Journal of Genetics and Genomics xxx (xxxx) xxx

Contents lists available at ScienceDirect

Journal of Genetics and Genomics

Journal homepage: www.journals.elsevier.com/journal-of-genetics-and-genomics/

Original research

Single-cell transcriptome atlas of the leaf and root of rice seedlings

Yu Wang [a, b, 1], Qing Huan [a, *, 1], Ke Li [a], Wenfeng Qian [a, b, *]

[a] State Key Laboratory of Plant Genomics, Institute of Genetics and Developmental Biology, Innovation Academy for Seed Design, Chinese Academy of Sciences, Beijing 100101, China
[b] University of Chinese Academy of Sciences, Beijing 100049, China

ARTICLE INFO

Article history:
Received 8 June 2021
Accepted 9 June 2021
Available online xxx

Keywords:
Single-cell RNA-seq
Rice seedling
Cell-to-cell heterogeneity
Plant tissue layer
Developmental trajectory
Abiotic stress

ABSTRACT

As a multicellular organism, rice flourishes relying on gene expression diversity among cells of various functions. However, cellular-resolution transcriptome features are yet to be fully recognized, let alone cell-specific transcriptional responses to environmental stimuli. In this study, we apply single-cell RNA sequencing to both shoot and root of rice seedlings growing in Kimura B nutrient solution or exposed to various abiotic stresses and characterize transcriptomes for a total of 237,431 individual cells. We identify 15 and nine cell types in the leaf and root, respectively, and observe that common transcriptome features are often shared between leaves and roots in the same tissue layer, except for endodermis or epidermis. Abiotic stress stimuli alter gene expression largely in a cell type-specific manner, but for a given cell type, different stresses often trigger transcriptional regulation of roughly the same set of genes. Besides, we detect proportional changes in cell populations in response to abiotic stress and investigate the underlying molecular mechanisms through single-cell reconstruction of the developmental trajectory. Collectively, our study represents a benchmark-setting data resource of single-cell transcriptome atlas for rice seedlings and an illustration of exploiting such resources to drive discoveries in plant biology.

Copyright © 2021, The Authors. Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, and Genetics Society of China. Published by Elsevier Limited and Science Press. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

A plant is composed of cells of various types, which function as a whole to secure successful vegetative growth. Since these cells carry mostly identical genetic information, the variation in functionality is created from the programmed regulation of transcriptional activities along the path of cell differentiation. Understanding heterogeneity in single-cell transcriptomes, therefore, will offer unique insights into plant physiology and development. However, such heterogeneity is often undetectable by the traditional bulk RNA sequencing (RNA-seq) analyses because the specific gene expression of a cell can be obscured by the transcriptomes of other cells.

Individual plant cells or cell types have been isolated with glass microcapillaries (Lieckfeldt et al., 2008; Han et al., 2017), laser microdissection (Jiao et al., 2009; Frank and Scanlon, 2015), ribosome tagging followed by translating ribosome affinity purification (Mustroph et al., 2009; Jiao and Meyerowitz, 2010), and fluorescence-activated cell sorting of reporter-expressed protoplasts (Birnbaum et al., 2003; Dinneny et al., 2008; Efroni et al., 2015). Combined with high-throughput transcriptome profiling approaches such as microarray and RNA-seq, these approaches have obtained promising results, such as the identification of sequence motifs for cell type-specific promoters (Jiao et al., 2009) and the detection of widespread heterogeneous and monoallelic gene expression (Han et al., 2017). However, the application of such approaches is generally limited to cells with prior knowledge of precise location in tissues or cell type-specific promoters.

The past decade has witnessed the development and application of single-cell RNA sequencing (scRNA-seq) techniques, which can characterize the transcriptome profiles of individual cells. Among a variety of scRNA-seq platforms, 10x Genomics encapsulates cells into individual "gel bead-in emulsion (GEM)" droplets using a microfluidic device, and therefore, the transcriptomes of cells of various functions can be sequenced in parallel (Zheng et al., 2017); the identities of individual cells can be revealed with following-up computational analyses (Efroni et al., 2015; Stuart et al., 2019). scRNA-seq has been mostly applied to animal tissues, albeit a fast-growing number of applications to plants were implemented since 2019. Early plant studies have been focused on the *Arabidopsis* root

* Corresponding authors.
E-mail addresses: qhuan@genetics.ac.cn (Q. Huan), wfqian@genetics.ac.cn (W. Qian).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.jgg.2021.06.001
1673-8527/Copyright © 2021, The Authors. Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, and Genetics Society of China. Published by Elsevier Limited and Science Press. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Please cite this article as: Y. Wang, Q. Huan, K. Li *et al.*, Single-cell transcriptome atlas of the leaf and root of rice seedlings, Journal of Genetics and Genomics, https://doi.org/10.1016/j.jgg.2021.06.001

(Denyer et al., 2019; Jean-Baptiste et al., 2019; Ryu et al., 2019; Shulse et al., 2019; Turco et al., 2019; Zhang et al., 2019; Rich-Griffin et al., 2020; Wendrich et al., 2020). In general, the application of scRNA-seq remains limited to only a few plant species, presumably due to the technical obstacles caused by the difficulty in protoplast isolation and the limited prior knowledge on cell identity and developmental biology.

Rice (*Oryza sativa*) not only provides the staple food for more than half of the world's population but is also a monocotyledonous model plant. It is well known that the fibrous root system of rice is developed from root meristematic cells, which can undergo cell division to self-renew or differentiate into specialized tissues (Rebouillat et al., 2009; Zeng et al., 2016). From the outermost to the innermost layers, the rice root is mainly composed of epidermal, cortical and endodermal, and vascular cells, which are differentiated from their respective initial cells. Some epidermal cells are further specialized into root hairs in the zone of maturation, facilitating water and nutrient absorbance. Procambial cells, which are differentiated from vascular initial cells, can further develop into the vascular system, consisting of xylem and phloem.

Rice leaves are developed from the shoot meristematic cell and are mainly composed of epidermal, mesophyll, and vascular cells, which are presumably developed from their respective initial cells (Zeng et al., 2016). The mesophyll is a chlorophyll-containing parenchymal (thin-walled) cell type specialized for photosynthesis. Although mesophyll cells are further differentiated into palisade and spongy parenchyma in dicotyledonous leaves, the morphologic distinction is not apparent in monocotyledons, including rice (Chonan, 1978). Another parenchymal cell type is the large parenchymal cell, which barely has chloroplasts, but participates in the formation of bundle sheath—the parenchymal cell layer packed around the vascular bundle—and the connection between the abaxial vascular bundle and the adaxial side (Zeng et al., 2016; Zhou et al., 2017).

In spite of the extensive knowledge that has been gained for rice development, the transcription map of the plant body remains to be precisely drawn at cellular resolution, and the dynamic traveling waves of gene expression over the course of development are yet to be fully characterized. For example, although the majority of mesophyll cells are plausibly differentiated from mesophyll initial cells, some mesophyll cells have been reported as differentiated from vascular initial cells (Zeng et al., 2016); it remains to be investigated if these mesophyll cells would preserve some transcriptome features of their developmental origin. It is also unknown how abiotic stresses can alter gene expression and rice development at cellular resolution. We reason that these questions will be answered, at least to some extent, with the application of scRNA-seq.

In this study, we first evaluated the applicability of scRNA-seq to rice seedlings, using *Arabidopsis* root cells as a spike-in control. We then identified major rice cell types using scRNA data obtained from roots or leaves based on the estimation of a weighted marker-based index of cell identity (MICI) for each cell and the spatial patterns of cell-specific genes in the *in situ* hybridization assays. Implementing scRNA-seq to rice whole seedlings, we further explored the transcriptional and developmental activities in response to three abiotic stresses and validated some conclusions with additional small-scale experiments. Our study showcases the power of applying scRNA-seq to rice and paves the way for the construction of the plant cell atlas (Rhee et al., 2019).

Results

Quality controls for the application of scRNA-seq to rice seedlings

Protoplasting and encapsulating, two critical steps of plant scRNA-seq library preparation, have been successfully applied to roots of the dicotyledonous model plant *Arabidopsis thaliana* (Denyer et al., 2019; Jean-Baptiste et al., 2019; Ryu et al., 2019; Shulse et al., 2019; Turco et al., 2019; Zhang et al., 2019; Wendrich et al., 2020). To evaluate if a similar experimental strategy can be extended to the monocotyledonous model plant *Oryza sativa*, we hydroponically grew a *japonica* rice cultivar (Nipponbare) in Kimura B nutrient solution for two weeks and harvested both the proximal shoots (~8 cm in length above the root-shoot junction, including mainly leaves and but probably also the shoot apex surrounded by leaf sheaths) and tips of the primary and crown roots (~2 cm in length, where lateral roots were not visible but their primordium might exist; Fig. 1A). By harvesting only these fast-growing tissues, we increased the proportion of meristematic cells, which would be otherwise too scarce. We isolated protoplasts through cell wall digestion in an enzyme solution containing both cellulase and macerozyme.

Our pilot protoplasting experiments revealed that the cell wall of rice seedlings was more recalcitrant to enzyme digestion than that of *Arabidopsis* roots. To increase protoplasting efficiency for rice tissues, we followed a previous protocol that promotes the infiltration of the enzyme solution into plant tissues by vacuum (Evrard et al., 2012). Briefly, we applied vacuum for 10 min after adding rice tissues to the enzyme solution and extended the incubation time from the conventional duration of 1 h for *Arabidopsis* (Birnbaum et al., 2005) to 2 h.

Presumably, the vacuum and extended enzyme incubation during protoplasting could lead to overdigestion of the cell wall and disruption of the cell membrane, releasing RNA of dead cells into cell suspension. Such background RNA molecules require attention because they could then be randomly encapsulated into individual GEM droplets, blurring the expression heterogeneity among individual cells. To assess the quality of the isolated protoplasts, we stained rice protoplasts with trypan blue, a dye that can selectively color dead protoplasts blue. The fraction of protoplasts maintaining cell membrane integrity was (66/79=) 84% (Fig. S1), indicating that our modified experimental procedure isolated mostly alive protoplasts.

Seeking to determine if the background RNA released from the small proportion (1−84% = 16%) of dead protoplasts could obscure the single-cell transcriptomic feature of alive protoplasts, we mixed 200 roots of one-week-old *Arabidopsis* seedlings—as a spike-in control—with four rice seedlings, isolated protoplasts from the mixed sample, and performed scRNA-seq using the 10x Genomics and Illumina platforms (Fig. 1A). To assess if one of the major features, the species origin, of a cell can be inferred from its transcriptome, we mapped sequencing reads to a concatenated reference of the rice and *Arabidopsis* genomes (Fig. 1B). Among the 3260 plant cells being sequenced, 2042 and 1049 cells had >90% of transcripts from the rice and *Arabidopsis* genomes, respectively (Fig. 1C; Table S1). They added up to ([2042 + 1049]/3260=) 94.8% of cells. A similar percentage ([3005 + 851]/4031 = 95.7%) was also observed in a biological replicate (Fig. S2A). These observations confirm that despite the presence of some background RNA in the protoplast suspension, the obtained single-cell transcriptomes can provide valuable information for the detection of the gene expression features of individual plant cells.

Protoplast isolation is known to alter some genes' expression levels in *Arabidopsis* (Birnbaum et al., 2003). It is conceivable that the non-native expression information of such genes could lead to inaccuracy in the downstream computational analysis that clusters cells based on transcriptomes. To test if protoplasting also alters the expression levels of some rice genes, we performed bulk RNA-seq before and after protoplasting for the *Arabidopsis*-spiked rice samples (Fig. 1A), each with two biological replicates. Although the expression levels in the two biological replicates were highly correlated among rice genes ($\rho > 0.95$ in both treatments, Spearman's correlation, Fig. S2B and S2C), the expression levels were less

