

Single-cell RNA sequencing analysis of the embryogenic callus clarifies the spatiotemporal developmental trajectories of the early somatic embryo in *Dimocarpus longan*

Shuting Zhang¹, Chen Zhu¹, Xueying Zhang¹, Mengyu Liu¹, Xiaodong Xue¹, Chunwang Lai¹, Xu Xuhuan^{1,2}, Yukun Chen¹, Zihao Zhang¹, Zhongxiong Lai^{1,*}  and Yuling Lin^{1,*} 

¹Institute of Horticultural Biotechnology, Fujian Agriculture and Forestry University, Fuzhou 350002, China, and

²Institut de la Recherche Interdisciplinaire de Toulouse, Toulouse 31300, France

Received 17 July 2022; revised 16 May 2023; accepted 23 May 2023; published online 26 May 2023.

*For correspondence (e-mail buliang84@163.com; laizx01@163.com).

SUMMARY

Plant embryogenic calli (ECs) can undergo somatic embryogenesis to regenerate plants. This process is mediated by regulatory factors, such as transcription factors and specifically expressed genes, but the precise molecular mechanisms underlying somatic embryogenesis at the single-cell level remain unclear. In this study, we performed high-resolution single-cell RNA sequencing analysis to determine the cellular changes in the EC of the woody plant species *Dimocarpus longan* (longan) and clarify the continuous cell differentiation trajectories at the transcriptome level. The highly heterogeneous cells in the EC were divided into 12 putative clusters (e.g., proliferating, meristematic, vascular, and epidermal cell clusters). We determined cluster-enriched expression marker genes and found that overexpression of the epidermal cell marker gene *GDSL ESTERASE/LIPASE-1* inhibited the hydrolysis of triacylglycerol. In addition, the stability of autophagy was critical for the somatic embryogenesis of longan. The pseudo-timeline analysis elucidated the continuous cell differentiation trajectories from early embryonic cell division to vascular and epidermal cell differentiation during the somatic embryogenesis of longan. Moreover, key transcriptional regulators associated with cell fates were revealed. We found that *ETHYLENE RESPONSIVE FACTOR 6* was characterized as a heat-sensitive factor that negatively regulates longan somatic embryogenesis under high-temperature stress conditions. The results of this study provide new spatiotemporal insights into cell division and differentiation during longan somatic embryogenesis at single-cell resolution.

Keywords: ERF6, heat stress, longan, scRNA-seq, somatic embryogenesis.

INTRODUCTION

Plants form various organs throughout their life cycle, including roots, stems, leaves, and flowers, which are associated with an integrated embryo structure that is critical for generating progeny. The plant embryo developmental process is referred to as embryogenesis. In *Arabidopsis thaliana* (Arabidopsis), embryogenesis starts from a single totipotent cell (i.e., zygote) (Ten Hove et al., 2015; Zhou et al., 2020). The zygote undergoes asymmetrical cell division, generating two daughter cells, an apical cell in the upper tier and a basal cell in the lower tier. The basal cell undergoes fewer cell divisions and can transform into the suspensor and hypophysis (Ten Hove et al., 2015; Zhou et al., 2020). The apical cell undergoes continuous division and differentiation processes that result in a mature embryo that comprises diverse basic tissue types, including the epidermis, meristematic tissue, hypophysis,

ground tissue, and vascular tissue (Ten Hove et al., 2015; Zhou et al., 2020). During post-embryonic development, these tissues continue to differentiate to form distinct structures (e.g., epidermal layer, xylem, phloem, and secondary cell wall) (Ten Hove et al., 2015; Zhou et al., 2020).

Although cell division and differentiation in the embryos of most flowering plants may appear to be chaotic and random events, the zygote that differentiates into a mature embryo will always contain these basic cell types (Ten Hove et al., 2015). However, the early zygotic embryos of most flowering plants are small (i.e., undetectable to the naked eye) and difficult to obtain. This is an obstacle to studying the mechanisms underlying embryo development in most plant species. Alternatively, somatic embryogenesis has been established in numerous plant species as a model system for studying embryo development and plant regeneration *in vitro*. Totipotent somatic cells in plant

tissues can be de-differentiated into embryogenic callus (EC) cells in response to growth regulators, including hormones, by applying plant tissue culture technology. Additionally, ECs may undergo somatic embryogenesis and ultimately develop into a whole plant via a process that is highly similar to zygotic embryogenesis. To date, the somatic embryogenesis system has been extensively applied to study the molecular basis of plant regeneration in many species that have zygotic embryos that are difficult to obtain (Horstman et al., 2017).

The molecular mechanisms of zygotic embryogenesis and somatic embryogenesis in plants have been widely investigated. Several embryogenesis-related marker genes are reportedly involved in plant embryo development, including *WUSCHEL-RELATED HOMEOBOX* (*WOX*) (Haecker et al., 2004), *BABY BOOM* (*BBM*), *LEAFY COTYLEDON 1* (*LEC1*), *LEC2* (Horstman et al., 2017), *AGAMOUS-LIKE 15* (*AGL15*) (Horstman et al., 2017), *ARABINOGALACTAN PROTEIN* (*AGP*) (Zhong et al., 2011), *LIPID TRANSFER PROTEIN 1* (*LTP1*) (Potocka et al., 2012), *GERMIN-LIKE PROTEIN* (*GLP*) (Mathieu et al., 2006), and *PEROXIDASE 1* (*POD1*) (Li et al., 2011). Moreover, complex transcriptional regulatory networks control the production of key cell types during the plant developmental process. For example, *PROTODERMAL FACTOR 2* (*PDF2*) expression is restricted to the protoderm in the early globular embryo stage (Abe et al., 2003); *LACCASE* (*LAC*) can affect vascular tissue development (Berthet et al., 2011; Q. Zhao et al., 2013); and *LTPG* can transport the lipids inside epidermal cells to the epidermal layer to form a cuticular layer (Kim et al., 2012).

Despite the extensive research on the molecular mechanism mediating plant embryogenesis, a comprehensive analysis of all EC cells has not been conducted, resulting in the limited characterization of the changes in the continuously differentiating EC cells. Previous related studies usually involved bulk RNA sequencing (RNA-seq) analyses to clarify the molecular functions of pooled cell populations, but this approach revealed very little about the characteristics of specific cell types. High-throughput single-cell RNA-seq (scRNA-seq) technology was recently developed to enable researchers to investigate gene expression patterns in individual cells. This approach facilitates the simultaneous analysis of thousands of cells in one experiment, thereby providing unique insights into gene expression, cell heterogeneity, and dynamic cell differentiation trajectories (Potter, 2018). Accordingly, scRNA-seq technology has been widely applied to conduct animal and medical research, including the characterization of animal embryos (Wagner et al., 2018). Recently, it has also been applied to study various plant species, including *Arabidopsis* (Zhai & Xu, 2021; Zhang et al., 2019), *Oryza sativa* (rice) (Liu et al., 2020), *Lycopersicon esculentum* (tomato) (Tian et al., 2020), *Zea mays* (maize) (Nelms & Walbot, 2019; Xu et al., 2021), *Arachis hypogaea* (peanut) (Liu et al.,

2021), *Camellia sinensis* (tea plant) (Wang et al., 2022), and *Catharanthus roseus* (Sun et al., 2023), with these studies focusing on the roots, shoots, ears, callus, and seedling cotyledon tissues. To date, scRNA-seq technology has not been used to examine zygotic or somatic embryo tissues in horticultural plants.

Dimocarpus longan Lour. (longan), which originated in Southeast Asia, is an important Sapindaceae woody fruit tree. However, collecting the longan zygotic embryo in the early developmental stage is a major challenge that hinders the elucidation of the molecular mechanisms underlying zygotic embryo development. Therefore, the longan somatic embryogenesis system has been used to study longan embryo development *in vitro* (Guan et al., 2016). Longan somatic embryogenesis is initiated by an EC, which is induced by the immature zygotic cotyledon embryo. Longan ECs consist of a certain proportion of embryonic cells, which can differentiate into mature somatic embryos through the effects of the synthetic auxin 2,4-dichlorophenoxyacetic acid (2,4-D) in Murashige and Skoog (MS) medium (Lai & Chen, 1997). The gene expression trends during the early somatic embryogenesis of longan were determined in previous research (Chen, Xu, et al., 2020; Lin et al., 2017). However, in these studies, the gene expression profiles were analyzed at specific developmental stages in limited pooled cell populations, and the specific cell types in the somatic embryos of longan remain unclear.

In the present study, we generated a high-resolution single-cell gene expression atlas of longan EC. A total of 12 cell clusters were detected. We identified four cell types in the EC and revealed cluster-enriched expression marker genes in the main cell types. The profiling and ordering of individual cells enabled us to reconstruct the continuous cell differentiation trajectories from early embryonic cell division to vascular and epidermal cell differentiation. Furthermore, precise spatiotemporal expression patterns and molecular mechanisms were determined. The identified cluster-enriched expression marker genes and transcription factor (TF) regulatory networks may be useful for future molecular and cellular investigations of somatic embryogenesis in longan.

RESULTS

scRNA-seq and identification of cell clusters in the longan EC

The longan EC was derived from an immature zygotic cotyledon embryo (40–50 days after flowering) cultured on MS solid medium containing 2.0 mg L⁻¹ 2,4-D for 6 weeks at 25°C in darkness (Lai et al., 1997). After sequencing, a total of 28 727 EC cells from two biological replicates were isolated and profiled via droplet-based scRNA-seq (Figure 1a), which generated approximately 518 946 126 reads (about 18 064 reads per cell). The median gene count and unique molecular identifier (UMI) count per cell were 874 and

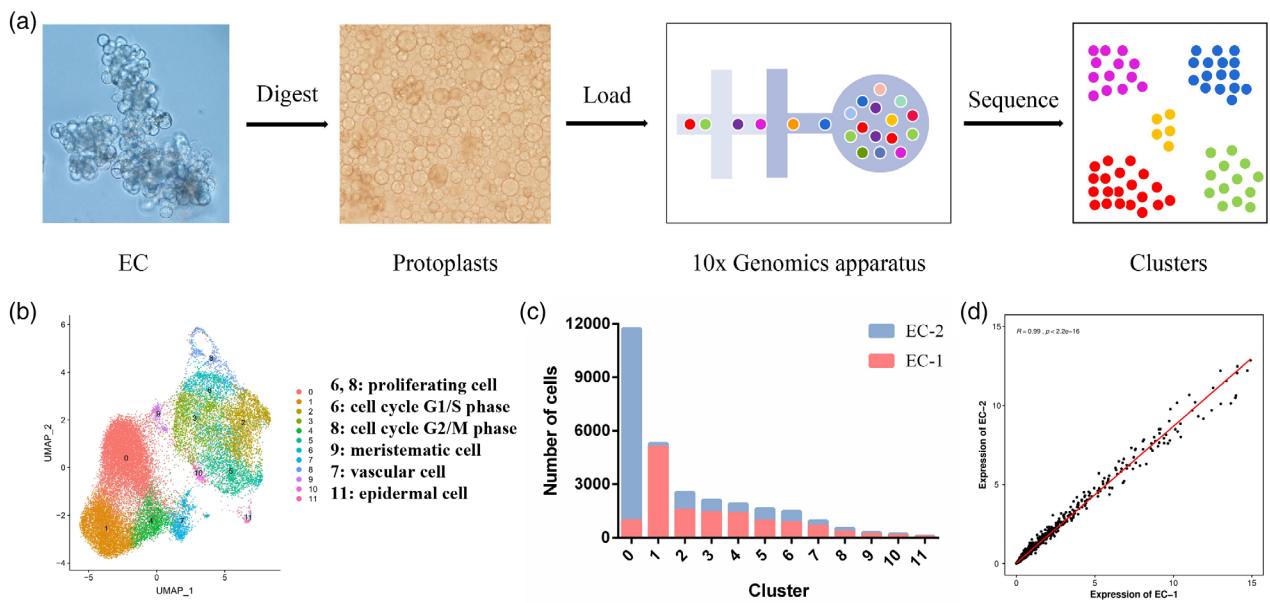


Figure 1. Single-cell RNA-seq analysis of the embryogenic callus and identification of cell clusters.

(a) Overview of the scRNA-seq analysis of the embryogenic callus. Protoplasts were isolated from the longan embryogenic callus. The scRNA-seq libraries generated using the 10x Genomics platform were analyzed by high-throughput sequencing. Each dot denotes a single cell.
 (b) Visualization of the embryogenic callus cell clusters using the UMAP algorithm. Each dot denotes a single cell. Cell clusters are differentiated by color.
 (c) Number of cells in each cluster calculated using two biological replicates. Red and blue indicate the EC-1 and EC-2 replicates, respectively.
 (d) Correlation of the gene expression between scRNA-seq replicates EC-1 and EC-2 ($R = 0.99$). EC, embryogenic callus.

1330, respectively. After filtering the data, 28 724 cells were retained for further analyses. Distinct EC cell populations were identified from the single-cell transcriptome data on the basis of an unsupervised clustering analysis of gene expression patterns performed using Seurat software (Satija et al., 2015). Twelve distinct clusters were identified, with the number of cells ranging from 99 (cluster 11) to 11 735 (cluster 0) (Figure 1b,c). The scRNA-seq data revealed that the two biological replicates (EC-1 and EC-2) had a similar number of cells in each cluster, except for cluster 0 and cluster 1 (Figure 1c). Specifically, clusters 0 and 1 in EC-1 had 940 (7.00%) and 5079 (37.81%) cells, respectively, whereas clusters 0 and 1 in EC-2 had 10 795 (70.60%) and 208 (1.36%) cells, respectively. However, the cluster-enriched expression of the genes in each cluster was highly correlated between replicates ($R = 0.99$; Figure 1d). Thus, the cells in the two replicates were grouped into the same cluster (0 and 1). Consistent with our findings, in an earlier study on the *Arabidopsis* vegetative shoot apex, the scRNA-seq results also indicated that the number of cells in some clusters (e.g., clusters 11 and 21) differed substantially (Zhang et al., 2021). Because of the spatiotemporal differences in gene expression among the cells of the same tissue, the cell growth states also varied, resulting in a different number of cells in two replicates. To determine how well the gene expression patterns revealed by scRNA-seq reflected the gene expression in the longan EC, we compared the scRNA-seq and bulk

RNA-seq data for the upregulated marker genes. Pearson correlation analysis indicated that the gene expression data generated by the scRNA-seq and bulk RNA-seq analyses were highly correlated (i.e., correlation coefficient greater than 0.8) (Figure S1). Therefore, the scRNA-seq data appeared to accurately reflect the transcriptome of the longan EC.

To determine the potential cell types in the clusters, we analyzed the upregulated genes in the 12 clusters. The number of upregulated genes ranged from 47 (cluster 1) to 534 (cluster 11) (Figure 2a). We selected the top 20 upregulated differentially expressed genes (DEGs) as cell type-enriched marker genes (Figure 2b; Table S1). According to the visualization of the data using the uniform manifold approximation and projection (UMAP) (Becht et al., 2019) algorithm and the upregulated marker genes, the proliferating cells (clusters 6 and 8), meristematic cells (cluster 9), vascular cells (cluster 7), and epidermal cells (cluster 11) (Figure 1b) were identified in the longan EC.

