

# Single-nucleus RNA and ATAC sequencing analyses provide molecular insights into early pod development of peanut fruit

Yuanyuan Cui<sup>1,3</sup>, Yanning Su<sup>2,3</sup>, Jianxin Bian<sup>1</sup>, Xue Han<sup>1</sup>, Haosong Guo<sup>1,2</sup>, Zhiyuan Yang<sup>1</sup>, Yijun Chen<sup>1</sup>, Lihui Li<sup>1</sup>, Tianyu Li<sup>1</sup>, Xing Wang Deng<sup>1,2</sup> and Xiaoqin Liu<sup>1,\*</sup>

<sup>1</sup>Peking University Institute of Advanced Agricultural Sciences, Shandong Laboratory for Advanced Agricultural Sciences at Weifang, Shandong 261325, China

<sup>2</sup>School of Advanced Agricultural Sciences, Peking University, Beijing 100083, China

<sup>3</sup>These authors contributed equally to this article.

\*Correspondence: Xiaoqin Liu (xiaoqin.liu@pku-iaas.edu.cn)

<https://doi.org/10.1016/j.xplc.2024.100979>

## ABSTRACT

Peanut (*Arachis hypogaea* L.) is an important leguminous oil and economic crop that produces flowers aboveground and fruits belowground. Subterranean fruit-pod development, which significantly affects peanut production, involves complex molecular mechanisms that likely require the coordinated regulation of multiple genes in different tissues. To investigate the molecular mechanisms that underlie peanut fruit-pod development, we characterized the anatomical features of early fruit-pod development and integrated single-nucleus RNA-sequencing (snRNA-seq) and single-nucleus assay for transposase-accessible chromatin with sequencing (snATAC-seq) data at the single-cell level. We identified distinct cell types, such as meristem, embryo, vascular tissue, cuticular layer, and stele cells within the shell wall. These specific cell types were used to examine potential molecular changes unique to each cell type during pivotal stages of fruit-pod development. snRNA-seq analyses of differentially expressed genes revealed cell-type-specific insights that were not previously obtainable from transcriptome analyses of bulk RNA. For instance, we identified *MADS-box* genes that contributes to the formation of parenchyma cells and gravity-related genes that are present in the vascular cells, indicating an essential role for the vascular cells in peg gravitropism. Overall, our single-nucleus analysis provides comprehensive and novel information on specific cell types, gene expression, and chromatin accessibility during the early stages of fruit-pod development. This information will enhance our understanding of the mechanisms that underlie fruit-pod development in peanut and contribute to efforts aimed at improving peanut production.

**Key words:** peanut, pod development, single-nucleus RNA-seq, single-nucleus ATAC-seq, *in situ* RNA hybridization

**Cui Y., Su Y., Bian J., Han X., Guo H., Yang Z., Chen Y., Li L., Li T., Deng X.W., and Liu X.** (2024). Single-nucleus RNA and ATAC sequencing analyses provide molecular insights into early pod development of peanut fruit. *Plant Comm.* **5**, 100979.

## INTRODUCTION

Geocarpy is a reproductive strategy for plant adaptation to the ecological environment (Van der Pijl, 1982). Peanut (*Arachis hypogaea* L.), an economically important plant worldwide, is a typical geocarpic species of the family Leguminosae. After flowers are produced aboveground, the ovary stalk (referred to as the peg) elongates and pushes the ovary belowground, where the fruit pod develops (Kumar et al., 2019). Because of the distinctive way in which the peanut fruit pod forms, the yield of peanut is heavily dependent on the generation and

gravitropic growth of the peg, as well as the expansion of the fruit pod in the soil.

Gravitropic growth is an important characteristic of the peanut peg. Once fertilization occurs and the peg is formed, the peg bends downward in response to gravity. Rapid intercalary

---

Published by the Plant Communications Shanghai Editorial Office in association with Cell Press, an imprint of Elsevier Inc., on behalf of CSPB and CEMPS, CAS.

## Plant Communications

meristem (im) division at the ovary base then promotes peg elongation until the peg is buried in the soil (Ziv, 1981; Liu et al., 2020). Several studies have shown that auxin plays an important role in peg gravitropism. Excision of the ovule region results in a dramatic decrease in the gravitropic response of the peg, but gravitropic growth can be restored by exogenous application of the auxin indole-3-acetic acid (IAA) to the excised tip (Moctezuma and Feldman, 1998). Through use of a monoclonal antibody raised against IAA, auxin was shown to be located in the meristematic region, the area encircling the seeds, and the cortex and epidermis region of the elongation zone during downward growth (Moctezuma, 1999). Zamski and Ziv performed a series of physiological experiments demonstrating that darkness and mechanical stress are important factors that affect the enlargement of peanut pods in the soil (Ziv and Zamski, 1975; Zamski and Ziv, 1976).

Previous studies have shown that various hormones and light-related genes play important roles in peanut fruit-pod development (Sun et al., 2013; Chen et al., 2016; Kumar et al., 2019). However, bulk transcriptome analyses measure the transcription of all genes in tissues and yield average expression levels across all cells. The transcription of key genes from key cell populations may therefore be masked. The complexity of cell types in the peanut peg makes it difficult to characterize the mechanism of peanut fruit-pod development using bulk RNA sequencing (RNA-seq) data (Shushu and Cutter, 1990). Fortunately, single-cell mRNA sequencing (scRNA-seq) enables the study of developmental events at a cellular resolution (Shahan et al., 2021). This technology has been applied to a variety of plants, including *Arabidopsis*, tomato, rice, and maize, and has provided a more detailed view of the transcriptional state of key genes (Kajala et al., 2021; Ortiz-Ramírez et al., 2021; Shahan et al., 2021; Zhang et al., 2021a). Single-cell assay for transposase-accessible chromatin with sequencing (ATAC-seq) has also been developed to study cell-type-specific chromatin accessibility in tissue samples that contain heterogeneous cell populations, in combination with single-cell transcriptome analysis (Xu et al., 2021; Grandi et al., 2022). Single-cell ATAC-seq has been used in plants to analyze the chromatin state of single cells in *Arabidopsis*, wheat, and peanuts (Dorrity et al., 2021; Farmer et al., 2021; Liu et al., 2023; Zhang et al., 2023). However, in most plants, scRNA-seq or ATAC-seq is typically restricted to simple young tissues such as leaves and roots.

In this study, we used isolated nuclei from three different stages of fruit-pod development, namely the aerial peg (Aerpeg), the subterranean peg (Subpeg), and the expanded pod (Exppod), to obtain single-cell gene expression and chromatin accessibility data at the same time. These data, combined with *in situ* RNA hybridization validation and pseudo-time-series analysis, revealed key genes that regulate peg gravitropism and pod expansion, providing in-depth insights into peanut fruit-pod development.

## RESULTS

### Cell morphology at three stages of pod development

Peanuts first bloom aboveground. After fertilization, peanut pegs elongate downward, pushing the ovary into the ground, where the

### Single nucleus RNA-seq and ATAC-seq of peanut pods

fruit develops until reaching maturity (Figure 1A and 1B). When Aerpegs appear, they quickly grow toward the ground. The mean transverse diameter of the Aerpeg, measured adjacent to the seed farthest from the tip, was 756.94, and the mean transverse diameter of the corresponding seed itself was 209.46  $\mu\text{m}$  (Figure 1C). Longitudinal sections revealed a large amount of mechanical tissue distributed around the periphery of the vascular bundle (Figure 1D). The associated cells were elongated and embedded in each other, enhancing the stability of the Aerpeg. The im was located in the upper part of the ovary and was composed of tiny, closely arranged cells with concentrated cytoplasm (Figure 1D). The seed consisted mainly of the seed coat, and the cells at the tip of the peg were closely arranged, with a layer of epidermal cells vertically perpendicular to the second layer of cells from the surface (Figure 1D).

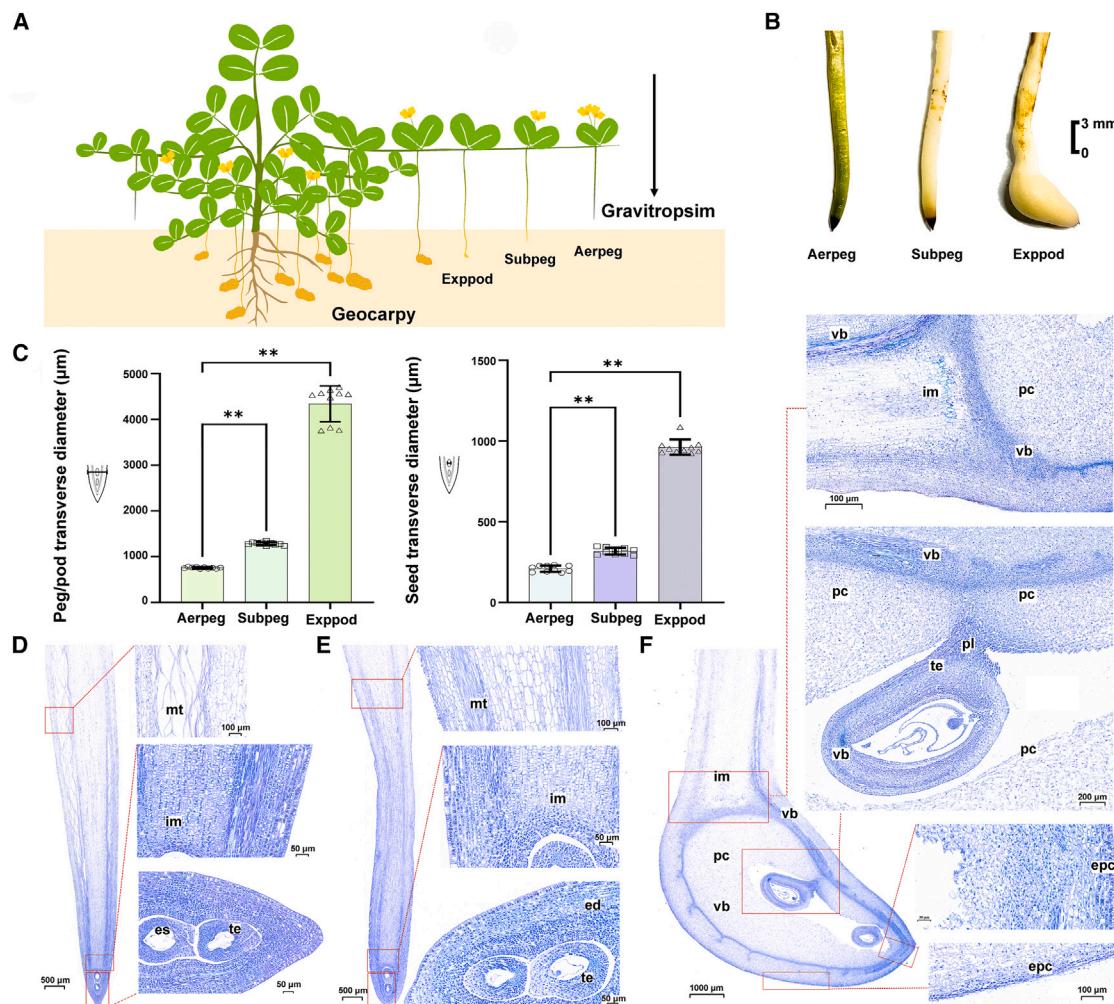
When the peg was buried in the soil but had not yet expanded, the color of the Subpeg was milky white (Figure 1B). The mean transverse diameter of the peg, measured adjacent to the seed farthest from the tip (1284.78  $\mu\text{m}$ ), and of the corresponding seed itself (318.07  $\mu\text{m}$ ) were 69.73% and 51.85% larger than those of the Aerpeg, respectively (Figure 1C). At this stage, the mechanical tissue changed from having a scattered arrangement to being clustered into bundles, the im stopped dividing, and the cells became larger. The cells at the tip of the subterranean peg, including those in the seeds, were more loosely arranged than those in the Aerpeg (Figure 1E).

As the peg began to expand, the cells located in the im became larger and longer, and the cell walls became thinner. A conspicuous ring of thick vascular bundles formed around the seeds. Close to the vascular bundle, many parenchyma cells developed to make up the major part of the developing pericarp. At this point, a significant placenta (pl) could be observed, and the seed farther from the tip developed faster than the seed at the tip (Figure 1F). The transverse diameters of the peg (4338.93  $\mu\text{m}$ ) and the seed farther from the tip (962.12  $\mu\text{m}$ ) were 237.72% and 202.49% larger than those of the Subpeg, respectively (Figure 1B). In addition, finer vascular bundles were visible in the seed coat, and the epidermal cells changed from a tight longitudinal arrangement to a loose lateral arrangement, presumably to meet the need for rapid expansion of the pod (Figure 1F).

### Statistical analyses of the single-nucleus data

Single nuclei were labeled using 10x Genomics barcode technology and reverse transcribed to obtain cDNA. The cDNA underwent quality checks and was found to be intact and normal, meeting the requirements for library construction (Supplemental Figure 1). We then proceeded with library construction and sequencing to obtain information on gene expression and chromatin activity (Supplemental Table 1).

Each of the three sample types was replicated twice, showing high correlation between the two replicates per sample based on the snRNA-seq data ( $R_{\text{Aerpeg}} = 0.76$ ;  $R_{\text{Subpeg}} = 0.84$ ;  $R_{\text{Exppod}} = 0.80$ ). Correlations between bulk RNA-seq data obtained previously (Zhang et al., 2016) and the snRNA-seq data obtained here further confirmed the reliability of the snRNA-seq results ( $R_{\text{Aerpeg}} = 0.73$ ;  $R_{\text{Subpeg}} = 0.73$ ;  $R_{\text{Exppod}} = 0.71$ ) (Supplemental



**Figure 1. Anatomical characteristics of peanut fruit pods at three developmental stages.**

(A) Schematic diagram illustrating the process of belowground fruiting in peanut. Peanut flowers grow aboveground, whereas the fruit develops belowground. Aerpeg (aerial peg), Subpeg (subterranean peg), and Exppod (expanded pod) represent the first three stages of pod development.

(B) Images of pegs/fruit pods at three developmental stages. Scale bar, 3 mm.

(C) Transverse diameters of pegs and seeds (the maximum transverse diameter of the upper side of the seed) at the Aerpeg, Subpeg, and Exppod developmental stages. A *t*-test was used to assess the significance of differences. \*\* $p \leq 0.01$ .

(D) Section of Aerpeg (scale bar, 500  $\mu$ m) showing, from top to bottom, the mechanical tissue (mt; scale bar, 100  $\mu$ m), the intercalary meristem (im; scale bar, 50  $\mu$ m), and the seeds (es, endosperm; te, testa; scale bar, 50  $\mu$ m).

