

Spatial transcriptome and single-cell RNA sequencing reveal the molecular basis of cotton fiber initiation development

Jun Zhang^{1,2,†}, Rui Chen^{1,†}, Fan Dai^{1,†}, Yue Tian^{3,†}, Yue Shi¹, Ying He¹, Yan Hu¹ and Tianzhen Zhang^{1,*}

¹Zhejiang Provincial Key Laboratory of Crop Genetic Resources, the Advanced Seed Institute, Plant Precision Breeding Academy, College of Agriculture and Biotechnology, Zhejiang University, Hangzhou, China,

²Institute of Horticulture, Zhejiang Academy of Agricultural Sciences, Hangzhou, Zhejiang, China, and

³College of Biotechnology, Jiangsu University of Science and Technology, Zhenjiang, China

Received 19 July 2024; revised 22 January 2025; accepted 10 February 2025.

*For correspondence (e-mail cotton@zju.edu.cn).

†These authors contributed equally to this work.

SUMMARY

Recent advances in single-cell transcriptomics have greatly expanded our knowledge of plant development and cellular responses. However, analyzing fiber cell differentiation in plants, particularly in cotton, remains a complex challenge. A spatial transcriptomic map of ovule from –1 DPA, 0 DPA, and 1 DPA in cotton was successfully constructed, which helps to explain the important role of sucrose synthesis and lipid metabolism during early fiber development. Additionally, single-cell RNA sequencing (scRNA-seq) further highlighted the cellular heterogeneity and identified clusters of fiber developmental marker genes. Integration of spatial and scRNA-seq data unveiled key genes *SVB* and *SVBL* involved in fiber initiation, suggesting functional redundancy between them. These findings provide a detailed molecular landscape of cotton fiber development, offering valuable insights for enhancing lint yield.

Keywords: *Gossypium hirsutum* L., single-cell transcriptomics, fiber differentiation, spatial transcriptome, *SVBL*, *SVB*.

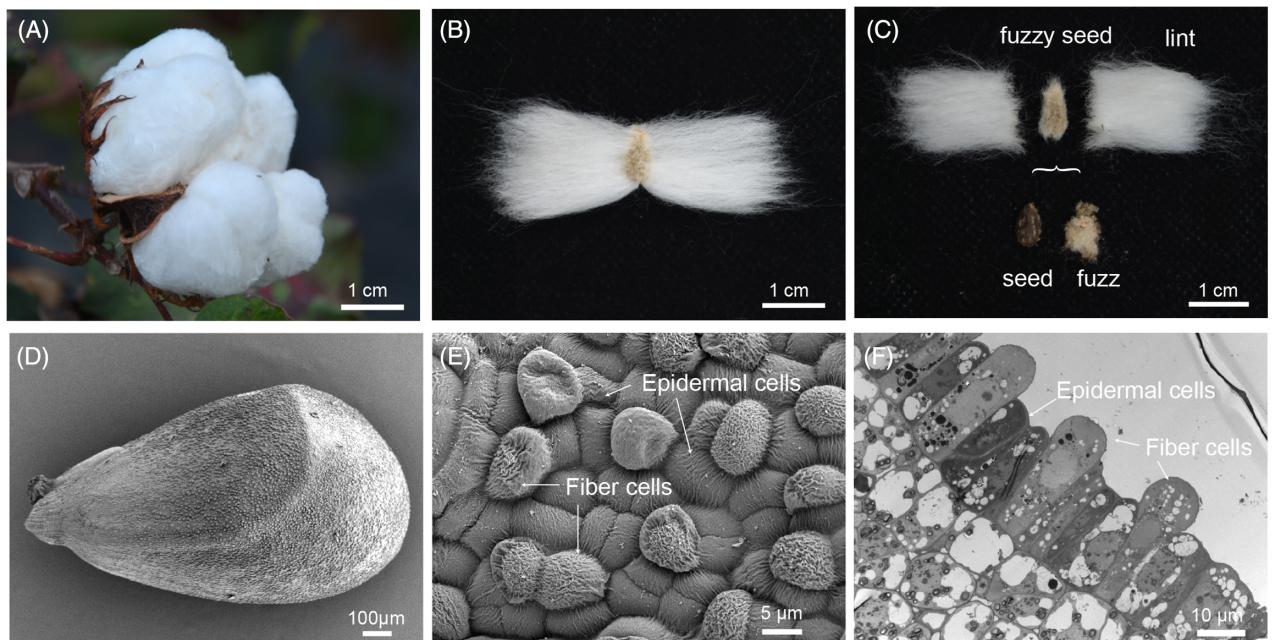
INTRODUCTION

Single-cell sequencing (scRNA-seq) approaches are quickly changing our perspective on biological processes such as cell differentiation and associated gene networks by increasing the spatiotemporal resolution of our analyses to the individual cell level. scRNA-seq has already been used in plants to identify at high spatial resolution the involvement of previously unknown factors that influence cell type-specific processes and cell stage transitions (Seyfferth et al., 2021). The first effective high-throughput generation of a single-cell plant transcriptome was from *Arabidopsis* root protoplasts (Ryu et al., 2019). Subsequently, many plant developmental processes have been investigated at single-cell resolution, such as root cell types (Denyer et al., 2019; Zhang et al., 2019), developing leaf cells (Procko et al., 2022), stomatal lineage cells (Liu et al., 2020), female gametophytes in *Arabidopsis* (Song et al., 2020), root cells in rice (Zhang, Chen, Liu, et al., 2021), cotton fiber (Qin et al., 2022; Wang et al., 2023), developing ears (Xu et al., 2021), and polar stem cells (Chen et al., 2021) in maize.

Conducting scRNA-seq necessitates the dissociation of tissues into solitary cells for microfluidic handling,

inevitably leading to the loss of crucial data regarding the spatial provenance of those cells. Meanwhile, many methodologies aimed at defining this information, such as *in situ* hybridization, are susceptible to introducing experimental bias. Spatial transcriptomics, a novel technique in the realm of plant biology, preserves the spatial information of cells and holds promise in mitigating some of the intrinsic limitations associated with scRNA-seq. Already, spatial transcriptomics has been instrumental in elucidating regenerative mechanisms in tomato callus tissue buds (Song et al., 2023), the developmental processes of maize and barley endosperm (Fu et al., 2023; Peirats-Llobet et al., 2023), the growth characteristics of poplar trees (Du et al., 2023), and the formation of floral organs in orchids (Liu et al., 2022). These studies help us to understand the complexity of plant cells as distinct entities, which has long been an active area of interest in plant biology.

Cotton is the most abundant fiber crop worldwide. Like *Arabidopsis* leaf trichomes, the fibers (seed trichomes) of cotton are single-celled plant hairs, but cotton has developed a unique network for fiber development (Tian & Zhang, 2021). In *Arabidopsis*, transcription factors (TFs)

**Figure 1.** Cotton fiber and fiber cells.

(A–C) Cotton fiber consists of fuzz fiber and lint fiber.

(D, E) Scanning electron microscopy images showing the epidermal surface of a 0 DPA ovule.

(F) Transmission electron microscopy image showing the epidermal surface of a 0 DPA ovule. Arrows indicate the initiating epidermal cells and fiber cells. Scale bars = 1 cm (A–C), 100 μ m (D), 5 μ m (E), 10 μ m (F).

such as homeodomain-leucine zipper IV subfamily factor GLABRA2 (*GL2*) and TRANSPARENT TESTA GLABRA1 (*TTG1*) are required to initiate leaf trichrome development (Rerie et al., 1994). The cotton homeodomain-leucine zipper IV gene *GaHOX1* can rescue the *Arabidopsis* glabrous mutant *gl2-2* and inhibit the development of epidermal hairs in wild-type *Arabidopsis* (Guan et al., 2008); the related gene *GhHOX3* is one of the most important TFs regulating cotton fiber elongation (Shan et al., 2014). The MYB TF family is likewise an important group of genes, being involved in almost the entire plant growth and development process. In *Arabidopsis*, mutations in the MYB family member *GL1* result in leaves with no epidermal hairs (Larkin et al., 1994). Overexpression of the cotton homolog *GaMYB2* (Stracke et al., 2001), which is highly expressed in the early stage of cotton fiber development, restores the hairless phenotype of *Arabidopsis gl1* mutants (Wang et al., 2004). Meanwhile, antisense-mediated suppression of *GhMYB109* leads to the retardation of fiber initials and shorter fibers (Pu et al., 2008). Notably, the MYBMIXTA-like (MML) subgroup of R2R3MYBs is expanded in the Malvaceae and comprises a Malvaceae-specific family that promotes epidermal fiber cell differentiation (Stracke et al., 2001). Transcriptome analysis has indicated that 10 MML TFs are specifically expressed in certain cotton fiber developmental stages (Zhang et al., 2015).

Cotton fibers are divided into fuzz and lint, both of which differentiate from the epidermal cells of the seed. Lint fiber differentiates 3 days before anthesis, and fuzz fiber 3–5 days post anthesis (DPA) (Stewart, 1975) (Figure 1A–E). The MML family member *GhMYB25* (*GhMML7*) was found to be highly expressed in fiber-bearing ovules at 0–2 DPA. RNA interference (RNAi) disrupting its expression results in a significant decrease in leaf epidermal hairs, while its overexpression leads to improved lint yield (Walford et al., 2011). Our map-based cloning and functional study on *GhMML3* (*GhMYB25-like*), which is highly expressed in –1 to 3 DPA ovules, revealed that its natural antisense transcripts produce small RNAs, degrade *GhMML3*, and lead to the dominant *N₁* fuzzless mutant phenotype. These data indicate *GhMML3* to be a key TF that controls cotton fuzz fiber initiation (Wan et al., 2016). Further supporting this function is the observation that RNAi suppression of *GhMML3* results in a nearly fiberless phenotype (Walford et al., 2011). Notably, the fiber cells at 0 DPA are small (~10 μ m in diameter) and intermingled with epidermal cells (Figure 1F), and hence, bulk transcriptome and laser-capture microdissection (LCM) cannot easily distinguish the two cell types at that early stage (Ando et al., 2021). The strategy of enzymatic dissociation and scRNA-seq offers the possibility of separately capturing fiber cells and epidermal cells.

The final lint yield in cotton is largely determined by the quantity of epidermal cells capable of transforming into fibers during the initiation stage of fiber development (Tian & Zhang, 2021). Thus, lint yield can be enhanced by increasing the proportion of epidermal cells that evolve into lint fibers. Consequently, it is crucial to identify and characterize the genes that play roles in the fiber initiation stage so as to comprehend the biological functions and genetic interactions that underpin fiber development and ultimately augment lint yield. In pursuit of this goal, we first constructed a spatial transcriptomic landscape of wild-type cotton at three key stages: –1 DPA, 0 DPA, and 1 DPA. Additionally, we conducted scRNA-seq on ovule epidermal cells of J668 (considered the linted fuzzy wild-type). The results revealed WT exhibits divergent gene expression patterns during fiber differentiation and protrusion, highlighting the significant role of MMLs in regulating the initiation of both cotton fuzz and lint fibers. Finally, our establishment of a spatial transcriptomic and scRNA-seq database (<http://39.98.92.180:6116/>; http://cotton.zju.edu.cn/sc_browser/cottonfiber/) along with the developmental trajectories of cotton fiber are an instrumental step in elucidating molecular mechanisms that could enhance cotton fiber yield. This advancement will address key challenges in the field of agricultural biotechnology.