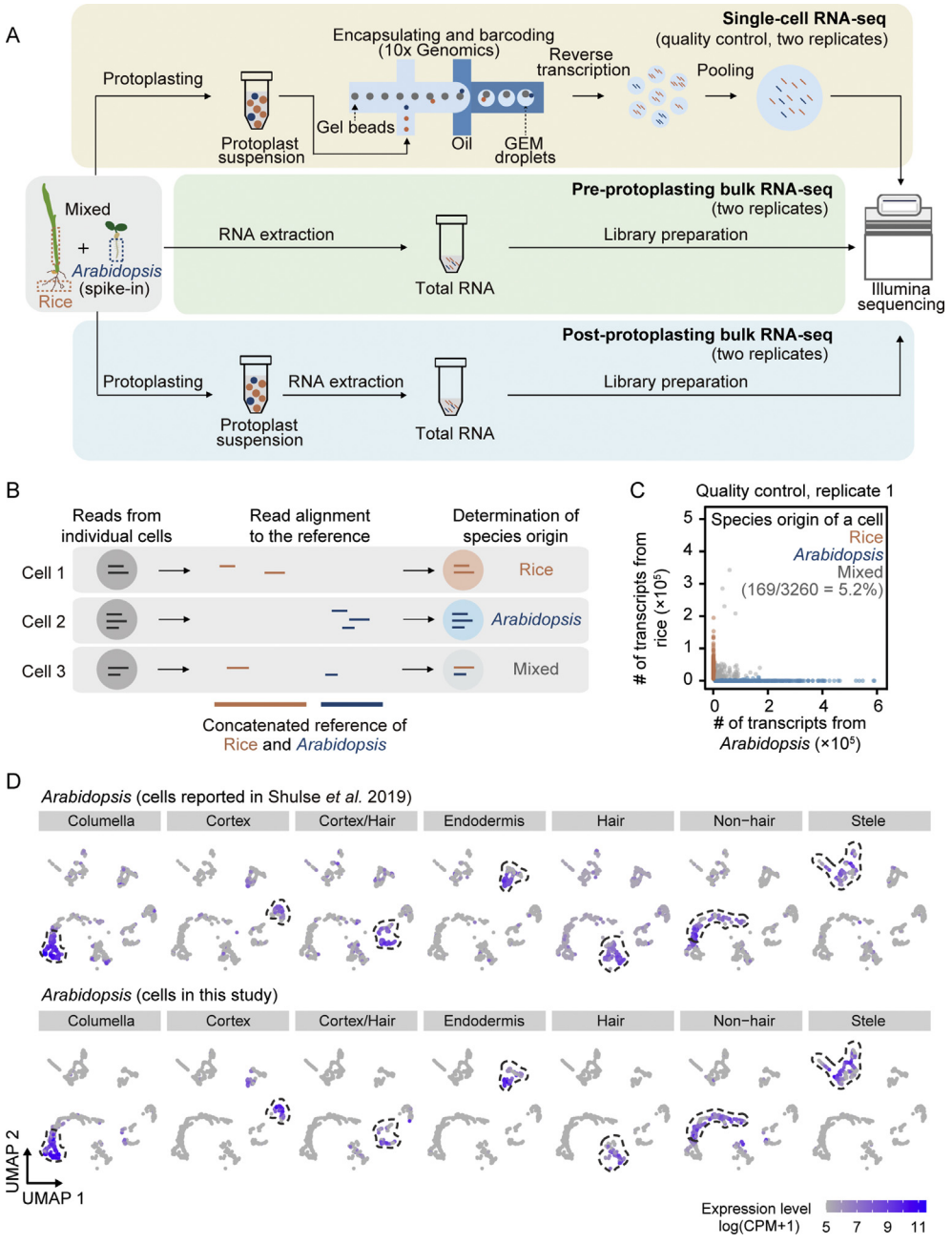

Fig. 1. Quality control experiments for the application of scRNA-seq to rice seedlings. **A**: A schematic shows the workflow to validate the application of scRNA-seq to rice seedlings. The tissues from rice seedlings (shown in the salmon boxes) were mixed with the roots of *Arabidopsis* seedlings (shown in the blue box); the latter served as a spike-in control. Isolated protoplasts were loaded onto the 10x Genomics platform, where individual protoplasts were encapsulated into "gel beads in emulsion" (GEM) droplets. After cell lysis and reverse transcription inside individual GEM droplets, they were pooled for the amplification of complementary DNA and the construction of single-cell RNA-seq libraries for Illumina sequencing (top panel). The accompanied bulk RNA-seq was also performed, with (bottom panel) or without (middle panel) protoplasting, on the mixed tissues from rice and *Arabidopsis*. Two biological replicates were conducted for each experiment, and for each replicate, four rice seedlings were mixed with the whole roots of 200 one-week-old *Arabidopsis* seedlings. **B**: A schematic shows the computational workflow to determine the species origin of a cell based on its transcriptome. Sequencing reads from each cell (those sharing the same cell barcode) were aligned to a concatenated reference genome of rice and *Arabidopsis*. The species origin of a cell was assigned to a species if >90% of its transcripts were from the corresponding genome. The cell was otherwise determined as "mixed", which is possibly caused by the background RNA contamination. Colors of reads and cells indicate the determined species origin. **C**: A scatter plot shows the number of detected *Arabidopsis* transcripts vs. rice transcripts for each cell. Blue and salmon dots represent cells that were determined as Arabidopsis and rice cells, respectively. Overall, 5.2% cells were determined as "mixed" cells, which are indicated by gray dots. The result of replicate 1 is shown; that of the other biological replicate is shown in Fig. S2A. **D**: The expression of seven cell type-specific genes in *Arabidopsis* roots identified in the previous single-cell transcriptome analysis (Shulse et al., 2019). The cell type-specific genes are *AT5G54370*, *AT1G12090*, *AT4G12545*, *AT1G05260*, *AT3G54580*, *AT1G28290*, and *AT1G12080* from left to right. Single-cell transcriptome data from the two studies were integrated as described in the main text and Fig. S3 and were plotted onto the same two-dimensional UMAP space. Dots represent individual cells, colored according to the expression levels of the corresponding cell type-specific genes. The dashed lines were used to circle the putative cell population of the corresponding cell type on the top panels; they were copied and pasted to the bottom panels. CPM, transcript count of a gene per one million transcripts in the cell.

correlated between the pre- and post-protoplasting bulk RNA-seq data ($\rho = 0.62$, $P < 2.2 \times 10^{-16}$, Spearman's correlation, the mean expression of two replicates was used here and hereafter; Fig. S2D), indicating that the expression level of some rice genes changed upon protoplasting.

To identify rice genes differentially expressed on protoplasting, we compared gene expression levels between the pre- and post-protoplasting bulk RNA-seq data. Since *Arabidopsis* roots were spiked (Fig. 1A), we first tested as a positive control if previously reported protoplasting-inducing *Arabidopsis* genes (Birnbaum et al., 2003) would show increased expression levels on protoplasting in our data. Indeed, we observed elevated expression levels for these genes during protoplasting ($P < 2.2 \times 10^{-16}$, paired Mann-Whitney *U* test, pre- *vs*. post-protoplasting, $N = 342$; Fig. S2E), bolstering our confidence for detecting differentially expressed genes (DEGs) on protoplasting.

We defined genes exhibiting consistent differential expression on protoplasting [$\log_2$ (fold change) > 3 or < -3] in both biological replicates as protoplasting-sensitive genes (Fig. S2F and S2G; Tables S2 and S3). Under this definition, we identified 2231 *Arabidopsis* and 6344 rice protoplasting-sensitive genes. As expected, when these genes were excluded, the Spearman's correlation coefficient between the pre- and post-protoplasting RNA-seq data increased from 0.76 and 0.62 to 0.82 and 0.81, among *Arabidopsis* and rice genes, respectively (Fig. S2D, S2H, S2I, S2J), indicating that the rest 17,126 and 20,538 expressed genes represent a more protoplasting-robust gene set for *Arabidopsis* and rice, respectively. The expression information of the protoplasting-sensitive genes was consequently excluded from the downstream dimensionality reduction and cell clustering.

It is worth noting that the 10x Genomics platform captures only a fraction of protoplasts loaded onto the microfluidic chip (3260/6000 = 54% in replicate 1 and 4031/6000 = 67% in replicate 2; Fig. 1A). Therefore, the detected single-cell transcriptomes may not be competent to fully represent the loaded protoplast population owing, for example, to protoplast size—biased encapsulating. Considering such potential biases, we combined the transcriptomes of all captured cells ("pseudobulk" hereafter) and determined if it can be used to recover the global expression profile obtained from the postprotoplasting bulk RNA-seq.

In the pseudobulk, we detected the expression of 26,793 and 28,064 out of the 38,869 nontransposable element-related rice genes in two replicates (68.9% and 72.2%, respectively; Table S1), a percentage similar to that in the accompanied postprotoplasting bulk RNA-seq experiments (58.9% and 63.1%; Table S4). Furthermore, the gene expression levels in the pseudobulk were highly correlated with the pre- and post-protoplasting bulk RNA-seq (Spearman's correlation coefficient $\rho = 0.81$ and 0.92, $N = 20{,}836$ and 21,282, respectively, $P < 2.2 \times 10^{-16}$ for both; Fig. S2K and S2L), confirming that scRNA-seq captured a representative population of rice seedling cells.

As we mixed *Arabidopsis* roots with rice samples for scRNA-seq (Fig. 1A), in addition to rice cells, we also obtained single-cell transcriptomes for a total of 1900 *Arabidopsis* root cells (Table S1). We reasoned that these *Arabidopsis* data could serve as a positive control to evaluate if the scRNA-seq data generated from our modified protoplasting protocol can recapitulate the previously reported single-cell transcriptome atlas for *Arabidopsis* roots. To this end, we corrected the technical/batch effects generated during library preparation (Stuart et al., 2019) by computationally integrating these 1900 *Arabidopsis* single-cell transcriptomes with those obtained previously for the 2056 *Arabidopsis* cells isolated from 1-week-old roots (Shulse et al., 2019).

Here, we briefly describe the data normalization and integration pipeline (Fig. S3); detailed parameters can be found in the Materials and Methods section. To control for the variation in the cellular sequencing depth among cells, we employed "sctransform" to calculate the normalized gene expression level as the Pearson residual estimated from the regularized negative binomial regression (Hafemeister and Satija, 2019). We selected 3000 genes with the highest among-cell variability in expression within samples, which will be used for the downstream dimensionality reduction. To ensure coherent cell clustering among samples, we performed canonical correlation analyses and captured the most correlated gene expression features to align cells across samples (Butler et al., 2018; Stuart et al., 2019). After that, we searched for the mutual nearest neighbors (Haghverdi et al., 2018) and used them as "anchors" to further transform the gene expression data sets. We performed principal component analysis on the transformed data and, based on the first 100 principal components, projected cells of different samples onto the same two-dimensional space using uniform manifold approximation and projection (UMAP) (McInnes et al., 2018).

We found that the *Arabidopsis* root cells detected in our study largely overlaid with the previously reported single-cell transcriptome atlas on the UMAP graph (Fig. 1D). All root cell types identified in the previous study (Shulse et al., 2019)—including columella, cortex, cortex/hair, endodermis, hair, nonhair, and stele—can be recovered from our scRNA-seq data, and the expression of the marker genes that were previously identified being specifically expressed in certain cell types remained confined in the corresponding cell groups in our study (Fig. 1D). These observations verified the reliability of our experimental and computational procedures for single-cell transcriptome analyses.

Cell type assignment using MICI

To obtain a comprehensive single-cell transcriptome atlas for rice seedlings, we increased the number of rice cells loaded onto the 10× Genomics microfluidic chips (Fig. 2A). Specifically, we obtained single-cell transcriptomes for a total of 55,328 rice cells from four biological replicates (labeled as Seedling-1 to Seedling-4; Table S1), and for each replicate, the loaded protoplasts were isolated from a mixture of leaves and root tips of 16 rice seedlings. To dissect the organ origin (leaf vs. root) of individual cells, we also separately sequenced the single-cell transcriptomes for proximal shoots and root tips, each with two biological replicates (Fig. 2A); among them, samples Leaf-1 and Root-1 were collected from 16 rice seedlings, and samples Leaf-2 and Root-2 were from another 16 rice seedlings.

After normalizing and integrating the scRNA-seq data across samples as described previously (Fig. S3), we applied the Louvain method—an unsupervised algorithm that iteratively extracts cell communities with the goal of optimizing modularity (Blondel et al., 2008)—to identify cell clusters. We detected 29 cell clusters; each contained a variable proportion of cells from 0.6% to 9.9% (Fig. 2B). Biological replicates largely overlaid with each other on the UMAP graphs for leaf, root, or seedling samples (Fig. S4A–S4C), verifying that our scRNA-seq data are reproducible. We then performed a data-driven analysis to identify specifically expressed genes in each cell cluster: genes exhibiting the largest expression fold difference between the cluster being queried and all other clusters (Fig. S5, and the top five genes for each cluster are provided in Table S5).

We next sought to assign a cell type to each cell cluster based on the single-cell transcriptomes obtained from the leaf and root samples (Figs. 2C and 3A). We noticed that most clusters were composed of cells from both leaves and roots (Fig. S4D), and therefore, we assigned a leaf and a root cell type to each cell cluster, except that only a leaf cell type was assigned to Clusters 23 and 24 due to the paucity of root cells (< 5% root cells) in the respective cluster. We combined cells from two replicates because the cell

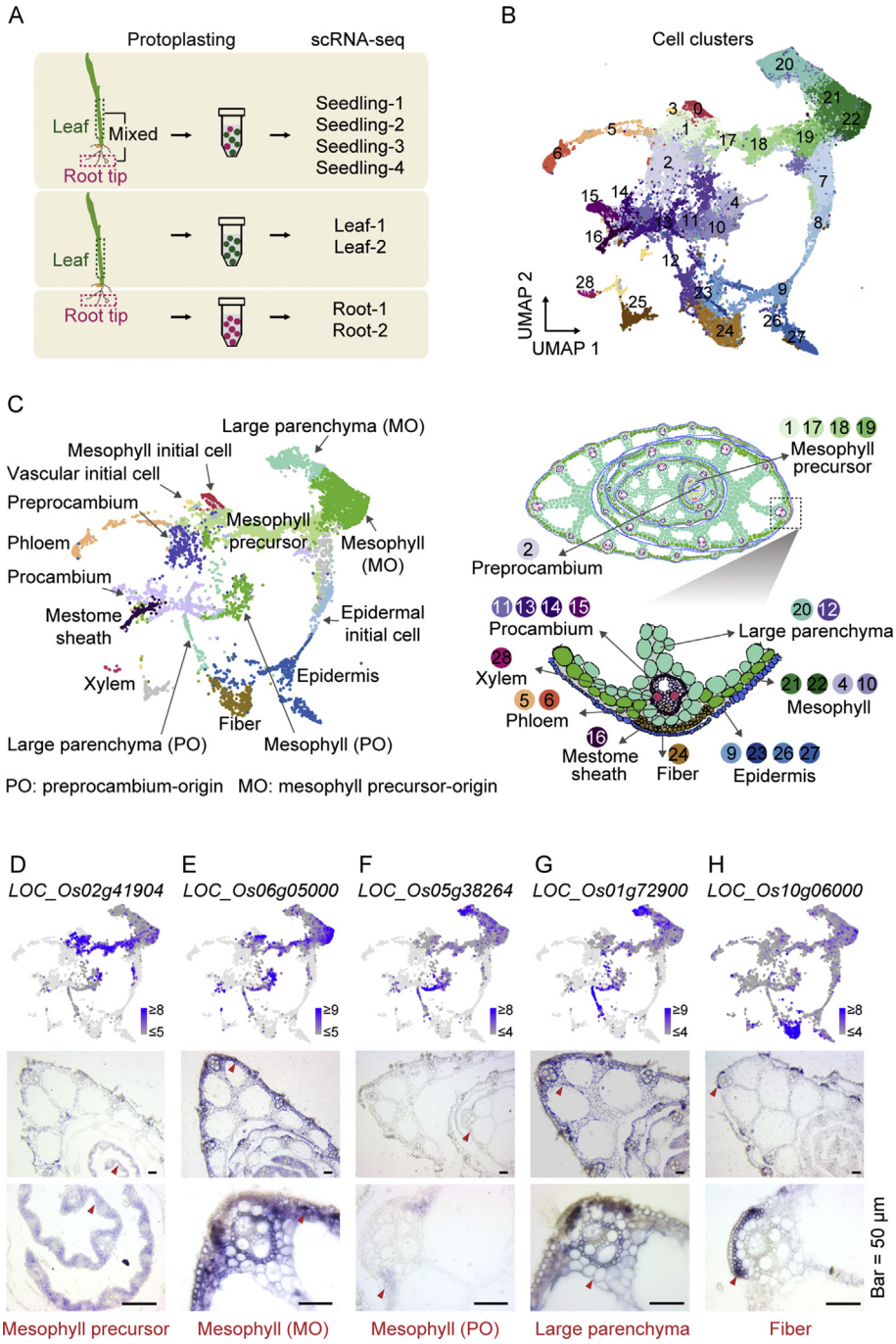

Fig. 2. Single-cell transcriptome atlas for the rice leaf. **A**: A schematic shows the scRNA-seq samples for proximal shoots (labeled as Leaf-1 and Leaf-2), root tips (Root-1 and Root-2), and mixed tissues (Seedling-1, Seedling-2, Seedling-3, and Seedling-4) of two-week-old rice seedlings. **B**: Visualization of the 29 cell clusters on the UMAP graph. Data integration and dimensionality reduction were performed based on the scRNA-seq data for two leaf, two root, and 16 seedling samples; note that in addition to the four seedling samples mentioned in panel (**A**), 12 seedling samples that would be collected later in this study (upon three stress treatments) were also included, in an effort to make all UMAP graphs in this study readily comparable to each other. Dots represent individual cells from two leaf, two root, and four seedling samples in panel (**A**), colored according to their cell clusters. **C**: Visualization of the 15 leaf cell types on the UMAP graph; dots represent individual cells from the two leaf samples, colored according to their cell types (left panel). The right panel shows the anatomic schematic of the rice leaf; cells are colored according to their cell types, and circled cluster numbers are colored according to (**B**). Preprocambium and mesophyll precursors are labeled on younger leaves but may also exist in the margin regions of older leaves. Initial cells (Clusters 0, 3, and 8) and unknown cells (Clusters 7 and 25) are not labeled. **D**–**H**: The UMAP visualization and RNA *in situ* hybridization assays for five cluster(s)-specific genes in the rice leaf. Cells are colored according to the normalized gene expression levels, log (CPM + 1) on the UMAP graph. In the *in situ* hybridization images, the identified cell types (showing on the bottom) are indicated by red triangles; the length of a scale bar represents 50 μm. In panels (**D**–**G**), the MICI-defined nonmesophyll cells are colored light gray to highlight the expression specificity of a gene among mesophyll cells (Clusters 1, 4, 10, 12, and 17–22).

population proportions were positively correlated between biological replicates (Pearson's correlation coefficient $r = 0.92$ and 0.58, $P = 9.4 \times 10^{-13}$ and 9.9×10^{-4} for root and leaf samples, respectively; Fig. S4E and S4F). As a negative control, the Pearson's correlation between the root and leaf samples was confirmed insignificant ($r = 0.06$, $P = 0.75$; Fig. S4G).