Proliferating and meristematic cell population characteristics in the longan EC

The cluster 6 and 8 cells formed the proliferating cell population, in which the cell cycle-related marker genes were predominantly expressed, including the first gap/synthesis (G1/S) phase-specific gene *G1/S-SPECIFIC CYCLIN-D 3* (*CYCD3*; cluster 6) and the second gap/mitotic (G2/M) phase-specific genes *CYCLIN-DEPENDENT KINASE B2;1*

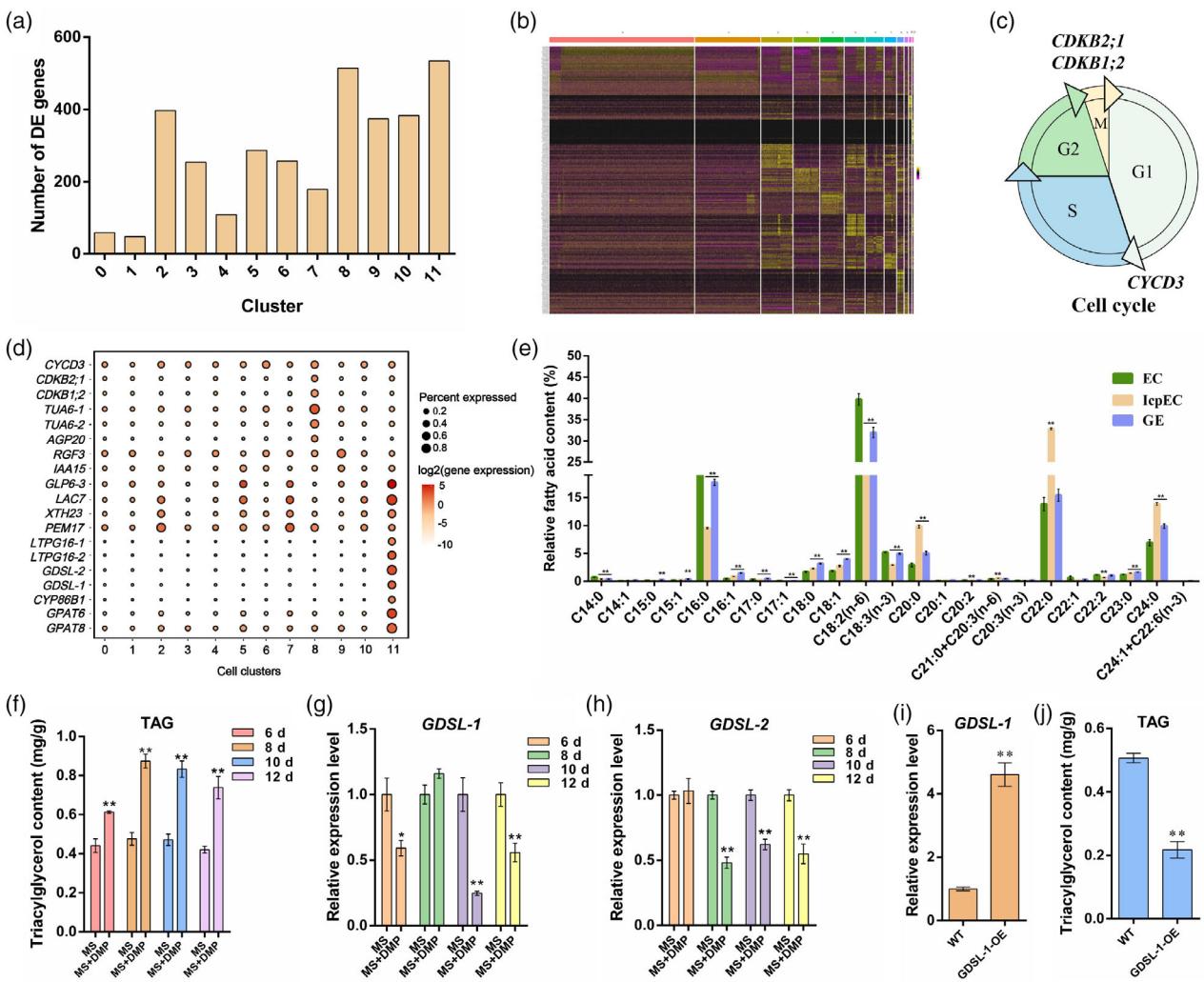


Figure 2. Functional characterization of genes with cluster-enriched expression.

(a) Number of upregulated differentially expressed genes (DEGs) in each cluster.

(b) Heatmap presenting the expression of the top 20 upregulated genes in each cluster. The gene list is given in Table S1.

(c) Cell cycle-related marker genes in clusters 6 and 8. Cell cycle phases are differentiated by color. G1, first gap phase; S, synthesis phase; G2, second gap phase; M, mitosis phase.

(d) Expression patterns of representative marker genes with cluster-enriched expression. Dot diameter and color indicate expression levels. The full names of selected genes are given in the main text.

(e) Fatty acid compositions and relative contents in the early somatic embryos (embryogenic callus, incomplete compact pro-embryogenic culture, and globular embryo stages). EC, embryogenic callus; IcpEC, incomplete compact pro-embryogenic culture; GE, globular embryo. Student's *t*-test, ***P* < 0.01.

(f) Triacylglycerol content changes in the longan early somatic embryos after diphenyl methylphosphonate treatment. Embryogenic callus cultured on MS medium without (mock) or with 10 µM diphenyl methylphosphonate to induce somatic embryos. Samples were harvested after 6, 8, 10, and 12 days to detect the triacylglycerol content. MS, Murashige and Skoog medium without additives; MS + DMP, Murashige and Skoog medium containing 10 µM diphenyl methylphosphonate. Student's *t*-test, ***P* < 0.01.

(g, h) *GDSL-1* and *GDSL-2* expression patterns in the early somatic embryos of longan treated with diphenyl methylphosphonate as determined by qRT-PCR. Somatic embryos cultured on MS medium without (mock) or with 10 µM diphenyl methylphosphonate. Samples were harvested after 6, 8, 10, and 12 days to detect the gene expression patterns. MS, Murashige and Skoog medium without additives; MS + DMP, Murashige and Skoog medium containing 10 µM diphenyl methylphosphonate. Student's *t*-test, **P* < 0.05, ***P* < 0.01.

(i) *GDSL-1* expression pattern in the *GDSL-1*-overexpressing embryogenic callus as determined by qRT-PCR. *GDSL-1*-overexpressing embryogenic callus cultured on MS medium. WT, wild type; *GDSL-1*-OE, *GDSL-1*-overexpressing callus. Student's *t*-test, ***P* < 0.01.

(j) Triacylglycerol content changes in the *GDSL-1*-overexpressing embryogenic callus. WT, wild type; *GDSL-1*-OE, *GDSL-1*-overexpressing callus. Student's *t*-test, ***P* < 0.01.

(*CDKB2;1*; cluster 8) and *CDKB1;2* (cluster 8) (Figure 2c,d; Figure S2) (Ascencio-Ibáñez et al., 2008; De Almeida Engler et al., 2009; Menges et al., 2002). Consistent with this

result, *CDKB2;1* was identified as a proliferating cell marker gene in a previous scRNA-seq analysis of the *Arabidopsis* vegetative shoot apex (Zhang et al., 2021). During the cell

cycle M phase, *TUBULIN ALPHA CHAIN 6-1 (TUA6-1)* and *TUA6-2* (Figure S3) are associated with the formation of microtubules in the spindle apparatus (Wright & Hunter, 2003), whereas *AGP20* (Figure S4) mediates cell division and cell plate formation (Yu & Zhao, 2012; Zhong et al., 2011). The upregulated expression of *TUA6-1*, *TUA6-2*, and *AGP20* in cluster 8 suggested that a number of cluster 8 cells were in the cell cycle M phase. Thus, the cluster 6 and 8 cells were designated as the proliferating cell population, with the cluster 6 cells in the G1/S phase and the cluster 8 cells in the G2/M phase.

Because of the upregulated expression of the meristem-specific gene *ROOT MERISTEM GROWTH FACTOR 3 (RGF3)* (Figure 2d; Figure S5) (Shinohara et al., 2016) in cluster 9, we defined the cells in this cluster as the meristematic cell population. Both *RGF2* and *RGF3* were designated as root meristematic cell marker genes following the scRNA-seq analysis of the Arabidopsis root (Zhang et al., 2019), reflecting the utility of *RGF* genes as root meristematic cell marker genes. Additionally, the expression of the auxin response gene *INDOLE-3-ACETIC ACID INDUCIBLE 15 (IAA15)* was upregulated in the meristematic cell population (Figure 2d; Figure S6), which was in accordance with the reported enhanced expression of the auxin response genes *IAA20* and *IAA33* in Arabidopsis root meristematic cells (Zhang et al., 2019). Accordingly, auxin may be important for maintaining cell states in the root meristem. These observations implied that cluster 9 in the longan EC consisted of the root-like meristematic cell population and that auxin may be involved in meristematic cell development.

Vascular cell population characteristics in the longan EC

Cluster 7, which was designated as the vascular cell population, included cell wall-associated marker genes with upregulated expression levels, including *GLP6-3*, *LAC7*, *XYLOGLUCAN ENDOTRANSGLUCOSYLASE/HYDROLASE 23 (XTH23)*, and *PECTIN METHYLESTERASE 17 (PME17)* (Figure 2d), implying that cell wall-related biological processes were active in cluster 7. In plant vascular tissues, the cell wall is modified by a secondary wall thickening process, during which lignin accumulates substantially. Recent studies demonstrated that *LAC* genes are important for lignin polymerization (Q. Zhao et al., 2013; Y. Zhao et al., 2015). A lack of lignification in the tissues of the Arabidopsis *lac11 lac4 lac17* triple mutant results in arrested vascular development (Q. Zhao et al., 2013; Y. Zhao et al., 2015). In the current study, we detected a lignin polymerization gene (*LAC7*) (Figure 2d; Figure S7) (Q. Zhao et al., 2013) that was overrepresented in cluster 7, providing further evidence that cluster 7 was comprised of vascular cells. According to the KEGG analysis, the ‘phenylpropanoid biosynthesis’ pathway was enriched among the genes in cluster 7, including 10 upregulated

DEGs (Figures S8 and S9; Table S2). Among these 10 DEGs, seven POD-encoding genes and *CYTOCHROME P450 FAMILY 98 SUBFAMILY A (CYP98A)* were specifically involved in lignin biosynthesis (Figure S9), indicative of the lignin synthesis in cluster 7 cells and further confirming that cluster 7 consisted of vascular cells.

Epidermal cell population characteristics in the longan EC

Cluster 11 was defined as the epidermal cell population. The upregulated expression of the epidermal cell marker genes *LTPG16-1*, *LTPG16-2*, and *GDSL ESTERASE/LIPASE (GDSL)* was specific to cluster 11 (Figure 2d) (Kim et al., 2012; Zhang et al., 2021). Consistent with these findings, an earlier scRNA-seq-based investigation of the epidermal cell population in the Arabidopsis vegetative shoot apex detected the increased expression of *LTP* family genes (*LTP1*, *LTPG1*, and *LTPG6*) and *GDSL* genes (Zhang et al., 2021). Cuticular waxes and cuticles, which cover the outer epidermal surface of land plants, mainly consist of aliphatic compounds, including fatty acids, hydrocarbons, aldehydes, alcohols, ketones, and glycerols (Suh et al., 2005). Phenylpropanoids, flavonoids, and terpenoids are also cuticular components (Suh et al., 2005). Cuticular waxes and cuticles are synthesized exclusively within the plant epidermal cell, after which they are transported to the outer surface of the epidermis via ADENOSINE TRIPHOSPHATE BINDING CASSETTE (ABC) and LTP transporters (Bird et al., 2007; DeBono et al., 2009; Kumar et al., 2016; Suh et al., 2005). In the Arabidopsis *ltpg1 ltpg2* double mutant, the composition and total abundance of the epidermal cell cuticular wax are significantly altered and the cuticular layer structure of the epidermal cell is modified substantially (Kim et al., 2012). Furthermore, *sn-GLYCEROL-3-PHOSPHATE ACYL-TRANSFERASE 6 (GPAT6)* and *GPAT8* (Figure 2d; Figure S10) were specifically expressed in cluster 11. In barley (*Hordeum vulgare*), these two genes are involved in the biosynthesis of the epidermal cuticle, thereby protecting spikelets from Fusarium head blight (Kumar et al., 2016). These results suggest that *LTPG* and *GPAT* genes are indispensable for the formation of epidermal cuticular waxes and cuticles.

According to the KEGG results, ‘fatty acid metabolism’, ‘fatty acid biosynthesis’, ‘fatty acid degradation’, and ‘fatty acid elongation’ were the enriched pathways among the genes in cluster 11 (Figure S8), suggesting the fatty acids in the cluster 11 cells were actively being metabolized. Because the upregulated genes in cluster 11 were associated with fatty acid metabolism-related pathways, we analyzed the fatty acid components in the longan early somatic embryos (i.e., EC, incomplete compact preembryogenic culture, and globular embryo stages). We detected higher relative contents of octadecadienoic acid ($C_{18:2}$ (n-6)), docosanoic acid ($C_{22:0}$), hexadecanoic acid

(C_{16:0}), tetracosanoic acid (C_{24:0}), eicosanoic acid (C_{20:0}), α -linolenic acid (C_{18:3 (n-3)}), octadecenoic acid (C_{18:1}), and octadecanoic acid (C_{18:0}) in the early somatic embryos (Figure 2e; Table S3). Among these compounds, C₁₆ and C₁₈ fatty acids are the main fatty acid types involved in epidermal cuticular wax and cuticle synthesis (Li-Beisson et al., 2013). The high proportions of C₁₆ and C₁₈ fatty acids further indicated that they were the main fatty acids contributing to the production of epidermal cuticular waxes and cuticles in longan somatic embryos. These observations suggested that epidermal cuticular wax and cuticle were synthesized in the epidermal cells and then exported by LTP transporters to the outer surface, where they accumulated to form the cuticular waxes and cuticles of longan. Overall, these results revealed that cluster 11 contained epidermal cells.

Overexpression of *GDSL* accelerated the hydrolysis of triacylglycerol in the embryonic callus of longan

The abovementioned results showed that the expression of the marker genes *GDSL-1* and *GDSL-2* was upregulated in cluster 11. Earlier research confirmed *GDSLs* are lipid hydrolytic enzymes with broad substrate specificity, such as acetyl and butyl esterase, sinapine esterase, fatty acyl ester hydrolase, and triacylglycerol (TAG) hydrolase (Akoh et al., 2004; Aparato & Suh, 2022; Lee et al., 1997). As an important energy source in plants, reserve lipids are stored as TAG in oil bodies, which are formed in the endoplasmic reticulum (Murphy, 2012). Thus, the oil body decomposition inhibitor diphenyl methylphosphonate (DMP), which preserves TAG by preventing the breakdown of oil bodies (Brown et al., 2013), was added to the MS solid medium to confirm whether *GDSL* genes are involved in lipid hydrolysis during the early somatic embryogenesis of longan. The TAG content increased significantly in the DMP-treated early somatic embryos (Figure 2f), revealing that DMP may inhibit the hydrolysis of TAG in longan. The expression levels of the marker genes *GDSL-1* and *GDSL-2* in the early somatic embryos treated with DMP were also analyzed. Compared with the control, *GDSL-1* expression tended to decrease in the early somatic embryos after 6, 10, and 12 days of DMP treatment (Figure 2g). The *GDSL-2* expression level also tended to decrease in the early somatic embryos after 8, 10, and 12 days of DMP treatment (Figure 2h), suggesting that both of these genes may be involved in the hydrolysis of TAG. To verify that *GDSL* genes participate in TAG hydrolysis in longan, *GDSL-1* was selected for stable transformation of the longan EC (Figure 2i). The TAG content decreased significantly in the *GDSL-1*-overexpressing cell line (Figure 2j), revealing that *GDSL-1* influences TAG hydrolysis in longan. Overall, *GDSL-1* and *GDSL-2* may encode proteins that affect the early somatic embryogenesis of longan by participating in TAG hydrolysis.

Autophagy was activated in the longan EC

The scRNA-seq analysis revealed the enrichment of an 'autophagy' pathway in the EC (Figure S8). More specifically, *Autophagy related gene 8C-1 (ATG8C-1)* expression was upregulated in clusters 0, 1, 9, and 10 (Figure S11), and *ATG8I*, *ATG4A*, and *ATG8C-2* expression levels were upregulated in clusters 2, 5, 10, and 11 (Figure S11). These findings implied that the degradation of proteins and organelles via autophagy in the somatic embryo of longan helped maintain cellular homeostasis or initiate cellular apoptosis. To explore whether autophagy can affect the early somatic embryogenesis of longan, rapamycin (RAPA) (Yin et al., 2017) and 3-methyladenine (3-MA), which inhibit and promote autophagy, respectively (Wu et al., 2010), were added to the MS solid medium, after which cell division and differentiation in the early developing somatic embryos were observed (Figure S12). Compared with the control, both RAPA and 3-MA inhibited the cell differentiation in the early developing somatic embryos of longan. Hence, the EC did not differentiate into a globular embryo after 12 days of RAPA or 3-MA treatment. These results indicated that the stability of cell autophagy activity is critical for the early somatic embryogenesis of longan. Excessive or insufficient autophagy activity will alter the early somatic embryogenesis of longan.

