BMC Plant Biology

RESEARCH

Open Access

Identification of key regulators for the development of *Prunus* leaves using single-cell RNA sequencing

Miao Li[1,2], Xiaolan Gao[1,2], Peixian Nie[1,2], Guixiang Li[1,2], Qingtao Gong[1,2], Xiaomin Dong[1,2] and Anning Zhang[1,2*]

Abstract

Background Leaves require the coordinated regulation of multiple types of cells to accumulate energy for plants. Transcriptome analysis at the tissue level cannot analyze the heterogeneity of cells or reveal the molecular regulatory mechanisms of different cell types. Recently, single-cell sequencing technology has become an important tool for analyzing cell heterogeneity, identifying rare cell populations, and drawing dynamic development tracks at a single-cell resolution. However, single-cell sequencing has not been conducted in *Prunus* leaf, limiting our understanding of key regulatory factors and the establishment of peach leaf cell lineages.

Results In this study, we performed single-cell RNA sequencing on young leaves in two varieties of rootstock almond and cultivated red leaf winter peach. A total of 11,132 high-quality leaf cells from both varieties were divided into seven cell clusters based on reported marker genes in *Arabidopsis*. Potentially useful and representative markers were obtained in each cell cluster. Reconstructing the developmental trajectory of procambial cells showed that *PpAPL*, expressed in phloem cells, was a good candidate gene involved in phloem development. In the epidermis cell cluster, *PpCER1* was found to increase the leaf wax content, thereby enhancing drought resistance in almond leaves.

Conclusions Our findings reveal the heterogeneity of leaf cells and identify key regulators in peach and almond, providing a theoretical basis for peach molecular breeding to enhance drought resistance.

Keywords Peach, Leaf, ScRNA-seq, Development, Drought resistance

Background

Leaves are important organs through which plants sense the external environment, and their morphology and structure are strongly susceptible to change under different ecological conditions [1]. After hundreds of millions of years of evolutionary selection, higher plants have developed complex leaf structures to adapt to the changing environment. The upper and lower epidermal cells mainly contain single-cell layers of compact pavement cells and guard cells, which control gas exchange and transpiration between plants and the environment [2]. The palisade and spongy mesophyll are located inside the epidermis and represent the main site for chlorophyll

*Correspondence:
Anning Zhang
Shandongpeach@163.com
[1]National Key Laboratory of Efficient Utilization of Nutrient Resources, Shandong Academy of Agricultural Sciences, Jinan, Shandong Province 250000, China
[2]Shandong Institute of Pomology, Longtai Road No.66, Taian City, Shandong Province 271000, China

synthesis and photosynthesis [3]. Vascular bundles composed of phloem and xylem cells are embedded in the mesophyll cells and are responsible for substance transmission [4].

In recent years, single-cell RNA sequencing (scRNA-seq) technology has developed rapidly, and researchers have made progress in the study of plant leaves using scRNA-seq. For example, Liu et al. [5] assembled global dynamic molecular profiles of stomatal lineage cell development and reported that *AtBRKY33* and *AtBPCs* were involved in the early stage of stoma development in *Arabidopsis*. In addition, Liu et al. [6] identified the novel regulators *AtCDF5* and *AtRGA*, which participate in the early development and function of leaf veins in the cotyledon of *Arabidopsis*. Previous research in allotetraploid peanut (*Arachis hypogaea* L.) revealed the critical transcription factors (TFs) in the leaf blade [7]. In *Arabidopsis*, Kim et al. [8, 9] obtained the single-cell transcriptome profile of the leaf phloem. Sun et al. [10] generated a single-cell transcriptional map of leaves of Chinese cabbage in response to high temperature. With the deepening of scRNA-seq research on leaves, additional cell type markers will be generated, benefiting research on currently underrepresented species.

The NAC-family transcription factor *Altered Phloem Development* (*APL*) serves as a master determinant of phloem cell identity. While *AtAPL* (*AT1G79430*) is indispensable for phloem specification in *Arabidopsis* [11], functional validation of its orthologs remains scarce. Notably, no ortholog has been functionally characterized in *Prunus*. Here, we identified *PpAPL* as the functional ortholog of *AtAPL* in peach, delineated its exclusive expression in developing phloem lineages via along the pseudo-temporal developmental trajectory, and established the foundation for *PpAPL*-mediated phloem development in *Prunus*. *ECERIFERUM1* (*CER1*) encodes a critical decarbonylase catalyzing the conversion of very-long-chain aldehydes to alkanes, a rate-limiting step in cuticular wax biosynthesis essential for minimizing non-stomatal water loss [12]. Functional orthologs of *CER1*, such as *Arabidopsis AtCER1*, *Brassica rapa BrCER1*, rice *OsCER1*, and wheat *TaCER1*, are experimentally validated to catalyze very-long-chain alkane biosynthesis and exhibit conserved epidermal-specific expression patterns critical for cuticular wax deposition [12–15]. Spatial expression studies consistently localize *CER1* to epidermal pavement cells and trichomes—the primary sites of wax deposition [12]. However, in *Prunus*, the quantitative contributions of *CER1* to epidermis-mediated drought resilience remain uncharacterized. In this work, we demonstrate that *PpCER1*-dependent alkane biosynthesis is indispensable for cuticle integrity and water conservation under drought conditions, providing a theoretical basis for enhancing drought resistance in *Prunus*.

Almond (*Prunus amygdalus L.*) trees, native to the mountainous regions of Central Asia, mainly grow in hot and dry environments and exhibit the characteristics of heat and drought resistance [16, 17]. Almond has been applied as a high-quality drought-resistant plant material in scientific research and field production, with almond widely used as a rootstock for peach (*Prunus persica* L.) production. Compared with other non-drought-resistant varieties, almond leaves have developed unique structures to adapt to drought environments, with specialized features including tightly arranged epidermal cells, a thick stratum corneum, veins with a strong transportation capacity, and mesophyll cells containing a high proportion of bound water [18, 19]. In this study, we analyzed the leaves of two closely related trees, namely, extremely drought-resistant rootstock almond (Com) and the cultivated peach variety 'red leaf winter peach' (RLW), at the single-cell level. The results revealed the major cell types found in peach leaves as well as specific marker genes, which can serve as references for further research on leaf development. This study addresses the shortage of research regarding the regulatory mechanism of leaf development in *Prunus* at the single-cell level.

Results

ScRNA-seq of *Prunus* leaves

To perform scRNA-seq of almond and peach leaves, protoplasts were isolated from young Com leaves and RLW shoots. Initially, we obtained 5286 Com cells, with 94,711 reads per cell and 1116 median genes, and 8245 RLW cells, with 38,215 reads per cell and 2436 median genes (Figs. S1–2), for 10 × Genomics scRNA-seq. Then, we conducted quality control at both the single-cell and gene levels, with genes considered eligible when gene number > 200, UMI > 1000, and $\log_{10}$GenesPerUMI > 0.7 (Fig. S3). Finally, a total of 11,132 high-quality cells (Com: 4059; RLW: 7073) were obtained for further analysis. To eliminate the batch effect between two samples in this project, the MNN [20] dimensionality reduction algorithm was applied. Then, based on the MNN dimensionality reduction, the UMAP algorithm was employed to visualize the clustering of single-cell populations, and 10 optimal clusters were ultimately generated using the SNN clustering algorithm [21]. As no scRNA-seq research has been reported in *Prunus* leaves, no markers were available to identify the cell types involved in leaf development. Our strategy was to search for orthologous genes in the two *Prunus* genomes based on several well-characterized markers for *Arabidopsis* leaves, which were employed to identify the cell types in *Prunus* leaves. Finally, six cell clusters and one unknown cluster were generated based on the related marker genes (Fig. 1,Table 1). The largest cell cluster was dominated by mesophyll cells (clusters 1–4), driven by the enrichment

