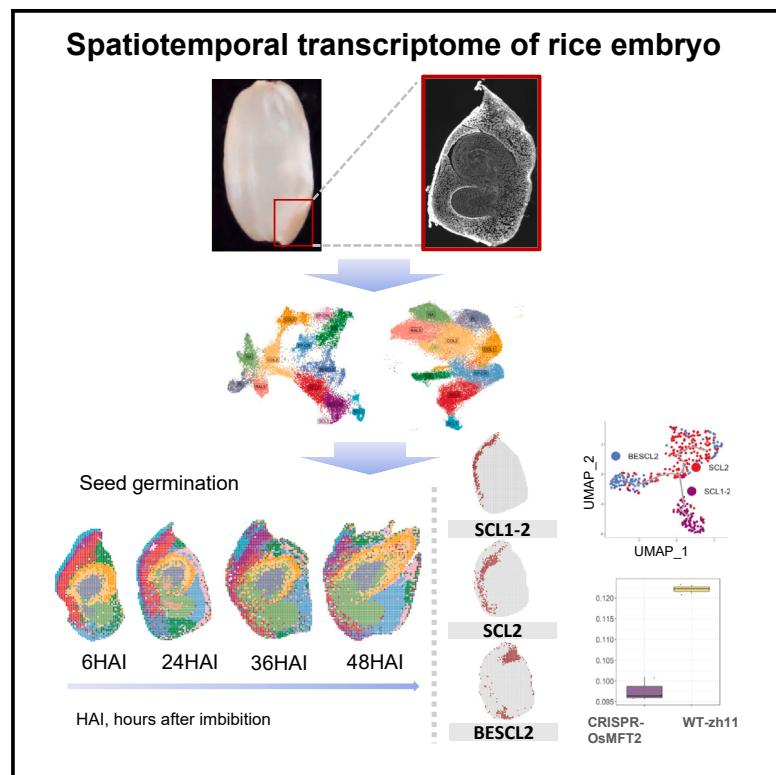


Developmental Cell

Spatiotemporal transcriptomic landscape of rice embryonic cells during seed germination

Graphical abstract



Authors

Jie Yao, Qinjie Chu, Xing Guo, ...,
Xun Xu, Huan Liu, Longjiang Fan

Correspondence

liuhuan@genomics.cn (H.L.),
fanlj@zju.edu.cn (L.F.)

In brief

Yao et al. establish the spatiotemporal transcriptomic cell census of rice seed embryo during germination. Employing cell bin and segmentation, they identify transcriptional disparities among embryo cells and delineate scutellum parenchyma cell developmental trajectory, as well as dynamics of pathways related to cell, hormones, and nutrition metabolism during germination.

Highlights

- Establish a rice embryo germination spatiotemporal transcriptomic census
- Develop AI-based plant cell wall segmentation for precise cell-type identification
- Identify scutellum cell subtypes and cell-type-specific gene expression pattern
- OsMFT2 specifically expressed in SCL2, with decreasing trends during seed germination



Article

Spatiotemporal transcriptomic landscape of rice embryonic cells during seed germination

Jie Yao,^{1,2,9} Qinjie Chu,^{1,9} Xing Guo,^{3,4,9} Wenwen Shao,^{3,4,9} Nianmin Shang,¹ Kang Luo,⁵ Xiaohan Li,¹ Hongyu Chen,¹ Qing Cheng,¹ Fangyu Mo,² Dihuai Zheng,¹ Fan Xu,⁶ Fu Guo,² Qian-Hao Zhu,⁷ Shuiguang Deng,⁵ Chengcai Chu,⁸ Xun Xu,³ Huan Liu,^{3,*} and Longjiang Fan^{1,2,10,*}

¹Institute of Crop Science & Institute of Bioinformatics, Zhejiang University, Hangzhou 310058, China

²Hainan Institute, Zhejiang University, Sanya 572025, China

³BGI Research, Shenzhen 518103, China

⁴BGI Research, Wuhan 430074, China

⁵College of Computer Science and Technology & Polytechnic Institute, Zhejiang University, Hangzhou 310015, Zhejiang, China

⁶College of Agronomy and Biotechnology, Southwest University, Chongqing 400715, China

⁷CSIRO, Agriculture and Food, Canberra, ACT 2601, Australia

⁸Guangdong Laboratory for Lingnan Modern Agriculture, State Key Laboratory for Conservation and Utilization of Subtropical Agro-Bioresources, College of Agriculture, South China Agricultural University, Guangzhou 510642, China

*These authors contributed equally

¹⁰Lead contact

*Correspondence: liuhuan@genomics.cn (H.L.), fanlj@zju.edu.cn (L.F.)

<https://doi.org/10.1016/j.devcel.2024.05.016>

SUMMARY

Characterizing cellular features during seed germination is crucial for understanding the complex biological functions of different embryonic cells in regulating seed vigor and seedling establishment. We performed spatially enhanced resolution omics sequencing (Stereo-seq) and single-cell RNA sequencing (scRNA-seq) to capture spatially resolved single-cell transcriptomes of germinating rice embryos. An automated cell-segmentation model, employing deep learning, was developed to accommodate the analysis requirements. The spatial transcriptomes of 6, 24, 36, and 48 h after imbibition unveiled both known and previously unreported embryo cell types, including two unreported scutellum cell types, corroborated by *in situ* hybridization and functional exploration of marker genes. Temporal transcriptomic profiling delineated gene expression dynamics in distinct embryonic cell types during seed germination, highlighting key genes involved in nutrient metabolism, biosynthesis, and signaling of phytohormones, reprogrammed in a cell-type-specific manner. Our study provides a detailed spatiotemporal transcriptome of rice embryo and presents a previously undescribed methodology for exploring the roles of different embryonic cells in seed germination.

INTRODUCTION

Seeds are biologically and economically vital because they protect and disseminate genetic information, ensure the continuation of plant life, promote plant diversity and ecological balance, and provide food for human beings and feed for livestock. Embryo, endosperm, and seed coat are the three major components of a seed, each playing a distinguishable yet essential role in seed viability. The rice (*Oryza sativa*) seed embryo consists of scutellum (SC), coleoptile (CO), plumule (PL), radicle (RA), and epiblast-coleorhiza (EP-CR) region without clear demarcation.^{1,2} Meanwhile, in rice root and leaf, three body layers are established from stem cells. Layer one (L1) is the single-cell epidermal layer; the root cortex/endodermis and the leaf mesophyll make up layer two (L2); layer 3 (L3) is where vascular tissue occurs.³

Seed germination is the first step of establishing a new life and involves intricate and dynamic biological processes, including significant morphological changes, cell division, and cell-state transition. Through germination, the seed embryo provides the genetic information and developmental abilities essential for proper morphogenesis of shoot and root tissues, which involves the activation of many intercellular metabolic pathways, intense signal transduction, and frequent energy synthesis.² Seed germination begins with water imbibition and ends with RA emergence. Rice seed imbibition can be divided into 0–24, 24–48, and 48–72 h after imbibition (HAI), three phases based on the features of water uptake and biochemical changes.⁴ The most significant transcriptome alterations occur during the first two phases (0–24 and 24–48 HAI). Phase I (0–24 HAI) is characterized by rapid water absorption and DNA repair, and phase II (24–48 HAI) is characterized by mitochondrion synthesis and



the translation of stored mRNA.⁵ Each cell type of embryo is speculated to have specialized functions in seed germination by response to environmental stimuli. Nevertheless, we know little about how the functions of different types of embryonic cells are regulated during seed germination and how they communicate with one another to build the complex form of new life.

Single-cell RNA sequencing (scRNA-seq) enables interrogating transcriptomic changes at a single-cell resolution, which enhances our understanding and knowledge of plant body cell composition and cell lineage development.^{6–8} However, application of scRNA-seq in analysis of rice embryos is still a challenge because of (1) biased gene expression induced by the protoplast preparation process,⁹ (2) biased cell-type ratios due to protoplast resistance,¹⁰ and (3) loss of spatial information of cells as a result of tissue dissociation.¹¹ The cutting-edge technologies of spatially resolved transcriptomes provide opportunity to characterize spatially distinct cell types, especially those without known marker genes, and to assign functions to cells in the proper locations. Several high-throughput *in situ* gene expression profiling techniques (e.g., the 10 \times Genomics Visium platform, slide-seq, and high-definition spatial transcriptomics (HDST)) have been developed in recent years.¹² The latest spatial transcriptomic technique, spatially enhanced resolution omics sequencing (Stereo-seq), can identify the spatial and temporal transcriptomic changes and interactions of cells with subcellular resolution (about 500 nm) and large field of view (up to cm scale).¹³ Despite that the advantages of Stereo-seq have been demonstrated by recent studies in establishing the spatiotemporal transcriptome atlas of mouse organogenesis¹⁴ and in identifying region-specific cell subtypes in *Arabidopsis* leaves,¹¹ its application in plant spatiotemporal transcriptomic study is still in the infancy stage, and no study has integrated Stereo-seq with scRNA-seq in plants.

In this study, we utilized Stereo-seq and scRNA-seq to examine the dynamic spatial-temporal transcriptomic landscapes of germinating rice embryos by using samples collected from four germination time points (i.e., 6, 24, 36, and 48 HAI). We also developed an automated cell-segmentation model based on deep learning to enhance annotation of cell types. To validate the identified marker genes through Stereo-seq and scRNA-seq, we incorporated *in situ* hybridization (ISH) and generated knockout lines of marker genes. Our results, including embryonic transcriptomic changes by bulk RNA-seq and seed germination rate of the mutation line of the marker gene OsMFT2 for the PL region, support our annotation about marker genes for cell types. Our study provides valuable insights into the gene expression patterns, cellular architecture, and developmental pathways of germinating rice embryos, as well as tools for analysis of spatiotemporal transcriptome in plants.

RESULTS

High-quality Stereo-seq and scRNA-seq data from rice embryos

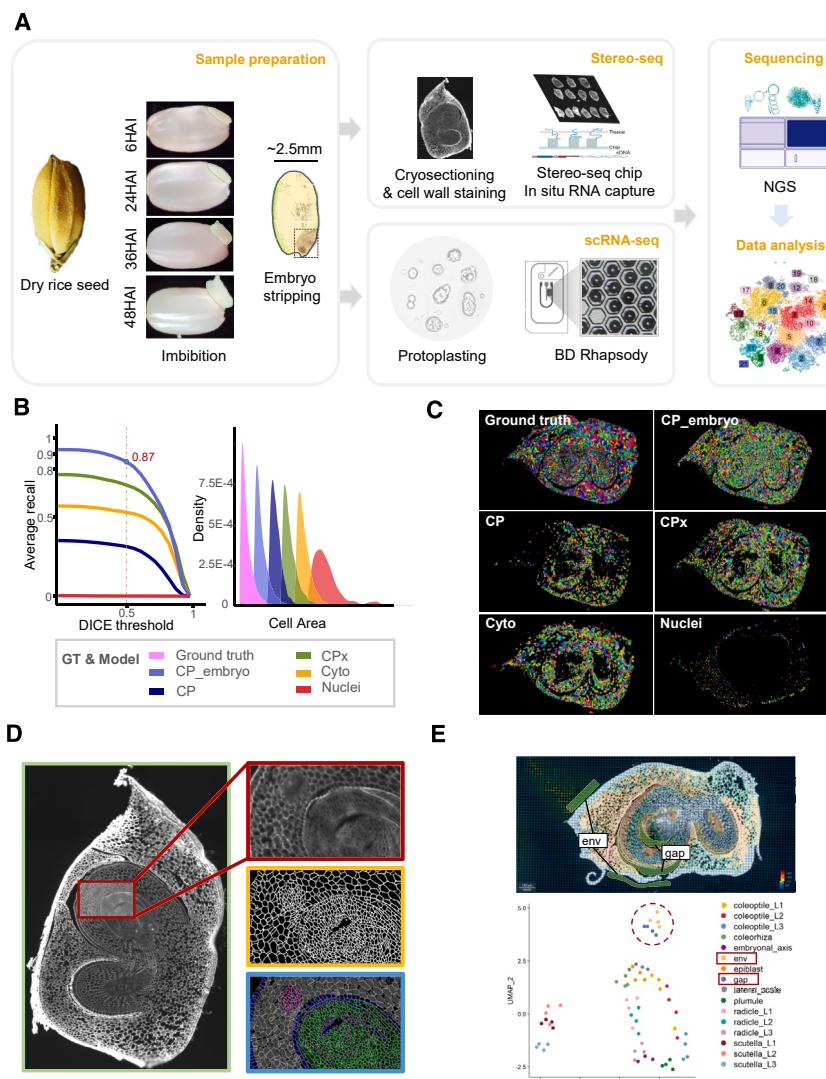
Rice embryos collected from 6, 24, 36, and 48 HAI were used in the generation of Stereo-seq transcriptomic data. We prepared cryosections of rice embryos at a 10- μ m thickness to capture roughly a single-cell layer of tissue. Then, we utilized the fluorescent brightener (FB) staining technique to visualize the cell walls

and placed multiple sections on one Stereo-seq chip for RNA permeabilization. Each DNA nanoball (DNB) spot on the chip has a diameter of 220 nm, and the center-to-center distance between neighboring spots is 500 nm, allowing capture of transcripts at the subcellular level (Figure 1A). The RNA from a total of 52 embryo tissue sections captured in 4 chips was collected. The transcript sequencing data were generated and visualized in the Spatial Transcript Omics (STOmics) visualization system. The FB images and Stereo-seq transcriptomic data can be browsed via our interactive data portal at <http://ibi.zju.edu.cn/REGSTA/>.