We intended to develop a computational approach for the cell type assignment to cell clusters based on the identity of each cell:

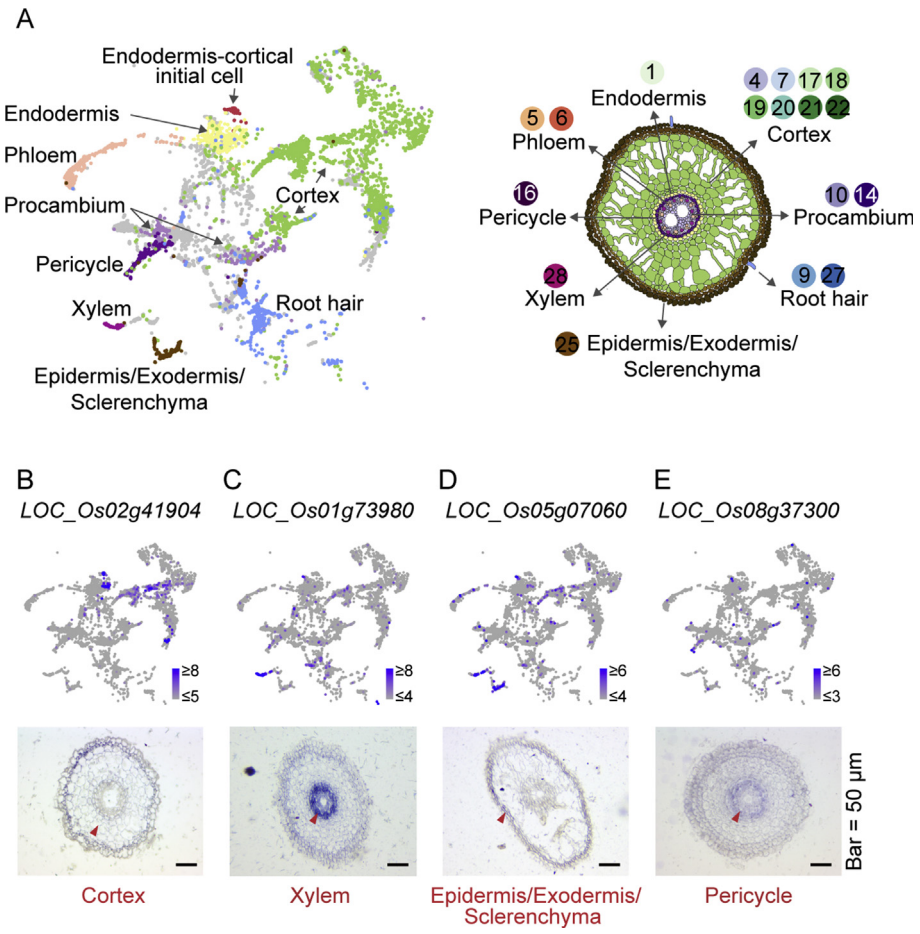

Fig. 3. Single-cell transcriptome atlas for the rice root. **A**: Visualization of the nine root cell types on the UMAP graph; dots represent individual cells from the two root samples, colored according to their cell types (left panel). Initial cells (cluster 0) and unknown cells (clusters 2, 3, 8, 11–13, 15, and 26) are not labeled. **B**–**E**: The UMAP visualization and RNA *in situ* hybridization assay for the expression of four cluster(s)-specific genes in the rice root.

first determining the cell type for each cell and then assigning a cell type to a cluster being queried according to the majority. Although the previously defined index of cell identity can estimate for each cell the propensity of being a cell type based on the reported cell type-specific transcriptomes (Efroni et al., 2015), the approach is likely not appropriate for rice because the cell type-specific transcriptomes remain relatively scarce for rice tissues compared with the wealth of such data for *Arabidopsis* roots, with only a few exceptions (Jiao et al., 2009; Takehisa et al., 2012).

To overcome this hurdle, we developed an index, MICI (see the Materials and Methods section and Fig. S6 for details), for each cell to gauge its propensity for being a predefined cell type, which was the weighted sum of the scaled expression levels of all marker genes of a cell type. Specifically, we applied logarithm transformation to reduce the skewness of the raw expression data (in the unit of transcript count of a gene per one million transcripts in a cell, CPM) and scaled among all cells to standardize the range of the expression level across genes. We further determined the weight of each marker gene from the variance in the cluster-averaged expression level among cell clusters: the greater variation among clusters, the more informative a marker for cell type determination.

We compiled from previous studies a list of 40 marker genes for 13 leaf cell types (initial cell, epidermis, guard cell, trichome, mesophyll, preprocambium, procambium, fiber, mestome sheath, phloem, phloem precursor, xylem, and xylem precursor) and 15 marker genes for five root cell types (initial cell, procambium, phloem, root hair, and xylem precursor) and the lateral root primordium (Table S6). Note that we simply followed the conventional use of these terms, which are not necessarily mutually exclusive; for example, the guard cell, trichome, and root hair are subtypes of epidermal cells, and both the xylem precursor and phloem precursor are subtypes of procambial cells (Zeng et al., 2016; Liu et al., 2020). We individually labeled the expression levels of these marker genes on the UMAP graphs, and visually, many of them indeed showed some level of cell cluster specificity in the leaf and/or root samples (Fig. S7).

For each given cell whose transcriptome was obtained, we calculated a MICI value for each of the predefined cell types aforementioned. We designated the cell type corresponding to the highest MICI value to the cell being queried, except that the cell type was considered "undetermined" if none of the MICI values of a cell was >2 for any cell type. The cell type of the largest cell population of a cell cluster was assigned to the cell cluster (Fig. S6), and if the largest cell population was "undetermined", the cell cluster was assigned as "unknown".

Fine assignment of leaf cell types

Implementing MICI, we successfully assigned seven leaf cell types—initial cell, mesophyll, epidermis, mestome sheath, phloem, procambium, and xylem—to 26 cell clusters (Fig. S8A; Table S7). The cell type of the remaining three clusters was "unknown" and some cell types were assigned to numerous cell clusters. For example,

mesophyll was assigned to Clusters 1, 4, 10, and 17−22 (Fig. S8A; Table S7). Moreover, mesophyll represented the largest cell population of Cluster 12 except for the "undetermined" cells. These 10 clusters made up (5191/9256=) 56.1% of cells detected in leaf samples. While mesophyll cells are expected to be reasonably abundant in the leaf, we speculated that the transcriptomic variation among these clusters, as reflected by the various coordinates on the UMAP graph—which is capable of preserving the global data structure (McInnes et al., 2018)—may bear some genuine developmental differences.

In particular, mesophyll cells gathered into two "superclusters" on the UMAP graph (Fig. S8A)—Supercluster 1 consisting of Clusters 1 and 17−22, and Supercluster 2 consisting of Clusters 4, 10, and 12—suggesting the presence of genome-wide transcriptional differentiation among mesophyll cells. We first focused on Supercluster 1 since it made up the majority (4461/5191 = 85.9%) of the MICI-defined mesophyll cells. To dissect the subtypes within the supercluster, we performed *in situ* hybridization on cluster(s)-specific genes to localize the cells being queried into the cross-section of leaf tissues. *LOC_Os02g41904* (highly expressed in Clusters 1 and 17−19) was more expressed in inner and younger leaves, protected by the sheaths of older leaves (Fig. 2D, the negative control shown in Fig. S9), whereas *LOC_Os06g05000* (highly expressed in Clusters 21−22) were more expressed in older leaves outside (Fig. 2E), indicating that Clusters 21−22 are mainly composed of more mature mesophyll cells that are differentiated from mesophyll precursors represented by the cells in Clusters 1 and 17−19 (Fig. 2C).

We were particularly interested in mesophyll Supercluster 2, which was spatially closer to procambial cells on the UMAP graph (Fig. S8A). The *in situ* hybridization assay revealed that a Cluster10−specific gene, *LOC_Os05g38264*, was specifically localized to the bilateral region of vascular bundles (Fig. 2F), echoing the previous report that preprocambial cells transiently exist in this region during leaf development and will differentiate into mesophyll cells if they do not ultimately participate in the formation of bundle sheaths (Zeng et al., 2016). Considering this differentiation lineage, we speculated that Cluster 10 might represent mesophyll cells differentiated from preprocambial cells. Similarly, Cluster4 cells were spatially close to procambial cells on the UMAP graph but also expressed mesophyll marker genes that encode the small subunits of RuBisCo (Tsutsumi et al., 2006), RBCS1 (*LOC_Os02g05830*), RBCS2 (*LOC_Os12g17600*), RBCS3 (*LOC_Os12g19381*), and RBCS4 (*LOC_Os12g19470*; Fig. S7). We consequently assigned Clusters 4 and 10 as mesophyll cells of preprocambium origin (PO; Fig. 2C).

So far, we have identified the mesophyll of mesophyll-precursor origin (MO; Clusters 21 and 22) and that of PO (Clusters 4 and 10). To investigate the common and distinct features for them, we identified their respective expression specificity compared with other leaf cells (Fig. S10A). Among the 957 and 536 genes specifically expressed in the mesophyll of mesophyll precursor and that of PO, respectively, 92 genes were shared, significantly greater than the random expectation (odds ratio = 6.6, $P < 2.2 \times 10^{-16}$, Fisher's exact test; Fig. S10B). The common genes included *RBCS4* and *Lhcb1.1* (*LOC_Os01g52240*, encoding light-harvesting chlorophyll *a*/b-binding protein b1.1) and were enriched in the gene ontology term "oxidation reduction" (false discovery rate = 0.02; Fig. S10C). Meanwhile, a vast majority of these specifically expressed genes were not shared between the mesophyll cells of the two origins (Fig. S10B). For example, the known procambium marker gene *4-COUMARATE:COENZYME A LIGASE 3* (*4CL3*, *LOC_Os02g08100*) (Gui et al., 2011) was broadly expressed in the mesophyll of PO but not in those of MO (Fig. S10D), indicating that transcription constraints imposed by the cell differentiation lineage may help to explain the apparent partitioning of mesophylls on the UMAP graph, despite their common function of photosynthesis.

The *in situ* hybridization assay also revealed that the transcripts of a Cluster 12− and Cluster 20−specific gene, *LOC_Os01g72900*, were largely localized to the chloroplast-free large parenchymal cells (including the bundle sheath) instead of the chloroplast containing the mesophyll cells (Fig. 2G). Visually we noticed that Clusters 12 and 20 were spatially well separated on the UMAP graph (Fig. 2B), echoing the two differentiation lineages of the large parenchymal cells, the MO and the PO (Zeng et al., 2016). Following the same logic as in the developmental origins of mesophyll, we speculated that Cluster 12 might represent the large parenchymal cells of PO that participate in the formation of the bundle sheath; in contrast, Cluster 20 might represent the large parenchymal cells of MO that constitute the ground tissues of leaves as well as bundle sheaths (Fig. 2C). Their initial assignment as mesophyll cells by MICI (Fig. S8A; Table S7) may simply reflect the fact that both mesophyll and large parenchyma are parenchymal.

In addition to mesophyll cells, epidermal cells were also assigned to a few clusters by MICI (Clusters 9, 23−24, and 26−27; Fig. S8A; Table S7), which made up (958/9256=) 10.4% of leaf cells. A computationally identified candidate marker gene for Cluster 24, *LOC_Os10g06000*, exhibited specific expression in the abaxial side of vascular bundles in the *in situ* hybridization assay (Fig. 2H), which helped us assign the fiber cell (a type of sclerenchyma) to Cluster 24 (Fig. 2C). These cells were initially assigned to the epidermis by MICI, probably because of some common transcriptomic features shared by the outermost few layers of cells in the leaf.

Initial cells were assigned to three clusters by MICI (Clusters 0, 3, and 8; Fig. S8A; Table S7). We manually curated Cluster 3 to putative vascular initial cells according to their specific expression of *WUSCHEL-LIKE HOMEOBOX 4* (*WOX4*, *LOC_Os04g55590*; Fig. S7), which is known to be expressed in the precursors of vascular bundles (Zeng et al., 2016). Considering that Cluster 0 was spatially closer to mesophyll precursors on the UMAP graph and Cluster 8 was closer to epidermal cells, we assigned them as putative mesophyll and epidermal initial cells, respectively (Fig. 2C).

