

# ScRNA-seq reveals dark- and light-induced differentially expressed gene atlases of seedling leaves in *Arachis hypogaea* L.

Quanqing Deng<sup>1,†</sup> , Puxuan Du<sup>1,†</sup>, Sunil S. Gangurde<sup>2,†</sup>, Yanbin Hong<sup>1,†</sup>, Yuan Xiao<sup>3</sup>, Dongxiu Hu<sup>1</sup>, Haifen Li<sup>1</sup>, Qing Lu<sup>1</sup>, Shaoxiong Li<sup>1</sup>, Haiyan Liu<sup>1</sup>, Runfeng Wang<sup>1</sup>, Lu Huang<sup>1</sup>, Wenyi Wang<sup>4</sup>, Vanika Garg<sup>5</sup>, Xuanqiang Liang<sup>1,\*</sup>, Rajeev K. Varshney<sup>4,\*</sup> , Xiaoping Chen<sup>1,\*</sup> and Hao Liu<sup>1,\*</sup> 

<sup>1</sup>Guangdong Provincial Key Laboratory of Crop Genetic Improvement, South China Peanut Sub-Center of National Center of Oilseed Crops Improvement, Crops Research Institute, Guangdong Academy of Agricultural Sciences, Guangzhou, Guangdong Province, China

<sup>2</sup>International Crops Research Institute for the Semi-Arid Tropic, Hyderabad, India

<sup>3</sup>School of Public Health, Wannan Medical College, Wuhu, Anhui Province, China

<sup>4</sup>College of Agriculture, South China Agricultural University, Guangzhou, Guangdong Province, China

<sup>5</sup>WA State Agricultural Biotechnology Centre, Centre for Crop and Food Innovation, Food Futures Institute, Murdoch University, Murdoch, Western Australia, Australia

Received 30 August 2022;

revised 22 January 2024;

accepted 23 January 2024.

\*Correspondence (Tel +86 020 87511794; fax +86 020 85514269; email liuhao@gdaas.cn (H.L.); Tel +86 020 85514243; fax +86 020 85514269; email chenxiaoping@gdaas.cn (X.C.); Tel +61 893602608; fax +61 9360 6303; email rajeev.varshney@murdoch.edu.au (R.V.); Tel +86 020 87597312; fax +86 020 85514269; email liangxuanqiang@gdaas.cn (X.L.)

†These authors contributed equally to this article.

**Keywords:** scRNA-seq, peanut leaf, single cell, light, gene atlases.

## Summary

Although the regulatory mechanisms of dark and light-induced plant morphogenesis have been broadly investigated, the biological process in peanuts has not been systematically explored on single-cell resolution. Herein, 10 cell clusters were characterized using scRNA-seq-identified marker genes, based on 13 409 and 11 296 single cells from 1-week-old peanut seedling leaves grown under dark and light conditions. 6104 genes and 50 transcription factors (TFs) displayed significant expression patterns in distinct cell clusters, which provided gene resources for profiling dark/light-induced candidate genes. Further pseudo-time trajectory and cell cycle evidence supported that dark repressed the cell division and perturbed normal cell cycle, especially the *PORA* abundances correlated with 11 TFs highly enriched in mesophyll to restrict the chlorophyllide synthesis. Additionally, light repressed the epidermis cell developmental trajectory extending by inhibiting the growth hormone pathway, and 21 TFs probably contributed to the different genes transcriptional dynamic. Eventually, peanut *AHL17* was identified from the profile of differentially expressed TFs, which encoded protein located in the nucleus promoted leaf epidermal cell enlargement when ectopically overexpressed in *Arabidopsis* through the regulatory phytohormone pathway. Overall, our study presents the different gene atlases in peanut etiolated and green seedlings, providing novel biological insights to elucidate light-induced leaf cell development at the single-cell level.

## Introduction

Light and darkness are environmental factors that significantly regulate plant growth and development. Seeds of the majority of terrestrial plants develop as etiolation seedlings under dark conditions, breaking through the soil to capture aerial light, which in turn promotes chloroplast maturation while slowing hypocotyl elongation (Shi *et al.*, 2018; Wit *et al.*, 2016; Wu *et al.*, 2020). Hence, light works as a switch to mediate photomorphogenesis and skotomorphogenesis, which is capable of activating the photosynthesis reaction, but directly represses seedling development (Deepika *et al.*, 2020; Lau and Deng, 2012). Seedling de-etiolation to green cotyledons is an important stage to study seedling reversal from dark growth condition to a light environment (Liu *et al.*, 2017; Wu *et al.*, 2020). Overall, plant leaves play a key role in heterotrophic to autotrophic growth under light condition, which involves various regulatory transcription factors and phytohormones, and light circumstances mediate distinct transcriptional regulation of seedling development, cellular differentiation and tissue or organ development (Favero *et al.*, 2021; Wit *et al.*, 2016). Despite

using many biological techniques in several molecular biology studies to understand plant and light interaction, the distinct light environment modulated cell atlases at the single-cell level have not been developed. Elucidating the cellular heterogeneity of different light condition may unravel the molecular mechanism of plant growth and development.

In general, bulk transcriptome sequencing (Bulk RNA-seq) has become a routine method to understand the average expression levels of all cells in a given sample (Efroni *et al.*, 2015; Gala *et al.*, 2021). However, cellular heterogeneity, with spatiotemporal expression differences between cell types and different developmental stages has been overlooked. Recently, single-cell transcriptome (scRNA-seq) has emerged and deployed to describe transcriptional changes in different organisms at the single-cell level resolution, which facilitated the study of intercellular gene interaction mechanisms, cell fate selection and developmental trajectories, and advancing the discovery of new cell types (Liu *et al.*, 2021; Ryu *et al.*, 2019). Separation of protoplasts from plant cells is difficult due to cell wall obstruction, which limits the large-scale application of scRNA-seq in plants. Despite the technical barriers, none of them can hardly deter researchers in the field of

plant genomics from applying scRNA-seq. Besides, scRNA-seq was initially used for revealing dynamic transcriptional expression at root developmental processes in *Arabidopsis*, which presented highly heterogeneous results compared to traditional bulk RNA-seq (Ryu *et al.*, 2019). scRNA-seq was used to analyse process of lateral root formation and develop a model of lateral root explaining sequential split of main root tissues and stem cells in *A. thaliana* (Serrano-Ron *et al.*, 2021). In *Arabidopsis*, scRNA-seq of leaves revealed that the phloem cells were significantly different from the companion cells, which contributed to the differences in sugar transport and metabolism in leaf veins (Kim *et al.*, 2021). Additionally, scRNA-seq was used to map the transcriptome profiles of leaves and roots of rice to promote differential gene expression studies and various rice developmental changes under abiotic stresses (Wang *et al.*, 2021b). Though the majority of scRNA-seq studies were conducted on model plants, there are few reports on the use of scRNA-seq in legume crops.

Peanut (*Arachis hypogaea* L.) is an important leguminous oilseed and food crop cultivated across the world for edible oil and high-quality protein (Hong *et al.*, 2015, 2021). The complex genome of allotetraploid cultivated peanuts has hindered the depth of peanut research (Liu *et al.*, 2020a). Cultivated peanuts evolved from two wild progenitors and both share the responsibility for the genetic regulation of multiple traits (Gangurde *et al.*, 2020). The availability of genome sequences of cultivated peanuts accelerated genomics research and breeding applications (Chen *et al.*, 2019; Zhuang *et al.*, 2019). Light and darkness serve as important environmental conditions for the development of peanut pods, which form after above-ground fertilization with pegs extending into the soil, the latter providing dark circumstances and mechanical friction (Cui *et al.*, 2022). According to multi-omics integration, the geocarpy was connected to gene expression involved in the gravity response, phytohormone production and phytochrome interaction factors (Chen *et al.*, 2016; Liu *et al.*, 2020b; Zhang *et al.*, 2018). Therefore, light development biological study enabled to provide novel insight for demonstrating the peanut tissues development features. Notably, peanut tissue ontology regarding transcriptional dynamic illustrating has entered a single-cell era, such as by using scRNA-seq approaches in conjunction with high-throughput multi-omics findings discovered the role of peanut AHL23 (AT-HOOK MOTIF NUCLEAR LOCALIZED PROTEIN 23) in leaf expansion on single-cell resolution (Liu *et al.*, 2021).

Leaves are primary sites of photosynthesis made up of three different tissue types such as epidermis, mesophyll and vasculature. However, the absolute role of photosynthesis in plant autotrophy has not been investigated on single-cell resolution. Therefore, to decipher the transcriptome landscape at the single-cell resolution we used scRNA-seq in peanut seedlings grown in light/dark conditions. Temporal gene expression profiles established at individual cell level using light/dark morphogenetic-grown peanut seedlings. Additionally, leaf photosynthesis-specific cell types were identified to provide new insights into peanut cell biology and development.

## Results

### The phenotype of peanut seedlings grown under light and dark conditions

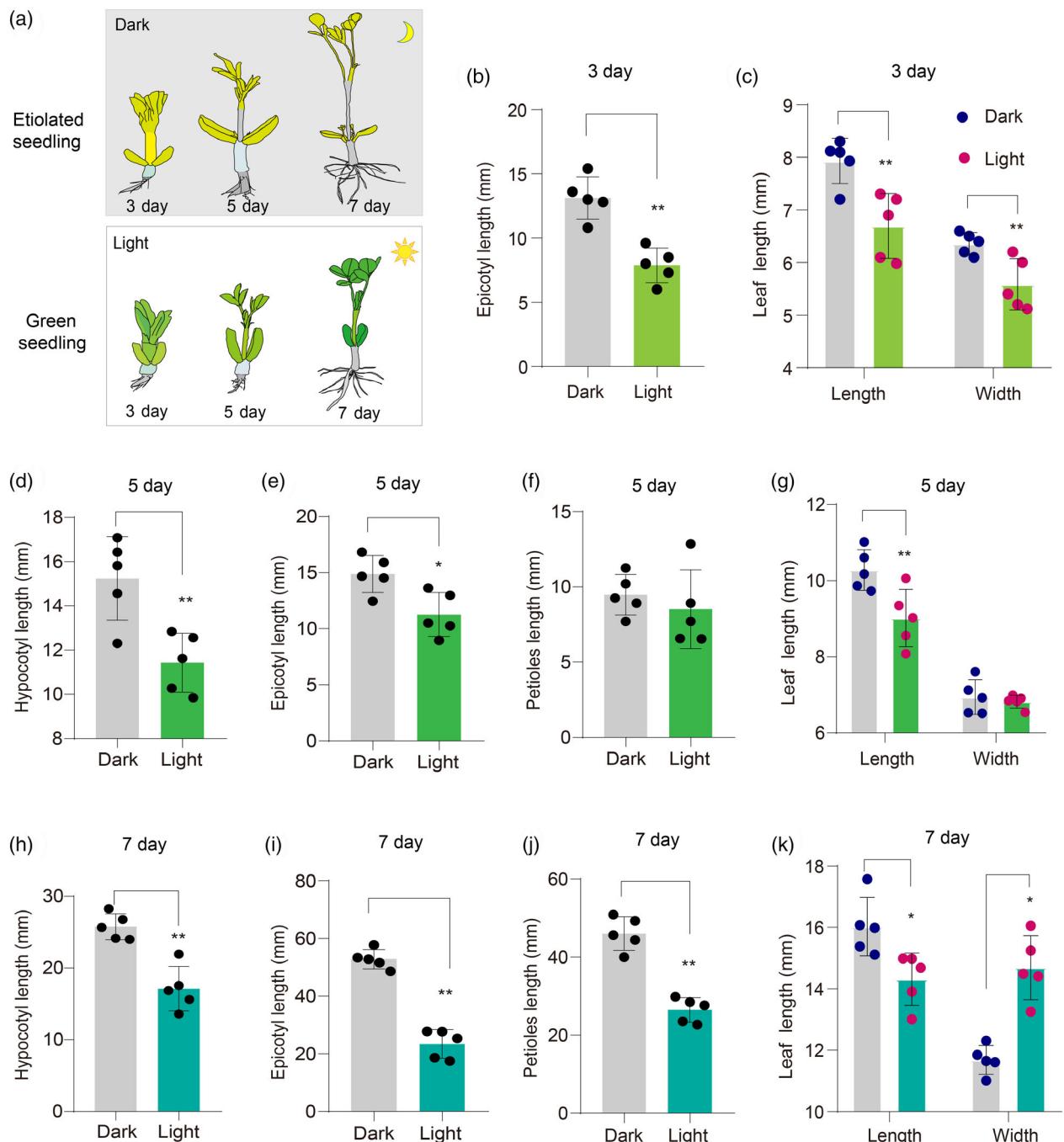
Peanut seedling phenotypes were recorded after seed germination within 1 week after planting (Figure 1a, Figure S1). On

the 3rd day, shoot apical meristem (SAM) gradually formed the shoot leaf bud, but the leaf blade was tightly closed and not expanded. Measurement of young epicotyl length and leaf area suggested that the seedling grown under the dark showed faster growth than the light-modulated seedling (Figure 1b,c, Figure S1). In addition, during the first true leaf development, the hypocotyl, epicotyl and leaf blade lengths were longer in seedlings grown in the dark than in the seedlings grown under the light. However, the leaf blade width and length of petioles did not show any significant differences on the 5th day between seedlings grown under light and dark (Figure 1d–g, Figure S1). The dark seedling significantly promoted the growth of plant epicotyl, hypocotyl and petioles elongation. The first true leaf of a light seedling was expanded to capture the sunlight with a larger leaf blade as compared to dark seedlings. However, the etiolated seedlings exhibited rolled leaf along with the leaf vein, indicating that dark condition represses the leaf expansion (Figure 1h–k, Figure S1). Therefore, the morphology observations of peanut seedlings grown under dark and light conditions demonstrated that light represses plant seedling growth. These findings provided basic information for further investigation of the mechanism of peanut plants grown under distinct light conditions.