The mean number of the unique raw mRNA molecular per Bin50 (50 \times 50 DNB bins) on the four chips was 1,200, and the mean unique gene features for every Bin50 was 583 based on the Stereo-seq data (Table S1). To minimize the incompleteness of cell-type sampling and batch effects, for each time point, only the sections with a comprehensive morphological structure and optimal gene expression pattern were selected for further analysis. As a result, a total of 14 embryo sections (4, 4, 3, and 3 from 6, 24, 36, and 48 HAI, respectively) were retained for subsequent analysis. We assessed the molecular identifiers' (MIDs) distribution of the 14 sections to confirm the transcript signals in the target tissue areas. Nearly all the MIDs overlapped with the embryo regions (Figure S1A, the middle graph of each sample group). Gene expression levels in the PL and RA areas were significantly higher, compared with other areas, during the primary stages of germination (Figure S1A, the right plot of each sample group), consistent with previously established knowledge.^{16,17} For each germination time point, a high degree of correlation was evident among the Stereo-seq replicates (Pearson's correlation coefficient ≥ 0.9) (Figure S1B), demonstrating the stability of the Stereo-seq technology and the validity of the data. Between different time points, a slightly lower correlation was noticed between 6 HAI and other time points, suggesting a significant change in gene expression from 6 to 24 HAI.

We also performed scRNA-seq with rice embryos collected from the four germination time points (i.e., 6, 24, 36, and 48 HAI) (Figure 1A) by using protoplasts from ~200 rice embryos at each time point. The average protoplast viability rate was 87.64%. scRNA-seq data were obtained for a total of 27,099 embryo cells (Table S3). The data were filtered at both cell and gene levels (see quality control in STAR Methods), resulting in a pool of 27,070 cells with 37,458 genes for further analysis. During the time course of seed germination, the mean molecules per cell increased from 5,458 (6 HAI) to 10,422 (48 HAI), and the mean gene feature per cell increased from 1,940 (6 HAI) to 3,117 (48 HAI), suggesting that the gene expression and metabolism in the embryo were gradually activated during germination (Table S3). High correlation was observed among the four scRNA-seq datasets from the four germination time points (Figure S1B). The correlation coefficient between scRNA-seq and Stereo-seq was approximately 0.5, likely due to the impact of enzymatic digestion during protoplast preparation (Figure S1B).

In brief, we generated high-quality transcriptomic data from 14 rice embryo sections by Stereo-seq and from 27,070 rice embryo cells by scRNA-seq at 4 germination time points, allowing us to generate the spatiotemporal transcriptomic atlas and dynamic gene expression pattern in germinating rice embryo.

**Figure 1. An overview of this study**

(A) Workflow of Stereo-seq and scRNA-seq data acquisition from germinating rice embryos. Rice embryos were stripped from 6-, 24-, 36-, and 48-HAI seeds, and half of them were used in Stereo-seq. Embryos were paraffin-embedded and sectioned. The sections were stained for cell wall observation. *In situ* RNA capture was performed on the sections loaded onto the Stereo-seq chip. Finally, cDNA amplification, library construction, and sequencing were conducted. Meanwhile, embryos of the four time points were used in scRNA-seq.

(B) Performance of different cell-segmentation models. The models of “nuclei,” “cyto,” CP, and “CPx” used images of nuclei-stained, cytoplasm fluorescent, or diverse cells as the training dataset,¹⁵ respectively. The CP_embryo is an updated model by training the CP model, using the rice embryo cell wall images generated by this study. Left panel: the DICE threshold measures similarity between the predicted and the ground-truth masks, with 1.0 being a perfect match. Right panel: predicted cell area distribution relative to the ground truth. The CP_embryo model had the closest cell area distribution relative to the ground truth and the highest average recall of cells predicted from the ground truth.

(C) Comparison of cell-segmentation results from the same image by different models.

(D) An example of cell segmentation by CP_embryo. The right-top panel displays details of cell structure in the plumule (PL) and coleoptile (CO) areas. The right-middle panel displays the predicted cell-segmentation result, and the right-bottom one shows the comparison of predicted cell-segmentation result with the ground truth, respectively.

(E) The top panel illustrates examples of lasso sampling in cell gaps (gap) and environment (env) regions, and the bottom panel displays the expression correlation analysis among cell gaps, environment, and various other cell regions. Different colors refer to different cell types.

Automated cell segmentation of rice embryos

In order to effectively capture RNAs from each cell, we established a pipeline for automatic segmentation of rice embryo cells (Figure S1C). We first utilized the FB staining technique to visualize the cell walls of the samples. To evaluate the accuracy of cell-segmentation model, a proprietary dataset was created using our in-house images that have clear cell wall staining. The dataset was also used to train a highly efficient and precisely automated cell-segmentation model. A set of metrics, including the recall curve of DICE coefficient¹⁸ and the parameters of the predicted vs. the ground truth in terms of cell area, was used to assess the efficacy of the cell-segmentation models (Figure 1B; see details below).

Of the five models tested, the “CP_embryo” model, which was constructed by integrating the pre-trained “CP” model provided by CellPose2.0¹⁵ and the FB images from our in-house rice embryo sections, had the best performance. When the threshold was set at 0.5, more than 87% of the cells in the test image set were predicted accurately by the CP_embryo model

(Figures 1B, left and 1C). Furthermore, the model also exhibited the closest match in cell area distribution when compared with the ground truth (Figures 1C, right and 1D). We registered images of the cell-segmentation results and the expression data from the STOmics visualization system to obtain spatial transcriptomic information at single-cell resolution. The cell-segmentation approach could obtain a higher number of captured cells but lower number of detected genes per cell, compared with the bin-based method (see next section for details).

We noted that employing the cell-segmentation method aids in mitigating the influence of tissue voids on data analysis. For example, expression in the gap region more closely resembles environmental expression rather than dispersed tissue region expression (e.g., Figure 1E). Meanwhile, cell segmentation allows for a more accurate assessment of the variations brought about by the cell sizes in different tissue regions of the rice seed embryo. The cells in rice embryos exhibit a diverse array of sizes. The SC L2 had the largest cells, with a typical diameter of around 30 μm , while the PL and RA regions included the

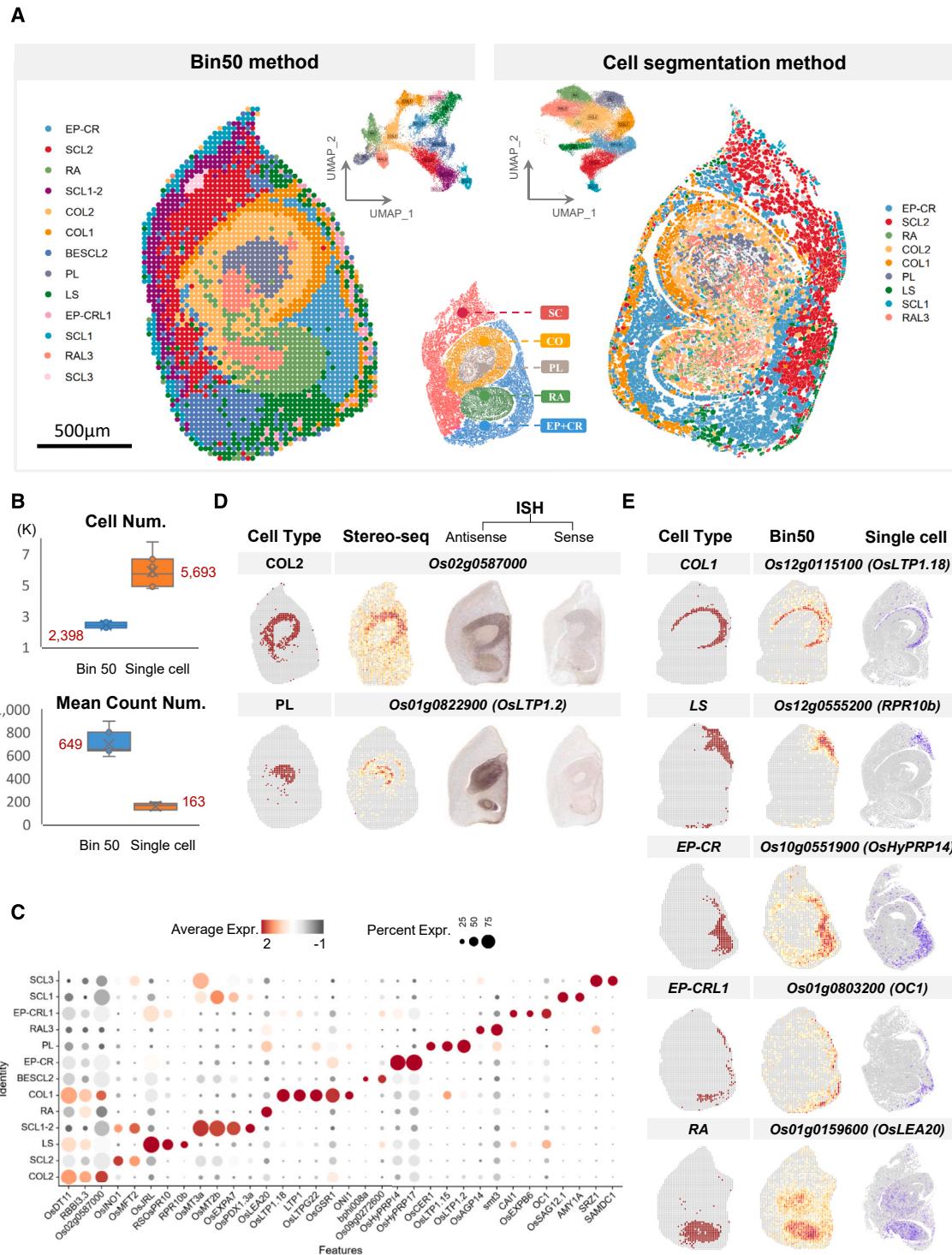


Figure 2. Annotation of cell types of rice embryo

(A) An example of cell clustering and cell-type annotation based on the 24-HAI spatial transcriptomic data. The middle panel shows the five anatomical regions of rice embryo, including scutellum (SC), coleoptile (CO), plumule (PL), radicle (RA), and epiblast-coleorhiza (EP-CR). The left and right panels show the UMAP clustering results of embryo cells and their spatial locations and cell-type annotation, using bin-based and cell segmentation-based method, respectively. Details about the cell types were described in the main text. Scale bar represent 500 μm.

(B) Comparison of the number of cells per sample and the count per cell produced by the Bin50 and single-cell segmentation method. Error bars represent standard deviation (SD) calculated from independent slides of embryo samples.