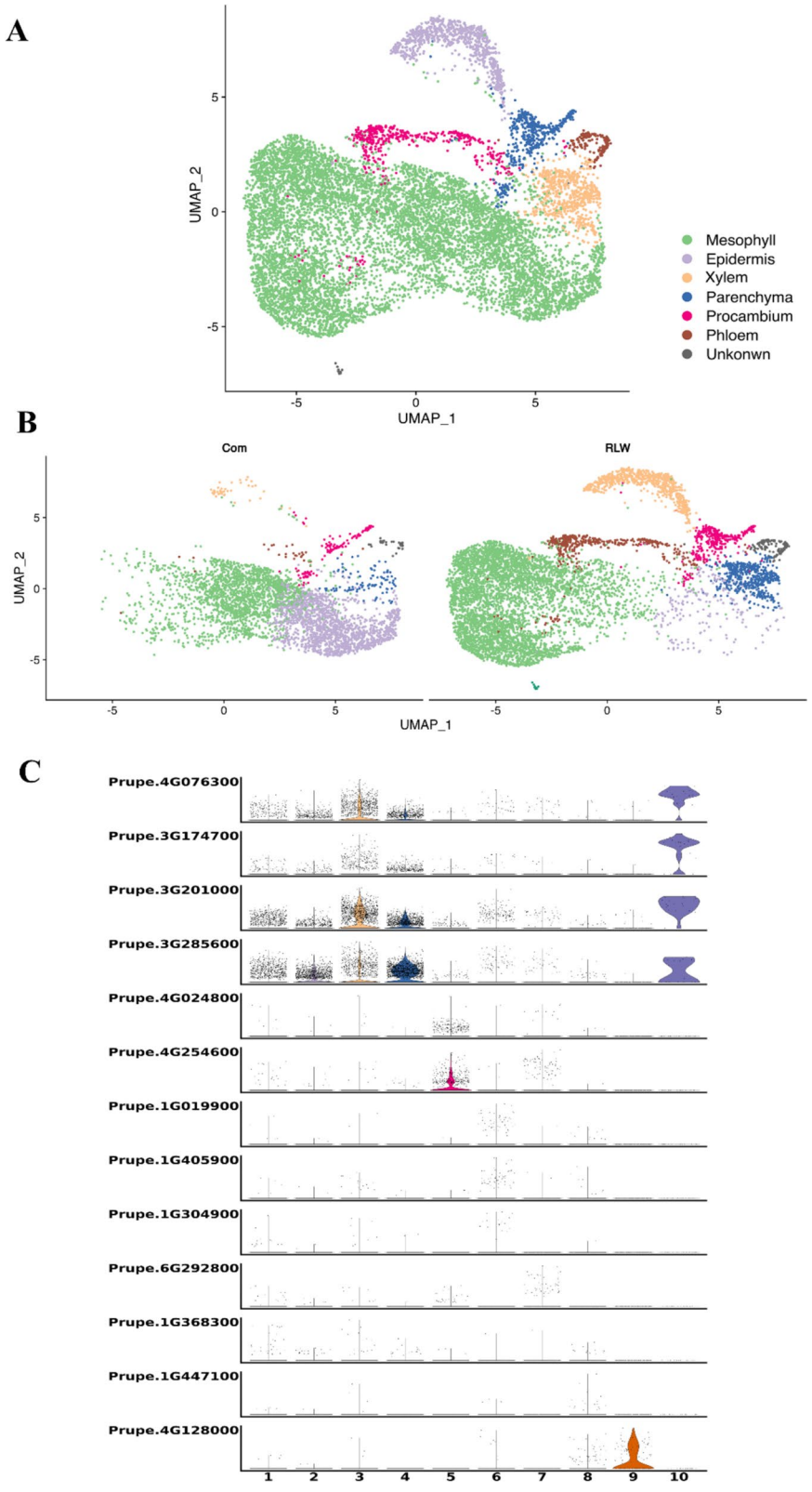

Fig. 1 **A** Identification of cell types according to known marker genes in each cell cluster in young leaves of two *Prunus* species, Com and RLW. **B** Seven cell types for two *Prunus* species, Com and RLW. Com, almond; RLW, red leaf winter peach. Each dot represents a single cell. Colors represent different cell clusters. **C** Violin plots displaying the expression levels of peach orthologous genes corresponding to known *Arabidopsis* marker genes across clusters 1 to 10. The height of each vertical line represents the level of gene expression, while the width of the violin plot represents the number of cells expressing that gene

Table 1 Ten reported markers used to identify cell clusters and their orthologous peach genes

Gene	Reported marker	Cell cluster	Reference	Orthologous gene in peach and protein sequence similarity (%)
AT1G29910	CAB3	Mesophyll	Kim et al., 2021 [9]	Prupe.4G076300 (89.1%)
AT2G05100	LHCB2.1	Mesophyll	Kim et al., 2021 [9]	Prupe.3G201000 (91.7%)
AT3G01500	CA1	Mesophyll	Kim et al., 2021 [9]	Prupe.3G285600 (63.3%)
AT4G21750	AtML1	Epidermis	Kim et al., 2021 [9]	Prupe.4G024800 (79.9%)
AT1G51500	CER5	Epidermis	Kim et al., 2021 [9]	Prupe.4G254600 (74.8%)
At2G15050	LTP7	Parenchymal	Liu et al., 2021 [7]	Prupe.6G292800 (48.8%)
AT1G19850	MP	Procambium	Kim et al., 2021 [9]	Prupe.1G368300 (60.0%)
AT4G32880	HB-8	Procambium	Kim et al., 2021 [9]	Prupe.1G447100 (83.1%)
AT5G19530	ACL5	Xylem	Kim et al., 2021 [9]	Prupe.1G405900 (80.1%)
AT1G07640	OBP2	Phloem	Sun et al., 2022 [10]	Prupe.6G079500 (44.7%)

of known mesophyll markers *CAB3*, *LHCB2.1*, and *CA1* (Figs. S4–6) [8, 9, 22, 23]. Epidermal cells were detected in cluster 5 based on the presence of epidermis-specific TFs *ML1* and *CER5* (Figs. S7–8) [8, 9, 24, 25]. Xylem cells were enriched in cluster 6 based on the marker ACL5 (Fig. S9) [8, 9]. The marker LTP7 showed that parenchymal cells were gathered in cluster 7 (Fig. S10) [7]. Cluster 8 mainly contained procambial cells and was identified by the markers MP.1 and HB 8.1 (Figs. S11–12) [3, 26, 27]. Phloem cells were enriched in cluster 9 and identified using the marker OBP2 (Fig. S13) [10].

Identification of the represented marker genes in each cell cluster

As little progress has been made in the scRNA-seq of *Prunus* leaves, there are no currently available bulk marker genes. We screened marker genes for each cell cluster based on the scRNA-seq of two samples, Com and RLW. To screen for significantly upregulated marker genes in cell clusters, 14,525 marker genes were generated using the presto test method [28]. The top 10 most highly expressed genes in each cluster were selected, and their expression profiles were used to construct a heatmap (Fig. 2; Table S1). The representative genes with the highest expression in each cluster are presented in the UMAP map shown in Fig. 3.

To display the potential TFs involved in regulating the development of *Prunus* leaves, we investigated the TFs that were specifically expressed in each cell cluster (Fig. 4). The mesophyll and epidermis contained numerous TFs, indicating that the formation and development processes of the mesophyll and epidermis rely on a complex regulatory network of TFs. Based on the Com and RLW samples, we screened for TFs that were specifically enriched in the epidermis. As the epidermis serves as the primary barrier against external stress factors, numerous TFs play roles in resisting pathogens and enhancing tolerance within the epidermis. We discovered 11 potential TFs related to disease resistance or environment tolerance in the network of the epidermis (Table S2), supporting the vital role of TFs in the stress resistance function of the epidermis.

Developmental trajectory in the procambium

To elucidate the differentiation process of leaf vein cells, we reconstructed the trajectory of the relationship among the procambium, phloem, and xylem in RLW leaves and obtained a continuous pseudo-time series beginning at procambial cells (Fig. 5A–B). The major putative developmental trajectory displayed a distinct branch and defined the two probable fates of procambial cells. Along the pseudo-temporal developmental trajectory, the differentiation of xylem cells occurred significantly earlier than that of phloem cells. A pseudo-time developmental trajectory among the procambium, phloem, and xylem in Com was also constructed, revealing a similar trajectory (Fig. S14).

To identify key genes controlling the transition of differentiation in procambial cells, heatmap analysis of the pseudo-temporal expression dynamics of specific genes was performed for both RLW (Fig. 5C; Table S3; Fig. S15) and Com (Table S4). After filtering the differentially expressed genes (DEGs) enriched in the xylem and phloem relative to other RLW cell types, 1097 genes, including 28 TFs (gene module 5), were detected to be specifically highly expressed in the xylem cluster, while 322 genes, including 22 TFs (gene module 8), were found to be specifically expressed in the phloem cluster (Table S3). By analyzing the specifically expressed genes in gene module 8 of RLW, we found that a putative MYB factor, *PpAPL* (*Prupe.4G128000*), displayed a highly overlapping pseudo-time trajectory in the phloem cluster (Fig. 5D–E), and *PpAPL* exhibited similar expression patterns in the Com profile (Fig. S15; Table S4). Examining the GUS activity in terms of *PpAPL pro::GUS* revealed enrichment in the phloem of the leaf veins (Fig. 5G). To further investigate the functional characteristics of these signature genes in each module, RLW was subjected to Gene Ontology (GO) functional enrichment analysis. The signature genes of module 8 were mainly enriched in 10 biological processes, which contained one term,