### Peanut leaf scRNA-seq and major cell clusters assessment

Despite current knowledge of light-dark-induced seedling germination being massively dissected with phenotype observation also multi-omics revealed functional genes regulating photomorphogenesis, the single-cell transcriptome dynamics during leaf cell differentiation under dark and light environment has not yet been in peanuts. In the present study, to decipher the transcriptome profiles under specific cell cluster levels, seedling leaves served as an ideal material to characterize the single-cell gene expression landscapes. Here, we developed a highly efficient protoplast isolation protocol by optimizing the formulation proportion of cellulase and pectinase in cell wall digest solution and extracted the intact protoplast from the leaf blade of a one-week-old seedling (Figure 2a, Figure S2). A total of 800 cells/ $\mu$ L (dark) and 682 cells/ $\mu$ L (light) were isolated and nearly  $4 \times 10^4$  cells were loaded into the microfluidic platform with barcode ligation and library construction.

Two groups under dark and light conditions captured a total of 13 429 and 12 209 cells respectively, after filtering duplicates, low-quality cells and genes, respectively (Figure 2d, Table S1). The sequencing data for light and dark cells generated 403 282 586 and 420 629 752 raw reads. Furthermore, the raw reads were mapped to the *A. hypogaea* reference genome (Table S1). Data quality control showed that 1960 and 2507 filtered UMI (unique molecular index) and 1348 and 1461 genes were allocated to the per cells under dark and light conditions, respectively (Tables S1, S2, Figure S3). Subsequently, the FPKM values were applied to identify a total of 49 882 genes whose expression abundance was subsequently anchored in a circular plot consisting of 20 peanut chromosomes to illustrate the single-cell gene expression profiles of leaf cells (Figure 2h, Table S3). The previously reported marker genes for different leaf cells were used to identify two groups of high-quality cells, which helped to identify the types of cell clusters, and the data results were visualized using principal component analysis and t-distributed stochastic neighbour embedding (t-SNE) and uniform manifold approximation and projection (UMAP) methods (Figure 2b, Table S3).



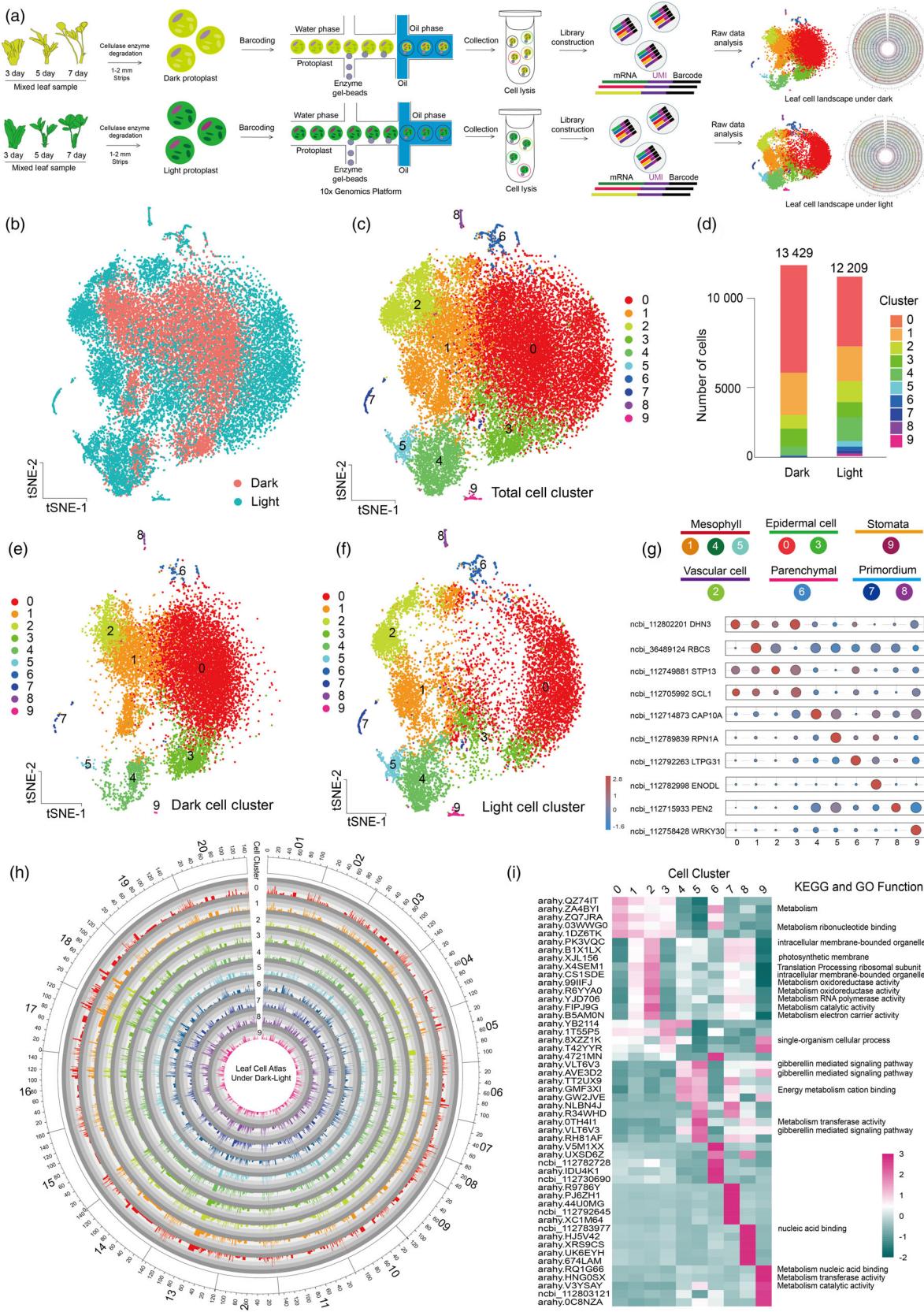
**Figure 1** Phenotypic variations between peanut seedlings grown under light and dark conditions. (a) Growth overview of etiolated (grown in the dark) and green (grown under light) seedlings on the 3rd, 5th, and 7th day. (b, c) Epicotyl length, leaf length and width under dark and light-grown seedlings on the 3rd day (\*\*P < 0.01). (d–g) Hypocotyl, epicotyl and petiole lengths, leaf length and width under dark and light-grown seedlings on the 5th day (\*P < 0.05, \*\*P < 0.01). (h–k) Hypocotyl, epicotyl and petiole length, leaf length and width of dark- and light-grown seedlings on the 7th day (\*P < 0.05, \*\*P < 0.01).

Figure S4). A total of 13 429 and 11 296 high-quality cells from two samples were divided into 10 cell clusters (Figure 2c,e,f). Among them, the number of single cells in each cell cluster was ranging from 2 to 7508 for dark samples, and from 86 to 4392 for light samples, and the majority of them belonged to epidermal cells (Figure 2d). To identify the cell types of identified cell clusters, we first collected a batch of marker genes reported for

specific leaf cells in the peanut genome and then visualized using t-SNE (Figure 2h, Figure S5, Table S4). The largest cell clusters were classified as epidermal cells and were detected by predominantly expressed genes in these cell populations like the *DEHYDRIN DHN3-LIKE (DHN3)* and *SCARECROW-LIKE PROTEIN 1 (SCL1)* (Li et al., 2020; Ma et al., 2014). The expression distribution of *EARLY NODULIN-LIKE PROTEIN 1 (ENOD1)* and *PROBABLE*

GAMMA-SECRETASE SUBUNIT PEN-2 ISOFORM X1 (PEN2) was used to determine the leaf primordia (Michalina et al., 2014; Nicolas et al., 2014). Additionally, the RIBULOSE 1,5-

### BISPHOSPHATE CARBOXYLASE/OXYGENASE LARGE SUBUNIT (CHLOROPLAST) (RBCS), CHLOROPHYLL A-B BINDING PROTEIN CP24 10A, CHLOROPLASTIC (CAP10A) and 26S PROTEASOME



**Figure 2** Steps involved in scRNA-seq and cell cluster classification of 1-week-old etiolated and green seedlings of peanut. (a) A brief protocol of isolation of peanut leaf protoplasts, 10 × genomics platform loading and genome anchoring. (b) tSNE plot illustrated leaf cell clusters of etiolated and green seedlings. Each dot represented a cell. The cells from light- and dark-grown seedlings were represented in blue and red colour, respectively. (c) 10 colours represented 10 distinct cell clusters identified scRNA-seq. (d) Bar plots represented the number of cells captured in etiolated and green seedlings and their distribution in 10 clusters. (e, f) Leaf cell clusters of etiolated and green seedlings, respectively. (g) Cell clusters 0–9 belonged to the respective cell types, and the bubble plot illustrates the expression pattern and distribution of cluster marker genes. (h) Circos plot represented the genome-wide expression of DEGs in peanut leaf cells by scRNA-seq. 0–9 numbers from the outer to inner tracks represent different cell clusters. Overall the circos plot represented the expression of DEGs from all cell clusters on 20 chromosomes in the peanut genome. (i) The heat map showed the top five DEGs with the highest expression levels for each cell cluster.

*NON-ATPASE REGULATORY SUBUNIT 2 HOMOLOG A (RPN1A)* involved in photosynthesis were highly enriched in mesophyll cells (Wang et al., 2021b). Vascular cells enriched with transcripts based on high expression levels of *SUGAR TRANSPORT PROTEIN 13 (STP13)* (Kim et al., 2021). Stomata guard cells show a unique expression profile of its marker gene *PROBABLE WRKY TRANSCRIPTION FACTOR 30 (WRKY30)* in cell cluster 9 (Liu et al., 2020c). Meanwhile, the *NON-SPECIFIC LIPID-TRANSFER PROTEIN-LIKE PROTEIN (LTPG31)* is highly expressed in parenchymal cells (Ibáñez et al., 2019).

Cell-specific tissue isolation and real-time PCR were performed to validate the true expression levels of marker genes in different cell clusters (Figure S6). And six cell types including mesophyll, epidermis, vascular, parenchymal, primordium and stomata cells were grouped according to cluster-specifically expressed differentially expressed genes (DEGs) (Figure 2g). Importantly, the distribution of gene expression was more concentrated for all cells in the seedlings under dark condition compared to the seedlings under light condition, which contributed to demonstrating the distribution of gene expression in different cell types under dark/light condition (Figure 2e, f). Furthermore, to gain insights into the molecular function of each cell cluster, the top five genes with the highest expression level in each cluster were selected for GO or KEGG functional annotation (Figure 2i, Figure S7, Table S5). From KEGG, the enrichment pathways in major cell clusters mainly involved metabolism and energy metabolism (Table S5). We proved that scRNA-seq could accelerate the resolution of functional genomics and identification of related genes by distinguishing cell types through marker genes and the annotation of enriched genes, and therefore could serve as a valuable resource for functional genomics studies in peanuts.

#### Identifying DEGs and Transcription Factors in distinct cell clusters and light/dark treated seedling leaf

In order to identify core-DEGs in cell clusters and developmental trajectories of peanut leaf cells, pseudo-time trajectories were applied to analyse differentially expressed genes in each cell cluster. Initially, relevant up-regulated DEGs distributed in the range 18 to 1755 were assessed in each cluster and core differentially expressed genes across cell clusters were counted by Venn diagram to identify the DEGs among cell subpopulations (Figure 3a,b, Tables S2, S6). Due to the transcription factors (TFs) providing transcriptional dynamic of gene expression, we focused on the identification of important TFs. Similarly, simultaneous screening for core TFs and 94 TFs was concurrently selected to visualize their expression patterns using a bubble heat map and KEGG functional annotation. Based on their conserved structural domains these TFs were involved in important pathways such as photoperiod, hormone pathway, JA pathway, auxin pathway,

ethylene pathway, signal transduction, environment adaptation and organismal systems (Figure 3c,d, Table S7).