(legend continued on next page)

smallest cells, with a diameter of only about 10 μm (Figure S1D). It is evident that the actual cell count per Bin50 in the PL region is 9.20, and in contrast, the SC parenchyma hosts 0.88 cells (Figure S1E). Regarding the diverse cell sizes in rice embryos, the Bin50 method (each bin measuring approximately 25 μm)¹⁴ was adopted as the reference method.

Spatial clustering and cell-type classification of rice embryo

Based on histological characteristics and spatial information, the rice embryos were first divided into five anatomical regions (PL, RA, CO, SC, and EP-CR) (Figure 2A, bottom of the middle panel). To dissect cell-type composition, we first conducted an unsupervised clustering analysis of spatial data using Seurat, based solely on gene expression.¹⁹ Taking the 24-HAI embryonic sections as a reference, we conducted a bin-based cell clustering analysis and identified 13 distinct groups based on the clustering results in a uniform manifold approximation and projection (UMAP) plot (Figure 2A). These clusters were then mapped onto the embryonic sections based on their spatial information, and the 13 cell clusters could be well annotated as different cell types in the five anatomical regions: the PL, CO epidermis/layer 1 (COL1), CO parenchyma/layer 2 (COL2), RA, RA vascular/L3 (RAL3), EP-CR, EP-CR epidermis/layer 1 (EP-CRL1), lateral scale (LS), SC epidermis/layer 1 (SCL1), SC parenchyma/layer 2 (SCL2), SC L1-L2 interlayer (SCL1-2), SC vascular/L3 (SCL3), and both ends of SC parenchyma/layer 2 (BESCL2). A good consistency between cell-type identification and spatial arrangement across different sections was achieved (Figures 2A, left and S2A). We observed that the spatial proximity of cell types in the UMAP diagram was often congruent with their proximity in physical coordinates. Based on other HAI embryonic sections, one more cell type, root cap (RC), was further identified, which resulted in a total of 14 cell types annotated by this study. We speculate that the RC cell type exists in all four time points of rice seed embryo examined by this study. However, due to the angle of sectioning, not all sections have a view of the RC.

Based on the cell-segmentation data of the same 24-HAI samples, we also performed unsupervised clustering analysis and spatial annotation and compared the results with the results obtained by the Bin50 method. The number of cells captured by the cell-segmentation approach ($n = 5,693$) was higher than that captured by the Bin50 method ($n = 2,398$) (Figure 2B). The average number of detected genes per cell was significantly lower in the cell-segmentation approach ($n = 163$) than in the Bin50 method ($n = 649$). Cell-segmentation data, based clustering results, identified 9 of the 13 cell types mentioned above (Figure 2A, right panel). These results suggest that both methods can generally be used to define cell types, based on spatial transcriptomic data. The Bin50 method identified significant variations in the number of RNA molecules detected between different cell types. For instance, in SCL3, the median RNA count per Bin50 unit is 952.5, while in SCL1, it is only 317 (Figure S2B,

left panel). In contrast, the cell-segmentation method yielded uniform RNA counts across different cell types, indicating that cell-segmentation enhances the comparability of gene expression analysis among cells of different sizes (Figure S2B, right panel). In other hands, considering the PL, where a single Bin50 contains 9.2 cells, the Bin50 method identified 876 RNA counts per PL region. Further calculations identified that the true RNA count detected per cell is approximately 95, close to the count of 85 obtained using the cell-segmentation method. This further demonstrates the precision of our cell-segmentation approach.

To ensure the fidelity of our cell-type annotation, we investigated the molecular traits of each spatially resolved cell type. We validated the expression patterns of some cell-type marker genes reported in previous studies, such as *qLTG3-1* (*Os03g0103300*) (Figure S2C), which is known to be specifically expressed in the CR and EP regions.²⁰ Additionally, we identified many marker genes by analyzing the spatial transcriptome data. Analyses of the differentially expressed genes (DEGs) among the 13 cell types confirmed that each major cell type expressed a unique set of marker genes (the expression patterns of the cardinal markers were displayed in Figure 2C). Further, the results of RNA ISH well validated the spatial distribution of the marker genes identified by the Stereo-seq (Figure 2D).

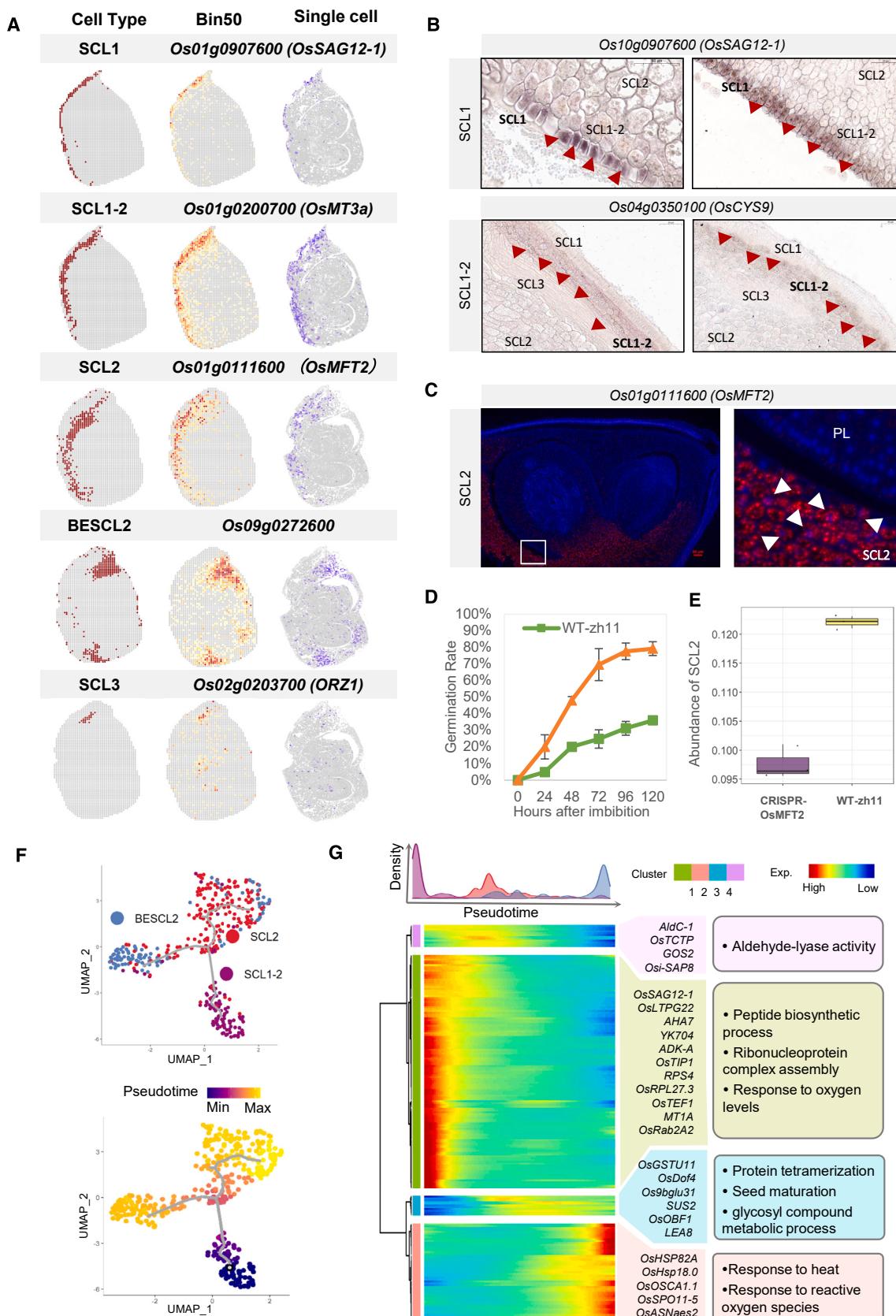
We observed that the genes encoding plant lipid transfer proteins (LTPs) exhibited cell-type-specific expression patterns. For instance, *OsLTP1.15* (*Os12g0114500*) and *OsLTP1.2* (*Os01g0822900*), which are involved in regulating cell division and elongation,²⁰ were explicitly expressed in PL; *OsLTP2.6* (*Os06g0705400*) was mainly expressed in COL2; and *OsHyPRP14/20* (*Os10g0551900*, *Os10g0552800*) were specifically expressed in EP-CR. The results from fluorescence *in situ* hybridization (FISH) also demonstrated that *OsHyPRP14* is specifically and highly expressed in the EP-CR region (Figure S2D). *OsLTP1.18* (*Os12g0115100*), *LTP1* (*Os11g0115400*), and *OsLTPG22* (*Os08g0532800*), which were involved in storage of seed lipids,²¹ were highly expressed in COL1. These results suggest that different LTP-encoding genes play distinct roles in germinating rice seeds by modulating lipid metabolism and transport in the corresponding cell types (Figures 2E and S2C).

Several other genes were also specifically expressed in one embryonic region. For example, *OsLEA20* (*Os01g0159600*), encoding a late embryogenesis abundant (LEA) protein, was highly expressed in RA (Figure 2E). Fatty acid elongase-coding gene *ONI1* (*Os03g0181500*), which is related to shoot development,²² had a specific expression pattern in COL1. Moreover, a homolog of *CER1* (*OsCER1*), which regulates anther development and plastid differentiation in rice,²³ showed high expression levels in the PL of germinating embryos. Gene Ontology (GO) enrichment analysis indicated that pathways such as “DNA package complex” ($p = 4.39\text{E}-4$) and “chromatin assembly or disassembly” ($p = 2.19\text{E}-4$) were enriched in PL, implying a high proliferative capacity of this region (Figure S2E).

(C) Dotplot showing the expression pattern of the selected putative cell-type marker genes. The circle size indicates the relative percentage of cells expressing the marker genes, and the color represents the relative expression level.

(D) The spatial expression pattern and RNA ISH validation results of two selected genes.

(E) Visualization of the spatial expression pattern of the selected marker genes in different cell types based on Bin50 and single-cell unit. The deeper the color, the higher the value.



Cell subdivision and trajectory inference of the SC region

Notably, the cells of the SC region were clustered into five groups that could be further annotated as five cell types (SCL1, SCL1-2, SCL2, BESCL1-2, and SCL3), based on the spatial transcriptomic data. Using the markers, we confidently identified SCL1 and SCL3, the two known cell types (Figure 3A). In SCL1, genes encoding cysteine protease *OsSAG12-1* (*Os01g0907600*) and alpha-amylase glycoprotein *AMY1A* (*Os02g0765600*) were highly expressed (Figures 3A and S2C), consistent with starch breakdown in the endosperm and transportation of nutrient from endosperm to embryo. We utilized ISH experiments to substantiate the specific high expression of *OsSAG12-1* in SCL1 (Figure 3B). SCL3 represents plant vascular tissue within the SC. Due to the limitations of the sectioning angle, this cell type is hard to be identified. We were lucky to observe SCL3 in some of our sections (Figure 3A). Several marker genes such as *SRZ1* (*Os03g0836200*), encoding a zinc-finger protein, and *SAMDC1* (*Os09g0424300*), encoding the S-adenosylmethionine decarboxylase, were also identified in SCL3 (Figure 2C), and both exhibited high expression levels in various parts of the embryo's vascular tissue.