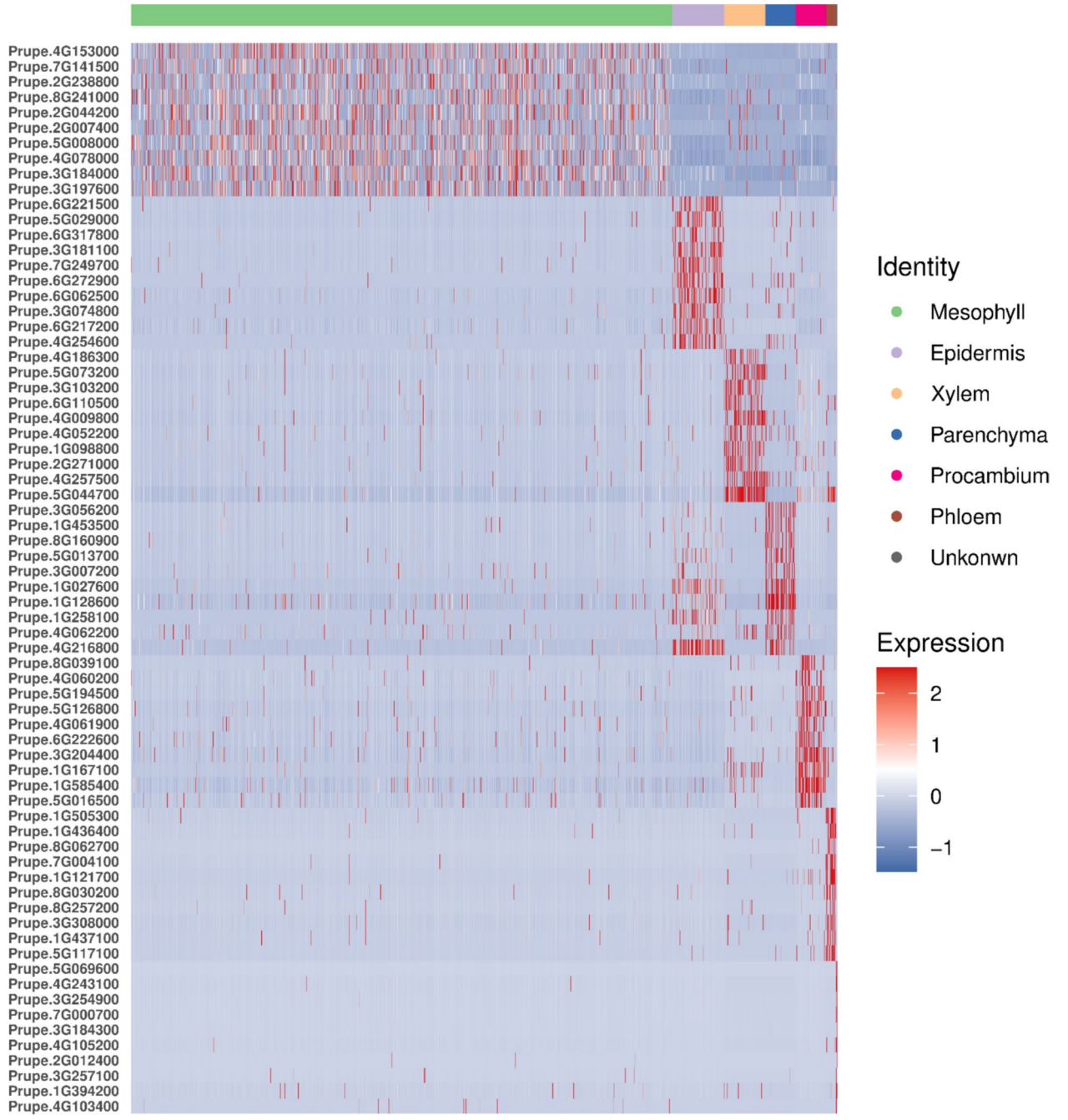

Fig. 2 Heatmap of expression levels of the representative marker genes in each cluster. Representative markers for each cluster are shown on the left; the length of the colors in the upper part of the figure represents the number of cells within that cluster; the color bar indicates the relative gene expression level

designated "phloem development", associated with five enriched genes, *Prupe.1G241500*, *Prupe.1G241400*, *Prupe.2G241500*, *Prupe.4G128000*, and *Prupe.5G162100*, which was consistent with the distribution of phloem cells in the procambium branch (Fig. 6).

Candidate core genes that positively regulate drought tolerance in Com leaves

To investigate the molecular mechanism of drought resistance in almond leaves, we analyzed the epidermal cells of Com and RLW to identify DEGs. A total of 1404 DEGs ($P < 0.05$ and FoldChange > 1.5; Table S5) were generated from Com vs. RLW. Among these DEGs, 596 upregulated DEGs were identified in Com cells. GO and Kyoto

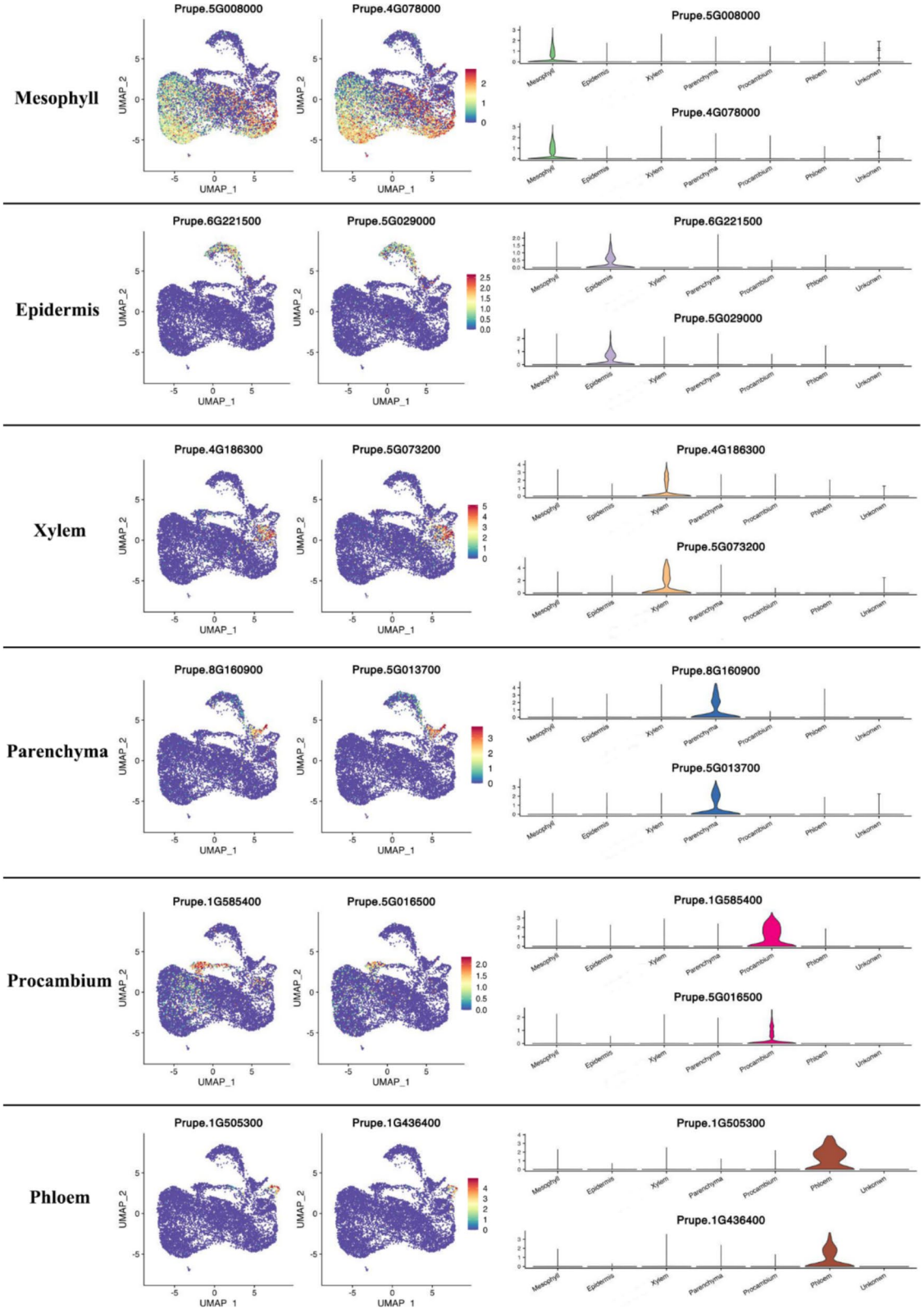

Fig. 3 Unified manifold approximation and projection (UMAP) plot (left) and violin plot (right) showing some top marker genes for different cell types. The corresponding color bar indicates the relative gene expression level. In the violin plot, the height of the vertical lines represents the gene expression, and the width of the violin plot represents the number of cells expressing that gene

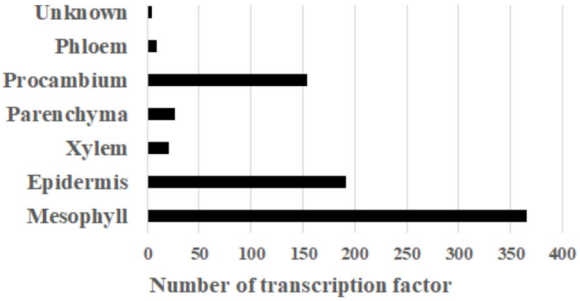

Fig. 4 Number of transcription factors (TFs) in each cluster in two *Prunus* species, Com and RLW. Com, almond; RLW, red leaf winter peach

Encyclopedia of Genes and Genomes (KEGG) analyses were conducted to further predict and analyze the gene functions of upregulated DEGs in Com. GO enrichment analysis indicated that the response to stimulus, cellular process, metabolic process, membrane, organelle, binding, and catalytic activity terms predominated (Fig. S16). Based on the results of KEGG analysis, the enriched pathways were mostly involved in lipid metabolism, energy metabolism, and transcription (Fig. S17). To identify candidate critical genes enhancing drought resistance in the epidermis of Com, we conducted functional predictions for 596 DEGs from the epidermal cells of Com vs. RLW. Except for most unknown genes, we inferred that an upregulated gene might enhance the drought resistance of Com. The candidate gene *Prupe.2G112600* (renamed *PpCER1*) is the orthologous gene of *AtCER1*, which is involved in the synthesis of alkanes and enhances drought resistance in *Arabidopsis* [12, 29].