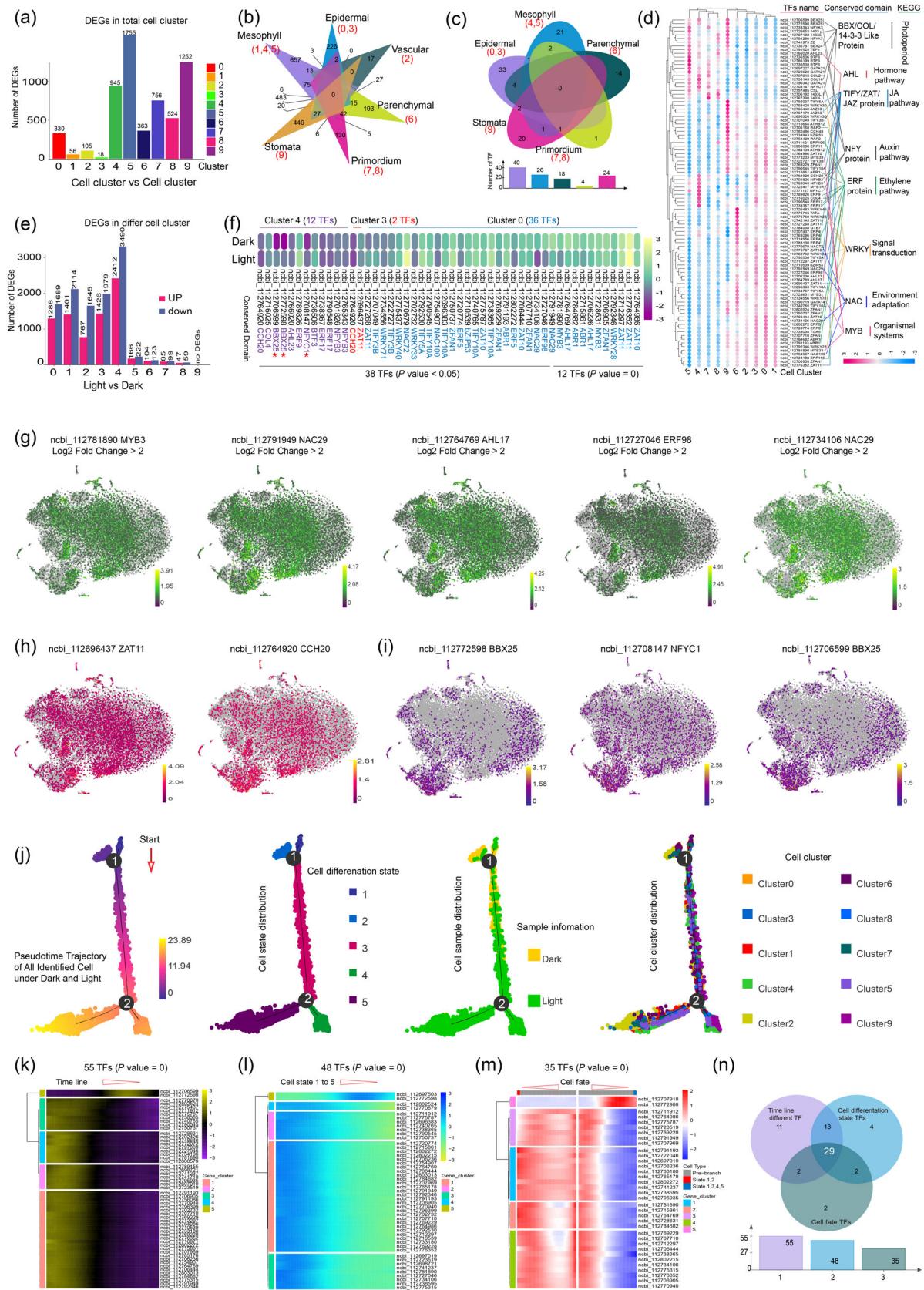
Furthermore, DEGs between samples were analysed to narrow down the number of target DEGs and TFs. By comparing the DEGs in etiolated and green samples, the up- and down-regulated DEGs were assessed in distinct cell clusters (Figure 3e, Table S8). A total of 50 core TFs associated with plant-pathogen interaction and phytohormone signal transduction identified in cell clusters 0, 3 and 4 were by intergroup expression difference analysis in cell sub-clusters of etiolated and green seedlings. In addition, 10 target TFs were short-listed based on fold change, *P*-value and characterized by a t-SNE plot (Figure 3f-i, Table S9). The expression patterns of 10 core TFs per cell revealed that the transcripts *MYB3*, *NAC29*, *AHL17*, *ERF98*, *ZAT11* and *CCH20* were highly expressed in most of the cell types of etiolated seedlings, while *BBX25* and *NFYC1* were highly expressed in green seedlings (Figure 3f-i, Table S9). Identification of core DEGs provides a gene resource for next illustrating critical genes modulated different processes between peanut cells differentiation in the distinct light environment.

#### Light promoted the leaf cell differentiation in the development trajectory

Subsequently, the spatial and temporal distribution of DEGs expression in all cells of etiolated and green seedlings was investigated using pseudo-time analysis and five cell states of leaf cell differentiation development were delineated (Figure 3j, Table S10). 12 696 DEGs were determined the cell differentiation with two branch nodes in the main stem (Figure 3j, Table S11). To understand the gene expression dynamics of the cell differentiation trajectory, we performed a hierarchical clustering analysis of 55 TFs across a pseudo-time line, which was highly expressed in etiolated seedlings (Figure 3k) and cell states 1–3 to involve in plant-pathogen interaction and plant hormone signal pathways (Table S12). Furthermore, 11 598 DEGs were described to correlate with differentiation states, specifically as the differentiation trajectory starts with cells in states 1 to 5 (Figure 3j, Table S13). Additionally, 48 TFs were associated with cell state distribution with these distinct states (Figure 3l, Table S14). On the other hand, 10 250 DEGs were investigated to understand the cell differentiation orientation along with the main stem of the pseudo-time trajectory in etiolated growth to de-etiolated morphogenesis seedling cell clusters (Figure 3j, Table S15). Of which 35 TFs specifically participated in branch node 1 to determine the cell fate (Figure 3m, Table S16). Importantly, 29 core TFs were filtered from three statuses of TFs, related to the developmental trajectory, cell differentiation state and cell fate, and visualized by Venn diagram and expression tendency heatmap (Figure 3n, Figure S8). Based on their KEGG functional annotation, these TFs were involved in hormone

regulation, signal transduction and environment adaptation pathways (Table S17). Overall results uncovered the process of cell differentiation and development between etiolated and green

seedling leaves of peanuts and identified crucial DEGs and TFs in this biological process. To explore whether the single-cell expressions of core TFs reflected the truth tissue transcription



**Figure 3** Details of DEGs in various cell clusters and pseudo-time trajectory analysis of cell types in 1-week-old leaves of etiolated and green peanut seedlings. (a) The number of DEGs in each cell cluster. (b, c) Venn diagram of core and specific DEGs and TFs in all cell types. (d) The bubble heat map illustrated the clustering and KEGG pathway enrichment analysis of core-TFs in each cell cluster. (e) Bar plots of up- and down-regulated DEGs in each cell cluster. (f) The heat map represented the average expression level of core-TFs in each cell cluster under dark and light treatments. (g–i) Feature plots of expression distribution of 10 core-TFs across all cell clusters. (j) Distribution of DEGs into cell differentiation state, sample information and cell cluster along the pseudo-time progression of leaf cell development. (k) Clustering and expression patterns of 55 TFs of cell differentiation along the pseudo-time progression. (l) Expression of 48 TFs involved in cell differentiation states. (m) Heatmap illustrated 35 DEGs involved in the determination of cell fate. The horizontal coordinate was the pseudo-time point (the pseudo-time point increases gradually from the middle to both sides), and the vertical coordinate was the branching differential genes. (n) The Venn diagram indicated core and specific TFs from cell differentiation, cell state differentiation and cell fate.

pattern during etiolated and green seedling leaves development, bulk RNA-seq was further utilized to verify their expression, which showed these 15 TFs consisted in both RNA-seq profiles (Figure S9, Table S18). Next, 4364 DEGs (2029 up- and 2335 down-regulated) were assessed along with the leaf expansion under dark and light conditions, and 51 core TFs related to phytohormone regulatory pathways were screened (Figure S9, Table S19). The distinct gene expression patterns switching from dark to light growth condition indicated that light drove the core-TFs expressions to mediate leaf cell differentiation.

Overall, the cell distribution in the etiolated seedling was clustered in branch node 1, whereas the green seedling was enriched in branch node 2, indicating a dramatic difference in the cell developmental trajectory of the two conditions (dark and light). DEGs and TFs expression tendency provided transcriptome dynamic, of which photosynthesis DEGs still dominated the differentiation difference, especially in the procedure of dark cells evolving into light cells at the cell state 3. Therefore, light seems like a catalyst that enables accelerate the differentiation speed from etiolated- convert into green-seedling cells.

#### scRNA-seq revealed the dark condition perturbed the leaf cell cycle

Light intensely alters plant growth and development through regulations of phytohormones, which synergistically regulate the progression of the cell cycle. Cell cycle phases reflected transcriptional heterogeneity during scRNA-seq (Hsiao et al., 2020). In present study, we investigated the expression of genes during phase G1 (9), G1S (433), S (609), G2M (144), M (0) and non-cycling (491) phases of the cell cycle (Figure S10a, Table S20). In addition, the gene expression analysis for all cells of each cluster under light and dark conditions were presented separately for each phase of the cell and non-cell cycle using circos plots and observed that the majority of dark leaf cells were in the S phase (Figure 4a,b, Figure S10b,c, Table S20). Furthermore, we also performed RNA velocity analysis of DEGs in each cell cluster using velocity vorticity, which was visualized by drawing an arrow for each cell, thus linking the actual state to the estimated future state (Liu et al., 2022). Furthermore, the RNA velocity dynamics generated five vectors pointing to the different trajectory fields indicating the direction of cell development (Figure S10d). As expected from the analysis, the etiolated seedling cells showed little variation in RNA velocity (short or no arrows) of epidermal cell clusters (vectors 1 and 5), while the transcriptional profile of the green seedling cells showed a larger variation (Figure S10e,f). Whereas, the top five DEGs of each sub-cluster across the phases of the cell cycle were visualized using bubble heat maps to indicate the distribution of upregulated genes during the cell cycle based on two parameters: gene expression level and cell percentage (Figure 4c, Table S21). Cell

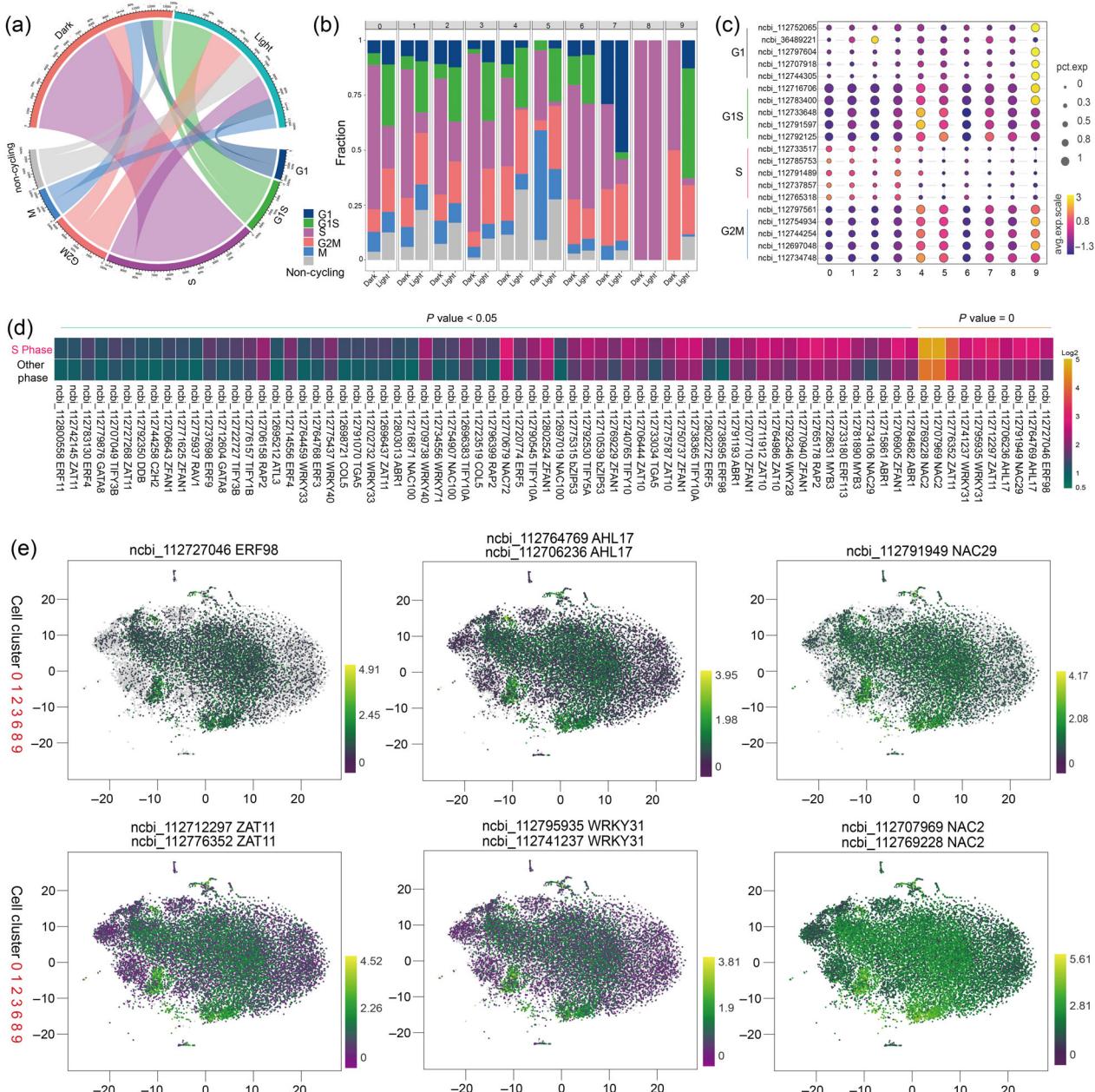
cycle and RNA velocity analysis revealed the long-term dark condition disrupted the normal cell cycle process when comparing green seedlings. We speculated that non-autotrophic dark seedlings not executing the photosynthesis reaction were a major factor, which indirectly reduced the inner dynamic of cell differentiation.

Due to the dark-seedling cells dropping into the S phase (Figure 4a), we next explored the TFs in the DEGs profile. 74 TFs also illustrated the differential expression patterns between the S-phase and other phases of the cell cycle, which facilitated transcriptome heterogeneity between cell cycles of seedlings grown under dark and light conditions (Figure 4d). The distribution density of 10 TFs ( $P$ -value = 0) with high expression in specific cell clusters was displayed using tSNE plots (Figure 4e). Interestingly, the core TFs, including AHL17, ERF98, NAC29, WRKY31 and ZAT11 were highly expressed in cells of etiolated seedlings, whereas NAC2 was dynamically expressed in both seedlings. Here, we demonstrated the utility of the scRNA-seq technique in revealing cellular transcriptional heterogeneity throughout the phases of the cell cycle and identified key TFs specifically and consistently expressed with the former cell type heterogeneity (Figure 4e). This result provides a potential TFs-pool for further demonstrating the functional details of the critical genes-regulated cell cycle difference between the dark- and light-seedling leaves.

#### Dark environment repressed the mesophyll cell differentiation and chlorophyllide synthesis

The obvious phenotype between etiolated- and green-seedling is leaf colour, to understand the mechanism of dark environment repressing the chlorophyllide production based on single-cell level. Mesophyll cells (Figure 5a) were extracted to rebuild the development trajectory, which pseudo-time trajectory contained three cell differentiation states, and dark cells mainly enriched in state 1 (Figure 5b, Table S22–S24). Further 631 core-DEGs were identified from cell trajectory, state and fate DEGs profiles (Figure 5c), which involved ribosome and photosynthesis pathways (Figure 5d). Of these, 22 TFs responded to the gene regulation pathway to modulate mesophyll differentiation (Figure 5e,f). The results implied that light could enhance the mesophyll differentiation ability, but dark produced an opposite impact.

