



# Single-cell transcriptome atlas reveals somatic cell embryogenic differentiation features during regeneration

Huihui Guo ,<sup>†</sup> Li Zhang ,<sup>†</sup> Haixia Guo ,<sup>†</sup> Xiwang Cui , Yupeng Fan , Tongtong Li , Xiushan Qi , Tongdi Yan , Aiyun Chen , Fengjuan Shi , Fanchang Zeng \*  
Tongdi Yan , Aiyun Chen , Fengjuan Shi , Fanchang Zeng \*

State Key Laboratory of Crop Biology, Shandong Agricultural University, Tai'an 271018, China

\*Author for correspondence: fczeng@sdau.edu.cn

<sup>†</sup>Equal contribution to this work.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (<https://academic.oup.com/plphys/pages/General-Instructions>) is Fanchang Zeng (fczeng@sdau.edu.cn).

## Abstract

Understanding somatic cell totipotency remains a challenge facing scientific inquiry today. Plants display remarkable cell totipotency expression, illustrated by single-cell differentiation during somatic embryogenesis (SE) for plant regeneration. Determining cell identity and exploring gene regulation in such complex heterogeneous somatic cell differentiation have been major challenges. Here, we performed high-throughput single-cell sequencing assays to define the precise cellular landscape and revealed the modulation mode of marker genes during embryogenic differentiation in cotton (*Gossypium hirsutum* L.) as the crop for biotechnology application. We demonstrated that nonembryogenic calli (NEC) and primary embryogenic calli (PEC) tissues were composed of heterogeneous cells that could be partitioned into four broad populations with six distinct cell clusters. Enriched cell clusters and cell states were identified in NEC and PEC samples, respectively. Moreover, a broad repertoire of new cluster-specific genes and associated expression modules were identified. The energy metabolism, signal transduction, environmental adaptation, membrane transport pathways, and a series of transcription factors were preferentially enriched in cell embryogenic totipotency expression. Notably, the SE-ASSOCIATED LIPID TRANSFER PROTEIN (SELTP) gene dose-dependently marked cell types with distinct embryogenic states and exhibited a parabolic curve pattern along the somatic cell embryogenic differentiation trajectory, suggesting that SELTP could serve as a favorable quantitative cellular marker for detecting embryogenic expression at the single-cell level. In addition, RNA velocity and Scissor analysis confirmed the pseudo-temporal model and validated the accuracy of the scRNA-seq data, respectively. This work provides valuable marker-genes resources and defines precise cellular taxonomy and trajectory atlases for somatic cell embryogenic differentiation in plant regeneration.

## Introduction

At the forefront of the most compelling scientific puzzles, somatic cell totipotency is one of the challenging problems in biology. Plants display a remarkable capacity for somatic cell totipotency, as demonstrated by somatic-to-embryogenic cell transition (Vogel 2005; Roeder et al. 2021; Xu et al. 2021). As a notable illustration of totipotency, single somatic cell

differentiation during plant somatic embryogenesis (SE) provides an ideal system for the investigation of embryogenic differentiation.

In the process of plant SE for regeneration, some somatic cells start to divide, becoming totipotent, and then enter the new pathway (Fehér et al. 2002). From the somatic to embryogenic states of development, crucial processes called the transition and induction of embryogenic competence

take place. This step is the most important, but at the same time it is less understood (Verdeil et al. 2001). In the early stage of cell embryogenesis, it is difficult to trace and identify cells capable of embryogenesis and regeneration (Guo et al. 2019a; Ogura et al. 2023). Therefore, it is critical to dissect the specific cellular markers associated with the acquisition of embryogenic competence in such highly refined systems.

Changes in cell fate and the direction of differentiation rely on the erasing of the genetic developmental program and switching on of a distinct one (Kurczynska et al. 2012). The key regulatory genes of SE for regeneration have been identified. Previous reports showed that AUXIN RESPONSE FACTOR (ARF) (Su et al. 2016), LEAFY COTYLEDON (LEC) (Rupps et al. 2016), WUSCHEL (WUS) (Elhiti et al. 2010; Li et al. 2019), SOMATIC EMBRYOGENESIS RECEPTOR KINASE (SERK) (Hu et al. 2005), BABY BOOM (BBM) (Rupps et al. 2016), WUSCHEL-RELAYED HOMEBOX (WOX) (Rupps et al. 2016; Zhai and Xu 2021; Ogura et al. 2023), and ARABIDOPSIS RESPONSE REGULATOR (ARR) (Zhai and Xu 2021) participated and played a decisive role in SE for plant regeneration. Genes related to stress, such as SERK1, ABSCISIC ALDEHYDE SYNTHESIS ENZYME 2 (ABA2), ABSCISIC ACID INSENSITIVE 3 (ABI3), JASMONATE ZIM-DOMAIN 1 (JAZ1), and LATE EMBRYOGENESIS ABUNDANT PROTEIN 1 (LEA1) were also demonstrated to regulate SE for plant regeneration (Jin et al. 2014).

Given the complex nature of somatic cell differentiation, it remains difficult to distinguish the various cell types during embryogenic competence acquisition and plant regeneration initiation. Previous bulk RNA-sequencing techniques necessarily underestimated and restricted the identification of complexity and diversity of the cell types in plant somatic cell embryogenic differentiation. Despite extensive studies, a comprehensive analysis of the cell identities in regeneration initiation is lacking, limiting our ability to characterize and delineate the somatic cell differentiation trajectories prospectively. Recent advances in single-cell RNA-sequencing (scRNA-seq) technology are revolutionizing molecular studies of heterogeneous tissues, enabling the systematic identification of cell-type populations and revealing the cell differentiation trajectory underlying key developmental processes (Islam et al. 2014; Macosko et al. 2015; Efroni et al. 2016; Ziegenhain et al. 2017; Shulse et al. 2019; Zhang et al. 2019; Roszak et al. 2021). Unlike bulk RNA sequencing, scRNA-seq achieves a higher degree of resolution to characterize and elucidate the many properties of subpopulations of heterogeneous groups of cells in plants (Birnbaum 2018; Mironova and Xu 2019; Rich-Griffin et al. 2020; Zhang et al. 2021). In cotton (*Gossypium hirsutum* L.), the constructions of single-cell expression profiles of pigment glands (Long et al. 2023; Sun et al. 2023), anthers (Li et al. 2023), fiber (Qin et al. 2022), and leaves (Lin et al. 2023) provided a theoretical framework to explore the molecular mechanisms of cell developmental processes.

Here, we performed high-throughput single-cell transcriptomics and functional analyses based on reverse genetics

during embryogenic competence acquisition in cotton. It is the crop for plant biotechnology in cell/genetic engineering, which commonly achieves plant regeneration through the SE pathway. The profiling and ordering of individual cells enable us to define cellular taxonomy and reconstruct continuous cell differentiation trajectories with high accuracy. Remarkably, SE-ASSOCIATED LIPID TRANSFER PROTEIN (SELTP) dose-dependently marks cell types with distinct embryogenic states, with a remarkable peaky curve expression feature along somatic cell embryogenic differentiation trajectory. Moreover, the identification of cell cluster-specific marker genes provides a valuable resource for future functional analysis of plant somatic cell embryogenic differentiation in the context of plant regeneration.

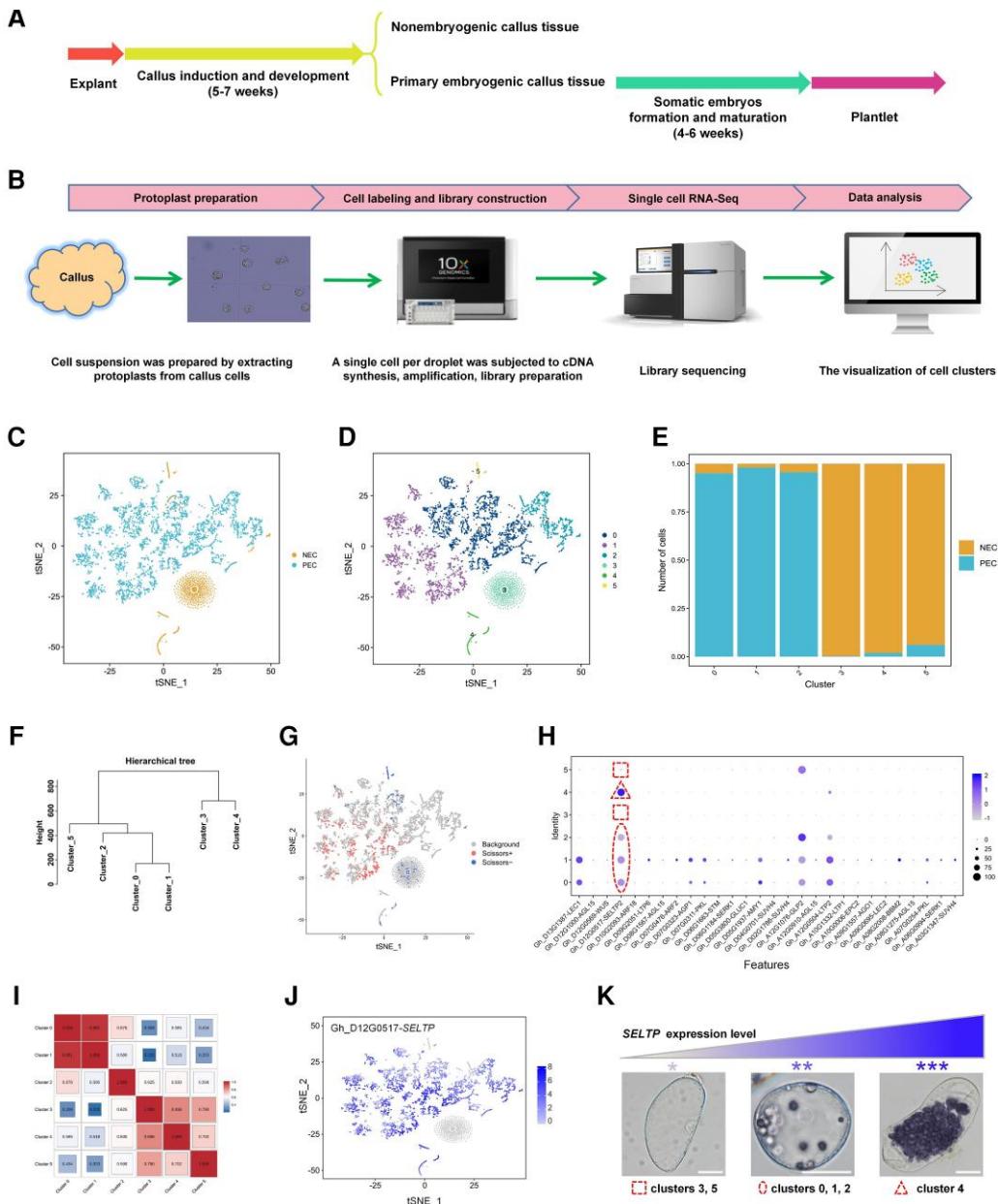
## Results

### Discovery of cell clusters in callus with distinct embryogenic competence

We have established an efficient SE regeneration system for collecting and identifying cultures of different developmental stages (Fig. 1A), as described in our previous series work (Guo et al. 2019a, 2019b). To distinguish cell types in the highly heterogeneous cell populations of calli tissues cultured in the complex and coordinated totipotency fate transition process, a high-throughput scRNA-seq assay was performed on *G. hirsutum* YZ-1, a variety with a high regeneration ability (Jin et al. 2006). Using the established efficient regeneration initiation system, we obtained two typical tissues in distinct states: nonembryogenic calli (NEC) and primary embryogenic calli (PEC) cultured for 56 d (Fig. 1A). The NEC and PEC tissues were protoplasted separately and subjected to droplet-based scRNA-seq using the 10X Genomics scRNA-seq platform (Fig. 1B).

The *t*-distributed stochastic neighborhood embedding (*t*-SNE) tool was used to visualize and explore the datasets for cell clustering relying on known markers (Supplementary Table S1). The results revealed six distinct clusters in the cell populations of the NEC and PEC tissues (Fig. 1, C and D). Notably, clusters 0, 1, and 2 marked PEC tissue, while NEC tissue was marked by cells in clusters 3, 4, and 5 (Fig. 1E). Furthermore, clusters 3 and 4 were distinguished from clusters 0, 1, 2, and 5 by expression hierarchical clustering (Fig. 1F). In order to confirm the scRNA-seq data, we further applied Scissor to combine scRNA-seq data and bulk RNA-seq data of NEC and PEC samples (Sun et al. 2022). In the single-cell dataset, Scissor would select cells that were most highly associated with NEC or PEC of bulk RNA-seq (Fig. 1G). The result indicated that all of 379 Scissor+ cells (100%) and 79 of 94 Scissor- cells (84%) were consistent with the single cell in PEC and NEC samples, respectively (Fig. 1, C and G), providing an important guarantee of the reliability of our scRNA-seq data.