PpCER1 positively regulates the cuticular wax of the leaf epidermis to enhance drought resistance

To investigate the role of *PpCER1* in drought resistance, the OE vector pBI121-*PpCER1* was constructed and transferred into *A. thaliana.* Fourteen *PpCER1*-OE transgenic lines were generated. After drought treatment, WT plants could not maintain normal plant morphology and exhibited withered and dead leaves, displaying a typical symptom of drought damage, whereas *PpCER1*-OE lines 4 and 13 maintained healthy morphology with luxuriant green leaves, showing obvious resistance to drought damage (Fig. 7A). The physiological indicators of drought resistance indicated that under drought treatment, the catalase (CAT) activity and the contents of proline (PRO) and soluble protein in *PpCER*-OE leaves increased compared with those in WT leaves, while the malondialdehyde (MDA) content of *PpCER*-OE leaves decreased (Fig. 7B). The distribution of wax on the leaves was observed using a scanning electron microscope. The result showed that the coverage degree of epidermal wax in transgenic *PpCER1*-OE line 4 was significantly higher than that in WT leaves (Fig. 7C). To assess the effect of *PpCER1* expression on wax biosynthesis, the cuticular wax composition of *PpCER1*-OE line 4 and WT leaves was analyzed in detail. The result showed that the total cuticular wax contents of *PpCER1*-OE leaves were significantly higher than those of WT leaves (Fig. 7D). In detail, six main components of alkanes were identified in the leaf epicuticular waxes of WT and *PpCER1*-OE lines, with carbon atom numbers ranging from C15 to C25 (Fig. 7D). The six components comprised 2,7,10-trimethyldodecane, hexadecane, undecylcyclohexane, nonadecane, 1-iodoeicosane, and docosyloxytrimethylsilane. Except for 2,7,10-trimethyldodecane, the contents of the other five alkane components were significantly higher in *PpCER1*-OE lines than in WT plants.

Discussion

In recent years, scRNA-seq has become a popular research strategy, enabling significant progress in revealing the formation mechanism and key regulators of complex plant tissues and organs [8, 9, 30–32]. Progress has been reported in the application of scRNA-seq in *Arabidopsis* leaves, and some special cell types have been identified, including vein-related cells, stomata-related cells, and phloem-related cells [5, 6, 8, 9].

Although some available cell type markers from *Arabidopsis* were of great assistance in classifying *Prunus*. leaf cell types in this study, other *Arabidopsis* markers proved to hold little application value in *Prunus*. Peach orthologous genes of the spongy-mesophyll marker *AtAFO* [8, 9] were highly expressed in all clusters from 1 to 4 and could not be used to distinguish spongy mesophyll cells from palisade mesophyll cells. Peach orthologous genes of the *Arabidopsis* stomata marker *AtBPCs* [5] were not specifically expressed in any cluster. We speculate that the expression of key genes in some cell clusters, such as the spongy mesophyll, palisade mesophyll, and stomata, has strong specificity in *Prunus* and lacks matching information in the reference database, making it difficult to identify cell clusters using scRNA-seq at present. In scRNA-seq, suitable markers generally possess characteristics such as expression specificity, detectability (moderate or high expression levels), and the stability of cross-samples [5, 6, 8, 9]. In the present study, we successfully classified and identified cell clusters by searching for orthologous genes of *Arabidopsis* markers. However, the violin plots of certain orthologous genes, including *Prupe.6G292800* (parenchyma marker) and *Prupe.1G447100* (procambium marker), exhibited specific, high-abundance but low-cell-count expression (Fig. 1C), which was not conducive to subsequent applications and research. To better identify leaf single-cell clusters in future research, this study screened 10 marker genes for each cell cluster that displayed specific, high-abundance expression and high-cell-count

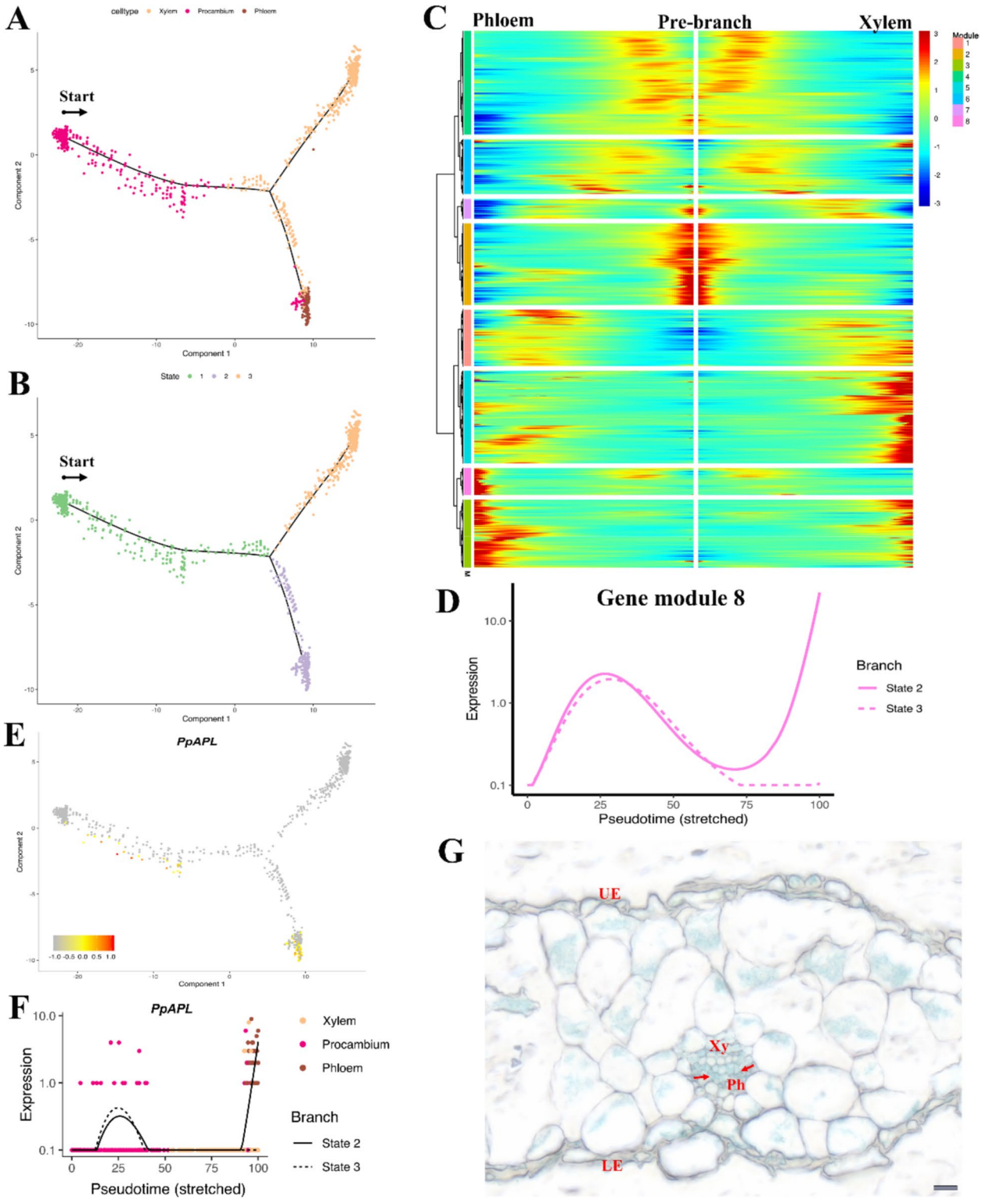

Fig. 5 (See legend on next page.)

expression. For example, *AtHDG5*, an orthologous gene of the epidermal top 10 marker *Prupe.5G029000*, was predominantly expressed in the epidermal layer of shoot meristems and organs [33, 34], indicating that *Prupe.5G029000* was also specially expressed in the epidermal layer in *Prunus* leaves, which was consistent with our analysis. *AtIAA14*, an orthologous gene of the top 10 parenchymal marker *Prupe.1G027600*, was found

(See figure on previous page.)