Apart from the known SCL1 and SCL3, the spatial transcriptomic data suggested the presence of three more cell types, i.e., SCL1-2, SCL2, and BESCL2, in the SC parenchyma (Figure 3A). The cellular composition of SCL2 includes multiple layers of parenchyma cells, located closer to the PL and RA within the SC. We identified several marker genes for the SCL2 cell type, such as *Os11g0496500* and *OsMFT2* (*Os01g0111600*). FISH verified a distinctively high expression of *OsMFT2* in the SCL2 region (Figure 3C). *OsMFT2* is recognized as part of the phosphatidylethanolamine-binding protein (PEBP) family, known for its role in modulating abscisic acid (ABA) signaling-mediated seed germination.²⁴ Subsequent time-series analysis through Stereo-seq also highlighted a consistent decrease of *OsMFT2* expression during the germination process from 6 to 48 HAI (Figure S4D; see details in the next section). To delve deeper into the impact of *OsMFT2* on the germination process, we generated an *OsMFT2* knockout line. Germination phenotype experiments indicated that the *OsMFT2* knockout lines showed faster germination, compared with the wild type (WT) (Figure 3D), i.e., reached a 30% germination rate in approximately 30 h, while the WT required about 96 h. This result suggested that the *OsMFT2* plays a crucial role in alleviating seed dormancy in rice. Additionally, embryonic bulk RNA-seq was performed on both the *OsMFT2* knockout lines and the WT after seed imbibition, and CIBERSORTx was employed to deconvolute the cell abundance composition based on the

bulk RNA-seq expression data. The results indicated a significantly reduced abundance of SCL2 cells in the *OsMFT2* knockout lines, compared with the WT (Figure 3C). This further substantiated the role of *OsMFT2* as a marker gene for the SCL2 region and suggested its potential influence on the development of the specific area. This example illustrates that leveraging spatial transcriptomics data enabled us to unearth key genes associated with germination from rice embryos and assisted us in pinpointing the critical functional sites of these genes.

The cell area in the BESCL2 and SCL2 regions was slightly larger than that in SCL1-2. BESCL2 was observed at both ends of the SCL2 layer and appeared discontinuous on the section but had a similar transcriptome profile. The cluster with this distribution was observed in multiple embryos, and several specifically expressed genes (e.g., *Os10g0361300* and *Os09g0272600*) were identified in the cell type (Figures 3A and S2C). SCL1-2 was located mainly in between the SC epidermis layer (SCL1) and the parenchyma layer (SCL2) and had a significantly high expression of *OsCYS9* (*Os04g0350100*) and *OsMT3a/2b* (*Os01g0200700* and *Os05g0111300*) (Figure S2C). *OsCYS9* is a member of the cystatin family that has been reported to play an important role in rice seed germination.²⁵ According to previous studies, the expression level of cystatin gene is higher in seeds at early germination stages and decreases dramatically when germination occurs. This expression profile contributes to the increased activity of cysteine protease during seed germination to convert seed storage protein into energy required for germination.²⁶ Here, we also employed ISH to corroborate the distinct expression pattern of *OsCYS9* (*Os04g0350100*) in SCL1-2 (Figure 3B).

To better define the transcriptome differences among the three cell types of SC parenchyma, we further analyzed their potential developmental trajectories using the integrated scRNA-seq and Stereo-seq data (Figures S3A and S3B). Histone-related genes are conspicuously overexpressed in SCL1-2, compared with two other types of parenchyma cells in the SC (Figures 3F and S3C). A set of DEGs and enriched pathways conforming to the trajectory inference were defined (Figures 3G and S3D). We observed that the early stages of development are predominantly characterized by the SCL1-2 cell type, accompanied by the peptide biosynthetic process. During the intermediate states of pseudo-time, pathways associated with aldehyde-lyase activity and seed maturation successively manifest enrichment. Subsequently, pathways related to environmental responses become enriched toward the terminal end of the pseudo-time

Figure 3. Cell types and trajectory inference of scutellum region (24 HAI)

- (A) Cell distribution of the five cell types in the scutellum region and the expression pattern of the key marker gene of the corresponding cell type at 24 HAI.
- (B) *In situ* hybridization detection of *OsSAG12-1* (*Os01g0907600*) and *OsCYS9* (*Os04g0350100*).
- (C) FISH validation result of *OsMFT2* (*Os01g0111600*).
- (D) Germination rate of the *OsMFT2* knockout line and the wild type (WT). Error bars represent standard deviation (SD) from three independent experiments.
- (E) Cell abundance of SCL2 in the *OsMFT2* knockout line and WT based on cell-type deconvolution of their embryonic bulk RNA-seq data. Error bars represent standard deviation (SD) from three independent experiments.
- (F) Trajectory inference analysis of SCL1-2, SCL2, and BESCL2. Cells are colored by cell type (left) or developmental pseudo-time (right).
- (G) Expression heatmap and clustering of the differentially expressed genes (DEGs) along a pseudo-time progression of scutellum parenchyma cells (left); the top DEGs of the cluster showing on the left panel and the enriched Gene Ontology (GO) terms (middle); and scatterplots showing the pseudo-time dynamics of the expression of the selected top DEGs (right).

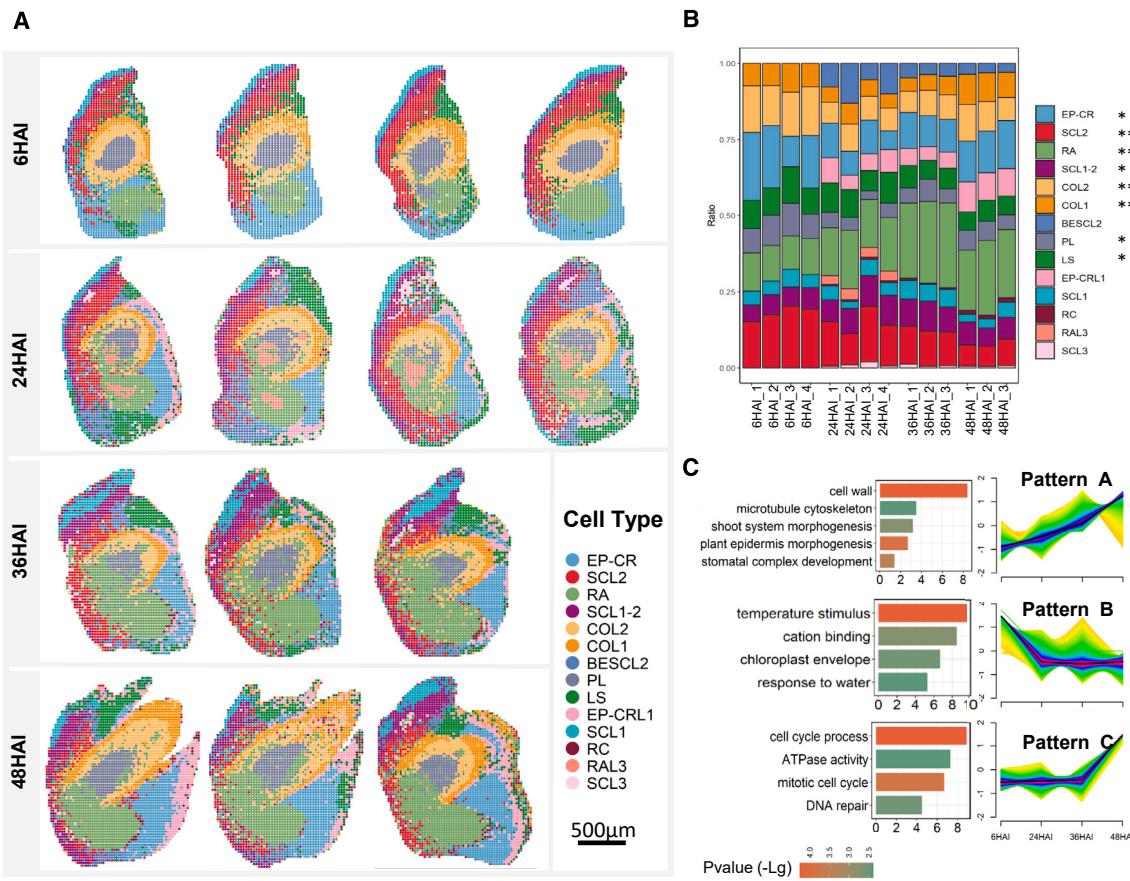


Figure 4. Spatiotemporal transcriptomic landscape of rice embryo during seed germination

(A) Spatial visualization of cell types identified in the rice embryo sections by Stereo-seq at the four germination time points (6, 24, 36, and 48 HAI). Each image of the same time point represents one replicate. Scale bar represent 500 μm.

(B) Barplot showing the cell count ratio of each cell type identified by Stereo-seq at different samples. The same cell type in different samples is represented by the same color. * indicates a significant change ($p < 0.05$) in the proportion of cell types at different germination time points, and ** indicates an extremely significant change ($p < 0.01$). Line plot depicting standardized spatial expression signal by time-series analysis. Three typical expression patterns, patterns A, B and C, in 6 to 48-HAI embryos are shown, with the enriched representative GO terms of the genes from each pattern shown on left.

(C) Representative gene expression patterns of Stereo-seq data by time-series analysis (right) and the GO enriched pathways corresponding to the expression pattern (left).

axis, predominantly constituted by the BESCL2 cell type during the phase.

Dynamics of rice embryonic cells during seed germination

We applied the spatial data of the 14 embryo sections from 6, 24, 36, and 48 HAI to investigate the cellular and molecular dynamics associated with embryo germination (Figure 4A). Unsupervised clustering analysis identified similar cell composition in all analyzed sections (Figure 4A), except for the 6-HAI sections where the BESCL2, SCL3, and EP-CRL1 were indistinguishable, and RAL3 could only be recognized on the 24-HAI sections. Cell ratio between sections from the same time point had no significant difference, implying the stability of the experimental and sequencing techniques. Further, we investigated the discrepancies in cell proportions between different germination time points, and we found significant variation ($p \leq 0.01$) in SCL2, RA, COL1, and COL2 for cell types (Figures 4B, S4A, and

S4B). A decreasing trend in cell proportion was evident for SCL2 and LS from 6 to 48 HAI, while a pronounced increase in the proportion of RA cells was observed from 6 to 48 HAI (Figure 4B).

We further conducted a comprehensive evaluation of the expression patterns of the identified marker genes at the four time points (Figures S4C and S4D). The expression pattern of some marker genes, such as *OsMFT2*, *OsSAG12-1*, *OsCER1*, and *OsPRP14*, also suggested their activation time point during seed germination. For instance, the expression of *OsMFT2* decreased in the SCL2 cells from 6 to 24 HAI and then stabilized (Figure S4D). *OsMFT2* encoding a multi-pass transmembrane protein is involved in the transportation of mannitol, a type of sugar alcohol, from the endosperm to the embryo. This transportation is important for seed germination as mannitol is an energy source for the developing embryo.²⁴ Additionally, *OsMFT2* and its partners GF14h and OREB1 regulate germination rate by controlling ABA-responsive genes.²⁷

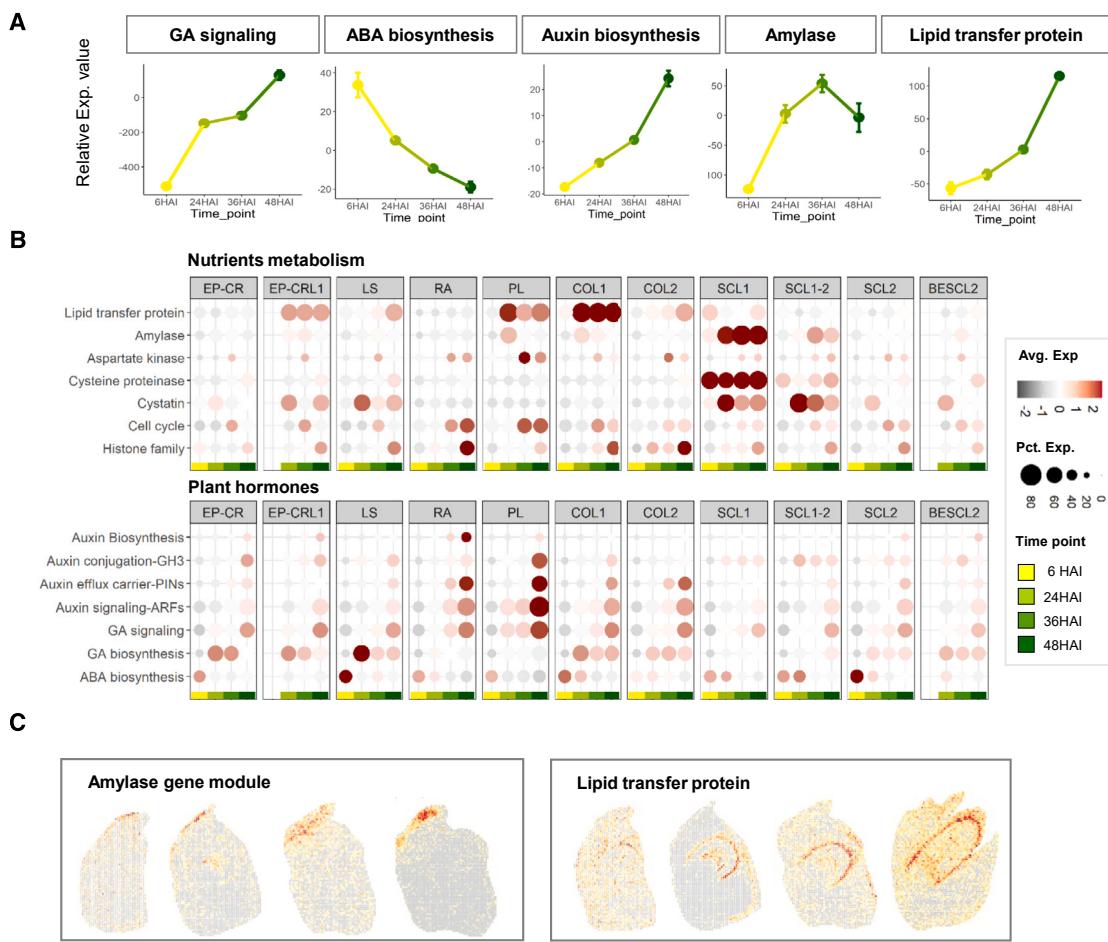


Figure 5. Expression pattern of the genes related to nutrients and plant hormones at the cell-type resolution