RESULTS

Spatial transcriptomic features of the cotton ovule epidermis during continuous development

Spatial transcriptomics using *in situ* sequencing techniques can unveil the heterogeneity within tissues. At –1 DPA, 0 DPA, and 1 DPA, epidermal fiber cells in the cotton ovule exhibit distinct states: no protrusions, protrusions, and early elongation (Figure S1). We conducted spatial transcriptome sequencing with the 10x Genomics Visium platform on lint-bearing ovule cross sections at these three developmental stages to elucidate the cellular transcriptomes (Figure 2A; Figure S2A,B). A total of 825 spots were initially obtained. After removing drifting spots, 778 spots remained, with a median of 4181 genes sequenced per spot, covering a total of 39,267 genes (Table S1). Overall, spots from the three developmental stages separated into 15 clusters (Figure S2C), to which we applied linear regression and UMAP visualization (Figure S2D). Cell types were assigned to the sequenced spots, including epidermis (E), epidermis and fiber cells (EF), fiber cells (F), inner integument (II), outer integument (OI), nucellus (N), and degenerated nucellus (DN) (Figure 2B; Figure S3). We observed specific expression of the genes EARLY NODULIN-LIKE PROTEIN 20 (*ENODL20*, *GH_D04G0473*), ERF/AP2 transcription factor 59 (*ERF59*, *GH_A06G0401*), AVRPPHB SUSCEPTIBLE1-LIKE 14 (*PBL14*, *GH_D13G2178*), and *MYB68* (*GH_D07G0252*) in the nucellus, inner integument,

outer integument, and degenerated nucellus, respectively (Figure 2C; Figure S3; Table S2). Meanwhile, high expression at the outermost layers of the ovule was evident for sucrose synthase 4 (*GH_D06G0856*), *MYB25* (*GH_D04G2108*), and *GhEXPA2* (*GH_D10G1269*) at –1 DPA, 0 DPA, and 1 DPA, respectively (Figure 2D) (Li et al., 2016; Machado et al., 2009).

Gene Ontology (GO) enrichment analysis revealed that, compared to 0 DPA and 1 DPA, –1 DPA uniquely exhibited enrichment in “autophagy,” “sucrose synthase activity,” and “sucrose metabolic process.” Additionally, the term “fatty acid biosynthetic process” showed increasing enrichment as the developmental stages progressed (Figure 2E; Table S2). Fatty acid metabolism has been reported important during fiber initiation (Zou et al., 2022); for example, the *FIDDLEHEAD* (*FDH*) gene encodes a putative lipid biosynthetic enzyme (Pruitt et al., 2000), and *Arabidopsis* *fdh* mutant plants have fewer trichomes on rosette leaves (Yephremov et al., 1999). We found cotton *FDH* genes (*GH_A13G2177*, *GH_D13G2158*) to be mainly expressed in epidermal and fiber cells (Figure S4), implying important roles of these genes during fiber initiation. Autophagy has likewise been reported relevant to fiber development, with specific inhibition of cotton heat shock protein 70 and 90 (*GhHSP70*, *GhHSP90*) resulting in autophagy in the epidermal layer of the ovule, which leads to fewer fiber initiations and retardation of fiber elongation (Sable et al., 2018). Concerning sucrose metabolism, sucrose synthase expression has previously been revealed necessary for initiation of cotton fiber (Ruan et al., 2005). We observe expression of sucrose synthase 4 (*GH_D06G0856*) to be closely associated with epidermal and fiber cells (Figure S4), supporting its importance during fiber initiation.

Enrichment of the terms “translational elongation,” “small ribosomal subunit,” and “GTPase activity” pathways played a significant role during fiber development stages (0 DPA and 1 DPA) (Figure 2E; Table S3). The translation elongation factor 1A (*eEF1A*) is important in protein synthesis, catalyzing the binding of aminoacyl-tRNA to the ribosome A-site by a GTP-dependent mechanism (Xu et al., 2007). Molecular characterization and expression analysis of cotton *GhEF1As* suggest that these genes are necessary to meet the high demand for protein synthesis during fiber cell elongation (Xu et al., 2007). Notably, *EF1B* (*GH_A08G2198*) was primarily expressed in epidermal and fiber cells, supporting its potential involvement in fiber cell initiation (Figure S4; Table S4). Enrichment of ribosome-related terms in fiber cells likewise reflects the upregulation of protein translation/synthesis in preparation for subsequent protrusion and expansion. For example, *RPL16A* is expressed in developing pollen and the early phase of lateral root initiation (Williams & Sussex, 1995), which shows an expression profile similar to that of fiber

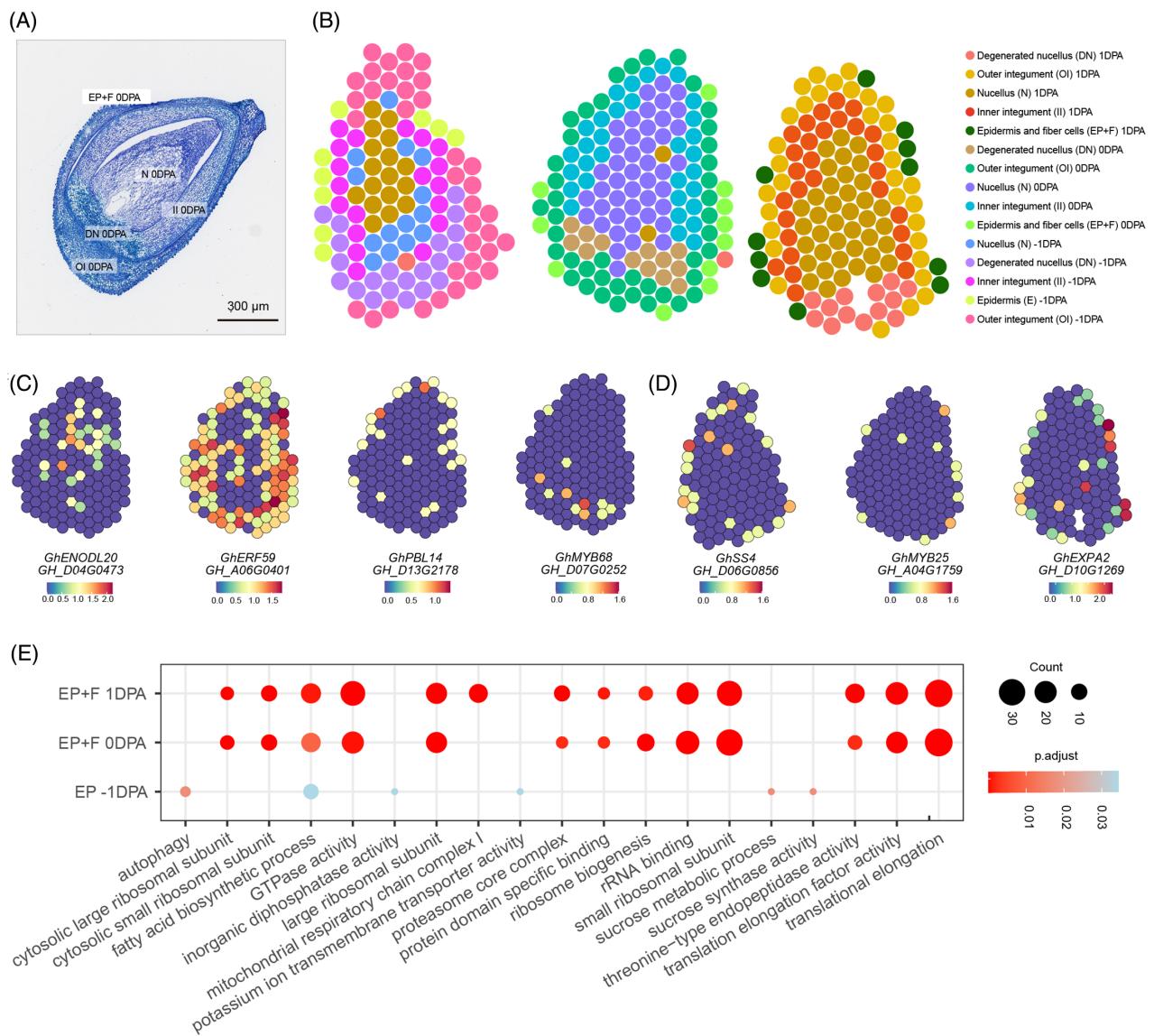


Figure 2. Spatial transcriptomic map of cotton ovules.

(A) Paraffin section of 0 days post anthesis (0 DPA) cotton ovules in transverse section, revealing the ovule structural composition including the epidermis and fiber cells (EP + F), pure fiber cells (F), inner integument (II), outer integument (OI), nucellus (N), and degenerated nucellus (DN).

(B) Spatial transcriptomics of cotton ovules at -1 DPA, 0 DPA, and 1 DPA, respectively, illustrate the distinct transcriptional profiles. Sequenced spots were categorized into different cell types: epidermis (E), epidermis and fiber cells (EF), fiber cells (F), inner integument (II), outer integument (OI), nucellus (N), and degenerated nucellus (DN).

(C) Analysis of gene expression revealed unique patterns for specific genes within various ovule components. The genes EARLY NODULIN-LIKE PROTEIN 20 (*ENODL20*, *GH_D04G0473*), ERF/AP2 transcription factor 59 (*ERF59*, *GH_A06G0401*), AvrPphB SUSCEPTIBLE1-LIKE 14 (*PBL14*, *GH_D13G2178*), and *MYB68* (*GH_D07G0252*) exhibited distinct expression patterns in the nucellus, inner integument, outer integument, and degenerated nucellus, respectively.

(D) Heightened expression in the outermost layers of the ovule was evident for the gene's sucrose synthase 4 (*GH_D06G0856*), *MYB25* (*GH_D04G2108*), and *GhEXPA2* (*GH_D10G1269*) at -1 DPA, 0 DPA, and 1 DPA, respectively.

(E) Gene Ontology (GO) enrichment analysis of the epidermis and fiber cells at -1 DPA, 0 DPA, and 1 DPA highlighted specific biological processes and molecular functions associated with these developmental stages.

cells during fiber cell initiation. Finally, GTPases are conserved molecular switches built according to a common structural design (Bourne et al., 1991), and overexpression of the GTPase protein GhROP6 has been shown to significantly enhance fiber length (Xi et al., 2023). We found the

GTPase *EF1ALPHA* (*GH_A01G1068*) to be mainly expressed in epidermal and fiber cells (Figure S4; Table S4), implying it may have an effect on fiber initiation.