Fig. 5 **A**–**B** Cell clusters and differentiation state distributions of red leaf winter peach (RLW) following the pseudo-time trajectory of procambium development. State 1, state 2, and state 3 are discrete cell states corresponding to distinct stages along the cell differentiation trajectory. Component 1, pseudo timeline; component 2, differences in cell subtype. **C** Heatmap showing the expression of branch-dependent genes over the pseudo-time trajectory. The colors on the left represent different gene modules. The branch point (shown in the middle of the heatmap) is the beginning of the pseudo-time trajectory, while both sides of the heatmap represent the end of pseudo-time. The color bar indicates the relative expression level. **D** Expression pattern of two differentiation states in gene module 8 of the RLW heatmap along the pseudo-time trajectory. **E***PpAPL* expression superimposed onto the pseudo-time trajectory, with colors indicating higher gene expression. **F** Expression of *PpAPL* in three cell clusters (procambium, phloem, and xylem) in RLW and two differentiation states along the pseudo-time trajectory. **G** GUS activity of *PpAPL* pro::GUS. UE, upper epidermis; LE, lower epidermis; Ph, phloem; Xy, xylem; bar = 10 μm

to be mainly expressed in the root pericycle, which was composed of a single layer of parenchymal cells [35], suggesting that *Prupe.1G027600* holds potential as a parenchymal marker in peach. The above findings underscore that *Prunus* species have unique characteristics compared with *Arabidopsis*, making it necessary to perform the single-cell sequencing of *Prunus* leaves.

After the parenchymal cells construct the basic tissue of the leaf, the procambium gradually differentiates to form vascular tissues to meet the plant's needs for substance transportation and mechanical support [6, 36]. The consistency of the pseudo-time trajectories for both Com and RLW suggests that this developmental trend may be a characteristic manifestation of the procambial cells in peach and almond, providing insight into the development of procambial cells. Xylem cells differentiated significantly earlier than phloem cells, indicating that this order was related to the priority needs of *Prunus* for water transportation. We infer that during the early stages of leaf development, water transport channels (xylem) need to be established rapidly to support cell expansion and the initiation of photosynthesis. In contrast, the enhancement of the phloem, which is responsible for the translocation of organic substances, can be conducted at a later stage as the output of photosynthetic products becomes more crucial following the full expansion of leaves [37, 38]. Based on the DEG profiles of RLW and Com, *PpAPL* was identified from gene modules to exhibit specific expression in the phloem cluster. GUS staining confirmed that *PpAPL* was highly expressed in the phloem of peach and almond leaf veins. It is known that *AtAPL* contributes to both promoting phloem differentiation and repressing xylem differentiation during vascular development [11]. As an orthologous gene of *AtAPL* in peach and almond, the results imply that *PpAPL* is a good candidate gene involved in phloem development, potentially enhancing the understanding of the dynamic process of the distribution of photosynthetic products, the precise transmission paths of signaling substances, and stress response mechanisms under adverse conditions.

We were able to identify DEGs at the cell cluster level, which made the transcriptome analysis more sensitive and accurate than the transcriptome analysis of the whole leaf. Through the heterologous expression of *PpCER1* in *Arabidopsis*, we demonstrated that there was a positive correlation between the expression of *PpCER1* and the synthesis of alkanes in the wax, indicating that *PpCER1* was a key enzyme regulating alkane synthesis. Because alkanes are the main components of cuticular wax [39], the enhanced drought resistance of transgenic *PpCER1*-OE lines was due to the increased content of epidermal wax on the leaves and the reduction of water loss under drought stress. This effect was directly linked to the physiological protective responses observed in transgenic plants, including elevated CAT activity and reduced MDA levels. The reinforced wax layer not only minimized water evaporation but also suppressed reactive oxygen species accumulation through maintaining cellular turgor pressure, which ultimately enabled the plants to retain a healthy morphology. Therefore, *PpCER1* confers drought resistance in plants by regulating alkane biosynthesis in cuticular wax, offering a valuable molecular target for drought-tolerant plant breeding.

Conclusions

This study obtained new insights into the cell clusters of peach and almond leaves and identified novel marker genes for further research. Pseudo-temporal analysis revealed further details regarding the differentiation process of the procambial cells. The results showed that the vital gene *PpCER1* regulates the cuticular wax of the leaf epidermis to enhance drought resistance, helping to elucidate the mechanism of drought resistance in the epidermis of almond and providing a molecular basis for the subsequent breeding of new drought-resistant varieties.

Methods

Plant materials and growth conditions

Com and RLW were collected from the Tianping Lake experimental base, Shandong Institute of Pomology, Taian city, Shandong Province, China (117.08°E, 36.20°N). The two accessions were cultivated in the field under standard local management practices for 12 years. RLW was propagated by bud grafting onto 'Ye Mao Tao' rootstock in 2013. The young leaves of almond and RLW shoots were employed for leaf scRNA-seq experiments. Young leaf samples were collected from the developing leaf buds on detached branches grown in MS medium (pH 5.8) containing added indole-3-butyric

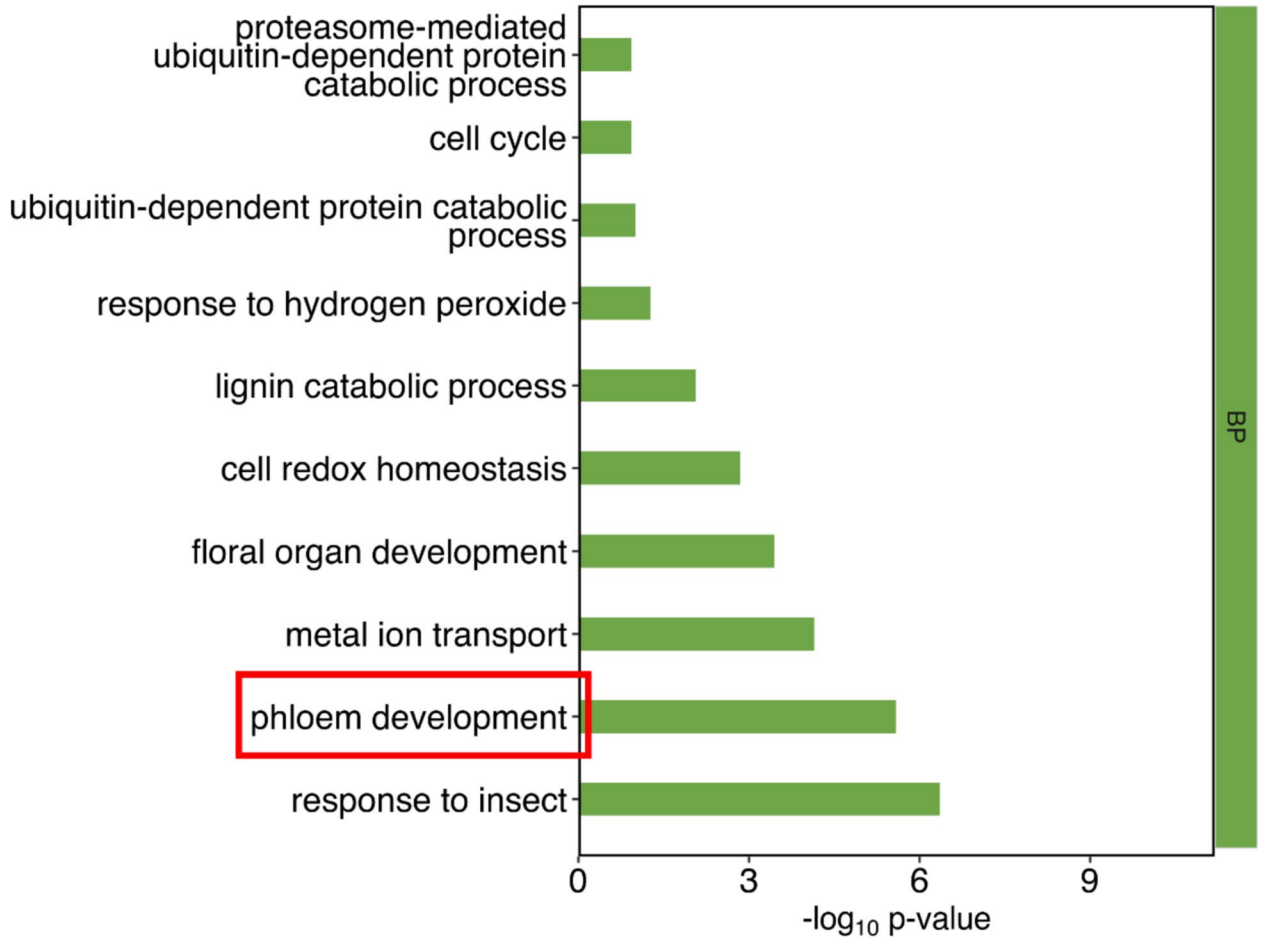

Fig. 6 Top 10 Gene Ontology (GO) terms in the biological process category for highly expressed genes in module 8 of the heatmap for red leaf winter peach (RLW). BP, biological process

acid (IBA) (0.1 mg/L) and 6-benzylaminopurine (6-BA) (1 mg/L) at 24 °C (16 h light/8 h dark, light intensity of 80 mmol m^{-1} s^{-1}). Transgenic *Arabidopsis thaliana* (Col-0) plants were grown in a greenhouse at 22 °C (16 h light/8 h dark, 60% relative humidity).