Reconstruction of the continuous cell differentiation trajectories in the longan EC

Because scRNA-seq analysis is applicable for analysis of cells in intermediate states, we extrapolated the continuous cell differentiation trajectories over the whole developmental process. To correctly order the EC cells along a reconstructed differentiation trajectory, we used Monocle2 to determine the unclear cell fates. We selected the identified cell populations in clusters 7, 8, 9, and 11 to construct the cell differentiation trajectory. Cells from clusters 8 and 9 were assembled at the beginning of the pseudotime (Figure 3a,b). Consistent with this assembly, the expression levels of multiple cell cycle-related genes, including *TUA6-1*, *CYCLIN B1;3 (CYCB1;3)*, *CYCB2;2*, *CDKB2;1*, *CYCB2;4*, *CYCB1;4*, and *CELL DIVISION CYCLE PROTEIN CODING GENE 2 (CDC2)*, were upregulated in cluster 8 (Figure 3c,d; Figure S13; Table S4), and most of these cell cycle-related genes were specific to the G2/M phase (Van Leene et al., 2010). In addition, the representative somatic embryogenesis marker genes, such as *AGP20* (cluster 8) (Zhong et al., 2011) and *LTP4* (cluster 9), had increased expression levels in the respective clusters (Figure 4a). Moreover, the expression of most of the top 20 upregulated genes in clusters 8 (18) and 9 (13) tended to decrease significantly from the EC stage to the globular embryo stage according to the bulk RNA-seq data for longan early somatic embryogenesis (Figure 4b–

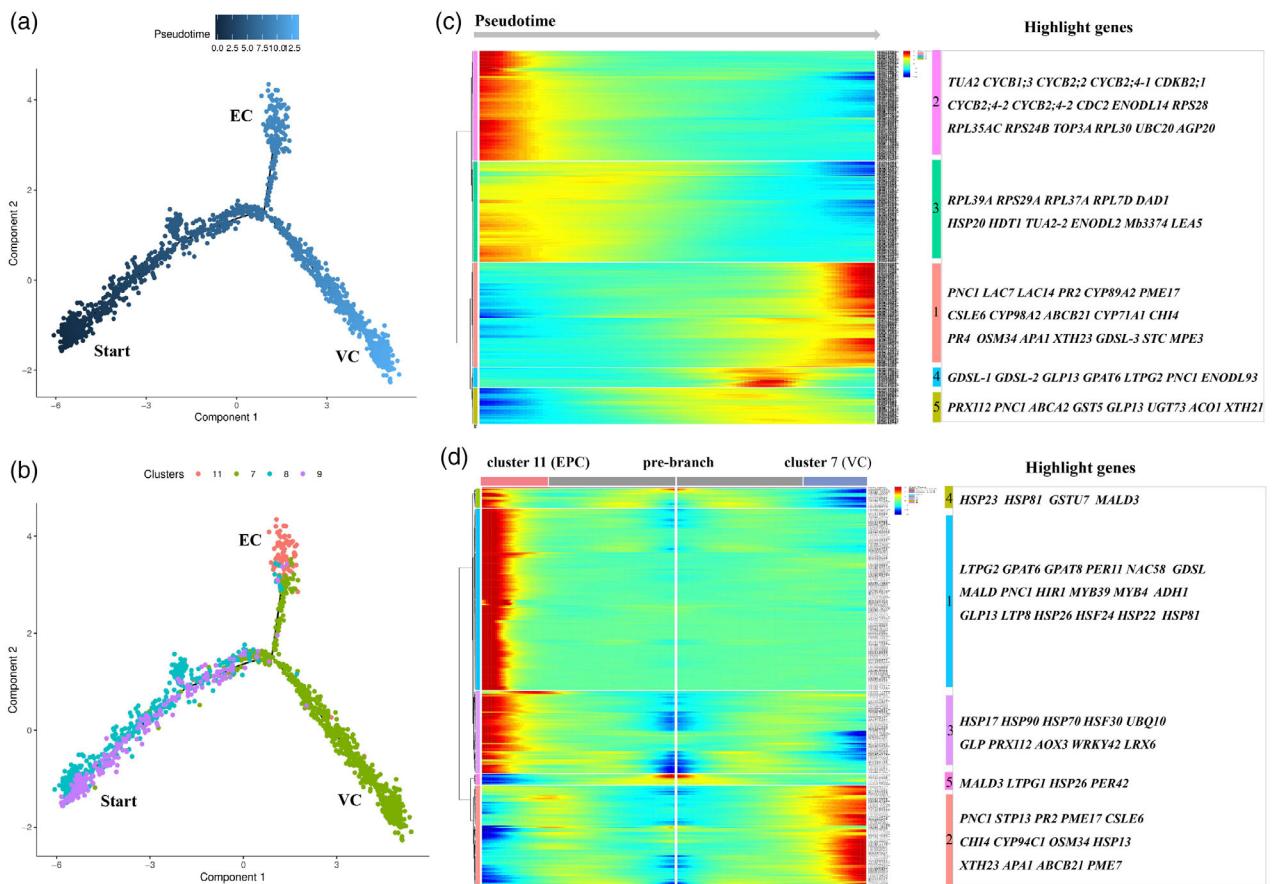


Figure 3. Continuous cell differentiation trajectory of the embryogenic callus.

(a, b) Simulation of the continuous differentiation trajectory of the vascular and epidermal cells over the pseudo-timeline. 'Start' denotes the beginning of the pseudo-timeline. The pseudo-timeline increases along with the differentiation trajectory. VC, vascular cell; EPC, epidermal cell. Each dot denotes a single cell.

(c) Heatmaps presenting the expression of pseudo-timeline-dependent genes over the pseudo-timeline. The pseudo-timeline increases along with the direction indicated by the arrow. The representative pseudo-timeline-dependent genes are presented on the right. The colored bar indicates relative expression levels. The gene list is given in Table S4.

(d) Heatmaps presenting the expression of branch-dependent genes over the pseudo-timeline. The pre-branch (in the middle of the heatmap) indicates the beginning of the pseudo-timeline. The pseudo-timeline extends on both sides of the heatmap. The representative branch-dependent genes are presented on the right. The colored bar indicates relative expression levels. VC, vascular cell; EPC, epidermal cell. The gene list is given in Table S4.

d; Figure S14; Table S5), suggesting that these genes in clusters 8 and 9 encode proteins with important functions during the early somatic embryogenesis of longan, especially in the EC stage. In Arabidopsis early pro-embryos, the upregulated genes in the apical cell lineage in the one-cell embryo stage are mainly related to the 'mitotic cell cycle' pathway (Zhou et al., 2020), indicating that cell division is a key characteristic of early embryonic cells. Hence, the cluster 8 cells likely formed the early embryogenic cell group in longan. Additionally, in contrast to the one-cell embryo stage, the upregulated Arabidopsis genes in the 32-cell embryo stage are mainly associated with the 'meristem development' pathway (Zhou et al., 2020). Furthermore, the shoot and root apical meristem cells form in the early globular embryo stage (Ten Hove et al., 2015), implying that meristem development-related genes are actively expressed at the relatively late stage of

early pro-embryo development. The upregulated expression of *RGF3* in cluster 9 suggested the cells in this cluster represent the root apical-like meristem cell population during the early somatic embryogenesis of longan. Thus, the enhanced expression of the genes related to the cell cycle and the meristem in clusters 8 and 9 was in accordance with the gene expression changes in Arabidopsis early pro-embryos. On the basis of these findings, clusters 8 and 9 likely comprised the early embryonic cell populations contributing to the early somatic embryogenesis in longan. Notably, the expression of the histone deacetylation modification gene *HISTONE DEACETYLASE 1* (*HDT1*) was upregulated in cluster 8, suggesting that the encoded protein may have important roles in the early embryonic cell group. In Arabidopsis, *HDT1* accumulates in somatic embryos and the induced expression of *HDT1* may be correlated with early somatic embryo

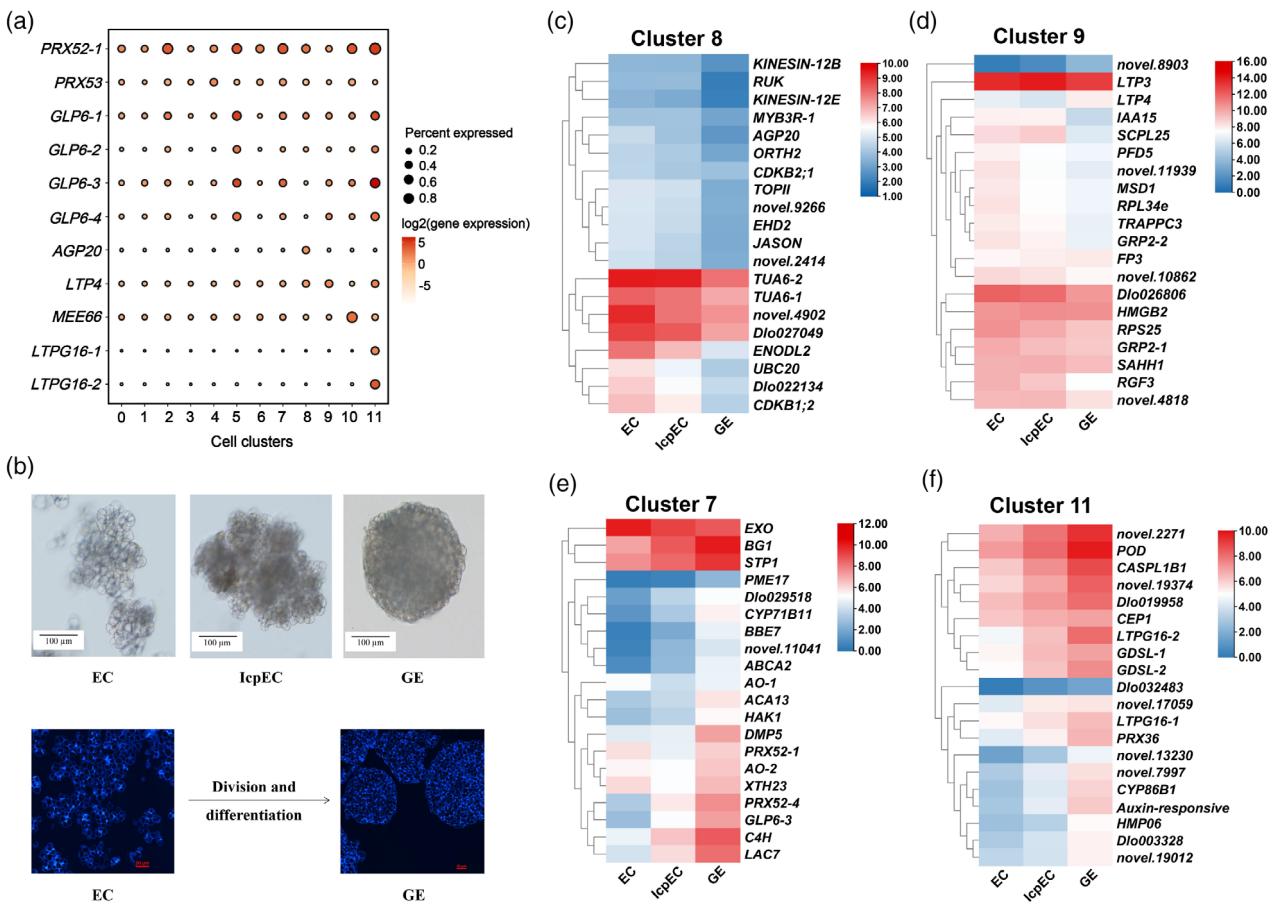


Figure 4. Expression patterns of somatic embryogenesis-related genes with cluster-enriched expression. (a) Expression patterns of somatic embryogenesis-related marker genes in the single-cell transcriptome. Dot diameter and color indicate expression levels. The full names of selected genes are given in the main text. (b) Morphological characteristics of the longan early somatic embryos. Photographs in the upper row were taken using a microscope (Olympus). Photographs of the paraffin sections in the lower row were taken using a fluorescence microscope (Nikon). Paraffin section thickness = 4 µm. Sections were counterstained with DAPI to visualize the cell nucleus (blue). EC, embryogenic callus; IcpEC, incomplete compact pro-embryogenic culture; GE, globular embryo. Bars = 100 and 50 µm. (c–f) Heatmaps presenting the expression patterns of the top 20 upregulated genes (clusters 8, 9, 7, and 11) during early somatic embryogenesis (embryogenic callus, incomplete compact pro-embryogenic culture, and globular embryo stages). The log₂-transformed (FPKM + 1) values were used to construct the heatmap. The colored bar indicates relative expression levels. EC, embryogenic callus; IcpEC, incomplete compact pro-embryogenic culture; GE, globular embryo.

development (Zhou et al., 2004). Therefore, we predicted that histone deacetylation may influence the early somatic embryo development in longan.

A continuous cell differentiation trajectory was observed for clusters 8, 9, 7, and 11. Cluster 7 was associated with the middle part of the pseudo-timeline (Figure 3a, b). The expression levels of the vascular cell-associated genes were upregulated in cluster 7, including *POD*, *LAC*, *PECTINESTERASE* (*PME*), and *CELLULOSE SYNTHASE-LIKE E* (*CSLE*) (Figure 3c,d; Figure S13; Table S4), implying that the vascular cells were undergoing lignin deposition and cell wall thickening. In addition, the expression of representative somatic embryogenesis marker genes, including *PEROXIDASE 52* (*PRX52*) and *GLP6* (Mathieu et al., 2006), increased in cluster 7 (Figure 4a). Additionally, the

expression levels of 16 of the top 20 upregulated genes in cluster 7 tended to increase significantly from the EC stage to the globular embryo stage according to the bulk RNA-seq results (Figure 4e; Figure S14; Table S5). Furthermore, the 10 DEGs related to the ‘phenylpropanoid biosynthesis’ pathway had significantly upregulated expression levels from the EC stage to the globular embryo stage (Figure S15; Table S6), indicating that the cluster 7 cells were differentiating during the somatic embryogenesis of longan. In the globular embryo stage, lignin biosynthesis and secondary cell wall thickening may be occurring in cells. On the basis of their increased expression in the globular embryo stage, we speculated that cluster 7 contained the differentiated cell populations mediating the later stage of somatic embryogenesis in longan.

Unlike clusters 8, 9, and 7, cluster 11 was associated with a later stage in the pseudo-timeline (Figure 3a,b). The genes related to the formation of the epidermis and fatty acid metabolism (i.e., *LTP* and *GDSL*) were specifically expressed in cluster 11 (Figure 3c,d; Figure S13; Table S4), further revealing the active metabolism of fatty acids in the epidermal cell population. Interestingly, cluster 11 was scattered on the UMAP plot (Figure 1b), and the top 20 upregulated genes were specifically expressed in cluster 11 (Figure S16), probably because the cluster 11 epidermal cells were terminally differentiated with specialized transcriptional patterns. The representative somatic embryogenesis marker genes *LTPG16-1* and *LTPG16-2* (cluster 11) (Vroemen et al., 1996) had upregulated expression levels in cluster 11 (Figure 4a). Additionally, the expression of 19 of the top 20 upregulated expression genes in cluster 11 tended to increase significantly from the EC stage to the globular embryo stage according to the bulk RNA-seq analysis (Figure 4f; Figure S14; Table S5). Thus, we predicted that the cluster 11 cells were also differentiating during the somatic embryogenesis of longan. Considered together, these results indicated that clusters 8 and 9 comprised the early somatic embryogenesis-related cell populations, whereas clusters 7 and 11 contained the somatic embryogenesis-related cell populations that differentiated relatively late. Hence, the scRNA-seq analysis provided new insights for reconstructing the continuous differentiation of embryonic cells. Furthermore, unlike the bulk RNA-seq analysis, it enabled an exploration of the transcriptional regulators of the cells in intermediate states.

A highly interconnected TF-target gene regulatory network coordinates somatic embryogenesis-related cell differentiation in longan

To explore TF gene expression dynamics during embryonic cell differentiation, we constructed a regulatory network consisting of differentially expressed TFs and their target genes that were differentially expressed in clusters 8, 9, 7, and 11 (Figure 5; Table S7). A total of 39 differentially expressed TF genes were identified. Key highly connected TF genes were detected throughout the cell differentiation process, including *PHR1-like 1* (*PHL1*), *ETHYLENE-RESPONSIVE FACTOR 114* (*ERF114*), *ERF6*, *MYB9*, *MYB4*, *NAC58*, and *WRKY75* (Figure 5a). Among the differentially expressed TFs, the genes encoding *NAC58*, *MYB9*, and *MYB4* were specifically expressed in cluster 11 (Table S7), suggesting they may contribute to the epidermal cell differentiation during the somatic embryogenesis of longan. To further elucidate the TF-target gene regulatory network during embryonic cell differentiation, we used the identified DEGs in the continuous cell differentiation trajectory of clusters 8, 9, 7, and 11 as the candidate target genes of the differentially expressed TFs. A total of 90 genes were targeted by 14 TFs (Figure 5b; Table S7). We

examined the regulatory relationship of two TF-target gene pairs (*NAC83-XTH23* and *NAC58-MYB9*) using longan protoplasts in a transient overexpression system (Figure 5c-g). The *XTH23* expression level was increased significantly in the *NAC83*-overexpressing protoplasts (Figure 5d,e), confirming that *NAC83* promotes the expression of *XTH23*. In addition, *NAC58* overexpression increased the expression of *MYB9* (Figure 5f,g), indicating that *NAC58* promotes the expression of *MYB9*. Recent research revealed that overexpression and silencing of *MYB9A1* and *MYB9A2* in *Phalaenopsis aphrodite* (Orchidaceae) lead to the flattening of epidermal cells and changes to the epicuticular wax composition and content (Lu et al., 2022), suggesting that *MYB9* can regulate the formation of epidermal cells and epidermal wax. In the current study, *NAC58* and *MYB9* were specifically expressed in cluster 11, reflecting the importance of the *NAC58-MYB9* relationship for the differentiation of epidermal cells during the somatic embryogenesis of longan; however, this process should be characterized more thoroughly in future studies. Overall, our network revealed the complex mechanism mediating the transition of early embryonic cells to differentiated cells during the somatic embryogenesis of longan.