(A) Expression trend of the genes related to nutrients and plant hormones in 6- to 48-HAI embryos. The relative expression level was generated based on pseudo-bulk RNA analysis by integrating Stereo-seq and scRNA-seq datasets. Error bars represent standard deviation (SD) from independent samples.

(B) Dotplot showing the expression level of genes encoding proteins involved in nutrient metabolism (top) and biosynthesis and signaling of phytohormones (bottom) in different cell types of germinating rice embryos (6–48 HAI from left to right). The circle size indicates the relative percentage of cells expressing the gene, and the color represents the relative expression level.

(C) The spatial expression pattern of the genes encoding amylase and lipid transfer protein in germinating rice embryos (6–48 HAI, from left to right).

We also assessed the molecular dynamics of rice embryos from 6 to 48 HAI using Mfuzz²⁸ (Figures 4C and S4E). Genes of the pathways related to cell wall composition, microtubule cytoskeleton, stomatal development, shoot system morphogenesis, and epidermis morphogenesis showed a continuously increased expression pattern during imbibition (pattern A). The expression level of the genes enriched for the pathways encompassing responding to temperature and water stimulation, cation binding, and chloroplast envelope featured a significant decrease from 6 to 24 HAI and then remained relatively consistent from 24 to 48 HAI (pattern B). The third expression profile (pattern C) was observed in the genes enriched for cell-cycle process, ATPase activity, and mitotic cell cycle and DNA repair, which exhibited a relatively consistent expression level from 6 to 36 HAI, followed by a sharp spike from 36 to 48 HAI (Figure 4C). These results imply an enhanced cell proliferation or cell-cycle progression after 36 HAI.

Specificity of cell types in regulation of nutrients and hormones

Plant hormones regulate nutrient metabolism during rice seed germination. The above analysis and previous research²⁶ identified several significant physiological events during seed germination, including the hydrolysis of soluble sugars in the endosperm and their transport into the embryo. By integrating the Stereo-seq and scRNA-seq data (so-called pseudo-bulk RNA), we observed a significant increase in the expression level of genes involved in auxin biosynthesis and gibberellin (GA) signaling and related to the transport of lipids, from 6 to 48 HAI. During the same time period, the expression level of ABA biosynthesis-related genes decreased significantly. The transcripts of the genes encoding amylase significantly increased from 6 to 36 HAI and then declined (Figure 5A; for the relevant gene sets, see Tables S4 and S5). These results were in line with those observed in analysis of bulk RNA-seq data of the whole embryo.^{29,30}

Stereo-seq data afforded us an opportunity to investigate the roles of each cell type in the regulation of nutrient metabolism and hormone biosynthesis during seed germination. Genes encoding alpha-amylase and the cystatin family proteins were significantly highly expressed in the SCL1 and SCL1-2 cells (Figure 5B). These types of cells link the endosperm with the embryo and showed higher expression levels of amylase-coding genes on the endosperm side, compared with the RA side (Figure 5C). Alpha-amylase plays a key role in amylase hydrolysis and the transfer of nutrients between endosperm and embryo.³¹ We performed a time-series analysis of cell types within the SC region (Figure S5A). The expression pattern (a continuous upward trend from 6 to 48 HAI) identified not only previously disclosed pathways but also a previously unexplored set of pathways. For example, glucosidase, glucose transmembrane transporter, and hydrolase (GO:0015926, GO:0005355) were activated at 36 HAI, suggesting enhanced starch hydrolysis and nutrient transportation in SCL1. In the SCL1-2, in addition to the polysaccharide metabolic process (GO:0005976), signals associated with embryonic organ and epithelial tube morphogenesis (GO:0060562, GO:0048562) were enriched, suggesting a high differentiation potential of the SCL1-2 cells.

During rice seed germination, hormones such as auxin, cytokinin, GA, and ABA regulate the expression of specific genes to regulate germination and ensure proper growth and development of seedling.³² We examined the expression pattern of the genes regulated by plant hormones in different cell types during seed germination. We found that the genes related to auxin biosynthesis, PINs (auxin efflux carriers), and GH3 (auxin conjugation) were significantly expressed in PL and RA (Figure 5B). Transcription factors (TFs) play a key role in hormone biosynthesis. In the PL and RA cells, several TFs (e.g., NAC, LBD, HSF, GATA, and C3H) showed a significantly gradual decrease while others (e.g., E2F/DP, CAMTA, ARF, and AP2) displayed a reverse trend of expression throughout the germination course (Figure S5B). Meanwhile, the LS and EP-CR cells exhibited an elevated expression level of the genes involved in GA signaling. Interestingly, genes related to ABA biosynthesis were notably expressed in the SCL2 cells with a gradual downward trend from 6 to 48 HAI.

DISCUSSION

By integrating Stereo-seq and scRNA-seq datasets generated from four time points of germinating rice embryos, we built a spatial and temporal transcriptomic atlas of germinating rice embryos at the single-cell resolution and obtained an in-depth understanding of the subcellular transcriptomes of rice embryos during the germination process. The single-cell spatiotemporal transcriptomic data allowed us to successfully distinguish the detailed cell composition (a total of 14 cell types) of the 5 anatomical regions (PL, RA, SC, CO, and EP-CR) of rice embryo and to identify the noteworthy and subtle differences in transcriptomic profiles and spatial location of each cell type. Particularly, a granular examination of the cells in the SC region led to refined SC cell types and their functional characterization. Our results provided insights into the gene expression patterns of individual cell types involved in various cellular and molecular processes, including biosynthesis and signaling of phytohormones, starch hydrolysis,

and nutrient transport between different tissues. The spatiotemporal transcriptome profile of germinating rice embryos laid the foundation for deeper exploration of the functions of different tissues and cell types in seed germination in future.

No studies on the tissue-specific expression of *OsMFT2* in seeds have been reported yet. In this study, we discovered that *OsMFT2* is one of the marker genes for the SCL2 cell type by innovatively employing Stereo-seq data (Figure 3A). The MOTHER OF FT AND TFL1 (MFT) gene family has been found to be closely associated with seed dormancy and germination in rice and several other crops. *OsMFT2* is implicated in the modulation of ABA signaling-mediated seed germination in rice, interacting with three bZIP TFs (*OsbZIP23*, *OsbZIP66*, and *OsbZIP72*) and enhancing their binding activities to the promoter of ABA-responsive genes. This interaction modulates ABA sensitivity during seed germination and post-germination growth.²⁴ Ectopic overexpression of a soybean MFT homolog (*GmMFT*) and two cotton homologs (*GhMFT1* and *GhMFT1*) in *Arabidopsis* inhibited seed germination, all through participation in GA and ABA signaling.^{33,34} Recent reports suggested that the expression of *MFT2* genes could be regulated by a network of TFs through multiple seed-specific *cis*-acting RY motifs in the *MFT2* promoters, indicating a coordinated regulatory mechanism during seed germination.³⁵ The specific role of *OsMFT2*'s expression in SCL2 remains to be further explored.

The mechanisms of seed dormancy and germination have long been a focus of research into crop genetics and breeding. Our study enhances the understanding of genes related to dormancy and germination by providing high-resolution temporal and cell-type-specific expression details. For instance, during the germination of rice seeds, *ABI3* (*Os08g0101000*) plays a crucial role in modulating the ABA signaling pathway, integral for maintaining seed dormancy and regulating germination.³⁶ Our results indicate that during germination, *ABI3* is predominantly expressed in the embryonic region, with an upward expression generally (Figures S5C and S5D). *ABI3* is recognized as a central element in the signaling network of rice embryos during seed germination, seen to collaborate with other factors, orchestrating the regulation of gene expression pivotal for seed germination.³⁷ The gene *SDR4* (*Os07g0585700*) is instrumental in regulating seed dormancy in rice.³⁸ Our observations identified that *SDR4* was specifically expressed in the SC region and the embryonic axis of the embryo, with its expression peak at 6 HAI (Figures S4C and S4D). Our studies contribute to an understanding of the cell-type specificity of genes and regulatory pathways involved in rice germination and dormancy.

The plant cell walls pose a challenge for conducting single-cell transcriptome studies.^{39–41} However, in this study, we leveraged the presence of the cell walls to our advantage by utilizing them in single-cell segmentation based on cell wall-stained high-definition images of rice embryo sections. Integrating this approach with Stereo-seq and scRNA-seq allowed us to generate a comprehensive and accurate *in situ* spatial transcriptional profile at a real single-cell level, overcoming the technical limitations of distinguishing cell types in previous single-cell technologies. We developed an automated cell-segmentation protocol based on deep learning, which is applicable in other plant species, thus offering a valuable tool for future spatial transcriptome analysis in plants. But more accurate cell segmentation needs to address

the limiting factors such as the number of genes detected in each cell and diffusion of intracellular RNAs during permeabilization. With the advance of spatial transcriptome sequencing technologies, we anticipate a significant increase in the number of sequencing reads or unique molecular identifiers (UMIs) from individual cells and more accurate vertical RNA permeabilization in future, which will significantly improve the power of automated cell segmentation, leading to high-quality preprocessing of spatial transcriptome data. We believe that automated cell segmentation will become a common and fundamental protocol in analysis of plant spatial transcriptome data, particularly in the integration of spatial transcriptome and scRNA-seq data.

scRNA-seq offers a higher throughput in detecting the genes in individual cells, as compared with spatial transcriptome technique. Despite the potential benefits, integration of these two types of data has proven to be challenging.⁴² Based on this study, there are at least a couple of limiting factors for the integrative analysis. First, protoplasting-induced highly expressed genes (ppDEGs) were enriched in scRNA-seq-based clusters (Figure S3A). The protoplasting process likely impacted transcript expression in germinating rice embryos, leading to a pronounced technical effect of scRNA-seq due to the 3-h digestion of plant cell walls with cellulases and pectinases. Second, the bias of sequencing cell space or population by the two methods allows for only a low percentage (about 10% based on our data) of the cells from scRNA-seq to be integrated into the clearly annotated Stereo-seq-based cell clusters.

In conclusion, we established a spatiotemporal transcriptomic atlas of germinating rice embryos with the help of the Stereo-seq technique and the automated cell-segmentation model built based on deep learning. Our study identified intricate cell-type-level transcriptional dynamics during seed germination and identified two previously unreported cell types in the SC regions and their marker genes. The spatiotemporal transcript resource at single-cell resolution generated by this study will be valuable for unraveling the molecular basis underlying seed development and germination and provide molecular solution for enhancing seed vigor.