Furthermore, although Cluster 2 was assigned as "unknown" by MICI, 14.5% of its cells expressed *FILAMENTOUS FLOWER* (*FIL*, *LOC_Os02g42950*; Fig. S7), a known marker gene for the developing phloem in young shoots (Liu et al., 2007). Consistently, Cluster 2 was localized close to phloem and procambium on the UMAP graph (Fig. S8A), suggesting that it is composed of preprocambial cells to be developed into the phloem. More broadly, 10.9% of Cluster 2 cells expressed *CYCLIN-DEPENDENT KINASE B2; 1* (*CDKB2; 1*, *LOC_Os08g40170*), a known marker gene for the shoot apical meristem and leaf primordia (Umeda et al., 1999), and 25.4% expressed the procambium marker *4CL3* (Fig. S7). Taken together, we speculated that Cluster 2 most likely represented preprocambial cells (Fig. 2C).

Procambial cells were assigned to Clusters 11, 13−15, and 25 by MICI (Figure S8A; Table S7), which made up (1005/9256=) 10.9% of leaf cells. Because Cluster 25 was specialized at the transcriptome level, as reflected by its spatial separation from the other four clusters on the UMAP graph (Fig. 2B), and contained the smallest fraction of procambial cells among these five clusters (Table S7), we conservatively curated the cell type of Cluster 25 to "unknown." Ultimately, cell types were assigned to 27 of the 29 cell clusters in the leaf (Fig. 2C).

Fine assignment of root cell types

Using MICI, we only assigned three root cell types—phloem, procambial, and endodermis-cortical initial cell—to six clusters (Figure S8B; Table S8). The cell types of most cell clusters (21/27) were "unknown," reflecting the scarcity of marker genes in the rice root when we initiated this study. To identify additional root cell types,

we performed *in situ* hybridization assays on the computationally identified cluster-specific genes to localize the cells being queried into the cross-section of the root.

Our *in situ* hybridization assay localized the transcripts of *LOC_Os02g41904*, a gene specifically expressed in Clusters 17, 18, and 19, to the cortex (Fig. 3B); as a result, we assigned cortical cells to these clusters (Fig. 3A). As we prepared our manuscript, two recent studies reported a total of seven additional marker genes for the cortex (Liu et al., 2021; Zhang et al., 2021a). Among them, *LOC_Os10g40510*, *LOC_Os10g40520*, and *LOC_Os03g04310* showed specific expression, particularly in Clusters 4, 7, and 20–22 (Fig. S11A), and consequently, we also assigned cortical cells to these clusters (Fig. 3A).

Following a similar logic, the transcripts of a Cluster 28–specific gene, *LOC_Os01g73980*, were localized to the xylem in our *in situ* hybridization assay (Fig. 3C), so the xylem cell was assigned to Cluster 28 (Fig. 3A). Consistently, a recent study also reported *LOC_Os01g73980* as a root xylem marker gene (Liu et al., 2021), bolstering our confidence in accurate curation of root xylem cells. Besides, a Cluster 25–specific gene, *LOC_Os05g07060*, was expressed in the outermost three layers of cells—epidermis, exodermis, and sclerenchyma—as shown in the *in situ* hybridization image (Fig. 3D), and these cell types were assigned to Cluster 25 accordingly (Fig. 3A). Consistently, Cluster 25–specific genes *LOC_Os06g38960* and *LOC_Os08g02300* (Fig. S11B) were recently reported being expressed in the epidermis (Liu et al., 2021) and sclerenchyma (Zhang et al., 2021a), respectively.

Procambial cells were assigned to Clusters 10, 14, and 16 (Table S8), which made up the largest MICI-detected cell population in the root (714/1408 = 50.7%). Although procambial cells were generally expected to be abundant in the root, we speculated that they should not be homogenous: When additional marker genes become available, subtypes of cells could be identified. To test this idea, we performed *in situ* hybridization assay on a Cluster 16–specific gene, *LOC_Os08g37300*, and found it specifically expressed in the pericycle (Fig. 3E), the outermost cell layer of the vascular cylinder. Consistently, another Cluster 16–specific gene, *LOC_Os10g38040* (Fig. S11C), was recently reported as expressed in the pericycle (Zhang et al., 2021a). Taking all information into consideration, we assigned the pericycle to cluster 16.

In the two recent scRNA-seq papers on the rice root, additional marker genes were also reported for root cap, root hair, and endodermis (Liu et al., 2021; Zhang et al., 2021a), three cells types that so far have not been identified in our study. We screened all these marker genes for clues of cluster-specific expression in our scRNA-seq data (Fig. S11D–S11F). Among them, six genes reported as the markers of endodermis—*LOC_Os08g03450*, *LOC_Os01g16890*, *LOC_Os03g13170*, *LOC_Os05g39960*, *LOC_Os10g06680*, and *LOC_Os04g58760*—were unanimously expressed in Cluster 1 (Fig. S11D), so we assigned the endodermis to Cluster 1. Similarly, based on the expression of *LOC_Os07g35860* and *LOC_Os03g05640* (reported as root hair markers) and *LOC_Os03g12290* (reported as a marker of "near root hair" epidermis) in Clusters 9 and 27 (Fig. S11E), we assigned the root hair to these clusters. The five root cap marker genes reported by Liu et al. did not exhibit cluster-specific gene expression in our data (Fig. S11F; see also Discussion). Ultimately, cell types were assigned to 19 of the 27 cell clusters in the root (Fig. 3A).

Common transcriptomic features were shared between leaves and roots for the same tissue layer, except for endodermal and epidermal cells

Visually, we noticed that leaf and root cells did not gather into two groups on the UMAP graph (Fig. S4A–S4C) but were often blended, suggesting a lack of a global distinction between the leaf and root cells at the gene expression level. Indeed, leaf and root cells coexisted in most of the 29 cell clusters (Fig. S4D). This observation is not explainable by potential data overfitting during the integration of single-cell transcriptomes between the leaf and root samples (Fig. 2B) because algorithmically, single cells from the leaf and root samples were both mapped to the two-dimensional UMAP space defined by the whole seedling sample Seedling 1. Instead, the blending of leaf and root cells on the UMAP graph indicates that some leaf cells are transcriptomically more similar to root cells than to other leaf cells, and vice versa.

To identify the transcriptomically most similar leaf cell type for each root cell type and the other way round, we estimated the averaged transcriptome among all cells of a cell type for the leaf or root samples and calculated the pairwise Spearman's correlation coefficients for gene expression levels between each leaf and each root cell type among 3000 highly variable genes (Fig. 4A and 4B). We identified several cell types whose leaf and root cells were reciprocally most similar at the transcriptome level—phloem, xylem, and procambial cells. Besides, leaf mestome sheath and root pericycle were also reciprocally most similar to each other, echoing that they are both the outermost cell layer of the vascular cylinder and are both developed from vascular initial cells (Zeng et al., 2016).

The situation was more complicated for the ground tissue. On the one hand, the root cortical cells were most similar to mesophyll precursors and the mesophyll cells differentiated from them (Fig. 4A and 4B); this observation is not unexpected since mesophyll and cortical cells are well known as leaf and root parenchyma, respectively. Another somewhat expected observation is that the endodermis-cortical initial cells were most similar to mesophyll initial cells, mesophyll precursors, vascular initial cells, and epidermal initial cells (Fig. 4A and 4B), implying the shared transcriptomic feature of these fast-proliferating and self-renewing meristematic cells, regardless of the tissue layer. On the other hand, endodermis, a single layer of root cells delimiting the cortex from the vascular cylinder, is supposed to be the counterpart of the bundle sheath in the leaf but was transcriptomically most similar to meristematic cells (Fig. 4A and 4B), likely implicating its developmental potential in giving rise to lateral root primordia (Rebouillat et al., 2009).

For the epidermal tissue, the counterpart cells between leaves and roots were not readily identifiable from the transcriptome, as reflected by the absence of any >0.6 correlation coefficients observed for the root or leaf epidermis (Fig. 4B). We surmise that there are two mutually nonexclusive explanations for the observation. First, the root epidermal cells identified in the scRNA-seq data (Cluster 25) were a mixed population of epidermal, exodermal, and sclerenchyma cells (Fig. 3A and D). Exodermis and sclerenchyma are ground tissues (Rebouillat et al., 2009), and the blending of them could have obscured the major transcriptomic features of the genuine root epidermis. Second, it is also possible that the epidermal tissue is functionally more variable between leaves and roots than the ground or vascular tissues because the epidermis is the outermost layer of the plant body, directly facing the external environment, may it be air or soil.

The leaf and root cells are usually transcriptomically distinguishable even if they fall into the same cluster

Despite the overall similarity in gene expression between some leaf and root cells, we hypothesized that some DEGs would enable transcriptome-based distinction of leaf and root cells even if they belong to the same cluster. To test this idea, we calculated for each cluster the difference in each gene's expression level between leaf and root samples, estimated the corresponding *P* value from the Mann-Whitney *U* test among individual cells, and drew volcano plots for visualization (Fig. 4C and Figs. S12). Many genes appeared to be

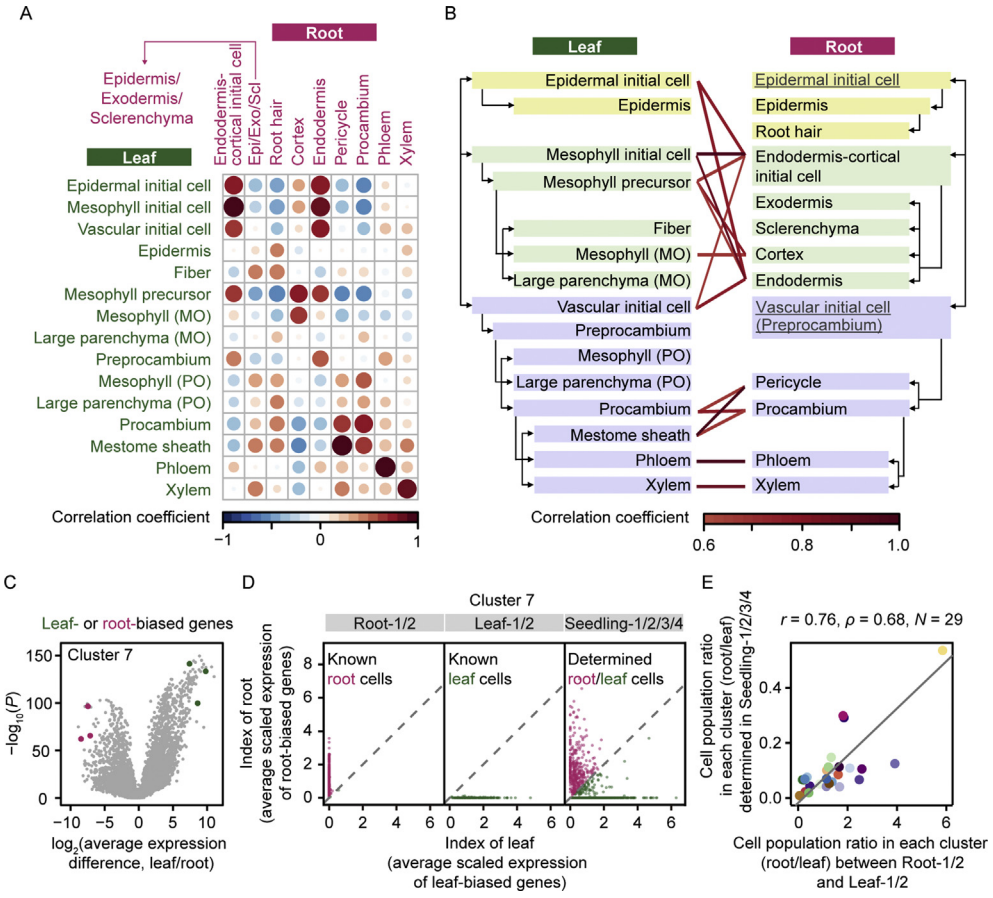

Fig. 4. The similarities and differences in transcriptomes between leaf and root cell types. **A**: The pairwise Spearman's correlation coefficients for the expression profile between a leaf cell type and a root cell type. This analysis was performed exclusively based on the single-cell transcriptomes from the two leaf and the two root samples, and the scRNA-seq data from two replicates of the same organ were combined. To estimate the correlation coefficient between two cell types, we retrieved the data integration–adjusted expression levels for the 3000 highly variable genes, averaged the gene expression level among all cells of a cell type, and calculated Spearman's correlation coefficients among the 3000 highly variable genes. The leaf and root cell types are labeled in green and magenta, respectively. The color and the area of each dot reflect Spearman's correlation coefficient (ρ) and $|\rho|$, respectively. **B**: The transcriptomic similarity between leaf and root cell types. The edges in the middle that connect two cell types represent >0.6 Spearman's correlation coefficients. The arrows on the sides represent the known cell lineage of rice development (Zeng et al., 2016). Cell types not identified in our scRNA-seq data are underlined and shown in gray. **C**: The selection of three leaf-biased genes and three root-biased genes for each cell cluster based on the single-cell transcriptomes of the leaf and root samples. The results of cluster 7 are shown here as an example, and those of other clusters are shown in Fig. S12. Each dot on the volcano plot represents a gene. For each gene, its expression ratio between the leaf and root samples is shown on the *x* axis, and the corresponding *P* value from the Mann-Whitney *U* test is shown on the *y* axis. Note that the fraction of cells expressing the gene was also considered for gene selection, but such information is not shown here (see Materials and Methods for details). Selected leaf-biased and root-biased genes are labeled in green and magenta, respectively. **D**: The determination of the organ origin (leaf vs. root) for individual cells in the seedling samples. Each dot represents a cell, and its average scaled expression level of the three selected leaf-biased (root-biased) genes is shown on the *x* axis (*y* axis). The minimum values of the indices of leaf and root were set as the origin of the coordinate system. Cells above, below, and on the y = x (dashed) line are determined as root (magenta), leaf (green), and unknown (gray) cells. The results of Cluster 7 are shown here as an example, and those of other clusters are shown in Fig. S13. The known leaf and root cells collected from the leaf and root samples are shown as positive controls. **E**: The correlation between the population ratio (root/leaf) of determined cells in the seedling samples and that of detected cells in the root and leaf samples, among the 29 cell clusters. Each dot represents one of the 29 cell clusters, colored according to Fig. 2B. The coefficients of both Pearson's and Spearman's correlation are shown on the top ($P < 1 \times 10^{-4}$ in both statistical tests). The straight line shows the standard major axis estimated from the "lmodel2" function in the lmodel2 package of *R*. The *x* and *y* axes differ by one order of magnitude, likely because a similar number of root and leaf protoplasts were sequenced in the leaf and root samples regardless of their relative abundance in the seedling samples.

differentially expressed between leaves and roots, indicating that leaf and root cells in the same cluster likely retain their respective transcriptomic features.