Overexpression of *ERF6* inhibited the differentiation of longan somatic embryos under high-temperature stress conditions

By analyzing the scRNA-seq results, we identified three ERF TF-encoding genes (i.e., *ERF6*, *ERF020*, and *ERF114*) with upregulated expression levels. The APETALA2/ERF (AP2/ERF) superfamily is one of the largest TF families in plants, modulating plant growth and development, stress resistance, and metabolism by modifying the expression of their target genes (Cai et al., 2014; Qin et al., 2007; Zhang, Yin, et al., 2020). In the present study, the longan EC was stably transformed with the *ERF6* overexpression (*ERF6-OE*) vector (i.e., pCAMBIA1301-35S-*ERF6-GUS*) to explore whether *ERF6* can affect the early somatic embryogenesis of longan (Figure 6). The upregulated expression of *ERF6* was detected in two *ERF6-OE* cell lines (i.e., *ERF6-OE1* and *ERF6-OE2*) (Figure 6a,b). The successful transformation of these transgenic lines was confirmed by GUS staining (Figure S17; Figure 6c). To observe the phenotypic changes in the transgenic cell lines, we selected the *ERF6-OE2* cell line for the induction of early somatic embryogenesis because its *ERF6* expression level was upregulated by more than 15 times (Figure 6b). Compared with the wild-type (WT) control, the *ERF6-OE* EC differentiated to form the incomplete compact pro-embryogenic culture, the early globular embryo, and the globular embryo after 6, 9, and 12 days of growth at 25°C (i.e., normal growth temperature) (Figure 6c; Figure S18), with a globular embryo formation rate of approximately 50% after 12 days of growth (Figure 6d), indicating that the longan EC was able to

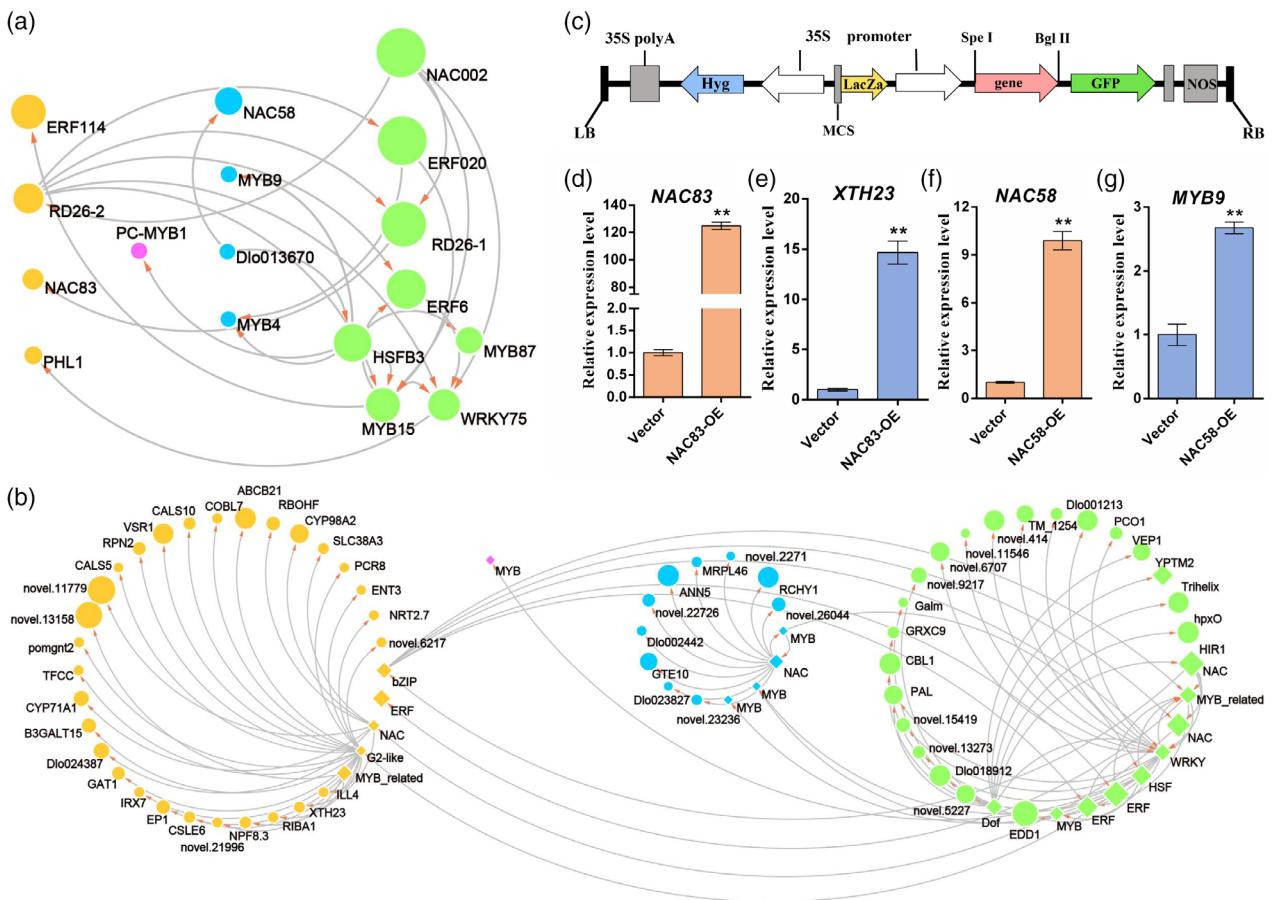


Figure 5. Regulatory networks comprising differentially expressed transcription factors and the corresponding upregulated differentially expressed target genes during somatic embryo cell differentiation.

(a) Regulatory network comprising transcription factors expressed dynamically across the pseudo-timeline. Node colors indicate different clusters. Node sizes represent gene expression levels.

(b) Transcription factor-target gene regulatory network constructed on the basis of the differentially expressed transcription factors and differentially expressed target genes across the pseudo-timeline. Node colors indicate different clusters. Node sizes represent gene expression levels. Rhombus-shaped and round nodes represent transcription factors and target genes, respectively.

(c) Diagram of the pCAMBIA1301-35S-GFP overexpression vector.

(d–g) Expression patterns of the NAC83-XTH23 and NAC58-MYB9 regulatory networks in the protoplasts transiently overexpressing NAC83 and NAC58 as determined by qRT-PCR. NAC83-OE, protoplasts transiently overexpressing NAC83; NAC58-OE, protoplasts transiently overexpressing NAC58. Student's t-test, **P < 0.01.

undergo normal somatic embryogenesis following *ERF6* overexpression.

As a temperature-sensitive fruit tree species, continuous exposure to high temperatures (34–38°C) will result in abnormal development of the longan embryo, resulting in arrested embryo development and fruit abortion, thereby leading to decreased fruit yield and quality (Nong et al., 2006). An analysis of the transcriptome datasets for longan ECs incubated at different temperatures (15°C, 25°C, and 35°C) revealed that *ERF6* was obviously upregulated in response to high-temperature stress (Figure S19), implying that *ERF6* may be important for the resistance of longan to excessive heat. To further clarify the functions of *ERF6* in longan, early somatic embryogenesis was induced in the *ERF6*-OE2 cell line incubated at 35°C (i.e., high-temperature

stress). The WT EC did not differentiate into the globular embryo after a 12-day incubation at 35°C (Figure 6c). Moreover, the *ERF6*-OE2 EC did not differentiate into the incomplete compact pro-embryogenic culture after 9 and 12 days under high-temperature stress (35°C). More specifically, compared with the WT control, the somatic embryos of *ERF6*-OE2 were disintegrated and approximately 67% of small loose cells as well as some elongated cells were detected (Figure 6c,e; Figure S18). Therefore, *ERF6* overexpression appeared to increase the inhibitory effect of heat stress during the early somatic embryogenesis of longan.

Melatonin (*N*-acetyl-5-methoxytryptamine; MT), which is a derivative of the essential amino acid tryptophan, reportedly increases the tolerance of plants to various abiotic stresses (e.g., heat, cold, and drought stresses) (Wang

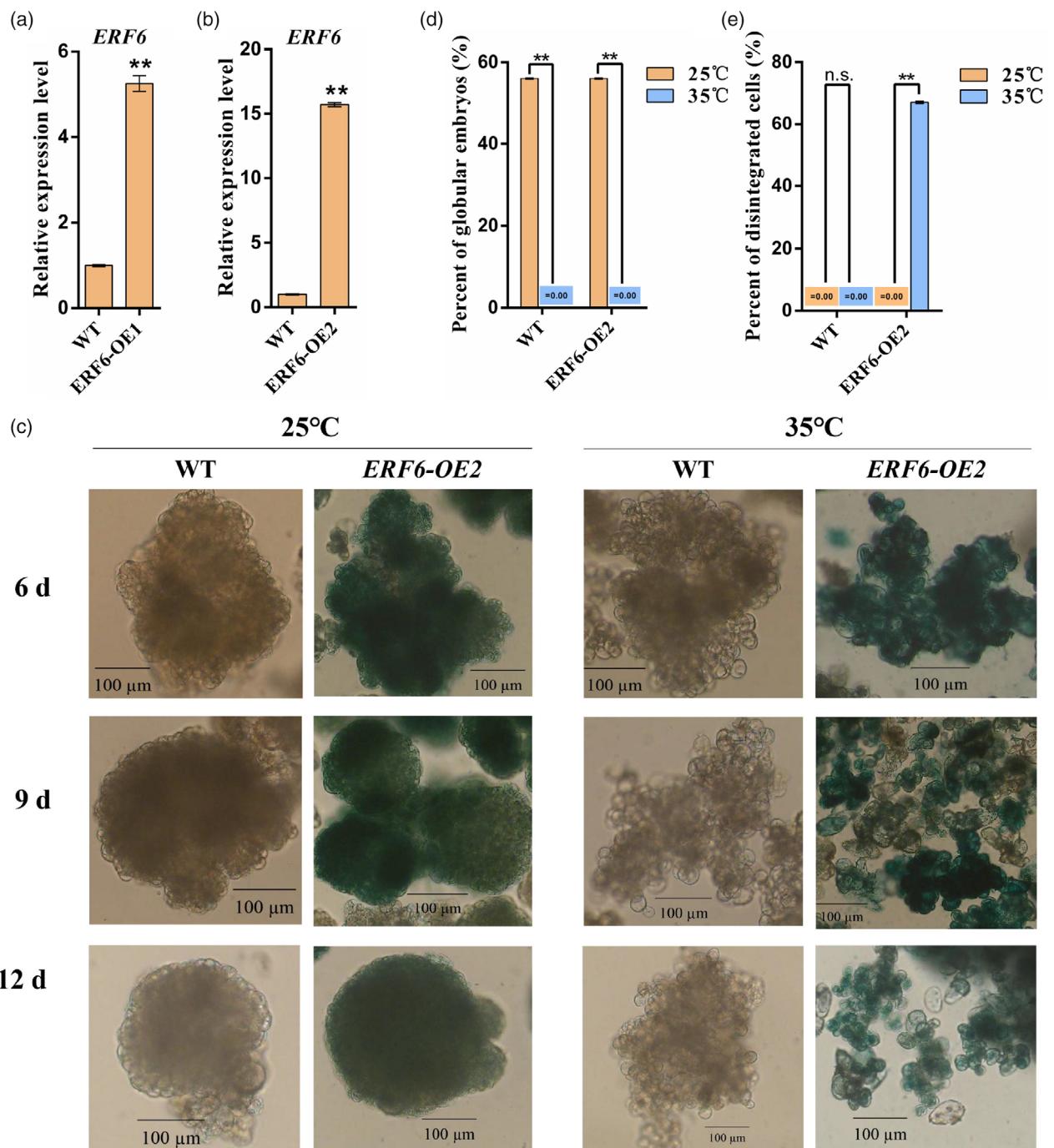


Figure 6. Gene expression patterns and phenotypes of WT and *ERF6*-overexpressing somatic embryos.

(a, b) *ERF6* expression pattern in the *ERF6*-overexpressing embryogenic callus as determined by qRT-PCR. *ERF6*-overexpressing embryogenic callus cultured on MS medium. WT, wild type; *ERF6*-OE1 and *ERF6*-OE2, *ERF6*-overexpressing embryogenic callus cell lines 1 and 2. Student's t-test, ** $P < 0.01$.

(c) Phenotypes of transgenic somatic embryos overexpressing *ERF6*. The transgenic embryogenic callus was cultured on MS solid medium under normal (25°C) and high-temperature stress (35°C) conditions to induce somatic embryos. Samples were harvested after 6, 9, and 12 days and the somatic embryo phenotypes were observed using a microscope (Olympus). Blue represents the GUS staining result. WT, wild type; *ERF6*-OE2, *ERF6*-overexpressing embryogenic callus cell line 2. Bar = 100 μ m.

(d) Percentage of globular embryos of transgenic somatic embryos overexpressing *ERF6*. The transgenic embryogenic callus was cultured on MS solid medium under normal (25°C) and high-temperature stress (35°C) conditions to induce somatic embryos. Samples were harvested after 12 days. WT, wild type; *ERF6*-OE2, *ERF6*-overexpressing embryogenic callus cell line 2. Student's t-test, ** $P < 0.01$.

(e) Percentage of disintegrated cells of transgenic somatic embryos overexpressing *ERF6*. The transgenic embryogenic callus was cultured on MS solid medium under normal (25°C) and high-temperature stress (35°C) conditions to induce somatic embryos. Samples were harvested after 12 days. WT, wild type; *ERF6*-OE2, *ERF6*-overexpressing embryogenic callus cell line 2. Student's t-test, ** $P < 0.01$.

et al., 2018). To further characterize the functions of *ERF6* under high-temperature stress conditions, the *ERF6*-OE EC was transferred to MS medium supplemented with 4.5 μ M MT (Figure 7). The cell disintegration phenomenon of the *ERF6*-OE somatic embryos was significantly improved under high-temperature stress conditions (Figure 7a; Figure S20). The somatic embryo differentiation process was similar to that observed for the heat-stressed (35°C) WT control, which differentiated into the incomplete compact pro-embryogenic culture, but did not differentiate further to produce the globular embryo, with the percentage of disintegrated cells decreasing to 14% (Figure 7b,c; Figure S20). Accordingly, MT treatment decreased the sensitivity of the *ERF6*-OE somatic embryos to heat stress.

To further explore the molecular mechanism underlying the effect of *ERF6* on the early somatic embryogenesis of longan under high-temperature stress conditions, we performed a comparative transcriptome analysis of the

WT, *ERF6*-OE1, and *ERF6*-OE2 ECs incubated at 25°C (room temperature) or 35°C (heat stress) and treated with MT. The DEGs in the *ERF6*-OE1 and *ERF6*-OE2 ECs were identified (Figure 8a). The Pearson correlation coefficient (greater than 0.9) revealed the high correlation among the biological replicates (Figure 8b). Therefore, the two biological replicates were combined for the subsequent analysis. The transcriptome analysis confirmed that *ERF6* was more highly expressed in the *ERF6*-OE ECs than in the WT control and that *ERF6* expression was upregulated in all of the samples exposed to heat stress (Figure 8c). Moreover, *ERF6* expression was downregulated in the heat-stressed samples treated with MT (Figure 8c). This result was consistent with the observed phenotypic changes, suggesting that MT inhibited *ERF6* expression, thereby significantly affecting somatic embryogenesis under high-temperature stress conditions. In addition, the GO annotations indicated that the expression levels of the DEGs associated with the

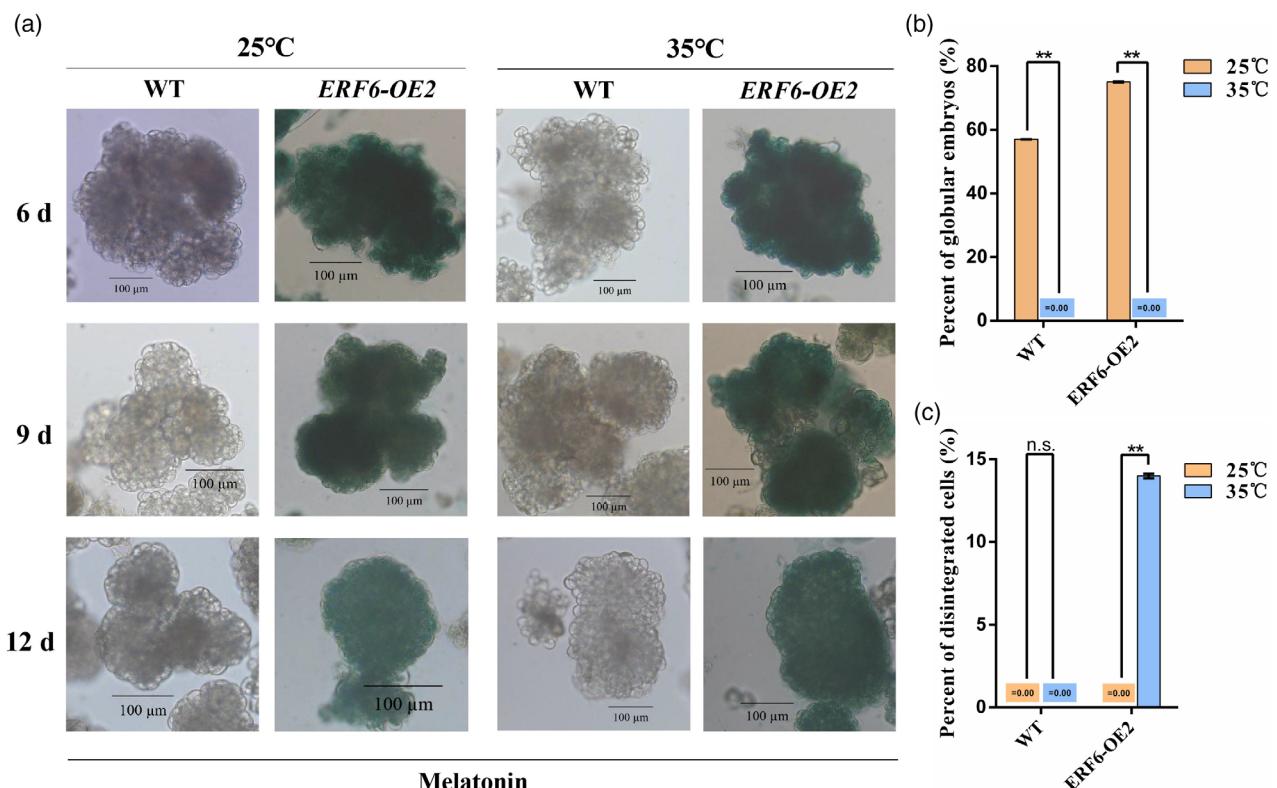


Figure 7. Phenotypes of the WT and *ERF6*-overexpressing somatic embryos treated with melatonin.