Limitations of the study

The rice seed embryo samples utilized in this study commenced from 6 HAI, omitting the use of dry seeds initiated at 0 HAI. Our extensive experimentation identified the excessive brittleness of dry seeds at 0 HAI, rendering them prone to fragmentation and precluding the acquisition of intact embryonic structures. Fortunately, the samples spanning from 6 to 48 HAI has adequately encompassed the most crucial physiological activities during rice embryo germination. Furthermore, the enzymatic digestion of the cell wall unavoidably influences the cellular transcriptome during scRNA-seq. Despite implementing measures to mitigate and minimize the protoplasting effect, this inadvertently resulted in the loss of numerous cells. Substituting single-nucleus RNA-seq and integrating it with spatial transcriptomics might offer a solution to circumvent this issue.

STAR METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
 - Lead contact
 - Materials availability
 - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
 - Plant material and germination condition
- METHOD DETAILS
 - Generation of transgenic plants
 - Tissue cryosection
 - Stereo-seq tissue processing and imaging
 - Stereo-seq library construction and sequencing
 - Rice embryo protoplast preparation
 - Single-cell library preparation and sequencing
 - In situ hybridization
- QUANTIFICATION AND STATISTICAL ANALYSIS
 - Cell segmentation of rice embryos
 - scRNA-seq and Stereo-seq data analysis
 - Gene module analysis
 - Time series expression pattern analysis
 - GO enrichment analysis
 - Cell type deconvolution

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.devcel.2024.05.016>.

ACKNOWLEDGMENTS

This study was supported by National Key R&D Program of China (2023ZD04073), Hainan Yazhou Bay Seed Laboratory (B23YQ2002), Guangdong Laboratory of Lingnan Morden Agriculture (AGIS-ZDXM202203), and China National GeneBank (CNGB; <https://www.cngb.org/>).

AUTHOR CONTRIBUTIONS

Conceptualization, L.F. and H.L.; methodology, J.Y., X.G., W.S., K.L., F.O., and Q.C.; formal analysis, J.Y. and N.S.; investigation, J.Y., X.L., H.C., and D.Z.; writing – original draft, J.Y.; writing – review & editing, L.F., Q.-H.Z., and Q.C.; resources, F.X. and F.G.; supervision, L.F., H.L., S.D., C.C., and X.X.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: April 24, 2023

Revised: February 15, 2024

Accepted: May 14, 2024

Published: June 6, 2024

REFERENCES

1. Guo, F., Han, N., Xie, Y., Fang, K., Yang, Y., Zhu, M., Wang, J., and Bian, H. (2016). The miR393a/target module regulates seed germination and seedling establishment under submergence in rice (*Oryza sativa* L.). *Plant Cell Environ.* 39, 2288–2302. <https://doi.org/10.1111/pce.12781>.
2. Jiao, Y., Tausta, S.L., Gandotra, N., Sun, N., Liu, T., Clay, N.K., Ceserani, T., Chen, M., Ma, L., Holford, M., et al. (2009). A transcriptome atlas of rice cell types uncovers cellular, functional and developmental hierarchies. *Nat. Genet.* 41, 258–263. <https://doi.org/10.1038/ng.282>.
3. Zeng, M., Hu, B., Li, J., Zhang, G., Ruan, Y., Huang, H., Wang, H., and Xu, L. (2016). Stem cell lineage in body layer specialization and vascular patterning of rice root and leaf. *Sci. Bull.* 61, 847–858. <https://doi.org/10.1007/s11434-015-0849-1>.
4. Damaris, R.N., Lin, Z., Yang, P., and He, D. (2019). The Rice Alpha-Amylase, Conserved Regulator of Seed Maturation and Germination. *Int. J. Mol. Sci.* 20, 450–467. <https://doi.org/10.3390/ijms20020450>.

5. Howell, K.A., Millar, A.H., and Whelan, J. (2006). Ordered assembly of mitochondria during rice germination begins with pro-mitochondrial structures rich in components of the protein import apparatus. *Plant Mol. Biol.* 60, 201–223. <https://doi.org/10.1007/s11103-005-3688-7>.
6. Liu, Q., Liang, Z., Feng, D., Jiang, S., Wang, Y., Du, Z., Li, R., Hu, G., Zhang, P., Ma, Y., et al. (2021). Transcriptional landscape of rice roots at the single-cell resolution. *Mol. Plant* 14, 384–394. <https://doi.org/10.1016/j.molp.2020.12.014>.
7. Zhang, T.-Q., Chen, Y., Liu, Y., Lin, W.-H., and Wang, J.-W. (2021). Single-cell transcriptome atlas and chromatin accessibility landscape reveal differentiation trajectories in the rice root. *Nat. Commun.* 12, 2053. <https://doi.org/10.1038/s41467-021-22352-4>.
8. Wang, Y., Huan, Q., Li, K., and Qian, W. (2021). Single-cell transcriptome atlas of the leaf and root of rice seedlings. *J. Genet. Genomics* 48, 881–898. <https://doi.org/10.1016/j.jgg.2021.06.001>.
9. Bai, Y., Liu, H., Lyu, H., Su, L., Xiong, J., and Cheng, Z.-M.M. (2022). Development of a single-cell atlas for woodland strawberry (*Fragaria vesca*) leaves during early *Botrytis cinerea* infection using single cell RNA-seq. *Hortic. Res.* 9, uhab055. <https://doi.org/10.1093/hr/uhab055>.
10. Kim, J.-Y., Symeonidi, E., Pang, T.Y., Denyer, T., Weidauer, D., Bezrutczyk, M., Miras, M., Zöllner, N., Hartwig, T., Wudick, M.M., et al. (2021). Distinct identities of leaf phloem cells revealed by single cell transcriptomics. *Plant Cell* 33, 511–530. <https://doi.org/10.1093/plcell/koaa060>.
11. Xia, K., Sun, H.-X., Li, J., Li, J., Zhao, Y., Chen, L., Qin, C., Chen, R., Chen, Z., Liu, G., et al. (2022). The single-cell stereo-seq reveals region-specific cell subtypes and transcriptome profiling in *Arabidopsis* leaves. *Dev. Cell* 57, 1299–1310.e4. <https://doi.org/10.1016/j.devcel.2022.04.011>.
12. Liao, J., Lu, X., Shao, X., Zhu, L., and Fan, X. (2021). Uncovering an Organ's Molecular Architecture at Single-Cell Resolution by Spatially Resolved Transcriptomics. *Trends Biotechnol.* 39, 43–58. <https://doi.org/10.1016/j.tibtech.2020.05.006>.
13. Liu, C., Li, R., Li, Y., Lin, X., Zhao, K., Liu, Q., Wang, S., Yang, X., Shi, X., Ma, Y., et al. (2022). Spatiotemporal mapping of gene expression landscapes and developmental trajectories during zebrafish embryogenesis. *Dev. Cell* 57, 1284–1298.e5. <https://doi.org/10.1016/j.devcel.2022.04.009>.
14. Chen, A., Liao, S., Cheng, M., Ma, K., Wu, L., Lai, Y., Qiu, X., Yang, J., Xu, J., Hao, S., et al. (2022). Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell* 185, 1777–1792.e21. <https://doi.org/10.1016/j.cell.2022.04.003>.
15. Pachitariu, M., and Stringer, C. (2022). Cellpose 2.0: how to train your own model. *Nat. Methods* 19, 1634–1641. <https://doi.org/10.1038/s41592-022-01663-4>.
16. Lu, Y., Ye, X., Guo, R., Huang, J., Wang, W., Tang, J., Tan, L., Zhu, J.-K., Chu, C., and Qian, Y. (2017). Genome-wide Targeted Mutagenesis in Rice Using the CRISPR/Cas9 System. *Mol. Plant* 10, 1242–1245. <https://doi.org/10.1016/j.molp.2017.06.007>.
17. Liew, L.C., Narsai, R., Wang, Y., Berkowitz, O., Whelan, J., and Lewsey, M.G. (2020). Temporal tissue-specific regulation of transcriptomes during barley (*Hordeum vulgare*) seed germination. *Plant J.* 101, 700–715. <https://doi.org/10.1111/tpj.14574>.
18. Dice, L.R. (1945). Measures of the Amount of Ecologic Association Between Species. *Ecology* 26, 297–302. <https://doi.org/10.2307/1932409>.
19. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. *Cell* 184, 3573–3587.e29. <https://doi.org/10.1016/j.cell.2021.04.048>.
20. Li, R., Xia, J., Xu, Y., Zhao, X., Liu, Y.-G., and Chen, Y. (2014). Characterization and genetic mapping of a Photoperiod-sensitive dwarf 1 locus in rice (*Oryza sativa* L.). *Theor. Appl. Genet.* 127, 241–250. <https://doi.org/10.1007/s00122-013-2213-7>.
21. Sakai, H., Lee, S.S., Tanaka, T., Numa, H., Kim, J., Kawahara, Y., Wakimoto, H., Yang, C.C., Iwamoto, M., Abe, T., et al. (2013). Rice Annotation Project Database (RAP-DB): an integrative and interactive database for rice genomics. *Plant Cell Physiol.* 54, e6. <https://doi.org/10.1093/pcp/pcs183>.
22. Tsuda, K., Akiba, T., Kimura, F., Ishibashi, M., Moriya, C., Nakagawa, K., Kurata, N., and Ito, Y. (2013). ONION2 fatty acid elongase is required for shoot development in rice. *Plant Cell Physiol.* 54, 209–217. <https://doi.org/10.1093/pcp/pcs169>.
23. Ni, E., Zhou, L., Li, J., Jiang, D., Wang, Z., Zheng, S., Qi, H., Zhou, Y., Wang, C., Xiao, S., et al. (2018). OsCER1 Plays a Pivotal Role in Very-Long-Chain Alkane Biosynthesis and Affects Plastid Development and Programmed Cell Death of Tapetum in Rice (*Oryza sativa* L.). *Front. Plant Sci.* 9, 1217. <https://doi.org/10.3389/fpls.2018.01217>.
24. Song, S., Wang, G., Wu, H., Fan, X., Liang, L., Zhao, H., Li, S., Hu, Y., Liu, H., Ayaad, M., et al. (2020). OsMFT2 is involved in the regulation of ABA signaling-mediated seed germination through interacting with OsbZIP23/66/72 in rice. *Plant J.* 103, 532–546. <https://doi.org/10.1111/tpj.14748>.
25. Wang, W., Zhao, P., Zhou, X.-M., Xiong, H.-X., and Sun, M.-X. (2015). Genome-wide identification and characterization of cystatin family genes in rice (*Oryza sativa* L.). *Plant Cell Rep.* 34, 1579–1592. <https://doi.org/10.1007/s00299-015-1810-0>.
26. Hwang, J.E., Hong, J.K., Je, J.H., Lee, K.O., Kim, D.Y., Lee, S.Y., and Lim, C.O. (2009). Regulation of seed germination and seedling growth by an *Arabidopsis* phytocystatin isoform, AtCYS6. *Plant Cell Rep.* 28, 1623–1632. <https://doi.org/10.1007/s00299-009-0762-7>.
27. Yoshida, H., Hirano, K., Yano, K., Wang, F., Mori, M., Kawamura, M., Koketsu, E., Hattori, M., Ordonio, R.L., Huang, P., et al. (2022). Genome-wide association study identifies a gene responsible for temperature-dependent rice germination. *Nat. Commun.* 13, 5665. <https://doi.org/10.1038/s41467-022-33318-5>.
28. Kumar, L., and E Futschik, M. (2007). Mfuzz: a software package for soft clustering of microarray data. *Bioinformation* 2, 5–7. <https://doi.org/10.6026/97320630002005>.
29. Gong, D., He, F., Liu, J., Zhang, C., Wang, Y., Tian, S., Sun, C., and Zhang, X. (2022). Understanding of Hormonal Regulation in Rice Seed Germination. *Life (Basel)* 12, 1021–1040. <https://doi.org/10.3390/life12071021>.
30. Sano, N., Lounifi, I., Cueff, G., Collet, B., Clément, G., Balzergue, S., Huguet, S., Valot, B., Galland, M., and Rajjou, L. (2022). Multi-Omics Approaches Unravel Specific Features of Embryo and Endosperm in Rice Seed Germination. *Front. Plant Sci.* 13, 867263. <https://doi.org/10.3389/fpls.2022.867263>.
31. Szewińska, J., Simińska, J., and Bielawski, W. (2016). The roles of cysteine proteases and phytocystatins in development and germination of cereal seeds. *J. Plant Physiol.* 207, 10–21. <https://doi.org/10.1016/j.jplph.2016.09.008>.
32. Ali, F., Qanmber, G., Li, F., and Wang, Z. (2022). Updated role of ABA in seed maturation, dormancy, and germination. *J. Adv. Res.* 35, 199–214. <https://doi.org/10.1016/j.jare.2021.03.011>.
33. Li, W.-Y., Chen, B.-X., Chen, Z.-J., Gao, Y.-T., Chen, Z., and Liu, J. (2017). Reactive Oxygen Species Generated by NADPH Oxidases Promote Radicle Protrusion and Root Elongation during Rice Seed Germination. *Int. J. Mol. Sci.* 18, 110. <https://doi.org/10.3390/ijms18010110>.
34. Yu, X., Liu, H., Sang, N., Li, Y., Zhang, T., Sun, J., and Huang, X. (2019). Identification of cotton MOTHER OF FT AND TFL1 homologs, GhMFT1 and GhMFT2, involved in seed germination. *PLoS One* 14, e0215771. <https://doi.org/10.1371/journal.pone.0215771>.
35. Utsugi, S., Kawahigashi, H., Tagiri, A., Kikuchi, R., Mishina, K., Morishige, H., and Nakamura, S. (2023). Conserved cis-acting motifs and localization of transcripts and proteins of MFT2 in barley and rice. Preprint at bioRxiv.
36. Chen, H.-C., Cheng, W.-H., Hong, C.-Y., Chang, Y.-S., and Chang, M.-C. (2018). The transcription factor OsbHLH035 mediates seed germination and enables seedling recovery from salt stress through ABA-dependent