In each cell cluster, the distribution of known marker genes linked with specific cell types was analyzed to annotate the



**Figure 1.** Cell cluster analysis of single-cell transcriptomes from complex calli during embryogenic differentiation initiation. **A)** The schematic illustration of somatic embryogenesis (SE) regeneration system with representative developmental stages. Different colors represent different stages of development. **B)** Workflow used for scRNA-Seq to illustrate transcriptomes from individual callus cells. **C, D)** t-SNE visualization of nonembryogenic calli (NEC) and primary embryogenic calli (PEC) tissues (**C**) as well as the six cell clusters (**D**). Each dot denotes a single cell. Colors denote corresponding tissues/clusters. **E)** Cell numbers of each cluster in NEC and PEC tissues. **F)** Hierarchical tree of cell clusters. The closer the relationship is, the easier it is to classify the cluster as a clade. **G)** The t-SNE visualization of the Scissor selected cells. The red and blue dots represent Scissor+ and Scissor- cells, associated with PEC and NEC samples, respectively. **H)** Expression pattern of representative cell-type marker genes. The dot plots indicate the fraction of cells in each cluster (dot size) expressing a given marker and the level of marker gene expression (dot intensity) for 28 genes known to exhibit preferential expression in distinct tissue/cell types. Three expression patterns with distinct levels of *SELTP* expression were identified in the six cell clusters, which are indicated with different dotted outlines (high level, triangle; intermediate level, elliptical; low level, square). The bubble color represents the average expression abundance of marker genes in the cell subgroup, and the darker the bubble color, the higher the average expression level of marker genes in that subgroup. **I)** Heatmap of correlations among cell clusters. The darker the red color, the higher the correlation, while the darker the blue color, the lower the correlation. **J)** t-SNE visualization of expression profiles of the *SELTP* gene for different cell types. Color intensity indicates the relative transcript level for the *SELTP* gene in each cell. **K)** *SELTP* marked three cell types with distinct embryogenic competence in a dose-dependent manner. *SELTP* exhibited three remarkable corresponding expression patterns in three types of cells: highest in cluster 4, intermediate in clusters 0, 1, and 2, and lowest in clusters 3 and 5 (**H**). The dose effect of *SELTP* was putatively correlated with the appropriate amyloplast accumulation pattern. Bars = 20  $\mu$ m. The darker the blue color, the higher the level of expression.

cluster and determine the cell type identity of the cluster (Fig. 1H). Cell clusters 0 and 1 showed high association, i.e. similar gene expression patterns, and were identified as the same cell type (Fig. 1I). These two clusters predominantly expressed genes involved in embryogenesis, such as *LEAFY COTYLEDON1* (*LEC1*) (Kwong et al. 2003; Min et al. 2015), *ARABINOGALACTAN-PROTEIN1* (*AGP1*) (Lucau-Danila et al. 2010; Poon et al. 2012), *ALPHA-AMYLASE1* (*AMY1*) (Guo et al. 2019a), and *BBM2* (Boutilier et al. 2002; Horstman et al. 2017), and they were annotated as embryogenic cells (ECs). We annotated cluster 2 as a proembryogenic cells (Pro-ECs) population because amyloplast-related *SELTP* (Guo et al. 2019a) and *GERMIN-LIKE PROTEIN2* (*GLP2*) (Mahdavi-Darvari et al. 2015; Izuno et al. 2020) were preferentially represented in this cluster. This cell type is described as the crucial cell type for embryogenic competence acquisition during embryogenic transition and plant regeneration initiation (Guo et al. 2019a). Known embryogenic marker genes were absent in cluster 3 and cluster 5, and these cells were identified as dedifferentiated cells (DDCs) in the early dedifferentiation stage (Zeng et al. 2006; Guo et al. 2019a). Cluster 4 cells were annotated as NECs based on the finding that few known embryogenic-related genes were expressed in this cluster, and the *SELTP* gene was expressed with excessive content, corresponding to the low embryogenic differentiation ability. Taken together, these results provide an atlas of the four cell populations in the callus during SE initiation. The two typical callus tissues (NEC and PEC) cultured exhibited distinct specific cell types. scRNA-seq enabled us to identify new cell types in the callus tissue in embryogenic differentiation initiation.

### *SELTP* marks cell types with distinct embryogenic expression in a dose-dependent manner

With special attention to the previously identified embryogenesis-related gene *SELTP* (Guo et al. 2019a), interestingly, the distribution of the *SELTP* transcripts in the six cell clusters showed three distinct expression patterns: expression was highest in cluster 4 (specific to NEC tissue), intermediate in clusters 0, 1, and 2 (specific to PEC tissue), and lowest in clusters 3 and 5 (specific to NEC tissue) (Fig. 1H). The t-SNE visualization of *SELTP* gene expression profiles for different cell types showed the same patterns (Fig. 1J, Supplementary Table S2). Taken together, the results appeared to reflect that *SELTP* regulated embryogenic differentiation initiation in a dose-dependent manner and could serve as a more effective marker refined for the detection of cell types with distinct embryogenic expression (Fig. 1K), enabling an early diagnosis of embryogenic competence.

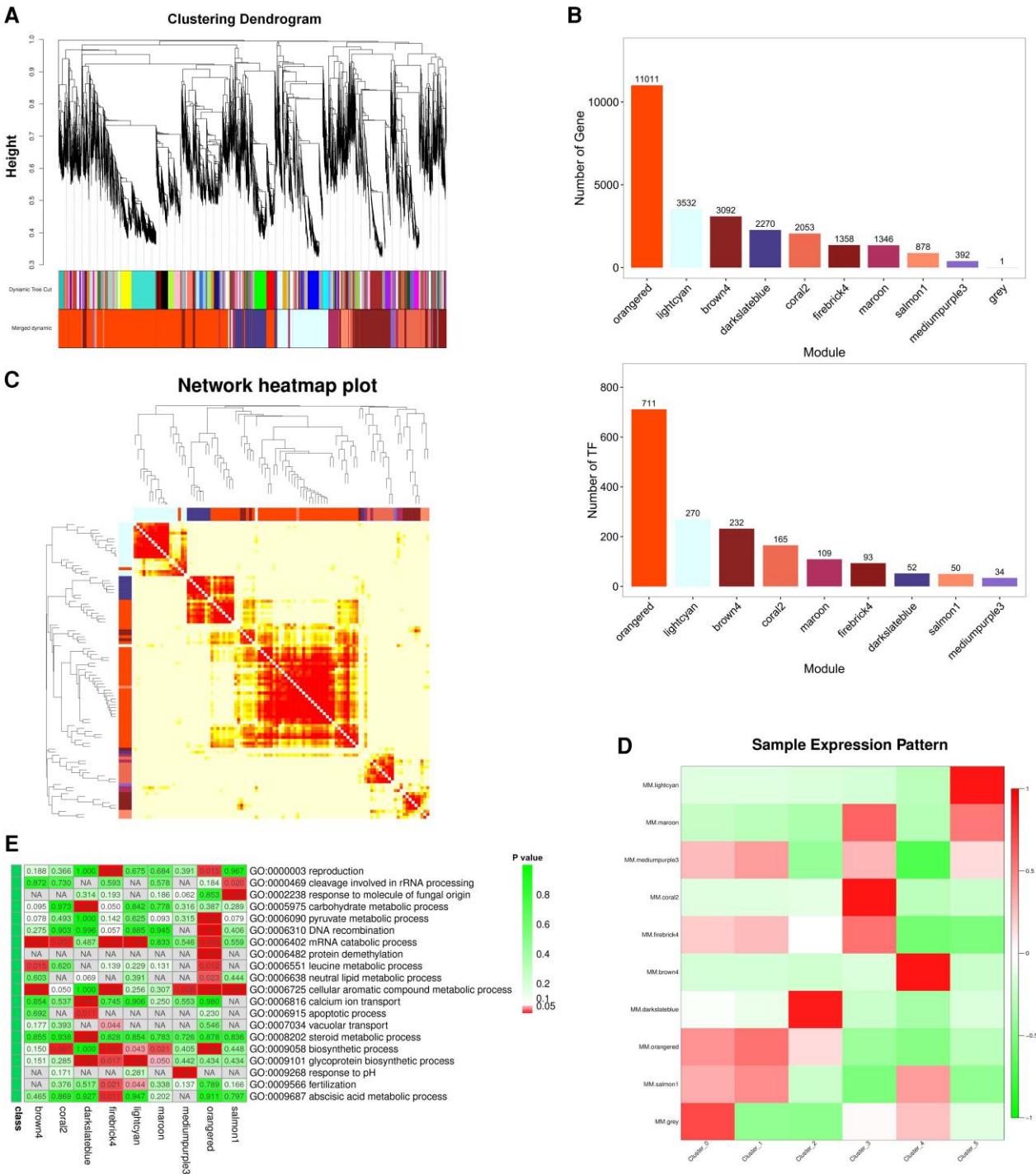
### Identification of cluster-specific marker genes and associated modules involved in controlling plant regeneration initiation

The assignment of clusters by scRNA-seq enabled us to identify a number of cluster-specific or enriched markers for

regeneration activation. We systematically and comprehensively identified cluster-enriched genes and performed functional annotation analysis (Supplementary Table S3). The top five marker genes that were highly expressed in each target cluster are presented in Supplementary Fig. S1A. The *AMY1*, *BBM2*, *CHITINASE 10* (*Cht10*), *FIBROIN HEAVY CHAIN* (*Fib-H*), *REPETITIVE-PROLINE-RICH PROTEIN* (*PRP*), and  $\beta$ -1,3-*GLUCANASE* (*Gluc*) genes were preferentially represented in cell clusters 0 to 5, respectively (Supplementary Fig. S1B). Gene Ontology (GO) functional enrichment analysis revealed that the embryogenic-specific cell clusters (clusters 0, 1, and 2) showed gene signatures corresponding to a series of biological processes that were substantially active relative to those in nonembryogenic-specific cell clusters (clusters 3, 4, and 5). Remarkably, specific genes in cell cluster 2 (proembryogenic cells, the critical cell type for embryogenic differentiation) were dominantly enriched in the GO molecular function of salicylic acid-mediated signaling pathway, cellular response to salicylic acid stimulus, and transmembrane transporter-related activity (Supplementary Fig. S1C).

Then, cluster-specific marker genes with similar patterns of expression were grouped into modules via hierarchical average linkage clustering using the weighted gene coexpression network analysis (WGCNA) package in R (Supplementary Table S4). A total of 10 modules were identified (Fig. 2A). Each module was assigned a unique color identifier, with the poorly connected genes colored gray. The numbers of genes and transcription factors (TFs) in each module are shown in Fig. 2B. The correlation coefficient and interactions of these modules were analyzed (Fig. 2C). In the topological overlap matrix (TOM) plot, increasing color intensity indicates higher connectivity (coexpression similarity) among genes in the network. Genes with the greatest connectivity indexes represent network “hubs” and are located in the center of individual modules (Fig. 2C). To identify modules that were substantially associated with different cell clusters, modules with common expression patterns were subjected to interaction analysis. The relevance between the module and the cell cluster was tested by calculating the gene expression pattern of each module in each cell cluster, and modules with higher expression levels were considered to have more connections with the corresponding cell clusters. The results identified the gray, orange, darkslateblue, coral2, brown4, and lightcyan modules as most substantially associated with cell clusters 0 to 5, respectively (Fig. 2D and Supplementary Fig. S2). Notably, the darkslateblue module specifically correlated to cluster 2 and the orange module commonly correlated to clusters 0, 1, and 2.

Functional enrichment analyses were performed for the module genes. Notably, in the darkslateblue module (closely correlated with cluster 2/proembryogenic cells, the critical cell type for embryogenic differentiation), the carbohydrate metabolic process, calcium ion transport biological process, apoptotic process, steroid metabolic process, and glycoprotein biosynthetic process were identified as substantially represented terms (Fig. 2E).