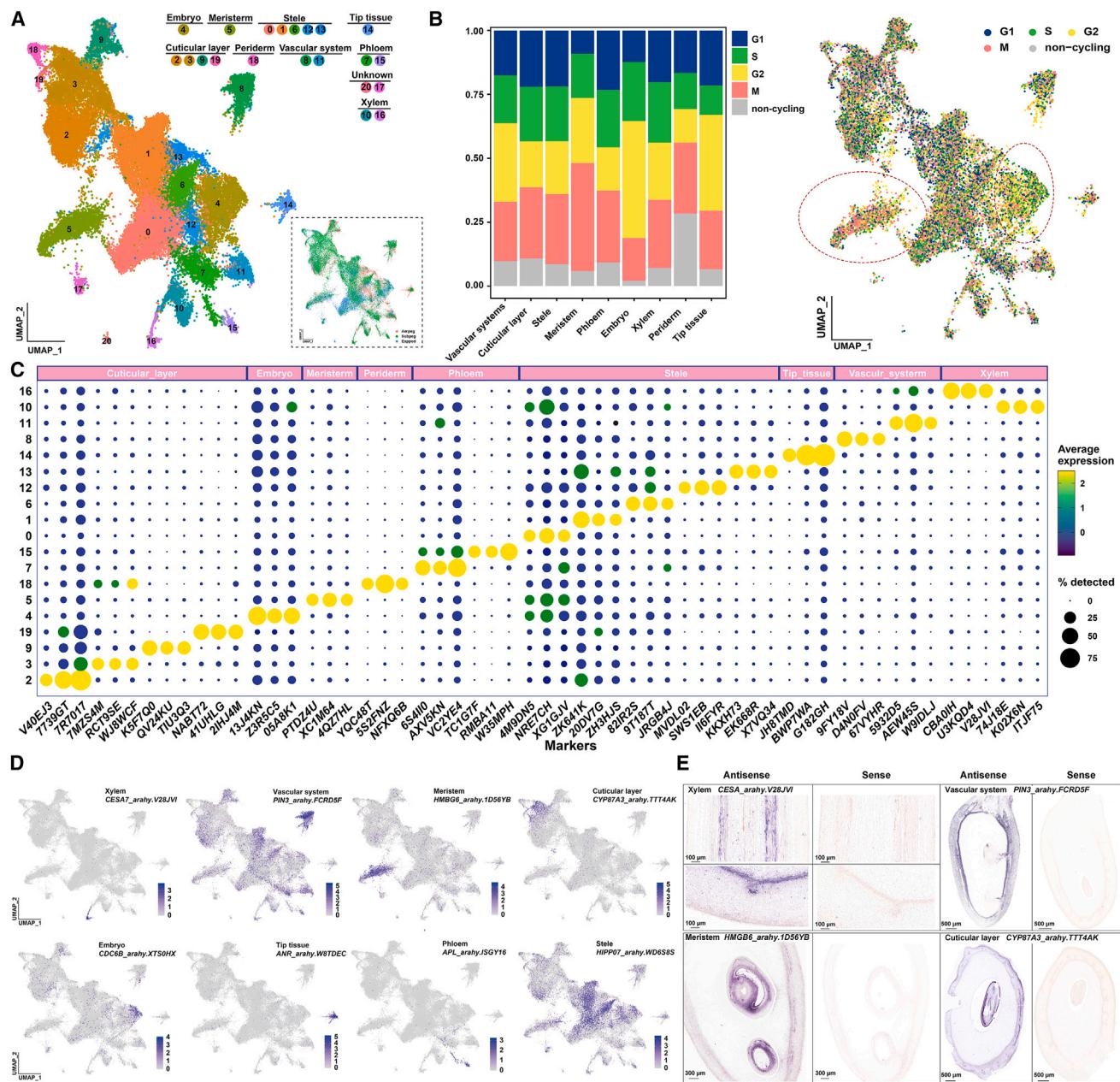
(E) Section of Subpeg (scale bar, 500  $\mu$ m) showing, from top to bottom, the gathering mt (scale bar, 100  $\mu$ m), the im (scale bar, 50  $\mu$ m), and the endodermis (ed) and seed (scale bar, 50  $\mu$ m).

(F) Section of Exppod (scale bar, 1000  $\mu$ m) (pc, parenchymal cells; vb, vascular bundle; pl, placenta) showing, from top to bottom, the im (scale bar, 100  $\mu$ m), the upper side of the seed (scale bar, 200  $\mu$ m), the closely arranged epidermal cells (epc) at the tip of the fruit pod, and the loosely arranged epc at the bottom of the fruit pod (scale bar, 100  $\mu$ m).

**Figure 2.** For the snRNA-seq data, the median number of genes per cell across the samples ranged from 845 to 2265. For the single-nucleus ATAC-seq (snATAC-seq) data, the median number of peaks per cell across the samples ranged from 5110 to 9210 (*Supplemental Table 1*). A total of 42 436 genes were shared between the snRNA-seq and snATAC-seq datasets, constituting 75.86% of all genes identified in these two omics analyses (*Supplemental Figure 2*). After exclusion of dead cells and doublets (see section “methods”), we obtained high-quality transcriptome data, including 13 230, 20 030, and 16 636 qualified single cells from Aerpeg, Subpeg, and Exppod samples, respectively (*Supplemental Figure 3*).

### Identification of cell clusters

A total of 49 896 cells from Aerpeg, Subpeg, and Exppod were clustered using snRNA-seq data and visualized using uniform manifold approximation and projection (UMAP); 21 distinct clusters were identified (*Figure 2A*). Analysis of differentially expressed genes (DEGs;  $\log_2$  fold change [ $FC$ ]  $\geq 0.36$ ,  $p \leq 0.01$ , and expressed in more than 25% of the cells in a cluster) between single-cell clusters and all other cell clusters showed that the number of upregulated DEGs in each cluster ranged from 151 to 1250 (*Supplemental Figure 3*; *Supplemental Table 2*). Cell-cycle analysis revealed two clusters with well-defined cell-cycle phases, cluster 4 (S phase) and cluster 5 (M phase) (*Figure 2B*). Most



**Figure 2. Cell-type annotation and cell-cycle analysis of the peanut peg/pod cells.**

**(A)** Visualization of 21 cell clusters using UMAP. Dots, individual cells,  $n = 49\,896$ ; color, different clusters. The 21 cell clusters were identified as 10 cell types: embryo, meristem, stele, tip tissue, cuticular layer, periderm, vascular system, phloem, xylem, and unknown cells.

**(B)** Cell-cycle analysis of all single cells. G1, growth 1 phase; S, synthesis phase; G2, growth 2 phase; and M, mitosis phase. Left: bar chart illustrating the cell-cycle distribution within the cell population of various cell types. Right: cell-cycle information is mapped onto the UMAP map for a visual representation of the distribution. The red circle marks the meristem, primarily in the M stage, and the embryo, primarily in the S stage.

**(C)** Dot plot showing the expression (normalized by Z score) of the top three cluster-specific marker genes. The size of the dot indicates the percentage of cells within a cell type in which the gene is expressed, and the color indicates the average expression level. The top three marker genes are shown in yellow.

**(D)** UMAP plot showing the representative marker genes of different cell types. Additional marker genes are shown in [Supplemental Figure 5](#), and references for the marker genes are summarized in [Supplemental Table 3](#). Expression level =  $\log(1 + (\text{UMI}_A \div \text{UMI}_{\text{Total}}) \times 10\,000)$ , where  $\text{UMI}_A$  is the number of UMIs of gene A in the target cell,  $\text{UMI}_{\text{Total}}$  is the sum of the numbers of all UMIs in the target cell, and log indicates the natural logarithm with base e.

**(E)** RNA *in situ* hybridization of CESA7 (scale bar, 100  $\mu\text{m}$ ), PIN3 (scale bar, 500  $\mu\text{m}$ ), HMGB6 (scale bar, 300  $\mu\text{m}$ ), and CYP87A3 (scale bar, 500  $\mu\text{m}$ ), which are marker genes for the xylem, vascular system, meristem, and cuticular layer, respectively.

clusters exhibited significant differential expression of marker genes, and the top cluster-specific marker genes with specific expression patterns are shown in Figure 2C.

Because homologous genes exhibit similarity in both sequence and function and because there is limited information on most peanut genes, information on marker genes from *Arabidopsis* and rice obtained from PlantCellMarker (<https://www.tobaccodb.org/pcmdb/homePage>) (Jin et al., 2022) was used to interpret the peanut data. All homologous genes used for cell-type identification in peanut exhibited high protein similarity with marker genes in the corresponding species (Supplemental Tables 3 and 4). The homology of specific key marker genes in peanut received additional support from a maximum likelihood phylogenetic tree of candidates with 1000 bootstrap replicates (Supplemental Figure 4). To enhance the credibility of our findings, we performed RNA *in situ* hybridization for seven selected genes, and the results of these experiments are shown in Figure 2E and Supplemental Figure 5.

Cluster 5 showed specific expression of *HMGB6* (*high mobility group B protein*), *KIN10A*, and *ARM repeat superfamily protein*, which are marker genes of the root meristem (Figure 2D and Supplemental Figure 5; Supplemental Table 3) (Zhang et al., 2019). Thus, cluster 5 was inferred to represent meristem cells. Cluster 11 cells were distinguished by the expression of genes required for DNA recombination, repair, and replication, namely *DMT1* (*DNA (cytosine-5)-methyltransferase 1*) and *RECQL1* (*ATP-dependent DNA helicase Q-like 1*) (Supplemental Figure 5; Supplemental Table 3), which are expressed mainly in the embryo (Xiao et al., 2006; Nayak et al., 2013). Therefore, cluster 11 was inferred to consist of embryo cells. The cell cycles of these two clusters aligned with our predicted cell types, as meristem cells (cluster 5) were predominantly in the M phase, and embryonic cells (cluster 4) were mainly in the S phase (Figure 2B).

Cluster 10 and 16 were classified as xylem because of the specific expression of xylem marker genes, including *CESA* (*cellulose synthase*), *NAC073* (a NAM, ATAF and CUC1/2 transcription factor gene), and *UTX1* (*UDP-xylose transporter 1*) (Figure 2D, Supplemental Figure 5; Supplemental Table 3) (Ryu et al., 2019; Zhang et al., 2019; Jayawardhane et al., 2020; Wendrich et al., 2020). Clusters 7 and 15, which showed specific expression of *PME32* (*pectinesterase/pectinesterase inhibitor 32*) (Ryu et al., 2019), *ABCC11* (*C-type ABC transporter 1*) (Song et al., 2014), and *APL* (*Altered phloem development*) (Bonke et al., 2003), were identified as phloem. Clusters 8 and 11 were identified as vascular cell types on the basis of the expression of known vascular markers: *ABCB15* (*ATP-Binding Cassette B15*) (Kaneda et al., 2011), *WAT1* (*Walls Are Thin 1*) (Ranocha et al., 2010; Ryu et al., 2019), and *ACL5* (*Arginine Decarboxylase-Like 5*) (Ryu et al., 2019) (Supplemental Table 3). Cluster 14 was distinguished by the expression of the anthocyanin synthesis-related genes *CHI* (*chalcone isomerase*), *ANS* (*anthocyanidin synthase*), and *ANR* (*anthocyanidin reductase*) (Figure 2D; Supplemental Figure 5). *ANS* and *ANR* are key genes in the synthesis of anthocyanins, and these genes were assumed to be expressed in the dark purple peg tips (Figure 1B).

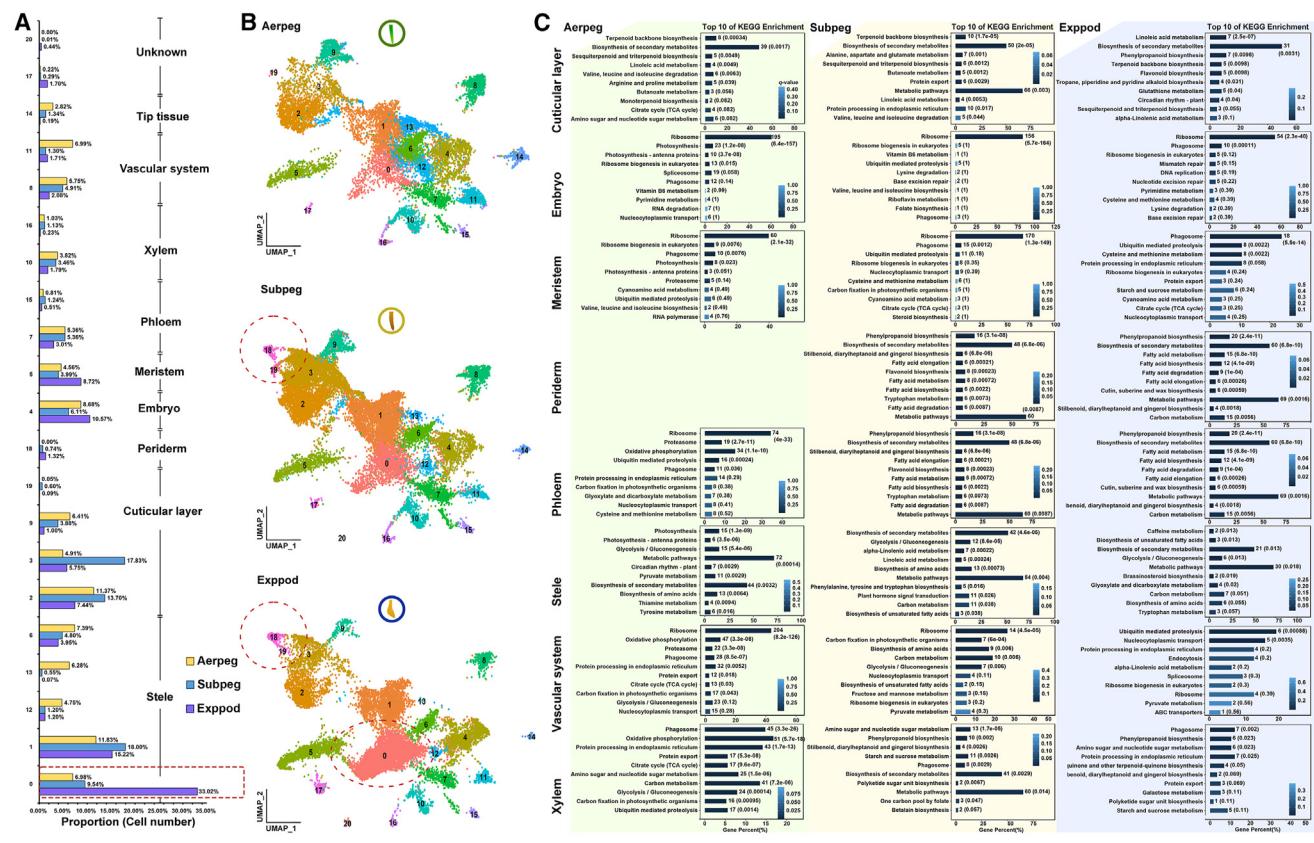
*KCS1* (*3-ketoacyl-CoA synthase 1*) (Denyer et al., 2019), *FAR3* (*Fatty acyl-CoA reductase 3*) (Rowland et al., 2006), and *HHT1*

(*Omega-hydroxypalmitate O-feruloyl transferase*) (Molina et al., 2009), which are involved in the biosynthesis of suberin and wax in the root periderm (Supplemental Figure 5; Supplemental Table 3), were specifically expressed in cluster 18 cells. These cells were therefore identified as periderm. Clusters 2, 3, 9, and 19 were adjacent to cluster 18 and were inferred to represent the cuticular layer because they expressed genes related to epidermal development, such as *PDR* (*pleiotropic drug resistance protein*) (Bessire et al., 2011), *NAC021* (Ryu et al., 2019), and *MYB102* (Supplemental Figure 5; Supplemental Table 3) (Zhang et al., 2019). Clusters 0, 1, 6, 12, and 13 were inferred to comprise stele cells because genes expressed in the root stele and pericycle, such as *HIPP07* (*heavy metal transport/detoxification superfamily protein*) and *SBT2.6* (*subtilisin-like protease SBT2.6*), were predominantly expressed in these clusters (Jean-Baptiste et al., 2019; Long et al., 2021) (Figure 2D, Supplemental Figure 5; Supplemental Table 3).