Preparation of leaf samples for scRNA-seq

Leaf tip regions were collected into a 50-mL centrifuge tube using a wash solution (8% mannitol), which was added with the dissociation solution and then digested for 1 h at 24 °C on a track oscillator. A 40-μm cell filter was placed at the top of the centrifuge tube, and the filtered protoplasts were suspended in the centrifuge tube. After centrifugation, the supernatant was discarded, and sufficient solution was retained to cover the protoplast precipitate. Protoplasts were resuspended in the wash solution, and the washing procedure was repeated twice. A blood cell count plate was used to determine the concentration of protoplasts, which were then placed on ice for subsequent analysis.

Quality control of scRNA-seq data and gene quantification

Quantitative quality control using Cell Ranger software revealed that the high-quality cell count distribution of Com was 5286, while that of RLW was 8245. After excluding dual cells, multicellular cells, and apoptotic cells, the final cell count distributions were 4059 (Com) and 7073 (RLW). The raw data generated through high-throughput sequencing were subjected to data quality statistical analysis using 10 × Genomics software Cell Ranger (v5.0.0, https://www.10xgenomics.com/). Compared to the reference genome, this software quantifies high-throughput single-cell transcriptome data by identifying barcode markers that distinguish cells in the sequence and unique molecular identifier (UMI) markers for different mRNA molecules in each cell, thereby obtaining high-quality cell counts, gene medians, sequencing saturation, and other quality control statistical information.

Gene quantitative quality control and data preprocessing

Based on the preliminary quality control using Cell Ranger, the data underwent further quality control

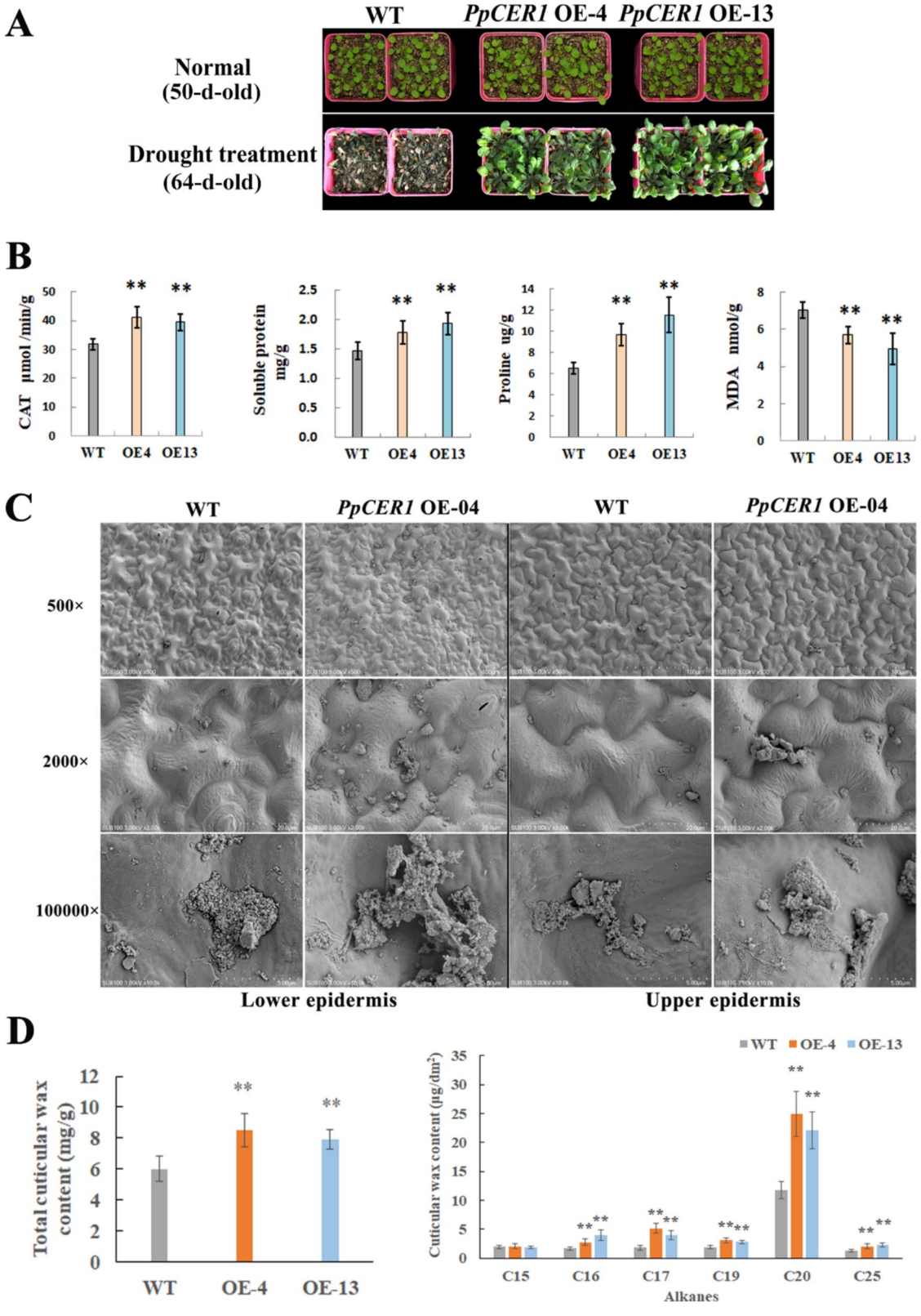

Fig. 7 (See legend on next page.)

(See figure on previous page.)

Fig. 7 **A** Phenotypes of wild-type (WT) *Arabidopsis* and transgenic lines overexpressing *PpCER1* in normal conditions (50-d-old) and 14 d after drought treatment. Bar = 1 cm. **B** Physiological indicators of WT and *PpCER1*-overexpression (OE) lines after drought treatment: catalase (CAT) activity (left); soluble protein (middle left); proline (PRO) content (middle right); and malondialdehyde (MDA) content (right). **C** Scanning electron microscopy images of the upper and lower epidermis of WT *Arabidopsis* and transgenic line *PpCER1*-OE-4. The numbers in each line represent the magnification factor; the bar values corresponding to each line are 100 μm, 20 μm, and 5 μm. **D** Total wax content of WT *Arabidopsis* and transgenic *PpCER1*-OE lines (left); content of six alkane components in WT *Arabidopsis* and transgenic *PpCER1*-OE lines (right). Data presented are the means ± standard deviation of three replicates. Significant differences among different groups are determined using Student's t-tests (*$P < 0.05$, **$P < 0.01$)

processing in the Seurat (v4.0.0) software package [40]. We filtered out low-quality cells based on the distribution of indicators such as nUMI, nGene, and the percentage of mitochondrial genes. Specific quality control thresholds were set to retain cells with gene numbers greater than 200, UMI numbers greater than 1000, and a log10GenesPerUMI value greater than 0.7 using Doublet-Finder (v2.0.2) software to remove dual cells [41].

Dimension reduction and clustering analysis

The mutual nearest neighbor (MNN) dimensionality reduction algorithm was employed to eliminate batch effects [20]. Then, based on the dimensionality reduction results obtained using MNN, the unified manifold approximation and projection (UMAP) algorithm was used to visualize the clustering of single-cell clusters. Finally, the shared nearest neighbor (SNN) clustering algorithm was used to obtain the optimal cell clustering.

Marker gene identification

The Find-All-Markers function (test.use = presto) in the Seurat package [40] was employed for the identification of marker genes, which involved identifying genes that were upregulated for each cell classification relative to other cell populations. The presto test method was employed to perform a difference test between the specified cell population and all other cell populations with the screening conditions of a logfc.threshold greater than 0 and a min.pct (gene expression proportion in all cells) value greater than 0.25 to obtain all marker genes for each cell population. The identified marker genes were visualized using the Vln-Plot and Feature-Plot functions in Seurat.

Identification of DEGs and enrichment analysis

The Find-Markers function (test.use = presto) in the Seurat package [40] was used to screen for DEGs, retaining genes with significant differences at $P < 0.05$ and a difference multiple of greater than 1.5 times. Benjamini–Hochberg [42] was applied for multiple testing adjustments of P values during the selection of DEGs, after which the corresponding q-values were obtained. GO enrichment [43] and KEGG pathway [44] enrichment analysis of DEGs were performed using R based on the hypergeometric distribution.

Pseudo-time analysis

Based on the expression patterns of key genes, Monocle software [45] was used to examine the dynamic changes during the process of temporal development. First, cells displaying significant variations in expression were selected for spatial dimensionality reduction based on their expression profiles, after which a minimum spanning tree (MST) was constructed. The MST was used to identify the longest path representing the differentiation trajectory of cells with similar transcriptional features.