Beyond enrichment analyses, we identified some important transcription factors, such as *GhMYB25-like*

(Wan et al., 2016), *GhMYB25* (Machado et al., 2009), and *GhHOX3* (Shan et al., 2014), as specifically expressed in epidermal and fiber cells (Figure S4; Table S4). Additionally, phytohormones are known to play important roles during fiber development; specifically, in vitro experiments have demonstrated that unfertilized ovules require auxin and gibberellin to produce fiber (Beasley & Ting, 1974). We found genes related to auxin (*GH3.1*) and gibberellin (*GASA4*) to be preferentially expressed in epidermal and fiber cells (Figure S4; Table S4) (Di et al., 2021; Rubinovich & Weiss, 2010), implying them to play important roles during fiber initiation. Furthermore, we identified a lipid transport protein (*LTPG1*) and lipid metabolism-related enzymes (*E6*, *PEL3*) to be specifically expressed in epidermal and fiber cells (Figure S4; Table S4) (Deng et al., 2016; Marks et al., 2009). We believe these genes to be expressed in preparation for the polar elongation of fiber cells.

The single-cell landscape of the cotton ovule epidermis

We observed the developmental stages of ovule epidermis in fuzzy-linted WT plants at −1 DPA, 0 DPA, and 1 DPA using scanning electron microscopy (SEM). Our findings reveal that, at 0 DPA, fiber cells exhibit a continuous transitional state, encompassing various cell types such as epidermal cells, early fiber cells, and fiber cells (Figure S1). Therefore, 0 DPA serves as a critical time point for capturing single cells in different states, making it particularly significant for simulating fiber cell differentiation, initiation, and investigating the dynamic development of fiber cells. Single-cell RNA sequencing (scRNA-seq) was conducted on 0 DPA ovules of the fuzzy-linted WT.

The 10× Genomics Chromium platform was used to capture and build the cDNA library (Figure 3A). A total of 7898 cells in WT were captured with an average of 1640 genes per cell. After filtration, 7566 cells remained in WT, accounting for 96.41% of the original cells (Figure S5; Table S5). Cell transcriptomic profiles were predicted in an unsupervised manner in the absence of marker genes; this yielded 12 clusters, which were then analyzed using the unified manifold approximation and projection (UMAP) algorithm for visualization (Figure 3B). Clusters 1 and 3 have more cells (Figure 3C).

A significant correlation between bulk RNA sequencing (RNA-seq) and scRNA-seq was observed in WT, with a Pearson's *R* value of 0.66 ($P < 2.2e-16$). Notably, the bulk RNA-seq samples were collected simultaneously with the scRNA-seq samples. These notable correlations highlight the exceptional caliber of the scRNA-seq data (Figure S6). Fiber developmental marker genes *GhMML3_D12* and *GhHOX3* were found to be highly expressed in Clusters 1 and 3 (Figure 3D–G); these results suggested Clusters 1 and 3 to be fiber cells. Upon scoring based on cell cycle

genes, Clusters 2 and 9 were categorized as proliferating cells (Figure S7; Table S7). Enhanced annotation of single-cell subpopulations can be realized through the integrated analysis of single-cell and spatial transcriptomic data. Here, such analysis identified Clusters 4 and 5 to signify late outer integument cells; Cluster 10 as indicative of early outer integument cells; Clusters 6 and 7 to represent inner integument cells; and Cluster 11 as encompassing late inner integument cells. Notably, the integrated analysis linked Cluster 1 with early fiber initiation cells, but not Cluster 3 (Figure S8).

Potential fiber initiation development genes *SVB*, *SVBL*

In *Arabidopsis*, the tissue specificity and intracellular distribution of the *Smaller Trichomes with Variable Branches* (*SVB*) family member DUF538 are very similar to *SVB-like* (*SVBL*) (Yu et al., 2021). The *svb svbl* double mutant in *Arabidopsis* displayed a more severe trichome phenotype than either the *svb* or *svbl* single mutants, with severe defects in trichome initiation, growth, and branching (Yu et al., 2021). The cotton *SVBL* (*GH_D08G2327*, *GH_D13G2020*, *GH_A13G2046*) gene is highly expressed in the epidermal and fiber cells of ovule ST (Figure 4). In WT single-cell clusters 1 (early fiber cells) and 3 (fiber cells), not only were the same three *SVBL* genes highly expressed, but also two *SVB* genes (*GH_D13G1275*, *GH_A13G1345*) were observed (Figure 4; Table S6). This suggests that these genes might be involved in cotton fiber initiation.

We selected key genes for validation using the "Zheda XiaoJin" variety, which is characterized by its short growth cycle, small plant size, and high space utilization (Zhang et al., 2023). Using the model cotton "Zheda XiaoJin," we performed virus-induced gene silencing (VIGS) for *SVBL* and *SVB* (Figure 5). RNA interference suppression of *MML3* (*GhMYB25-like*) resulted in cotton plants with fiberless seeds, so we chose silenced *MML3* as a positive control (Walford et al., 2011). Quantitative PCR results indicated effective silencing of each gene in the different treatments (Figure S9). Scanning electron microscopy results showed that the 0 DPA ovule epidermis of the empty vector negative control (pTRV2:00) and the positive control (pTRV2:MML3) had (233.3 ± 55.2) and (2 ± 1) fiber cell protrusions, respectively. In contrast, silencing *SVB* alone resulted in (160.7 ± 11.7) fiber cells, and silencing *SVBL* alone resulted in (145.7 ± 18.6) fiber cells. Although the number of fiber cells decreased when either gene was silenced, the difference was not significant (Figure 5). However, simultaneous silencing of both *SVBL* and *SVB* significantly reduced the number of fiber cells to (69.7 ± 22.9) , showing a significant difference compared to the control (Figure 5). This study demonstrates that *SVBL* and *SVB* in cotton also affect fiber cell initiation and that functional redundancy exists between them in cotton.

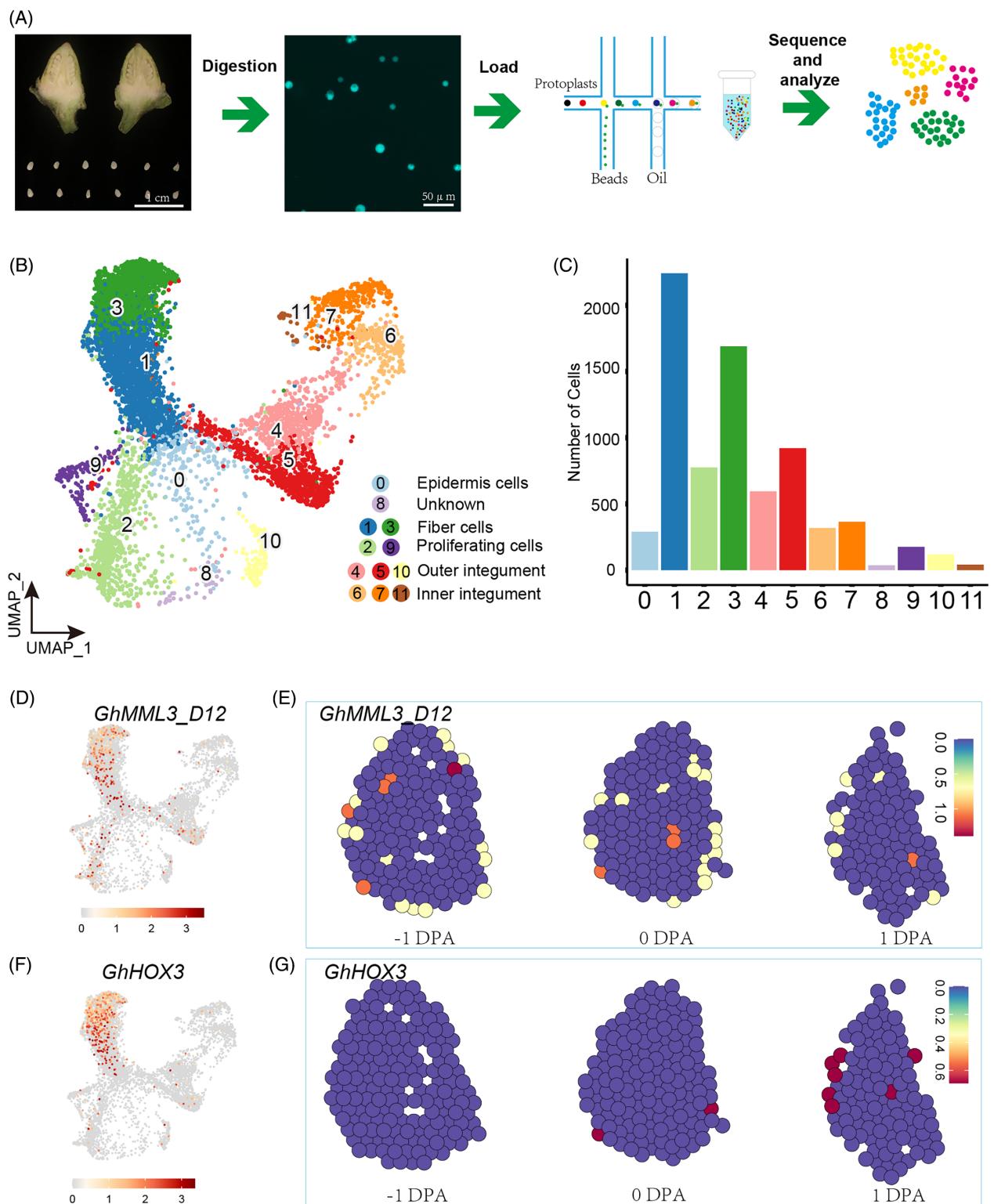


Figure 3. Single cell landscape of WT.

(A) Workflow of scRNA-seq. Protoplasts isolated from 0 DPA ovules were loaded into a 10× Genomics Chromium Controller.

(B) UMAP visualization of 12 clusters derived from 7566 cells in WT. Each dot denotes a single cell.

(C) The number of cells in each cluster.

(D, E) Expression of *GhMML3_D12* in single-cell and spatial transcriptomics.

(F, G) Expression of *GhHOX3* in single-cell and spatial transcriptomics.