**Figure 2.** WGCNA of cluster-specific marker genes involved in controlling somatic cell differentiation. **A)** Clustering dendrogram of the cell cluster-specific marker genes. Dissimilarity was determined based on topological overlap, together with assigned module colors. The 10 coexpression modules are shown in different colors. **B)** Histogram of gene and TF numbers in each module identified in (A). The x-axis represents each module, and the y-axis represents the number of genes and TFs. **C)** Heatmap visualization of the gene TOM. Light colors represent low overlap, and progressively darker red coloration represents greater overlap. Blocks of darker colors along the diagonal are modules. The gene dendrogram and module assignments are also shown along the left side and at the top. **D)** Associations of modules and the cell clusters. Each row corresponds to a module; each column corresponds to a cell cluster. Colors ranging from green through white to red indicate low to high correlations. MM, module membership. **E)** Functional GO enrichment analysis of genes in the darkslateblue module (closely correlated with specific cell cluster 2, the critical cell type for embryogenic differentiation). NA, not available.

## Pseudo-time trajectory of somatic cell differentiation during plant regeneration initiation

Because callus tissue contains a mixture of cells in varying developmental stages, we reasoned that our experiment should have captured a representative snapshot of their differentiation. The Monocle program was used to place cells with developmental states in a pseudo-time order along a maturation trajectory based on their gene expression profiles. To resolve the progression of cells during plant somatic cell embryogenic differentiation, we performed a pseudo-time analysis, which revealed a linear ordering of cells that reflected the cluster arrangement. Then, the cell trajectories of seven differentiation states (Fig. 3A), six cell clusters (Fig. 3, B and C), and two callus types (NEC and PEC) (Fig. 3, D and E) were visualized.

The DDCs in cluster 3 were given the earliest developmental stage assignment and were designated as the root of the trajectory. It suggests a gradual transition from the cells in cluster 3 to embryogenic-related cell classes (clusters 0, 1, and 2) or nonembryogenic-related cell classes (clusters 4 and 5). In addition, previously reported key factors governing somatic cell differentiation, including SELTP, LEC1, AGP1, AUXIN RESPONSE FACTOR 2 (ARF2), GLP2, and AMY1, were sequentially expressed along the cell differentiation trajectory (Supplementary Fig. S3A).

Three branches were identified in the cell differentiation trajectory direction: branch 1 (states 1, 2, 3, and 4 and states 1, 2, 3, and 5), branch 2 (states 1, 2, 3, 4, and 5 and states 1, 2, and 6), and branch 3 (states 1, 2, 3, 4, 5, and 6 and states 1 and 7) (Fig. 3A). Notably, cell state 7, involving SELTP expression, was preferentially present in the PEC samples (Fig. 3, A, D, and E and Supplementary Fig. S3A), and cell state 1 was preferentially present in the NEC samples (Fig. 3, A, D, and E). Cell cluster 2 included a similar proportion of cell state 1 and cell state 7 as the transition cell type from the nonembryogenic to embryogenic phase (Fig. 3C).

## Stepwise temporal progression of individual genes for somatic cell embryogenic differentiation

We next sought to visualize the continuous program of gene expression changes that occurs in each cell along regeneration initiation trajectory of somatic cells. To identify a broad repertoire of genes related to the process of somatic embryogenic differentiation, we analyzed the differentially expressed (DE) genes in various types of differentiated cells (Supplementary Table S5). We identified the highly dispersed DE genes that fell into seven distinct gene groups (1 to 7) and reflect successive waves of gene expression across pseudo-time (Fig. 3F). Genes predominantly expressed at the beginning of the developmental trajectory (gene groups 1 and 2) correspond to early cell fate determinants. At the end of the trajectory, the genes overrepresented in gene groups 3 and 4 correspond to late cell fate determinants. Among the subset of genes that showed a dynamic pattern of expression across pseudo-time, the SELTP gene was expressed in the

middle of the trajectory (in gene groups 5 and 7) (Fig. 3F and Supplementary Fig. S3B). Particularly, SELTP exhibited a remarkable expression feature of parabolic curve, which appeared to reflect that SELTP homeostasis could be involved in embryogenic differentiation. Additionally, the branch-dependent DE genes were analyzed and hierarchically clustered to identify the key genes controlling the three branches (Fig. 3G and Supplementary Figs. S3, C and D; Supplementary Table S6).

Our analysis reveals dynamic stepwise temporal progression in somatic cell embryogenic differentiation. The pseudo-time trajectory thus highly precisely reflects the temporal expression changes of individual genes during cell differentiation in plant regeneration initiation, providing a refined view of the changes a cell undergoes during its transition from a somatic cell to a fully differentiated cell.

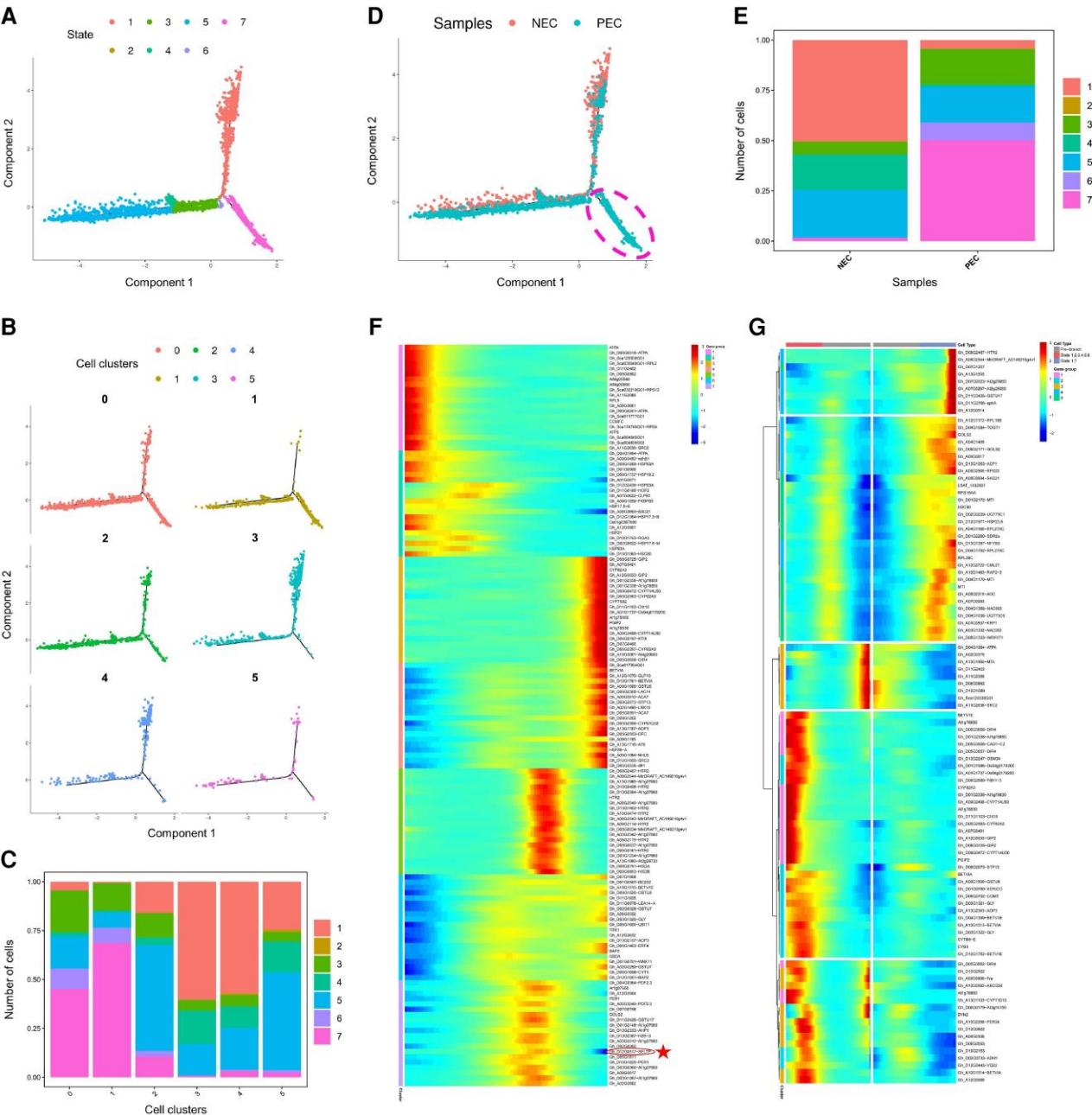
## The cellular pathways of the cluster-marker genes related to somatic cell embryogenic differentiation

In the embryogenesis-related cell clusters (clusters 0, 1, and 2), we further identified the upregulated DE genes and examined their potential functions and cellular pathways according to Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses. In KEGG analysis, we observed that the marker genes of three clusters were enriched in carbohydrate metabolism, energy metabolism, signal transduction, environmental adaptation, and membrane transport pathways (Fig. 4A). Additionally, feature plots of the representative marker genes were analyzed in cell clusters 0, 1, and 2 (Fig. 4B and Supplementary Fig. S1B). Interestingly, genes encoding AMY1 and Cht10 were strongly expressed in clusters 0/1 and 0/1/2, respectively, suggesting that amylase and chitinase activity might be involved in embryogenic differentiation. Meanwhile, BBM2 was specifically expressed in cluster 1.

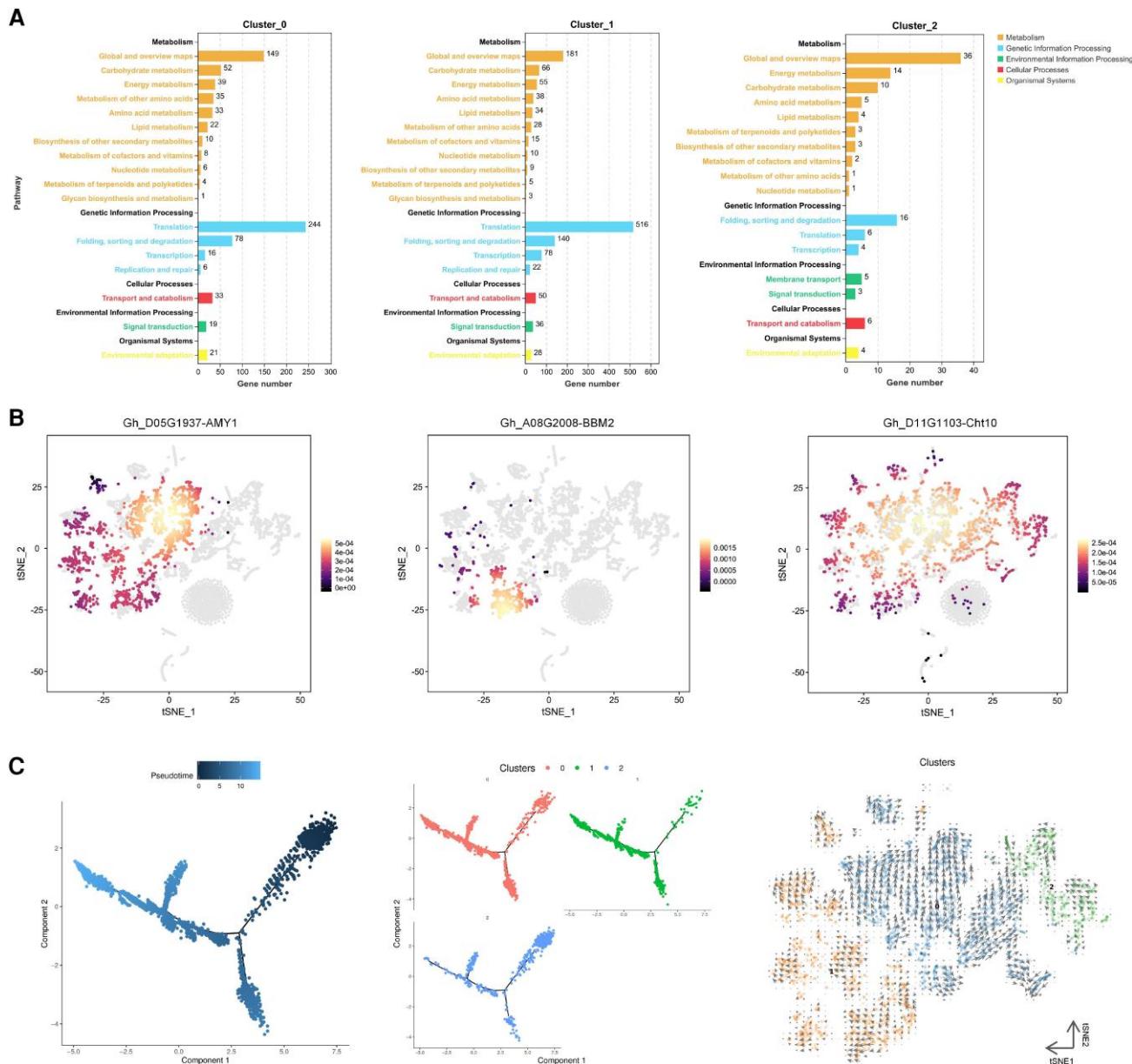
To further test the pseudo-temporal model, we performed both pseudo-time trajectory analysis and RNA velocity analysis for cell clusters 0, 1, and 2. The RNA velocity was visualized by plotting an arrow for each cell linking the actual state to the estimated future state and can be applied as a procedure to verify the cell differentiation direction in the pseudo-time trajectory (La Manno et al. 2018; Zywitzka et al. 2018; Bergen et al. 2020). The results indicated that the cell differentiation directions obtained by pseudo-time trajectory and RNA velocity analysis methods were basically consistent with each other (Fig. 4C).