To distinguish the leaf and root cells within a cell cluster for the seedling samples, we defined the index of leaf and the index of root for each cell to gauge the likelihood of a cell being isolated from leaves or roots, respectively. Specifically, for each cluster, we estimated a leaf-biased expression score for each gene from the scRNA-seq data of the leaf and root samples, taking both the gene's average expression level and the fraction of cells expressing the gene into consideration (see the Materials and Methods for details). Based on the biased expression scores, we defined three leaf-biased genes and three root-biased genes for each cell cluster (Fig. 4C and Figs. S12; Table S9). We calculated the index of leaf (or root) as the average scaled expression level of the three leaf (or root)-biased genes, and according to the values of the two indices, we partitioned cells into leaf and root cells for the seedling samples (Fig. 4D and Figs. S13).

Applying this method to the seedling data, we estimated the organ origin of each cell for each cell cluster (Figs. S12 and S13). Among cell clusters, the estimated cell population ratio of organ origins (leaf vs. root) was positively correlated with the ratio estimated from the leaf and root samples ($r = 0.76$, $P = 2 \times 10^{-6}$, $N = 29$, Pearson's correlation, $\rho = 0.68$, $P = 5 \times 10^{-5}$, Spearman's correlation; Fig. 4E), bolstering our confidence for implementing scRNA-seq to whole seedlings in the rest of this study.

Abiotic stress-induced transcriptomic responses often act in a cell type–specific manner

After the construction of the single-cell transcriptome atlas of rice seedlings in the Kimura B nutrient solution, we realized that scRNA-seq provides an exciting opportunity to address the long-standing question of potential heterogeneity among cell types in response to abiotic stresses (Dinneny et al., 2008). To this end, we adjusted NH_4^+, NaCl, or Fe^{3+} to various concentrations in a pilot experiment (Fig. 5A), on the basis of the Kimura B nutrient solution. We did not observe apparent variation in the plant architecture except for the high-salinity treatment (Fig. 5A) and therefore decided to perform single-cell RNA-seq experiments in the most extreme conditions used in the pilot experiment: The two-week-old rice seedlings were treated with an NH_4^+-free medium for five days ("low nitrogen"), with 200 mM additional NaCl for three days ("high salinity"), or with an Fe^{3+}-free medium for five days ("iron deficiency"). We isolated protoplasts from the mixed proximal shoots and root tips from 16 rice seedlings for each biological replicate, and four biological replicates were performed for each condition. We obtained transcriptomes for a total of 159,262 cells under these three treatments (Table S1).

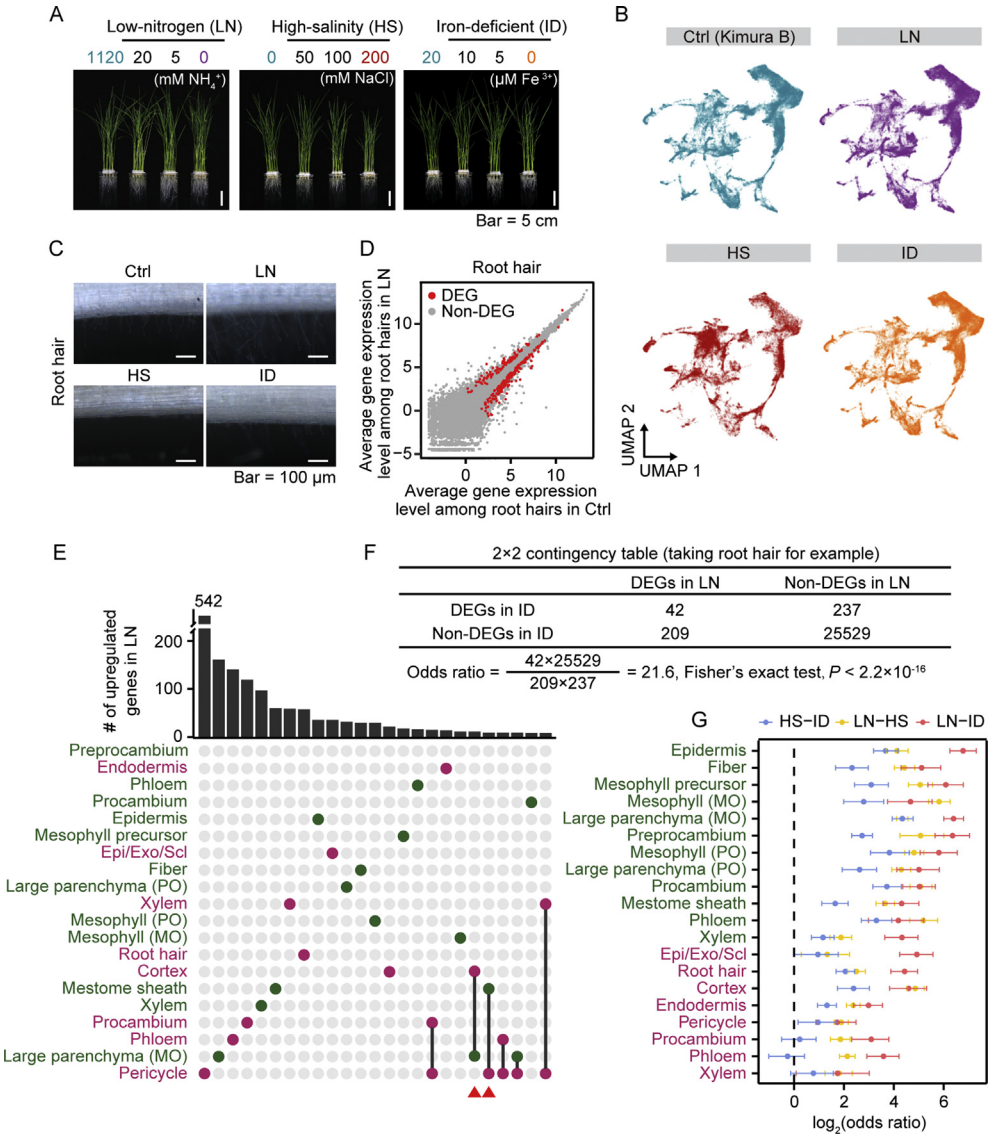

Fig. 5. Cell type–specific gene expression changes on abiotic stresses. **A**: The two-week-old rice seedlings on gradient abiotic stresses. [NH_4^+], [NaCl], and [Fe^{3+}] were individually adjusted to the concentration labeled on the top, on the basis of the Kimura B solution, in which the respective concentrations of these chemicals are shown in cyan. The conditions used for scRNA-seq experiments are labeled in color. Scale bars represent 5 cm. **B**: UMAP visualization of seedling samples in four conditions: control (Ctrl, i.e., the four seedling samples), low nitrogen (LN), high salinity (HS), and iron deficient (ID). Dots represent individual cells, colored according to the growing conditions. For each condition, the scRNA-seq data from four biologically replicates are combined. **C**: The morphology of the root hairs on abiotic stresses. Scale bars represent 100 μm. **D**: Scatter plot shows genes that are differentially expressed on the low-nitrogen treatment (red dots) in root hair cells. The scatter plots of DEGs in other cell types and on other treatments are shown in Fig. S14. Note that initial cells were not included in this analysis because of their relatively small numbers of cells, thereby limited statistical power for the detection of DEGs. **E**: The UpSet plot shows the intersection sizes of upregulated genes among cell types in response to the low-nitrogen stress. To be specific, the height of the bar on the top of each column reflects the number (intersection size) of upregulated genes shared by the cell type(s) filled in the matrix. Only the columns with the intersection size ranked top 25 are shown. Cell types of roots and leaves are labeled in magenta and green, respectively. The red triangles on the bottom indicate the columns representing the intersection between the counterpart cell types between leaves and roots. The UpSet plot was generated using the "upset" function in the UpSetR package of *R*. The UpSet plots for downregulated DEGs and other conditions are shown in Fig. S16. **F** and **G**: The independence of DEGs on two abiotic stresses for each cell type. The root hair cells were used as an example for the 2 × 2 contingency table (**F**). The odds ratios and the corresponding *P* values were estimated from Fisher's exact tests. The error bars in (**G**) represent the 95% confident intervals of the corresponding odds ratios.

All major cell types defined in the standard Kimura B nutrient solution were also detected when treated with the three abiotic stresses (Fig. 5B), suggesting that these abiotic stresses did not cause fundamental variation in cell type transcriptomic features. Nevertheless, morphologic changes in some cell types were observed under the microscope. For example, root hairs were significantly elongated when rice seedlings were treated with low nitrogen (Fig. 5C), indicating the presence of gene expression alteration in some cells. To test if such alternation acts in a cell type−specific manner, we lumped together all cells of the same type from the four replicates of each condition and identified DEGs from these single-cell transcriptomes (Tables S10–S12). Specifically, DEGs were defined as genes with a >1.5-fold difference in the average expression level between two conditions, with the corresponding *P* value in the Mann-Whitney *U* test <0.01, and being expressed in at least 10% of cells in the condition showing the higher average expression level. DEGs identified in the root hair upon the low-nitrogen treatment are shown in Fig. 5D as an example, and those in the other cell types are shown in Fig. S14. Two cases of differential gene expression were further validated using *in situ* hybridization (Fig. S15).

To determine the level of transcriptional heterogeneity among cell types in response to abiotic stress, we drew two UpSet plots for each stress treatment—one for upregulated and the other for downregulated genes—to illustrate the number of shared stress-responding genes among cell types. We observed that the majority of transcriptional responses took place in individual cell types (Fig. 5E and Figs. S16). For example, among the 3383 genes whose expression was upregulated upon the low-nitrogen stress in at least one cell type, 76% (2563 genes) were upregulated in only one cell type (Fig. 5E). These data were consistent with the previous report in *Arabidopsis* roots that the expression of 49% of DEGs was only altered in a single cell type on sucrose removal (Shulse et al., 2019). These observations together indicate that expression responses on abiotic stress likely operate in a cell type−specific manner.

Nevertheless, shared transcriptional responses across cell types did exist. To identify common transcriptional responses among cell types, we focused on the intersection sets of genes in the UpSet plots. We found that some common responses occurred in the counterpart cells between leaves and roots identified in Fig. 4A and 4B, for example, between the large parenchyma in the leaf and the cortex in the root and between the mestome sheath in the leaf and the pericycle in the root (marked with red triangles in Fig. 5E). These observations indicate that cells of the same tissue layer between leaves and roots exhibit not only similar transcriptome features but also common transcriptional regulation in response to environmental stress.

Transcriptional responses are shared among various abiotic stresses

We next sought to investigate if cells use the same transcriptional regulation program to respond to various abiotic stresses. To this end, we determined if the identities of responding genes are independent between two stresses using Fisher's exact tests. Root hair was used to exemplify the computational procedure in Fig. 5F: We counted the number of DEGs in response to the low-nitrogen and iron-deficiency stresses and their intersection, filled up a 2×2 table, and calculated the odds ratio, the *P* value, and the confidence intervals. A significantly >1 ($P < 0.05$) odds ratio indicates that the same set of genes tend to be used to respond to various stresses, whereas a significantly < 1 ($P < 0.05$) odds ratio indicates that instead, different genes tend to be used.

As shown in Fig. 5F, the expression levels of 251 (or 279) genes were altered in root hairs in response to low nitrogen (or iron deficiency). If the responses have been independent between the two treatments, we expect two genes to be shared. The observation of 42 shared genes means a ~20-fold enrichment, indicating that DEGs in response to low nitrogen also tend to play a role in iron deficiency. This observation was not a special case, as globally we observed that the odds ratios were mostly significantly >1 in various cell types and between any two of the three conditions examined in this study (Fig. 5G). Looking closer, we noticed that the significant ($P < 0.01$, Fisher's exact test) enrichment in the gene ontology term "response to stress" (GO:0006950) was not shared across conditions in the majority of the cell types (except in the large parenchyma and procambium); in sharp contrast, "translation" (GO:0006412) or "generation of precursor metabolites and energy" (GO:0006091) was frequently shared across conditions, indicating that the stress-induced changes in the fundamental cellular processes are likely accountable for the common transcriptional responses across various abiotic stresses. Collectively, these observations imply that rice cells tend to respond to various abiotic stresses using a common transcriptional regulation strategy.

The proportion of cell populations changes in response to environmental stresses

Visually we noticed that rice leaves were pale and yellow (i.e., chlorosis) when treated with high salinity, a phenomenon that was not apparent for the seedlings treated with iron deficiency or low nitrogen in our experiments (Fig. 6A). The green color of leaves is known to be caused by chlorophyll, which exists in the chloroplast of mesophyll cells. We therefore speculate that there are two mutually nonexclusive explanations for the observed chlorosis: reduced proportion of mesophyll cells and reduced abundance of chlorophyll molecules per mesophyll cell. We next tested these explanations individually using the scRNA-seq data.

To test if the cell population has altered in response to stress, we estimated the cell proportion for each cell type in each sample and compared it between rice seedlings growing in Kimura B nutrient solution (as a control) and exposed to the three abiotic stresses (Fig. S17). The population of mesophyll cells of MO substantially decreased upon the high-salinity treatment ($P = 0.03$, Mann-Whitney *U* test; Fig. 6B) but remained largely unchanged in the iron deficiency and low nitrogen conditions ($P = 0.89$ and 0.89; respectively, Fig. 6B), indicating that the chlorosis observed on the high-salinity treatment is at least partly attributable to the proportional reduction in the mesophyll cell population.

To test if chlorophyll abundance per mesophyll cell has also decreased on the high-salinity treatment, we retrieved from the single-cell transcriptomes the expression level of *CAB2R* (*LOC_Os01g41710*), which encodes a major chlorophyll *a*/b-binding protein that functions in the light-harvesting complex in chloroplasts and can serve as a marker gene for mesophyll (Tsutsumi et al., 2006). *CAB2R* exhibited a >16-fold reduction in messenger RNA abundance on the high-salinity treatment (Fig. 6C), but its expression level remained largely unchanged when treated with iron deficiency or low nitrogen (fold change = 1.48 or 0.90, respectively). Similar patterns were also observed for other genes encoding the subunits of the light-harvesting complex or the RuBisCo complex (Fig. S18), suggesting that the reduced chlorophyll abundance per mesophyll cell likely also accounts for the observed chlorosis on the high-salinity treatment.