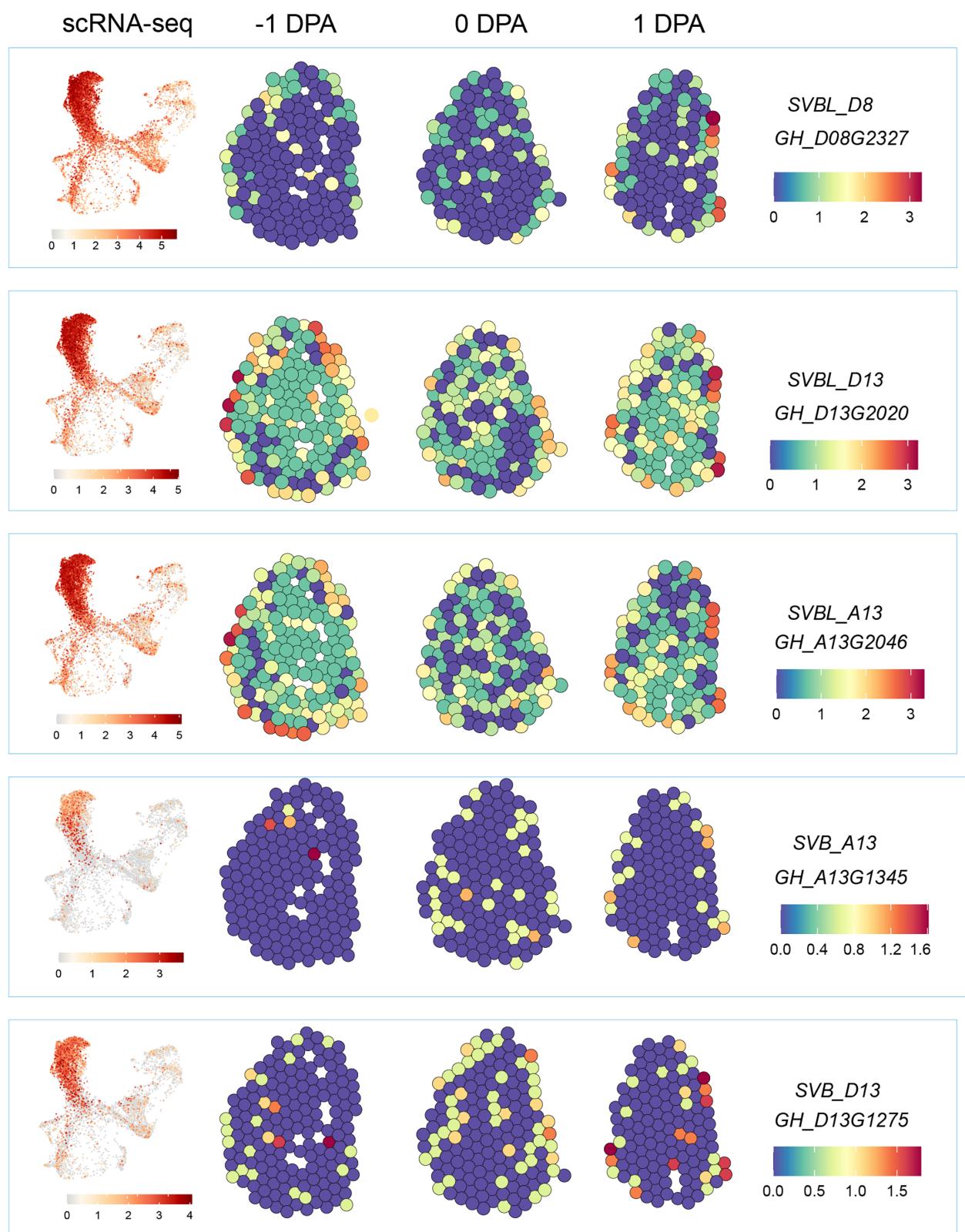


Figure 4. The expression pattern of the *SVB* and *SVBL* genes in the single-cell and spatial localization of the ovule. The expression of five genes, *SVBL* (*SVBL_A13*, *SVBL_D13*, and *SVBL_D8*) and *SVB* (*SVB_A13* and *SVB_D13*), was shown in the UMAP plot of the single-cell transcriptomics data and across three developmental stages (-1 DPA, 0 DPA, and 1 DPA) in the spatial transcriptomics of the ovule.

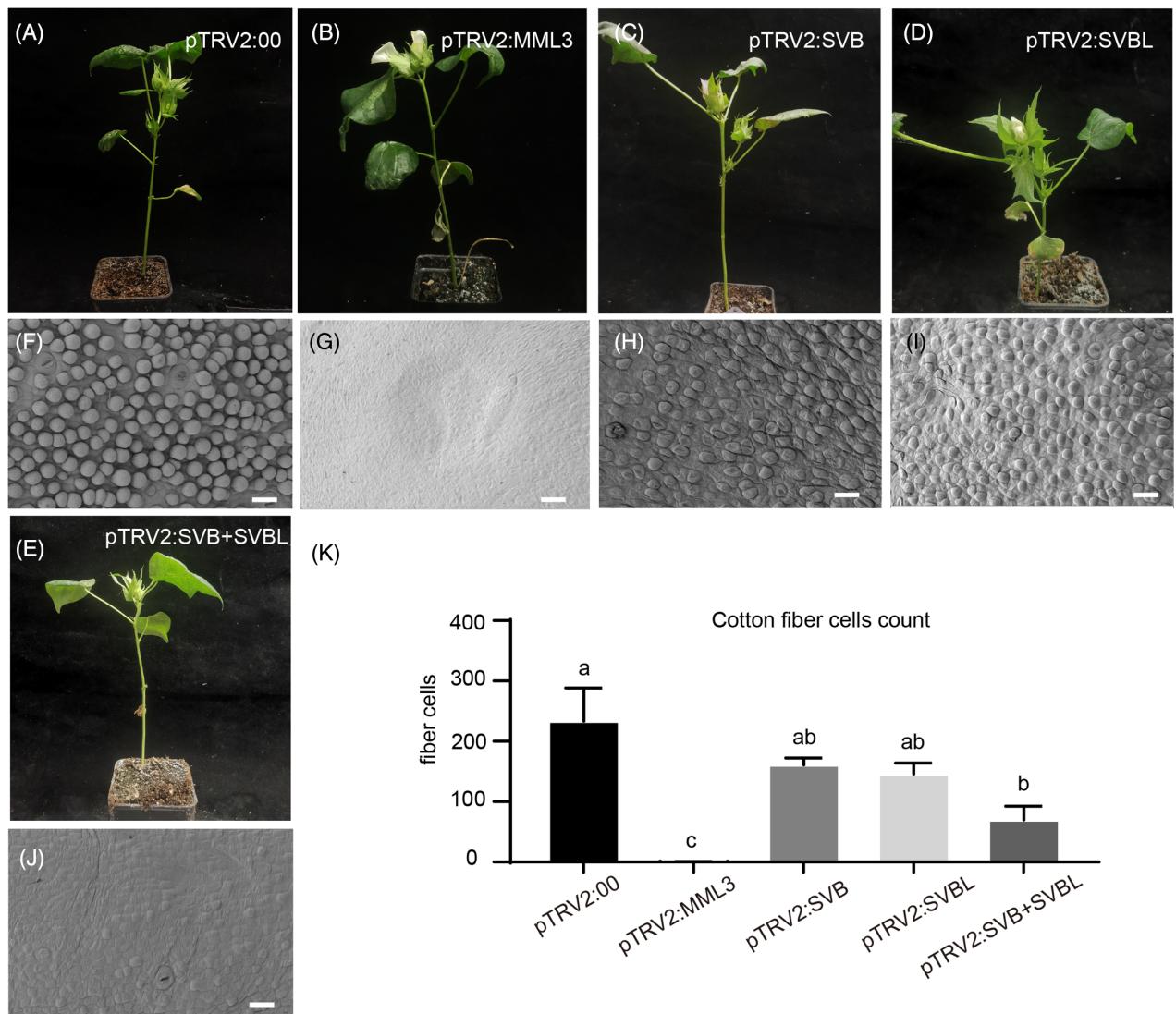


Figure 5. Functional validation of *SVB* and *SVBL* genes in "Zheda Xiaojin."

(A, B) The empty vector negative control and *MML3* positive control, respectively.

(C, D) Silencing of *SVB* and *SVBL*, respectively.

(E) Co-silencing of *SVB* and *SVBL*.

(F–J) SEM results of the respective silenced genes, scale = 20 μm.

(K) Statistical count of fibroblast numbers under SEM view (3450 μm²). Significant differences tested using Dunnett's test are represented with different letters ($n = 3$, $P < 0.05$).

The developmental transition from epidermal to fiber cell trajectories in WT

Cotton fibers are produced from ovule epidermal cells through differentiation and elongation. As this process determines the number of epidermal cells that eventually differentiate into fibers, it is particularly important for cotton fiber output. To investigate the development of these immature epidermal cells and identify key differentiation nodes, we performed an analysis focused on Clusters 0, 1, 2, and 3 from the WT to identify changes in the transition of naive epidermal cells into fiber cells. WT Cluster 2 was

comprised of cells undergoing active division (identified by expression of, e.g., histone-associated genes), epidermal cells (*GH3.1*), and fiber cells (*LTP*) (Table S6). The developmental trajectories of the seven WT re-clusters were predicted using SCORPIUS and CytoTRACE, which yielded the following sequence: re-cluster WT4, re-cluster WT5, re-cluster WT1, re-cluster WT6, re-cluster WT0, re-cluster WT2, and re-cluster WT3 (Figure 6). Further analysis using the Mfuzz method yielded six distinct patterns of gene expression associated with the proliferation of fiber cells and epidermal cells (Figure S10). In particular, genes linked

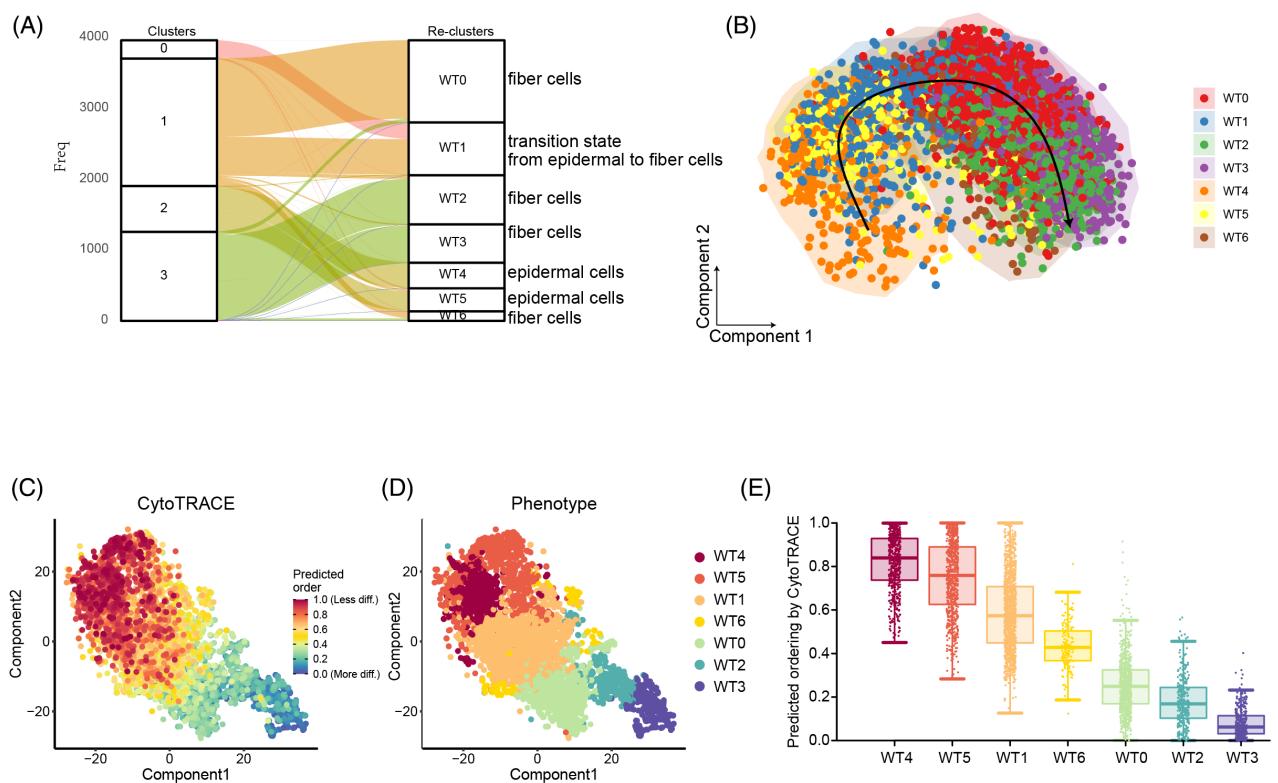


Figure 6. Dynamic developmental status of fiber cells.