## Identification of TFs and module-marker genes involved in somatic cell embryogenic differentiation

Previous research has revealed that SE is regulated by a complex network of TFs (Wen et al. 2020). In our study, a set of preferentially expressed TFs, including ERF, WRKY, NAC, bHLH, MYB, C2H2, and bZIP, were identified in the darkslateblue and orange modules, which specifically correlated to cluster 2 and commonly correlated to clusters 0/1/2, respectively (Fig. 5A). These critical SE-associated TFs might



**Figure 3.** Pseudo-time analysis of the cell differentiation trajectory. **A)** Distribution of cells in various differentiation states in cell trajectories. Different colors represent different states of differentiation. **B)** Distribution of cells from the six clusters in cell trajectories. Different colors represent different cell clusters. **C)** Proportions of cells in different differentiation states in the six cell clusters. Different colors represent different cell states. **D)** Distribution of cells from NEC and PEC samples in cell trajectories. Different colors represent different samples. The dashed-line circle indicates the cell state and the differentiation trajectory branch specifically found in PEC. **E)** Proportion of cells with different differentiation states in NEC and PEC samples. Different colors represent different cell states. **F)** Heatmap of DE genes over pseudo-time. Rows represent points in pseudo-time, and columns represent gene expression levels. Columns are grouped based on gene expression similarity, resulting in seven clusters (indicated on the left). Genes in clusters 1 and 2 were expressed early in pseudo-time, and genes in clusters 3 and 4 were expressed late in pseudo-time. Different colors in the heatmap represent the levels of gene expression. The *SELTP* gene targeted in this study is circled and marked with a red star. **G)** Heatmap of branch 3-dependent DE genes over pseudo-time. Rows represent the point in pseudo-time (pseudo-time gradually increases from the middle to both sides), and columns represent gene expression levels. The left and right sides are each one branch. Gene groups represent the gene sets with similar expression patterns. Different colors in the heatmap represent the levels of gene expression.

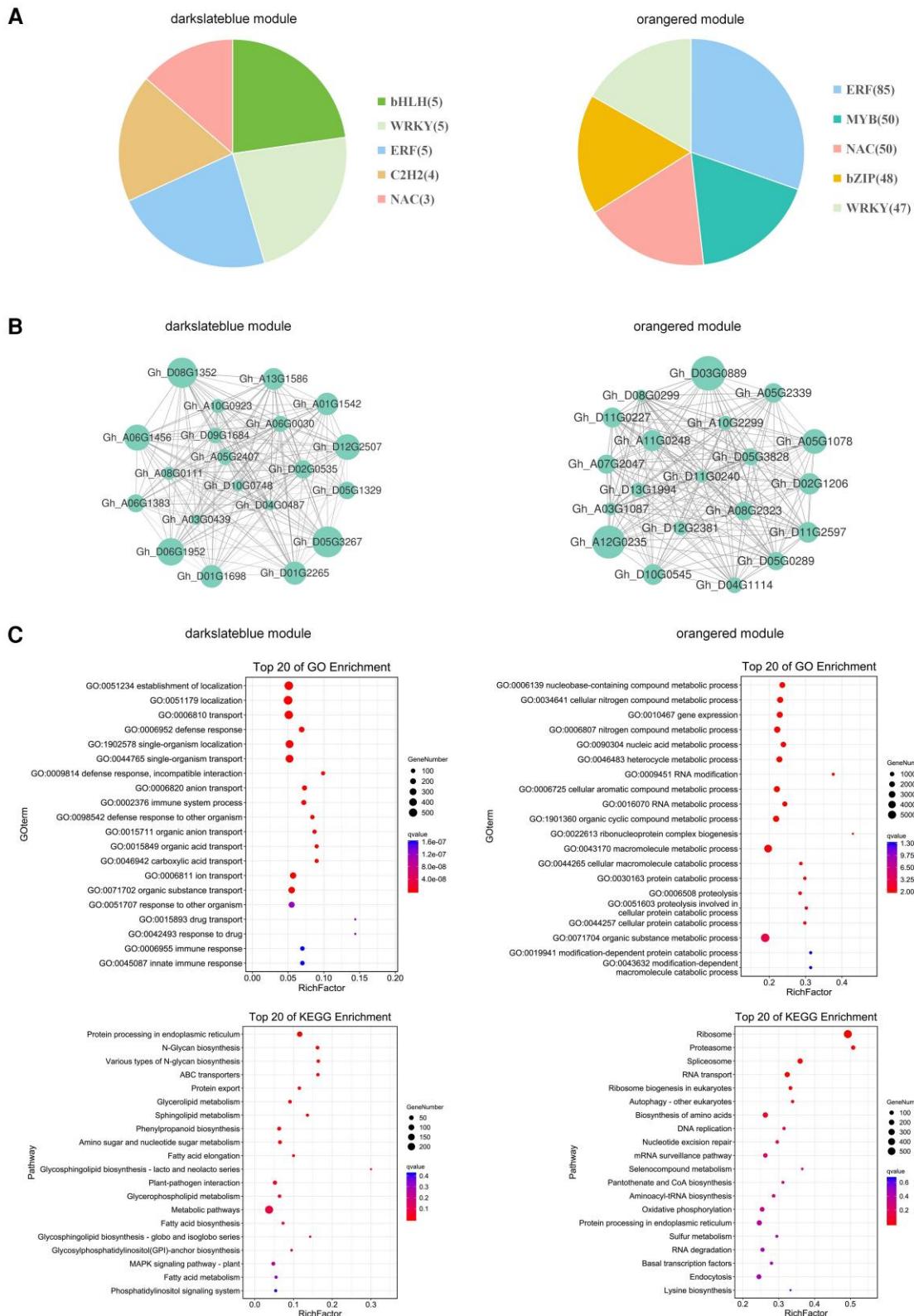


**Figure 4.** The cellular pathways of the cluster-marker genes related to somatic cell embryogenic differentiation. **A)** KEGG enrichment analysis of the DE genes in the cell clusters 0, 1, and 2. Each column represents a pathway, and the height of the column represents the number of genes contained in this pathway. The legend indicates the pathways represented by different column colors. **B)** t-SNE graphs showing the distribution and expression level of representative marker genes in clusters 0, 1, and 2, respectively. Yellow represents high expression levels, and purple represents low expression levels. **C)** Pseudo-time trajectory and RNA velocity analysis of clusters 0, 1, and 2. (Left) Pseudo-time distribution of cell trajectories of clusters 0, 1, and 2. Different points represent different cells, and the darker the point color, the smaller the pseudo-time, and the earlier the development period. (Center) Distribution of cells from the clusters 0, 1, and 2 in cell trajectories. Different colors represent different cell clusters. (Right) RNA velocity plotted in t-SNE space of clusters 0, 1, and 2. The direction of state transitions and the extent of change in RNA dynamics are indicated by the vectors (arrows) and their lengths, respectively. Long arrows indicate that cells have either initiated or terminated the transcription of many different genes, while short arrows or no arrows indicate cells with small changes in transcriptional activity.

act as a link between stress/hormone responding and the developmental pathway in embryogenic differentiation.

Meanwhile, we also identified the top 20 most highly connected hub genes in the darkslateblue and orangered module (*Supplementary Table S7*). These hub genes included SUGAR PHOSPHATE/PHOSPHATE TRANSLOCATOR, RECEPTOR-LIKE

SERINE/THREONINE-PROTEIN KINASE, TRANSMEMBRANE 9 SUPERFAMILY MEMBER 8 (TMN8), CALLOSE SYNTHASE 9 (CALS9) genes and CALCIUM-DEPENDENT PROTEIN KINASE 3 (CPK3) (*Fig. 5B; Supplementary Table S7*), which were associated with the enrichment pathways involved in embryogenic differentiation.



**Figure 5.** Identification of TFs and module-marker genes involving in somatic cell embryogenic differentiation. **A)** Top five TF families in the darkslateblue and orangered module. The number of TF families were shown in the bracket. TF families were represented by different colors. **B)** Network relationships among the top 20 genes with the highest connectivity in the darkslateblue and orangered module. A larger node size indicates a more significant DE gene, and a smaller node size indicates a less significant DE gene. **C)** Top 20 GO terms (biological process) and pathways of darkslateblue and orangered module. The circle size represents enriched gene numbers. Red and blue color represent significant and insignificant enrichments, respectively. *q* Value represents the degree of significant enrichments on GO terms and KEGG pathways.

Additionally, to provide insights into the cellular pathways in the darkslateblue and orangered modules, we performed GO and KEGG analyses (Fig. 5C). The results showed that the enrichment pathways were generally consistent with KEGG analysis of the clusters related to somatic cell embryogenic differentiation (Figs. 4A and 5C).

These results above suggested that the collection of callus cell type-specific or enriched genes can accelerate gene discovery and serve as a valuable resource for studying plant somatic cell embryogenic differentiation and plant regeneration in the future.

### Evident cell division/proliferation features in regeneration initiation

One of the key themes of plant development is differentiation-associated cell division (Sablowski 2007). Generally, stem cells and progenitor cells with the competence for proliferation and regeneration have the ability to divide; that is, they show an active cell cycle. Activity of cell division/proliferation was directly reflected by cell cycle scoring (Supplementary Fig. S4A; Supplementary Table S8). A substantial majority of cells in the PEC tissue presented an active cell cycle (G1 phase, S phase or G2/M phase; Supplementary Fig. S4, B and C), revealing heterogeneity and substantially high division/proliferation activity. In contrast, in NEC tissue, the vast majority of the cells were in the noncycling phase (Supplementary Fig. S4, B and C). In addition, the majority of cells in clusters 0 and 1 were undergoing division, while cells in clusters 2, 3, 4, and 5 were in the noncycling phase (Supplementary Fig. S4, D to F). We further explored the expression patterns of cell cycle stage-specific genes during plant regeneration initiation. The results indicated that ETHYLENE-RESPONSIVE TRANSCRIPTION FACTOR 5 (ERF5), HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN Q (SYNCRIP), HISTONE H2B (HIS2B), and SNAKIN-2 (SN2) were mainly expressed in the G1, S, G2, and M phase, respectively (Supplementary Fig. S4G). The cell cycle analyses indicated that the proliferation status were obviously different between the NEC- and PEC-staged cell populations. We proposed that differences in the regulation of cell cycle transitions could underlie embryogenic cell developmental allochrony.