Generation of transgenic *Arabidopsis* and drought resistance analysis

The vector pBI121 was linearized, and the GUS gene was removed through SmaI and SacI double digestion. The coding sequence of *PpCER1* was amplified from *P. amygdalus* L. using sequence-specific primers fused with orthologous recombination junctions (FW: 5'-GGACT CTAGAGGATCCCCGGG-3'; RV: 5'-CGATCGGGG-AAATTCGAGCTC-3') recombined into the linearized pBI121 through orthologous recombination reactions, and the amplified sequence was then transformed into *Agrobacterium* strain GV3101. *Arabidopsis* transformation was conducted according to the protocol of Zhang et al. [46]. *Arabidopsis* seedlings were transplanted 14 days after sowing. We ensured that each treatment included at least 20 individual plants, with 5 plants in each pot, and watering was performed every 4 days. A 14-day drought treatment was initiated 50 days after sowing, during which no water was supplied. Plant phenotypes were observed following drought treatment. The leaves of 20 individual plants were collected from each treatment for the measurement of their drought tolerance indices.

PRO content was detected using a plant PRO extraction kit (Suzhou Mengxi Biomedical Technology Co., Ltd., Suzhou, China) according to the protocol of Mmf et al. [47]. In brief, leaf samples were collected and ground into powder with liquid nitrogen. The leaf powder samples were homogenized in extraction solution at a ratio of 0.1 (g): 1 (mL). Each sample was then extracted by shaking at 90 °C for 10 min. After centrifugation at 10,000 rpm for 10 min, the cooled supernatant was transferred to a microplate reader for analysis at a wavelength of 520 nm. The PRO content (μg/g) was calculated as (determined value – control value – 0.0047)/0.0209 × V1/(sample mass × V1/V2), where V1 is the volume of sample

added to the reaction system, and V2 is the volume of extract added.

MDA content was detected using a plant MDA extraction kit (Suzhou Mengxi Biomedical Technology Co., Ltd., Suzhou, China). In brief, leaf samples were collected and ground into powder with liquid nitrogen. Powdered leaf samples were homogenized in extraction solution at a ratio of 0.1 (g): 1 (mL). After centrifugation at 10,000 rpm at room temperature for 10 min, the supernatant was collected, mixed with the reaction solution at a ratio of 100 (μL): 300 (μL), and heated at 92 °C for 30 min. The mixture was then cooled to room temperature and centrifuged at 12,000 rpm for 10 min. Next, 200 μL of supernatant was transferred to a 1-mL cuvette. Optical density (OD) values at 532 and 600 nm were obtained. The MDA content (μg/g) was calculated as $(A532 - A600 + 0.0055) \div 0.025 \times$ volume of the extraction solution/sample mass.

CAT activity was measured using a plant CAT extraction kit (Suzhou Mengxi Biomedical Technology Co., Ltd., Suzhou, China) according to the protocol of Johansson et al. [48]. Briefly, leaf samples were collected and ground into powder with liquid nitrogen. The powdered leaf samples were homogenized in extraction solution at a ratio of 0.1 (g): 1 (mL). After centrifugation at 12,000 rpm and 4 °C for 10 min, 10 μL of the supernatant was mixed with 190 μL of reaction solution and transferred to the microplate reader for analysis at a wavelength of 240 nm. The CAT activity (μmol/min/g) was calculated as $(A1 - A2) \times 918$/sample mass, where A1 is the absorbance value at 10 s, and A2 is the absorbance value at 70 s.

The soluble protein content was measured using the Coomassie Brilliant Blue Method Micro-detection Kit for Protein Content (Suzhou Mengxi Biomedical Technology Co., Ltd., Suzhou, China). In brief, leaf samples were collected and ground into powder with liquid nitrogen. The leaf powder samples were homogenized in extraction solution at a ratio of 0.1 (g): 1 (mL). After centrifugation at 12,000 rpm and 4 °C for 10 min, 100 μL of the supernatant was mixed with 100 μL of reaction solution and transferred to the microplate reader for analysis at a wavelength of 620 nm. DdH_2O was employed as a control. The soluble protein (mg/g) content was calculated as $(A1 - A2 - 0.008) \div 6.0074 \times$ volume of the extraction solution/sample mass, where A1 is the absorbance value of the sample, and A2 is the absorbance value of the control. Each experiment was independently repeated three times.

Cuticular wax analysis

Determination of total wax content An appropriate amount of the sample was added to chloroform and extracted at room temperature under shock treatment at 120 rpm for 20 min. After the chloroform had completely evaporated, the surface wax content was calculated by measuring the mass of the container before and after the test and determining the difference.

Wax composition analysis Chloroform (25 mL) was added to 65-d-old leaves collected from WT and *PpCER1*-overexpression (OE) lines, and extraction was conducted at 60 °C. The extract was collected, 40 μL of n-octadecane standard solution (1350 ppm) was added, and the resulting solution was concentrated to a volume of 1 mL under a nitrogen stream. The solution was then transferred to an injection vial and then dried under a nitrogen stream, followed by the addition of 100 μL of N,O-Bis(trimethylsilyl)trifluoroacetamide (BSTFA) and 100 μL of pyridine, and derivatization was conducted at 70 °C for 60 min. After drying, the product was diluted to a constant volume using 1.5 mL of chloroform, filtered through a membrane, and then injected into the Agilent gas chromatography-mass spectrometry (GC–MS) 8860-5977B instrument. The parameters of the instrument were as follows: chromatographic column model: HP-5MS UI, 30 m × 0.250 mm × 0.5 μm; flow rate: 1.0 mL/min; column temperature: held at 50 °C for 5 min, then increased to 315 °C at a rate of 5 °C/min, and held for 20 min; injection port temperature: 290 °C; transfer line temperature: 315 °C; and scanning mode: ion source temperature: 230 °C and quadrupole temperature: 150 °C.

Statistical analyses of physiological indicators

Two-tailed Student's t-tests were performed to analyze the significance of differences. All statistical analyses were conducted in Statistical Product and Service Solutions (SPSS) software version 10.0 [49].

Abbreviations

RNA-seq	RNA sequencing
scRNA-seq	Single-cell RNA sequencing
TFs	Transcription factors
HSPs	Heat-shock proteins
Com	Almond
RLW	Ed leaf winter peach
DEGs	Differentially expressed genes
GO	Gene Ontology
KEGG	Kyoto Encyclopedia of Genes and Genomes
CAT	Catalase
PRO	Proline
MDA	Malondialdehyde

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-026-08348-6.

Supplementary Material 1.

Supplementary Material 2.

Supplementary Material 3.

Supplementary Material 4.

Supplementary Material 5.

Supplementary Material 6.

Supplementary Material 7.

Supplementary Material 8.

Acknowledgements
We thank LetPub (www.letpub.com.cn) for its linguistic assistance during the preparation of this manuscript.

Authors' contributions
A.N.Z. and M.L. planned and designed the research. M.L., X.L.G., P.X.N., G.X.L., Q.T.G., and X.M.D. performed experiments, analyzed data, and drew conclusions based on the results. A.N.Z. and M.L. wrote the manuscript. All authors read and approved the final manuscript.

Funding
This study was supported by the Natural Science Foundation of Shandong Province (ZR2023QC056), the Projects of the Key Research and Development Program of Shandong Province (2023TZXD088), the Shandong Provincial Agricultural Improved Variety Project (2025LZGC041), the China Agriculture Research System (CARS-30-Z-08), the Scientific Research and Innovation Fund of Shandong Institute of Pomology (2023GSKY08), the Innovation Project of Shandong Academy of Agricultural Sciences (CXGC2025A06), and the Shandong Improved Agricultural Seed Project (2024LZGCQY021).

Data availability
The raw single-cell RNA sequences of almond (Com) and peach (RLW) obtained in this study have been uploaded to the National Center for Biotechnology Information (NCBI) under accession numbers SRR32586024 and SRR32741306, respectively.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 8 November 2025 / Accepted: 6 February 2026
Published online: 12 February 2026