To confirm the accuracy of cell-type identifications, the expression of eight marker genes was examined by *in situ* RNA hybridization: *CESA7* (*Cellulose synthase A7*), *PIN3* (*Pin-formed 3, encoding an auxin transporter*), *HMGB6* (*high mobility group box 1*), *CYP87A3* (*Cytochrome P450 87A3*), *protein DEK-like*, *YAB5* (*YABBY transcription factor 5*), *NAC021*, and *KCS1* (*3-ketoacyl-CoA synthase 1*). *CESA7*, *PIN3*, *HMGB6*, and *CYP87A3* were predominantly expressed in the xylem, vascular system, meristem, and cuticular layer, respectively (Figure 2E). *Protein DEK-like* was expressed mainly in the embryo; *YAB5* was expressed primarily in the seed, embryo, meristem, and vascular system; *NAC021* was expressed mainly in the cuticular layer and periderm; and *KCS1* was expressed mainly in the periderm. Notably, these genes were found to be expressed in precisely the same locations as initially predicted (Figure 2E and Supplemental Figure 5).

We also examined the chromatin-enriched peaks of the eight marker genes shown in Figure 2D across distinct cell types. However, we did not observe any significant cell-type-specific peaks in the gene bodies or their upstream and downstream regions (Supplemental Figure 6). One possible explanation is that the snATAC-seq results did not correlate well with the snRNA-seq results. Another possibility is that the sequencing data from single-cell snATAC-seq were insufficient to detect the presence of differential peaks.

To investigate the relationship between gene expression and chromatin accessibility, snATAC-seq data were used to cluster peanut pod cells at the chromatin level, subsequently correlating them with cell types identified by snRNA-seq. The cell types in snATAC-seq exhibited expression of specific marker genes (Supplemental Figure 7A). To examine the correlation between snATAC-seq and snRNA-seq data, the gene activities of snRNA-seq cell-type-specific marker genes were analyzed according to snATAC-seq. The gene activity of cell-type-specific marker genes in the cuticular layer, stele, and vascular system was consistent with their gene expression (Supplemental Figure 7B). This might be because the capture rate of accessible chromatin regions was not sufficiently high at the single-cell level. Consequently, only cell types with a substantial number of cells, such as the cuticular layer, stele, and vascular system, demonstrated consistency with snRNA-seq gene expression, whereas this consistency was challenging to detect in cell types with a lower cell count.



**Figure 3. Changes in cell type and KEGG analysis during pod development.**

**(A)** Changes in the proportion of cells within each cell cluster across three developmental stages. The red box marks cluster 0, where the number of cells increased significantly with pod development.

**(B)** UMAP plot visualizing the dynamic changes of 21 cell clusters in Aerpeg, Subpeg, and Exppod during pod development. The red circle marks cluster 18 (periderm), which emerges in Subpeg and Exppod. The red oval marks cluster 0 in Exppod, which is the cell group with the largest proportion in Exppod.

**(C)** KEGG analysis of the upregulated DEGs in each cell type within Aerpeg, Subpeg, and Exppod. The top 10 pathways are displayed in the figure. The y axis represents the pathways, and the x axis indicates the percentage of these pathways relative to the total number of DEG pathways. The color intensity reflects the q value, with darker colors indicating smaller q values. The values displayed beside each bar represent the number and q value of the respective pathway.

To further assess the relationship between snRNA-seq and snATAC-seq data, we integrated the top three marker genes from both snRNA-seq and snATAC-seq (Figure 2C and Supplemental Figure 7A) and used them to examine correlations between gene expression and gene activity across various clusters and cell types (Supplemental Figure 8). Most cell types and clusters showed positive correlations, supporting the idea that chromatin accessibility is related to the expression patterns of at least several marker genes, consistent with previous studies in *Arabidopsis* (Farmer et al., 2021). The Pearson analysis showed stronger correlations than the Spearman analysis, which may be due to Spearman's more stringent nature (Supplemental Figure 8A). When all samples were considered, Pearson's analysis indicated that correlations were stronger in the embryo, meristem, periderm, cuticular layer, stele, and vascular system than in other cell types. When the three samples were analyzed separately using Pearson and Spearman correlations, they showed a correlation pattern similar to that observed across all samples (Supplemental Figure 8B).

Transcription factors control chromatin and transcription by recognizing specific DNA sequences to form complex systems that direct genome expression. Motif analysis was performed

through snATAC-seq at the cell-type level. The cell-type-specific known motifs were confirmed with MEME-AME (<https://meme-suite.org/meme/doc/ame.html>) (Supplemental Table 5). Remarkably, we discovered that five cell types harbored motifs distinct from those found in other cell types: the cuticular layer had 14 such motifs, primarily ERF (Ethylene Response Factor); the stele had 12, predominantly TCP (Teosinte Branched1/Cycloidea/Proliferating Cell Factor); the meristem had 11, primarily MYB (Myeloblastosis) and IDD (Indeterminate Domain); the embryo had nine, mainly ERF; and the vascular system had seven, mainly MYB (Supplemental Figure 9). These known motif sites can serve as a reference for screening of transcription factors specific to each cell type.

### Characterization of cell types and cell numbers during fruit-pod development

To further understand the changes in cell type and proportion during fruit-pod development, we calculated changes in the proportion of each cell type and visualized the clustered cells from Aerpeg, Subpeg, and Exppod samples separately. The proportion of cells in cluster 0 increased significantly during peg penetration and expansion. Notably, cluster 18 (periderm)

emerged as a newly prominent cluster following peg penetration (Figure 3A and 3B).

Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis of marker genes in each cell type for each sample showed that the marker genes of the cuticular layer were significantly enriched in the biosynthesis of secondary metabolites, which may help the epidermis to respond to changes in the environment (Figure 3C). The marker genes associated with the embryo and meristem were significantly enriched in ribosome-related processes. The robust protein synthesis observed aligned with their active cell-cycle status. Following penetration, the marker genes associated with the periderm and phloem showed significant enrichment in phenylpropanoid biosynthesis and fatty-acid-related processes. This suggests that these two cell types perform substantial lipid and wax synthesis after penetration. Interestingly, the marker genes associated with the vascular system shifted from ribosome-related processes to ubiquitin-mediated proteolysis following pod expansion. This suggests a transition to protein degradation processes in vascular cells during pod enlargement (Figure 3C).

### Differentiation trajectories in the development of peanut epidermal cells

Epidermal cells play an important role in adapting the gynophore to new environments. Two types of epidermal cells, namely clusters 2, 3, 9, and 19 (the cuticular layer) and cluster 18 (the periderm) were analyzed using pseudo-time analysis (Figure 4A). The Monocle R package defines different segments of the cell lineage as distinct states of cellular differentiation. Branching occurs in the lineage when cells make differentiation fate choices during development. The upside branch 2 was exclusively enriched in cells from Subpeg and Exppod, with periderm cells specifically enriched in upside branch 2. The cell fate transition from branch 1 (cuticular layer) to branch 2 (periderm) defined the development of periderm cells (Figure 4B).

DEGs between cell fate 1 (from branch 1 [cuticular layer] to branch 2 [periderm]) and cell fate 2 (from branch 1 [cuticular layer] to branch 3 [cuticular layer]) were analyzed to identify key genes related to the development of periderm cells (Figure 4C). Among these DEGs, four KCS (3-ketoacyl CoA synthase) genes were identified as being highly expressed in periderm cells. Their expression changed during the process of cuticular layer development (Figure 4C and 4D, Supplemental Figure 10A). Chromatin-enriched peaks across distinct cell types for two KCS genes were also analyzed, but we did not observe any significant cell-type-specific peaks (Supplemental Figure 10B).

Very-long-chain fatty acids are essential precursors of cuticular wax and aliphatic suberin, and KCS is the key enzyme in very-long-chain fatty acid biosynthesis (Lee et al., 2009; Huang et al., 2023). Consequently, upregulation of KCS gene expression might help to facilitate the biosynthesis of cuticular wax, contributing to the adaptation of subterranean pegs to the belowground environment. We transiently overexpressed *AhKCS1* (AYAJ7Z) in tobacco leaves, and a subcellular localization assay showed that it was primarily located on membranes (Figure 4E). In addition, owing to increased wax accumulation, leaves of *AhKCS1*-

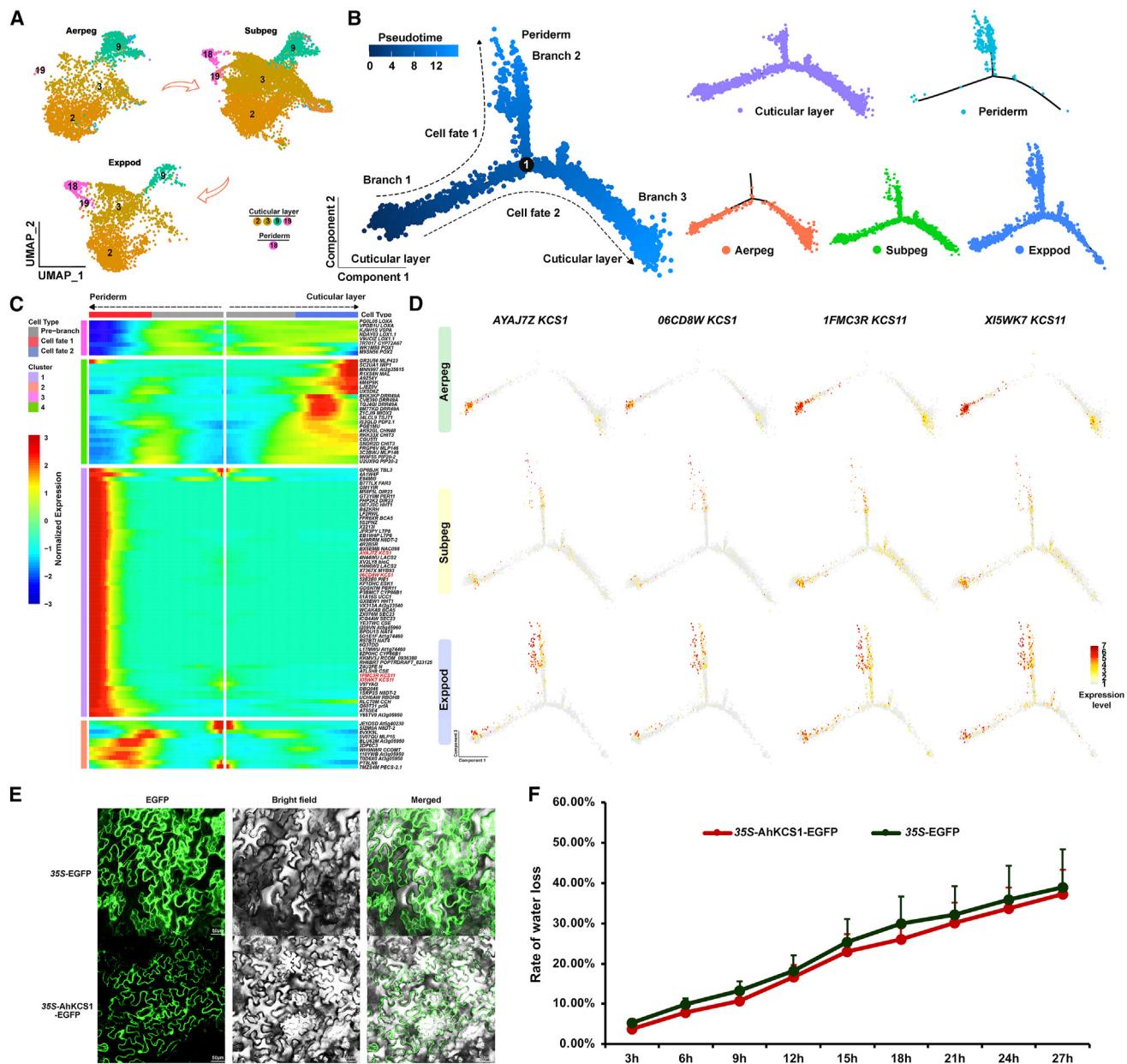
overexpressing plants had a lower rate of water loss than those of the empty-vector control (Figure 4F). We performed qPCR to confirm this pattern during pod development and revealed upregulation of these two genes after peg penetration (Supplemental Figure 10C). In examining the epidermis, we used our ATAC data to identify motifs unique to each of the three samples in the cuticular layer (Supplemental Figure 10D; Supplemental Table 7). Remarkably, the top three motifs specific to the cuticular layer of Aerpeg showed potential binding sites for bHLH (Basic Helix-Loop-Helix) transcription factors (Supplemental Figure 10E), underscoring the pivotal role of bHLHs in the development and maintenance of aboveground peanut pegs.

We also identified pseudo-temporal DEGs during the process of epidermal cell development (Supplemental Table 6). KEGG analysis of the top 500 DEGs revealed a significant enrichment of ribosome-related genes (Supplemental Figure 11A). This result could be attributed to the presence of a large number of ribosomal proteins among the top 500 DEGs (Supplemental Table 6). Hormone-related genes and transcription factors were extracted from the top 500 DEGs, as these genes may play a role in the development of the epidermis for environmental adaptation (Supplemental Figure 11B and 11C). Among these genes, two NAC021 genes were specifically expressed only in epidermal cells (Supplemental Figure 11C and 11D), consistent with the results of *in situ* RNA hybridization (Supplemental Figure 5). NAC021 genes were predominantly expressed in epidermal cells of Subpeg, as confirmed by qPCR (Supplemental Figure 11E). In *Arabidopsis*, NAC021 is induced by auxin and mediates auxin signaling to promote lateral root development (Xie et al., 2000). The ATAC data revealed open chromatin in the promoter regions of both NAC021 genes, approximately -143 bp upstream of the ATG, enriched with the TATA box, which potentially contributed to gene expression (Supplemental Figure 11F).

### Differentiation trajectories of parenchyma cells within the peanut shell

Expansion of the peanut pod begins with the development of parenchyma cells inside the shell, creating space for subsequent growth of the seeds, especially for stele cells (Figure 1F). We compared the proportion changes in each cluster of cells during development of the stele. We determined that cluster 0 was associated with parenchyma cells during pod development. Pseudo-time analysis was performed on cells from cluster 0 (Figure 5A and 5B). Auxin is a key component in seed development (Cao et al., 2020). Interestingly, among the top 200 pseudo-time DEGs, auxin-related genes, including *AUX28* (auxin/indole-3-acetic acid 28), *IAA9* (indole-3-acetic acid 9), *YAB5*, and *ARF* (auxin response factor), were identified and demonstrated an upregulation pattern with the progression of pseudo-time (Figure 5C; Supplemental Table 8). A dot plot was used to visualize the expression of hormone-related genes and transcription factors from among the top 200 pseudo-time DEGs in different clusters and samples (Figure 5D).