### SELTP regulates the acquisition of embryogenic competence

We further explored and verified the molecular function of SELTP (Guo et al. 2019a) in embryogenic competence acquisition. In the current study, first, we quantitatively examined the SELTP expression pattern in tissues at different developmental stages during SE for plant regeneration; the examined tissues were hypocotyls (HY), dedifferentiated calli (DDC), proembryogenic masses (PEM), primary embryogenic calli (PEC), globular embryos (GE), and cotyledon embryos (CE) (Fig. 6A). SELTP was expressed predominantly in PEM, which indicated that it was associated with embryogenic differentiation. RNA in situ hybridization (RNA-FISH) also revealed that SELTP was strongly expressed in PEM (Fig. 6B). This

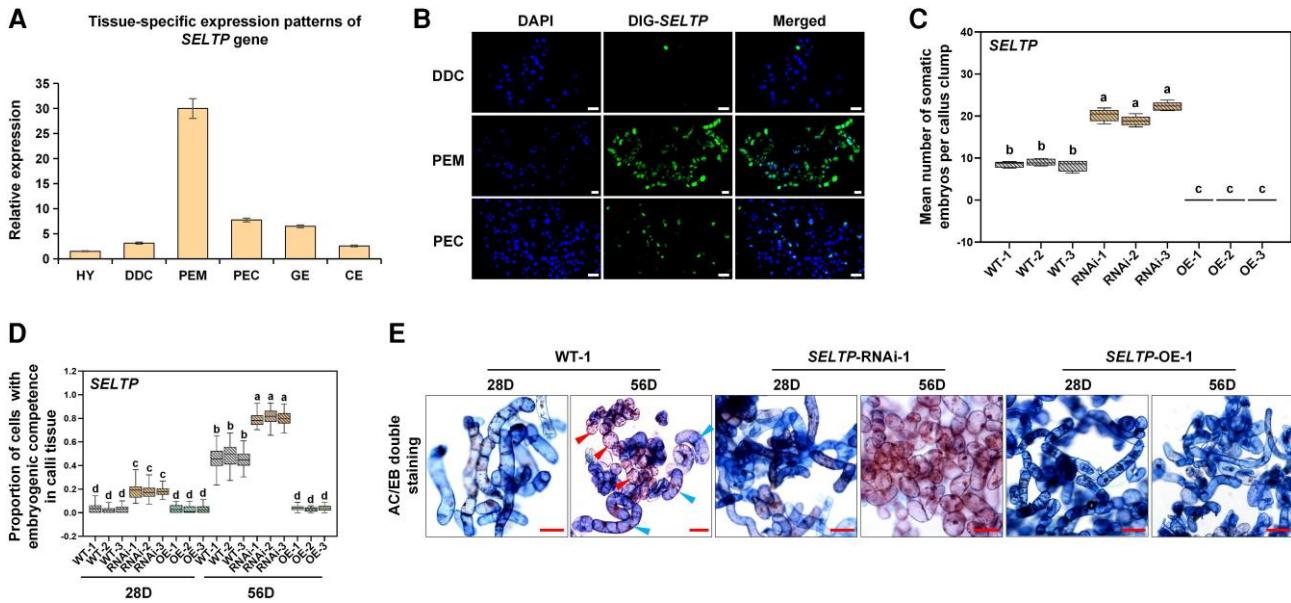
result is consistent with a previous report showing that lipid transfer proteins were markedly activated in PEM (Guo et al. 2019a), suggesting the crucial role of SELTP in embryogenic competence acquisition during plant regeneration initiation.

Next, to examine the biological functional role of SELTP in somatic cell embryogenic competence acquisition, genetic transformation was performed to obtain overexpression (OE) and RNAi knockdown lines. Thereafter, the effects of the OE and gene silencing of SELTP on regulating plant regeneration initiation were investigated. We identified the SE phenotype at the tissue level. The result revealed that the mean number of somatic embryos per callus clump from PEC in the SELTP-RNAi lines were higher than that in the wild type (Fig. 6C). However, the transgenic calli overexpressing SELTP did not undergo embryogenic differentiation (Fig. 6C). To identify the phenotypic characteristics of transgenic callus lines at the cytological level, acetocarmine/Evans blue was used to observe the cell masses with embryogenic competence. The results showed that in the SELTP-RNAi lines, the proportion of cells with embryogenic competence in calli tissue (Fig. 6, D and E) was significantly greater than that in wild type at both the 28D (DDC) and 56D (PEC) stages, with obviously distinct differentiation state. Conversely, we observed significant decreases in cell masses with embryogenic competence in the OE lines (Fig. 6, D and E). Taken together, the results indicated that SELTP silencing promotes embryogenic differentiation. In contrast, SELTP-OE resulted in cell differentiation with loss of embryogenic competence.

## Discussion

### Precise cellular taxonomy and trajectories in complex heterogeneous somatic cell differentiation

scRNA-seq is a powerful tool for dissecting cellular heterogeneity, identifying new cell types, and characterizing the differentiation trajectories of calli tissues cultured in the complex and coordinated totipotency fate transition process (Ryu et al. 2019; Zhang et al. 2019, 2021). Cotton is the crop for plant biotechnology in cell/genetic engineering. In our current work, the scRNA-seq is applied to yield a precise single-cell atlas and refine the coexpression network of the ordered cell atlas in embryogenic differentiation initiation in cotton. We provide a comprehensive cellular landscape in plant somatic cell embryogenic differentiation. In this regard, our scRNA-seq assay reveals that the cultured typical NEC and PEC are composed of heterogeneous populations and together could be partitioned into four broad populations with six transcriptionally distinct cell clusters, reflecting and supporting the common concern on distinguishing cell types in the complex calli tissues cultured. The NEC and PEC tissues contain distinct specific cell clusters. Additionally, a specific branch with a specific cell state in the direction of the embryogenic differentiation trajectory is identified in PEC-staged calli, and the gene group participates in the trajectory of regeneration initiation.



**Figure 6.** *SELTP* functions in somatic cell differentiation during plant regeneration initiation. **A)** Relative transcript levels of *SELTP* in tissues representing different developmental stages. HY, hypocotyls. DDC, dedifferentiated calli cultured for 28 d. PEM, proembryogenic mass cultured for 49 d. PEC, primary embryogenic calli cultured for 56 d. GE, globular embryos. CE, cotyledon embryos. Values are presented with mean  $\pm$  SD ( $n = 3$ ). **B)** RNA in situ hybridization analysis of *SELTP* expression in developing calli at different stages. The immunofluorescence signals of *SELTP* (labeled with digoxin DIG, green) indicate expression locations. Bars = 20  $\mu$ m. DAPI, 2-(4-Aminophenyl)-6-indolecarbamidine dihydrochloride. **C)** Mean number of somatic embryos per PEC clump in *SELTP* transgenic lines. WT, wild type. OE: overexpression. Boxplots display median (line), interquartile range (box), whiskers (extending 1.5 times the interquartile range). Different lowercase letters indicate significant differences between groups (Tukey's test,  $P < 0.05$ ).  $n = 20$ . **D)** Proportion of cells with embryogenic competence in *SELTP* transgenic lines. Boxplots display median (line), interquartile range (box), whiskers (extending 1.5 times the interquartile range). Different lowercase letters indicate significant differences between groups (Tukey's test,  $P < 0.05$ ).  $n = 10$ . **E)** Identification of embryogenic cell masses by acetocarmine/Evans blue double staining in *SELTP* transgenic lines. Embryogenic cell masses are stained red by acetocarmine (marked by red arrow), and nonembryogenic/dedifferentiated cell masses are stained blue by Evans blue (marked by blue arrow). Bars = 50  $\mu$ m.

By sequencing the transcriptome of individual cells, we can now measure the output of the regulatory network without the loss of information caused by the signal averaging in bulk RNA sequencing technologies (Zhang et al. 2021). In our study, scRNA-seq enabled us to identify single-cell transcriptomes representing distinct cell types, including relatively rare cells. Of the identified four cell populations, cluster 2 cells (Pro-ECs) can be considered the transition state toward embryogenesis identity. Meanwhile, pseudo-time analysis demonstrated the successive waves of gene expression occurring in somatic cell totipotency, which reflects that somatic-to-embryogenic differentiation follows finely developmental trajectories, with progression steps beyond the commonly described nonembryogenic and embryogenic state cells. These findings reveal precise gene expression dynamics and further demonstrate the advantage of scRNA-seq to capture and delineate the hidden transition cell states along the differentiation trajectory underlying embryogenic competence acquisition and plant regeneration.

### The identified new marker genes and regulatory pathways related to somatic cell differentiation

Previous studies have revealed single-cell resolution transcriptomic landscapes during plant regeneration (Zhai and

Xu 2021; Ogura et al. 2023; Song et al. 2023; Zhang et al. 2023; Zhu et al. 2023). On the basis of analyses at the single-cell level, the researchers found that the tissue structure of *Arabidopsis* (*Arabidopsis thaliana*) callus on callus-inducing medium is similar to that of the root primordium or root apical meristem, and the middle cell layer with quiescent center-like transcriptional identity exhibits the ability to regenerate organs (Zhai and Xu 2021). In addition, the single-cell transcriptome based on Quartz-Seq2 has shown that *WOX13* plays key roles in determining the cellular identity of callus cell population (Ogura et al. 2023). Other studies revealed the heterogeneity and functional diversity of cells in tomato (*Solanum lycopersicum* L.) callus and analyzed that light promotes shoot regeneration by inducing chlorenchyma cell development and coordinating sugar signaling (Song et al. 2023). Moreover, the key ethylene response factor *ERF6*, which is a negative regulator of SE under high-temperature stress conditions, was screened based on the single-cell transcriptome of longan (*Dimocarpus longan* Lour.) (Zhang et al. 2023). It is worth noting that Zhu et al. studied a high-resolution single-cell transcriptomic landscape of hypocotyl tissue from the highly regenerable cotton genotype Jin668 and the recalcitrant TM-1. The study revealed that hormone response-related genes, including

*LAX2*, *LAX1* and *LOX3*, exhibit different expression patterns in the primary xylem and cambium region of Jin668 and TM-1. Moreover, it showed that *LAX2*, *LAX1* and *LOX3* play important roles in callus proliferation and plant regeneration (Zhu et al. 2023). In our current work, at the single-cell level, we additionally identified seven TFs involved in stress response and hormone signaling, which may be crucial for embryogenic differentiation. Furthermore, we found that the classic embryogenesis-related gene *BBM2* was specifically expressed in cluster 2 (Pro-ECs, the critical cell type for embryogenic differentiation). In addition, the marker genes *AMY1* and *Cht10* were strongly expressed in clusters 0/1 and 0/1/2, respectively, suggesting that amylase and chitinase activity might be involved in embryogenic differentiation. These results were consistent with and extended previous reports. The valuable resources provided by the recent studies indicate more effective and accurate strategies for screening candidate genes for SE research, which will be helpful to further explore the mechanisms of SE and the determination of single-cell fate.

### The *SELTP* homeostasis could be involved in embryogenic competence acquisition

Our results support the hypothesis that *SELTP* marks cell types with distinct embryogenic states in a dose-dependent manner at the single-cell level. The *SELTP* gene could serve as a favorable quantitative cellular marker for the detection of cell types with distinct embryogenic expression. These results extend our previous study (Guo et al. 2019a) and represent cellular insights into the *SELTP* gene as the refined, precise marker in early diagnosis of embryogenic competence.

Meanwhile, *SELTP* was expressed predominantly in the PEM stage before embryogenic calli formation and exhibited an obviously decreased pattern in the PEC stage (Fig. 6A). Meanwhile, the single-cell analysis results (Fig. 1, H and K) also confirmed that the expression of *SELTP* was at a relatively low level in embryogenic calli. Additionally, *SELTP* exhibited a remarkable parabolic curve expression feature along somatic cell embryogenic differentiation trajectory (Supplementary Fig. S3B). The homeostasis of *SELTP* expression level is considered to be involved in embryogenic differentiation. The embryogenic differentiation initiation process requires a certain amount of *SELTP* expression, but excessively low and high levels of *SELTP* could have a negative impact on that.

### scRNA-seq resource for understanding plant cell embryogenic differentiation and cell engineering/biotechnology application

Compared with previous noteworthy reports of scRNA-seq on tissue of *A. thaliana* callus (Zhai and Xu 2021; Ogura et al. 2023), tomato callus (Song et al. 2023), hypocotyl tissue from cotton (Zhu et al. 2023), and embryogenic calli of woody plant species *D. longan* (Zhang et al. 2023), the dataset of single-cell transcriptomes reported here provides a detailed

view of gene expression maps in distinct two differentiation states (NEC and PEC) during embryogenic competence acquisition in cotton. Our comprehensive single-cell atlas offers a spatiotemporal perspective and advances the field of plant somatic cell totipotency expression during embryogenic initiation which was not previously achievable. It provides important insights and information for understanding somatic cell totipotency in plant regeneration.

Additionally, we systematically define cell cluster-specific genes that can be used as markers to identify a cell/cells whose fate changes from a somatic to an embryogenic state, which is critical for cell manipulation during cell engineering. The identified cell clusters, together with the cluster-specific marker genes, pave the way for us to analyze the developmental and physiological functions of the new cell types during the embryogenic differentiation initiation process. More importantly, the expression profiling of cells undergoing somatic differentiation enables us to reconstruct the continuous cell differentiation trajectories of different cultured cell types at the single-cell level.