(a) Phenotypes of transgenic somatic embryos overexpressing *ERF6* treated with melatonin. The transgenic embryogenic callus was cultured on MS solid medium supplemented with 4.5 μ M melatonin under normal (25°C) and high-temperature stress (35°C) conditions to induce somatic embryos. Samples were harvested after 6, 9, and 12 days and the somatic embryo phenotypes were observed using a microscope (Olympus). Blue represents the GUS staining result. WT, wild type; *ERF6*-OE2, *ERF6*-overexpressing embryogenic callus cell line 2. Bar = 100 μ m.

(b) Percentage of globular embryos of transgenic somatic embryos overexpressing *ERF6* treated with melatonin. The transgenic embryogenic callus was cultured on MS solid medium supplemented with 4.5 μ M melatonin under normal (25°C) and high-temperature stress (35°C) conditions to induce somatic embryos. Samples were harvested after 12 days. WT, wild type; *ERF6*-OE2, *ERF6*-overexpressing embryogenic callus cell line 2. Student's *t*-test, ** $P < 0.01$.

(c) Percentage of disintegrated cells of transgenic somatic embryos overexpressing *ERF6* treated with melatonin. The transgenic embryogenic callus was cultured on MS solid medium supplemented with 4.5 μ M melatonin under normal (25°C) and high-temperature stress (35°C) conditions to induce somatic embryos. Samples were harvested after 12 days. WT, wild type; *ERF6*-OE2, *ERF6*-overexpressing embryogenic callus cell line 2. Student's *t*-test, ** $P < 0.01$.

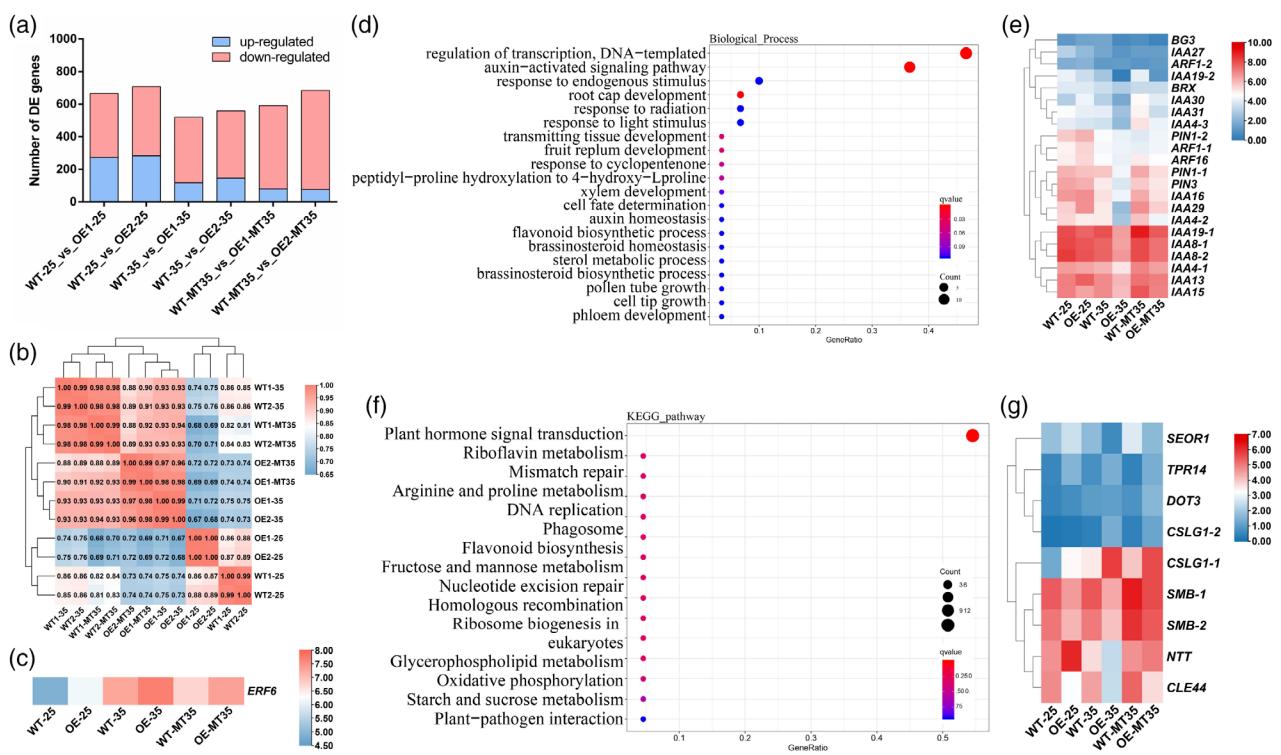


Figure 8. Analysis of differentially expressed genes in the *ERF6*-overexpressing embryogenic callus.

WT and *ERF6* transgenic embryogenic calli were cultured on MS solid medium supplemented with 0 or 4.5 µM melatonin under normal (25°C) and high-temperature stress (35°C) conditions. Samples were harvested after a 24-h treatment for the bulk RNA-seq analysis.

(a) Number of differentially expressed genes (DEGs) between WT and *ERF6*-overexpressing embryogenic callus cell lines after the high-temperature and melatonin treatments.

(b) Pearson correlation analysis of the bulk RNA-seq samples to assess the reproducibility of the biological replicates.

(c) *ERF6* expression patterns in the *ERF6*-overexpressing embryogenic callus transcriptome dataset. The log₂-transformed (FPKM + 1) values were used to construct the heatmap. The colored bar indicates relative expression levels.

(d) GO enrichment analysis of the upregulated DEGs between the OE-35 and OE-MT35 cell lines. The size of the dots represents the number of genes associated with the GO term and the color of the dots represents the *q* value.

(e) Expression patterns of the 'auxin-activated signaling' pathway DEGs in the *ERF6*-overexpressing embryogenic callus transcriptome dataset. The log₂-transformed (FPKM + 1) values were used to construct the heatmap. The colored bar indicates relative expression levels.

(f) KEGG enrichment analysis of the upregulated DEGs between the OE-35 and OE-MT35 cell lines. The size of the dots represents the number of genes associated with the enriched KEGG pathway and the color of the dots represents the *q* value.

(g) Expression patterns of the plant growth and development pathway DEGs in the *ERF6*-overexpressing embryogenic callus transcriptome dataset. The log₂-transformed (FPKM + 1) values were used to construct the heatmap. The colored bar indicates relative expression levels. WT-25, wild-type embryogenic callus cultured under normal temperature (25°C) conditions; OE-1-25 and OE-2-25, *ERF6*-overexpressing embryogenic callus cell lines 1 and 2 cultured under normal temperature (25°C) conditions; WT-35, wild-type embryogenic callus cultured under high-temperature stress (35°C) conditions; OE-1-35 and OE-2-35, *ERF6*-overexpressing embryogenic callus cell lines 1 and 2 cultured under high-temperature stress (35°C) conditions; WT-MT35, wild-type embryogenic callus treated with melatonin and cultured under high-temperature stress (35°C) conditions; OE-1-MT35 and OE-2-MT35, *ERF6*-overexpressing embryogenic callus cell lines 1 and 2 treated with melatonin and cultured under high-temperature stress (35°C) conditions.

'auxin-activated signaling' pathway were downregulated in the *ERF6*-OE EC exposed to heat stress (Figure S21; Table S8), whereas the expression of most of these genes was upregulated in the heat-stressed *ERF6*-OE EC treated with MT (Figure 8d,e; Tables S9 and S10). Hence, *ERF6* may mediate the differentiation of longan somatic embryos under high-temperature stress conditions by inhibiting the expression of 'auxin-activated signaling' pathway genes. Additionally, the expression levels of the DEGs related to plant growth and development (e.g., 'root cap development', 'transmitting tissue development', and 'fruit replum development'), 'phagosome', 'nucleotide excision

repair', and 'mismatch repair' pathways were upregulated in the heat-stressed *ERF6*-OE EC treated with MT (Figure 8d,f,g; Tables S9, S11, and S12), but they were significantly downregulated in the *ERF6*-OE EC exposed to heat stress (Figure 8g), suggesting that *ERF6* inhibited the differentiation of longan somatic embryos under high-temperature stress conditions by inhibiting the expression of genes mediating plant growth and development. The MT treatment restored this process by restricting *ERF6* expression. Overall, the results indicated that *ERF6* is sensitive to high temperatures and negatively affects the somatic embryogenesis of longan exposed to heat stress.

DISCUSSION

Generation of a longan EC cell atlas

In this study, our scRNA-seq analysis helped clarify the cell types in the longan EC. We identified 12 cell clusters containing proliferating cells, meristematic cells, vascular cells, and epidermal cells. Each cluster was defined by the expressed marker genes or by the enriched biological processes. The identified cell types were similar to the findings of an earlier study in which the scRNA-seq analysis of the *Arabidopsis* callus detected 10 cell clusters, of which clusters 0–5 comprised callus cells, including epidermis-like cells, lateral root cap-like cells, and vascular initial-like cells (Zhai & Xu, 2021). Our scRNA-seq data for the longan EC provide the foundation for future investigations into longan embryo development at the cellular level. On the basis of the identified TFs and their target genes dynamically expressed during cell differentiation, a highly interconnected network regulating the somatic embryogenesis of longan was established. The results of this study may be useful for future research into longan embryo development and the pluripotency of embryonic cells in longan.

Interestingly, our scRNA-seq results revealed that clusters 5 and 7 were enriched in vascular cell marker genes, including *GLP* genes and *LAC7*. Additionally, 'phenylpropanoid biosynthesis' was an enriched KEGG pathway among the upregulated genes in these two clusters. We predicted that cluster 5 may also consist of a vascular cell population. However, the genes related to the 'glutathione metabolism' pathway were enriched only in cluster 5, whereas the genes related to the 'amino sugar and nucleotide sugar metabolism' pathway were enriched only in cluster 7. Both clusters 6 and 8 were enriched with upregulated cell cycle-related genes and the 'ribosome' KEGG pathway genes, but some of the genes in cluster 8 were specifically involved in the 'phagosome' pathway. These findings indicated that the cells in clusters 5 and 7 and the cells in clusters 6 and 8 have similar characteristics, but the cells were also unexpectedly heterogeneous. Considered together, these results reflect the substantial cell heterogeneity in the long EC.

Stability of cell autophagy level is critical for the somatic embryogenesis of longan

Autophagy is a conserved biological process in all eukaryotes that ensures proper cellular development. Autophagy-related genes mediate the degradation of cytoplasmic components to maintain cellular homeostasis or cellular apoptosis (Hu et al., 2019). Moreover, autophagy modulates various biological processes, including cellular proliferation and differentiation, intracellular renewal, and metabolic regulation (Galluzzi & Green, 2019). It is also critical for stem cell maintenance and affects cell differentiation (Vessoni et al., 2012). The recent clarification of the

transcriptional regulation of autophagy-related genes in early human embryonic cells at single-cell resolution revealed that these genes are activated in the early human embryo and encode proteins with fundamental roles affecting early embryo development (Song et al., 2022). In the present study, 'autophagy' was an enriched pathway in different cell types during the somatic embryogenesis of longan. The results showed that a stable autophagy level is necessary for longan somatic embryogenesis. Hence, scRNA-seq can quickly screen out important biological processes related to tissue development, thereby providing the basis for future studies on the role of various biological processes such as lignin and fatty acid metabolism during the somatic embryogenesis of longan.

scRNA-seq facilitates the screening of key genes affecting the somatic embryogenesis of longan

Longan somatic embryogenesis marker genes were identified in an earlier study (Chen, Xu, et al., 2020). However, a bulk RNA-seq analysis of specific developmental stages failed to reveal the molecular basis of continuous cell differentiation or the most significant factors. The marker genes and TFs with cluster-enriched expression identified during our scRNA-seq analysis were used as new marker genes and TFs for establishing a continuous transcriptional regulatory network governing the somatic embryogenesis of longan. We precisely characterized the functions of the upregulated genes and TFs in the somatic embryogenesis-related cell populations and identified numerous marker genes and TFs, such as *AGP20*, *TUA6*, *UBIQUITIN-CONJUGATING ENZYME 20* (*UBC20*), *GLP*, *LAC7*, *GDSL*, *LTP*, *MEE66*, *POD*, *ERF*, *MYB*, and *NAC* genes. Fatty acid and TAG contents increase during plant embryo development and seed maturation processes, thereby ensuring there is a sufficient supply of energy. Transcriptional reprogramming is an important mechanism affecting the oil accumulation process. In the *Arabidopsis myb89* and *myb89-1* mutants, oil and major fatty acids accumulate significantly in seeds, suggesting that MYB89 is a crucial inhibitor of seed oil accumulation (Li et al., 2017). In *Arabidopsis*, MYB96 can activate the expression of *acyl-CoA:diacylglycerol acyltransferase 1* (*DGAT1*) and *phospholipid:diacylglycerol acyltransferase 1* (*PDAT1*) in the TAG biosynthesis pathway and regulate seed oil accumulation (Lee et al., 2018). In the present study, we functionally characterized *GDSL* genes, which were the marker genes of epidermal cells in the EC of longan. The *GDSL* genes were involved in the hydrolysis of TAG during the early somatic embryogenesis of longan, but the functions of the proteins encoded by these genes during longan somatic embryogenesis will need to be more thoroughly characterized in future studies. Additionally, our scRNA-seq analysis revealed the upregulated expression of a *MATERNAL EFFECT EMBRYO ARREST* gene (*MEE66*). The *Arabidopsis*

mee66 mutant exhibits defective embryo development during the two-cell embryo stage. Furthermore, MEE45 controls seed size by activating AINTEGUMENTA, whereas other MEEs have different key functions influencing endosperm development and embryogenesis (Li et al., 2021; Pagnussat et al., 2005). Therefore, *MEE66* may encode a key regulator of longan somatic embryogenesis.