Two *YAB5* genes and three *MADS*-box transcription factor genes (two *AGL5* and one *AGL1*) were specifically expressed in cluster 0 and were upregulated during pod development (Figure 5D and



**Figure 4. Pseudo-time analysis of epidermal cells.**

(A) UMAP visualization highlighting the epidermal cells (cuticular layer and periderm) in three samples for pseudo-time analysis. Clusters 2, 3, 9, 18, and 19 were used.

(B) Monocle 2 analysis reveals the differentiation trajectory of epidermal cells, specifically the cuticular layer and periderm cells (left). The distribution of epidermal cells in cell trajectories for various cell types (cuticular layer and periderm) and samples (right). This section visually represents the distribution of cell types (upper part: cuticular layer and periderm) and the three samples (lower part) on a pseudo-timeline.

(C) Hierarchically clustered heatmap showing the top 100 DEGs, based on branch-dependent analysis, which determine cell differentiation toward the periderm (branch 2) or cuticular layer (branch 1). These genes were divided into four groups according to their expression patterns.

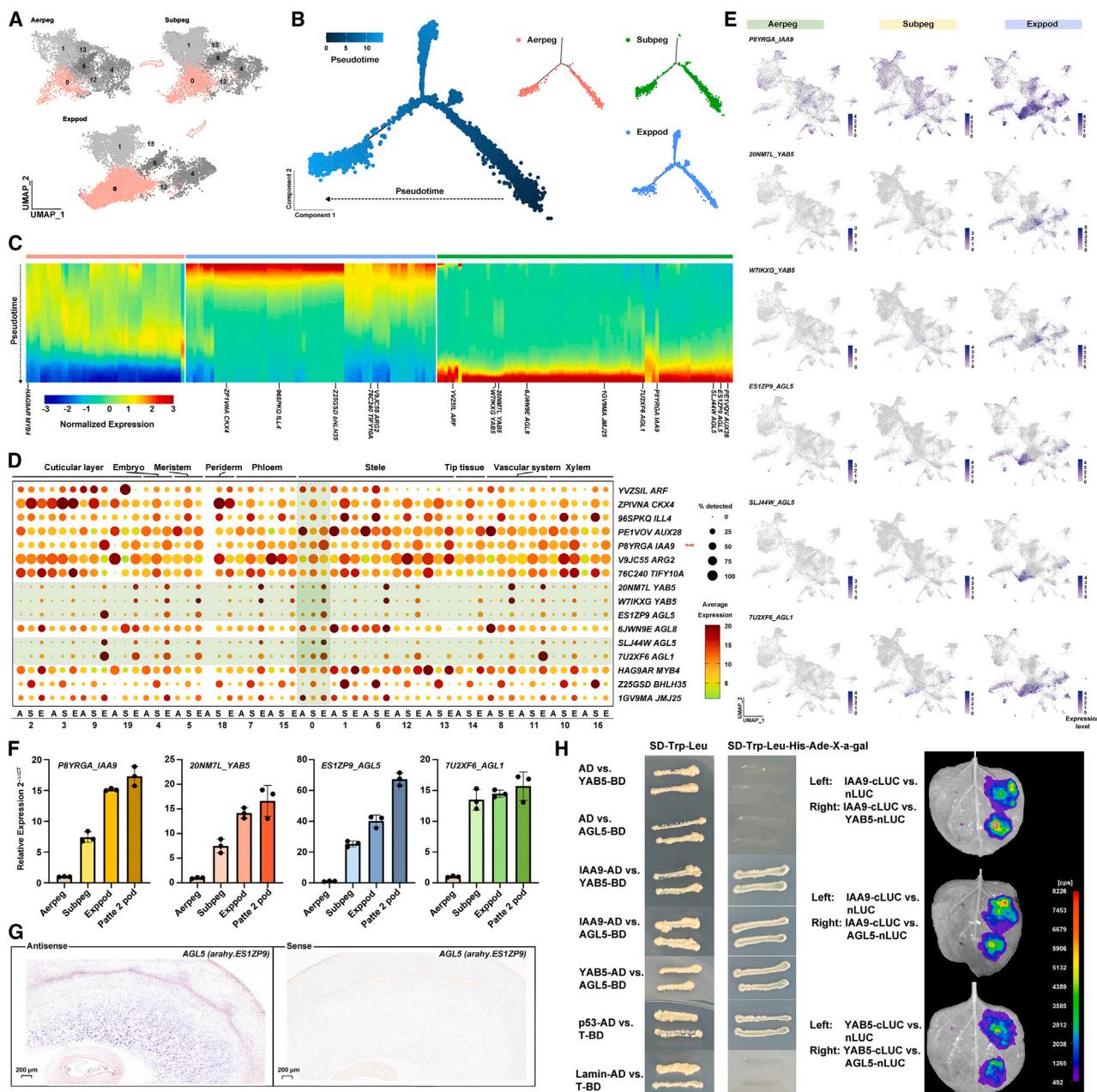
(D) Cell trajectory showing the expression of four KCS genes in three samples. The method used to calculate expression levels is the same as that in Figure 2D.

(E) Subcellular localization of AhKCS1.

(F) Measurements of leaf water loss showed that leaves transiently expressing AhKCS1 lost water more slowly than leaves expressing the empty-vector control.

5E). *IAA9* was also upregulated during pod development (Figure 5D and 5E). Interestingly, *IAA9* was also the key gene identified as affecting pod development in our previous bulk transcriptome study (Cui et al., 2022). qPCR results for the aforementioned

genes confirmed this expression pattern, revealing their upregulation during pod development (Figure 5F). We analyzed the chromatin-enriched peaks for the *IAA9*, *YAB5*, *AGL1*, and *AGL5* genes across the stele in three different samples

**Figure 5. Pseudo-time analysis of cells in cluster 0 (parenchyma cells).**

- (A) UMAP visualization emphasizing cluster 0 in three samples for pseudo-time analysis. Cells from cluster 0 are highlighted in pink.
- (B) Illustration of the cell distribution in cell trajectories across different samples. The left side displays the differentiation trajectory of cluster 0 cells, and the right side illustrates the distribution of cluster 0 cells in cell trajectories from different samples (Aerpeg, Subpeg, and Expod).
- (C) Heatmap showing the top 200 significant DEGs along pseudo-time. Genes related to plant hormones and TFs are marked. The genes were divided into three groups according to their expression patterns. The detailed gene list is provided in [Supplemental Table 8](#).
- (D) Dot plot depicting the expression of plant hormone-related and TF genes identified from the top 200 significant DEGs along pseudo-time. The size of the dot indicates the percentage of cells expressing the gene within a cell type, and the color indicates the average expression level. Five genes shaded in green are highly expressed in cluster 0, with expression increasing during pod development: two YAB5s, two AGL5s, and one AGL1.
- (E) UMAP plot showing the expression of DEGs (two YAB5s, two AGL5s, and one AGL1) expressed preferentially in cluster 0, as identified in the dot plot. The method used to calculate expression levels is the same as that in [Figure 2D](#).
- (F) Relative gene expression of IAA9, YAB5, AGL5, and AGL1 during pod development as measured by qPCR.
- (G) RNA *in situ* hybridization of AGL5; scale bar, 200  $\mu$ m.
- (H) Interactions among IAA9, YAB5, and AGL5 as demonstrated in yeast two-hybrid (left) and luciferase complementation (right) assays. The left side represents the control group, and the right side represents the experimental group.

## Plant Communications

but did not observe any significant sample-specific peaks ([Supplemental Figure 12](#)).

In *Arabidopsis*, AGL5 and AGL1 are associated with the formation of the endosperm and inner seed coat layer ([Ehlers et al., 2016](#)). *In situ* RNA hybridization showed a distinct AGL5 signal in the parenchyma cell tissue of the inner wall of the shell near the seed during pod expansion ([Figure 5G](#)). Subcellular localization assays showed that AGL5, YAB5, and IAA9 were all localized in the nucleus ([Supplemental Figure 13](#)). Because of their similar gene expression patterns and localization, we investigated the relationships among AGL5, YAB5, and IAA9. Yeast two-hybrid and luciferase complementation assays demonstrated their interactions with each other ([Figure 5H](#)), suggesting that they may function together as a complex. Because *YAB5* and *IAA9* are auxin-related genes that are upregulated during pod development, we speculated that auxin plays a significant role in the development of parenchyma cells, with *YAB5* and *IAA9* serving as key positive regulators.

In a further exploration of stele tissue using ATAC data, we identified 36 distinct motifs in the enlarged pod, 24 of which were classified as ERF related ([Supplemental Figure 14](#)). This predominance of ERF-related motifs strongly suggests their key importance in the development and enlargement of subterranean pods. This finding offers support to a previous hypothesis that expansion of underground pods is governed by changes in ethylene production, a response triggered by mechanical pressure ([Shlamovitz et al., 1995](#)).

### Gravitropism-related genes specifically expressed in vascular tissue

Gravitropism is a characteristic feature of peanut pegs and is necessary for the geocarpy of peanut. Interestingly, the top 10 marker genes in cluster 8 included auxin- and gravitropism-related genes such as *NPY5* (*naked pins YUC mutant 5*) and *SGR6* (*shoot gravitropism 6*). This drew our attention to this particular cluster ([Figure 6A](#)). *PHOT1* (*phototropin 1*), which encodes a blue-light photoreceptor, was also specifically expressed in cluster 8.

Genes related to auxins, gravitropism, and light that were specifically expressed in cluster 8 in the three samples ( $\log_2FC \geq 3$  in any of the samples) were selected for further investigation to determine whether their expression changed during fruit-pod development ([Supplemental Figure 15A](#)). With the exception of *ABC15*, all of these auxin-, gravitropism-, and light-related genes in cluster 8 were downregulated during fruit-pod development ([Supplemental Figure 15A](#)), and their expression patterns measured by qPCR were similar to those obtained by snRNA-seq ([Supplemental Figure 15B](#)).

Pseudo-time analysis was performed for the cells of cluster 8 to identify more genes that regulate peg gravitropism or pod development. The cells of Aerpeg were found at the start of the pseudo-time path (branch 1), whereas cells of Exppod were distributed in branches 2 and 3 ([Figure 6B](#)). According to similarities in their expression patterns, branch-determined DEGs were divided into four groups (groups A–D) ([Figure 6C](#)). Cell fate 1, progressing from branch 1 to 2, predominantly

### Single nucleus RNA-seq and ATAC-seq of peanut pods

contained cells from Aerpeg and Subpeg. Highly expressed genes at the end of cell fate 1 (group A), including *PIN3*, *PIN4*, two *SGR6s*, and two *PHOT1s*, might therefore serve as primary regulators of peg gravitropism in Aerpeg and Subpeg. By contrast, cell fate 2, progressing from branch 1 to branch 3, was exclusively enriched in cells from Subpeg and Exppod. Consequently, highly expressed genes at the end of cell fate 2 (group C) might act as primary regulators of pod development ([Figure 6C](#)). This group includes *CKX4* (*cytokinin dehydrogenase 4*) and *BGLU* (*beta-glucosidase*), suggesting their potential involvement in pod development through cytokinin and regulation of the cell wall.

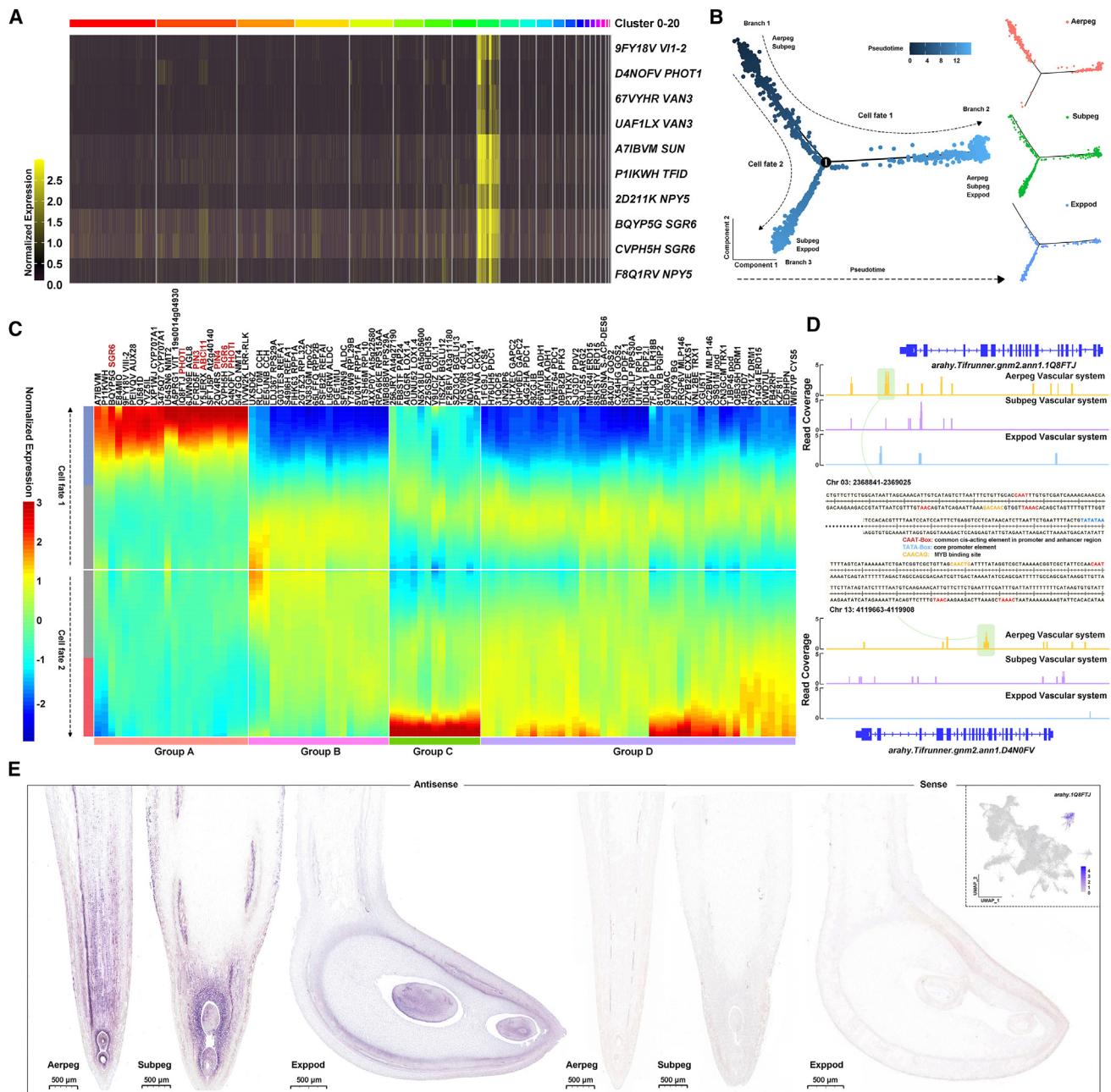
Given the role of *PIN* genes in auxin transport and gravity response, we examined the expression of all previously identified *PIN* genes in peanut across different samples and cell types ([Bian et al., 2023](#)) ([Supplemental Figure 16](#)). This analysis revealed distinct expression patterns for 21 genes. Notably, the four identified *PIN* genes (*FCRD5F*, *L76SLB*, *2QVARS*, *72M571*) in cluster 8 exhibited the highest expression levels among all *PIN* genes. In addition, four other *PIN* genes (*K2GLG1*, *13W9YS*, *953E9B*, *LAPE95*) were predominantly identified in the phloem and upregulated during pod development, suggesting that they may contribute to pod enlargement ([Supplemental Figure 16](#)).