These findings provide a rich framework and serve as a valuable resource for understanding embryogenic differentiation during plant regeneration at a single-cell resolution, which will facilitate future work aimed at revealing the mechanism of enigmatic somatic cell totipotency expression in plants. Moreover, the nature and knowledge of somatic cell embryogenic differentiation discovered in our current study will guide future strategies for modern cell engineering, with practical value in asexual reproduction and biotechnology breeding applications in crops.

## Materials and methods

### Plant material and growth conditions

The crop for plant biotechnology and cell/genetic engineering application, upland cotton (*G. hirsutum* cultivar YZ-1), was used in this study. It commonly achieves plant regeneration through the SE pathway. We have established an efficient regeneration system for collecting and identifying cultures of different developmental stages, as described in our previous series work (Zeng et al. 2006; Guo et al. 2019a, 2019b). Mature seeds were sterilized in 0.1%  $\text{HgCl}_2$  (w/v) for 8 min and were then rinsed three to four times with distilled water. The seeds were then germinated in Murashige and Skoog (MS) medium supplemented with 3% (w/v) sucrose and 0.25% (w/v) phytigel. Hypocotyl explants (0.5 to 1.0 cm) from 7-d-old seedlings were cultured in MS plus B5 vitamin (MSB) medium containing  $0.45 \mu\text{mol L}^{-1}$  2,4-dichlorophenoxyacetic acid (2,4-D) and  $0.46 \mu\text{mol L}^{-1}$  kinetin (KT). DDC were maintained in MSB medium for 5 to 6 wk at 28 °C in a 16/8 h light/dark photoperiod and were then subcultured in fresh MSB medium without hormones. Following an additional 2 to 4 wk of growth, the somatic-to-embryogenic transition progressed to the point where it induced the development of PEC. Usually, the embryogenic calli are firm and light yellow.

The cells usually formed small and compact clusters, which would be developed further into somatic embryos. NEC, however, are friable, semitransparent, and yellow or gray. These formed relatively larger and loosely held clusters that cannot develop into somatic embryos because they do not have the preembryogenic determined cells responsible for this action.

### scRNA-seq library construction and sequencing

To identify cell clusters in the cultured heterogeneous callus, scRNA-seq was performed in two tissues with distinct differentiation states, NEC and PEC cultured for 56 d. This stage was chosen because the embryogenic differentiation initiation of callus regularly occurs in this specific stage, corresponding to the time window when somatic embryos are formed. Cells were isolated from the NEC and PEC tissue to be tested using a short (1 to 2 h) cell-wall digestion, followed by two filtrations through a 40  $\mu\text{m}$  screen. It is more difficult to prepare protoplasts from NEC tissue, which results from the fragile structure and the large cellular vacuole of nonembryogenic cells. After cell viability identification reached the test standard, cellular suspensions were loaded on a 10X Genomics GemCode single-cell instrument that generates single-cell Gel Bead-In-EMlusion (GEMs). Libraries were generated and sequenced from the cDNAs with Chromium Next GEM Single Cell 3' Reagent Kits v3.1. A single cell per droplet was subjected to cDNA synthesis, amplification, library preparation, read alignment, and expression calling. The cDNA libraries were sequenced on the Illumina sequencing platform NovaSeq 6000 by Genedenovo Biotechnology Co., Ltd (Guangzhou, China). The reference genome file was downloaded from the CottonGen database ([https://www.cottongen.org/species/Gossypium\\_hirsutum/nbi-AD1\\_genome\\_v1.1](https://www.cottongen.org/species/Gossypium_hirsutum/nbi-AD1_genome_v1.1)).

### Cell clustering, annotation and, cluster-specific marker gene identification

The 10X Genomics Cell Ranger software (version 3.1.0) was used for quality control, alignment, and counts quantification. The cells by gene matrices for each sample were individually imported to Seurat version 3.1.1 for downstream analysis. To further filter abnormal cells, cells with unusually high numbers of UMIs ( $>20,000$ ) or unusually high numbers of genes ( $>8,000$ ) were filtered out. After removing unwanted cells from the dataset, data combination and batch effect correction of the two samples were performed using Harmony. After prefILTERing at both cell and gene levels, we successfully captured 1,000 and 4,059 cells, with median numbers of 140 and 2,799 expressed genes per cell for NEC and PEC, respectively (Supplementary Table S10). In total, we detected transcripts of 30,577 and 50,565 genes in our scRNA-seq of NEC and PEC, respectively (Supplementary Table S10). To evaluate the effect of protoplasting genes on cell clustering, the proportion of the protoplasting genes was calculated and plotted. The proportion of protoplasting

genes in each cell cluster was below 3%, which had little effect on the clustering (Supplementary Table S10). Subsequently, the cells were clustered using the Louvain algorithm, and the clustering results were visualized using t-SNE. Clusters were identified using the Seurat function "FindClusters" with "resolution = 0.5". The cell type of each cell cluster was manually defined by known marker genes, which are listed in Supplementary Table S1. Significant upregulated genes that had to be expressed in more than 25% of the cells belonging to the target cluster ( $P\text{-value} \leq 0.01$ ) were identified as cluster-specific marker genes. The Seurat R package was used to assign a cell cycle score to each cell based on the 100 marker genes for the G1/S phase, 113 marker genes for the S phase, 133 marker genes for the G2/M phase, 151 marker genes for the M phase, and 106 marker genes for the M/G1 phase. Cells with the highest score of less than 0.3 were identified as noncycling cells.

### Integration of phenotype-associated bulk expression data and single-cell data by scissor

Scissor integrates phenotype-associated bulk expression data and single-cell data by first quantifying the similarity between each single cell and each bulk sample (Sun et al. 2022). The R package Scissor (v2.1.0) was used to identify the NEC and PEC cell subpopulations. In actual operation, we set the family = "gaussian", cutoff = "0.2" by a Gaussian linear regression model. In order to display more cells, we set the alpha = "0.1".

### Cluster-associated modules identification and construction of cell trajectories

The co-expression networks for the cluster-specific marker genes were constructed using WGCNA (v1.47) package in R. To find out biologically significant modules, module eigen-genes were used to calculate the correlation coefficient with samples. Intramodular connectivity (K.in) and module correlation degree (MM) of each gene were calculated using the R package of WGCNA, and genes with high connectivity tended to be hub genes. The networks were visualized using Cytoscape\_3.3.0. Cell trajectory was analyzed using matrix of cells and gene expressions by Monocle (Version2.6.4). We identified key DE genes related to the differentiation process with  $\text{FDR} < 1e^{-5}$  and grouped genes with similar trends in expression. Monocle develops BEAM to test for branch-dependent gene expression. Enrichment analyses of GO terms and KEGG pathways of DE genes were performed using R based on the hypergeometric distribution.

### RNA velocity analysis

Loom files were generated for each single cell using velocito (v0.17.13) (La Manno et al. 2018) with options -c and -U to indicate that each BAM represents an independent cell and reads are counted instead of molecules (UMIs), respectively. The individual loom files were subsequently merged using the combine function from the loompy Python module. AnnData object was created from the h5ad file using

the scvelo python module (Bergen et al. 2020) for RNA velocity analysis. Highly variable genes were identified, and the corresponding spliced and unspliced RNA counts were normalized and log2-transformed using the scvelo.pp.filter\_and\_normalize function. Next, the first- and second-order moments were computed for velocity estimation using the scvelo.pp.moments function. The velocities (directionalities) were computed based on the stochastic model as defined in the scvelo.t1.velocity function, and the velocities were subsequently projected on the t-SNE embeddings generated from Seurat above.

### Tissue expression pattern analysis of SELTP

To examine the expression pattern of SELTP during the SE process, reverse transcription quantitative PCR (RT-qPCR) was performed to detect the expression levels of SELTP in different developmental staged tissues, including hypocotyls (HY), DDC cultured for 28 d, PEM cultured for 49 d, PEC cultured for 56 d, GE, and CE. RNA was isolated using an RNAPrep pure Plant Kit. cDNA was synthesized using a One-Step gDNA Removal and cDNA Synthesis SuperMix. The RT-qPCR was performed with Ultra SYBR Mixture (Low ROX, TransGen Biotech, Beijing, China) using an Applied Biosystems 7500 Real-Time PCR System (Thermo Fisher, USA). GhUB7 gene (Cottongen: Gh\_A11G0969.1) was used as a reference to normalize the target gene expression values. Three biological replicates were performed independently. The related primers are listed in Supplementary Table S11.

### RNA fluorescence in situ hybridization (RNA-FISH) assay

In order to detect the tissue expression pattern of the SELTP gene during embryogenic differentiation, SELTP was labeled with digoxin in the RNA-FISH assay. The tissue was dehydrated by gradient alcohol and was paraffin-embedded. The paraffin was sliced through the slicer and roasted for 2 h in a 62° oven. The sections were soaked in two changes of xylene for 15 min each. The cells were dehydrated in two changes of pure ethanol for 5 min each. Then, the samples were dehydrated in gradient ethanol of 85% and 75% ethanol 5 min for each. Wash in DEPC dilution. According to the tissue fixation time, the slices were boiled in the retrieval solution for 10 to 15 min and naturally cooled. The objective tissue was marked with liquid blocker pen according to the characteristics of tissues. Proteinase K (20 µg/mL) working solution was added to cover objectives, which were incubated at 37 °C for 20 min. The sections were washed with pure water, and then washed three times with PBS (pH 7.4) in a rocker device for 5 min each. Three percent of methanol-H<sub>2</sub>O<sub>2</sub> was added, and the mixture was incubated in dark at room temperature for 15 min. The slides were washed in PBS (pH 7.4) three times for 5 min each with gentle agitation. Prehybridization solution was added to each section and the sections were incubated for 1 h at 37 °C. The prehybridization solution was removed, the SELTP probe hybridization solution was added at a concentration of 1 µM, and the sections

were incubated in a humidity chamber and hybridized overnight at 42 °C. The hybridization solution was removed. Wash sections with 2×SSC for 10 min at 37 °C. Blocking solution (rabbit serum) was added to the section, which were incubated at room temperature for 30 min. The mouse anti-digoxin-labeled peroxidase (anti-DIG-HRP) was added, the blocking solution was removed, and anti-DIG-HRP was added. The mixture was incubated at 37 °C for 40 min, after which the sections were washed in PBS four times for 5 min each. The sections were dried slightly, and freshly prepared TSA chromogenic reagent was added to the marked tissue. The mixture was allowed to react in the dark for 5 min at room temperature. Then, the sections were washed in PBS three times for 10 min each. The cells were incubated with DAPI for 8 min in the dark and then mounted. The fluorescence signals for DIG (excitation 488 nm, emission 505 to 550 nm) were detected. Experiments were repeated three times independently. The related probe information is listed in Supplementary Table S11.

### Vector construction and transformation

The coding sequence of SELTP (Cottongen: Gh\_D12G0517) was obtained from amplification of PEC tissue and was subcloned into the pEarleyGate203 vector by Gateway Technology (Invitrogen, Carlsbad, CA) for generating overexpression lines (Earley et al. 2006). The unique transcript fragment of SELTP was inserted into pb7GWIWG2 (II) interference vector to construct the SELTP-RNAi recombinant plasmid (Helliwell et al. 2002). These constructs were transformed into *Agrobacterium tumefaciens* strain LBA4404 for genetic transformation of cotton hypocotyls. The positive transgenic lines were preliminary screened (Spiegelman et al. 2018). Then, the successful transformation and integrity of the target gene were further confirmed via PCR analysis (Yang et al. 2023). Ultimately, homozygous T3 generation transgenic lines were obtained through self-pollination. The related primers are listed in Supplementary Table S11.

### Cytological phenotype identification of transgenic lines

To identify the phenotypic characteristics during embryogenic differentiation initiation at the cytological level, staining with 1% I<sub>2</sub>-KI was used to examine the accumulation level of cellular amyloplasts under a Nikon Ni-U fluorescence microscope (Nikon Corp., Tokyo, Japan). I<sub>2</sub>-KI staining indicates the accumulation of amyloplasts, which probes the embryogenic competence during callus embryogenic differentiation.