In Arabidopsis, *ERF6* overexpression adversely affects leaf growth and cell proliferation by modulating gibberellin and DELLA signaling. More specifically, *ERF6* overexpression leads to decreases in the leaf area and plant size as well as abnormal rosette development, indicating that *ERF6* expression is negatively correlated with Arabidopsis growth (Dubois et al., 2013). When both *ERF6* and *ERF11* are overexpressed in Arabidopsis, the decrease in plant size due to *ERF6* overexpression is undetectable and the rosette area is comparable to that of the WT control. Thus, *ERF6* or *ERF11* overexpression has detrimental effects on plant growth, but co-overexpression of these two genes leads to relatively normal growth (Dubois et al., 2015). In the current study, we observed that the EC differentiated to form the globular embryo in *ERF6*-OE samples incubated at a normal temperature. However, under heat stress conditions, *ERF6* overexpression inhibited the expression of genes in 'auxin-activated signaling' and plant growth and development pathways, resulting in considerably inhibited somatic embryo differentiation. The scRNA-seq analysis of Arabidopsis revealed that the upregulated expression of the auxin biosynthesis gene *TRYPTOPHAN AMINOTRANSFERASE OF ARABIDOPSIS1* (*TAA1*) promotes endogenous auxin production in calli, with the resulting high auxin level potentially contributing to the maintenance of the pluripotency of callus cells (Zhai & Xu, 2021). Moreover, we previously determined that the endogenous auxin level is high during the somatic embryogenesis of longan, indicating that auxin accumulation is essential for the differentiation of somatic embryos (Lai & Chen, 2002). Overall, our study showed that *ERF6* is a heat-sensitive TF in longan and that endogenous auxin homeostasis may be critical for the somatic embryogenesis of longan.

Since the cell division and differentiation in the somatic embryos of most flowering plants may appear to be chaotic and random events (Ten Hove et al., 2015), unlike other plant tissues (e.g., leaf, stem, and root), the distribution of the same type of cells in the somatic embryo is disordered and difficult to distinguish. Moreover, longan ECs are loose and it is difficult to fix structures, making it difficult to verify cell clusters in the somatic embryos by *in situ* hybridization or GFP marker lines. In general, the precise function of the marker genes can be confirmed in regenerating transgenic plants of longan, but we still could not construct an effective system for regenerating transgenic plants of longan. This process should be more thoroughly characterized in future studies.

Nevertheless, compared with bulk RNA-seq, scRNA-seq may be more useful for studying longan somatic embryogenesis because there is no loss of information due to the averaging of signals and it can quickly identify cell type-specific regulators of somatic embryogenesis. The functionally redundant genes could be excluded from our scRNA-seq results to precisely verify the functions of these regulatory factors via reverse genetics. The data generated in these studies may be relevant for elucidating the cell fate determination during longan somatic embryogenesis.

Collectively, the scRNA-seq results described herein provide new spatiotemporal insights into cell division and differentiation in the longan EC (Figure 9). The identified cell clusters together with the upregulated genes and TFs may be relevant for designing new approaches to clarify the developmental and physiological processes involved in the somatic embryogenesis of longan. Moreover, the generated gene expression map will enable us to investigate how embryonic cells differentiate into vascular and epidermal cells. Because of the limited number of identified somatic embryogenesis-related marker genes in longan, we were unable to identify more cell types in the EC. Nevertheless, the results of our scRNA-seq analysis of the longan EC may provide the foundation for future molecular and cellular investigations on the mechanisms mediating longan somatic embryogenesis and zygotic embryogenesis.

EXPERIMENTAL PROCEDURES

Plant materials

Cultivated longan 'Hong He Zi' ('HHZ') EC was cultured as previously described (Lai & Chen, 1997). The longan EC was derived from an immature zygotic cotyledon embryo (40–50 days after flowering) cultured on MS solid medium containing 2.0 mg L⁻¹ 2,4-D for 6 weeks at 25°C in darkness (Lai et al., 1997). To proliferate cells, the EC was cultured on MS solid medium containing 1.0 mg L⁻¹ 2,4-D for 20 days before being harvested. Longan embryogenic suspension cell lines were established as follows. The ECs proliferated for 20 days were transferred to MS liquid medium containing 1.0 mg L⁻¹ 2,4-D, 0.5 mg L⁻¹ kinetin, and 5.0 mg L⁻¹ AgNO₃ and then harvested after 6 days. The harvested materials were transferred to MS liquid medium containing 1.0 mg L⁻¹ 2,4-D and then harvested after 6 days. The embryogenic suspension cells were used for isolating the protoplasts needed for the subsequent scRNA-seq and transient expression analyses. Two biological replicates were prepared for the scRNA-seq analysis (i.e., EC-1 and EC-2). For the bulk RNA-seq analysis, qRT-PCR verification, and fatty acid composition determination, longan early somatic embryos at the EC, incomplete compact protembryogenic culture, and globular embryo stages were cultured on MS solid medium containing 1.0, 0.5, and 0.1 mg L⁻¹ 2,4-D, respectively. Samples were harvested after 22 days. For the DMP treatment, the longan EC was cultured on MS solid medium containing 0 or 10 μM DMP. Samples were harvested after 6, 8, 10, and 12 days. For RAPA treatment, the longan EC was cultured on MS solid medium containing 0, 0.1, 0.5, or 1.0 mg L⁻¹ RAPA. Samples were harvested after 6, 9, and 12 days. For 3-MA treatment, the longan EC was cultured on MS solid medium containing

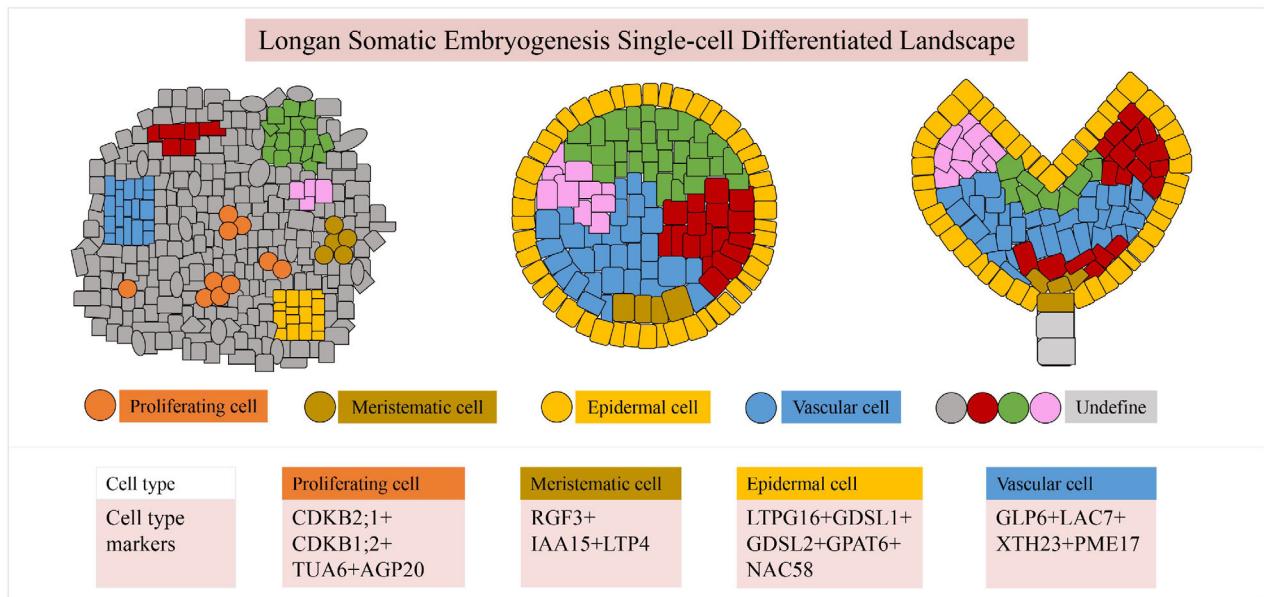


Figure 9. Longan somatic embryogenesis single-cell transcriptional landscape.

Probable longitudinal cross-sections of a developing longan somatic embryo. Cells are colored according to the distinct cell types identified in this study. Representative regulators of distinct cell types in the longan somatic embryo are indicated.

0, 0.1, 0.5, or 1.0 μM 3-MA. Samples were harvested after 6, 9, and 12 days. For the comparative transcriptome analysis, the WT and *ERF6* transgenic ECs were cultured on MS solid medium containing 0 or 4.5 μM MT under normal (25°C) and high-temperature stress (35°C) conditions. Samples were harvested after 24 h.

Protoplast isolation

The longan embryogenic suspension cells were digested for 12 h at 25°C in darkness in an RNase-free enzyme solution (2.0 ml MS liquid medium containing 20.0 g L⁻¹ sucrose and 1.0 mg L⁻¹ 2,4-D and 2.0 ml enzyme mixture containing 12.7% mannitol, 0.011% NaH₂PO₄·2H₂O, 0.36% CaCl₂·2H₂O, 0.12% MES, 1.5% macerozyme, and 1.5% cellulase R10). The protoplasts isolated from the EC were filtered using a cell strainer (40 μm diameter), washed four times with 8% mannitol, and centrifuged (500 g min^{-1}) for 5 min at 25°C. The viability of the collected protoplasts was determined by staining with trypan blue. The protoplast isolation was considered acceptable if more than 90% of the protoplasts were viable. After adjusting the concentration to 1000–5000 protoplasts μl^{-1} , the protoplasts were processed according to the 10x Genomics Single Cell Protocol (10x Genomics, Pleasanton, CA, USA).

scRNA-seq library construction and sequencing

Approximately 29 000 protoplasts of longan ECs were loaded onto a Single Cell Chip (10x Genomics). The scRNA-seq libraries of longan ECs (EC-1 and EC-2) were constructed using a Chromium Controller (10x Genomics) and the Chromium Single Cell 3' Kit (version 3) (10x Genomics). Cell suspensions of longan EC on the Single Cell Chip were loaded on the Chromium Controller to generate single-cell gel-beads-in-emulsion (GEMs). The scRNA-seq libraries of longan EC were prepared using the Chromium Single Cell Library Kit (version 3) (PN-1000078; 10x Genomics) and the Chromium Single Cell 3' GEM Gel Bead Kit (version 3) (PN-1000076; 10x Genomics). Libraries were then sequenced using an Illumina HiSeq 2500 platform (Genedenovo Biotechnology Co., Ltd., Guangzhou, China). The quality of the DNA libraries was

evaluated using the DNA 1000 assay Kit (Agilent Technologies, Santa Clara, CA, USA) and a 2100 Bioanalyzer (Agilent Technologies). The quantitative analysis and pooling were performed using an ABI StepOnePlus Real-Time PCR System (Life Technologies, Carlsbad, CA, USA). The raw scRNA-seq dataset of EC-1 and EC-2 contained Read1, Read2, P5, P7, and i7 index reads. The Single Cell 3' 10x Barcode (16 bp) and UMI (10 bp) were encoded in Read1. Read2 was added during the construction of sequencing libraries to sequence the cDNA fragment. The final libraries of EC-1 and EC-2 contained the P5 and P7 primers for the Illumina bridge amplification. A total of 597 616 669 reads were obtained.

Data quality control and gene expression quantification

The raw scRNA-seq dataset of EC-1 and EC-2 was first analyzed using Cell Ranger (10x Genomics). The Cell Ranger aligner 'STAR' was used to map the reads to the longan genome. The longan genome and GFF annotation files were downloaded from the 'HHZ' longan third-generation genome database (GenBank BioProject ID: PRJNA792504). Cell Ranger aligned the reads to annotated transcripts. Reads uniquely mapped to the 'HHZ' longan transcriptome were used for counting UMIs. The number of UMIs in the genes in each longan EC cell represented the quantitative data for the genes. Of the generated reads, 83.6% were aligned to the 'HHZ' longan reference genome. Seurat was used for data normalization (with the 'Normalization' function) and to ensure the read depth was the same for the two libraries (EC-1 and EC-2) before merging. After equalizing, 518 946 126 reads, 18 064 mean reads per cell, 874 median genes per cell, and 1330 median UMIs per cell were included in further analyses.

Cell clustering analysis

We used the Seurat R package to analyze cell clustering. To exclude the unqualified cells (deads and doublets), we filtered the cells according to the following criteria: 500 < gene counts < 5000 per cell, UMI counts < 20 000 per cell, and percentage of mitochondrial genes < 10%. After filtering, 28 724 cells were retained

for downstream analyses. The gene expression levels were first normalized using the 'LogNormalize' function of Seurat. The normalized data then underwent a principal component analysis (PCA). The significant PCs ($P < 1e-5$) were used for the cell clustering analysis. Cell clustering was implemented according to the graph-based clustering approach of Seurat. The cell clustering dataset was visualized using the UMAP algorithm.

Identification of upregulated DEGs in each cluster

The likelihood-ratio test was used to compare all cells and identify the upregulated DEGs in each cluster. More specifically, the following criteria were applied to identify DEGs: $P \leq 0.01$, $\log_2(\text{fold change}) \geq 0.360674$ (fold change of the average gene expression levels between two groups), and the proportion of cells in which the gene is detected in a specific cluster $> 25\%$.

Marker gene analysis

The top 20 upregulated DEGs in each cluster were selected as the marker genes for further analyses. The expression of key marker genes was visualized using bubble diagrams prepared using TBtools (Chen, Chen, et al., 2020) and OmicShare Tools (<https://www.omicshare.com/tools/>). Bubble diagrams were constructed on the basis of the \log_2 -transformed (expression level + 1) values. The phylogenetic trees were constructed using MEGA11 software with the maximum likelihood algorithm to identify the homologous marker genes between longan and *Arabidopsis*.

KEGG pathway enrichment analysis

The KEGG pathway enrichment analysis was performed using the KEGG database (<https://www.kegg.jp/>) to clarify the biological functions of the upregulated DEGs in each cluster. The significantly enriched KEGG pathways ($q \leq 0.05$) were determined.

Pseudo-timeline analysis

The Monocle2 (version 2.6.4) package was used to analyze single-cell trajectories and to elucidate the cell differentiation process according to the matrix comprising cells and gene expression data (Trapnell et al., 2014). The upregulated DEGs between cell groups were identified and their significance was determined using Monocle2. The key upregulated DEGs related to cell development and differentiation of longan were identified on the basis of a false discovery rate (FDR) of $< 1e-5$. Genes with similar expression trends were grouped together because they might have common biological functions. Branching Expression Analysis Modeling, which applies generalized linear modeling, was used to analyze the pseudo-timeline-dependent or branch-dependent gene expression dynamics. The upregulated DEGs among the pseudo-timeline and branches were filtered (FDR $< 1e-5$).

TF-target gene regulatory network analysis

The upregulated DEGs revealed by the pseudo-timeline analysis were aligned to sequences in the *Arabidopsis* genome database. After selecting the optimally aligned *Arabidopsis* genes as homologs of the longan genes ($e < 1e-5$), the corresponding TFs were used as candidate TFs in this study. Next, the differentially expressed TFs detected following the pseudo-timeline analysis were included in the analysis of interactions conducted using STRING software (<https://string-db.org/>).

The 2000-bp genomic sequence upstream of the start codon was extracted to analyze the binding regions for the differentially expressed TFs. The motif meme of *Arabidopsis* TFs was extracted

to predict the target genes. More specifically, the FIMO program (<https://meme-suite.org/meme/tools/fimo>) was used for predicting target genes. On the basis of the differentially expressed TFs and target genes resulting from the pseudo-timeline analysis, the TF-target gene pairs were selected for further analyses (correlation coefficient > 0.7). The TF-TF and TF-target gene regulatory network diagrams were drawn using Cytoscape.

Bulk RNA-seq library preparation and sequencing

Total RNA was extracted from longan early somatic embryos (EC, incomplete compact pro-embryogenic culture, and globular embryo stages) using TRIzol reagent according to the manufacturer-recommended procedure (Invitrogen, Carlsbad, CA, USA). The quality of the extracted RNA samples was assessed using a NanoPhotometer® spectrophotometer (IMPLEN, Westlake Village, CA, USA) and a 2100 Bioanalyzer (Agilent Technologies). After the mRNA enrichment step using poly-T oligo-attached magnetic beads, sequencing libraries were constructed using the NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, Ipswich, MA, USA) and then sequenced using the Illumina HiSeq X Ten sequencer (Frasergen Bioinformatics Co., Ltd., Wuhan, China). The final library size and quality were checked using a 2100 Bioanalyzer (Agilent Technologies). Quantitative analysis of the libraries was performed using a Qubit 3.0 Fluorometer (Life Technologies). Gene expression was calculated and normalized to fragments per kilobase of exon model per million mapped reads (FPKM) values. DESeq2 was used to identify the DEGs ($\log_2(\text{fold change}) > 1$, $\text{FDR} < 0.05$). The bulk RNA-seq analysis was completed using three biological replicates.

Determination of fatty acid composition and relative content

The fatty acid composition of longan early somatic embryos was determined by analyzing methyl esters using a 7890A gas chromatography system (Agilent Technologies). Briefly, 20 μl internal standard, 2 ml 5% concentrated sulfuric acid/methanol solution, and 300 μl methylbenzene were added to 0.2 g EC, incomplete compact pro-embryogenic culture, and globular embryo samples. The mixtures were placed in a 95°C water bath for 60 min. After cooling to room temperature, 2 ml 0.9% NaCl and 1 ml *n*-hexane were added to the mixtures, which were then centrifuged. For each sample, the supernatant was collected, dried using nitrogen gas, and redissolved in 100 μl *n*-hexane. The supernatant was obtained by centrifugation and transferred to a sample injection bottle. An HP-FFAP column (30 m \times 0.25 mm \times 0.25 μm) and a flame ionization detector were used to determine the fatty acid composition. Gas chromatography analysis was performed using the following parameters: injector and detector temperatures were set at 260°C and 280°C, respectively; the split ratio was 20:1; and the initial column temperature was initially 150°C, increased to 210°C (10°C min $^{-1}$), maintained for 8 min, increased to 230°C (20°C min $^{-1}$), and maintained for 6 min. Each chromatographic peak represented the relative content of the corresponding fatty acid component. The fatty acid determination was completed using three biological replicates.