(A) The re-clustering outcomes of WT, specifically subcluster WT0–6 with a resolution of 0.6. Before re-clustering, cluster 0 was epidermal cells, cluster 2 was dividing epidermal cells, and clusters 1 and 3 were fiber cells. After re-clustering, WT4 and WT5 were identified as epidermal cells, WT1 as the transition state from epidermal to fiber cells, and WT6, WT0, WT2, and WT3 as fiber cells.

(B) SCORPIUS trajectory analysis showing the relationships between different fiber re-clusters (clusters WT0–6 from the WT samples).

(C) The direction of differentiation in nearly every evaluated dataset, recovered by CytoTRACE.

(D) Phenotypic clustering of WT subgroup cells by CytoTRACE.

(E) The direction of differentiation in nearly every evaluated dataset, recovered by CytoTRACE. The CytoTRACE method was utilized to forecast the capacity of clusters WT0–6 to differentiate, resulting in the establishment of the developmental sequence as WT4, WT5, WT1, WT6, WT0, WT2, and WT3.

to fiber cell growth and epidermal cell development exhibited distinct initiation times (Table S8).

Gene Ontology enrichment analysis of Mfuzz trend groups found the group WT_C3, which represents early proliferative cells, to be enriched for the terms “nucleosome,” “protein heterodimerization activity,” and “nucleosome assembly.” All of these terms mostly involve proteins of the histone superfamily and are indicative of rapidly dividing cells (Table S9). In addition, WT_C3 showed early and elevated expression of genes associated with fiber initiation, such as *GhMML3-Dt*, auxin response factor 2 (*GhARF2b*), and the cell fate determinant *GhCPC*. This may indicate the early differentiation of epidermal cells into fiber cells (Table S7). Meanwhile, the epidermal cells in WT_C5 showed substantial enrichments for the GO terms “metal ion transport,” “response to stress,” and “sucrose synthase activity” (Table S9), and for genes associated with the production of fiber, such as *SS3* (Ruan et al., 2003), *HDA5* (Kumar et al., 2018), *JAZ2* (Hu et al., 2016), actin depolymerizing factor (*ADF*) (Wang

et al., 2009), and *PIN1* (Zhang, Zeng, et al., 2017) (Table S9). The groups WT_C2, WT_C4, and WT_C6 represent a range of transcriptional initiation, as indicated by enrichment of translation-associated GO pathways such as “large ribosomal subunit,” “small ribosomal subunit,” “translational elongation,” and “ribosomes.” This suggests an increase in protein translation/synthesis in fiber cells, preparing them for protrusion and expansion (Table S9). We also observed exclusive expression of some important TFs, such as *GhMYB25-like* (*GhMML3*) (Walford et al., 2011), *GhMYB25* (Machado et al., 2009), *GhMYB109* (Pu et al., 2008), a homeodomain–leucine zipper 1 (*GhHD1*) (Walford et al., 2012), PROTODERMAL FACTOR1 (*PDF1*) (Deng et al., 2012), and *GhHOX3* (Shan et al., 2014) in fiber cells rather than epidermal cells, suggesting these factors to be crucial in the initiation of fiber cells (Table S9). Overall, the group WT_C2 appeared to occupy a transitional state between epidermal and fiber cells, containing early fiber development genes such as homeodomain–leucine zipper 1(*HD1*), PROTODERMAL FACTOR1 (*PDF1*), and *MYB*

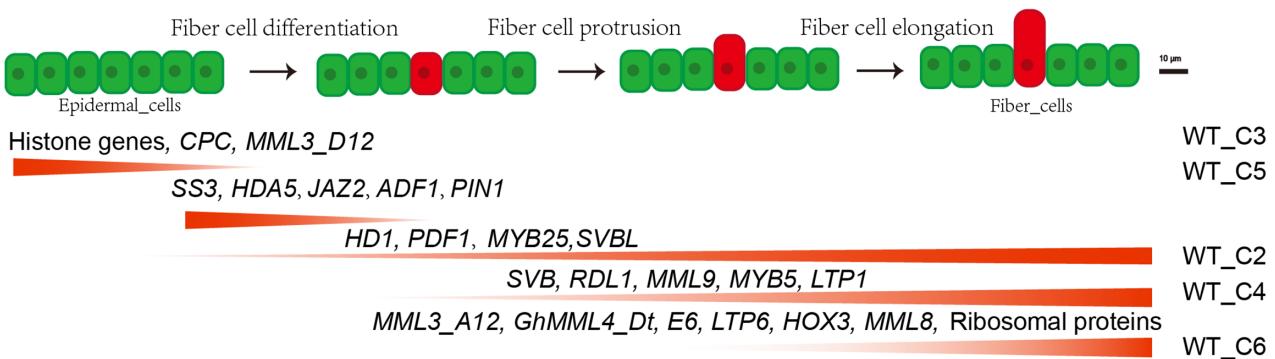


Figure 7. Contrasting dynamic patterns of fiber development between the wild type.

During the fiber differentiation and development phase of WT samples, genes demonstrate diverse expression patterns that ultimately result in typical fiber initiation and elongation.

transcription factors 25 (*MYB25*) (Figure 6; Table S9). Meanwhile, WT_C4 and WT_C6 both exhibited rapid upregulation of genes linked to fiber cell extension and elongation, such as *RD22-like1* (*RDL1*), *E6*, and *HOX3* (Figure 7; Table S8). Finally, expression of “lipid metabolism” genes like *E6* (John & Keller, 1996), *PEL3* (Zhu et al., 2012), *GDSL* (Nigam et al., 2014), *LTP3* (Zhu et al., 2012), and *LTPG1* (Deng et al., 2016) was restricted to fiber cells, indicating mobilization of fatty acid synthesis and transportation in preparation for fiber cell elongation (Table S8). This is consistent with a prior report that lipid metabolism pathways are significantly upregulated during fiber cell elongation (Gou et al., 2007).

DISCUSSION

Hindrance to the use of single-cell sequencing in plants is the complexity of plant cell types, which are difficult to annotate. Despite the widespread use of model plants such as *Arabidopsis* and poplar, there still remain some cell subsets that cannot be annotated based on current knowledge (Chen et al., 2021; Zhang, Chen, & Wang, 2021). *In situ* hybridization serves as a valuable technique for spatially mapping gene expression; however, its application is constrained by the limitations of experimental platforms, coupled with the notably low efficiency of individual gene hybridization.

Spatial transcriptomics technology has been used to analyze the spatial distribution of gene expression within tissues, with a focus on the positional information of transcripts. We obtained and analyzed spatial transcriptomic data from six ovule tissues at each of three different developmental phases. These tissues included the epidermis, fiber cells, inner integument, outer integument, nucellus, and degenerated nucellus. Currently, the 10x Genomics Visium platform is restricted in spatial resolution to the tissue/structural level, rather than the individual cell level. Nonetheless, our research revealed discrete expression clusters of specific genes in separate sections. The

biological meaning of these clusters is yet uncertain, and further research is required to clarify the importance of these signals.

Molecular investigations of fiber initiation will provide valuable data for improving our understanding of plant cell fate and subsequently increasing lint yield in cotton. In a previous report, *GhMML3* (*GhMYB25-like*) was highly expressed in –1 to 3 DPA ovules, and its inhibition by RNAi was shown to result in seeds with a nearly fiberless phenotype, whereas other epidermal hairs were not affected (Walford et al., 2011). It has been identified that two fuzzless loci, *N₁* and *n₂*, can lead to naked seeds. Wan et al. (2016) performed map-based cloning of the dominant fuzzless mutant *N₁* gene encoded by *GhMML3_A12*. By comparing the genotypes of different fuzzless mutants, the *n₂* locus was additionally mapped to chromosome 26 (D12), and it was proposed that *n₂* might be *MML3_D12* (Chen et al., 2020).

Although cotton fiber cells become visible on the ovule surface at anthesis, the differentiation of these selected ovule epidermal cells begins several days prior. Pseudo-time analysis of scRNA-seq data could locate single cells through a trajectory based on a biological process, such as cell differentiation. In the present work, we have distinguished epidermal and fiber cells in the 0 DPA ovule (Figure 1). By quantifying and comparing gene expression in the distinguished epidermal and fiber cells after the branch node, we determined most phytohormone-related genes to be preferentially expressed in epidermal cells (Table S6); accordingly, we believe that fiber cell differentiation requires phytohormone synthesis in and transportation from epidermal cells, especially auxin, which has previously been confirmed to transport from epidermal cells into fiber cells (Zhang, Zeng, et al., 2017). Subsequently, many key TFs are activated for fiber cell initiation, especially members of the MYB (*GhMYB25-like*, *GhMYB109*, *GhMML4*, *GhMML9*) and HD-ZIP (*GhHD1*, *GhPDF1*, *GhPDF2*) families. We believe

phytohormones and TFs coordinate to promote cotton fiber initiation in a sophisticated and subtle manner.

GO enrichment analysis showed epidermal cells to be closely associated with “response to stress,” “CTP synthase activity,” “metal ion transport,” and “chitin binding” (Table S7), and furthermore that most highly expressed genes are involved in responses to stress stimuli, like *WRKY33* (*GH_D04G1498*), *LOX3* (*GH_D10G0585*), *HMP53* (*GH_A08G2273*), and *CML27* (*GH_D12G2915*) (Birkenbihl et al., 2012; Yang et al., 2020). The large number and high levels of expressed stress-related genes imply them to play important roles during epidermal cell differentiation; precisely how these genes contribute to the process remains to be discovered.

Likewise, lipid transport proteins (LTP1, LTP3, LTPG1, LTPG6) and lipid-metabolism-related enzymes (E6, PEL3, GDSL) were specifically identified in fiber cells. We believe these genes are upregulated in preparation for the polar elongation of the fiber cell. The transcription factor *GhHOX3* likewise exhibited preferential expression in fiber cells and is known as one of the most important regulators during fiber elongation. Cotton fiber development is a complicated process consisting of four consecutive but overlapping stages: fiber cell differentiation and initiation, rapid elongation or primary cell wall synthesis, secondary cell wall synthesis, and maturation. We believe that genes are gradually activated in the course of this process; thus, by monitoring the dynamics of gene expression throughout the developmental process, we can infer and distinguish the different stages. Although we took samples at a time point of 0 DPA, pseudo-time analysis helps us to examine the continuous and dynamic process based on gene expression throughout fiber development (Figure 7). We also found a subset of cells to be in an intermediate state between fiber and epidermal cells; we believe some of these cells will further differentiate into fiber cells as fiber development proceeds.