### Cytochemical identification of cell masses with embryogenic competence

To generally distinguish embryogenic and nonembryogenic/dedifferentiated callus, acetocarmine/Evans blue double staining assay was performed as described previously (Steiner et al. 2016; Guo et al. 2019b). Callus samples were placed on a slide, stained with 2% (w/v) acetocarmine for

2 min, and rinsed with distilled sterile water three times. Then, the samples were stained with 0.5% (w/v) Evans blue for 30 s, washed again to remove the excess dye and mounted with cover slides. Finally, callus cells were examined under a Nikon Ni-U microscope with the CFI60 Image Capture System (Nikon Corp., Tokyo, Japan). In our previous reports, acetocarmine/Evans double staining assay was used as the assessment approach for testing the embryogenic competence of cell mass in callus. Cell mass with embryogenic competence can be stained red with acetocarmine, while dedifferentiated/nonembryogenic cell mass can be stained blue with Evans blue (Gupta and Durzan 1987; Steiner et al. 2016; Guo et al. 2019b). This procedure had been verified based on the molecular markers (BBM2, AGP1, LEC1, etc.) associated with embryogenic competence in our previous investigation. In addition, embryogenic competence can be further identified by cell morphology (round and small) and cytological characteristics (cellular amyloplasts state, dense cytoplasm, etc.) reported in previous work (Gupta and Durzan 1987; Steiner et al. 2016; Guo et al. 2019a).

### Light microscopy analysis

The images for I<sub>2</sub>-KI and acetocarmine/Evans blue staining were taken under a Nikon Ni-U microscope (Ni-U, Nikon Corp., Tokyo, Japan). Cultured tissues were observed and imaged using the Nikon SMZ25 stereomicroscope and the NIS Elements software (SMZ25, Nikon Corp., Tokyo, Japan).

### Statistical analysis

Three transgenic lines of each ectopic transformation were selected for the phenotypically statistical analysis. The proportion of cells in calli tissue with embryogenic competence (checked by quantitative testing of molecular marker genes, acetocarmine/Evans blue staining, and cytological morphology identification of small cells with round shape features) is the embryogenic mass number to the total cell mass number in each field of view. Ten fields of views were counted for each line with randomly selected samples. Independent experiments were repeated five times. The PEC sample with the same growth condition was cut into 0.5 × 0.5 cm<sup>2</sup> in size and inoculated on the medium. The number of somatic embryos per PEC clump was determined after 4 wk of culture. Twenty clumps were randomly selected for each transgenic and WT callus line. The mean number of somatic embryos per callus clump was calculated by dividing the total number of somatic embryos produced by the number of PEC masses that produce somatic embryos. This was repeated five times for each transgenic and WT callus line. Significant differences of statistical data described above were determined by Tukey's multiple comparisons test,  $P < 0.05$ . Values are presented with mean ± standard deviation (SD).

### Accession numbers

Gene sequences from this article can be found in the Cottongen database (<https://www.cottongen.org/>) under

accession numbers: SELTP (Gh\_D12G0517), AMY1 (Gh\_D05G1937), BBM2 (Gh\_A08G2008), Cht10 (Gh\_D11G1103), Fib-H (Gh\_Sca004839G02), PRP (Gh\_A02G0379), and Gluc (Gh\_D09G0636).

### Acknowledgments

Acknowledgments to Dr. Fang Liu, Dr. Xiongming Du, Dr. Yinhua Jia, and Dr. Chaojun Zhang in germplasm center in the Chinese Academy of Agricultural Sciences (CAAS) for providing the germplasm material.

### Author contributions

F.C.Z. and H.H.G. conceived and designed the research project. H.H.G., L.Z., H.X.G., and Y.P.F. performed tissue and cell sampling. H.H.G., L.Z., H.X.G., Y.P.F., T.T.L., A.Y.C., and F.J.S. performed all morphological, cellular, molecular, cytochemical, functional and scRNA-seq experiments. H.H.G., L.Z., H.X.G., X.W.C., X.S.Q., T.D.Y., and F.C.Z. analyzed the data and wrote the article.

### Supplementary data

The following materials are available in the online version of this article.

**Supplementary Figure S1.** Identification of cluster-specific marker genes for plant regeneration initiation.

**Supplementary Figure S2.** Heatmap of gene expression patterns of each cluster-associated module.

**Supplementary Figure S3.** Expression patterns of individual genes over pseudo-time during plant regeneration initiation.

**Supplementary Figure S4.** Comparative analysis of the cell cycle patterns in callus cell sets.

**Supplementary Table S1.** Known marker genes for cell clustering.

**Supplementary Table S2.** The expression of SELTP in each cell cluster.

**Supplementary Table S3.** The specific marker genes identified in each cell cluster.

**Supplementary Table S4.** The cluster-specific marker genes and the associated modules.

**Supplementary Table S5.** The DE genes in different state cells.

**Supplementary Table S6.** The branch-dependent DE genes along pseudo-time cell trajectory.

**Supplementary Table S7.** The highly connected hub genes in each module.

**Supplementary Table S8.** Cell cycle stage-specific genes in callus cell sets.

**Supplementary Table S9.** Key reagent resources used in this study.

**Supplementary Table S10.** Basic statistics information of scRNA-seq results.

**Supplementary Table S11.** Sequences of primers and probe used in this study.

## Funding

This project was supported by the National Natural Science Foundation of China (32001589, 32372142), Modern Agro-industry Technology Research System of Shandong Province (SDAIT-03-02), and Young Elite Scientists Sponsorship Program by CAST (2022QNRC001).

*Conflict of interest statement.* The authors declare no competing interests.

## Data availability

All relevant data can be found and available within the manuscript and supplemental information.

## References

- Bergen V, Lange M, Peidli S, Wolf FA, Theis FJ. Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat Biotechnol.* 2020;**38**(12):1408–1414. <https://doi.org/10.1038/s41587-020-0591-3>
- Birnbaum KD. Power in numbers: single-cell RNA-Seq strategies to dissect complex tissues. *Annu Rev Genet.* 2018;**52**(1):203–221. <https://doi.org/10.1146/annurev-genet-120417-031247>
- Boutilier K, Offringa R, Sharma VK, Kieft H, Ouellet T, Zhang L, Hattori J, Liu CM, van Lammeren AA, Miki BL, et al. Ectopic expression of BABY BOOM triggers a conversion from vegetative to embryonic growth. *Plant Cell.* 2002;**14**(8):1737–1749. <https://doi.org/10.1105/tpc.001941>
- Buels R, Yao E, Diesh CM, Hayes RD, Munoz-Torres M, Helt G, Goodstein DM, Elsik CG, Lewis SE, Stein L, et al. JBrowse: a dynamic web platform for genome visualization and analysis. *Genome Biol.* 2016;**17**(1):66. <https://doi.org/10.1186/s13059-016-0924-1>
- Earley KW, Haag JR, Pontes O, Opper K, Juehne T, Song K, Pikaard CS. Gateway-compatible vectors for plant functional genomics and proteomics. *Plant J.* 2006;**45**(4):616–629. <https://doi.org/10.1111/j.1365-313X.2005.02617.x>
- Efroni I, Mello A, Nawy T, Ip PL, Rahni R, DelRose N, Powers A, Satija R, Birnbaum KD. Root regeneration triggers an embryo-like sequence guided by hormonal interactions. *Cell.* 2016;**165**(7):1721–1733. <https://doi.org/10.1016/j.cell.2016.04.046>
- Elhiti M, Tahir M, Gulden RH, Khamiss K, Stasolla C. Modulation of embryo-forming capacity in culture through the expression of *Brassica* genes involved in the regulation of the shoot apical meristem. *J Exp Bot.* 2010;**61**(14):4069–4085. <https://doi.org/10.1093/jxb/erq222>
- Fehér A, Pasternak T, Ötvös K, Miskolczi PC, Dudit D. Induction of embryogenic competence in somatic plant cells: a review. *Biologia.* 2002;**51**:5–12. <https://www.researchgate.net/publication/232221295>
- Guo HH, Guo HX, Zhang L, Fan YJ, Fan YP, Zeng FC. SELTP-assembled battery drives totipotency of somatic plant cell. *Plant Biotechnol J.* 2019a;**17**(7):1188–1190. <https://doi.org/10.1111/pbi.13107>
- Guo HH, Wu JF, Chen CX, Wang HM, Zhao YL, Zhang CJ, Jia YH, Liu F, Ning TY, Chu ZH, et al. Identification and characterization of cell cultures with various embryogenic/regenerative potential in cotton based on morphological, cytochemical, and cytogenetical assessment. *J Integr Agr.* 2019b;**18**(1):1–8. [https://doi.org/10.1016/S2095-3119\(17\)61876-8](https://doi.org/10.1016/S2095-3119(17)61876-8)
- Gupta PK, Durzan DJ. Biotechnology of somatic polyembryogenesis and plantlet regeneration in loblolly pine. *Nat Biotechnol.* 1987;**5**(2):147–151. <https://doi.org/10.1038/nbt0287-147>
- Helliwell CA, Wesley SV, Wielopolska AJ, Waterhouse PM. High-throughput vectors for efficient gene silencing in plants. *Funct Plant Biol.* 2002;**29**(10):1217–1225. <https://doi.org/10.1071/FP02033>
- Horstman A, Li M, Heidmann I, Weemen M, Chen B, Muino JM, Angenent GC, Boutilier K. The BABY BOOM transcription factor activates the LEC1-ABI3-FUS3-LEC2 network to induce somatic embryogenesis. *Plant Physiol.* 2017;**175**(2):848–857. <https://doi.org/10.1104/pp.17.00232>
- Hu H, Xiong L, Yang Y. Rice SERK1 gene positively regulates somatic embryogenesis of cultured cell and host defense response against fungal infection. *Planta.* 2005;**222**(1):107–117. <https://doi.org/10.1007/s00425-005-1534-4>
- Islam S, Zeisel A, Joost S, La Manno G, Zajac P, Kasper M, Lönnberg P, Linnarsson S. Quantitative single-cell RNA-Seq with unique molecular identifiers. *Nat Methods.* 2014;**11**(2):163–166. <https://doi.org/10.1038/nmeth.2772>
- Izuno A, Maruyama TE, Ueno S, Ujino-Ihara T, Moriguchi Y. Genotype and transcriptome effects on somatic embryogenesis in *Cryptomeria japonica*. *PLoS One.* 2020;**15**(12):e0244634. <https://doi.org/10.1371/journal.pone.0244634>
- Jin F, Hu L, Yuan D, Xu J, Gao W, He L, Yang X, Zhang X. Comparative transcriptome analysis between somatic embryos (SEs) and zygotic embryos in cotton: evidence for stress response functions in SE development. *Plant Biotechnol J.* 2014;**12**(2):161–173. <https://doi.org/10.1111/pbi.12123>
- Jin S, Zhang X, Nie Y, Guo X, Liang S, Zhu H. Identification of a novel elite genotype for *in vitro* culture and genetic transformation of cotton. *Biol Plantarum.* 2006;**50**(4):519–524. <https://doi.org/10.1007/s10535-006-0082-5>
- Kurczynska EU, Potocka I, Dobrowolska I, Kulinska-Lukaszek K, Sala K, Wrobel J. Cellular markers for somatic embryogenesis. In: Sato K, editors. *Embryogenesis*. London: IntechOpen; 2012. p. 307–332.
- Kwong RW, Bui AQ, Lee H, Kwong LW, Fischer RL, Goldberg RB, Harada JJ. LEAFY COTYLEDON1-LIKE defines a class of regulators essential for embryo development. *Plant Cell.* 2003;**15**(1):5–18. <https://doi.org/10.1105/tpc.006973>
- La Manno G, Soldatov R, Zeisel A, Braun E, Hochgerner H, Petukhov V, Lidschreiber K, Kastriti ME, Lönnberg P, Furlan A, et al. RNA velocity of single cells. *Nature.* 2018;**560**(7719):494–498. <https://doi.org/10.1038/s41586-018-0414-6>
- Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics.* 2008;**9**(1):559. <https://doi.org/10.1186/1471-2105-9-559>
- Li J, Wang M, Li Y, Zhang Q, Lindsey K, Daniell H, Jin S, Zhang X. Multi-omics analyses reveal epigenomics basis for cotton somatic embryogenesis through successive regeneration acclimation process. *Plant Biotechnol J.* 2019;**17**(2):435–450. <https://doi.org/10.1111/pbi.12988>
- Li Y, Ma H, Wu Y, Ma Y, Yang J, Li Y, Yue D, Zhang R, Kong J, Lindsey K, et al. Single-cell transcriptome atlas and regulatory dynamics in developing cotton anthers. *AdvSci.* 2023;**11**(3):e2304017. <https://doi.org/10.1002/advs.202304017>
- Lin JL, Chen L, Wu WK, Guo XX, Yu CH, Xu M, Nie GB, Dun JL, Li Y, Xu B, et al. Single-cell RNA sequencing reveals a hierarchical transcriptional regulatory network of terpenoid biosynthesis in cotton secretory glandular cells. *Mol Plant.* 2023;**23**:00322–2. <https://doi.org/10.1016/j.molp.2023.10.008>
- Long L, Xu FC, Wang CH, Zhao XT, Yuan M, Song CP, Gao W. Single-cell transcriptome atlas identified novel regulators for pigment gland morphogenesis in cotton. *Plant Biotechnol J.* 2023;**21**(6):1100–1102. <https://doi.org/10.1111/pbi.14035>
- Lucau-Danila A, Laborde L, Legrand S, Huot L, Hot D, Lemoine Y, Hilbert JL, Hawkins S, Quillet MC, Hendriks T, et al. Identification of novel genes potentially involved in somatic embryogenesis in chicory (*Cichorium intybus* L.). *BMC Plant Biol.* 2010;**10**(1):122. <https://doi.org/10.1186/1471-2229-10-122>
- Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, Tirosh I, Bialas AR, Kamitaki N, Martersteck EM, et al. Highly parallel genome-wide expression profiling of individual cells using