Determination of the TAG content

The TAG content of longan early somatic embryos was determined using a commercial TAG content determination kit (Comin Biotechnology, Suzhou, China). Briefly, 500 μl Reagent 1 was added to a mortar containing 0.1 g longan somatic embryos, which were ground in an ice bath and then centrifuged

(8000 g min^{-1}) for 10 min. A 50- μl aliquot of the supernatant was transferred to a new centrifuge tube, after which Reagents 2 and 3 were added (475 μl each). The mixtures were left undisturbed at room temperature for 20 min before measuring their absorbance at 505 nm using a microplate reader (Infinite M200 PRO; Tecan, Mannedorf, Switzerland). The control solution consisted of 50 μl Reagent 1, 475 μl Reagent 2, and 475 μl Reagent 3. The TAG content (mg g^{-1}) was calculated using the following formula: $1.533 \times (\Delta A + 0.0199)/W$, where ΔA refers to the difference in absorbance between the experimental and control solutions and W represents the sample quantity (g). The TAG content was determined using three biological replicates.

Construction of the overexpression vectors

The *NAC83* and *NAC58* coding sequences in longan EC were amplified by PCR and then inserted into separate pCAMBIA-1302-35S vectors at the *Bgl*II and *Spel* restriction sites using the In-Fusion® HD Cloning Kit (Takara, San Jose, CA, USA). The generated pCAMBIA1302-35S-NAC83 and pCAMBIA1302-35S-NAC58 recombinant plasmids were inserted into protoplasts isolated from longan EC. The *GDSL-1* and *ERF6* coding sequences in longan EC were amplified by PCR and then inserted into separate pCAMBIA-1301-35S vectors at the *Bam*HI and *Sall* restriction sites using the In-Fusion HD Cloning Kit (Takara). The generated pCAMBIA1301-35S-*GDSL-1* and pCAMBIA1301-35S-*ERF6* recombinant plasmids were used for longan EC stable transformation. Details regarding the amplified sequences are provided in Table S13.

Transient transformation longan protoplasts

A polyethylene glycol (PEG)-based method was used for the transient transformation of protoplasts. Briefly, 1 μg recombinant plasmid was added to a centrifugation tube. Sterile ddH₂O was then added to reach 20 μl , after which 200 μl protoplast suspension and then an equal volume of PEG solution (40% PEG 400, 0.2 M mannitol, and 0.1 M CaCl₂) were added. The prepared solution was mixed gently and then left undisturbed for 5.5 min before 440 μl W5 buffer (0.154 M NaCl, 0.125 M CaCl₂, 0.005 M KCl, 0.005 M glucose, and 0.002 M MES) was added to stop the reaction. The solution was centrifuged (100 g min^{-1} for 1 min) and the supernatant was removed. After adding 440 μl W5 buffer, the sample was centrifuged again and the abovementioned steps were repeated again. Finally, the transformed protoplasts were dissolved in 1 ml W5 buffer. The protoplast suspensions were transferred to a 24-well plate and incubated at 25°C for 24 h on a shaker (50 r min⁻¹), after which the samples were collected.

Stable transformation of the longan EC

Agrobacterium tumefaciens strain EHA105 cells were transformed with the pCAMBIA1301-35S-*GDSL-1* and pCAMBIA1301-35S-*ERF6* recombinant plasmids according to the freeze-thaw method. The rapidly proliferating longan EC was transferred into the *A. tumefaciens* culture (OD₆₀₀ = 0.7–0.8). After a 30-min infection, the EC was transferred to MS solid co-culture medium containing 30 g L⁻¹ sucrose for a 3-day incubation, after which it was transferred to MS solid medium containing 1.0 mg L⁻¹ 2,4-D, 20 g L⁻¹ sucrose, and 300 mg L⁻¹ cefotaxime. The proliferating EC was harvested after 20 days and transferred to MS solid selection medium containing 1.0 mg L⁻¹ 2,4-D, 20 g L⁻¹ sucrose, 300 mg L⁻¹ cefotaxime, and 20 mg L⁻¹ hygromycin and subcultured every 30 days until a resistant EC formed. The resistant EC was detected on the basis of GUS staining (Huayueyang Biotechnology, Beijing, China) and PCR amplification. For the

differentiation of the early somatic embryos, the WT and transformed ECs were transferred to MS solid medium containing 300 mg L⁻¹ cefotaxime and then incubated at 25°C and 35°C. Samples were harvested after 6, 9, and 12 days. For MT treatment, WT and transformed ECs were transferred to MS solid medium containing 4.5 μM MT and 300 mg L⁻¹ cefotaxime and then incubated at 25°C and 35°C. Samples were harvested after 6, 9, and 12 days. The morphological characteristics of the WT and transformed somatic embryos were examined using a microscope (Olympus, Southborough, MA, USA).

Comparative transcriptome analysis

WT and *ERF6* transgenic ECs cultured on MS solid medium containing 0 or 4.5 μM MT under normal (25°C) and high-temperature stress (35°C) conditions were collected, and the total RNA was extracted for RNA sequencing using TRIzol reagent (Invitrogen). The quality of the extracted RNA samples was assessed using a Nanodrop2000 (Thermo Fisher, Waltham, MA, USA) and an Agilent 2100 Bioanalyzer (Agilent Technologies). After the mRNA enrichment step using poly-T oligo-attached magnetic beads, sequencing libraries were constructed using the Hieff NGS Ultima Dual-mode mRNA Library Prep Kit for Illumina (Yeasen Biotechnology, Shanghai, China) and then sequenced using an Illumina NovaSeq 6000 platform (Biomarker Technologies Co., Ltd., Wuhan, China). The final library size and quality were checked using a Qsep400 and a Qubit Fluorometer (Life Technologies). Gene expression was calculated and normalized to FPKM values. DESeq2 was used to identify the DEGs ($\log_2(\text{fold change}) > 1$, $p < 0.01$). GO and KEGG enrichment analyses of the DEGs were conducted using clusterProfiler in R. The RNA-seq analysis was completed using two biological replicates.

RNA extraction and gene expression analysis

Total RNA isolation, first-strand cDNA synthesis, and qRT-PCR analysis were performed as previously described (Zhang, Zhu, et al., 2020). *UBIQUITIN* (*UBQ*), *BETA-ACTIN* (*ACTB*), and *EUKARYOTIC ELONGATION FACTOR 1-ALPHA* (*EF-1 α*) were used as internal reference controls to normalize gene expression levels. The qRT-PCR data were analyzed according to the $2^{-\Delta Ct}$ method. The qRT-PCR analysis was performed using three biological replicates. The transcriptional data (FPKM values) were obtained from the longan early somatic embryogenesis transcriptome datasets to analyze the expression change of the cluster-enriched expression genes. The longan EC temperature stress (15°C, 25°C, and 35°C) transcriptome dataset was previously obtained by us (GenBank BioProject ID: PRJNA889670). In brief, the rapidly proliferating longan EC was transferred to 15°C, 25°C, and 35°C conditions, and samples were harvested after 24 h to use for bulk RNA-seq. The log₂-transformed (FPKM + 1) values were used to construct the heatmaps. Details regarding the amplified sequences are provided in Table S13.

Statistical analyses

The significance of the differences among groups was determined using a Student's *t* test and one-way analysis of variance (ANOVA) followed by the Duncan test.

AUTHOR CONTRIBUTIONS

SZ, CZ, ML, and XX performed the experiments; SZ, CZ, XZ, CL, XX, YC, and ZZ analyzed the data; YL, SZ, and ZL wrote the article with contributions from the other authors; all authors contributed to the design of the project.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (31672127), the Natural Science Foundation of Fujian Province (2020J01543), the Constructions of Plateau Discipline of Fujian Province (102/71201801101), and the Technology Innovation Fund of Fujian agriculture and forestry university (CXZX2019033S and CXZX2018078). We thank Guangzhou Genedenovo Biotechnology Co., Ltd. for assisting in sequencing and bioinformatics analysis. We thank Liwen Bianji (Edanz) (www.liwenbianji.cn/) for editing the English text of a draft of this manuscript.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflicts of interest.

DATA AVAILABILITY STATEMENT

The scRNA-seq and RNA-seq data (BioProjects PRJCA016297, PRJCA016314, and PRJCA016317) were deposited in the Beijing Institute of Genomics Data Center (<http://bigd.big.ac.cn>).

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Figure S1. Pearson correlation analysis of scRNA-seq and bulk RNA-seq data ($R > 0.8$). EC, embryogenic callus.

Figures S2–S7. Phylogenetic trees constructed using the maximum likelihood method reflecting the relationships among the marker genes between longan and Arabidopsis.

Figure S8. KEGG pathway enrichment analysis of upregulated differentially expressed genes in the 12 cell clusters. The colored bar and numbers indicate the q values; $q \leq 0.05$ was set as the threshold for identifying significantly enriched KEGG pathways.

Figure S9. Distribution of upregulated differentially expressed genes involved in the phenylpropanoid biosynthesis pathway. Gene names in the colored text box indicate the differentially expressed genes revealed by our scRNA-seq data.

Figure S10. Phylogenetic tree constructed using the maximum likelihood method reflecting the relationships among the GPATs of longan and Arabidopsis.

Figure S11. Expression patterns of the upregulated genes associated with the 'regulation of autophagy' pathway in the single-cell transcriptome. Dot diameter and color indicate expression levels.

Figure S12. Phenotypes of the early somatic embryo in response to rapamycin and 3-methyladenine treatments. For rapamycin treatment, embryogenic calli were cultured on MS solid medium containing 0, 0.1, 0.5, or 1.0 mg L⁻¹ rapamycin. Samples were harvested after 6, 9, and 12 days and the somatic embryo phenotypes were observed using a microscope (Olympus). MS, Murashige and Skoog medium without additives; MS + RAPA, Murashige and Skoog medium containing 0, 0.1, 0.5, or 1.0 mg L⁻¹ rapamycin. Bar = 100 µm. The red arrow indicates typical globular embryos. For 3-methyladenine treatment, embryogenic calli were cultured on MS solid medium containing 0, 0.1, 0.5, or 1.0 µM 3-methyladenine. Samples were harvested after 6, 9, and 12 days and the somatic embryo phenotypes were observed using a microscope (Olympus). MS, Murashige and Skoog medium without additives; MS + 3-MA, Murashige and Skoog medium

containing 0, 0.1, 0.5, or 1.0 mg L⁻¹ 3-methyladenine. Bar = 100 µm. The red arrow indicates typical globular embryos.

Figure S13. Representative expression dynamics over the pseudo-timeline of the marker genes with cluster-enriched expression. The colored bar indicates relative expression levels. Each dot denotes a single cell.

Figure S14. Verification of the representative top 20 upregulated genes during early somatic embryogenesis (embryogenic callus, incomplete compact pro-embryogenic culture, and globular embryo stages) by qRT-PCR analysis. EC, embryogenic callus; IcpEC, incomplete compact pro-embryogenic culture; GE, globular embryo. One-way ANOVA followed by the Duncan test, $P < 0.01$.

Figure S15. Expression patterns of upregulated differentially expressed genes related to the phenylpropanoid biosynthesis pathway during early somatic embryogenesis (embryogenic callus, incomplete compact pro-embryogenic culture, and globular embryo stages). The log₂-transformed (FPKM + 1) values were used to construct the heatmap.

Figure S16. Expression patterns of the top 20 upregulated genes in cluster 11 in the single-cell transcriptome. Dot diameter and color indicate expression levels.

Figure S17. GUS staining of *ERF6*-overexpressing transgenic embryogenic callus. WT, wild type; *ERF6*-OE1 and *ERF6*-OE2, *ERF6*-overexpressing embryogenic callus cell lines 1 and 2. Bar = 100 µm.

Figure S18. Phenotypes of WT and *ERF6*-overexpressing somatic embryos. The transgenic embryogenic callus was cultured on MS solid medium under normal (25°C) and high-temperature stress (35°C) conditions to induce somatic embryos. Samples were harvested after 6, 9, and 12 days and the somatic embryo phenotypes were observed using a microscope (Olympus). Blue represents the GUS staining result. WT, wild type; *ERF6*-OE2, *ERF6*-overexpressing embryogenic callus cell line 2. Bar = 100 µm.

Figure S19. Expression patterns of *ERF6* in the transcriptome datasets for different temperatures (15°C, 25°C, and 35°C). The log₂-transformed (FPKM + 1) values were used to construct the heatmap.

Figure S20. Phenotypes of the WT and *ERF6*-overexpressing somatic embryos treated with melatonin. The transgenic embryogenic callus was cultured on MS solid medium supplemented with 4.5 µM melatonin under normal (25°C) and high-temperature stress (35°C) conditions to induce somatic embryos. Samples were harvested after 6, 9, and 12 days and the somatic embryo phenotypes were observed using a microscope (Olympus). Blue represents the GUS staining result. WT, wild type; *ERF6*-OE2, *ERF6*-overexpressing embryogenic callus cell line 2. Bar = 100 µm.

Figure S21. GO term enrichment analysis of the downregulated DEGs between the WT-35 and OE-35 cell lines. The size of the dots represents the number of genes associated with the GO term and the color of the dots represents the q value. WT-35, wild-type embryogenic callus cultured under high-temperature stress (35°C) conditions; OE-35, *ERF6*-overexpressing embryogenic callus cell line cultured under high-temperature stress (35°C) conditions.

Table S1. Top 20 upregulated marker genes in each cluster.

Table S2. KEGG enrichment analysis of the upregulated differentially expressed genes in 12 cell clusters.

Table S3. Fatty acid compositions and relative contents in the early somatic embryos (embryogenic callus, incomplete compact pro-embryogenic culture, and globular embryo stages).

Table S4. Upregulated differentially expressed genes over the pseudo-timeline of the cell differentiation trajectory.

Table S5. FPKM values of the top 20 upregulated genes in the early somatic embryos (embryogenic callus, incomplete compact pro-embryogenic culture, and globular embryo stages).

Table S6. FPKM values of the upregulated differentially expressed genes related to the phenylpropanoid biosynthesis pathway in the early somatic embryos (embryogenic callus, incomplete compact pro-embryogenic culture, and globular embryo stages).

Table S7. Differentially expressed transcription factors and the corresponding upregulated differentially expressed target genes over the pseudo-timeline of the cell differentiation trajectory.

Table S8. Downregulated differentially expressed genes between WT-35 and OE-35 associated with enriched GO terms.

Table S9. Upregulated differentially expressed genes between OE-35 and OE-MT35 associated with enriched GO terms.

Table S10. FPKM values of the ‘auxin-activated signaling’ pathway DEGs in different *ERF6*-overexpressing embryogenic calli.

Table S11. Upregulated differentially expressed genes between OE-35 and OE-MT35 associated with enriched KEGG pathways.

Table S12. FPKM values of the plant growth and development pathway DEGs in different *ERF6*-overexpressing embryogenic calli.

Table S13. Details regarding the primer sequences used in this study.