EXPERIMENTAL PROCEDURES

Plant materials

The transgenic receptor material was created from *Gossypium hirsutum* acc. J668 (Li et al., 2019). Cotton plants were cultured in nutrient soil in controlled environment rooms at 25°C with a 16 h light/8 h dark photoperiod.

Tissue processing and spatial transcriptome scRNA-seq data analysis

Plant tissue samples were delicately rinsed with cold PBS and positioned cut-side down in a plastic mold filled with optimal cutting temperature (OCT, Tissue-Tek, catalog no. 4583), then snap-frozen using chilled isopentane. Cryosections were prepared at a thickness of 10 µm using a Leica cm1950. These sections were mounted onto chilled Visium Tissue Optimization Slides (catalog no. 3000394, 10× Genomics) and Visium Spatial Gene Expression

Slides (catalog no. 2000233, 10× Genomics). Adhesion was achieved by gently warming the slide’s reverse side. Following this, the tissue sections were fixed in chilled methanol and stained as recommended by the Visium Spatial Gene Expression User Guide (catalog no. CG000239 Rev A, 10× Genomics) and the Visium Spatial Tissue Optimization User Guide (catalog no. CG000238 Rev A, 10× Genomics). For gene expression assays, samples were subjected to a permeabilization process for an optimal duration of 6 min, determined through tissue optimization time-course studies (Figure S2A). Histological examinations were performed under bright-field conditions using a 20× objective on a Pannoramic MIDI scanner (3D HISTECH). Raw image compilations were facilitated by the Pannoramic MIDI software and saved as both low- and high-resolution .tiff files. In tissue optimization studies, fluorescent imagery was captured using a TRITC filter and a 20× objective lens.

The preparation of 10× libraries was conducted in accordance with the guidelines provided in the Visium Spatial Gene Expression User Guide (CG000239: https://assets.ctfassets.net/an68im79xiti/3pyXucRaiKWcscXy3cmRHL/a1ba41c77cbf60366202805ead8f64d7/CG000239_VisiumSpatialGeneExpression_UserGuide_Rev_A.pdf). Afterwards, the libraries were adjusted to a concentration of 300 picomolar (pM) and sequenced using the NovaSeq 6000 System manufactured by Illumina. Sequencing utilized the NovaSeq S4 Reagent Kit (200 cycles, catalog no. 20027466, Illumina) and consisted of four steps: read 1, which involved 28 cycles; i7 index read, 10 cycles; i5 index read, also 10 cycles; and read 2, 91 cycles. The sequencing was carried out by Majorbio Co., Ltd, located in Shanghai, China. The resulting data, which consisted of raw FASTQ files and histological images, were processed using the Space Ranger program version 1.2.2, which utilized STAR (Dobin et al., 2013) to align the genome with the Cell Ranger cotton reference genome (Hu et al., 2019). Quality metrics for each spot were assessed using the SeuratV4.0 R Bioconductor package (Stuart et al., 2019). Regions containing unique molecular identifiers (UMIs) that yielded a count of less than 200 were omitted from the analysis. Gene expression from each voxel was normalized using the sctransform method in Seurat (Hafemeister & Satija, 2019), which utilizes regularized negative binomial models to address technological errors and preserve biological variation. Following that, the 30 most significant principal components were identified and employed in the construction of a K-nearest neighbors (KNN) graph. The clusters were visualized using a 2D map created by employing both t-distributed stochastic neighbor embedding (t-SNE) and Uniform Manifold Approximation and Projection (UMAP) techniques. Subsequently, differentially expressed genes (DEGs) between samples or clusters were identified using Seurat’s FindMarkers function, which employed a likelihood ratio test. Differential expression was considered significant when the absolute log₂ fold change was larger than 0.25 and the Q-value was less than or equal to 0.05. The spatial transcriptomics website visualizes data using the SpatialViewR package (Mohanty et al., 2024).

Single-cell sequencing experimental methods and data analysis

Tissue samples of 0 DPA ovules were immediately collected after resection and dissociated into a single-cell suspension. The enzymatic solution was prepared with 1.5% (w/v) Cellulase R10, 0.75% (w/v) Macerozyme R-10, 1% (w/v) hemicellulase, 0.4 M mannitol, 20 mM KCl, 20 mM MES (pH 5.7), 10 mM CaCl₂, and 0.1% (w/v) BSA. For the dissociation, samples were added to a 50-mL centrifuge tube containing the enzymatic hydrolysis solution and

incubated at 35 rpm in the dark for 2 h at 28°C. Afterwards, the enzymatic hydrolysis reaction was terminated, and the solution was filtered through 75 and 40 µm cell strainers. The final flow-through was centrifuged at 200 rcf for 2 min to pellet the protoplasts, and the supernatant was removed. The pellet was then resuspended in DPBS containing 8% mannitol and 0.04% BSA, and the resuspension was transferred to a centrifuge tube containing 2 mL of 20% and 2 mL of 40% Percoll solution and centrifuged at 400 rcf for 20 min. Subsequently, the middle layer was carefully aspirated with a dropper and filtered using a 40-µm cell strainer. The filtrate was centrifuged at 200 rcf for 2 min, the supernatant was removed, and the protoplasts were resuspended. Finally, the cells were stained with 0.002% (w/v) FDA solution for 2 min at room temperature in the dark and observed under a microscope (200×) to ensure the proportion of viable cells reached 80% (Figure S2). The cells were counted using a 25 × 16 hemocytometer. After this confirmation, the final cell concentration was adjusted to 1200 cells/µL. Single cells were selected using microfluidic technology based on the 10× Genomics platform. DNA libraries were amplified by PCR, and high-throughput sequencing was performed using the Illumina sequencing platform's double-terminal sequencing mode.

Sequencing results were converted to FASTQ format using the Illumina bcl2fastq software. Sample demultiplexing, barcode processing, and single-cell 3' gene counting were performed using the Cell Ranger pipeline (<https://www.10xgenomics.com/support-software/cell-ranger/latest>; version 5.0.1) as was alignment of scRNA-seq data were aligned to the cotton reference genome TM-1 V2.1 (Hu et al., 2019). The Cell Ranger output was loaded into Seurat v. 4.0.1 (Satija et al., 2015) and used for dimensional reduction, clustering, and analysis. Overall, cells passed the quality control threshold: All genes expressed in fewer than three cells were removed, the number of genes expressed per cell was between 400 and 5000, UMI counts were less than 150, and mitochondrial-DNA derived genes comprised <10% of transcripts. As a further quality control step, doublets (two or multiple cells in one oil droplet) in each scRNA-seq dataset were detected with DoubletFinder (v. 2.0.3) (McGinnis et al., 2019) using 0.25 as the number of artificial doublets (pN). To identify the optimal neighborhood size (pK), we executed the function "paramSweep_v3" using the parameter "PCs = 1:20," and the maximum pK value was selected as an optimal pK parameter. To estimate the number of expected real doublets (nExp), we assumed a doublet formation rate of 7.5%, and the nExp value was then adjusted according to the homotypic doublet proportion.

To visualize the data, we used Seurat to reduce the dimensionality of all 7566 cells and UMAP to project the cells into 2D space. This process included using the LogNormalize method of the "Normalization" function of Seurat to calculate expression values, which normalized values were then utilized in principal component analysis (PCA). The top 10 PCs were employed in clustering and UMAP analysis. Clusters were identified using the weighted shared nearest neighbor graph-based clustering method, and marker genes for each cluster were identified using the "bimod" method (likelihood-ratio test) with default parameters via the FindAllMarkers function in Seurat. This function selects marker genes expressed in more than 25% of the cells in a cluster and for which the average log2 (fold change) greater than 0.25. The Cell-Trek method was used to perform a joint analysis of single-cell transcriptomics and spatial transcriptomics (Wei et al., 2022).

Pseudo-time analysis

We characterized the potential process of cell function changes and determined potential lineage differentiation along the

epidermal cell –fiber cell trajectory. Specifically, we used the differential GeneTest function in the Monocle 2 v2.22.0 package (Qiu et al., 2017) to construct single-cell developmental trajectories in pseudo-time order based on genes having differential expression between clusters. Dimension reduction was performed using the DDRTree approach implemented in the reduce dimension function. Cells mapped to the two-dimensional space were visualized as a minimum spanning tree using the plot_cell_trajectory function in Monocle 2. Trajectory reconstruction analysis of cotton fiber subclusters was performed using the R packages SCORPIUS (1.0.8) (Cannoodt et al., 2016) and CytoTRACE (0.3.3) (Gulati et al., 2020). Finally, the temporal trends in the expression of candidate marker genes obtained by SCORPIUS and clusters were analyzed using Mfuzz (Kumar & Futschik, 2007).

RNA in situ hybridization

We conducted mRNA in situ hybridization as previously described (Pei et al., 2021). Briefly, the fifth internodes of 0 DPA ovules were collected and immediately fixed in an FAA solution (10% formaldehyde, 5% acetic acid, 47.5% ethanol) for 48 h. The fixed tissue was dehydrated through a graded alcohol series (70, 85, 90, 95, and 100%) and a HistoClear series, embedded in paraplast, and cut into 10 µm sections that were mounted on ProbeOn Plus Slides. Probes were labeled using a DIG RNA Labeling Kit (Roche). Primer sequences for all genes are listed in Table S10.

Scanning and transmission electron microscopy

Scanning electron microscopy (SEM) and transmission electron microscopy (TEM) were performed as previously described (Zhang, Ruan, et al., 2017). Briefly, cotton seeds (0 DPA) were fixed in 2.5% (v/v) glutaraldehyde for 12 h, then dehydrated in a step-graded ethanol series. For SEM, the samples were further dried in a Hitachi HCP-2 critical point dryer and observed in a Hitachi SU-8010 SEM. For TEM, samples were washed in buffer and embedded in LR White resin (medium grade; Alltech) through a step-graded series. Infiltrated seeds were then polymerized in gelatin capsules at 70°C for ~2 h under nitrogen gas. Sections (90 nm thick) were cut with an ultramicrotome (EMUC7; Leica), collected on nickel mesh, stained with uranyl acetate for 10 min and lead citrate for 5 min, then washed with 0.01 M PBS (six times, 3 min each) and water (four times, 3 min each). Finally, the sections were viewed under a Hitachi H-7650 transmission electron microscope at 80 kV to observe the structure.

ACCESSION NUMBERS

All sequencing data generated in this study have been submitted to the SRA database (<https://www.ncbi.nlm.nih.gov/bioproject/>) with BioProject identification PRJNA869296 (Reviewer link: <https://dataview.ncbi.nlm.nih.gov/object/PRJNA869296?reviewer=o6svnj0ggehhqt40jnm2vpdn0r>). A supplementary online web server (http://cotton.zju.edu.cn/sc_browser/cottonfiber/) and a detailed user manual for the server (Supplementary Text) have also been developed to facilitate the use of our datasets. We created a download link (<http://cotton.zju.edu.cn/download.html>) for the scRNA-seq expression matrix. The bulk RNA-seq data of different fiber development stages are derived from our previously published article by Hu et al. (Hu et al., 2019). The code has been shared on GitHub to ensure transparency and reproducibility of the results (<https://github.com/ZhangJun->

CHN/Original-code-for-analysis-Spatial-Transcriptome-and-Single-Cell-RNA-Seq). Other data can be found within the manuscript and supporting materials.