So far, we have shown using the scRNA-seq data a concordant reduction, on the high-salinity treatment, in both the proportion of mesophyll cell population and the abundance of chlorophyll-binding proteins per mesophyll cell. Seeking to validate these observations with a second experimental approach, we analyzed mesophyll protoplasts isolated from rice leaves using flow cytometry, based on the red autofluorescence emitted from their chlorophyll (Fig. 6D). The

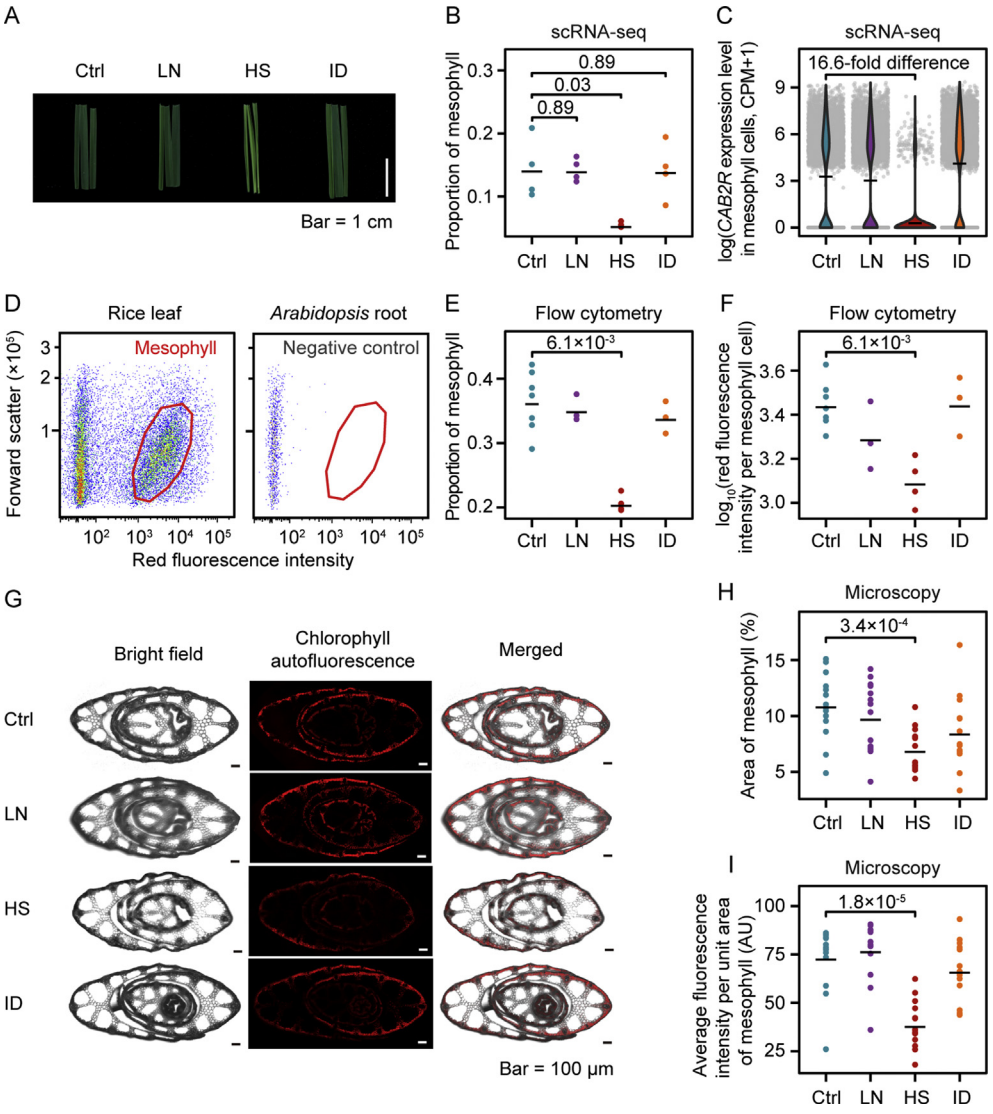

Fig. 6. Changes in mesophyll cell population and chlorophyll abundance on abiotic stresses. **A**: The color of the rice leaves on abiotic stresses. The scale bar represents 1 cm. **B**: The change in mesophyll cell population on abiotic stresses, detected by scRNA-seq. The proportion of mesophyll cells among all leaf cells is shown in dots for each scRNA-seq sample, colored according to the respective conditions. The short horizontal line indicates the mean expression level of four replicates. *P* values are given by the Mann-Whitney *U* tests. The changes in other cell populations are shown in Fig. S17. **C**: The violin plots show the expression of *CAB2R* in mesophyll cells in the four conditions, detected by scRNA-seq. Each dot represents an individual cell (four replicates were combined), and the short horizontal line indicates the mean expression level among all cells. The differential expression is significant on all three conditions ($P < 2.2 \times 10^{-16}$, Mann-Whitney *U* tests). **D**: The gate (the same red polygon in both plots) used for the flow cytometry analysis to identify mesophyll cells. In the density plots, the forward scatters of the detected events are plotted against the red fluorescence intensities. Red (or blue) dots represent a relatively high (or low) density of events within the cell population. The protoplasts isolated from *Arabidopsis* roots are used as the negative control. **E**: Similar to panel **B**, the change in mesophyll cell population on abiotic stresses, detected by flow cytometry. **F**: The median red fluorescence intensity among mesophyll cells identified by flow cytometry in each replicate. *P* values were given by the Mann-Whitney *U* tests (only <0.01 *P* values are labeled). The short horizontal line indicates the mean expression level among all replicates of each condition. The numbers of replicates in panels (**E**) and (**F**) are 7, 3, 4, and 3 for the four conditions. **G**: Microscopic images of the cross-section of rice seedlings grown in four environmental conditions. **H**: Similar to Panel (**E**), the fraction of the cell area showing chlorophyll autofluorescence within the total cell area defined in the bright field. **I**: Similar to Panel (**F**), the average fluorescence intensity per unit mesophyll cell area. The numbers of replicates in panels (**H**) and (**I**) are 15, 14, 15, and 13 for the four conditions, and *P* values were given by the Mann-Whitney *U* tests (only <0.01 *P* values are labeled).

flow cytometry data confirmed that both mesophyll cell population and chlorophyll abundance per mesophyll cell reduced in response to the high-salinity stress (Fig. 6E and F), and such response can be stimulated by much lower salt concentrations, as evident by a series of salt gradient treatments (Fig. S19).

It is worth noting that protoplasting efficiency may change on abiotic stress due, for example, to variable structure or chemical composition of the mesophyll cell wall; such variation, when occurs in a cell type–specific manner, could have confounded the results about cell populations based on scRNA-seq or flow cytometry. Considering this possibility, we further tested if the reduced mesophyll population on the high-salinity treatment can be detected by protoplasting-free approaches. Using microscopy (Fig. 6G), we observed both a reduction in the area of red autofluorescence (Fig. 6H) and a reduced red autofluorescence intensity per unit area exhibiting red autofluorescence (Fig. 6I), confirming the reduced mesophyll cell population and chlorophyll abundance per mesophyll cell in response to the high-salinity stress.

The transcriptional basis underlying the delay in development toward mesophyll cells on the high-salinity treatment

To understand the transcriptional basis underlying the high salinity—induced chlorosis, we ordered single cells according to their transcriptomes and reconstructed the time course expression profiles along with the developmental maturation of mesophyll cells (Trapnell et al., 2014). We focused on the mesophyll cells of MO because their cell population showed the greatest change on the high-salinity treatment (Fig. S17). We reconstructed the cell differentiation trajectory containing mesophyll initial cells, mesophyll precursors, and the mesophyll and large parenchyma of MO. The trajectory reveals that after mesophyll initial cells are differentiated into the mesophyll precursor, two major differentiation trajectories exist, toward mature mesophyll or toward large parenchyma (Fig. S20A and S20B). This observation is in accordance with the developmental trajectories reported in a previous study (Zeng et al., 2016).

We identified 7766 genes exhibiting significant expression variation, as mesophyll initial cells were differentiated toward mesophyll cells. These genes fell into three distinct gene classes and depicted waves of gene expression in a well-organized temporal order over the course of development (Fig. S20C). Genes expressed early in the pseudotime axis were associated with translation, related to the fast-proliferation nature of initial cells, whereas genes expressed late in the pseudotime axis were associated with photosynthesis (Fig. S20D), confirming that the reconstructed cell differentiation trajectory can be used to infer the temporal order of biological processes.

Upon the high-salinity treatment, the time course expression profiles along the developmental maturation of mesophyll cells were substantially disrupted (Fig. S20E). In particular, genes specifically expressed in the early, intermediate, and late stages in the typical developmental trajectory toward mesophyll were all suppressed upon the high-salinity treatment (Fig. S20F), potentially leading to the postponed proliferation, differentiation, and maturation toward mesophyll cells, respectively, and accounting for the reduced cell population of mesophyll and the chlorophyll abundance per mesophyll cell, in response to the high-salinity stress.

Discussion

In this study, we applied scRNA-seq to rice seedling samples in the baseline hydroponic growing condition (Kimura B solution) or under three abiotic stresses. We obtained single-cell transcriptomes for a total of 237,431 rice cells, identified major cell types using a hybrid approach combining both computational and experimental analyses (Figs. 2 and 3), and recognized transcriptome-based counterpart cells between leaves and roots (Fig. 4). We unveiled three basic principles for transcriptional responses of plant cells to abiotic stresses: The response is often cell type specific, is shared to some extent upon various abiotic stresses, and is sometimes common for the counterpart cells between leaves and roots (Fig. 5). Besides, the proportional change in cell populations may also participate in response to abiotic stresses through varied rates of cell proliferation and differentiation (Fig. 6 and Figs. S20). Although the molecular and developmental mechanisms should be further investigated for individual genes and cell types in the future, our study showcases the power of single-cell transcriptome analyses in understanding the global heterogeneity among plant cells.

We noted that some caveats existed in our single-cell transcriptome analyses. First, we discarded genes whose expression levels were altered during protoplasting for dimensionality reduction and cell clustering to avoid computational artifacts associated with protoplast isolation. However, these genes may also be differentially expressed among cells and therefore could have been informative (Fig. S21A). Nevertheless, we had confirmed that the removal of protoplasting-sensitive genes from data normalization and integration did not fundamentally influence the results of dimensionality reduction and cell clustering (Fig. S21B). While the impact of such a computational procedure remains to be thoroughly investigated, it can be experimentally alleviated with the development of more gentle protocols for protoplasting and thereby reducing the number of genes deemed to be excluded in data analysis.

Second, we developed in this study a computational strategy, the weighted MICI (Fig. S6), to assist in assigning the cell type to each cell cluster, but the validity of the strategy has yet to be independently confirmed. Note that although not highlighted in the Results section, we did compare our cell type assignment with two recently published atlases of rice roots (Liu et al., 2021; Zhang et al., 2021a). Specifically, we computationally integrated the 8538 single-cell transcriptomes in the two root samples with those obtained previously for the 10,968 root cells isolated from Nipponbare (Liu et al., 2021) and the 31,204 root cells isolated from ZH11 (Zhang et al., 2021a).

Visually, we noticed in Fig. S22A that the cells detected in our study largely overlaid with Liu et al. but was substantially different from a study by Zhang et al., possibly because of the differences in experimental materials and protoplasting protocols, as well as the conversion of gene identity across different annotations of the rice genome (Zhang et al., 2021a). Therefore, we put our focus on the comparison with Liu et al., who assigned a total of eight root cell types: metaxylem, root cap, root hair, epidermis (near root hair), epidermis, endodermis, stele, and cortex. Among them, six broadly correspond to the cell type terms used in our study (Fig. S22B). The expression of the marker genes reported by Liu et al. for these six cell types remained confined in our root cells of the corresponding UMAP locations (Fig. S22B), indicating a reliable data integration between the two studies.

To test if the cell type assignment was consistent between the two studies, we labeled for these cell types individually on the UMAP graph (Fig. S22C); the corresponding cell types between the two studies greatly overlaid (Fig. S22C). Consistently, when the transcriptomes of individual cells of the same type were combined, the average transcriptomes of the corresponding cell types between a study by Liu et al. and our study were highly correlated (Fig. S22D).

Four cell types identified in our study were not reported by Liu et al.: the initial cell, pericycle, phloem, and procambium. Three of them—the initial cell (under the name of meristem), pericycle, and phloem—were reported by Zhang et al. Because Zhang et al. did not provide cell type annotation for each cell, we labeled the expression of the marker genes reported by Zhang et al. for these three cell types on the UMAP graph (Fig. S23A). When our cells of these three cell types were labeled onto the UMAP graph, two out of the three cell types were substantially overlaid (the initial cell and phloem, Fig. S23B). Collectively, these observations verified the reliability of the cell type assignment in our study.

Third, it is worth noting that some cell types, such as the root cap, were not assigned to any cell clusters in our scRNA-seq data (Fig. 3). There are two potential explanations: (1) the cells of root caps were not captured in the experiments, possibly attributable to the tendency of the root cap to be disassociated from root tips (Kawata et al., 1979), or (2) these cells were experimentally captured, but we failed to identify them computationally. We noticed that a recent scRNA-seq analysis of rice roots had assigned the root cap to (31/10,968=) 0.28% cells in Nipponbare (Liu et al., 2021) and identified five marker genes for the root cap cells. To explore the reason for the absence of root cap in our data, we labeled the expression levels of these five marker genes individually on the UMAP graph. Some cells expressing these marker genes were detected, but none of the five marker genes were specifically expressed in any cell cluster (Fig. S11F), suggesting that the sporadic root cap cells captured in

our experiments did not show sufficient common transcriptome features to be computationally clustered.

In our study, the guard cell was not assigned to any cell clusters either (Fig. 2). To investigate the underlying causes, we labeled on the UMAP graph the expression of *SLAC1* (*LOC_Os04g48530*, Fig. S7), which encodes a slow anion channel expressed predominantly in the guard cell (Negi et al., 2008; Vahisalu et al., 2008). Cells expressing *SLAC1* were located within Cluster 9, which has been assigned to the epidermis, echoing the notion that the guard cell is a subtype of epidermal cells (Liu et al., 2020). Guard cells were not assigned to any cell cluster likely because the proportion of its population was relatively small, and their expression was not sufficiently specialized to form a unique cell cluster. Nevertheless, they can be computationally identified within Cluster 9 using the expression of its marker gene *SLAC1*. With the fast accumulating of additional marker genes, enabled by the broader application of scRNA-seq to plants, the curation of cell type or subtype will become more accurate.

The fourth caveat is that the detection of cell type−specific differential expression highly relies on the computational accuracy in cell clustering and cell type assignment both in the Kimura B solution and upon abiotic stress treatments. Although not highlighted in our results, note that we did conduct additional experiments to confirm the genome-wide pattern of differential gene expression detected by the scRNA-seq analysis. In particular, we performed bulk RNA-seq on the mesophyll cells sorted by flow cytometer (Fig. S24A), both in the Kimura B solution and upon the high-salinity treatment. The upregulation and downregulation of genes detected in the single-cell transcriptomes were also observed using bulk RNA-seq on the sorted mesophyll cells (Fig. S24B), confirming that our scRNA-seq has successfully gauged the genome-wide pattern of differential gene expression in a cell type−specific manner.