REFERENCES

- Abe, M., Katsumata, H., Komeda, Y. & Takahashi, T. (2003) Regulation of shoot epidermal cell differentiation by a pair of homeodomain proteins in *Arabidopsis*. *Development*, **130**, 635–643.
- Akoh, C.C., Lee, G.C., Liaw, Y.C., Huang, T.H. & Shaw, J.F. (2004) GDSL family of serine esterases/lipases. *Progress in Lipid Research*, **43**, 534–552.
- Aparato, V.P.M. & Suh, D.Y. (2022) The functions and applications of GDSL esterase/lipase proteins in agriculture: a review. *JSFA Reports*, **2**, 304–312.
- Ascencio-Ibañez, J.T., Sozzani, R., Lee, T.J., Chu, T.M., Wolfinger, R.D., Cella, R. et al. (2008) Global analysis of *Arabidopsis* gene expression uncovers a complex array of changes impacting pathogen response and cell cycle during geminivirus infection. *Plant Physiology*, **148**, 436–454.
- Becht, E., McInnes, L., Healy, J., Dutertre, C.A., Kwok, I.W.H., Ng, L.G. et al. (2019) Dimensionality reduction for visualizing single-cell data using UMAP. *Nature Biotechnology*, **37**, 38–44.
- Berthet, S., Demont-Caulet, N., Pollet, B., Bidzinski, P., Cézard, L., Le Bris, P. et al. (2011) Disruption of LACCASE4 and 17 results in tissue-specific alterations to lignification of *Arabidopsis thaliana* stems. *The Plant Cell*, **23**, 1124–1137.
- Bird, D., Beisson, F., Brigham, A., Shin, J., Greer, S., Jetter, R. et al. (2007) Characterization of *Arabidopsis* ABCG11/WBC11, an ATP binding cassette (ABC) transporter that is required for cuticular lipid secretion. *The Plant Journal*, **52**, 485–498.
- Brown, L.A., Larson, T.R., Graham, I.A., Hawes, C., Paudyal, R., Warriner, S.L. et al. (2013) An inhibitor of oil body mobilization in *Arabidopsis*. *New Phytologist*, **200**, 641–649.
- Cai, X.T., Xu, P., Zhao, P.X., Liu, R., Yu, L.H. & Xiang, C.B. (2014) *Arabidopsis* ERF109 mediates cross-talk between jasmonic acid and auxin biosynthesis during lateral root formation. *Nature Communications*, **5**, 5833.
- Chen, C., Chen, H., Zhang, Y., Thomas, H.R., Frank, M.H., He, Y. et al. (2020) TBTtools: an integrative toolkit developed for interactive analyses of big biological data. *Molecular Plant*, **13**, 1194–1202.
- Chen, Y., Xu, X., Liu, Z., Zhang, Z., XuHan, X., Lin, Y. et al. (2020) Global scale transcriptome analysis reveals differentially expressed genes involve in early somatic embryogenesis in *Dimocarpus longan* Lour. *BMC Genomics*, **21**, 4.
- De Almeida Engler, J., De Veylder, L., De Groot, R., Rombauts, S., Boudolf, V. & De Meyer, B. (2009) Systematic analysis of cell-cycle gene expression during *Arabidopsis* development. *The Plant Journal*, **59**, 645–660.
- DeBono, A., Yeats, T.H., Rose, J.K.C., Bird, D., Jetter, R., Kunst, L. et al. (2009) *Arabidopsis* LTPG is a glycosylphosphatidylinositol-anchored lipid transfer protein required for export of lipids to the plant surface. *The Plant Cell*, **21**, 1230–1238.
- Dubois, M., Skirycz, A., Claeys, H., Maleux, K., Dhondt, S. & De Bodt, S. (2013) ETHYLENE RESPONSE FACTOR6 acts as a central regulator of leaf growth under water-limiting conditions in *Arabidopsis*. *Plant Physiology*, **162**, 319–332.
- Dubois, M., Van den Broeck, L., Claeys, H., Van Vlierberghe, K., Matsui, M. & Inze, D. (2015) The ETHYLENE RESPONSE FACTORS ERF6 and ERF11 antagonistically regulate mannitol-induced growth inhibition in *Arabidopsis*. *Plant Physiology*, **169**, 166–179.
- Galluzzi, L. & Green, D.R. (2019) Autophagy-independent functions of the autophagy machinery. *Cell*, **177**, 1682–1699.
- Guan, Y., Li, S.G., Fan, X.F. & Su, Z.H. (2016) Application of somatic embryogenesis in woody plants. *Frontiers in Plant Science*, **7**, 938.
- Haecker, A., Groß-Hardt, R., Geiges, B., Sarkar, A., Breuninger, H., Herrmann, M. et al. (2004) Expression dynamics of WOX genes mark cell fate decisions during early embryonic patterning in *Arabidopsis thaliana*. *Development*, **131**, 657–668.
- Horstman, A., Bemer, M. & Boutilier, K. (2017) A transcriptional view on somatic embryogenesis. *Regeneration*, **4**, 201–216.
- Hu, Y.X., Han, X.S. & Jing, Q. (2019) Autophagy in development and differentiation. *Autophagy: Biology and Diseases*, **1206**, 469–487.
- Kim, H., Lee, S.B., Kim, H.J., Min, M.K., Hwang, I. & Suh, M.C. (2012) Characterization of glycosylphosphatidylinositol-anchored lipid transfer protein 2 (LTPG2) and overlapping function between LTPG/LTPG1 and LTPG2 in cuticular wax export or accumulation in *Arabidopsis thaliana*. *Plant & Cell Physiology*, **53**, 1391–1403.
- Kumar, A., Yogendra, K.N., Karre, S., Kushalappa, A.C., Dion, Y. & Choo, T.M. (2016) WAX INDUCER1 (HvWIN1) transcription factor regulates free fatty acid biosynthetic genes to reinforce cuticle to resist fusarium head blight in barley spikelets. *Journal of Experimental Botany*, **67**, 4127–4139.
- Lai, Z.X. & Chen, C.L. (2002) Changes of endogenous hormones during somatic embryogenesis of longan. *Chinese Journal of Tropical Crops*, **23**, 7.
- Lai, Z.X. & Chen, Z.G. (1997) Somatic embryogenesis of high frequency from longan embryogenic calli. *Journal of Fujian Agricultural University*, **26**, 271–276.
- Lai, Z.X., Pan, L.Z. & Chen, Z.G. (1997) Establishment and maintenance of longan embryogenic cell lines. *Journal of Fujian Agricultural University*, **2**, 33–40.
- Lee, H.G., Kim, H., Suh, M.C., Kim, H.U. & Seo, P.J. (2018) The MYB96 transcription factor regulates triacylglycerol accumulation by activating DGAT1 and PDAT1 expression in *Arabidopsis* seeds. *Plant & Cell Physiology*, **59**, 1432–1442.
- Lee, Y.L., Chen, J.C. & Shaw, J.F. (1997) The thioesterase I of *Escherichia coli* has arylesterase activity and shows stereospecificity for protease substrates. *Biochemical and Biophysical Research Communications*, **231**, 452–456.
- Li, D., Jin, C., Duan, S., Zhu, Y., Qi, S., Liu, K. et al. (2017) MYB89 transcription factor represses seed oil accumulation. *Plant Physiology*, **173**, 1211–1225.
- Li, H.J., Xue, Y., Jia, D.J., Wang, T., Hi, D.Q., Liu, J. et al. (2011) POD1 regulates pollen tube guidance in response to micropylar female signaling and acts in early embryo patterning in *Arabidopsis*. *The Plant Cell*, **23**, 3288–3302.
- Li, Y.J., Yu, Y., Liu, X., Zhang, X.S. & Su, Y.H. (2021) The *Arabidopsis* MATERNAL EFFECT EMBRYO ARREST45 protein modulates maternal auxin biosynthesis and controls seed size by inducing AINTEGUMENTA. *The Plant Cell*, **33**, 1907–1926.
- Li-Beisson, Y., Shorrosh, B., Beisson, F., Andersson, M.X., Arondel, V., Bates, P.D. et al. (2013) Acyl-lipid metabolism. *Arabidopsis book*, **8**, e0133.
- Lin, Y., Min, J., Lai, R., Wu, Z., Chen, Y., Yu, L. et al. (2017) Genome-wide sequencing of longan (*Dimocarpus longan* Lour.) provides insights into molecular basis of its polyphenol-rich characteristics. *GigaScience*, **6**, 1–14.
- Liu, H., Hu, D., Du, P., Wang, L., Liang, X., Li, H. 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 Biotechnology Journal*, **19**, 2261–2276.
- Liu, Q., Liang, Z., Feng, D., Jiang, S., Wang, Y., Du, Z. et al. (2020) Transcriptional landscape of rice roots at the single cell resolution. *Molecular Plant*, **14**, 1–11.
- Lu, H.C., Lam, S.H., Zhang, D., Hsiao, Y.Y., Li, B.J., Niu, S.C. et al. (2022) R2R3-MYB genes coordinate conical cell development and cuticular wax biosynthesis in *Phalaenopsis aphrodite*. *Plant Physiology*, **188**, 318–331.

- Mathieu, M., Lelu-Walter, M.A., Blervacq, A.S., David, H., Hawkins, S. & Neutelings, G.** (2006) Germin-like genes are expressed during somatic embryogenesis and early development of conifers. *Plant Molecular Biology*, **61**, 615–627.
- Menges, M., Hennig, L., Gruissem, W. & Murray, J.A.H.** (2002) Cell cycle-regulated gene expression in *Arabidopsis*. *Journal of Biological Chemistry*, **277**, 41987–42002.
- Murphy, D.J.** (2012) The dynamic roles of intracellular lipid droplets: from archaea to mammals. *Protoplasma*, **249**, 541–585.
- Nelms, B. & Walbot, V.** (2019) Defining the developmental program leading to meiosis in maize. *Science*, **364**, 52–56.
- Nong, W., Huang, J.L., Li, G.W. & Li, G.F.** (2006) Research on the reason of severely small size fruit phenomenon in longan. *Journal of Southern Agriculture*, **37**, 314–316.
- Pagnussat, G.C., Yu, H.J., Ngo, Q.A., Rajani, S., Mayalagu, S., Johnson, C.S. et al.** (2005) Genetic and molecular identification of genes required for female gametophyte development and function in *Arabidopsis*. *Development*, **132**, 603–614.
- Potocka, I., Baldwin, T.C. & Kurczynska, E.U.** (2012) Distribution of lipid transfer protein 1 (LTP1) epitopes associated with morphogenic events during somatic embryogenesis of *Arabidopsis thaliana*. *Plant Cell Reports*, **31**, 2031–2045.
- Potter, S.S.** (2018) Single-cell RNA sequencing for the study of development, physiology and disease. *Nature Reviews Nephrology*, **14**, 479–492.
- Qin, F., Kakimoto, M., Sakuma, Y., Maruyama, K., Osakabe, Y., Tran, L.S.P. et al.** (2007) Regulation and functional analysis of ZmDREB2A in response to drought and heat stresses in *Zea mays* L. *The Plant Journal*, **50**, 54–69.
- Satija, R., Farrell, J.A., Gennert, D., Schier, A.F. & Regev, A.** (2015) Spatial reconstruction of single-cell gene expression data. *Nature Biotechnology*, **33**, 495–502.
- Shinohara, H., Mori, A., Yasue, N., Sumida, K. & Matsubayashi, Y.** (2016) Identification of three LRR-RKs involved in perception of root meristem growth factor in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America*, **113**, 3897–3902.
- Song, S., Guo, Q., Zhu, Y., Yuan, P., Yan, Z., Yan, L. et al.** (2022) Exploring the role of autophagy during early human embryonic development through single-cell transcriptome and methylome analyses. *Science China Life Sciences*, **65**, 940–952.
- Suh, M.C., Samuels, A.L., Jetter, R., Kunst, L., Pollard, M., Ohlrogge, J. et al.** (2005) Cuticular lipid composition, surface structure, and gene expression in *Arabidopsis* stem epidermis. *Plant Physiology*, **139**, 1649–1665.
- Sun, S., Shen, X., Li, Y., Li, Y., Wang, S., Li, R. et al.** (2023) Single-cell RNA sequencing provides a high-resolution roadmap for understanding the multicellular compartmentation of specialized metabolism. *Nature Plants*, **9**, 179–190.
- Ten Hove, C.A., Lu, K.J. & Weijers, D.** (2015) Building a plant: cell fate specification in the early *Arabidopsis* embryo. *Development*, **142**, 420–430.
- Tian, C., Du, Q., Xu, M., Du, F. & Jiao, Y.** (2020) Single-nucleus RNA-seq resolves spatiotemporal developmental trajectories in the tomato shoot apex. *BioRxiv*. Available from: <https://doi.org/10.1101/2020.09.20.305029>
- Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M. et al.** (2014) The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nature Biotechnology*, **32**, 381–386.
- Van Leene, J., Hollunder, J., Eeckhout, D., Persiau, G., Van De Slijke, E., Stals, H. et al.** (2010) Targeted interactomics reveals a complex core cell cycle machinery in *Arabidopsis thaliana*. *Molecular Systems Biology*, **6**, 397.
- Vessoni, A.T., Muotri, A.R. & Okamoto, O.K.** (2012) Autophagy in stem cell maintenance and differentiation. *Stem Cells and Development*, **21**, 513–520.
- Vroemen, C.W., Langeveld, S., Mayer, U., Ripper, G., Jurgens, G., Van Kammen, A. et al.** (1996) Pattern formation in the *Arabidopsis* embryo revealed by position-specific lipid transfer protein gene expression. *The Plant Cell*, **8**, 783–791.
- Wagner, D.E., Weinreb, C., Collins, Z.M., Briggs, J.A., Megason, S.G. & Klein, A.M.** (2018) Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. *Science*, **360**, 981–987.
- Wang, Q., Wu, Y., Peng, A., Cui, J., Zhao, M., Pan, Y. et al.** (2022) Single-cell transcriptome atlas reveals developmental trajectories and a novel metabolic pathway of catechin esters in tea leaves. *Plant Biotechnology Journal*, **20**, 2089–2106.
- Wang, Y., Reiter, R.J. & Chan, Z.** (2018) Phytomelatonin: a universal abiotic stress regulator. *Journal of Experimental Botany*, **69**, 963–974.
- Wright, A.J. & Hunter, C.P.** (2003) Mutations in a β-Tubulin disrupt spindle orientation and microtubule dynamics in the early *Caenorhabditis elegans* embryo. *Molecular Biology of the Cell*, **14**, 4512–4525.
- Wu, Y.Y., Tan, H.L., Shui, G., Bauvy, C., Huang, Q., Wenk, M.R. et al.** (2010) Dual role of 3-methyladenine in modulation of autophagy via different temporal patterns of inhibition on class I and III sphingomyelinase 3-kinase. *Journal of Biological Chemistry*, **285**, 10850–10861.
- Xu, X., Crow, M., Rice, B.R., Li, F., Harris, B., Liu, L. et al.** (2021) Single-cell RNA sequencing of developing maize ears facilitates functional analysis and trait candidate gene discovery. *Developmental Cell*, **56**, 557–568.
- Yin, Z.Y., Yin, J., Huo, Y.F., Yu, J., Sheng, L.X. & Dong, Y.F.** (2017) Rapamycin facilitates fracture healing through inducing cell autophagy and suppressing cell apoptosis in bone tissues. *European Review for Medical and Pharmacological Sciences*, **21**, 4989–4998.
- Yu, M. & Zhao, J.** (2012) The cytological changes of tobacco zygote and proembryo cells induced by beta-glucosyl Yariv reagent suggest the involvement of arabinogalactan proteins in cell division and cell plate formation. *BMC Plant Biology*, **12**, 126.
- Zhai, N. & Xu, L.** (2021) Pluripotency acquisition in the middle cell layer of callus is required for organ regeneration. *Nature Plants*, **7**, 1453–1460.
- Zhang, J., Yin, X.R., Li, H., Xu, M., Zhang, M.X., Li, S.J. et al.** (2020) ETHYLENE RESPONSE FACTOR39-MYB8 complex regulates low-temperature-induced lignification of loquat fruit. *Journal of Experimental Botany*, **71**, 3172–3184.
- Zhang, S., Zhu, C., Lyu, Y., Chen, Y., Zhang, Z., Lai, Z. et al.** (2020) Genome-wide identification, molecular evolution, and expression analysis provide new insights into the APETALA2/ethylene responsive factor (AP2/ERF) superfamily in *Dimocarpus longan* Lour. *BMC Genomics*, **21**, 62.
- Zhang, T.Q., Chen, Y. & Wang, J.W.** (2021) A single-cell analysis of the *Arabidopsis* vegetative shoot apex. *Developmental Cell*, **56**, 1056–1074.
- Zhang, T.Q., Xu, Z.G., Shang, G.D. & Wang, J.W.** (2019) A single-cell RNA sequencing profiles the developmental landscape of *Arabidopsis* root. *Molecular Plant*, **12**, 648–660.
- Zhao, Q., Nakashima, J., Chen, F., Yin, Y., Fu, C., Yun, J. et al.** (2013) LACCASE is necessary and nonredundant with PEROXIDASE for lignin polymerization during vascular development in *Arabidopsis*. *The Plant Cell*, **25**, 3976–3987.
- Zhao, Y., Lin, S., Qiu, Z., Cao, D., Wen, J., Deng, X. et al.** (2015) MicroRNA857 is involved in the regulation of secondary growth of vascular tissues in *Arabidopsis*. *Plant Physiology*, **169**, 2539–2552.
- Zhong, J., Ren, Y., Yu, M., Ma, T., Zhang, X. & Zhao, J.** (2011) Roles of arabinogalactan proteins in cotyledon formation and cell wall deposition during embryo development of *Arabidopsis*. *Protoplasma*, **248**, 551–563.
- Zhou, C., Labbe, H., Sridha, S., Wang, L., Tian, L., Latoszek-Green, M. et al.** (2004) Expression and function of HD2-type histone deacetylases in *Arabidopsis* development. *The Plant Journal*, **38**, 715–724.
- Zhou, X., Liu, Z., Shen, K., Zhao, P. & Sun, M.X.** (2020) Cell lineage-specific transcriptome analysis for interpreting cell fate specification of proembryos. *Nature Communications*, **11**, 1366.