AUTHOR CONTRIBUTIONS

TZ and JZ conceived the project. JZ, RC, and YS prepared protoplasts. JZ and YH participated in constructing the vector. YS performed the RNA *in situ* hybridization experiments. FD and JZ analyzed the data of spatial transcriptomic and single-cell datasets. TZ, JZ, YT, and RC wrote the manuscript. All authors discussed the results and commented on the manuscript. The authors read and approved the final manuscript.

ACKNOWLEDGEMENTS

We thank the Agricultural Experiment Station at Zhejiang University for their support in the greenhouse material planting and management. We also extend our gratitude to the Bio-ultrastructure Analysis Lab of the Analysis Center of Agrobiology and Environmental Sciences at Zhejiang University for their support with scanning electron microscopy and transmission electron microscopy. This work was supported by the Fundamental Research Funds for the Central Universities (226-2022-00100), the NSFC (32130075, 32200284), and the Natural Science Foundation of Jiangsu Province (BK20210877).

CONFLICT OF INTEREST

The authors declare no competing financial interests.

DATA AVAILABILITY STATEMENT

All relevant data can be found within the manuscript and its supporting materials.

SUPPORTING INFORMATION

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

Figure S1. The continuous process of epidermal protuberance in wild-type cotton ovules is depicted.

Figure S2. The spatial transcriptomic atlas of continuously developing ovules comprises several components.

Figure S3. The spatial distribution of distinct clusters.

Figure S4. Differential gene expression in spatial transcriptomic maps.

Figure S5. Sample and data processing for single cell.

Figure S6. Correlation analysis of single-cell data with bulk RNA-seq and single-cell data cluster heat map.

Figure S7. Annotation of proliferating cell types.

Figure S8. Integration analysis of single-cell data and spatial transcriptomics.

Figure S9. Quantification of expression after viral silencing of SVB, SVBL genes.

Figure S10. The clusters are presented individually on WT using Mfuzz.

Table S1. Spatial transcriptome capture information.

Table S2. Ovule spatial transcriptome marker genes in three stages.

Table S3. Ovule spatial transcriptome marker genes GO enrichment in three stages.

Table S4. Differential genes of epidermal/fiber cells between ovules in three stages.

Table S5. Summary of cells in different clusters.

Table S6. Marker genes of different clusters in WT.

Table S7. GO enrichment of different clusters in WT.

Table S8. Mfuzz cluster analysis in WT.

Table S9. GO enrichment in WT after Mfuzz cluster analysis.

Table S10. Primers used in the study.

Text S1. User manual for the online web server.

Text S2. User manual for the online spatial transcriptome web server.

REFERENCES

- Ando, A., Kirkbride, R.C., Jones, D.C., Grimwood, J. & Chen, Z.J. (2021) LCM and RNA-seq analyses revealed roles of cell cycle and translational regulation and homoeolog expression bias in cotton fiber cell initiation. *BMC Genomics*, **22**, 309.
- Beasley, C.A. & Ting, I.P. (1974) Effects of plant growth substances on *in vitro* fiber development from unfertilized cotton ovules. *American Journal of Botany*, **61**, 188–194.
- Birkenbihl, R.P., Diezel, C. & Somssich, I.E. (2012) *Arabidopsis WRKY33* is a key transcriptional regulator of hormonal and metabolic responses toward *Botrytis cinerea* infection. *Plant Physiology*, **159**, 266–285.
- Bourne, H.R., Sanders, D.A. & McCormick, F. (1991) The GTPase superfamily: conserved structure and molecular mechanism. *Nature*, **349**, 117–127.
- Cannoodt, R., Saedens, W., Sichen, D., Tavernier, S., Janssens, S., Guillemins, M. et al. (2016) SCORPIUS improves trajectory inference and identifies novel modules in dendritic cell development. *bioRxiv*. <https://doi.org/10.1101/079509>
- Chen, W., Li, Y., Zhu, S., Fang, S., Zhao, L., Guo, Y. et al. (2020) A retrotransposon insertion in *GhMML3_D12* is likely responsible for the lintless locus *li3* of tetraploid cotton. *Frontiers in Plant Science*, **11**, 593679.
- Chen, Y., Tong, S., Jiang, Y., Ai, F., Feng, Y., Zhang, J. et al. (2021) Transcriptional landscape of highly lignified poplar stems at single-cell resolution. *Genome Biology*, **22**, 319.
- Deng, F., Tu, L., Tan, J., Li, Y., Nie, Y. & Zhang, X. (2012) *GbPDF1* is involved in cotton fiber initiation via the core cis-element HDZIP2ATATHB2. *Plant Physiology*, **158**, 890–904.
- Deng, T., Yao, H., Wang, J., Wang, J., Xue, H. & Zuo, K. (2016) *GhLTPG1*, a cotton GPI-anchored lipid transfer protein, regulates the transport of phosphatidylinositol monophosphates and cotton fiber elongation. *Scientific Reports*, **6**, 26829.
- Denyer, T., Ma, X., Klesen, S., Scacchi, E., Nieselt, K. & Timmermans, M.C.P. (2019) Spatiotemporal developmental trajectories in the *Arabidopsis* root revealed using high-throughput single-cell RNA sequencing. *Developmental Cell*, **48**, 840–852.e845.
- Di, D.W., Li, G., Sun, L., Wu, J., Wang, M., Kronzucker, H.J. et al. (2021) High ammonium inhibits root growth in *Arabidopsis thaliana* by promoting auxin conjugation rather than inhibiting auxin biosynthesis. *Journal of Plant Physiology*, **261**, 153415.
- Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S. et al. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15–21.
- Du, J., Wang, Y., Chen, W., Xu, M., Zhou, R., Shou, H. et al. (2023) High-resolution anatomical and spatial transcriptome analyses reveal two types of meristematic cell pools within the secondary vascular tissue of poplar stem. *Molecular Plant*, **16**, 809–828.
- Fu, Y., Xiao, W., Tian, L., Guo, L., Ma, G., Ji, C. et al. (2023) Spatial transcriptomics uncover sucrose post-phloem transport during maize kernel development. *Nature Communications*, **14**, 7191.
- Gou, J.-Y., Wang, L.-J., Chen, S.-P., Hu, W.-L. & Chen, X.-Y. (2007) Gene expression and metabolite profiles of cotton fiber during cell elongation and secondary cell wall synthesis. *Cell Research*, **17**, 422–434.