- and ABA-independent pathways, respectively. *Rice (N Y)* 11, 50. <https://doi.org/10.1186/s12284-018-0244-z>.
37. Park, J., Lee, N., Kim, W., Lim, S., and Choi, G. (2011). ABI3 and PILS collaboratively activate the expression of SOMNUS by directly binding to its promoter in imbibed *Arabidopsis* seeds. *Plant Cell* 23, 1404–1415. <https://doi.org/10.1105/tpc.110.080721>.
38. Zhao, B., Zhang, H., Chen, T., Ding, L., Zhang, L., Ding, X., Zhang, J., Qian, Q., and Xiang, Y. (2022). Sdr4 dominates pre-harvest sprouting and facilitates adaptation to local climatic condition in Asian cultivated rice. *J. Integr. Plant Biol.* 64, 1246–1263. <https://doi.org/10.1111/jipb.13266>.
39. Iqbal, M.M., Hurgobin, B., Holme, A.L., Appels, R., and Kaur, P. (2020). Status and Potential of Single-Cell Transcriptomics for Understanding Plant Development and Functional Biology. *Cytometry A* 97, 997–1006. <https://doi.org/10.1002/cyto.a.24196>.
40. Rich-Griffin, C., Stechemesser, A., Finch, J., Lucas, E., Ott, S., and Schäfer, P. (2020). Single-Cell Transcriptomics: A High-Resolution Avenue for Plant Functional Genomics. *Trends Plant Sci.* 25, 186–197. <https://doi.org/10.1016/j.tplants.2019.10.008>.
41. Seyfferth, C., Renema, J., Wendrich, J.R., Eekhout, T., Seurinck, R., Vandamme, N., Blob, B., Saeyns, Y., Helariutta, Y., Birnbaum, K.D., et al. (2021). Advances and Opportunities in Single-Cell Transcriptomics for Plant Research. *Annu. Rev. Plant Biol.* 72, 847–866. <https://doi.org/10.1146/annurev-applant-081720-010120>.
42. Longo, S.K., Guo, M.G., Ji, A.L., and Khavari, P.A. (2021). Integrating single-cell and spatial transcriptomics to elucidate intercellular tissue dynamics. *Nat. Rev. Genet.* 22, 627–644. <https://doi.org/10.1038/s41576-021-00370-8>.
43. Qiu, X., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H.A., and Trapnell, C. (2017). Reversed graph embedding resolves complex single-cell trajectories. *Nat. Methods* 14, 979–982. <https://doi.org/10.1038/nmeth.4402>.
44. Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D.M., Hill, A.J., Zhang, F., Mundlos, S., Christiansen, L., Steemers, F.J., et al. (2019). The single-cell transcriptional landscape of mammalian organogenesis. *Nature* 566, 496–502. <https://doi.org/10.1038/s41586-019-0969-x>.
45. Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P.-R., and Raychaudhuri, S. (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* 16, 1289–1296. <https://doi.org/10.1038/s41592-019-0619-0>.
46. Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16, 284–287. <https://doi.org/10.1089/omi.2011.0118>.
47. Ma, X., Zhu, Q., Chen, Y., and Liu, Y.-G. (2016). CRISPR/Cas9 Platforms for Genome Editing in Plants: Developments and Applications. *Mol. Plant* 9, 961–974. <https://doi.org/10.1016/j.molp.2016.04.009>.
48. Zhang, Y., Su, J., Duan, S., Ao, Y., Dai, J., Liu, J., Wang, P., Li, Y., Liu, B., Feng, D., et al. (2011). A highly efficient rice green tissue protoplast system for transient gene expression and studying light/chloroplast-related processes. *Plant Methods* 7, 30. <https://doi.org/10.1186/1746-4811-7-30>.
49. Sun, Y., Han, Y., Sheng, K., Yang, P., Cao, Y., Li, H., Zhu, Q.-H., Chen, J., Zhu, S., and Zhao, T. (2023). Single-cell transcriptomic analysis reveals the developmental trajectory and transcriptional regulatory networks of pigment glands in *Gossypium bickii*. *Mol. Plant* 23, 694–708. <https://doi.org/10.1016/j.molp.2023.02.005>.
50. Newman, A.M., Steen, C.B., Liu, C.L., Gentles, A.J., Chaudhuri, A.A., Scherer, F., Khodadoust, M.S., Esfahani, M.S., Luca, B.A., Steiner, D., et al. (2019). Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat. Biotechnol.* 37, 773–782. <https://doi.org/10.1038/s41587-019-0114-2>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Seeds of <i>Oryza sativa</i>	This study	N/A
Chemicals, peptides, and recombinant proteins		
OCT	Leica	4583
FluorescentBrightener 28	SIGMA	F3542-5G
Pepsin	Sigma	P7000
HCl	Dongjiang	N/A
0.1x SSC buffer	Thermo	AM9770
RNase inhibitor	NEB	M0314L
SuperScript II	Invitrogen	18064-014
Tissue Removal buffer	BGI	N/A
Exonuclease I	NEB	M0293L
KAPA HiFi Hotstart Ready Mix	Roche	KK2602
Qubit dsDNA Assay Kit	Thermo	Q32854
SDS buffer	BBI	A600485-0500
1x KAPA HiFi Hotstart Ready Mix	BGI	N/A
Stereo-Library-F primer	BGI	N/A
Stereo-Library-R primer	BGI	N/A
Nuclease-free H2O	Ambion	AM9937
Ampure XP Beads	Vazyme	N411-03
Ethanol	XILONGSCIENTIFIC	171225-02
Deposited data		
Raw scRNA-seq data	This study	CNGBdb: CNP0004152
Stereo-seq data: raw matrix, filtered feature-spot matrix, FB stained images, and cell annotation information	This study	STOmicsDB: STT0000049
Custom code	This study	https://github.com/YAOJ-bioin/REGSTA-script ; Zenodo: https://doi.org/10.5281/zenodo.1111577
REGSTA	This study	http://ibi.zju.edu.cn/REGSTA/
<i>Oryza sativa</i> Japonica Group annotation	IRGSP-1.0	https://plants.ensembl.org/Oryza_sativa/Info/Index
Oligonucleotides		
FISH primers		
OsMFT2:5'-CY3-CATGTGTACATGCGCC CAGGGTGGCCATCCATCCATATGCAT GCC-3'	This Study	N/A
OsHyPRP14: 5'-CY3-CTAGCCATCAT CGAATGAAATAGATAAGCTTAATTACG AATATTG-3'	This Study	N/A
RNA in situ hybridization primers		N/A
Os02g0587000:5'-DIG-CAAGGTGGAC GTGCAGGACTACTGGCGCG-3'	This Study	N/A
OsLTP1.2:5'-DIG-CTTGGGCGAT GGTGGAGCTATATTGTGCA-3'	This Study	N/A

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
OsSAG12.1:5'-DIG-CGCCTCGTCGGA CTCCAAATCCCTCTCGT-3'	This Study	N/A
OsCYS9:5'-DIG-GCCTTGTTGCTGCC ACGTCCCTGATC-3'	This Study	N/A
Software and algorithms		
Seurat v4.2.0	Hao et al. ¹⁹	https://github.com/satijalab/seurat
Mfuzz v2.52.0	Kumar and E Futschik ²⁸	https://bioconductor.org/packages/release/bioc/html/Mfuzz.html
Monocle v2.20.0	Qiu et al. ⁴³	https://bioconductor.org/packages/release/bioc/html/monocle.html
Monocle3 v0.2.3.0	Cao et al. ⁴⁴	https://cole-trapnell-lab.github.io/monocle3/
Harmony v0.1.0	Korsunsky et al. ⁴⁵	https://github.com/immunogenomics/harmony
clusterProfiler v4.0.5	Yu et al. ⁴⁶	https://bioconductor.org/packages/release/bioc/html/clusterProfiler.html
R	R Core	https://www.r-project.org/
Cellpose v2.0	Pachitariu and Stringer ¹⁵	https://www.cellpose.org/

RESOURCE AVAILABILITY**Lead contact**

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Longjiang Fan (fanlj@zju.edu.cn).

Materials availability

Rice mutant lines generated in this study may be requested from the [lead contact](#).

Data and code availability

All scRNA-seq raw data have been deposited to CNGBdb: CNP0004152; the Stereo-seq data, including the lassoed raw matrix, filtered feature-spot matrix, FB stained images, and cell annotation information, have been deposited into the Spatial Transcript Omics DataBase, STOmics DB: STT0000049. The custom code supporting the current study is available at <https://github.com/YAOJ-bioin/REGSTA-script>, and it can also be obtained via the Zenodo: <https://doi.org/10.5281/zenodo.11115774>. Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS**Plant material and germination condition**

Rice (*Oryza sativa*) Nipponbare embryos were harvested and utilized for Stereo-seq, scRNA-seq, and *in situ* detection procedures. The rice seeds were subjected to imbibition at 6, 24, 36, and 48 HAI, maintained in water at a constant temperature of 30°C, shielded from light, ensuring the water level reached two-thirds of the seed's height.

METHOD DETAILS**Generation of transgenic plants**

Zhonghua 11 (ZH11), a *japonica* rice variety was used for genetic transformation. The CRISPR/Cas9/sgRNA vector for OsMFT2 was constructed as previously described with some modifications.⁴⁷ The forward primer of MFT2U6a is GCCGctgttctattgtttggacc, and reverse primer of MFT2U6a is AAACggtc当地aaataagaacag. Briefly, the primers for OsMFT2 were mixed to 1 μM, heated to 90°C for 30 seconds, and annealed to room temperature, and then inserted into a U6a sgRNA intermediate plasmid at the Bsal site. The gRNA expression cassettes were obtained after two rounds of PCR using the linkage products as templates, where U-F/gRNA-R primers were used first, followed by B1'/B2 and B2'/BL primers. Subsequently, the products were inserted into pYLCRISPR/Cas9-MH binary vectors at Bsal sites to obtain the complete CRISPR/Cas9/sgRNA vectors.⁴⁷

Tissue cryosection

The water-imbibed embryos were manually separated from the germinating seeds and then embedded in pre-cooled OCT (optimal cutting temperature compound) and stored at a temperature of -80°C until ready for processing. Cryosections were cut at a thickness of 10 μm in a Leika CM1950 cryostat.