Our data revealed cell-type-specific variations in accessibility across the majority of major cell types, enabling us to identify sets of transcription factors associated with accessible chromatin regions. By comparing the chromatin accessibility of two *PHOT1s* in the vascular system, we found that chromatin accessibility decreased during pod development in the three samples ([Figure 6D](#)). Moreover, two distinct open regions were identified in the promoter of *1Q8FTJ* and the gene region in the Aerpeg sample. Both regions contained the CAAT box, which can enhance gene expression. In addition, they shared a CAACAG MYB-binding site, suggesting regulation by MYB transcription factors. The open region of the *1Q8FTJ* promoter also included a TATA box ([Figure 6D](#)). Next, the chromatin-enriched peaks across distinct cell types for auxin-, gravitropism-, and light-related genes were analyzed ([Supplemental Figure 17](#)). We found higher peaks in the promoter regions of *PHOT1* (*1Q8FTJ*) and *ABCI11* (*Y5JERX*) within the vascular system, but these were not particularly significant. *In situ* RNA hybridization confirmed that *PHOT1* was predominantly expressed in vascular cells ([Figure 6E](#)). This hinted at a possible correlation between light response and gravitropism.

Our ATAC-seq analysis across Aerpeg, Subpeg, and Exppod vascular tissues identified 7, 47, and two unique motifs, respectively. Motifs specific to PIF and SGR were found in Aerpeg, and motifs specific to HY5, HYH, and PIF1 were found in Subpeg ([Supplemental Figure 18](#)). These findings are significant, as PIF, HY5, and HYH are key transcription factors in light response, especially HY5 and HYH, which promote photomorphogenesis ([Han et al., 2020](#)). These results, highlighting motifs associated with light-responsive factors, align with and reinforce the critical role of light in pod development.

### Changes in the expression of light-related genes

Given that light is a crucial factor in peanut fruit-pod development ([Kumar et al., 2019](#)) and that *PHOT1* was identified as a marker

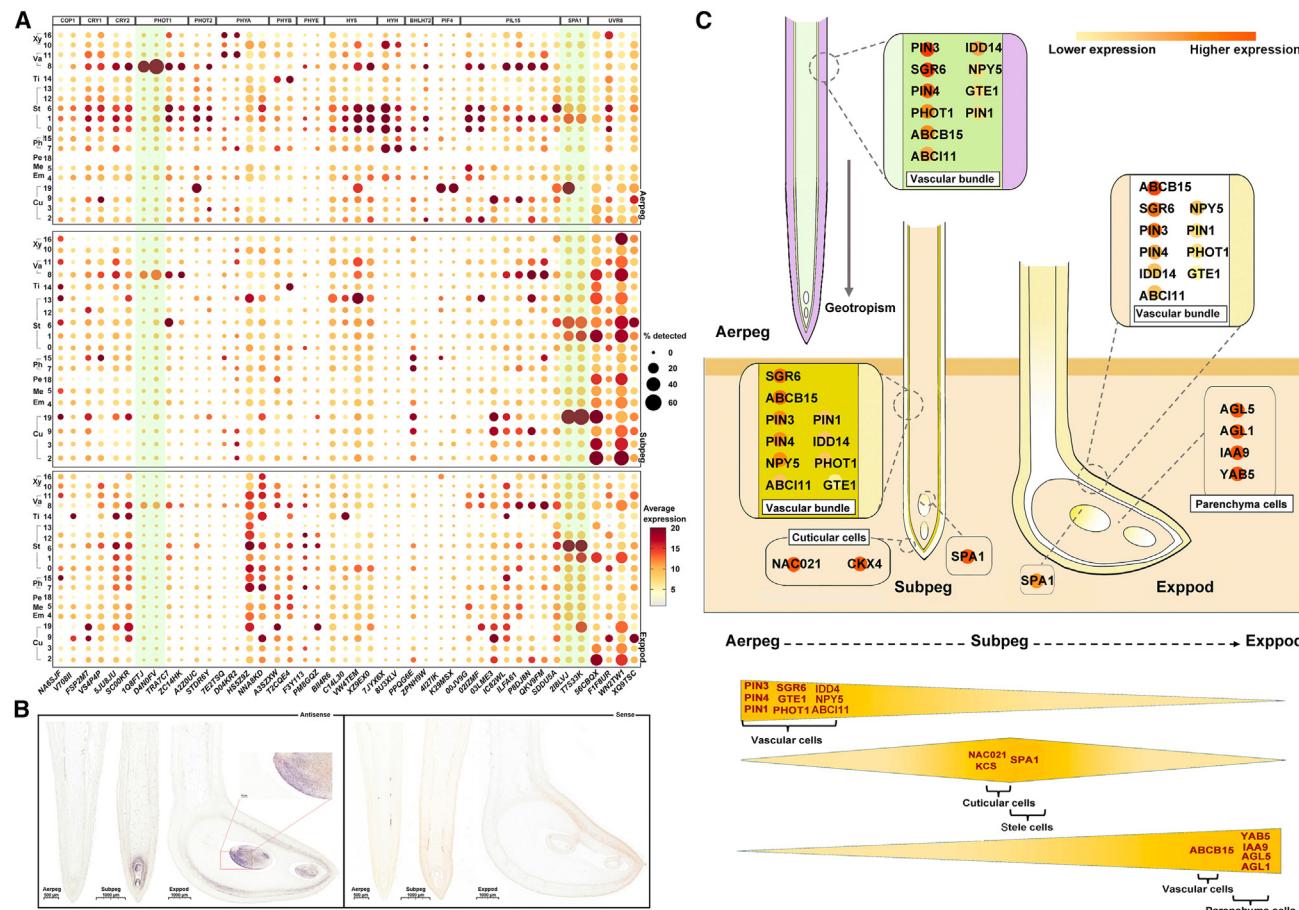


**Figure 6. Geotropism-related marker genes in cell cluster 8.**

- (A) The top 10 marker genes in cluster 8. In the heatmap, genes with high expression levels are represented in a prominent yellow color, making them easily distinguishable.
- (B) Pseudo-time distribution of cells in cluster 8. The left part displays the differentiation trajectory of cluster 8 cells, and the right part illustrates the distribution of cluster 8 cells in cell trajectories for different samples (Aerpeg, Subpeg, and Expod).
- (C) Hierarchically clustered heatmap of branch-dependent DEGs. Cell fate 2 determines cell differentiation toward the Subpeg and Expod. These genes were divided into four groups according to their expression patterns.
- (D) Evaluation of gene activity for two *PHOT1* genes in three samples, accompanied by the identification of motifs within the open regions, specifically within Aerpeg cells. The middle part of the figure displays the specific sequence of peaks, annotated with motifs.
- (E) RNA *in situ* hybridization of *PHOT1*; scale bar, 500 µm.

gene for cluster 8, it is necessary to investigate the specific expression patterns and levels of key genes related to light. Key light-related genes were identified through searches for homologous genes in *Arabidopsis* (Supplemental Figure 19). In this single-cell transcriptome dataset, *PHOT1* was a particular

gene of interest, as it was the only one of these genes with clear cell-cluster specificity, and its expression decreased with pod development (Figure 7A). Two SPAs also had interesting expression patterns: they were mainly expressed in clusters 1 and 6 in stelle cells and were upregulated during pod development.



**Figure 7. Expression of genes related to light and a diagram summarizing regulatory genes during fruit-pod development.**

(A) Dot plot showing the expression of key light-related genes during early pod development. The genes were identified by homologous alignment; the conserved domain is shown in [Supplemental Figure 19](#). The genes shaded in green showed significant changes in expression during early pod development.

(B) RNA *in situ* hybridization of SPA1 (*araby.2/8LVJ*) in three samples; scale bars, 500, 1000, and 1000 µm for Aerpeg, Subpeg, and Exppod, respectively.

(C) Schematic diagram detailing gene expression changes and positions during fruit-pod development (upper part) and identification of major genes expressed in different samples of fruit-pod development (lower part). GTE1, general TF group E6 1; PHOT1, phototropin 1; SGR6, shoot gravitropism 6; NPY5, naked pins YUC mutant 5; PIN, pin-formed; IDD14, indeterminate-domain 14; ABC, ATP-binding cassette; KCS, 3-ketoacyl-coenzyme A synthase; NAC, NAM, ATAF, and CUC1/2; SPA, suppressor of phyA; AGL5, agamous-like MADS-box protein 5; IAA9, indoleacetic acid-induced protein 9; YAB5, axial regulator YABBY 5.

They were primarily expressed in cluster 1 and 6 in Aerpeg ([Supplemental Figure 20A](#)). Their expression range expanded in Subpeg and Exppod, and they were also expressed in vascular and epidermal cells. The gene activity of the two SPAs was consistent with their gene expression ([Supplemental Figure 20A](#)). Analysis of chromatin-enriched peaks across various cell types and different samples for the two SPAs revealed that higher peaks were observed in the promoter regions of these genes in Subpeg ([Supplemental Figure 20B](#)).

The SPA expression pattern revealed by *in situ* RNA hybridization was consistent with the transcriptome results. Expression of this gene was not detected in Aerpeg, probably because its expression level was too low to be detected by *in situ* hybridization. SPA was expressed in and around the testa of Subpeg, whereas in Exppod, the hybridization signal was detected only in the testa ([Figure 7B](#)). SPA, also known as suppressor of PhyA, collaborates to suppress photomorphogenesis in *Arabidopsis thaliana* by

targeting key transcription factors and phytochrome A for degradation through the 26S proteasome system ([Zhu et al., 2008](#)). It is plausible that SPA promotes fruit-pod development by promoting the proteasome-mediated degradation of pod-enlargement inhibitors.

Expression of the two HYH (*elongated hypocotyl 5 homolog*) genes disappeared after peg penetration ([Figure 7A](#)). HYH, which is involved in the phyB signaling pathway ([Holm et al., 2002](#)), likely plays a role in the regulation of fruit-pod development through light signaling. The presence of light-related genes such as SPA and HYH further supported the notion that light is a crucial factor in the regulation of peanut fruit-pod enlargement at the molecular level. Two UVR8 (*UV resistance locus 8*) genes had the highest expression levels in Subpeg and Exppod ([Figure 7A](#)). We speculated that they might be associated with environmental adaptation and resistance, aiding in adaptation of the penetrating peg to the soil environment.

## DISCUSSION

In this study, we investigated cell types and gene expression profiles at three important stages of peanut fruit-pod development at a single-cell resolution. Gravitropism is an important response of the peanut peg, and we found that genes associated with gravitropism, including *PINs*, *NPY5s*, *SGR6s*, *IDD14*, *ABCI11*, and *ABCB15*, were specifically expressed in peg vascular cells (Figure 7C). *KCS* and *NAC021* were specifically expressed in cuticular cells and may be involved in the regulation of environmental adaptive differentiation. Two auxin-related genes, *YAB5* and *IAA9*, and two *MADS-box* genes, *AGL1* and *AGL5*, were specifically expressed in spongy parenchyma cells inside the pod shell of Eppod. Three highly expressed genes (*GTE1*, *PHOT1*, and *SPA1*) associated with light were identified and may be involved in pod development under dark conditions. *GTE1* and *PHOT1*, which were most highly expressed in vascular tissue, were downregulated during fruit-pod development. By contrast, *SPA1*, which was predominantly expressed around the seed coat, was upregulated during fruit-pod development (Figure 7C). Our study provides novel insights for future investigations of peanut fruit-pod development.

### Gravity and light-related genes associated with pod development

Genes related to auxin and gravitropic responses were specifically expressed in the vascular bundle of the peg (Figure 6). Among them, *NPY5s*, *SGR6s*, and *IDD14* are known to be related to gravitropic responses in *Arabidopsis*. *NPY5* acts as a substrate-specific adapter of an E3 ubiquitin-protein ligase complex and plays an essential role in root gravitropic responses (Li et al., 2011). *SGR6* is a membrane-associated protein that regulates shoot gravitropism by modulating vacuolar membrane dynamics in gravity-sensing cells (Hashiguchi et al., 2014). *SGR6* mutants show reduced gravitropism in inflorescence stems but normal gravitropism in hypocotyls and roots of *Arabidopsis* (Yamauchi et al., 1997). A single-cell analysis of the *Arabidopsis* vegetative shoot apex identified a suite of genes enriched in the shoot endodermis that are involved in the shoot gravitropic response; these included *SGR6*, consistent with our results (Zhang et al., 2021b). The expression pattern of *SRG6* in peanut implies that the peg has stem-like properties. *IDD14* is a transcription factor that regulates spatial auxin accumulation and thereby controls organ morphogenesis and gravitropic responses (Cui et al., 2013). Why these genes are specifically expressed in vascular bundles is an intriguing question. Vascular bundles are the main transport system in plants, and whether the mRNAs of these genes are transported through the vasculature is also worthy of future research.

Many years ago, researchers demonstrated through ingenious experiments that dark conditions are necessary for peanut pod development (Zamski and Ziv, 1976). However, the key genes that regulate this process are poorly understood. Bulk omics studies suggested that *phyA* (*phytochrome A*) may be a key gene, but this possibility has not been confirmed (Liu et al., 2020; Cui et al., 2022). Through single-cell sequencing, we identified three key genes, *GTE1*, *PHOT1*, and *SPA1*, that showed significant changes at the transcriptional level during fruit-pod development. *GTE1* and *PHOT1* were especially highly ex-

pressed in vascular cells, implying the existence of crosstalk between the light and gravitropism pathways. In *Arabidopsis*, *GTE1* is a transcriptional activator that promotes seed germination by negatively and positively regulating the abscisic acid and *phyA* signal transduction pathways, respectively (Duque and Chua, 2003). *PHOT1* is a protein kinase that acts as a blue-light photoreceptor in a signal transduction pathway responsible for photo-induced movements. *PHOT1* also contributes to chloroplast accumulation (Kagawa and Wada, 2002) and may thus be involved in greening of the light-exposed pod shell. *SPA1* is a component of the COP1/SPA E3 ubiquitin-protein ligase complex, which acts as a negative regulator of *phyA*-mediated de-etiolation of young seedlings and regulates circadian rhythm in plants (Ishikawa et al., 2006).

### Multiple signal pathways involved in development of the spongy parenchyma of the fruit pod

As the number of cells in the spongy parenchyma increases, the space for seed expansion gradually increases, and the seed begins to grow (Figure 1F). Thus, the onset of spongy parenchyma development is fundamental to pod enlargement. Three auxin-related genes (one *IAA9* gene and two *YAB5* genes) were especially highly expressed in spongy parenchyma cells (Figure 5). *IAA9* is considered to be a short-lived transcription factor; it interacts with ARFs and alters their ability to modulate expression of early auxin response genes (Liscum and Reed, 2002). In *Arabidopsis*, *YAB5* regulates the initiation of shoot apical meristem development in the embryo (Stahle et al., 2009). In flowering plants, there is clear evidence that *MADS-box* genes are involved in the development of reproductive structures (Ng and Yanofsky, 2001). In this study, three *MADS-box* genes (one *AGL1* gene and two *AGL5* genes) were specifically expressed in newly differentiated spongy parenchyma cells (Figure 5). These three genes have extremely high amino acid sequence similarity (Supplemental Figure 21). In *Arabidopsis*, At*AGL1* and At*AGL5* share 85% sequence similarity at the amino acid level, have similar expression patterns, and show functional redundancy. At*AGL1* and At*AGL5* are required for formation of the dehiscence zone (Savidge et al., 1995; Ferrández et al., 1999).