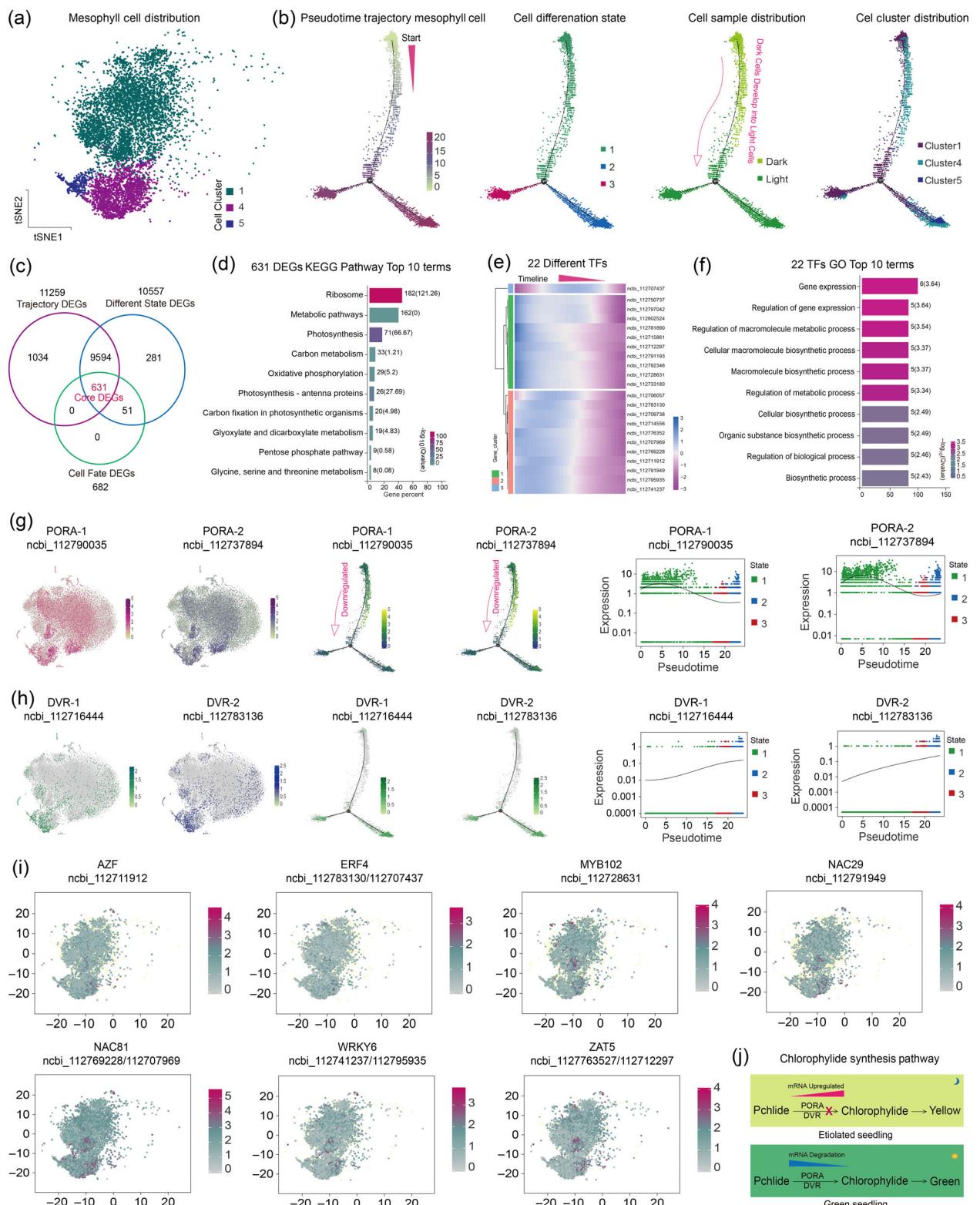
Furthermore, dark cell development repressing inhibited the chlorophyllide synthesis, especially the critical enzyme PORA (Protochlorophyllide Reductase, 2 genes from 2 sub-genome) transcription abundance was highly expressed in dark mesophyll clusters (Figure 5g), PORA expression level down-regulated in the trajectory of dark-cells converted into light-cells, and which intensively related to the regulation of cell differentiation from state 1 developed to 2–3. Additionally, other chlorophyllide



**Figure 4** Distribution of differentially expressed genes in each phase of the cell cycle between one-week-old etiolated and green seedlings leaves. (a) The circos diagram showed the correspondence between etiolated and green seedlings and the various phases of the cell cycle. G1, G1S, S, G2M and M were the various phases during the cell cycle. The different colours of the links inside the circle represented the different cell clusters that correspond to different phases of the cell cycle. The width of the links represented the number of cells. For instance, the thicker lines represented more cells. (b) The percent stacked bar chart indicated the proportion of cells in various phases of the cell cycle in each cell cluster. Various colours represented distinct phases of the cell cycle, and the length of the corresponding colour rectangle was proportional to the percentage of cells in that particular period. (c) Average expression levels of the top five DEGs in each cell cluster across the phases of the cell cycle. A larger bubble size represented a higher proportion of the sum of the expression of up-regulated genes in each phase. Dark colour of the bubble indicated a higher expression of the up-regulated genes. (d) The heatmap showed the average expression of up-regulated TFs in cells at S- and other phases. (e) The plots illustrated expression of representative TFs with a higher proportion (>50%) in the S-phase of the cell cycle in distinct cell clusters.

synthesis restricts enzyme *DVR* (Divinyl chlorophyllide an 8-Vinyl Reductase) mRNA presented a similar result with *PORA*, but not dominated the dark-cell differentiation (Figure 5h). The higher expression *PORA* was probably controlled by 11 TFs which exhibited similar expression tendency with *PORA* in dark mesophyll cells (Figure 5i). Accordingly, we proposed a model

(Figure 5j) was *PORA* and *DVR* expression level up-regulated in mesophyll under the dark condition to restrict the process of pchlde converted into chlorophyllide. Once the leaf accepted the light to degrade the *PORA* and *DVR* mRNA, etiolated-seedling leaf recovered green colour due to chlorophyllide synthesis.



**Figure 5** Developmental trajectory of mesophyll cells under different light conditions. (a) Mesophyll cells (cluster 1, 4, 5) distribution in t-SNE map. (b) Mesophyll cells pseudo-time developmental trajectory and differentiation states. (c) Core DEGs identification from the profiles of trajectory, cell states and cell fate. (d) 631 DEGs involved in the KEGG enrichment pathway. (e) 22 differentially expressed TFs in the profile of 631 DEGs. (f) 22 TFs enriched GO terms. (g) PORA expression pattern in cell distribution map, mesophyll development trajectory and cell states. (h) DVR expression pattern in cell distribution map, mesophyll development trajectory and cell states. (i) 11 TFs expression pattern in mesophyll cell distribution map. (j) Chlorophyllide synthesis model in a distinct light environment based on scRNA-seq result.

## Dark condition enhanced epidermal cell expansion by phytohormone pathway

The epidermal cells are model cells for the study of plant cell morphogenesis. Epidermal cells are affected by plant hormones and grow interlaced with each other, thus producing puzzle-like cells on a two-dimensional plane (Chen *et al.*, 2015). Herein, based on the results of the transmission electron microscopy and scanning electron microscopy of sections from leaf epidermal cells, we observed that the epidermal cell length and area of leaf from etiolated seedlings were significantly greater than those of green seedlings, with 7.78%–10.02% and 13.43%–24.11%, respectively, on the 3rd, 5th and 7th day (Figure 6a,b).

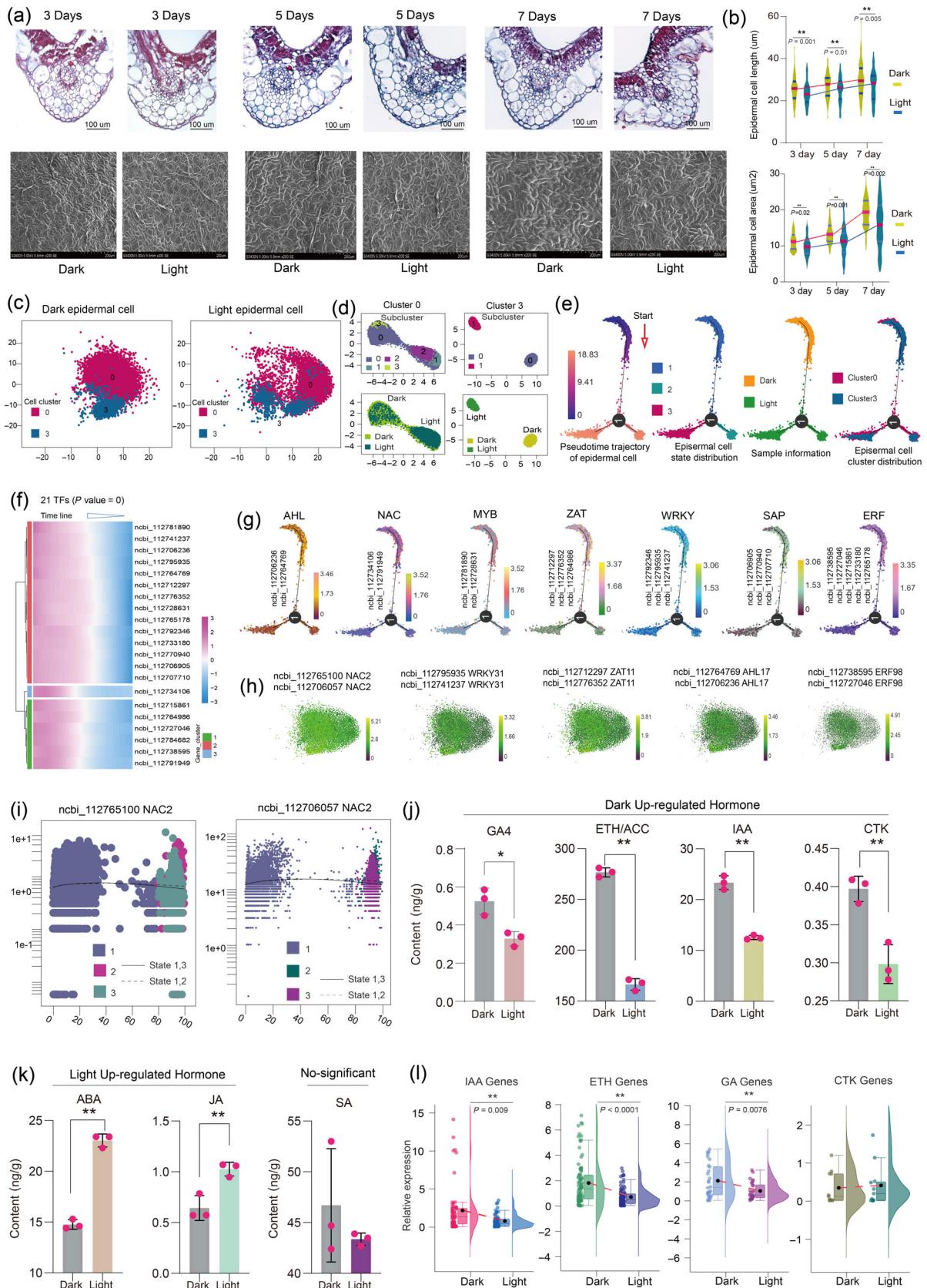
Furthermore, to explore the differentiation and development of epidermal cells in peanut leaf under dark and light conditions, we isolated epidermal cell populations especially in cell clusters 0 and 3, and reconstructed the epidermal cell developmental trajectories (Figure 6c, Table S3). The epidermal cell clusters 0 and 3 were re-aggregated to obtain 4 and 2 cellular sub-clusters (Figure 6d), respectively. The pseudo-time trajectory analysis to understand epidermal cell development identified a total of 8551 DEGs along the main stem of the timeline, whereas all cells were classified into three differentiation states and a total of 8191 DEGs were associated with these three distinct states were also characterized (Table S25, Figure 6e). Importantly, among three states, the cells of state 1 are distributed along the constructed developmental trajectory on the main stem near the 'start', and then the cells differentiated to produce two branches, which further differentiated into state 2 and state 3, respectively, and a total of 421 TFs were differentially expressed in these cells (Tables S26, S27, Figure 6e). Additionally, the epidermal cells of etiolated seedlings showed one branching point of the developmental trajectory concentrated at the developmental initial location, while the cells of green seedlings were clustered at another branching point, indicating diverse transcriptome re-programming between green and etiolated seedlings of peanut (Figure 6e). After filtering the TF profiles based on *P*-values, a total of 21 TFs were identified, while the expression levels of these 21 TFs gradually decreased with the pseudo-time line, and the functions of these transcripts were related to the regulation of plant hormones (Table S27, Figure 6f). Specifically, these TFs with homologous genes were classified into gene families such as *AHL*, *NAC*, *MYB*, *ZAT*, *WRKY*, *SAP* and *ERF*, and the expression and distribution of these TFs were visualized in t-SNE plot (Figure 6g,h). From the results, only two *NAC2* were differently expressed in the profile of epidermis cell differentiation state 1–2 versus state 1–3, which might be playing an important role in epidermal cell fate differentiation (Figure 6f–i, Table S26).

It is widely reported that plant hormones play various important roles during growth and development. The functions of plant hormones are regulated by light conditions, which is the basis for light regulation of plant growth and development. For instance, abscisic acid (ABA) inhibits seed germination, and gibberellin acid (GA) promotes seed germination and this process is regulated by light signals (Wit *et al.*, 2016). To further unravel this phenomenon, we investigated the phytohormone contents of peanut seedling leaves under dark and light conditions. As result, the seedlings grown in dark had significantly higher GA4, Ethylene/1-aminocyclopropanecarboxylic acid (ETH/ACC), indole-3-acetic acid (IAA) and cytokinin (CTK) contents as compared to light-grown seedlings. However, ABA and jasmonic acid (JA) contents were significantly higher in light-grown seedlings. Salicylic acid (SA) content was not significantly different between both light and dark-grown seedlings (Figure 6j,k). The relative expression of phytohormone-related gene sets also demonstrated that the expression of IAA, ETH and GA genes was significantly higher in etiolated peanut seedlings compared to green seedlings (Figure 6l, Table S28). In conclusion, a total of 21 TFs and 10 hub genes with their homologous were identified, which may be important candidate drivers for differences in peanut leaf epidermal cells in various light environments.