Last but not least, we have to stress that the detection of proportional change in cell populations based on the scRNA-seq data can be confounded by numerous factors. For example, we reported the reduction in mesophyll cell population upon the high-salinity treatment using both scRNA-seq (Fig. 6B) and flow cytometry (Fig. 6E). However, technically an alternative explanation is that in response to stress, mesophyll cells become harder to be isolated from the tissue or less likely to survive the protoplasting process. Although our results based on fluorescence microscopy do not rely on protoplast isolation (Fig. 6G and H), they are plausibly confounded by the reduced chlorophyll abundance per mesophyll cell (Fig. 6F and I), thereby reducing the sensitivity in the autofluorescence-based detection of mesophyll cells. Note that albeit not highlighted in the Results section, we also stained mesophyll and large parenchyma to blue and red, respectively, using the fast green/safranin O dyes (Abdellaoui et al., 2017); the reduction in the mesophyll cell population upon the high-salinity treatment was still observed ($P = 0.02$, Mann-Whitney *U* test, Fig. S26). Generally speaking, additional histological analyses would always be informative for drawing conclusions about the proportional change in cell populations based on the scRNA-seq data.

In this study, we observed that leaf and root often shared similar transcriptomic features within the same tissue layer, except for endodermal and epidermal cells (Fig. 4). Consistent with our observations, the transcriptomic similarity in vascular cells and dissimilarity in epidermal cells between the leaf and root samples were also recently reported in *Arabidopsis* (Zhang et al., 2021b). Yet, we did observe that leaf epidermal cells and root hairs coexisted in some cell clusters, such as Clusters 9 and 27 (Figs. 2 and 3). Nevertheless, when we labeled cells from the leaf and root samples on the same two-dimensional UMAP space, the leaf and root cells showed apparent spatial partitioning within Cluster 9 or 27 (Fig. S25), confirming the transcriptomic dissimilarity between the leaf and root epidermal cells (Fig. 4A and 4B).

Looking forward to the future, scRNA-seq analyses, together with single-cell epigenetic approaches (Buenrostro et al., 2015; Huan et al., 2018; Lai et al., 2018; Hainer et al., 2019) and the developing methods based on *in situ* reverse transcription that can keep the spatial information (Giacomello et al., 2017), will synergistically contribute to the characterization of the spatiotemporal transcriptome atlas of plants. Furthermore, the procedure for nucleus isolation is likely more uniform across tissues and treatments (Tian et al., 2020; Long et al., 2021), so RNA-seq or single-nucleus assays for transposase-accessible chromatin using sequencing could also exert a strong influence. These genome-wide approaches will not only promote our understanding of plant cell and developmental biology at the cellular resolution but also help breed more robust crops against environmental stresses (Rhee et al., 2019).

Materials and methods

Plant materials and growth conditions

Seeds of *Oryza sativa* ssp. *japonica* variety Nipponbare were germinated on wet sterile filter paper. Seedlings were grown in a growth chamber at 30°C (day)/28°C (night) under a 12-h photoperiod. Rice seedlings were grown in Kimura B nutrient solution for 14 days, which was the default condition in this study if not otherwise specified. The Kimura B solution consists of 0.56 mM $(NH_4)_2SO_4$, 0.18 mM KH_2PO_4, 0.27 mM K_2SO_4, 0.24 mM $Ca(NO_3)_2$, 0.55 mM $MgSO_4$, 0.02 mM Fe(II)-EDTA, 0.70 mM Na_2SiO_3, 9.14 μM $MnCl_2$, 46.9 μM H_3BO_3, 0.32 μM $CuSO_4$, 0.76 μM $ZnSO_4$, and 0.6 μM Na_2MoO_4, with the pH adjusted to 5.8. On the background of Kimura B solution, high-salinity (additional 200 mM NaCl), low-nitrogen ($(NH_4)_2SO_4$ dropped out), and iron-deficiency (Fe(II)-EDTA dropped out) solutions were prepared individually; rice seedlings were treated with these solutions for three days, five days, or five days, respectively, right before tissue collection on the 14th day. Fresh nutrient solutions were replaced every three days during the experiment.

As a spike-in control, seeds of *Arabidopsis thaliana* ecotype Columbia (Col-0) were surface sterilized in 75% ethanol for 1 min, followed by 10% bleach for 15 min and washing with sterile water until clean. Seeds were sown on 1 × Murashige and Skoog (MS) + 1% Sucrose (w/v) plates and treated at 4°C for 3 days. Seeds were further grown at 23°C (day)/23°C (night) under a 16-h photoperiod.

Protoplasting

Rice and *Arabidopsis* protoplasts were isolated using previously described protocols (Evrard et al., 2012; Jabnoune et al., 2015) with modifications. Briefly, proximal shoots (~8 cm in length above the root-shoot junction) of rice seedlings were finely chopped into 0.5−1 mm strips with freshly sharpened blades. The root tips (~2 cm in length) of rice seedlings or the whole root of *Arabidopsis* seedlings were chopped into small pieces in enzyme solution with flamed scissors without shredding. Plant materials were transferred immediately to 5 mL freshly prepared enzyme solution, including 1.5% (w/v) cellulase RS (Yakult), 0.75% (w/v) macerozyme R-10 (Yakult), 0.6 M mannitol, 1 mM $CaCl_2$, 10 mM MES (pH = 5.5), 5 mM 2-mercaptoethanol, and 0.1% BSA. The strips were then placed in a vacuum pump (at 55 kPa) for 10 min at room temperature, followed by incubation in the enzyme solution on a shaker set at 50 rpm. After 2-h gentle shaking, the digestion mixture was sequentially filtered through a 70 μm (FALCON, Cat # 352,350) and a 40 μm (FALCON, Cat # 352,340) nylon mesh sieve. Protoplasts were collected after soft washing twice (centrifugation at 100*g* for 3 min with a swinging bucket rotor) and were resuspended in 8% mannitol solution. The numbers of isolated protoplasts were measured with hemocytometer

counting chambers, and the viability of these protoplasts was assessed by 0.4% trypan blue staining. Ultimately, ~300,000 protoplasts were isolated for each sample. The concentration of protoplasts was adjusted to 1500−2000 cells/μL.

scRNA-seq

Construction and sequencing of the single-cell RNA-seq libraries

We constructed scRNA-seq libraries from isolated protoplasts, following the manufacturer's protocol of the 10x Genomics Chromium Single Cell 3′ Reagent Kits v3 (CG000183, RevC). The numbers of protoplasts loaded onto the 10x Genomics Chromium single-cell microfluidics device were ~6000, ~10,000, and ~20,000 for each replicate of the rice and *Arabidopsis* mixed samples, the rice leaf/root samples, and the rice seedling samples, respectively. Polymerase chain reaction (PCR) was performed for 11 cycles to amplify the pooled complement DNA. After ligating to the sequencing adapters, complement DNA was again amplified by PCR for 10−15 cycles, according to the DNA concentration of the template. scRNA-seq libraries were sequenced on the Illumina NovaSeq 6000 platform in the paired-end 150-nt mode. Services of library construction and sequencing were provided by Berry Genomics.

Sequencing read alignment and generation of single-cell gene expression matrices

We aligned scRNA-seq data and generated single-cell gene expression matrices using Cell Ranger (version 3.1.0; 10× Genomics) (Zheng et al., 2017). Sequencing reads of rice scRNA-seq samples were aligned to the rice reference genome (release 7 of the MSU Rice Genome Annotation Project, RGAP7) (Kawahara et al., 2013) with the embedded STAR (version 2.6.0a) (Dobin et al., 2013). Reads that share the same cell barcode, gene identity, and unique molecular identifier (UMI, one mismatch allowed) were combined into a UMI count in an effort to remove PCR duplicates during sequencing library preparation. Based on the distribution of the total UMI count within each GEM droplet among all cell barcodes, the valid cell barcodes were determined by Cell Ranger. Briefly, the cell barcodes with the total UMI count >10% of the total UMI count of the 99th percentile of the expected number of recovered cells were considered representing genuine cells. The expected cell number was estimated according to the 10x Genomics manufacturer's protocol (approximately 62.5% of the loaded cell number). Finally, a gene expression matrix for the genuine cells of each scRNA-seq sample was generated, of which each row represented a gene, and each column represented a valid cell barcode.

Sequencing reads of the rice and *Arabidopsis* mixed samples were aligned to the concatenated reference of RGAP7 and the *Arabidopsis* information resource 10 genome release (TAIR10) (Lamesch et al., 2012). To accurately classify the species origin of each transcript, we kept only uniquely mapped and 100% matched reads for the following analyses.

Normalization and integration for scRNA-seq data

After removing protoplasting-sensitive genes, we normalized and integrated gene expression matrices across scRNA-seq samples using the Seurat *R* package (version 3.2.3). Briefly, we normalized gene expression levels using function "SCTransform" to reduce the influence from variation in the cellular sequencing depth (Hafemeister and Satija, 2019). We selected 3000 most highly variable genes across samples using the function "SelectIntegrationFeatures." Finally, we employed the reference-based Seurat integration workflow to integrate multiple samples (Stuart et al., 2019), setting samples "one-week-old root" from Shulse et al. (2019) and seedling 1 as the references for *Arabidopsis* and rice, respectively.

To compare the single-cell transcriptome data for rice roots reported in our study and two recent studies (Liu et al., 2021; Zhang et al., 2021a), we downloaded the scRNA-seq data of the two studies from Gene Expression Omnibus under the accession number GSE146035 and Single Cell Expression Atlas under the accession number E-ENAD-52, respectively. To make the data from three sources directly comparable, we conversed the gene annotation of the rice genome in a study by Zhang et al. (2021) from Os-Nipponbare-Reference-IRGSP-1.0 to RGAP7, as used in our study and in a study by Liu et al. (2021). We normalized and integrated the gene expression matrices across three scRNA-seq data sets using the same parameters as described earlier, setting root 2 in our study as the reference.

Dimensionality reduction and cell clustering

We performed linear dimensionality reduction (Chen et al., 2019) of the integrated scRNA-seq data using function "RunPCA" in the Seurat *R* package. We chose the top 100 principal components and performed nonlinear dimensionality reduction for visualization, with function "UMAP" in the UMAP *R* package (version 0.2.7.0; under the parameters n_neighbors = 10, metric = "pearson2", and min_dist = 0.01) (McInnes et al., 2018). We further clustered cells according to the top 100 principal components based on the Louvain algorithm, using functions "FindNeighbors" and "FindClusters" (resolution = 0.75) in the Seurat *R* package.

Identification of cluster-specific genes

We identified the cluster-specific genes for leaf samples (two replicates combined) by function "FindAllMarkers" (min.pct = 0.25, logfc.threshold = 0.25) in the Seurat *R* package. The cluster-specific genes for root samples were similarly identified.

Assignment cell types to clusters by MICI

We compiled a list of cell type marker genes reported in previous studies (Table S6) and combined single cells from two leaf samples or two root samples. We developed a weighted MICI algorithm for cell type determination for each cell (Fig. S6). Specifically, we first estimated the weight for each reported marker gene from the variance among cell clusters in the logarithm-transformed, cluster-averaged expression level (in the unit of CPM) of each marker gene. All 29 clusters were used for the leaf, and 27 clusters were used for the root (excluding Clusters 23 and 24, whose cells were entirely assigned to the leaf). Then, the expression level of each marker gene was logarithm transformed and scaled among all cells detected in the leaf samples or root samples. For each cell, we calculated a MICI value for each cell type as the sum of the weighted expression levels of all marker genes for the cell type. The cell type of the highest MICI value was assigned to the cell being queried if the value was greater than 2; otherwise, the cell type was "undermined." Finally, the cell type (including "undetermined") of the largest cell population was assigned to the cell cluster being queried.

Identification of leaf and root cells in the whole seedling samples

We identified leaf and root cells in the whole seedling samples based on the expression of leaf-biased and root-biased genes, which were predetermined from the single-cell transcriptomes of leaf and root samples. To identify leaf-biased and root-biased genes for each cell cluster, we defined a leaf-biased expression score for each gene as the product of an expression-level score and a cell fraction score. The expression-level score of a gene was defined as the ratio of its average expression level among all cells in the cluster between Leaf-1 and Root-1 samples (or between Leaf-2 and Root-2). The

expression level was in the unit of CPM, adding one to all the values for the sake of logarithm transformation.

The cell fraction score of a gene was defined for each cell cluster as the ratio between the fraction of leaf cells that express the gene and the fraction root cells that express the gene; we arbitrarily assigned cell fraction <10% as 10% to avoid large faction scores for genes that are only expressed in a tiny fraction of cells but exhibit a great fold difference between leaf and root samples (e.g., 1% in an organ and 0.01% in another). We estimated the leaf-biased expression scores for both biological replicates, and the total logarithm-transformed score was used for selecting biased expression genes. For each cell cluster, three genes with the highest leaf-biased scores were defined as leaf-biased genes, and three genes with the lowest leaf-biased scores were defined as root-biased genes (Table S9). It is worth noting that in an effort to use the same set of genes to partition leaf and root cells for all four environmental conditions, we discarded genes with the average expression level deviating by twofold on any of the three treatments.

For each cell cluster in the whole seedling samples, we scaled expression levels for each predefined leaf-biased and root-biased gene. For each cell, we estimated the average scaled expression level of the three leaf-biased (root-biased) genes as the index of leaf (root) for the cell. Cells in the whole seedling samples with a greater index of leaf (or root) were classified as leaf (or root) cells. If these two indices are equal in a cell, the organ origin of the cell was not determined. The detailed annotation (organ and cell type) for each cell in each sample is provided in Table S13.

Identification of cell type-specific stress-responding genes

We combined cells of the same type from four scRNA-seq replicates of an environmental condition. We defined stress-responding genes as DEGs by function "FindMarkers" (min.pct = 0, logfc.threshold = 0) with >1.5-fold difference in the average expression level (in the unit of CPM) upon a stress treatment and $P < 0.01$ in the Mann-Whitney *U* test among individual cells. The intersection size of DEGs across cell types was visualized using the function "upset" in the UpSetR *R* package (version 1.4.0) (Lex et al., 2014).