In peanut, we observed interactions of the auxin-related proteins *YAB5* and *IAA9* with *AGL5* (Figure 5H). In rice, mediator subunits interact with specific transcription factors to regulate various processes, including seed development. Notably, OsMED14\_1 has been experimentally demonstrated to interact physically with the transcription factors *YAB5* and *MADS29* (Malik et al., 2020). Mediator functions as a complex co-activator with multiple subunits. Recognized components of the mediator complex include the head, middle, tail, and kinase modules. Normally, the tail module interacts with transcriptional activators or repressors, whereas the middle and head modules interface with RNA polymerase II and the transcriptional machinery. Consequently, the mediator complex serves as an intermediary, regulating gene transcription by mediating communication between transcription factors and the transcriptional machinery (Soutourina, 2018; Malik et al., 2020). On the basis of our discovery of a transcription factor complex in peanuts, we postulate the existence of a mediator complex responsible for orchestrating the development of peanut pods. Environmental signals could potentially be transmitted to

## Plant Communications

auxin-related proteins via light-related proteins. Moreover, there is speculation that auxin signaling-related proteins, either independently or in collaboration with AGL5 and AGL1 through the mediator complex, play a role in regulating and promoting pod enlargement.

In summary, this study mapped the dynamic molecular profiles of peg penetration and pod expansion, enabling a better understanding of the heterogeneity of peanut fruit-pod development. Identification of functional marker genes associated with gravitropism, plant hormone signals, and light signals can help us understand the mechanism of peanut geocarpy and facilitate future improvements in peanut quality and yield.

## METHODS

### Plant material and growth conditions

*A. hypogaea* cv. Tifrunner (Bertioli et al., 2019), a runner-type peanut cultivar, was planted at the Experimental Station of the Institute of Advanced Agricultural Sciences of Peking University and used as the experimental material. Peanut pegs/pods were collected from 30 individual plants at each of three developmental stages: Aerpeg (green or purple aerial-grown pegs), Subpeg (pegs that had been in the soil for about 2–3 days and in which white swelling was not detected), and Exppod (subterranean pegs that had expanded into small pods with a maximum transverse diameter of about 4–5 mm at about 4–7 days after soil penetration).

### Observation of anatomical characteristics

Peg/pod tips about 10 mm in length (Supplemental Figure 2), including seeds, were obtained at the Aerpeg, Subpeg, and Exppod stages, cleaned, and immediately placed into fixing fluid (G1110, Servicebio, Wuhan) for 24 h. The tissue was then dehydrated through an alcohol gradient, embedded in paraffin, and sectioned longitudinally to a thickness of about 10 µm, passing through the center of the tissue and the seed. After drying at 62°C for 2 h, sections were deparaffinized and hydrated in an ethanol gradient series (100% twice for 5 min each, 85% and 75% for 5 min each, and diethylpyrocarbonate water). The sections were placed into toluidine blue staining solution for about 2 min, then dried in an oven at 60°C after washing with running water. The dried sections were cleared with xylene for 5 min and sealed with neutral gum. The sections were observed under a microscope (Nikon DS-U3, Japan). Longitudinal sections of six representative pegs/pods, passing through the middle axis, were used to calculate the transverse diameter of the pegs/pods (across the upper seed) and the seeds (upper side, farthest from the tip).

### Library construction and single-nucleus RNA-seq and ATAC-seq

About 10 mm of tips containing seeds from the three samples, each from 30 individual plants per stage, were excised (Supplemental Figure 2). Nuclei were extracted using an improved liquid nitrogen method as described in the supplementary information. The three RNA-seq and ATAC-seq libraries for each of the three samples were constructed according to the instructions of the Chromium Next GEM Single Cell Multiome ATAC + Gene Expression Reagent Bundle (10x Genomics, PN-1000283, San Francisco, CA). In brief, the nuclear suspension was incubated with transposase mixture, which cut the open chromatin regions, to fragment the DNA, and a sequencing primer was ligated to the ends of the DNA fragments. Next, gel beads with barcode tags, two capture sequences (one for the RNA-seq library and one for the ATAC-seq library), and a single nucleus were wrapped in oil droplets to form gel beads in emulsion (GEMs). Subsequently, the gel beads were dissolved, and the nuclei broke, releasing DNA fragments, pre-mRNA molecules, and mature mRNA molecules, which were tagged at the ends.

### Single nucleus RNA-seq and ATAC-seq of peanut pods

Poly(dT) sequences were used to capture pre-mRNA and mature mRNA. After the GEM structure broke, all the cDNA was mixed for PCR amplification to obtain sufficient cDNA for library construction. The cDNA was digested into fragments of about 200–300 bp, and a standard sequencing library was obtained by PCR amplification after completing conventional second-generation sequencing library construction steps such as end repair, A-tailing, and adaptor ligation (Van Dijk et al., 2014). Spacer sequences, which enable barcode attachment to transposed DNA fragments for the ATAC library, were used to capture DNA fragments. After the GEM structure broke, all double-stranded DNA was mixed, and PCR amplification was performed to obtain sufficient DNA. The sequencing connector P7 and sample index were added to the DNA fragments by PCR to build a standard sequencing library (Supplemental Figure 2).

The constructed libraries were sequenced using the paired-end sequencing mode of the Illumina sequencing platform. For the snRNA-seq library, each read contained 16 bp of barcode information and a 10-bp unique molecular identifier (UMI) for determining the source of the cells and quantifying the amount of expression. For the snATAC-seq library, each read contained 16 bp of barcode information, which was used to match the sequencing data of two libraries generated from the same cell.

### Raw data preprocessing for snRNA-seq and snATAC-seq

The raw scRNA-seq data underwent quality control and alignment to the reference genome (version 2 of *A. hypogaea* cv. Tifrunner) using the official 10x Genomics analysis software, Cell Ranger ARC (<https://support.10xgenomics.com/single-cell-multiome-atac-gex/software/overview/welcome>). Supplemental Table 1 contains comprehensive Cell Ranger reports. The “filtered\_gene\_bc\_matrices” generated by 10x Genomics served as processed raw data for subsequent analyses.

Similarly, the raw snATAC-seq reads were processed and aligned to the reference genome (version 2 of *A. hypogaea* cv. Tifrunner) using Cell Ranger ATAC (v2.0). Detailed Cell Ranger reports are available in Supplemental Table 1. The filtered\_gene\_bc\_matrices produced by 10x Genomics served as the processed raw data for further analyses. The full-featured R package Signac (v1.1.1) (Stuart et al., 2021) was used to complete the subsequent analysis.

### Data quality control and batch effect correction

Doublets in each dataset were identified using DoubletFinder (v.2.0.3) (McGinnis et al., 2019). The UMI count matrix underwent processing with the Seurat R package (v.4.3.0) (Hao et al., 2021). To eliminate dead cells and dissociative RNAs, the following filtering criteria were applied: cells with fewer than 300 unique gene counts and genes expressed in fewer than three cells were excluded; cells with UMIs above 10 000 or below 500 were filtered out; and cells with unique gene counts greater than 5800 or fewer than 200 were removed from the analysis. After the removal of low-quality cells, batch effect correction was carried out using the Harmony algorithm (Korsunsky et al., 2019) to mitigate batch effects on cell clustering. The corrected Harmony feature space was used for cell clustering and visualization.

### Cell clustering, visualization, and identification of cell types

Linear dimensionality reduction of the integrated scRNA-seq data was performed using the “RunPCA” function in the Seurat R package. For principal component (PC) analysis, the scaled data were reduced to approximately 100 PCs (set n\_pcs = 100). UMAP dimensional reduction was employed for two-dimensional visualization using 60 PCs. Subsequently, cells were clustered on the basis of the top 60 PCs using the Louvain algorithm. This clustering process was accomplished using the functions “FindNeighbors” and “FindClusters” (resolution = 0.5) in Seurat.

For the “FindAllMarkers” function of Seurat, we set logfc.threshold = 0.25 and min.pct = 0.25 to identify cluster-enriched genes. Clusters were

## Single nucleus RNA-seq and ATAC-seq of peanut pods

classified and annotated on the basis of the known functions and expression patterns of genes enriched in each cluster ([Supplemental Table 2](#)). The expression of representative marker genes was visualized using the “DotPlot” function of Seurat.

The rank sum test in Seurat was used to analyze the upregulated DEGs in different cell clusters. The criteria for differential expression were  $\log_{2}FC \geq 0.36$ ,  $p \leq 0.01$ , and expression in more than 25% of cells in the cluster. The upregulated DEGs in cell clusters were further analyzed for cell-type identification. These upregulated DEGs were aligned to cell-type-specific gene databases of other plant species, such as *Arabidopsis* and rice (<https://www.tobaccodb.org/pcmbdb/homePage>) ([Jin et al., 2022](#)), to infer the cell type of each cluster in peanut. In addition, protein sequences of experimentally verified cell-type-specific genes from other plants (<https://www.tobaccodb.org/pcmbdb/homePage>) ([Jin et al., 2022](#)) were extracted and blasted against the peanut protein sequence database using BLASTP. The genes with the highest scores were selected from the blast results, and their sequence similarity was checked. Finally, expression of these homologous peanut genes was visualized by UMAP to aid in deducing the cell type corresponding to the cell cluster.

### Read alignment, peak scanning, and motif analysis

The clean reads from each sample were aligned to the reference genome using Bowtie (version 2) as described by [Langmead and Salzberg \(2012\)](#). Duplicate data files were removed using Picard (<http://broadinstitute.github.io/picard/>). All aligned paired mates were used for peak calling with MACS2 ([Zhang et al., 2008](#)) using the parameters: “–shift -75 –extsize 150 –nomodel –call-summits –nolambda –keep-dup all -p 0.01.” The MEME suite (<http://meme-suite.org/>) was used to detect motifs. Specifically, MEME-ChIP was used to scan motifs with high reliability through peak regions, and MEME-AME was used to confirm the existence of any specific known motifs.

### Cell-cycle analysis

The gene set for cell-cycle assessment was constructed according to ([Macosko et al., 2015](#)) and Cyclebase (<https://cyclebase.org>). Seurat was used to calculate a cell-cycle score for each cell based on the expression of marker genes in each cell-cycle stage. For a single cell, the difference between the average expression of the cycle characteristic gene set and the background gene set in the cell was taken as the cell-cycle score. Each cycle characteristic gene set was scored in turn. The stage with the highest score represented the cell-cycle stage, and the corresponding score was the cell-cycle score. If the score of a cell in each cell cycle was lower than 0.3, the cell was considered to be a non-cell-cycle cell ([Neftel et al., 2019](#)).

### Pseudo-time analysis of subclusters

Unsupervised ordering of subclusters was performed using the Seurat results as input to construct a tree-like differentiation trajectory. This was achieved using the DDRTree algorithm of the Monocle v2 R package ([Trapnell et al., 2014](#)), and the data were reduced to a two-dimensional plane. In the lower-dimensional space, the “orderCells” function was used to arrange the cells in pseudo-time on the basis of transcriptome correlation. In addition, we used the “plot\_cell\_trajectory” and “plot\_pseudo-time\_heatmap” functions to visualize the cell trajectory. The branch-dependent genes were analyzed using the “BEAM” function. We used “plot\_genes\_branched\_heatmap” to illustrate the bifurcation of gene expression along the two branches, and DEGs were then identified using the “~differentialGeneTest” function with false discovery rate (FDR)  $<1e-7$ . These genes were used as features to reconstruct cell trajectories with Seurat.

### Correlation analysis of single-cell and bulk data

For the bulk data, expression matrices were retrieved from the NCBI Sequence Read Archive (SRP064700) ([Zhang et al., 2016](#)). The

## Plant Communications

expression data were consolidated for both the single-cell data and bulk data, considering overlapping genes. The transcripts per million expression values for each gene in the bulk RNA-seq data were computed. For the correlation analysis, we calculated both Pearson and Spearman’s rank correlations using R.

### KEGG enrichment analysis

KEGG enrichment analysis of DEGs in various cell types of the three samples was performed using the KEGG database (<https://www.kegg.jp>). The resulting  $p$  values underwent FDR correction, with a threshold of FDR  $\leq 0.05$  used to define significantly enriched pathways.

### Phylogenetic reconstruction in peanut, soybean, rice, and *Arabidopsis*

The protein sequences of candidate markers ([Supplemental Table 3](#)) served as templates for BLASTP searches against all genes, applying an e-value threshold of 1e-5. Subsequently, the candidate sequences were aligned using MAFFT ([Katoh et al., 2002](#)) and trimmed using trimAL ([Capella-Gutierrez et al., 2009](#)). A phylogenetic tree of the candidates was constructed using the maximum likelihood method implemented in FastTree ([Price et al., 2010](#)) with 1000 bootstraps. The resulting trees were visualized using iTOL (<https://itol.embl.de>).

### In situ RNA hybridization

Sections used for *in situ* RNA hybridization were boiled in sodium citrate buffer (1 M, pH 6.0) for 10–15 min and naturally cooled. Proteinase K (20 µg/ml) was added, and the slides were incubated at 37°C for 30 min. The slides were washed three times with PBS (pH 7.4) for 5 min each. Pre-hybridization solution was added to each section, and slides were incubated for 1 h at 37°C. The pre-hybridization solution was discarded, and the slides were incubated in probe hybridization solution (500 nM) overnight at 42°C. Slides were washed with 2× saline sodium citrate (SSC) for 5 min at 37°C, 1× SSC twice for 5 min each at 37°C, and 0.5× SSC for 10 min at room temperature. Slides were incubated at 42°C for 3 h with a digoxin-labeled RNA probe and washed with SSC solution as described above. A blocking solution was added, and the slides were incubated at room temperature for 30 min. After removal of the blocking solution, the slides were incubated with an alkaline-phosphatase-conjugated immunoglobulin G fraction monoclonal mouse anti-digoxin antibody at 37°C for 40 min, then washed in Tris-buffered saline four times for 5 min each. The sections were stained with 5-bromo-4-chloro-3-indolyl phosphate/tetranitro blue tetrazolium chloride solution (SK20301, Coolaber, Beijing) and nuclear fast red solution. After drying, they were sealed with quick-drying sealing glue. The sections were observed under a microscope (Nikon DS-U3, Japan). Probe sequences are listed in [Supplemental Table 9](#).

### Real-time qPCR analysis

Four samples were used for qPCR analysis: Aerpeg, Subpeg, Exppod, and Pattee 2 pod ([Pattee et al., 1974](#)). RNA extraction and reverse transcription were performed as described previously ([Cui et al., 2019](#)). Primers were designed with Primer3 ([Untergasser et al., 2012](#)), and elongation factor 1B (*Arahy.E3HYWR*) served as the internal control ([Supplemental Table 9](#)). qPCR was performed on a QuantStudio 5 instrument (Applied Biosystems, CA, USA) using Hieff qPCR SYBR Green Master Mix (Low Rox Plus) (11202ES08, Shanghai, China) with three replicates. The reaction conditions were as follows: 95°C for 5 min, 40 thermal cycles of 95°C for 10 s, followed by 60°C for 30 s. The  $2^{-\Delta\Delta Ct}$  method was used to calculate relative gene expression ([Livak and Schmittgen, 2001](#)).