## Ectopic expression of peanut *AHL17* enhanced leaf growth in *Arabidopsis*

As mentioned above, we mined critical TFs, such as *AHL17*, *MYB3*, *NAC29*, *BBX25*, *ZAT11*, *NAC2*, *ERF98* and *WRKY31*, involved in cell development pseudo-time trajectory and various phases of the cell cycle. Furthermore, we extracted epidermal cell clusters for sub-clustering and identified the TFs involved in the regulation of important phytohormones. Due to transcriptome abundance of *AHL17* (Figure S11), we focused on elaborating the biological function of *AHL17* (*Arah.y.6X9KLT*) (Figure S12) with expression pattern (Figures 3g, 4e, 6h). *AHL17* was highly expressed in different peanut tissues (leaves, roots, stems, pegs, seeds and pods) at distinct growth stages, especially in leaves, pegs and seeds, indicating a clear tissue-specific expression pattern (Figure S13a,b). Furthermore, *AHL17* expression differed significantly between etiolated and green seedlings, peaking in the epidermal cells of etiolated seedlings (Figure S13c,d). The above expression resulted from tissue to single-cell level suggest that *AHL17* may have many specific functions regulating various tissue development. *AHL17* is an AT-hook motif nuclear-localized protein 17-like encodes an uncharacterized protein containing 310 amino acids (Figure 7a). Phylogenetic analysis indicated that *AHL17* was similar to *AtAHL1* and *AtAHL10* in group 3

**Figure 6** scRNA-Seq for comprehensive analysis of differentially expressed genes and transcription factors of the epidermal cells and hormone profiles of etiolated and green seedlings. (a) Transmission and scanning electron microscopy analysis of leaf epidermal cells. (b) The violin plots represented the length and area of leaf epidermal cells (\**P* < 0.05, \*\**P* < 0.01). (c) Distribution pattern of epidermal cell clusters 0 and 3. Each dot represented a cell. (d) The t-SNE plot showed sub-clusters after the second round of clustering within epidermal cell clusters 0 and 3, and their distribution in the dark and light treatments. (e) Analysis of pseudo-time differentiation trajectories of epidermal cells, four plots were based on pseudo time, cell state, sample type and cell cluster, respectively. 'Start' displayed the beginning of the pseudo time. (f) Clustering and expression kinetics of 21 TFs along the pseudo-time trajectory. (g) Pseudo-time trajectories of the expression distribution of core TF families (*AHL*, *NAC*, *MYB*, *ZAT*, *WRKY*, *SAP* and *ERF*) involved in epidermal cell development. (h) Expression patterns of *NAC2*, *WRKY31*, *ZAT11*, *AHL17* and *ERF98* in epidermal cells. (i) The scatter plot exhibited the branching trajectory of *NAC2* expression levels along with the pseudo-timeline. The horizontal coordinate was the pseudo-time axis; the vertical coordinate was the branch differential gene expression. Three colours represented three states of cell differentiation. The solid and dashed lines represented distinct branches. (j, k) Bar plots illustrated the up-regulated phytohormones under dark- and light-grown peanut seedlings (\**P* < 0.05, \*\**P* < 0.01). (l) Relative expression of phytohormone regulatory gene sets under dark- and light-grown peanut seedlings (\*\**P* < 0.01).



(Figure 7b). Subsequently, subcellular localization showed that the AHL17-GFP-anchored organelle was the nucleus in *Arabidopsis* protoplast cell, indicating that the AHL17 protein was mainly

localized in nucleus (Figure 7c). Additionally, the binding sites of AHL17 protein in the peanut genome were enriched by DNA affinity purification sequencing (DAP-Seq), and 5 short-motifs

were detected, of which AA(T)AAAATA was the main conserved motif of AHL17 protein (Figure S14, Tables S30, S31).

Moreover, *AHL17* was ectopically overexpressed in *Arabidopsis* to generate *AHL17-OE* transgenic lines and compared their phenotypes at various growth stages such as germination, seedling stage and flowering stage under light and dark conditions (Figure 7d–i). At the time of germination and seedling stages, the cotyledons of the *AHL17-OE* lines which were overexpressed, showed significantly larger length and width compared with the wild type (WT) and the empty plasmid (pBWA(V)HS), and *AHL17-OE* lines also exhibited a significantly higher rate of growth in terms of whole plant size (Figure 7d–g). While, at the flowering stage, *AHL17-OE* lines showed enhanced leaf size expansion and whole plant expansion as compared to WT and pBWA(V)HS (Figure 7h,i). The *AHL17-OE* lines showed significant early flowering in light seedlings, while no noticeable change in leaf size was observed in the *AHL17-OE* lines under dark condition compared to WT and pBWA(V)HS at the seedling stage. In addition, scanning electron microscopy observations of the leaves of WT, pBWA(V)HS and *AHL17-OE* showed that *AHL17* overexpression increased the area of epidermal cells at all growth stages as compared to WT and pBWA(V)HS (Figure 7j,k, Figures S15, S16). These results suggested that *AHL17* positively regulated epidermal cell enlargement and thus promoted leaf expansion leading to enhanced plant size. However, the mechanism of early flowering in *AHL17-OE* is not clear. Since scRNA-seq results suggested that *AHL17* is involved in the regulation of phytohormone pathways. To prove this, we collected overexpressed *AHL17-OE* *Arabidopsis* leaf tissues from three stages for phytohormone profiling using LC-MS. As compared to WT and pBWA(V)HS, the *AHL17-OE* lines showed significantly increased levels of IAA, CTK and the ethylene precursor 1-aminocyclopropane-1-carboxylic acid (ACC) (Figure 7l). Additionally, the truth expression levels of crucial genes in these phytohormone biosynthesis pathways were detected by qPCR (Figure S17). The expression of *YUCCA* (*Flavin Monooxygenase*) (IAA) and *ACS* (*ACC SYNTHASE*) (ETH) presented a high expression level in *AHL17-OE* lines when compared to WT and pBWA(V)HS grown under dark condition (Figure S17a,c). Cytokinin synthesis genes *CYP735A* (*CYTOCHROME P450 SUPERFAMILY PROTEIN*) and *LOG* (*LONELY GUY*) were particularly upregulated under dark condition (Figure S17b). Additionally, 245 genes were screened out based on DAP-Seq annotation, which was mainly associated with the synthesis of IAA, CTK and ETH (Table S32). Therefore, we hypothesized that *AHL17* played important role in regulation of phytohormone pathway and enhanced leaf epidermal cell enlargement (Figure 7m).

## Discussion

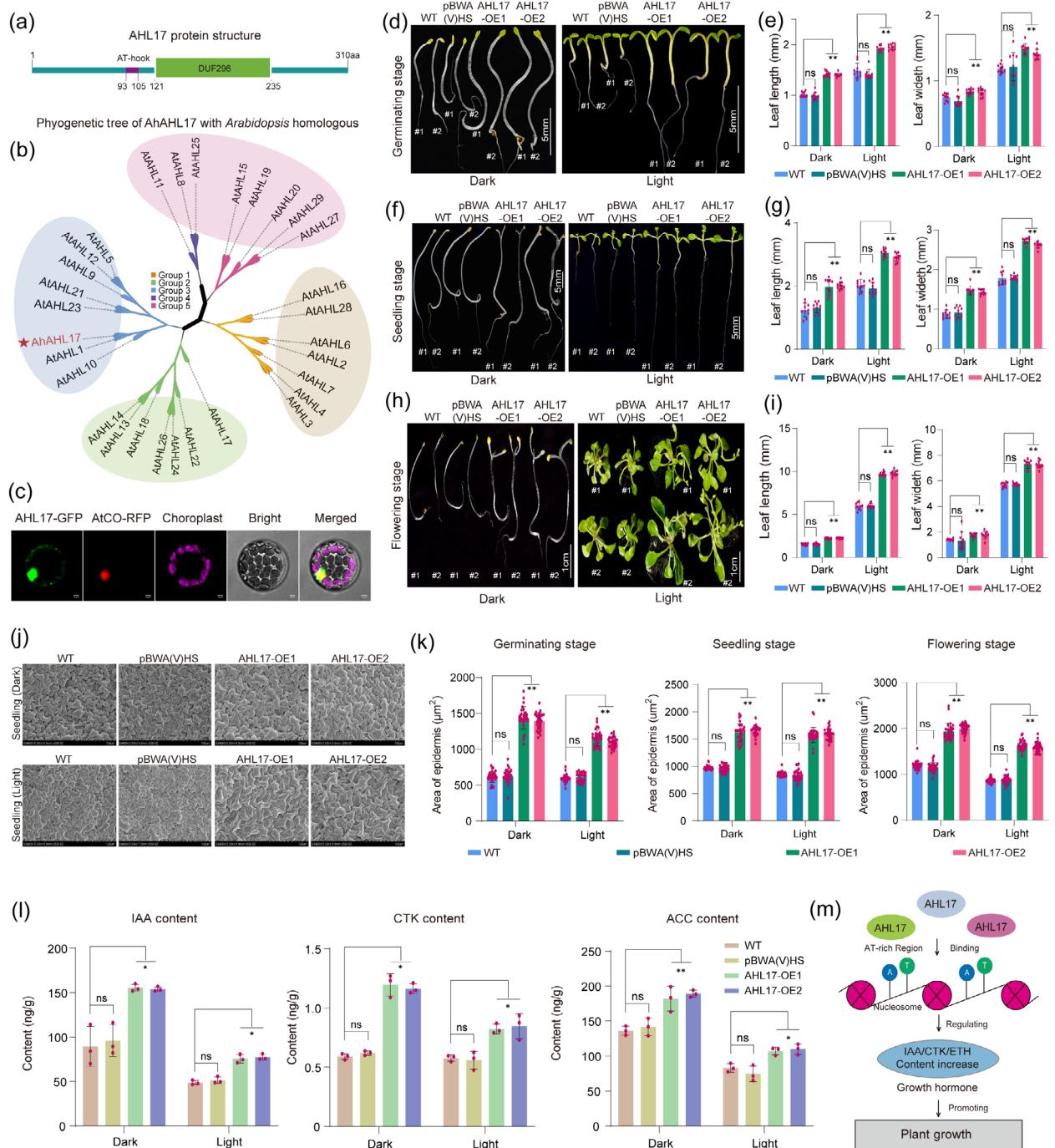
Peanut is completely distinguished from other crops due to their unique fruit development characteristics, aerial flowers and underground fruit development (Chen et al., 2013; Sinha et al., 2020). In our previous study, a transcriptome map for peanut pod developmental stages identified a total of 245 genes that responded to light stimulation, two involved in dark condition and six participated in de-etiolation (Chen et al., 2016). Furthermore, extra insights into transcriptome programming associated with peanut pods development were gained under dark condition, including TF regulation, phytohormone signalling and photosynthesis, which suggested that light suppressed aerial

pod expansion and embryonic development, while underground fruit expansion is due to the absence of light induction (Liu et al., 2020b). Importantly, our previous scRNA-seq study found that the cell differentiation fate of peanut leaf was regulated through the phytohormone accumulation pathway in etiolated seedlings (Liu et al., 2021). However, the present single-cell transcriptomic study describes the diversity of cellular clusters, cell development trajectories and cell cycle changes in leaves of peanut seedlings grown under dark and light conditions. scRNA-seq results revealed DEGs profiles in light-suppressed peanuts and explored the cellular and transcriptional heterogeneity in dark/light-grown leaf cells. Besides, our results provide new insights into the single-cell transcriptome gene atlases under dark/light morphogenesis in peanuts and its dynamic expression patterns during leaf development.

### scRNA-seq described the different gene atlas in peanut seedlings grown under the dark and light conditions

The irreversible transformation of plant seedlings from etiolated growth to de-etiolated morphogenesis after exposure to light triggers a variety of developmental programs in plant tissue (Sinclair et al., 2017). Several important studies have reported that the morphological transformation of plants and leaves is considered to be the crucial response of plants under light or dark. However, due to technical limitations, the cellular developmental processes using single-cell transcriptomics underlying this transformation have not been investigated (Wan et al., 2008; Wang et al., 2021a). Herein, plant scRNA-seq research has demonstrated the technical advantages of its large-scale high-throughput assays and the ability to analyse the heterogeneity of single-cell gene expressions patterns, such as cotyledon leaf vein development and stomata genealogy trajectories (Liu et al., 2022; Lopez-Anido et al., 2021). The enzymatic digestion method was used to obtain protoplasts and generated the scRNA-seq data through the microfluidics platform to construct the scRNA-seq profile in peanut etiolated and green seedling leaves. scRNA-seq described a gene atlas showing different expression patterns in light/dark treated seedlings, which provided a gene resource for further elucidating gene regulation triggered plant morphological variation under distinct light conditions.

