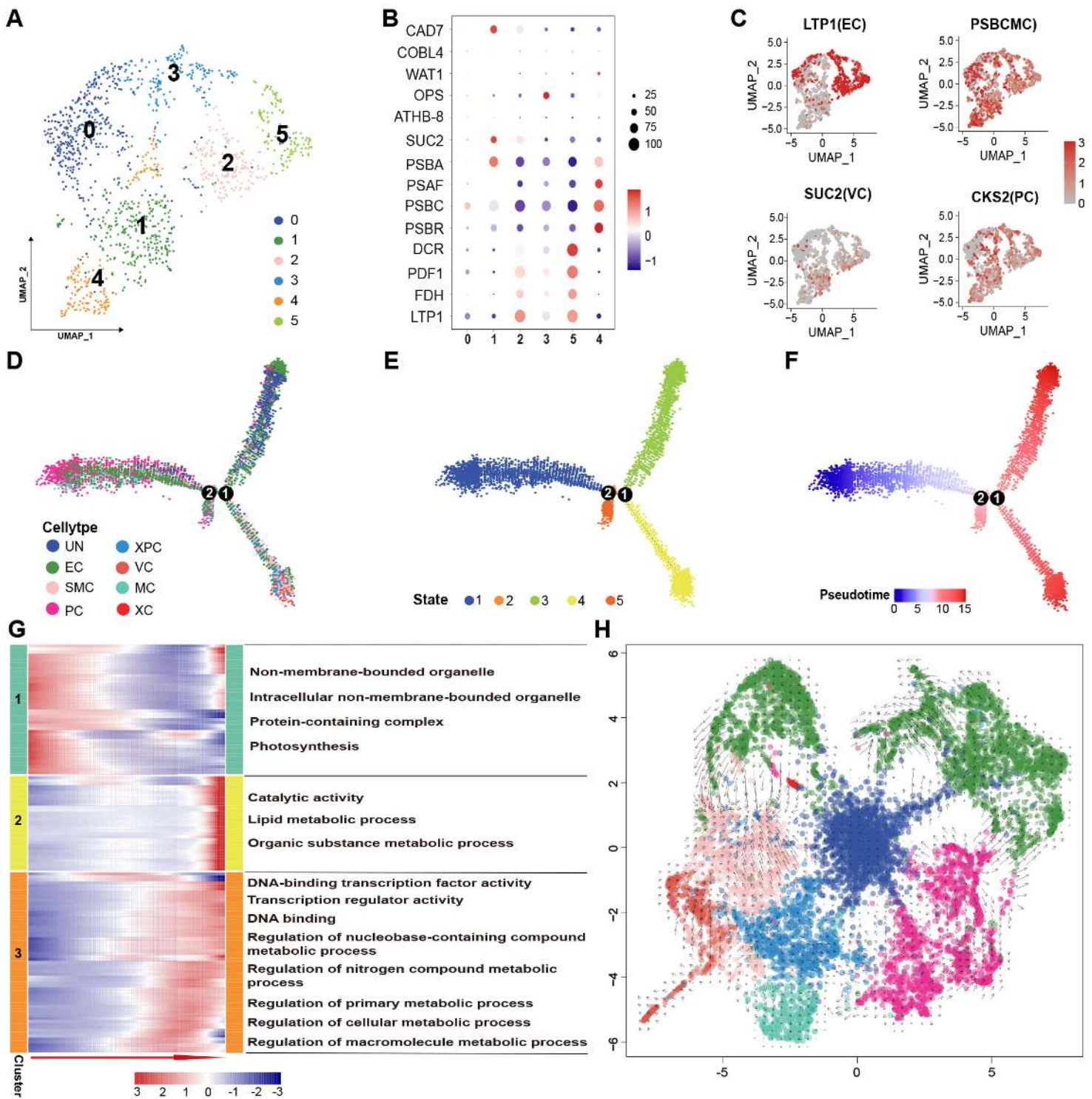
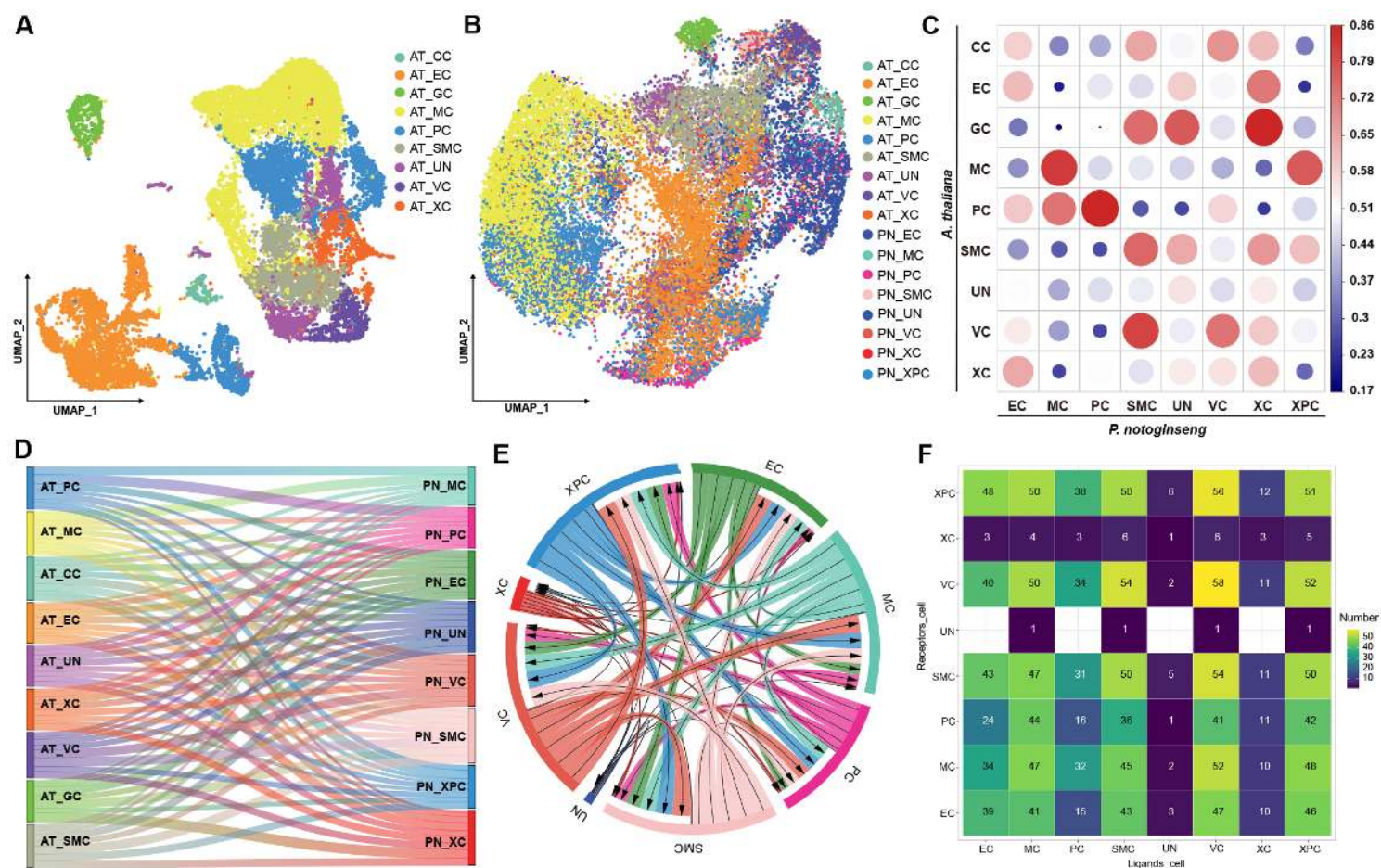
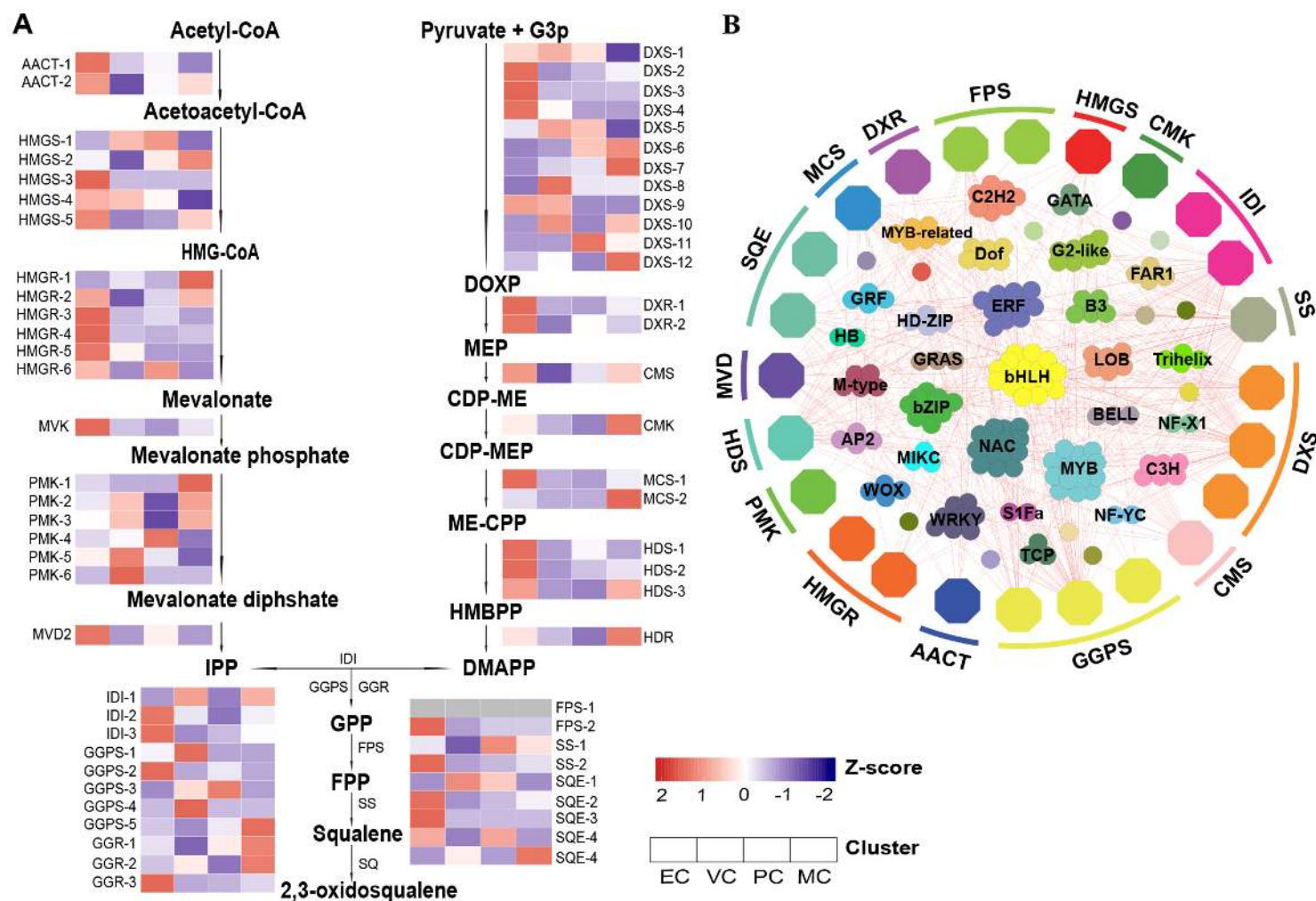


Figure 3

Legend not included with this version.







## Supplementary Files

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

- [Additionalfile1Fig.S1.pdf](#)
- [Additionalfile1Fig.S2.pdf](#)
- [Additionalfile1Fig.S3.pdf](#)
- [Additionalfile1Fig.S4.pdf](#)
- [Additionalfile1Fig.S5.pdf](#)
- [Additionalfile1Fig.S6.pdf](#)
- [Additionalfile1Fig.S7.pdf](#)
- [Additionalfile1TableS1.xlsx](#)
- [Additionalfile1TableS2.xlsx](#)

- [Additionalfile1TableS3.xlsx](#)
- [Additionalfile1TableS4.xlsx](#)
- [Additionalfile1TableS5.xlsx](#)
- [Additionalfile1TableS6.xlsx](#)
- [Additionalfile1TableS7.xlsx](#)
- [SupplementaryInformationLegends.docx](#)