### Subcellular localization and leaf water-loss determination

Using the ClonExpress II One Step Cloning Kit (C112, Novozymes, China), the CDS (Coding Sequence) sequences of *KCS1*, *IAA9*, *YAB5*, and *AGL5* without stop codons were inserted into the pCambia-1302 vector; see [Supplemental Table 9](#) for primer details. The recombinant and empty

## Plant Communications

control vectors were transformed into *Agrobacterium tumefaciens* strain GV3101 by the freeze-thaw method. *Agrobacterium* cultures were prepared and injected into tobacco leaves. After 48 h, GFP fluorescence signals were observed using a Nikon laser scanning confocal microscope (Nikon, Tokyo, Japan).

To measure leaf water loss, six leaves each from the experimental and control groups were cut with scissors and placed in the same Petri dish. Each set was weighed, and the initial weight was recorded as W1 (the weight before water loss). The leaves were then weighed every 3 h (recorded as Wn) until they showed obvious signs of wilting, at which point weighing was stopped. The leaves were then dried to constant weight at 65°C, and the dry weight W0 was recorded. The water-loss rate was calculated as (W1 – Wn)/(W1 – W0).

### Yeast two-hybrid and luciferase complementation assays

Yeast two-hybrid assays were performed using the Matchmaker Gold Yeast Two-Hybrid Library Screening System (catalog no. 630491, Clontech). Full-length cDNAs of *IAA9* and *YAB5* were ligated into the pGADT7-AD vector, and full-length cDNAs of *YAB5* and *AGL5* were ligated into the pGBTKT7-BD vector. Primers are detailed in *Supplemental Table 9*. All constructs were transformed into the yeast strain Y2H Gold, and yeast cells were cultured on SD-Leu/-Trp medium in accordance with the manufacturer's guidelines. Transformed colonies were then plated onto SD-Leu/-Trp/-His/-Ade medium with β-galactosidase to assess potential interactions.

For luciferase complementation assays, *IAA9* and *YAB5* were cloned into p1300-35s-cLUC (C-terminal Luciferase), and *YAB5* and *AGL5* were cloned into p1300-35s-nLUC (N-terminal Luciferase) for the expression of fusion proteins. Primers can be found in *Supplemental Table 9*. The cLUC and nLUC fusion plasmids were transformed into *Agrobacterium* strain GV3101 and infiltrated into tobacco leaves. Tobacco plants were transferred to darkness for 12 h and then to weak light, and luciferase activity was measured using an LB985 NightShade system (Berthold Technologies) after 2 days of growth.

### Statistical analysis

The Benjamini and Hochberg method was used to identify cell-type-specific genes or peaks, and the significance threshold was set to  $p < 0.01$ . For cell trajectory analysis, genes that were differentially expressed with changes in differentiation status and pseudo-time value were identified using an FDR  $< 1e-5$ . A negative binomial generalized linear model was fitted for different branches, and genes with different expression patterns were identified using FDR  $< 1e-7$ . For data presented in bar charts, statistical analysis was performed using GraphPad Prism; *t*-tests were used to compare the means of two groups ( $**p \leq 0.01$ ).

## DATA AND CODE AVAILABILITY

The raw sequence data reported in this paper have been deposited at NCBI (<https://www.ncbi.nlm.nih.gov>) and in the China National Genomics Data Center (<https://ngdc.cncb.ac.cn>). The accession numbers are PRJNA1068272 and PRJCA026274, respectively. A website has been established to host the single-cell transcriptome data in order to enhance its accessibility and utilization (<http://8.140.245.74:6677/>).

## SUPPLEMENTAL INFORMATION

Supplemental information is available at *Plant Communications Online*.

## FUNDING

This study was supported by grants from the Taishan Scholar Foundation of Shandong Province (tsqn202103161), the Natural Science Foundation of Shandong Province (ZR202103010405), the Key R&D Program of Shan-

## Single nucleus RNA-seq and ATAC-seq of peanut pods

dong Province, China (ZR202211070163), the Foundation of President of the Peking University Institute of Advanced Agricultural Sciences (ZR202211070163), and the Peanut Seed Industry Project in Shandong Province, China (2022LZGC007) to X.L. It was also funded by the National Natural Science Foundation of China (NSFC) Key Program (32230006) to X.W.D.

## AUTHOR CONTRIBUTIONS

X.L. conceived and designed this study. X.L. and X.W.D. supervised the research. Y.Cui., Y.S., and X.L. performed the experiments and analyzed the data. Y.Cui., Y.S., X.L., X.W.D., J.B., X.H., H.G., Z.Y., Y.Chen., L.L., and T.L. wrote the manuscript and prepared the revised version. All authors prepared, read, and approved the manuscript for publication.

## ACKNOWLEDGMENTS

No conflict of interest is declared.

Received: July 12, 2023

Revised: January 31, 2024

Accepted: May 22, 2024

Published: May 24, 2024

## REFERENCES

- Bertioli, D.J., Jenkins, J., Clevenger, J., Dudchenko, O., Gao, D., Seijo, G., Leal-Bertioli, S.C.M., Ren, L., Farmer, A.D., Pandey, M.K., et al. (2019). The genome sequence of segmental allotetraploid peanut *Arachis hypogaea*. *Nat. Genet.* **51**:877–884. <https://doi.org/10.1038/s41588-019-0405-z>.
- Bessire, M., Borel, S., Fabre, G., Carraça, L., Efremova, N., Yephremov, A., Cao, Y., Jetter, R., Jacquat, A.C., Métraux, J.P., and Nawrath, C. (2011). A member of the PLEIOTROPIC DRUG RESISTANCE family of ATP binding cassette transporters is required for the formation of a functional cuticle in *Arabidopsis*. *Plant Cell* **23**:1958–1970. <https://doi.org/10.1105/tpc.111.083121>.
- Bian, J., Cui, Y., Li, J., Guan, Y., Tian, S., and Liu, X. (2023). Genome-wide analysis of PIN genes in cultivated peanuts (*Arachis hypogaea* L.): identification, subcellular localization, evolution, and expression patterns. *BMC Genom.* **24**:629. <https://doi.org/10.1186/s12864-023-09723-5>.
- Bonke, M., Thitamadee, S., Mähönen, A.P., Hauser, M.T., and Helariutta, Y. (2003). APL regulates vascular tissue identity in *Arabidopsis*. *Nature* **426**:181–186, PMID: 14614507. <https://doi.org/10.1038/nature02100>.
- Cao, J., Li, G., Qu, D., Li, X., and Wang, Y. (2020). Into the Seed: Auxin Controls Seed Development and Grain Yield. *Int. J. Mol. Sci.* **21**:1662. <https://doi.org/10.3390/ijms21051662>.
- Capella-Gutiérrez, S., Silla-Martínez, J.M., and Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**:1972–1973. <https://doi.org/10.1093/bioinformatics/btp348>.
- Chen, X., Yang, Q., Li, H., Li, H., Hong, Y., Pan, L., Chen, N., Zhu, F., Chi, X., Zhu, W., et al. (2016). Transcriptome-wide sequencing provides insights into geocarpy in peanut (*Arachis hypogaea* L.). *Plant Biotechnol. J.* **14**:1215–1224. <https://doi.org/10.1111/pbi.12487>.
- Cui, D., Zhao, J., Jing, Y., Fan, M., Liu, J., Wang, Z., Xin, W., and Hu, Y. (2013). The *arabidopsis* IDD14, IDD15, and IDD16 cooperatively regulate lateral organ morphogenesis and gravitropism by promoting auxin biosynthesis and transport. *PLoS Genet.* **9**:e1003759. <https://doi.org/10.1371/journal.pgen.1003759>.
- Cui, Y., Bian, J., Lv, Y., Li, J., Deng, X.W., and Liu, X. (2022). Analysis of the Transcriptional Dynamics of Regulatory Genes During Peanut Pod Development Caused by Darkness and Mechanical Stress. *Front. Plant Sci.* **13**:904162. <https://doi.org/10.3389/fpls.2022.904162>.

## Single nucleus RNA-seq and ATAC-seq of peanut pods

Cui, Y., Wang, Z., Chen, S., Vainstein, A., and Ma, H. (2019). Proteome and transcriptome analyses reveal key molecular differences between quality parameters of commercial-ripe and tree-ripe fig (*Ficus carica* L.). *BMC Plant Biol.* **19**:146. <https://doi.org/10.1186/s12870-019-1742-x>.

Denyer, T., Ma, X., Klesen, S., Scacchi, E., Nieselt, K., and Timmermans, M.C.P. (2019). Spatiotemporal Developmental Trajectories in the Arabidopsis Root Revealed Using High-Throughput Single-Cell RNA Sequencing. *Dev. Cell* **48**:840–852.e5. <https://doi.org/10.1016/j.devcel.2019.02.022>.

Dorrity, M.W., Alexandre, C.M., Hamm, M.O., Vigil, A.L., Fields, S., Queitsch, C., and Cuperus, J.T. (2021). The regulatory landscape of Arabidopsis thaliana roots at single-cell resolution. *Nat. Commun.* **12**:3334. <https://doi.org/10.1038/s41467-021-23675-y>.

Duque, P., and Chua, N.H. (2003). IMB1, a bromodomain protein induced during seed imbibition, regulates ABA- and phyA-mediated responses of germination in Arabidopsis. *Plant J.* **35**:787–799. <https://doi.org/10.1046/j.1365-313x.2003.01848.x>.

Ehlers, K., Bhide, A.S., Tekleyohans, D.G., Wittkop, B., Snowdon, R.J., and Becker, A. (2016). The MADS Box Genes ABS, SHP1, and SHP2 Are Essential for the Coordination of Cell Divisions in Ovule and Seed Coat Development and for Endosperm Formation in Arabidopsis thaliana. *PLoS One* **11**:e0165075. <https://doi.org/10.1371/journal.pone.0165075>.

Farmer, A., Thibivilliers, S., Ryu, K.H., Schiefelbein, J., and Libault, M. (2021). Single-nucleus RNA and ATAC sequencing reveals the impact of chromatin accessibility on gene expression in Arabidopsis roots at the single-cell level. *Mol. Plant* **14**:372–383. <https://doi.org/10.1016/j.molp.2021.01.001>.

Ferrández, C., Pelaz, S., and Yanofsky, M.F. (1999). Control of carpel and fruit development in Arabidopsis. *Annu. Rev. Biochem.* **68**:321–354. <https://doi.org/10.1146/annurev.biochem.68.321>.

Grandi, F.C., Modi, H., Kampman, L., and Corces, M.R. (2022). Chromatin accessibility profiling by ATAC-seq. *Nat. Protoc.* **17**:1518–1552. <https://doi.org/10.1038/s41596-022-00692-9>.

Han, X., Huang, X., and Deng, X.W. (2020). The Photomorphogenic Central Repressor COP1: Conservation and Functional Diversification during Evolution. *Plant Commun.* **1**:100044. <https://doi.org/10.1016/j.xplc.2020.100044>.

Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., 3rd, Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. *Cell* **184**:3573–3587.e29. <https://doi.org/10.1016/j.cell.2021.04.048>.

Hashiguchi, Y., Yano, D., Nagafusa, K., Kato, T., Saito, C., Uemura, T., Ueda, T., Nakano, A., Tasaka, M., and Terao Morita, M. (2014). A unique HEAT repeat-containing protein SHOOT GRAVITROPISM6 is involved in vacuolar membrane dynamics in gravity-sensing cells of Arabidopsis inflorescence stem. *Plant Cell Physiol.* **55**:811–822. <https://doi.org/10.1093/pcp/pcu020>.

Holm, M., Ma, L.G., Qu, L.J., and Deng, X.W. (2002). Two interacting bZIP proteins are direct targets of COP1-mediated control of light-dependent gene expression in Arabidopsis. *Genes Dev.* **16**:1247–1259. <https://doi.org/10.1101/gad.969702>.

Huang, H., Yang, X., Zheng, M., Chen, Z., Yang, Z., Wu, P., Jenks, M.A., Wang, G., Feng, T., Liu, L., et al. (2023). An ancestral role for 3-KETOACYL-COA SYNTHASE3 as a negative regulator of plant cuticular wax synthesis. *Plant Cell* **35**:2251–2270. <https://doi.org/10.1093/plcell/koad051>.

Ishikawa, M., Kiba, T., and Chua, N.H. (2006). The Arabidopsis SPA1 gene is required for circadian clock function and photoperiodic flowering. *Plant J.* **46**:736–746. <https://doi.org/10.1111/j.1365-313X.2006.02737.x>.

## Plant Communications

Jayawardhane, K.N., Singer, S.D., Ozga, J.A., Rizvi, S.M., Weselake, R.J., and Chen, G. (2020). Seed-specific down-regulation of Arabidopsis CELLULOSE SYNTHASE 1 or 9 reduces seed cellulose content and differentially affects carbon partitioning. *Plant Cell Rep.* **39**:953–969. <https://doi.org/10.1007/s00299-020-02541-z>.

Jean-Baptiste, K., McFaline-Figueroa, J.L., Alexandre, C.M., Dorrity, M.W., Saunders, L., Bubb, K.L., Trapnell, C., Fields, S., Queitsch, C., and Cuperus, J.T. (2019). Dynamics of Gene Expression in Single Root Cells of Arabidopsis thaliana. *Plant Cell* **31**:993–1011. <https://doi.org/10.1105/tpc.18.00785>.

Jin, J., Lu, P., Xu, Y., Tao, J., Li, Z., Wang, S., Yu, S., Wang, C., Xie, X., Gao, J., et al. (2022). PCMDB: a curated and comprehensive resource of plant cell markers. *Nucleic Acids Res.* **50**:D1448–D1455. <https://doi.org/10.1093/nar/gkab949>.

Kagawa, T., and Wada, M. (2002). Blue light-induced chloroplast relocation. *Plant Cell Physiol.* **43**:367–371. <https://doi.org/10.1093/pcp/pcf049>.

Kajala, K., Gouran, M., Shaar-Moshe, L., Mason, G.A., Rodriguez-Medina, J., Kawa, D., Pauluzzi, G., Reynoso, M., Canto-Pastor, A., Manzano, C., et al. (2021). Innovation, conservation, and repurposing of gene function in root cell type development. *Cell* **184**:3333–3348.e19. <https://doi.org/10.1016/j.cell.2021.04.024>.

Kaneda, M., Schuetz, M., Lin, B.S.P., Chanis, C., Hamberger, B., Western, T.L., Ehling, J., and Samuels, A.L. (2011). ABC transporters coordinately expressed during lignification of Arabidopsis stems include a set of ABCBs associated with auxin transport. *J. Exp. Bot.* **62**:2063–2077. <https://doi.org/10.1093/jxb/erq416>.

Katoh, K., Misawa, K., Kuma, K.i., and Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**:3059–3066. <https://doi.org/10.1093/nar/gkf436>.

Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P.R., and Raychaudhuri, S. (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**:1289–1296. <https://doi.org/10.1038/s41592-019-0619-0>.

Kumar, R., Pandey, M.K., Roychoudhry, S., Nayyar, H., Kepinski, S., and Varshney, R.K. (2019). Peg biology: deciphering the molecular regulations involved during peanut peg development. *Front. Plant Sci.* **10**:1289. <https://doi.org/10.3389/fpls.2019.01289>.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**:357–359. <https://doi.org/10.1038/nmeth.1923>.