Based on the analysis, six cellular styles described as distinct cell clusters 0–9 were identified, especially the distinct cell clusters that exhibited obvious cell differentiation states under distinct light conditions. Light-grown cells clustered into more differentiation states than dark-grown cells represented that light condition enhanced the cell differentiation dynamic in peanut leaf. This biological process is attributed to the photosynthesis reaction in light-grown cells, which generates carbohydrates for autotrophic nutrition, alters the supply mode and gene expression pattern of original growth energy in every single light-cell compared with dark-grown cells, resulting in light-grown cells divided into more cell differentiation states. Besides, the gene expression landscape of each cell cluster was characterized by its marker genes, reflecting the considerable precision of scRNA-seq. Peanut, as a non-model plant, has few reported marker genes that can be used to identify cell types. Therefore, in this study, we used direct homologous between *Arabidopsis* and peanut genes to identify peanut cell types. The gene expression landscapes of each cell cluster were characterized by its marker genes demonstrating the feasibility and efficiency of the scRNA-seq



**Figure 7** Overexpression of peanut AHL17 gene in Arabidopsis and comparison of AtAHL17-OE with wild type. (a) Schematic diagram of AHL17 protein structure. (b) Phylogenetic tree of AhAHL17 homologous genes in Arabidopsis. (c) Subcellular localization of AHL17 protein in the nucleus of Arabidopsis protoplasts. Scale bar was 10 μm. (d, f, h) Comparison of phenotypes of WT, pBWA(V)HS and AHL17-OE lines at the stages of seed germination, seedling and flowering under dark and light conditions. (e, g, i) The measurement of leaf length and width in WT, pBWA(V)HS and AHL17-OE lines plants ( $n = 10$ ) (\* $P < 0.05$ , \*\* $P < 0.01$ ). (j) Scanning electron microscope (SEM) analysis of the size of epidermal cells in WT, pBWA(V)HS and AHL17-OE transgenic lines at the seedling stage. Scale bar was 100 μm. (k) Area of epidermal cells at germinating, seedling and flowering stage in WT, pBWA(V)HS, and AHL17-OE plants under dark and light conditions ( $n = 30$ ) (\* $P < 0.05$ , \*\* $P < 0.01$ ). (l) Phytohormone profiles of IAA, CTK and ACC content of WT, pBWA(V)HS and AHL17-OE plants were generated using LC-MS. The bar plots indicated the IAA, CTK and ACC content of WT, pBWA(V)HS and AHL17-OE plants under light and dark conditions (\* $P < 0.05$ ). (m) A putative model illustrated the biological function of peanut AHL17 protein in plant growth.

method. In addition, this method provides a reference for cell type identification in future studies by scRNA-seq in non-model species. Based on the identified cell types, we also

reported the marker genes for each peanut leaf cell type that are specifically expressed and likely have a key role in the regulation of leaf structure and function.

## Identification of the core TFs regulated the leaf cell development

Leaf cell development is a continuous process accompanied by significant transcriptional regulation, and physiological and metabolic processes (Bezrutczyk *et al.*, 2021). Here, the scRNA-seq could facilitate the understanding of plant organ development and the discovered key regulatory genes (Shaw *et al.*, 2021). Through the analysis of scRNA-seq, we observed that about 65.18% of cells of etiolated seedlings and 47.32% of green seedlings belong to leaf epidermal cells with phytohormones gene expression patterns (Figure 2). In contrast, the remaining cells were distributed among eight different cell clusters. Generally, the predominant number of epidermal cells provided an opportunity to identify critical pathways during peanut leaf blade morphogenesis and identified the main cell development processes and trajectories during leaf cell growth. Thereby, KEGG and Venn diagram analysis of DEGs and TFs specifically expressed in different cell types under dark and light conditions provided in-depth knowledge about the biological functions of various leaf cells. Importantly, 10 differentially expressed TFs from families *AHL*, *MYB*, *NAC*, *ERF*, *ZAT*, *CCH* and *BBX*, were found to be associated with photoperiod and phytohormone pathways (Figure 3d-i) (Mu *et al.*, 2022). Previous studies have shown that key regulators of leaf morphological transformation, such as *AHL*, *MYB*, *NAC* and *ERF*, were expressed at specific developmental stages of the leaf (Chen *et al.*, 2018; Lin *et al.*, 2003; Liu *et al.*, 2021; Singh *et al.*, 2016).

Furthermore, the results of the pseudo-time analysis showed that all cells can be sorted along a major developmental trajectory (Figure 3j). Analysis of the distribution of different cell states, samples and cell clusters showed that cells in different cell clusters were distributed relatively and independently along the pseudo-time trajectory. Expression tendency analysis of differentially distributed TFs yielded consistent and dynamic expression patterns of TFs in different cell clusters. The temporal trajectories were coordinated with their stage of leaf cell development, for instance, a shift from low to high expression levels for all TFs except *BBX25*, which displayed expression patterns from high to low (Figure 3g-n). Further cell cycle and RNA velocity analysis revealed that the epidermal cells of dark-grown seedlings have low RNA metabolic activity, which was compatible with their status as predominantly out during the S phase of the cell cycle (Figure 4a,b) (Liu *et al.*, 2022; Svensson and Pachter, 2018; Zywitsa *et al.*, 2018). In contrast, light-grown seedlings showed high RNA metabolic activity compared with dark-grown seedlings, indicating that the cells of light-grown seedlings have a high rate of cell differentiation compared with dark seedlings. Consequently, the above results suggested that scRNA-seq could reconstruct leaf cell differentiation trajectories with various phases of the cell cycle for accurate mining of crucial genes in etiolated and green seedlings.

## Light condition promoted the cell differentiation

Light repressed the seedling development by reducing the growth hormone content, scRNA-seq constructed cell development trajectory indicated that light accelerated the cell differentiation procedure during the dark cells evolved into light cells, similar phenomena commonly existed in mesophyll and epidermis. Interestingly, mesophyll contained plenty of chloroplastic to modulate the pchlde transformed into chlorophyllide, the last step restricts enzyme *PORA* should be highly expressed in

mesophyll clusters, but we also discovered that *PORA* expression highly enriched in the epidermis (Cluster 0). scRNA-seq result supported a novel putative illustration that the *PORA* mRNA firstly degraded in the epidermis under light stimulation, then the degradation signalling transduction into mesophyll to terminate *PORA* repressing influence on chlorophyllide synthesis. Additionally, once the light-induced photosynthesis reaction recovered to normal operation, the epidermis cell differentiation dynamic was enhanced by the light environment, which reduced the growth hormone content and closed the transcription of phytohormone-correlated TFs. Conclusively, scRNA-seq was far-beyond traditional tissue-bulk transcriptome and provided several novel biological insights for plant light developmental biology on single-cell resolution.

## *AHL17* mediated the phytohormone pathway to promote leaf epidermal cell enlargement

The *AHL* gene family characterized by the presence of an AT-hook motif and an unknown functional domain DUF296 is highly conserved in land plants (Bishop *et al.*, 2020). Previous studies have shown that *AHL* proteins regulate plant growth and developmental processes, including leaf expansion, flower enlargement, hypocotyl elongation (Liu *et al.*, 2021; Zhao *et al.*, 2013). Even though several *AHL* genes are involved in the regulation of plant hormones, such as ABA, JA and CTK (Bishop *et al.*, 2020; Wong *et al.*, 2019; Zhao *et al.*, 2013). However, functional exploration of the *AHL* protein family has focused on *Arabidopsis*, and few functions have been systematically described in peanuts. Here, we observed that leaf epidermal cell development in dark-grown versus light-grown seedlings was associated with the *AHL17* transcription factor. When *AHL17* was ectopically expressed in *Arabidopsis*, the *Arabidopsis AHL17-OE* lines showed a significant expansion of leaf cell area and increased leaf size. Additionally, scRNA-seq results indicated that *AHL17* was involved in the signalling of phytohormone. *AHL17* activates signalling pathways of growth phytohormones crucial for seedling epidermis development. Therefore, we proved that *AHL17* was crucial during the dark-to-light environment transition (Deepika *et al.*, 2020; Liu *et al.*, 2017).

Plant *AHLS* (*AHL22*, *AHL27*, *AHL20*, *AHL19*) modulated several aspects of plant development, which involved the homeostasis of phytohormones or the mediation of their responses, such as enhanced gibberellin induced adult leaf growth (*AHL25*), restricted petiole elongation by antagonizing the growth-promoting PIFs with auxin, ethylene, brassinosteroids pathway (*SOB3/AHL29*) and delayed the juvenile into adult vegetative phase change by stimulating cambium activity with cytokinin signalling (*AHL15*) (Favero *et al.*, 2020). However, according to the DAP-seq result, *AHL17* encoding protein locates in nucleus to bind the conserved motif of AA(T)AAAATA and activates downstream genes expression in peanut genome that involves in multiple phytohormone signalling pathway, including *IAA9*, *YUCCA5*, *LOG3*, *CYP735A2* and *ACS*, which are associated with the synthesis of auxin, cytokinin, ethylene (Table S32). Therefore, *AHL17* promotes the entire leaf expansion preferentially by depending upon diversity phytohormone, and externally improves the growth ration of epidermis under light condition. Additionally, skotomorphogenesis under dark-induced seedlings is positively regulated by ethylene signalling when treated with exogenous ethylene (Mazzella *et al.*, 2014). Similarly, our findings suggested that the changes in the ACC content of ethylene precursors in the hormonal profile of *AHL17-OE* lines are

consistent with this. *AHL17* may up-regulate the ACS and ACO (ACC OXIDASE) to accelerate epidermis cell expansion by increasing the ethylene content under dark circumstance, while this hypothesis still needs to be validated with further experiment data. Conclusively, peanut *AHL17* mediates the positive regulation of plant growth via the hormones pathway, which functional interpretation may help to accelerate peanut breeding and extend our understanding of *AHLS* molecular details in peanut.

In summary, our results firstly map the transcriptome profiles of dark and light-grown peanut seedlings, provide new marker genes for identifying peanut leaf cells for more systematic and extend accurate understanding of leaf cell developmental mechanisms. Finally, we believe that the scRNA-seq collaboration with other single-cell or spatial multi-omics will enable the acceleration of peanut individual-cell science research in the future.

## Materials and methods

### Plant material, growth conditions and phenotype assays

A popular peanut cultivar Hanghua2hao was used in the present study. After 3 h of immersion in water, peanut seeds were sterilized in 0.1% (v/v) bleach (5.25% sodium hypochlorite) for 15 min, washed in sterile water and sown on the soil. Seeds were grown under the conditions of complete darkness and 12 h of alternating light-dark cycles, respectively, in environmental chambers with a temperature of 25 °C and a humidity of 70% conditions, respectively. The morphological data for mesocotyl, stem, petioles and leaf blade lengths of seedlings were evaluated by vernier calliper at 3, 5 and 7 days after seed sowing. The phenotypic data were recorded for five seedlings for each condition/time.

### Peanut leaf collection and protoplast isolation

The peanut seedling leaves grown under light and dark induction were collected separately at 3, 5 and 7 days after germination. One-week-old leaf samples were cut into 1 mm strips and placed in centrifuge tubes containing 30 mL of digestion solution containing 3% cellulose (R-10) and 0.3% pectinase, 1.5% macerozyme, 0.25% Bovine Serum Albumin (BSA), 5 mM MES and 8% (w/v) mannitol without calcium ( $\text{Ca}^{2+}$ ) and magnesium ( $\text{Mg}^{2+}$ ) ions. With shaking at 40 rpm for 2 h at 25 °C. Cells were filtered with a 40 µm cell strainer. Cell activity (>90%) was detected by trypan blue staining and cell member was measured using a haemocytometer. Protoplasts were resuspended in 8% (w/v) mannitol solution in preparation for loading on the chromium controller.

### scRNA-seq library construction, sequencing and raw data analysis

10xGenomics single-cell transcriptome 3' gene expression libraries were constructed according to the manufacturer manual. Initially, approximately  $2 \times 10^4$  isolated single cells were packed into nanolitre-scale Gel Bead-In-Emulsions (GEMs) oil droplets. Then, the RNA of the GEMs droplet was reverse transcribed into cDNA. 3' gene expression libraries were constructed by PCR amplification and enzyme section segmentation. The scRNA-seq data was processed in Cell Ranger software ([support.10xgenomics.com](https://support.10xgenomics.com)) to obtain raw data by aligning sequencing reads to the peanut genome (GCF\_003086295 at [peanutbase.org](https://peanutbase.org)) (Bertioli et al., 2019), and the Cell Ranger was used to correct barcodes (Table S3). The Seurat R package (version 4.0.0) was used for quality control (QC) of raw data, cellular

dimensionality reduction, and clustering and grouping. Seurat normalizes and corrects the raw data based on QC metrics in order to obtain highly variable gene sets.

Furthermore, PCA feature selection and t-SNE dimensionality reduction analysis were used to visualize the data and cluster the cells. Seurat screens the differentially expressed genes (DEGs) by likelihood ratio test and identifies specific marker genes for each subgroup. Subsequently, gene expression fold changes of DEGs in divergent cell clusters were calculated by  $\log_2\text{FC}$  (fold change). The  $\log_2\text{FC}$  threshold for DEGs was identified as  $\log_2\text{FC} > 0.35$ , and  $P < 0.05$ . Meanwhile, transcription factors (TFs) were filtered from the above cell-type DEGs profile. In addition, the DEGs of each cell cluster between dark and light samples were identified by applying a cutoff  $\log_2\text{FC}$  of  $> 0.5$  and  $< -0.5$  ( $P < 0.05$ ) for up-regulated and down-regulated genes, respectively. Gene Ontology analysis was carried out for functional annotations of differentially expressed genes in different cellular clusters using the Gene Ontology database ([geneontology.org](https://geneontology.org)) (Ashburner et al., 2000). KEGG (Kanehisa and Goto, 2000) was used to identify the metabolic pathways or signal transduction pathways that were associated with significantly expressed DEGs. The top five genes with the highest expression were selected based on the results of DEGs analysis in each cell cluster. The expression of each marker gene was represented using heatmaps and bubble plots by using TBtools (Chen et al., 2020).

### Cell-specific tissue isolation and marker genes validation

Cell-specific tissue isolation was performed according to the previously described method (Wang et al., 2022). Leaf mid veins were cut off from blades as the vascular cells, and parenchymal cells surrounding the vascular system were dissociated under the stereomicroscope (Endo et al., 2016). The up and down epidermis layers were removed using tweezers as mixed epidermis population, and the mesophyll cells were isolated from the removed epidermis part by using tweezers slightly stripped exposed mesophyll tissue. Primordium cells were stripped from the shoot apical meristem, guard cells were isolated from the tweezers, and removed epidermis by using the capillary glass tube. All samples were frozen in liquid nitrogen for total mRNA extraction and reverse transcription. RNA was extracted from each cell group to construct libraries by following the SMART-seq protocol (SMART-Seq HT Kit, Takara). Quantitative real-time reverse transcription PCR (qRT-PCR) was used to measure the relative expression of marker genes for all tissue-specific cell types. The last cDNA library served as a template for detecting the gene expression level by applying conventional quantitative PCR with the ABI step one plus system. The epidermal cell population was used as the reference sample, and *Ah18S* was used as the internal reference control.