- nanoliter droplets. *Cell*. 2015;**161**(5):1202–1214. <https://doi.org/10.1016/j.cell.2015.05.002>
- Mahdavi-Darvari F, Noor NM, Ismanizan I.** Epigenetic regulation and gene markers as signals of early somatic embryogenesis. *Plant Cell Tiss Org*. 2015;**120**(2):407–422. <https://doi.org/10.1007/s11240-014-0615-0>
- Min L, Hu Q, Li Y, Xu J, Ma Y, Zhu L, Yang X, Zhang X.** LEAFY COTYLEDON1-CASEIN KINASE I-TCP15-PHYTOCHROME INTERACTING FACTOR4 network regulates somatic embryogenesis by regulating auxin homeostasis. *Plant Physiol*. 2015;**169**(4):2805–2821. <https://doi.org/10.1104/pp.15.01480>
- Mironova V, Xu J.** A single-cell view of tissue regeneration in plants. *Curr Opin Plant Biol*. 2019;**52**:149–154. <https://doi.org/10.1016/j.pbi.2019.09.003>
- Ogura N, Sasagawa Y, Ito T, Tameshige T, Kawai S, Sano M, Doll Y, Iwase A, Kawamura A, Suzuki T, et al.** WUSCHEL-RELATED HOMEOBOX 13 suppresses de novo shoot regeneration via cell fate control of pluripotent callus. *Sci Adv*. 2023;**9**(27):eadg6983. <https://doi.org/10.1126/sciadv.adg6983>
- Poon S, Heath RL, Clarke AE.** A chimeric arabinogalactan protein promotes somatic embryogenesis in cotton cell culture. *Plant Physiol*. 2012;**160**(2):684–695. <https://doi.org/10.1104/pp.112.203075>
- Qin Y, Sun M, Li W, Xu M, Shao L, Liu Y, Zhao G, Liu Z, Xu Z, You J, et al.** Single-cell RNA-Seq reveals fate determination control of an individual fibre cell initiation in cotton (*Gossypium hirsutum*). *Plant Biotechnol J*. 2022;**20**(12):2372–2388. <https://doi.org/10.1111/pbi.13918>
- Qiu X, Hill A, Packer J, Lin D, Ma YA, Trapnell C.** Single-cell mRNA quantification and differential analysis with census. *Nat Methods*. 2017;**14**(3):309–315. <https://doi.org/10.1038/nmeth.4150>
- Rich-Griffin C, Stechemesser A, Finch J, Lucas E, Ott S, Schäfer P.** Single-cell transcriptomics: a high-resolution avenue for plant functional genomics. *Trends Plant Sci*. 2020;**25**(2):186–197. <https://doi.org/10.1016/j.tplants.2019.10.008>
- Roeder AHK, Otegui MS, Dixit R, Anderson CT, Faulkner C, Zhang Y, Harrison MJ, Kirchhelle C, Goshima G, Coate JE, et al.** Fifteen compelling open questions in plant cell biology. *Plant Cell*. 2021;**34**(1):72–102. <https://doi.org/10.1093/plcell/koab225>
- Roszak P, Heo JO, Blob B, Toyokura K, Sugiyama Y, de Luis Balaguer MA, Lau WWY, Hamey F, Cirrone J, Madej E, et al.** Cell-by-cell dissection of phloem development links a maturation gradient to cell specialization. *Science*. 2021;**374**(6575):eaba5531. <https://doi.org/10.1126/science.aba5531>
- Rupp A, Raschke J, Rümmler M, Linke B, Zoglauer K.** Identification of putative homologs of *Larix decidua* to BABYBOOM (BBM), LEAFY COTYLEDON<sub>1</sub> (LEC<sub>1</sub>), WUSCHEL-related HOMEOBOX<sub>2</sub> (WOX<sub>2</sub>) and SOMATIC EMBRYOGENESIS RECEPTOR-like KINASE (SERK) during somatic embryogenesis. *Planta*. 2016;**243**(2):473–488. <https://doi.org/10.1007/s00425-015-2409-y>
- Ryu KH, Huang L, Kang HM, Schiefelbein J.** Single-cell RNA sequencing resolves molecular relationships among individual plant cells. *Plant Physiol*. 2019;**179**(4):1444–1456. <https://doi.org/10.1104/pp.18.01482>
- Sablowski R.** The dynamic plant stem cell niches. *Curr Opin Plant Biol*. 2007;**10**(6):639–644. <https://doi.org/10.1016/j.pbi.2007.07.001>
- Satija R, Farrell JA, Gennert D, Schier AF, Regev A.** Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol*. 2015;**33**(5):495–502. <https://doi.org/10.1038/nbt.3192>
- Shulse CN, Cole BJ, Ciobanu D, Lin J, Yoshinaga Y, Gouran M, Turco GM, Zhu Y, O’Malley RC, Brady SM, et al.** High-throughput single-cell transcriptome profiling of plant cell types. *Cell Rep*. 2019;**27**(7):2241–2247.e4. <https://doi.org/10.1016/j.celrep.2019.04.054>
- Song X, Guo P, Xia K, Wang M, Liu Y, Chen L, Zhang J, Xu M, Liu N, Yue Z, et al.** Spatial transcriptomics reveals light-induced chlorenchyma cells involved in promoting shoot regeneration in tomato callus. *Proc Natl Acad Sci U S A*. 2023;**120**(38):e2310163120. <https://doi.org/10.1073/pnas.2310163120>
- Spiegelman Z, Lee CM, Gallagher KL.** King is a plant-specific kinesin that regulates both intra- and intercellular movement of SHORT-ROOT. *Plant Physiol*. 2018;**176**(1):392–405. <https://doi.org/10.1104/pp.17.01518>
- Steiner N, Farias-Soares FL, Schmidt ÉC, Pereira ML, Scheid B, Rogge-Renner GD, Bouzon ZL, Schmitz D, Maldonado S, Guerra MP.** Toward establishing a morphological and ultrastructural characterization of proembryogenic masses and early somatic embryos of *Araucaria angustifolia* (Bert.) O. Kuntze. *Protoplasma*. 2016;**253**(2):487–501. <https://doi.org/10.1007/s00709-015-0827-0>
- Su YH, Liu YB, Zhou C, Li XM, Zhang XS.** The microRNA167 controls somatic embryogenesis in *Arabidopsis* through regulating its target genes ARF6 and ARF8. *Plant Cell Tiss Org*. 2016;**124**(2):405–417. <https://doi.org/10.1007/s11240-015-0903-3>
- Sun D, Guan X, Moran AE, Wu LY, Qian DZ, Schedin P, Dai MS, Danilov AV, Alumkal JJ, Adey AC, et al.** Identifying phenotype-associated subpopulations by integrating bulk and single-cell sequencing data. *Nat Biotechnol*. 2022;**40**(4):527–538. <https://doi.org/10.1038/s41587-021-01091-3>
- Sun Y, Han Y, Sheng K, Yang P, Cao Y, Li H, Zhu QH, Chen J, Zhu S, Zhao T.** Single-cell transcriptomic analysis reveals the developmental trajectory and transcriptional regulatory networks of pigment glands in *Gossypium bickii*. *Mol Plant*. 2023;**16**(4):694–708. <https://doi.org/10.1016/j.molp.2023.02.005>
- Verdeil J, Hocher V, Huet C, Grosdemange F, Escoute J, Ferriere NM, Nicole MR.** Ultrastructural changes in coconut calli associated with the acquisition of embryogenic competence. *Ann Bot*. 2001;**88**(1):9–18. <https://doi.org/10.1006/anbo.2001.1408>
- Vogel G.** How does a single somatic cell become a whole plant? *Science*. 2005;**309**(5731):86. <https://doi.org/10.1126/science.309.5731.86>
- Wen L, Li W, Parris S, West M, Lawson J, Smathers M, Li Z, Jones D, Jin S, Saski CA.** Transcriptomic profiles of non-embryogenic and embryogenic callus cells in a highly regenerative upland cotton line (*Gossypium hirsutum* L.). *BMC Dev Biol*. 2020;**20**(1):25. <https://doi.org/10.1186/s12861-020-00230-4>
- Xu M, Du Q, Tian C, Wang Y, Jiao Y.** Stochastic gene expression drives mesophyll protoplast regeneration. *Sci Adv*. 2021;**7**(33):eabg8466. <https://doi.org/10.1126/sciadv.abg8466>
- Yang Z, Liu Z, Ge X, Lu L, Qin W, Qanmber G, Liu L, Wang Z, Li F.** Brassinosteroids regulate cotton fiber elongation by modulating very-long-chain fatty acid biosynthesis. *Plant Cell*. 2023;**35**(6):2114–2131. <https://doi.org/10.1093/plcell/koad060>
- Zeng F, Zhang X, Zhu L, Tu L, Guo X, Nie Y.** Isolation and characterization of genes associated to cotton somatic embryogenesis by suppression subtractive hybridization and macroarray. *Plant Mol Biol*. 2006;**60**(2):167–183. <https://doi.org/10.1007/s11103-005-3381-x>
- Zhai N, Xu L.** Pluripotency acquisition in the middle cell layer of callus is required for organ regeneration. *Nat Plants*. 2021;**7**(11):1453–1460. <https://doi.org/10.1038/s41477-021-01015-8>
- Zhang S, Zhu C, Zhang X, Liu M, Xue X, Lai C, Xuhan X, Chen Y, Zhang Z, Lai Z, et al.** Single-cell RNA sequencing analysis of the embryogenic callus clarifies the spatiotemporal developmental trajectories of the early somatic embryo in *Dimocarpus longan*. *Plant J*. 2023;**115**(5):1277–1297. <https://doi.org/10.1111/tpj.16319>
- Zhang TQ, Chen Y, Wang JW.** A single-cell analysis of the *Arabidopsis* vegetative shoot apex. *Dev Cell*. 2021;**56**(7):1056–1074.e8. <https://doi.org/10.1016/j.devcel.2021.02.021>
- Zhang TQ, Xu ZG, Shang GD, Wang JW.** A single-cell RNA sequencing profiles the developmental landscape of *Arabidopsis* root. *Mol Plant*. 2019;**12**(5):648–660. <https://doi.org/10.1016/j.molp.2019.04.004>
- Zhu X, Xu Z, Wang G, Cong Y, Yu L, Jia R, Qin Y, Zhang G, Li B, Yuan D, et al.** Single-cell resolution analysis reveals the preparation for reprogramming the fate of stem cell niche in cotton lateral meristem. *Genome Biol*. 2023;**24**(1):194. <https://doi.org/10.1186/s13059-023-03032-6>

**Ziegenhain C, Vieth B, Parekh S, Reinius B, Guillaumet-Adkins A, Smets M, Leonhardt H, Heyn H, Hellmann I, Enard W.** Comparative analysis of single-cell RNA sequencing methods. *Mol Cell*. 2017;**65**(4): 631–643.e4. <https://doi.org/10.1016/j.molcel.2017.01.023>

**Zywitzka V, Misios A, Bunyatyan L, Willnow TE, Rajewsky N.** Single-cell transcriptomics characterizes cell types in the subventricular zone and uncovers molecular defects impairing adult neurogenesis. *Cell Rep*. 2018;**25**(9):e2458. <https://doi.org/10.1016/j.celrep.2018.11.003>