References

1. Roderick ML, Berry SL, Saunders AR, Noble IR. On the relationship between the composition, morphology and function of leaves. Funct Ecol. 1999;13:696–710.
2. von Groll U, Altmann T. Stomatal cell biology. Curr Opin Plant Biol. 2001;4:555–60.
3. Xia KK, Sun HX, Li J, Li JM, Zhao Y, Chen LC, et al. The single-cell stereo-seq reveals region-specific cell subtypes and transcriptome profiling in Arabidopsis leaves. Dev Cell. 2022;57:1299–310.
4. Svozil J, Gruissem W, Baerenfaller K. Proteasome targeting of proteins in Arabidopsis leaf mesophyll, epidermal and vascular tissues. Front Plant Sci. 2015;6:376.
5. Liu Z, Zhou Y, Guo J, Li J, Tian Z, Zhu Z, et al. Global dynamic molecular profiling of stomatal lineage cell development by single-cell RNA sequencing. Mol Plant. 2020;13:1178–93.
6. Liu Z, Wang J, Zhou Y, Zhang Y, Qin A, Yu X, et al. Identification of novel regulators required for early development of vein pattern in the cotyledons by single-cell RNA-sequencing. Plant J. 2022;110:7–22.
7. Liu H, Hu D, Du P, Wang L, Liang X, Li H, et al. Single-cell RNA-seq describes the transcriptome landscape and identifies critical transcription factors in the leaf blade of the allotetraploid peanut (*Arachis hypogaea L.*). Plant Biotechnol J. 2021;19:2261–76.
8. Kim JY, Symeonidi E, Pang T, Denyer T, Weidauer D, Bezrutczyk M, et al. Unique and distinct identities and functions of leaf phloem cells revealed by single cell transcriptomics. Plant Cell. 2021;33:511–30.
9. Kim JY, Symeonidi E, Pang TY, Denyer T, Frommer WB. Distinct identities of leaf phloem cells revealed by single cell transcriptomics. Plant Cell. 2021;33:511–30.
10. Sun X, Feng D, Liu M, Qin R, Li Y, Lu Y, et al. Single-cell transcriptome reveals dominant subgenome expression and transcriptional response to heat stress in Chinese cabbage. Genome Biol. 2022;23:262.
11. Bonke M, Thitamadee S, Mahonen AP, Hauser MT, Helariutta Y. APL regulates vascular tissue identity in Arabidopsis. Nature. 2003;426:181–6.
12. Bourdenx B, Bernard A, Domergue F, Pascal S, Leger A, Roby D, et al. Overexpression of Arabidopsis ECERIFERUM1 promotes wax very-long-chain alkane biosynthesis and influences plant response to biotic and abiotic stresses. Plant Physiol. 2011;156:29–45.
13. Wang C, Li Y, Xie F, Kuang H, Wan Z. Cloning of the brcer1 gene involved in cuticular wax production in a glossy mutant of non-heading Chinese cabbage (*brassica rapa* L. var. communis). Mol Breed. 2017;37:142.
14. Ni E, Zhou L, Li J, Jiang D, Wang Z, Zheng S, et al. OsCER1 plays a pivotal role in very-long-chain alkane biosynthesis and affects plastid development and programmed cell death of tapetum in rice (*Oryza sativa* L.). Front Plant Sci. 2018;9:1217.
15. Li T, Sun Y, Liu T, Wu H, An P, Shui Z, et al. *TaCER1-1A* is involved in cuticular wax alkane biosynthesis in hexaploid wheat and responds to plant abiotic stresses. Plant Cell Environ. 2019;42:3077–91.
16. Torrecillas A, Domingo R, Planes J, Sanchezblanco MJ, Alarcon JJ. Strategies for drought resistance in leaves of two almond cultivars. Plant Sci. 1996;118:135–43.
17. Yadollahi A, Arzani K, Ebadi A, Wirthensohn M, Karimi S. The response of different almond genotypes to moderate and severe water stress in order to screen for drought tolerance. Sci Hortic. 2011;129:403–13.
18. Rodriguez-Dominguez CM, Buckley TN, Egea G, de Cires A, Hernandez-Santana V, Martorell S, et al. Most stomatal closure in woody species under moderate drought can be explained by stomatal responses to leaf turgor. Plant Cell Environ. 2016;39:2014–26.
19. Oliveira I, Meyer A, Afonso S, Gonalves B. Compared leaf anatomy and water relations of commercial and traditional *prunus dulcis* (Mill.) cultivars under rain-fed conditions. Sci Hortic. 2018;229:226.
20. Haghverdi L, Lun ATL, Morgan MD, Marioni JC. Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. Nat Biotechnol. 2018;36:421–7.
21. Liu H, Wu X, Zhang S. Neighbor selection for multilabel classification. Neurocomputing. 2016;182:187–96.
22. Sawchuk MG, Donner TJ, Head P, Scarpella E. Unique and overlapping expression patterns among members of photosynthesisassociated nuclear gene families in Arabidopsis. Plant Physiol. 2008;148:1908.
23. Endo M, Shimizu H, Nohales MA, Araki T, Kay SA. Tissue-specific clocks in Arabidopsis show asymmetric coupling. Nature. 2014;515:419–22.
24. Pighin JA, Zheng H, Balakshin LJ, Goodman IP, Western TL, Jetter R, et al. Plant cuticular lipid export requires an ABC transporter. Science. 2004;306:702–4.
25. Takada S, Takada N, Yoshida A. ATML1 promotes epidermal cell differentiation in Arabidopsis shoots. Development. 2013;140:1919.
26. Gardiner J, Donner TJ, Scarpella E. Simultaneous activation of SHR and ATHB8 expression defines switch to preprocambial cell state in Arabidopsis leaf development. Dev Dyn. 2011;240:261–70.
27. Moller BK, Ten HCA, Xiang D, Williams N, Lopez LG, Yoshida S, et al. Auxin response cell-autonomously controls ground tissue initiation in the early Arabidopsis embryo. Proc Natl Acad Sci U S A. 2017;114:E2533–9.
28. Jeffrey MP, Davis JM. A comparison of marker gene selection methods for single-cell RNA sequencing data. Genome Biol. 2024;25:1474–7596.
29. Sakuradani E, Zhao L, Haslam TM, Kunst L. The CER22 gene required for the synthesis of cuticular wax alkanes in *Arabidopsis thaliana* is allelic to CER1. Planta. 2013;237:731–8.
30. Zhang TQ, Chen Y, Liu Y, Lin WH, Wang JW. Single-cell transcriptome atlas and chromatin accessibility landscape reveal differentiation trajectories in the rice root. Nat Commun. 2021;12:1–12.

31. Chen Y, Tong S, Jiang Y, AiF, Feng Y, Zhang J, Gong J, Qin J, Zhang Y, Zhu Y, Liu J, Ma Y. Transcriptional landscape of highly lignified poplar stems at single-cell resolution. Genome biology. 2021;22:1–22.
32. Long L, Xu FC, Wang CH, Zhao XT, Yuan M, Song CP, et al. Single-cell transcriptome atlas identified novel regulators for pigment gland morphogenesis in cotton. Plant Biotechnol J. 2023;21:1100–2.
33. Nakamura M, Katsumata H, Abe M, Yabe N, Komeda Y, Yamamoto KT, et al. Characterization of the class IV homeodomain-leucine zipper gene family in *Arabidopsis*. Plant Physiol. 2006;141:1363–75.
34. Kamata N, Okada H, Komeda Y, Takahashi T. Mutations in epidermis-specific HD-ZIP IV genes affect floral organ identity in *Arabidopsis thaliana*. Plant J. 2013;75:430–40.
35. Fukaki H, Tameda S, Masuda H, Tasaka M. Lateral root formation is blocked by a gain-of-function mutation in the SOLITARY-ROOT/IAA14 gene of *Arabidopsis*. Plant J. 2002;29:153–68.
36. Runions A, Tsiantis M, Prusinkiewicz P. A common developmental program can produce diverse leaf shapes. New Phytol. 2017;216:401–18.
37. Edwards J, Martin AP, Andriunas F, Offler CE, Patrick JW, McCurdy DW. GIGANTEA is a component of a regulatory pathway determining wall ingrowth deposition in phloem parenchyma transfer cells of *Arabidopsis thaliana*. Plant J. 2010;63:651–61.
38. Arun Chinnappa KS, Nguyen TTS, Hou J, WuY, McCurdy DW. Phloem parenchyma transfer cells in Arabidopsis-an experimental system to identify transcriptional regulators of wall ingrowth formation. Frontiers In Plant Science. 2013;4:102.
39. Kunst L, Samuels L. Plant cuticles shine: advances in wax biosynthesis and export. Curr Opin Plant Biol. 2009;12:721–7.
40. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol. 2018;36:411–20.
41. Mcginnis CS, Murrow LM, Gartner Z. DoubletFinder: Doublet detection insingle-cell RNA sequencing data using artificial nearest neighbors. Cell Syst. 2019;8:329-337.
42. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol. 1995;57:289–300.
43. Consortium GO. The gene ontology resource: 20 years and still going strong. Nucleic Acids Res. 2019;47:330–8.
44. Kanehisa M, Araki M, Goto S, Hattori M. KEGG for linking genomes to life and the environment. Nucleic Acids Res. 2008;36:480–4.
45. Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat Biotechnol. 2014;32:381–6.
46. Zhang X, Henriques R, Lin SS, Niu QW, Chua NH. *Agrobacterium*-mediated transformation of *Arabidopsis thaliana* using the floral dip method. Nat Protoc. 2006;1:641–6.
47. Mmf M, Ali EF. Evaluation of proline functions in saline conditions. Phytochemistry. 2017;140:52–68.
48. Johansson LH, Borg LAH. A spectrophotometric method for determination of catalase activity in small tissue samples. Anal Biochem. 1988;174:331–6.
49. Link J, Pachaly J. Intranarcotic infusion therapy: a computer interpretation using the program package SPSS (Statistical package for the social sciences). Infusions Ther Klin Ernahr. 1975;2:255–9.

Publisher's Note