### Pseudo-time trajectory, cell cycle and RNA velocity analysis

The pseudo-time trajectory analysis consisted of single-cell trajectories, differential expression and branches analysis. Single-cell trajectories were structured using a matrix of cells and their gene expression profiles with the plot\_cell\_trajectory of Monocle (Version 3.0) (Trapnell et al., 2014). Differential expression analysis in Monocle was used to identify key genes related to the development and differentiation process with  $\text{FDR} < 1 \times 10^{-5}$  and to cluster genes with similar expression trends. The branches occurred because cells execute alternative gene expression programs. Branches appeared in trajectories during development,

when cells made fate choices, one developmental lineage proceeded down one path, while the other lineage produced a second path.

To perform the cell cycle analysis, a single-cell gene expression atlas was used as the basic data, and then the AddModuleScore function from the Seurat R package was used to compute each cell dropped into the genome duplication phase (G1, S, G2, and M phase) by recounted the expression levels of cell cycle marker proteins in cell cycle analysis (Macosko *et al.*, 2015). Finally, the genes with a significant difference were selected according to  $\text{abs log}_2\text{FC} > 0.35$  and  $P < 0.05$ .

The RNA velocity analysis was performed using the Python script *velocyto.py* ([www.velocyto.org](http://www.velocyto.org)) with the default parameter, and the BAM files with the spliced and unspliced reads were used as inputs (La Manno *et al.*, 2018). Based on the spliced mRNAs, RNA Velocity is expressed in terms of the number of spliced mRNAs and time derivative. For visualization, velocity vectors were plotted as locally average vector fields on the t-SNE embeddings of our high-quality cells by the R package *velocyto.R* (v0.6) with the parameter (min.grid.cell.mass = 50, grid.n = 20).

#### Validation of biological function of *AHL17* gene

The *AHL17* gene was cloned from the peanut leaf cDNA library. The homologous *AHL17* gene was searched by NCBI protein BLAST and a phylogenetic tree was constructed using MEGA version 7 (Tamura *et al.*, 2013). The CDS sequence of *AHL17* was inserted into the plasmid vector containing green fluorescent protein pBWA(V)HS-GFP. The fusion expression vector pBWA(V)HS-*AHL17-GFP* with the *AHL17* gene was constructed (Table S29). pBWA(V)HS-*AHL17-GFP* vector was transformed into *Arabidopsis* protoplasts using polyethylene glycol (PEG) at 28 °C for 18–24 h in dark culture, and slides were prepared and placed under a laser confocal microscope (Nikon C2-ER, Japan) for observation. Subsequently, in order to generate transgenic plants, the CDS sequence of *AHL17* was cloned into the pBWA(V)HS plasmid vector to construct the 35S::*AHL17* plasmid, which was transformed into *Agrobacterium tumefaciens* strain EHA105 to infiltrate inflorescence using the *Agrobacterium*-mediated floral dipping method in *Arabidopsis* (Zhang *et al.*, 2006). Additionally, the empty pBWA(V)HS plasmid vector was generated using the same *Agrobacterium*-mediated transformation method in *Arabidopsis*. The true transformed T<sub>0</sub> seeds were selected by hygromycin (Hyg), and stable *AHL17* overexpressed (*AHL17-OE*) lines were obtained at T<sub>2</sub> generation. We used the *AHL17-OE* lines to record leaf phenotypes. *Arabidopsis* (ecotype Columbia) species were used as wild type to compare the phenotypes of transgenic *Arabidopsis* lines with wild type.

#### Phytohormones uptake and detection assay

Leaf samples were collected under light and dark conditions at each time point and ground with liquid nitrogen using a grinder (30 Hz, 1 min). The 50 mg powder for each sample was mixed with the isotopic internal standards and extracted with 1 mL of extraction solvent mixture (methanol/ultrapure water/formic acid = 15:4:1, v/v/v). The extraction solutions were concentrated and re-solubilized with 100 µL of 80% methanol, filtered through a 0.22 µm syringe membrane and placed in the injection vial. Qualitative and quantitative phytohormone profiles were determined using ultra-performance liquid chromatography–tandem mass spectrometry (UPLC–MS/MS) with isotopic internal standards from Gene Denovo Biotechnology Co., Ltd (Guangzhou, China). Each assay was performed in three replicates.

#### Leaf ultrastructure observation and measurement

The leaves of *Arabidopsis thaliana* were collected at germination, seedling and flowering stages for ultra-microscopic observation. The leaves of all stages of etiolated and green seedlings were cotyledon, except for the green seedlings at the flowering stage, whose leaves were the first pair of true leaves of the rosette. Peanut and *Arabidopsis* leaves were cut into 1 mm × 2 mm pieces and placed into glass vials containing 2.5% glutaraldehyde and closed with a rubber stopper, the leaves were pumped out of gas using a vacuum pump until the samples sank at the bottom of the vials. Electron scanning electron microscopy, each fixed sample was rinsed three times with phosphate buffer (pH 7.4, 0.1 mol/L) at room temperature for 15 min; dehydrated with 30%, 50%, 70%, 80% and 90% ethanol for 15 min, followed by 100% ethanol for three times for 15 min; dried with a carbon dioxide critical point desiccator (Samdri-PVT-3D, Tousimi), and then sprayed with gold by ion sputtering (MSP-2S, IXRF systems) and observed under scanning electron microscope (Hitachi S-3400 N, Hitachi, Japan). Transmission electron microscopy, each fixed sample was rinsed with buffer solution and dehydrated with 50%, 70%, 80%, 95% and 100% ethanol in sequence, and the alcohol in the samples was gradually replaced with propylene oxide. Finally, the samples were sectioned, stained with uranyl acetate and observed under transmission electron microscopy (Hitachi H-600IV, Japan). The length and area of leaf cells were calculated using ImageJ ([imagej.nih.gov](https://imagej.nih.gov)).

#### Acknowledgements

This work was supported by National Key Research and Development Project (2023YFD1202802), Guangdong Provincial Key Research and Development Program-Modern Seed Industry (2022B0202060004), The Open Competition Program of Top 10 Critical Priorities of Agricultural Science and Technology Innovation for the 14th Five-Year Plan in Guangdong Province (2022SDZG05), The National Natural Science Foundation of China (32172051, 32301869), Guangdong Basic and Applied Basic Research Foundation (2021A1515010811, 2023A1515010098), China Agriculture Research System of MOF and MARA (CARS-13), Technology Special Fund of Guangdong Province Agriculture and Rural Affairs Department (2019KJ136-02), Guangzhou Basic and Applied Basic Research Foundation (202201010281), Agricultural Competitive Industry Discipline Team Building Project of Guangdong Academy of Agricultural Sciences (202104TD), Special Fund for Scientific Innovation Strategy-Construction of High Level Academy of Agriculture Science (R2020PY-JX004, R2020PY-JG005 and R2021PY-QY003), The Foundation of Director of Crop Research Institute of Guangdong Academy of Agriculture Sciences (202101), The Science and Technology Planning Project of Guangdong Province (2021A0505030047), The Special Support Program of Guangdong Province (2021TX06N789), Rural Revitalization Strategy Special Foundation GDAAS Specialist Working Station (2023 Gong Zuo Zhan 05), Planting Industry Development Program of 2023 Rural Revitalization Strategy Special Project (2023-440000-51020300-9528), Science and Technology Planning Project of Heyuan City (Heyuan She Nong Da Zhuan Xiang 2022002), Science and Technology Planning Project of Qingyuan City (2023KJ002), Guangdong Province Science and Technology Program (2023B1212060038), Lianping Peanut Modern Agricultural Industrial Park Project, Guangdong Province Crop Modern Seed Industrial Park Project. Rajeev K Varshney Thanks Food Future

Institute, Murdoch University for The Start up Grant for Its Contribution To This Study.

## Conflict of interest

The authors declare that there is no conflict of interest.

## Author contributions

Quanqing Deng, Hao Liu and Yanbin Hong developed the protoplast isolation approach. Puxuan Du, Yuan Xiao and Dongxiu Hu performed experiment on scRNA-seq validation and *AHL17* function analysis. Qing Lu, Haifen Li and Wenyi Wang performed the scRNA-seq experiment. Shaoxiong Li, Lu Huang and Haiyan Liu analysed the scRNA-seq data. Runfeng Wang and Hao Liu processed the experimental figure. Xiaoping Chen facilitated funding. Xuanqiang Liang supervised the study. Rajeev K Varshney and Vanika Garg contributed to conceptualization, supervision, intellectual guidance, data analysis and interpretation to this study. Quanqing Deng, Sunil S. Gangurde, Rajeev Varshney and Hao Liu wrote the article, Hao Liu and Xiaoping Chen conceived the study.

## Data availability statement

Data supporting the discovering of our work are available within the paper and its Supplementary Information files. Raw sequencing data have been uploaded to China National Center for Bioinformation (CNCB) under Bio-project accession PRJCA010983 and PRJCA015968.

## References

- Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P. et al. (2000) Gene ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25–29.
- Bertioli, D.J., Jenkins, J., Clevenger, J., Dudchenko, O., Gao, D., Seijo, G. et al. (2019) The genome sequence of segmental allotetraploid peanut *Arachis hypogaea*. *Nat. Genet.* **51**, 877–884.
- Bezrutczyk, M., Zöllner, N.R., Kruse, C.P.S., Hartwig, T., Lautwein, T., Köhrer, K., Frommer, W.B. et al. (2021) Evidence for phloem loading via the abaxial bundle sheath cells in maize leaves. *Plant Cell* **33**, 531–547.
- Bishop, E.H., Kumar, R., Luo, F., Saski, C. and Sekhon, R.S. (2020) Genome-wide identification, expression profiling, and network analysis of AT-hook gene family in maize. *Genomics* **112**, 1233–1244.
- Chen, C., Chen, H., Zhang, Y., Thomas, H.R., Frank, M.H., He, Y. and Xia, R. (2020) TBtools: An integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* **13**, 1194–1202.
- Chen, J., Wang, F., Zheng, S., Xu, T. and Yang, Z. (2015) Pavement cells: a model system for non-transcriptional auxin signalling and crosstalks. *J. Exp. Bot.* **66**, 4957–4970.
- Chen, X., Lu, Q., Liu, H., Zhang, J., Hong, Y., Lan, H., Li, H. et al. (2019) Sequencing of cultivated peanut, *Arachis hypogaea*, yields insights into genome evolution and oil improvement. *Mol. Plant* **12**, 920–934.
- Chen, X., Yang, Q., Li, H., Li, H., Hong, Y., Pan, L., Chen, N. et al. (2016) Transcriptome-wide sequencing provides insights into geocarpy in peanut (*Arachis hypogaea* L.). *Plant Biotechnol. J.* **14**, 1215–1224.
- Chen, X., Zhu, W., Azam, S., Li, H., Zhu, F., Li, H., Hong, Y. et al. (2013) Deep sequencing analysis of the transcriptomes of peanut aerial and subterranean young pods identifies candidate genes related to early embryo abortion. *Plant Biotechnol. J.* **11**, 115–127.
- Chen, Y., Wu, P., Zhao, Q., Tang, Y., Chen, Y., Li, M., Jiang, H. et al. (2018) Overexpression of a phosphate starvation response AP2/ERF gene from *physcomitrella patens* in *Arabidopsis* alters root morphological traits and phosphate starvation-induced anthocyanin accumulation. *Front. Plant Sci.* **9**, 1186.
- 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.
- Deepika, A., Sagar, S. and Singh, A. (2020) Dark-induced hormonal regulation of plant growth and development. *Front. Plant Sci.* **11**, 581666.
- Efroni, I., Ip, P.L., Nawy, T., Mello, A. and Birnbaum, K.D. (2015) Quantification of cell identity from single-cell gene expression profiles. *Genome Biol.* **16**, 9.
- Endo, M., Shimizu, H. and Araki, T. (2016) Rapid and simple isolation of vascular, epidermal and mesophyll cells from plant leaf tissue. *Nat. Protoc.* **11**, 1388–1395.
- Favero, D.S., Kawamura, A., Shibata, M., Takebayashi, A., Jung, J.H., Suzuki, T., Jaeger, K.E. et al. (2020) AT-hook transcription factors restrict petiole growth by antagonizing PIFs. *Curr. Biol.* **30**, 1454–1466.
- Favero, D.S., Lambolez, A. and Sugimoto, K. (2021) Molecular pathways regulating elongation of aerial plant organs: A focus on light, the circadian clock, and temperature. *Plant J.* **105**, 392–420.
- Gala, H.P., Lanctot, A., Jean-Baptiste, K., Guizou, S., Chu, J.C., Zemke, J.E., George, W. et al. (2021) A single-cell view of the transcriptome during lateral root initiation in *Arabidopsis thaliana*. *Plant Cell* **33**, 2197–2220.
- Gangurde, S.S., Wang, H., Yaduru, S., Pandey, M.K., Fountain, J.C., Chu, Y., Isleib, T. et al. (2020) Nested-association mapping (NAM)-based genetic dissection uncovers candidate genes for seed and pod weights in peanut (*Arachis hypogaea*). *Plant Biotechnol. J.* **18**, 1457–1471.
- Hong, Y., Pandey, M.K., Liu, Y., Chen, X., Liu, H., Varshney, R.K., Liang, X. et al. (2015) Identification and evaluation of single-nucleotide polymorphisms in allotetraploid peanut (*Arachis hypogaea* L.) based on amplicon sequencing combined with high resolution melting (HRM) analysis. *Front. Plant Sci.* **6**, 1068.
- Hong, Y., Pandey, M.K., Lu, Q., Liu, H., Gangurde, S.S., Li, S., Liu, H. et al. (2021) Genetic diversity and distinctness based on morphological and SSR markers in peanut. *Agron. J.* **113**, 4648–4660.
- Hsiao, C.J., Tung, P., Blischak, J.D., Burnett, J.E., Barr, K.A., Dey, K.K., Stephens, M. et al. (2020) Characterizing and inferring quantitative cell cycle phase in single-cell RNA-seq data analysis. *Genome Res.* **30**, 611–621.
- Ibáñez, S., Ruiz-Cano, H., Fernández, M.Á., Sánchez-García, A.B., Villanova, J., Micó, J.L. and Pérez-Pérez, J.M. (2019) A network-guided genetic approach to identify novel regulators of adventitious root formation in *Arabidopsis thaliana*. *Front. Plant Sci.* **10**, 461.
- Kanehisa, M. and Goto, S. (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30.
- Kim, J.Y., Symeonidi, E., Pang, T.Y., Denyer, T., Weidauer, D., Bezrutczyk, M., Miras, M. et al. (2021) Distinct identities of leaf phloem cells revealed by single cell transcriptomics. *Plant Cell* **33**, 511–530.
- La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K. et al. (2018) RNA velocity of single cells. *Nature* **560**, 494–498.
- Lau, O.S. and Deng, X.W. (2012) The photomorphogenic repressors COP1 and DET1: 20 years later. *Trends Plant Sci.* **17**, 584–593.
- Li, X., Liu, Q., Feng, H., Deng, J., Zhang, R., Wen, J., Dong, J. et al. (2020) Dehydrin MtCAS31 promotes autophagic degradation under drought stress. *Autophagy* **16**, 862–877.
- Lin, W.C., Shuai, B. and Springer, P.S. (2003) The *Arabidopsis LATERAL ORGAN BOUNDARIES*-domain gene *ASYMMETRIC LEAVES2* functions in the repression of *KNOX* gene expression and in adaxial-abaxial patterning. *Plant Cell* **15**, 2241–2252.
- Liu, H., Hong, Y., Lu, Q., Li, H., Gu, J., Ren, L., Deng, L. et al. (2020a) Integrated analysis of comparative lipidomics and proteomics reveals the dynamic changes of lipid molecular species in high-oleic acid peanut seed. *J. Agric. Food Chem.* **68**, 426–438.
- Liu, H., Hu, D., Du, P., Wang, L., Liang, X., Li, H., Lu, Q. et al. (2021) Single-cell RNA-seq describes the transcriptome landscape and identifies critical transcription factors in the leaf blade of the allotetraploid peanut (*Arachis hypogaea* L.). *Plant Biotechnol. J.* **19**, 2261–2276.
- Liu, H., Liang, X., Lu, Q., Li, H., Liu, H., Li, S., Varshney, R. et al. (2020b) 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.