- Guan, X.Y., Li, Q.J., Shan, C.M., Wang, S., Mao, Y.B., Wang, L.J. et al.** (2008) The HD-zip IV gene *GaHOX1* from cotton is a functional homologue of the *Arabidopsis GLABRA2*. *Physiologia Plantarum*, **134**, 174–182.
- Gulati, G.S., Sikandar, S.S., Wesche, D.J., Manjunath, A., Bharadwaj, A., Berger, M.J. et al.** (2020) Single-cell transcriptional diversity is a hallmark of developmental potential. *Science*, **367**, 405–411.
- Hafemeister, C. & Satija, R.** (2019) Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biology*, **20**, 296.
- Hu, H., He, X., Tu, L., Zhu, L., Zhu, S., Ge, Z. et al.** (2016) *GhJAZ2* negatively regulates cotton fiber initiation by interacting with the R2R3-MYB transcription factor *GhMYB25-like*. *The Plant Journal*, **88**, 921–935.
- Hu, Y., Chen, J., Fang, L., Zhang, Z., Ma, W., Niu, Y. et al.** (2019) *Gossypium barbadense* and *Gossypium hirsutum* genomes provide insights into the origin and evolution of allotetraploid cotton. *Nature Genetics*, **51**, 739–748.
- John, M.E. & Keller, G.** (1996) Metabolic pathway engineering in cotton: biosynthesis of polyhydroxybutyrate in fiber. *Proceedings of the National Academy of Sciences of the United States of America*, **93**, 12768–12773.
- Kumar, L. & Futschik, M.E.** (2007) Mfuzz: a software package for soft clustering of microarray data. *Bioinformatics*, **2**, 5–7.
- Kumar, V., Singh, B., Singh, S.K., Rai, K.M., Singh, S.P., Sable, A. et al.** (2018) Role of *GhHDA5* in H3K9 deacetylation and fiber initiation in *Gossypium hirsutum*. *The Plant Journal*, **95**, 1069–1083.
- Larkin, J.C., Oppenheimer, D.G. & Marks, M.D.** (1994) The *GL1* gene and the trichome developmental pathway in *Arabidopsis thaliana*. *Results and Problems in Cell Differentiation*, **20**, 259–275.
- Li, J., Wang, M., Li, Y., Zhang, Q., Lindsey, K., Daniell, H. et al.** (2019) Multi-omics analyses reveal epigenomics basis for cotton somatic embryogenesis through successive regeneration acclimation process. *Plant Biotechnology Journal*, **17**, 435–450.
- Li, Y., Tu, L., Pettolino, F.A., Ji, S., Hao, J., Yuan, D. et al.** (2016) *GbEXPATR*, a species-specific expansin, enhances cotton fibre elongation through cell wall restructuring. *Plant Biotechnology Journal*, **14**, 951–963.
- Liu, C., Leng, J., Li, Y., Ge, T., Li, J., Chen, Y. et al.** (2022) A spatiotemporal atlas of organogenesis in the development of orchid flowers. *Nucleic Acids Research*, **50**, 9724–9737.
- Liu, Z., Zhou, Y., Guo, J., Li, J., Tian, Z., Zhu, Z. et al.** (2020) Global dynamic molecular profiling of stomatal lineage cell development by single-cell RNA sequencing. *Molecular Plant*, **13**, 1178–1193.
- Machado, A., Wu, Y., Yang, Y., Llewellyn, D.J. & Dennis, E.S.** (2009) The MYB transcription factor *GhMYB25* regulates early fibre and trichome development. *The Plant Journal*, **59**, 52–62.
- Marks, M.D., Wenger, J.P., Gilding, E., Jilk, R. & Dixon, R.A.** (2009) Transcriptome analysis of *Arabidopsis* wild-type and *gl3-sst sim* trichomes identifies four additional genes required for trichome development. *Molecular Plant*, **2**, 803–822.
- McGinnis, C.S., Murrow, L.M. & Gartner, Z.J.** (2019) DoubletFinder: doublet detection in single-cell RNA sequencing data using artificial nearest neighbors. *Cell Systems*, **8**, 329–337.e324.
- Mohanty, C., Prasad, A., Cheng, L., Arkin, L.M., Shields, B.E., Drolet, B. et al.** (2024) SpatialView: an interactive web application for visualization of multiple samples in spatial transcriptomics experiments. *Bioinformatics*, **40**, btae117.
- Nigam, D., Kavita, P., Tripathi, R.K., Ranjan, A., Goel, R., Asif, M. et al.** (2014) Transcriptome dynamics during fibre development in contrasting genotypes of *Gossypium hirsutum* L. *Plant Biotechnology Journal*, **12**, 204–218.
- Pei, Y., Zhang, J., Wu, P., Ye, L., Yang, D., Chen, J. et al.** (2021) *GoNe* encoding a class VIIIb AP2/ERF is required for both extrafloral and floral necrotic development in *Gossypium*. *The Plant Journal*, **106**, 1116–1127.
- Peirats-Llobet, M., Yi, C., Liew, L.C., Berkowitz, O., Narsai, R., Lewsey, M.G. et al.** (2023) Spatially resolved transcriptomic analysis of the germinating barley grain. *Nucleic Acids Research*, **51**, 7798–7819.
- Procko, C., Lee, T., Borsuk, A., Bargmann, B.O.R., Dabi, T., Nery, J.R. et al.** (2022) Leaf cell-specific and single-cell transcriptional profiling reveals a role for the palisade layer in UV light protection. *Plant Cell*, **34**, 3261–3279.
- Pruitt, R.E., Vielle-Calzada, J.-P., Ploense, S.E., Grossniklaus, U. & Lolle, S.J.** (2000) *FIDDLEHEAD*, a gene required to suppress epidermal cell interactions in *Arabidopsis*, encodes a putative lipid biosynthetic enzyme. *Proceedings. National Academy of Sciences. United States of America*, **97**, 1311–1316.
- Pu, L., Li, Q., Fan, X., Yang, W. & Xue, Y.** (2008) The R2R3 MYB transcription factor *GhMYB109* is required for cotton fiber development. *Genetics*, **180**, 811–820.
- Qin, Y., Sun, M., Li, W., Xu, M., Shao, L., Liu, Y. et al.** (2022) Single-cell RNA-seq reveals fate determination control of an individual fibre cell initiation in cotton (*Gossypium hirsutum*). *Plant Biotechnology Journal*, **20**, 2372–2388.
- Qiu, X., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H.A. et al.** (2017) Reversed graph embedding resolves complex single-cell trajectories. *Nature Methods*, **14**, 979–982.
- Rerie, W.G., Feldmann, K.A. & Marks, M.D.** (1994) The *GLABRA2* gene encodes a homeo domain protein required for normal trichome development in *Arabidopsis*. *Genes and Development*, **8**, 1388–1399.
- Ruan, Y.-L., Llewellyn, D.J. & Furbank, R.T.** (2003) Suppression of sucrose synthase gene expression represses cotton fiber cell initiation, elongation, and seed development. *Plant Cell*, **15**, 952–964.
- Ruan, Y.L., Llewellyn, D.J., Furbank, R.T. & Chourey, P.S.** (2005) The delayed initiation and slow elongation of fuzz-like short fibre cells in relation to altered patterns of sucrose synthase expression and plasmodesmata gating in a lintless mutant of cotton. *Journal of Experimental Botany*, **56**, 977–984.
- Rubinovich, L. & Weiss, D.** (2010) The *Arabidopsis* cysteine-rich protein GASA4 promotes GA responses and exhibits redox activity in bacteria and in *planta*. *The Plant Journal*, **64**, 1018–1027.
- Ryu, K.H., Huang, L., Kang, H.M. & Schiefelbein, J.** (2019) Single-cell RNA sequencing resolves molecular relationships among individual plant cells. *Plant Physiology*, **179**, 1444–1456.
- Sable, A., Rai, K.M., Choudhary, A., Yadav, V.K., Agarwal, S.K. & Sawant, S.V.** (2018) Inhibition of heat shock proteins *HSP90* and *HSP70* induce oxidative stress, suppressing cotton fiber development. *Scientific Reports*, **8**, 3620.
- 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.
- Seyfferth, C., Renema, J., Wendrich, J.R., Eekhout, T., Seurinck, R., Vandamme, N. et al.** (2021) Advances and opportunities in single-cell transcriptomics for plant research. *Annual Review of Plant Biology*, **72**, 847–866.
- Shan, C.M., Shangguan, X.X., Zhao, B., Zhang, X.F., Chao, L.M., Yang, C.Q. et al.** (2014) Control of cotton fibre elongation by a homeodomain transcription factor *GhHOX3*. *Nature Communications*, **5**, 5519.
- Song, Q., Ando, A., Jiang, N., Ikeda, Y. & Chen, Z.J.** (2020) Single-cell RNA-seq analysis reveals ploidy-dependent and cell-specific transcriptome changes in *Arabidopsis* female gametophytes. *Genome Biology*, **21**, 178.
- Song, X., Guo, P., Xia, K., Wang, M., Liu, Y., Chen, L. et al.** (2023) Spatial transcriptomics reveals light-induced chlorenchyma cells involved in promoting shoot regeneration in tomato callus. *Proceedings. National Academy of Sciences. United States of America*, **120**, e2310163120.
- Stewart, J.M.** (1975) Fiber initiation on the cotton ovule (*Gossypium hirsutum*). *American Journal of Botany*, **62**, 723–730.
- Stracke, R., Werber, M. & Weisshaar, B.** (2001) The R2R3-MYB gene family in *Arabidopsis thaliana*. *Current Opinion in Plant Biology*, **4**, 447–456.
- Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., 3rd et al.** (2019) Comprehensive integration of single-cell data. *Cell*, **177**, 1888–1902.
- Tian, Y. & Zhang, T.** (2021) MIXTAs and phytohormones orchestrate cotton fiber development. *Current Opinion in Plant Biology*, **59**, 101975.
- Walford, S.A., Wu, Y., Llewellyn, D.J. & Dennis, E.S.** (2011) *GhMYB25-like*: a key factor in early cotton fibre development. *The Plant Journal*, **65**, 785–797.
- Walford, S.A., Wu, Y., Llewellyn, D.J. & Dennis, E.S.** (2012) Epidermal cell differentiation in cotton mediated by the homeodomain leucine zipper gene, *GhHD-1*. *The Plant Journal*, **71**, 464–478.
- Wan, Q., Guan, X., Yang, N., Wu, H., Pan, M., Liu, B. et al.** (2016) Small interfering RNAs from bidirectional transcripts of *GhMML3_A12* regulate cotton fiber development. *The New Phytologist*, **210**, 1298–1310.
- Wang, D., Hu, X., Ye, H., Wang, Y., Yang, Q., Liang, X. et al.** (2023) Cell-specific clock-controlled gene expression program regulates rhythmic fiber cell growth in cotton. *Genome Biology*, **24**, 49.

- Wang, H.Y., Wang, J., Gao, P., Jiao, G.L., Zhao, P.M., Li, Y. et al. (2009) Down-regulation of *GhADF1* gene expression affects cotton fibre properties. *Plant Biotechnology Journal*, **7**, 13–23.
- Wang, S., Wang, J.W., Yu, N., Li, C.H., Luo, B., Gou, J.Y. et al. (2004) Control of plant trichome development by a cotton fiber MYB gene. *Plant Cell*, **16**, 2323–2334.
- Wei, R., He, S., Bai, S., Sei, E., Hu, M., Thompson, A. et al. (2022) Spatial charting of single-cell transcriptomes in tissues. *Nature Biotechnology*, **40**, 1190–1199.
- Williams, M.E. & Sussex, I.M. (1995) Developmental regulation of ribosomal protein L16 genes in *Arabidopsis thaliana*. *The Plant Journal*, **8**, 65–76.
- Xi, J., Zeng, J., Fu, X., Zhang, L., Li, G., Li, B. et al. (2023) *GhROP6* GTPase modulates auxin accumulation in cotton fibers by regulating cell-specific *GhPIN3a* localization. *Journal of Experimental Botany*, **74**, 265–282.
- Xu, W.L., Wang, X.L., Wang, H. & Li, X.B. (2007) Molecular characterization and expression analysis of nine cotton *GhEF1A* genes encoding translation elongation factor 1A. *Gene*, **389**, 27–35.
- 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.e556.
- Yang, T.H., Lenglet-Hilfiker, A., Stolz, S., Glauser, G. & Farmer, E.E. (2020) Jasmonate precursor biosynthetic enzymes LOX3 and LOX4 control wound-response growth restriction. *Plant Physiology*, **184**, 1172–1180.
- Yephremov, A., Wisman, E., Huijser, P., Huijser, C., Wellesen, K. & Saedler, H. (1999) Characterization of the *FIDDLEHEAD* gene of *Arabidopsis* reveals a link between adhesion response and cell differentiation in the epidermis. *Plant Cell*, **11**, 2187–2201.
- Yu, C.Y., Sharma, O., Nguyen, P.H.T., Hartono, C.D. & Kanehara, K. (2021) A pair of DUF538 domain-containing proteins modulates plant growth and trichome development through the transcriptional regulation of *GLABRA1* in *Arabidopsis thaliana*. *The Plant Journal*, **108**, 992–1004.
- Zhang, J., Si, Z., Chen, R., Liu, W., Shi, Y., Shi, Z. et al. (2023) A new model system for cotton indoor genetic and genomic research. *Science China Life Sciences*, **66**, 1444–1446.
- Zhang, M., Zeng, J.Y., Long, H., Xiao, Y.H., Yan, X.Y. & Pei, Y. (2017) Auxin regulates cotton fiber initiation via *GhPIN*-mediated auxin transport. *Plant & Cell Physiology*, **58**, 385–397.
- Zhang, T., Hu, Y., Jiang, W., Fang, L., Guan, X., Chen, J. et al. (2015) Sequencing of allotetraploid cotton (*Gossypium hirsutum* L. acc. TM-1) provides a resource for fiber improvement. *Nature Biotechnology*, **33**, 531–537.
- Zhang, T.Q., Chen, Y., Liu, Y., Lin, W.H. & Wang, J.W. (2021) Single-cell transcriptome atlas and chromatin accessibility landscape reveal differentiation trajectories in the rice root. *Nature Communications*, **12**, 2053.
- Zhang, T.Q., Chen, Y. & Wang, J.W. (2021) A single-cell analysis of the *Arabidopsis* vegetative shoot apex. *Developmental Cell*, **56**, 1056–1074.e1058.
- 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.
- Zhang, Z., Ruan, Y.L., Zhou, N., Wang, F., Guan, X., Fang, L. et al. (2017) Suppressing a putative sterol carrier gene reduces plasmodesmal permeability and activates sucrose transporter genes during cotton fiber elongation. *Plant Cell*, **29**, 2027–2046.
- Zhu, H., Lv, J., Zhao, L., Tong, X., Zhou, B., Zhang, T. et al. (2012) Molecular evolution and phylogenetic analysis of genes related to cotton fibers development from wild and domesticated cotton species in *Gossypium*. *Molecular Phylogenetics and Evolution*, **63**, 589–597.
- Zou, X., Ali, F., Jin, S., Li, F. & Wang, Z. (2022) RNA-seq with a novel glabrous-ZM24fl reveals some key lncRNAs and the associated targets in fiber initiation of cotton. *BMC Plant Biology*, **22**, 61.