Stereo-seq tissue processing and imaging

The Stereo-seq capture chips were thoroughly washed with RNase inhibitor-supplemented NF-H₂O and dried at room temperature. Afterwards, tissue sections were affixed onto the chip surface and incubated at 37°C for 4 minutes. Following the incubation, the tissue sections were subjected to fixation in methanol at -20°C for 40 minutes. After methanol fixation, the tissue sections were stained for 3 minutes in a fluorescent staining solution, composed of 0.1× SSC (Saline-sodium citrate), 1/200 nucleic acid dye (Thermo Fisher Q10212) and 2 U/ μl RNase inhibitor. Following removal of the staining solution, the chip was washed with a 0.1× SSC (saline-sodium citrate) buffer supplemented with RNase inhibitor-. Subsequently, the chip was imaged using a Motic Custom PA53 FS6 microscope and the fluorescence channel (FITC, 10× objective) was captured.

Stereo-seq library construction and sequencing

Before Stereo-seq library construction, the tissue sections mounted on the chip were permeabilized using 0.1% pepsin and the released RNAs were captured by DNB and used in reverse transcription and cDNA amplification.¹⁴ The PCR product was quantified using the Qubit dsDNA Assay Kit from Thermo (product code Q32854). A total of 20 ng of DNA was fragmented through the application of Tn5 transposase at a temperature of 55°C for 10 minutes. The reaction was then stopped with the addition of 0.02% SDS (sodium dodecyl sulfate), followed by gentle mixing at room temperature for 5 minutes. Subsequently, the fragmented products were amplified using a KAPA HiFi Hotstart Ready Mix (Roche, KK2602) in a total volume of 100 μl , with the addition of a set of 0.3 μM primer, i.e., Stereo-seq-Library-F (/5phos/CTGCTGACGTACTGAGAGG*C'A-3) and Stereo-seq-Library-R (5-GAGACGTTCTC GACTCAGCAGA-3). The resulting PCR products were purified using VAHTSTM DNA Clean Beads (0.6× and 0.15×) and utilized for the generation of DNBS. The Stereo-seq library was sequenced using a MGI DNBSEQ-Tx sequencer with a read length of 35 bp for read 1 and 100 bp for read 2.

Rice embryo protoplast preparation

Embryos were softly manually separated from seeds at the specific germination time points (6, 24, 36, 48 HAI). The embryos were placed in 0.8 mol/L mannitol solution and cut into small pieces. Clean rice embryo pieces were incubated in a digestion solution containing 1.5% cellulase "Onozuka RS", 0.5% macerase, 0.5 M mannitol, and 4 mM 2-(N-morpholino) ethanesulfonic acid, in a 20-ml Petri dish in a 28°C incubator (with gentle shaking at 80 rpm) for 180min. The solution was filtered through a 40- μm nylon mesh and centrifuged at 130g for 3 min. The supernatant was transferred to a new centrifugal tube (tube #1) and the precipitate was resuspended in W5⁴⁸ in the original tube (tube #2). NaCl solution (0.4 mol/L) was added to tube #1 to a final volume of 10 ml. Tube#1 was then centrifuged at 130g for 3min and the precipitate was resuspended in 500 μl 0.4 mol/L NaCl solution. In tube#2, an equal volume of 70% Ficoll solution was slowly topped up to the solution and the tube was then centrifuged with 80g in 3mins. After centrifugation, the cloudy material (protoplasts) near the liquid level was collected to a new centrifugal tube (tube #3). Tube #3 was subjected to the same process as aforementioned for tube #1. The protoplasts collected in tube #1 and tube #3 were mixed and centrifuged with 80g in 3mins. Following discarding the supernatant, the precipitate was resuspended in W5.

Single-cell library preparation and sequencing

The microbeads-captured single-cell library construction was performed with BD Rhapsody™ Single-Cell Analysis System by following the manufacturer's guidelines. The required volume of cell suspension was calculated, and the up-sample premix of the cells was configured for single cell separation and single cell mRNA was captured through the microplate chip. The magnetic beads that captured the single-cell mRNA were collected to a centrifugal tube. The captured mRNAs were reverse transcribed in a PCR instrument to achieve cDNA labeling. Briefly, double-stranded cDNA was first generated from the microbead-captured single-cell transcriptome in several steps, including reverse transcription, second-strand synthesis, end preparation, adapter ligation, and whole-transcriptome amplification. Then, the final cDNA library was generated from double-stranded full-length cDNA by random priming amplification using a BD Rhapsody cDNA Kit (BD Biosciences, 633773) and the BD Rhapsody Targeted mRNA and AbSeq Amplification Kit (BD Biosciences, 633774). Sequencing was done on the Illumina Novaseq platform with the PE150 mode. Single-cell sequencing was performed at the laboratory of Sinotech Genomics company (Shanghai, China) without any other information given.

In situ hybridization

The probe sequences for all genes were listed in key resources table.⁴⁹ In situ hybridization was performed by Wuhan Servicebio Technology Co., Ltd. Rice embryos (24HAI) were separated from seeds by hand and softened with a plant softening solution at 55°C. The softened embryos were rinsed with running water for 30 minutes, soaked in 15% ethanol for 2 hours, and dehydrated by treatment with a series of increasing concentration of ethanol (30%, 50%, 75%, 85%, 90%, 95% and 100%). Finally, the embryos were cleared with xylene, infiltrated and then entrapped in melted wax. The wax blocks were cut into sections with a thickness of about 5 μm . After adding pre-hybridization solution, the sections were incubated for 1 h at 37°C. Following discarding the

pre-hybridization solution, the sections were soaked in the probe hybridization solution and hybridized in a humidity chamber overnight at 37°C. In the next morning, the hybridization solution was removed, and the sections were washed with 2×SSC for 5 min. Cell nucleus staining was done by incubating the sections in DAPI for 8 min in the dark, and the stained sections were then mounted for observation. FISH was performed by Shanghai Gefan Biotechnology Co., Ltd, following the steps below. After fixing and dewaxing tissue sections, they were processed into the standard working solution in a humid chamber. A mixture of formamide and 5×SSC was applied to the sections, followed by a 10-minute treatment and washing with DEPC (Diethylpyrocarbonate) water. Subsequently, hydrochloric acid was added to the sections to neutralize charged proteins, and then washed with DEPC water. The tissue sections were treated with proteinase K at 37°C for 20 minutes and the reaction was stopped. Proteinase K was terminated by washing with glycine, followed by PBS washing. Tissue sections were then fixed with 4% PFA paraformaldehyde, followed by acetylation treatment and subsequent washing. Next, tissue sections were pre-treated with pre-hybridization solution and pre-hybridized at 65°C for 1 hour. Afterwards, 500ng/ml labeled probes were added and incubated in the dark at 65°C for 48 hours. Finally, nuclei were stained after washing, and the sections were sealed and observed.

QUANTIFICATION AND STATISTICAL ANALYSIS

Cell segmentation of rice embryos

To visualize the cell-cell boundary, we performed fluorescent brightener (FB) staining of the plant cell walls. Images of a total of 11 stained sections were manually annotated to distinguish cell wall boundary and a total of 68,380 cells of rice embryos were determined as ground-truth (GT). The annotated datasets were divided into training and test two parts for cell segmentation modeling. The models were trained using the Cellpose 2.0 software and optimized based on the pertained model.¹⁵

In order to determine the optimal model, centroids of each mask in the prediction and ground-truth were calculated. DICE¹⁸ coefficient was then calculated between the predicted masks and the ground-truth in each cell to evaluate the prediction accuracy. The DICE value ranges from [0,1].

$$DICE = \frac{2|X \cap Y|}{|X| + |Y|}$$

X representing the predicted cell image segmentation; Y representing the image of the GT cells.

A valid match was called if the cells in GT could be predicted at a given DICE threshold. The valid matches were defined as true positives (TP), while the leftover predicted masks were defined as false positives (FP). Recall was calculated at various DICE levels to evaluate the performance of different models.

$$\text{Recall} = \frac{TP}{TP+FN}$$

scRNA-seq and Stereo-seq data analysis

The scRNA-seq data were analyzed with the standard BD Rhapsody™ Whole Transcriptome Assay Analysis Pipeline on Seven Bridges (<https://www.sevenbridges.com>) according to the manufacturer's recommendations. Fastq files were generated using a DNBSEQ-T10 sequencer, and the raw data were processed according to the Stereo-seq method.¹⁴ The expression data of genes in the corresponding regions of tissue were obtained through the lasso function in the STOmics (<https://stereomap.cngb.org/>). The raw count matrices were aligned with the results of cell segmentation using in-house scripts. Meanwhile, we also generated Bin50 matrix (50 × 50 DNB bins) as the reference for the following analysis. Basic quality control of scRNA-seq and Stereo-seq count matrices was implemented using the R package Seurat (v4.0.0)¹⁹ to remove low-quality cells and genes based on the default parameter values. Specifically, we set the threshold for the minimum number of genes must express to 200 per cell (min.features=200). Additionally, a gene must be detected in at least three cells to be retained (min.cells=3). Furthermore, we discarded cells with a high percentage of mitochondrial genes (>5%).

The expression matrices were then normalized via the SCTransform function in Seurat (v4.1.0) in order to eliminate any effects arising from differences in sequencing depth. Harmony⁴⁵ was utilized for dataset integration. For dimension reduction of the data, principle component analysis (PCA) was conducted and the top 30 PCs (nPC=30) were retained for further analysis. Cluster analysis was performed using the FindClusters function. The results of the clustering were visualized through dimension reduction analysis UMAP. When annotating cell types, consideration was first given to the location information of the identified clusters, followed by a comparison of marker gene sets to arrive at the final cell type annotations. Genes differentially expressed across cell types or clusters were identified by function FindAllMarker with Wilcoxon Rank Sum test. For cell types, the following cutoffs were applied: logfc.threshold = 0.25, min.pct = 0.25.

Monocle2 (v2.18.0)⁴³ and Monocle3 (v0.2.3.0)⁴⁴ were used for trajectory analysis (<http://cole-trapnell-lab.github.io/monocle-release>). We first normalized and integrated data by SCTransform function and harmony, respectively, before inputting them to monocle2 and monocle3. For the analysis of SC region, we chose the SCL1-2 cluster as the starting point for trajectory analysis. The top marker genes of each cell type were computed by Seurat (v4.1.0). Finally, we performed dimensionality reduction by the DDRTree method and visualized it by plot_cell_trajectory function.

Gene module analysis

Gene module expression score is a measure of how much a cell expresses a predefined set of genes. We used this score to analyze the expression patterns of gene modules for plant hormone and substance metabolites in different cell types. We selected these gene sets manually from GO and other databases (the gene lists shown in [Table S3](#)). We calculated the gene module expression score for each cell type using the AddModuleScore function in the Seurat with the default parameters.

Time series expression pattern analysis

We conducted time series expression pattern analysis on full embryos and each individual cell types using the R package Mfuzz.²⁸ The Stereo-seq SCT normalized gene expression for cells from a same germination time point was aggregated to form a pseudo-bulk gene expression matrix. Prior to clustering, genes with an expression level < 0.25 at all-time points were removed. The normalized data of the remaining genes were then Z-score transformed before executing the mfuzz_cluster.

GO enrichment analysis

GO enrichment analysis was performed using R package clusterProfiler⁴⁶ with IRGSP-1.0 annotation as the background. The smaller the P-value is, the more the GO term is significantly enriched.

Cell type deconvolution

We conducted bulk RNA sequencing on both the OsMFT2 knockout line and the WT at 24HAI, with three replicates for each condition. Cell type abundance deconvolution was performed using CIBERSORTx,⁵⁰ employing the signature matrix generated from spatial transcriptomic expression matrices of rice embryo clusters obtained via Stereo-seq data at 24HAI. The normalized bulk RNA-seq results were utilized as the mixture matrix to estimate the abundance of constituent cell types within the mixed cell population.