- Liu, X., Li, Y. and Zhong, S. (2017) Interplay between light and plant hormones in the control of *Arabidopsis* seedling chlorophyll biosynthesis. *Front. Plant Sci.* **8**, 433.
- Liu, Z., Wang, J., Zhou, Y., Zhang, Y., Qin, A., Yu, X., Zhao, Z. et al. (2022) Identification of novel regulators required for early development of vein pattern in the cotyledons by single-cell RNA-sequencing. *Plant J.* **110**, 7–22.
- Liu, Z., Zhou, Y., Guo, J., Li, J., Tian, Z., Zhu, Z., Wang, J. et al. (2020c) Global dynamic molecular profiling of stomatal lineage cell development by single-cell RNA sequencing. *Mol. Plant* **13**, 1178–1193.
- Lopez-Anido, C.B., Vatén, A., Smoot, N.K., Sharma, N., Guo, V., Gong, Y., Anleu Gil, M.X. et al. (2021) Single-cell resolution of lineage trajectories in the *Arabidopsis* stomatal lineage and developing leaf. *Dev. Cell* **56**, 1043–1055.
- Ma, Z., Hu, X., Cai, W., Huang, W., Zhou, X., Luo, Q., Yang, H. et al. (2014) *Arabidopsis* miR171-targeted scarecrow-like proteins bind to GT cis-elements and mediate gibberellin-regulated chlorophyll biosynthesis under light conditions. *PLoS Genet.* **10**, e1004519.
- Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M. and Tirosh, I. (2015) Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214.
- Mazzella, M., Casal, J., Muschietti, J. and Fox, A. (2014) Hormonal networks involved in apical hook development in darkness and their response to light. *Front. Plant Sci.* **5**, 52.
- Michalina, S., Tomasz, S., Michał, M., Krzysztof, L., Ross, J., Gwyneth, I. and Przemysław, W. (2014) Gamma-secretase subunits associate in intracellular membrane compartments in *Arabidopsis thaliana*. *J. Exp. Bot.* **65**, 3015–3027.
- Mu, X.R., Tong, C., Fang, X.T., Bao, Q.X., Xue, L.N., Meng, W.Y., Liu, C.Y. et al. (2022) Feedback loop promotes sucrose accumulation in cotyledons to facilitate sugar-ethylene signaling-mediated, etiolated-seedling greening. *Cell Rep.* **38**, 110529.
- Nicolas, D., Boris, S. and Laurent, D.N. (2014) Emerging functions of Nodulin-Like proteins in non-nodulating plant species. *Plant Cell Physiol.* **55**, 469–474.
- 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.
- Serrano-Ron, L., Perez-Garcia, P., Sanchez-Corriñero, A., Gude, I., Cabrera, J., Ip, P.L., Birnbaum, K.D. et al. (2021) Reconstruction of lateral root formation through single-cell RNA sequencing reveals order of tissue initiation. *Mol. Plant* **14**, 1362–1378.
- Shaw, R., Tian, X. and Xu, J. (2021) Single-cell transcriptome analysis in plants: advances and challenges. *Mol. Plant* **14**, 115–126.
- Shi, H., Lyu, M., Luo, Y., Liu, S., Li, Y., He, H., Wei, N. et al. (2018) Genome-wide regulation of light-controlled seedling morphogenesis by three families of transcription factors. *P. Natl. Acad. Sci. USA* **115**, 6482–6487.
- Sinclair, S.A., Larue, C., Bonk, L., Khan, A., Castillo-Michel, H., Stein, R.J., Grolimund, D. et al. (2017) Etiolated seedling development requires repression of photomorphogenesis by a small cell-wall-derived dark signal. *Curr. Biol.* **27**, 3403–3418.
- Singh, S., Grover, A. and Nasim, M. (2016) Biofuel potential of plants transformed genetically with NAC family genes. *Front. Plant Sci.* **7**, 22.
- Sinha, P., Bajaj, P., Pazhamala, L.T., Nayak, S.N., Pandey, M.K., Chitikineni, A., Huai, D. et al. (2020) *Arachis hypogaea* gene expression atlas for fastigiat subspecies of cultivated groundnut to accelerate functional and translational genomics applications. *Plant Biotechnol. J.* **18**, 2187–2200.
- Svensson, V. and Pachter, L. (2018) RNA velocity: Molecular kinetics from single-cell RNA-seq. *Mol. Cell* **72**, 7–9.
- Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S. (2013) MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* **30**, 2725–2729.
- Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N.J. et al. (2014) The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386.
- Wan, Y.L., Eisinger, W., Ehrhardt, D., Kubitscheck, U., Baluska, F. and Briggs, W. (2008) The subcellular localization and blue-light-induced movement of phototropin 1-GFP in etiolated seedlings of *Arabidopsis thaliana*. *Mol. Plant* **1**, 103–117.
- Wang, H., Kong, F. and Zhou, C. (2021a) From genes to networks: The genetic control of leaf development. *J. Integr. Plant Biol.* **63**, 1181–1196.
- Wang, Q., Wu, Y., Peng, A., Cui, J., Zhao, M., Pan, Y., Zhang, M. et al. (2022) Single-cell transcriptome atlas reveals developmental trajectories and a novel metabolic pathway of catechin esters in tea leaves. *Plant Biotechnol. J.* **20**, 2089–2106.
- Wang, Y., Huan, Q., Li, K. and Qian, W. (2021b) Single-cell transcriptome atlas of the leaf and root of rice seedlings. *J. Genet. Genomics* **48**, 881–898.
- Wit, M.D., Galvão, V.C. and Fankhauser, C. (2016) Light-mediated hormonal regulation of plant growth and development. *Annu. Rev. Plant Biol.* **67**, 513–537.
- Wong, M.M., Bhaskara, G.B., Wen, T.N., Lin, W.D., Nguyen, T.T., Chong, G.L. and Verslues, P.E. (2019) Phosphoproteomics of *Arabidopsis* highly ABA-Induced1 identifies AT-Hook-Like10 phosphorylation required for stress growth regulation. *P. Natl. Acad. Sci. USA* **116**, 2354–2363.
- Wu, Q., Kuang, K., Lyu, M., Zhao, Y., Li, Y., Li, J., Pan, Y. et al. (2020) Allosteric deactivation of PIFs and EIN3 by microproteins in light control of plant development. *P. Natl. Acad. Sci. USA* **117**, 18858–18868.
- Zhang, X., Henrique, R., Lin, S.S., Niu, Q.W. and Chua, N.H. (2006) Agrobacterium-mediated transformation of *Arabidopsis thaliana* using the floral dip method. *Nat. Protoc.* **1**, 641–646.
- Zhang, Y., Sun, J., Xia, H., Zhao, C., Hou, L., Wang, B., Li, A. et al. (2018) Characterization of peanut phytochromes and their possible regulating roles in early peanut pod development. *PLoS One* **13**, e0198041.
- Zhao, J., Favero David, S., Peng, H. and Neff Michael, M. (2013) *Arabidopsis thaliana AHL* family modulates hypocotyl growth redundantly by interacting with each other via the PPC/DUF296 domain. *P. Natl. Acad. Sci. USA* **110**, 4688–4697.
- Zhuang, W., Chen, H., Yang, M., Wang, J., Pandey, M.K., Zhang, C., Chang, W.C. et al. (2019) The genome of cultivated peanut provides insight into legume karyotypes, polyploid evolution and crop domestication. *Nat. Genet.* **51**, 865–876.
- Zywitzka, V., Misios, A., Bunyatyan, L., Willnow, T.E. and Rajewsky, N. (2018) Single-cell transcriptomics characterizes cell types in the subventricular zone and uncovers molecular defects impairing adult neurogenesis. *Cell Rep.* **25**, 2457–2469.

## Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**Figure S1** Phenotypic investigation of etiolated (grown under dark) and green (grown under light) seedlings on the 3<sup>rd</sup>, 5<sup>th</sup> and 7<sup>th</sup> day.

**Figure S2** The protoplast isolation of etiolated (a) and green (b) peanut seedling leaves.

**Figure S3** The quality control of scRNA-seq in peanut seedling leaves.

**Figure S4** UMAP plot illustrated leaf cell clusters of etiolated and green seedlings.

**Figure S5** Circos plot represented genome wide expression of DEGs in peanut etiolated (a) and green (b) seedlings leaf cells by scRNA-seq.

**Figure S6** A qPCR examination for the marker genes in distinct cell populations.

**Figure S7** (a) Dot plots and (b) t-SNE plots for the top marker genes for cell clusters.

**Figure S8** The heatmap for core-TFs involved in developmental trajectory, cell differentiation state and cell fate.

**Figure S9** Bulk RNA-seq of one-week-old etiolated and green seedlings of peanut.

**Figure S10** Cell cycle analysis and RNA velocity dynamics using tSNE plot.

**Figure S11** Expression patterns of AHL17 in scRNA-Seq.

**Figure S12** Phylogenetic tree analysis of 64 AHL family genes in peanut.

**Figure S13** Expression and differential analysis of *AHL17* in different tissues and cells.

**Figure S14** Genome-wide identification of *AHL17* binding region by DAP-seq.

**Figure S15 and S16** Scanning electron microscope (SEM) to study the size of epidermal cell in WT, pbWA(V)HS, and *AHL17-OE* transgenic lines at germinating, and flowering stage.

**Figure S17** A qPCR examination for the key genes of phytohormones in WT, pbWA(V)HS, and *AHL17-OE* lines.

**Table S1** Statistics of preliminary mapping results of scRNA-seq data.

**Table S2** Statistical computation of cell clusters.

**Table S3** Total identified gene in peanut leaf.

**Table S4** The reported marker gene for cell clusters.

**Table S5** Top 50 novel marker genes in different cell clusters.

**Table S6** The number of DEGs in each cell cluster.

**Table S7** The core TFs in each cell cluster.

**Table S8** The DEGs of each cell cluster between dark and light samples.

**Table S9** The core TFs of each cell cluster between dark and light samples.

**Table S10** Statistical table of the distribution information of subpopulations in cell trajectories.

**Table S11** Trajectory analysis identified the DEGs along with the timeline.

**Table S12** Trajectory analysis identified the TFs along with the timeline.

**Table S13** Trajectory identified five states of all cell differentiation.

**Table S14** Identification of important TFs in all cell differentiation states.

**Table S15** Identification of DEGs in differ cell differentiation branches.

**Table S16** Identification of important TFs in different cell differentiation branches.

**Table S17** The core TFs in developmental trajectory, cell differentiation state and cell fate.

**Table S18** The core DEGs consisted in scRNA-seq and bulk RNA-seq profiles.

**Table S19** The DEGs in bulk RNA-seq.

**Table S20** Cell cycles of each cell.

**Table S21** The up-regulated expression of genes in each cell cycle.

**Table S22** Trajectory analysis identified the DEGs of mesophyll cells along with the timeline.

**Table S23** Trajectory identified three state of mesophyll cell differentiation.

**Table S24** Trajectory analysis identified the DEGs of mesophyll cells in differ cell differentiation branches.

**Table S25** Trajectory analysis identified the DEGs of epidermal cells along with the timeline.

**Table S26** Trajectory identified three states of epidermal cell differentiation.

**Table S27** Trajectory analysis identified the TFs of epidermal cells along with the timeline.

**Table S28** The phytohormone gene sets of peanut leaves.

**Table S29** Primers were used in this study.

**Table S30** Peak number statistics table.

**Table S31** The annotated table of *AHL17* peak-related genes.

**Table S32** The annotated table of *AHL17-halo* peak-related genes.

**Table S33** The cellular barcodes in dark and light samples.