Lee, S.B., Jung, S.J., Go, Y.S., Kim, H.U., Kim, J.K., Cho, H.J., Park, O.K., and Suh, M.C. (2009). Two Arabidopsis 3-ketoacyl CoA synthase genes, KCS20 and KCS2/DAISY, are functionally redundant in cuticular wax and root suberin biosynthesis, but differentially controlled by osmotic stress. *Plant J.* **60**:462–475. <https://doi.org/10.1111/j.1365-313X.2009.03973.x>.

Li, Y., Dai, X., Cheng, Y., and Zhao, Y. (2011). NPY genes play an essential role in root gravitropic responses in Arabidopsis. *Mol. Plant* **4**:171–179. <https://doi.org/10.1093/mp/ssq052>.

Liscum, E., and Reed, J.W. (2002). Genetics of Aux/IAA and ARF action in plant growth and development. *Plant Mol. Biol.* **49**:387–400. <https://doi.org/10.1007/978-94-010-0377-3>.

Liu, H., Guo, Z., Gangurde, S.S., Garg, V., Deng, Q., Du, P., Lu, Q., Chitikineni, A., Xiao, Y., Wang, W., et al. (2024). A Single-Nucleus Resolution Atlas of Transcriptome and Chromatin Accessibility for Peanut (*Arachis hypogaea* L.) Leaves. *Adv. Biol.* **8**:e2300410. <https://doi.org/10.1002/adbi.202300410>.

Liu, H., Liang, X., Lu, Q., Li, H., Liu, H., Li, S., Varshney, R., Hong, Y., and Chen, X. (2020). Global transcriptome analysis of subterranean

- pod and seed in peanut (*Arachis hypogaea* L.) unravels the complexity of fruit development under dark condition. *Sci. Rep.* **10**:13050. <https://doi.org/10.1038/s41598-020-69943-7>.
- Livak, K.J., and Schmitgen, T.D.** (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2-(Delta Delta CT) Method. *Methods* **25**:402–408. <https://doi.org/10.1006/meth.2001.1262>.
- Long, Y., Liu, Z., Jia, J., Mo, W., Fang, L., Lu, D., Liu, B., Zhang, H., Chen, W., and Zhai, J.** (2021). FlsnRNA-seq: protoplasting-free full-length single-nucleus RNA profiling in plants. *Genome Biol.* **22**:66. <https://doi.org/10.1186/s13059-021-02288-0>.
- Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R., Kamitaki, N., Martersteck, E.M., et al.** (2015). Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* **161**:1202–1214. <https://doi.org/10.1016/j.cell.2015.05.002>.
- Malik, N., Ranjan, R., Parida, S.K., Agarwal, P., and Tyagi, A.K.** (2020). Mediator subunit OsMED14\_1 plays an important role in rice development. *Plant J.* **101**:1411–1429. <https://doi.org/10.1111/tpj.14605>.
- McGinnis, C.S., Murrow, L.M., and Gartner, Z.J.** (2019). DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Syst.* **8**:329–337.e4. <https://doi.org/10.1016/j.cels.2019.03.003>.
- Moctezuma, E.** (1999). Changes in auxin patterns in developing gynophores of the peanut plant (*Arachis hypogaea* L.). *Ann. Bot.* **83**:235–242. <https://doi.org/10.1006/anbo.1998.0814>.
- Moctezuma, E., and Feldman, L.J.** (1998). Growth rates and auxin effects in graviresponding gynophores of the peanut, *Arachis hypogaea* (Fabaceae). *Am. J. Bot.* **85**:1369–1376. <https://doi.org/10.2307/2446395>.
- Molina, I., Li-Beisson, Y., Beisson, F., Ohlrogge, J.B., and Pollard, M.** (2009). Identification of an Arabidopsis feruloyl-coenzyme A transferase required for suberin synthesis. *Plant Physiol.* **151**:1317–1328. <https://doi.org/10.1104/pp.109.144907>.
- Nayak, N.R., Putnam, A.A., Addepalli, B., Lowenson, J.D., Chen, T., Jankowsky, E., Perry, S.E., Dinkins, R.D., Limbach, P.A., Clarke, S.G., et al.** (2013). An Arabidopsis ATP-dependent, DEAD-box RNA helicase loses activity upon IsoAsp formation but is restored by PROTEIN ISOASPARTYL METHYLTRANSFERASE. *Plant Cell* **25**:2573–2586. <https://doi.org/10.1105/tpc.113.113456>.
- Neftel, C., Laffy, J., Filbin, M.G., Hara, T., Shore, M.E., Rahme, G.J., Richman, A.R., Silverbush, D., Shaw, M.L., Hebert, C.M., et al.** (2019). An Integrative Model of Cellular States, Plasticity, and Genetics for Glioblastoma. *Cell* **178**:835–849.e21. <https://doi.org/10.1016/j.cell.2019.06.024>.
- Ng, M., and Yanofsky, M.F.** (2001). Function and evolution of the plant MADS-box gene family. *Nat. Rev. Genet.* **2**:186–195. <https://doi.org/10.1038/35056041>.
- Ortiz-Ramírez, C., Guillotin, B., Xu, X., Rahni, R., Zhang, S., Yan, Z., Coqueiro Dias Araujo, P., Demesa-Arevalo, E., Lee, L., Van Eck, J., et al.** (2021). Ground tissue circuitry regulates organ complexity in maize and *Setaria*. *Science* **374**:1247–1252. <https://doi.org/10.1126/science.ab2327>.
- Pattee, H.E., Johns, E.B., Singleton, J.A., and Sanders, T.H.** (1974). Composition changes of peanut fruit parts during maturation. *Peanut Sci.* **1**:57–62.
- Price, M.N., Dehal, P.S., and Arkin, A.P.** (2010). FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS One* **5**:e9490. <https://doi.org/10.1371/journal.pone.0009490>.
- Ranocha, P., Denancé, N., Vanholme, R., Freydier, A., Martinez, Y., Hoffmann, L., Köhler, L., Pouzet, C., Renou, J.P., Sundberg, B., et al.** (2010). Walls are thin 1 (WAT1), an *Arabidopsis* homolog of *Medicago truncatula* NODULIN21, is a tonoplast-localized protein required for secondary wall formation in fibers. *Plant J.* **63**:469–483. <https://doi.org/10.1111/j.1365-313X.2010.04256.x>.
- Rowland, O., Zheng, H., Hepworth, S.R., Lam, P., Jetter, R., and Kunst, L.** (2006). CER4 encodes an alcohol-forming fatty acyl-coenzyme A reductase involved in cuticular wax production in *Arabidopsis*. *Plant Physiol.* **142**:866–877. <https://doi.org/10.1104/pp.106.086785>.
- Ryu, K.H., Huang, L., Kang, H.M., and Schiefelbein, J.** (2019). Single-Cell RNA Sequencing Resolves Molecular Relationships Among Individual Plant Cells. *Plant Physiol.* **179**:1444–1456. <https://doi.org/10.1104/pp.18.01482>.
- Savidge, B., Rounsley, S.D., and Yanofsky, M.F.** (1995). Temporal relationship between the transcription of two *Arabidopsis* MADS box genes and the floral organ identity genes. *Plant Cell* **7**:721–733. <https://doi.org/10.1105/tpc.7.6.721>.
- Shahan, R., Nolan, T.M., and Benfey, P.N.** (2021). Single-cell analysis of cell identity in the *Arabidopsis* root apical meristem: insights and opportunities. *J. Exp. Bot.* **72**:6679–6686. <https://doi.org/10.1016/j.devcel.2022.01.008>.
- Shlamovitz, N., Ziv, M., and Zamski, E.** (1995). Light, dark and growth regulator involvement in groundnut (*Arachis hypogaea* L.) pod development. *Plant Growth Regul.* **16**:37–42. <https://doi.org/10.1007/BF00040505>.
- Shushu, D.D., and Cutter, E.G.** (1990). Growth of the gynophore of the peanut arachis hypogaea. 2. regulation of growth. *Can. J. Bot.* **68**:965–978. <https://doi.org/10.1139/b90-123>.
- Song, W.Y., Yamaki, T., Yamaji, N., Ko, D., Jung, K.H., Fujii-Kashino, M., An, G., Martinoia, E., Lee, Y., and Ma, J.F.** (2014). A rice ABC transporter, OsABCC1, reduces arsenic accumulation in the grain. *Proc. Natl. Acad. Sci. USA* **111**:15699–15704. <https://doi.org/10.1073/pnas.1414968111>.
- Soutourina, J.** (2018). Transcription regulation by the Mediator complex. *Nat. Rev. Mol. Cell Biol.* **19**:262–274. <https://doi.org/10.1038/nrm.2017.115>.
- Stahle, M.I., Kuehlich, J., Staron, L., von Arnim, A.G., and Golz, J.F.** (2009). YABBYs and the transcriptional corepressors LEUNIG and LEUNIG\_HOMOLOG maintain leaf polarity and meristem activity in *Arabidopsis*. *Plant Cell* **21**:3105–3118. <https://doi.org/10.1105/tpc.109.070458>.
- Stuart, T., Srivastava, A., Madad, S., Lareau, C.A., and Satija, R.** (2021). Single-cell chromatin state analysis with Signac. *Nat. Methods* **18**:1333–1341. <https://doi.org/10.1038/s41592-021-01282-5>.
- Sun, Y., Wang, Q., Li, Z., Hou, L., Dai, S., and Liu, W.** (2013). Comparative proteomics of peanut gynophore development under dark and mechanical stimulation. *J. Proteome Res.* **12**:5502–5511. <https://doi.org/10.1021/pr003504>.
- Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N.J., Livak, K.J., Mikkelsen, T.S., and Rinn, J.L.** (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**:381–386. <https://doi.org/10.1038/nbt.2859>.
- Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., and Rozen, S.G.** (2012). Primer3-new capabilities and interfaces. *Nucleic Acids Res.* **40**:e115. <https://doi.org/10.1093/nar/gks596>.
- Van der Pijl, L.** (1982). *Principles of Dispersal in Higher Plants* 214 (Springer-Verlag).
- van Dijk, E.L., Jaszczyzyn, Y., and Thermes, C.** (2014). Library preparation methods for next-generation sequencing: tone down the

## Single nucleus RNA-seq and ATAC-seq of peanut pods

- bias. *Exp. Cell Res.* **322**:12–20. <https://doi.org/10.1016/j.yexcr.2014.01.008>.
- Wendrich, J.R., Yang, B., Vandamme, N., Verstaen, K., Smet, W., Van de Velde, C., Minne, M., Wybouw, B., Mor, E., Arents, H.E., et al.** (2020). Vascular transcription factors guide plant epidermal responses to limiting phosphate conditions. *Science* (New York, N.Y.) **370**:eaay4970. <https://doi.org/10.1126/science.aay4970>.
- Xiao, W., Custard, K.D., Brown, R.C., Lemmon, B.E., Harada, J.J., Goldberg, R.B., and Fischer, R.L.** (2006). DNA methylation is critical for Arabidopsis embryogenesis and seed viability. *Plant Cell* **18**:805–814. <https://doi.org/10.1105/tpc.105.038836>.
- Xie, Q., Frugis, G., Colgan, D., and Chua, N.H.** (2000). Arabidopsis NAC1 transduces auxin signal downstream of TIR1 to promote lateral root development. *Genes Dev.* **14**:3024–3036. <https://doi.org/10.1101/gad.852200>.
- Xu, W., Wen, Y., Liang, Y., Xu, Q., Wang, X., Jin, W., and Chen, X.** (2021). A plate-based single-cell ATAC-seq workflow for fast and robust profiling of chromatin accessibility. *Nat. Protoc.* **16**:4084–4107. <https://doi.org/10.1038/s41596-021-00583-5>.
- Yamauchi, Y., Fukaki, H., Fujisawa, H., and Tasaka, M.** (1997). Mutations in the SGR4, SGR5 and SGR6 loci of *Arabidopsis thaliana* alter the shoot gravitropism. *Plant Cell Physiol.* **38**:530–535. <https://doi.org/10.1093/oxfordjournals.pcp.a029201>.
- Zamski, E., and Ziv, M.** (1976). Pod formation and its geotropic orientation in the peanut, *Arachis hypogaea* L., in relation to light and mechanical stimulus. *Ann. Bot.* **40**:631–636. <https://doi.org/10.1093/oxfordjournals.aob.a085173>.
- Zhang, L., He, C., Lai, Y., Wang, Y., Kang, L., Liu, A., Lan, C., Su, H., Gao, Y., Li, Z., et al.** (2023). Asymmetric gene expression and cell-type-specific regulatory networks in the root of bread wheat revealed by single-cell multiomics analysis. *Genome Biol.* **24**:65. <https://doi.org/10.1186/s13059-023-02908-x>.
- Zhang, T.Q., Chen, Y., Liu, Y., Lin, W.H., and Wang, J.W.** (2021). Single-cell transcriptome atlas and chromatin accessibility landscape reveal differentiation trajectories in the rice root. *Nat. Commun.* **12**:2053. <https://doi.org/10.1038/s41467-021-22352-4>.
- Zhang, T.Q., Chen, Y., and Wang, J.W.** (2021). A single-cell analysis of the *Arabidopsis* vegetative shoot apex. *Dev. Cell* **56**:1056–1074.e8. <https://doi.org/10.1016/j.devcel.2021.02.021>.
- Zhang, T.Q., Xu, Z.G., Shang, G.D., and Wang, J.W.** (2019). A Single-Cell RNA Sequencing Profiles the Developmental Landscape of *Arabidopsis* Root. *Mol. Plant* **12**:648–660. <https://doi.org/10.1016/j.molp.2019.04.004>.
- Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al.** (2008). Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**:R137. <https://doi.org/10.1186/gb-2008-9-9-r137>.
- Zhang, Y., Wang, P., Xia, H., Zhao, C., Hou, L., Li, C., Gao, C., Wang, X., and Zhao, S.** (2016). Comparative transcriptome analysis of basal and zygote-located tip regions of peanut ovaries provides insight into the mechanism of light regulation in peanut embryo and pod development. *BMC Genom.* **17**:606. <https://doi.org/10.1186/s12864-016-2857-1>.
- Zhu, D., Maier, A., Lee, J.H., Laubinger, S., Saijo, Y., Wang, H., Qu, L.J., Hoecker, U., and Deng, X.W.** (2008). Biochemical characterization of *Arabidopsis* complexes containing CONSTITUTIVELY PHOTOMORPHOGENIC1 and SUPPRESSOR OF PHYA proteins in light control of plant development. *Plant Cell* **20**:2307–2323. <https://doi.org/10.1105/tpc.107.056580>.
- Ziv, M.** (1981). Photomorphogenesis of the gynophore, pod, and embryo in peanut (*Arachis hypogaea* L.). *Ann. Bot.* **48**:353–359. <https://doi.org/10.1093/oxfordjournals.aob.a086133>.
- Ziv, M., and Zamski, E.** (1975). Geotropic responses and pod development in gynophore explants of peanut (*Arachis hypogaea* L.) cultured in vitro. *Ann. Bot.* **39**:579–583. <https://doi.org/10.1093/oxfordjournals.aob.a084968>.

## Plant Communications