Reconstruction of developmental trajectories from scRNA-seq data

The developmental trajectories from mesophyll initial cells (Cluster 0, from all four conditions, same below) to mesophyll precursor (Cluster 1, 17−19), mesophyll (Cluster 21, 22), and large parenchymal cells (Cluster 20) were reconstructed using the Monocle *R* package (version 2.14.0) (Trapnell et al., 2014). Note that for mesophyll and large parenchymal cells, only those of MO were included here, as cells of PO by definition should not be placed on this trajectory. We performed dimensionality reduction by function "reduceDimension" (reduction_method = "DDRTree") and ordered cells along the trajectory by function "orderCells". To analyze the dynamic gene expression pattern along the pseudotime, we selected 7766 significantly variable genes (false discovery rate < 0.01) in the four control seedling samples by function "differentialGeneTest".

The gene ontology enrichment analyses

The gene ontology enrichment analyses were performed for these genes on the Web site AgriGO2 (http://systemsbiology.cau.edu.cn/agriGOv2) (Tian et al., 2017).

Bulk RNA-seq

Total RNA was extracted with TRIzol from fresh plant materials or protoplasts. Strand-specific RNA-seq libraries were constructed, followed by high-throughput sequencing on Illumina NovaSeq 6000 platform in the paired-end 150-nt mode. Sequencing reads of low quality (>50% bases with a Phred score < 20) were filtered. After sequencing adapters being removed by fastp, the clean reads were aligned to the concatenated reference of RGAP7 and TAIR10 with STAR (–outFilterMismatchNmax 0, –outFilterMultimapNmax 1). Gene expression levels were estimated from the uniquely mapped reads using HTSeq (version 0.11.2) (Anders et al., 2015).

RNA *in situ* hybridization assay

RNA *in situ* hybridization assay was performed as previously described (Umeda et al., 1999), with minor modifications. Briefly, fresh leaves and roots of rice seedlings were fixed in FAA solution (50% ethanol, 5% acetic acid, and 3.7% formaldehyde), dehydrated, and embedded in paraffin (Sigma-Aldrich). The embedded samples were sliced into 7-μm sections, which were then mounted on slides. We selected target genes mainly from the identified cluster-specific genes (Fig. S5; Table S5) but also took the average gene expression level into consideration. The specific sequences of these target genes were used as probes (Table S14) and were cloned into the pSPT18 vector, which was linearized afterward. Probes were *in vitro* transcribed using Digoxigenin RNA labeling kit (Roche) and then applied on tissue sections for hybridization. After washing, the slides were incubated with an anti-digoxigenin antibody (Roche), the signal of which was then detected in the nitro-blue tetrazolium/5-bromo-4-chloro-3′-indolyphosphate stock solution (NBT/BCIP solution; Roche). Images were taken in the bright field mode using Olympus CX23.

Flow cytometry

Protoplasts isolated from leaves of two-week-old rice seedlings were loaded onto the flow cytometer (BD FACS Aria II cell sorter, using a 100 μm nozzle) at a rate of 2000−3000 events per second and fluid pressure of 20 psi. We set up the gate for sorting rice mesophyll cells based on the red fluorescence (695/40 nm) intensity and the forward scatter, and with *Arabidopsis* root protoplasts as the negative control. To prepare the bulk RNA-seq library for mesophyll cells, we sorted mesophyll cells into Qiagen RLT lysis buffer. The red fluorescence intensities of ~60,000 cells were recorded for each sample.

Microscopy

The primary shoots of rice seedlings were fixed with a half-cut styrofoam board and were sliced with a freshly sharpened blade. The emitted red autofluorescence was captured with a Leica confocal microscope (TCS SP5). The total red fluorescence intensity and the mesophyll cell area were estimated with ImageJ under default parameters. Their ratio was used to infer the average red fluorescence intensity per unit mesophyll cell area. The total cell area was estimated from bright-field images. The fraction of cell areas emitting red fluorescence was used to infer the proportion of the mesophyll cell population.

Cross-section tissue slices were also stained with the modified fast green/safranine O staining kit following the manufacturer's protocol (Solarbio, China) and were mounted with neutral balsam immediately. Images were taken with an optical microscope (Olympus CKX41). The areas of mesophyll and large parenchymal cells were measured with ImageJ.

CRediT authorship contribution statement

Yu Wang: Conceptualization, Investigation, Formal analysis, Data curation, Writing – original draft, Writing – review & editing. **Qing Huan**: Conceptualization, Investigation, Formal analysis, Validation, Data curation, Writing – original draft, Writing – review & editing, Funding acquisition. **Ke Li**: Methodology, Investigation. **Wenfeng Qian**: Conceptualization, Writing – original draft, Writing – review & editing, Supervision, Funding acquisition.

Data availability

All sequencing data have been deposited to the Genome Sequence Archive (Wang et al., 2017) in BIG Data Center (https://ngdc.cncb.ac.cn/gsa/), Beijing Institute of Genomics, Chinese Academy of Sciences, under the accession number CRA004082. Codes to analyze the data and generate figures are available at GitHub (https://github.com/Yuwang-art/scRNA-seq_in_rice_seedlings) and Zenodo (https://doi.org/10.5281/zenodo.4916334).

Conflict of interest

The authors declare that they have no competing interests.

Acknowledgments

We thank Dr. Xiao Chu, Dr. Ying Chen, Yan-Ming Chen, Yuliang Zhang, Min Dai, and Dr. Xiaofeng Gu for supports in data analyses, Ting Li and Yanbo Tian for supports in flow cytometry and fluorescence microscopy, Dr. Caihuan Tian for supports in protoplast isolation, and Dr. Taolan Zhao and Dr. Yuling Jiao for reading the manuscript. This work was supported by grants from the National Natural Science Foundation of China (31900229 to Q.H. and 31922014 to W.Q.).

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jgg.2021.06.001.

References

Abdellaoui, R., Boughalleb, F., Chebil, Z., Mahmoudi, M., Belgacem, A.O., 2017. Physiological, anatomical and antioxidant responses to salinity in the Mediterranean pastoral grass plant *Stipa lagascae*. Crop Pasture Sci. 68, 872–884.

Anders, S., Pyl, P.T., Huber, W., 2015. HTSeq–a Python framework to work with high-throughput sequencing data. Bioinformatics 31, 166–169.

Birnbaum, K., Jung, J.W., Wang, J.Y., Lambert, G.M., Hirst, J.A., Galbraith, D.W., Benfey, P.N., 2005. Cell type-specific expression profiling in plants via cell sorting of protoplasts from fluorescent reporter lines. Nat. Methods 2, 615–619.

Birnbaum, K., Shasha, D.E., Wang, J.Y., Jung, J.W., Lambert, G.M., Galbraith, D.W., Benfey, P.N., 2003. A gene expression map of the *Arabidopsis* root. Science 302, 1956–1960.

Blondel, V.D., Guillaume, J.-L., Lambiotte, R., Lefebvre, E., 2008. Fast unfolding of communities in large networks. J. Stat. Mech. Theor. Exp. 2008, P10008.

Buenrostro, J.D., Wu, B., Litzenburger, U.M., Ruff, D., Gonzales, M.L., Snyder, M.P., Chang, H.Y., Greenleaf, W.J., 2015. Single-cell chromatin accessibility reveals principles of regulatory variation. Nature 523, 486–490.

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., Satija, R., 2018. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol. 36, 411–420.

Chen, Y., Li, K., Chu, X., Carey, L.B., Qian, W., 2019. Synchronized replication of genes encoding the same protein complex in fast-proliferating cells. Genome Res. 29, 1929–1938.

Chonan, N., 1978. A comparative anatomy of mesophyll among the leaves of gramineous crops. Jpn. Agric. Res. Q. 12, 128–131.

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. Dev. Cell 48, 840–852 e845.

Dinneny, J.R., Long, T.A., Wang, J.Y., Jung, J.W., Mace, D., Pointer, S., Barron, C., Brady, S.M., Schiefelbein, J., Benfey, P.N., 2008. Cell identity mediates the response of *Arabidopsis* roots to abiotic stress. Science 320, 942–945.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., Gingeras, T.R., 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21.

Efroni, I., Ip, P.L., Nawy, T., Mello, A., Birnbaum, K.D., 2015. Quantification of cell identity from single-cell gene expression profiles. Genome Biol. 16, 9.

Evrard, A., Bargmann, B.O., Birnbaum, K.D., Tester, M., Baumann, U., Johnson, A.A., 2012. Fluorescence-activated cell sorting for analysis of cell type-specific responses to salinity stress in *Arabidopsis* and rice. Methods Mol. Biol. 913, 265–276.

Frank, M.H., Scanlon, M.J., 2015. Cell-specific transcriptomic analyses of three-dimensional shoot development in the moss *Physcomitrella patens*. Plant J. 83, 743–751.

Giacomello, S., Salmen, F., Terebieniec, B.K., Vickovic, S., Navarro, J.F., Alexeyenko, A., Reimegard, J., McKee, L.S., Mannapperuma, C., Bulone, V., et al., 2017. Spatially resolved transcriptome profiling in model plant species. Native Plants 3, 17061.

Gui, J., Shen, J., Li, L., 2011. Functional characterization of evolutionarily divergent 4-coumarate:coenzyme A ligases in rice. Plant Physiol. 157, 574–586.

Hafemeister, C., Satija, R., 2019. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. Genome Biol. 20, 296.

Haghverdi, L., Lun, A.T.L., Morgan, M.D., Marioni, J.C., 2018. Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. Nat. Biotechnol. 36, 421–427.

Hainer, S.J., Boskovic, A., McCannell, K.N., Rando, O.J., Fazzio, T.G., 2019. Profiling of pluripotency factors in single cells and early embryos. Cell 177, 1319–1329 e1311.

Han, Y., Chu, X., Yu, H., Ma, Y.-K., Wang, X.-J., Qian, W., Jiao, Y., 2017. Single-cell transcriptome analysis reveals widespread monoallelic gene expression in individual rice mesophyll cells. Sci. Bull. 62, 1304–1314.

Huan, Q., Zhang, Y., Wu, S., Qian, W., 2018. HeteroMeth: a database of cell-to-cell heterogeneity in DNA methylation. Genom. Proteom. Bioinform. 16, 234–243.

Jabnoune, M., Secco, D., Lecampion, C., Robaglia, C., Shu, Q., Poirier, Y., 2015. An efficient procedure for protoplast isolation from mesophyll cells and nuclear fractionation in rice. Bio-protocol 5.

Jean-Baptiste, K., McFaline-Figueroa, J.L., Alexandre, C.M., Dorrity, M.W., Saunders, L., Bubb, K.L., Trapnell, C., Fields, S., Queitsch, C., Cuperus, J.T., 2019. Dynamics of gene expression in single root cells of *Arabidopsis thaliana*. Plant Cell 31, 993–1011.

Jiao, Y., Meyerowitz, E.M., 2010. Cell-type specific analysis of translating RNAs in developing flowers reveals new levels of control. Mol. Syst. Biol. 6, 419.

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.

Kawahara, Y., de la Bastide, M., Hamilton, J.P., Kanamori, H., McCombie, W.R., Ouyang, S., Schwartz, D.C., Tanaka, T., Wu, J., Zhou, S., et al., 2013. Improvement of the *Oryza sativa* Nipponbare reference genome using next generation sequence and optical map data. Rice 6, 4.

Kawata, S.-i., Suzuki, S., Yamazaki, k., 1979. The detachment of the "primary root caps" in rice plants. Jpn. J. Crop Sci. 48, 303–310.

Lai, B., Gao, W., Cui, K., Xie, W., Tang, Q., Jin, W., Hu, G., Ni, B., Zhao, K., 2018. Principles of nucleosome organization revealed by single-cell micrococcal nuclease sequencing. Nature 562, 281–285.

Lamesch, P., Berardini, T.Z., Li, D., Swarbreck, D., Wilks, C., Sasidharan, R., Muller, R., Dreher, K., Alexander, D.L., Garcia-Hernandez, M., et al., 2012. The *Arabidopsis* Information Resource (TAIR): improved gene annotation and new tools. Nucleic Acids Res. 40, D1202–D1210.

Lex, A., Gehlenborg, N., Strobelt, H., Vuillemot, R., Pfister, H., 2014. UpSet: visualization of intersecting sets. IEEE Trans. Visual. Comput. Graph. 20, 1983–1992.

Lieckfeldt, E., Simon-Rosin, U., Kose, F., Zoeller, D., Schliep, M., Fisahn, J., 2008. Gene expression profiling of single epidermal, basal and trichome cells of *Arabidopsis thaliana*. J. Plant Physiol. 165, 1530–1544.

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.

Liu, H.L., Xu, Y.Y., Xu, Z.H., Chong, K., 2007. A rice *YABBY* gene, *OsYABBY4*, preferentially expresses in developing vascular tissue. Dev. Gene. Evol. 217, 629–637.

Liu, Z., Zhou, Y., Guo, J., Li, J., Tian, Z., Zhu, Z., Wang, J., Wu, R., Zhang, B., Hu, Y., et al., 2020. Global dynamic molecular profiling of stomatal lineage cell development by single-cell RNA sequencing. Mol. Plant 13, 1178–1193.

Long, Y., Liu, Z., Jia, J., Mo, W., Fang, L., Lu, D., Liu, B., Zhang, H., Chen, W., Zhai, J., 2021. FlsnRNA-seq: protoplasting-free full-length single-nucleus RNA profiling in plants. Genome Biol. 22, 66.

McInnes, L., Healy, J., Melville, J., 2018. UMAP: uniform manifold approximation and projection for dimension reduction. arXiv, 180203426.

Mustroph, A., Zanetti, M.E., Jang, C.J., Holtan, H.E., Repetti, P.P., Galbraith, D.W., Girke, T., Bailey-Serres, J., 2009. Profiling translatomes of discrete cell populations resolves altered cellular priorities during hypoxia in *Arabidopsis*. Proc. Natl. Acad. Sci. U. S. A 106, 18843–18848.

Negi, J., Matsuda, O., Nagasawa, T., Oba, Y., Takahashi, H., Kawai-Yamada, M., Uchimiya, H., Hashimoto, M., Iba, K., 2008. CO_2 regulator SLAC1 and its homologues are essential for anion homeostasis in plant cells. Nature 452, 483–486.

Rebouillat, J., Dievart, A., Verdeil, J.L., Escoute, J., Giese, G., Breitler, J.C., Gantet, P., Espeout, S., Guiderdoni, E., Périn, C., 2009. Molecular genetics of rice root development. Rice 2, 15–34.

Rhee, S.Y., Birnbaum, K.D., Ehrhardt, D.W., 2019. Towards building a plant cell atlas. Trends Plant Sci. 24, 303–310.

Rich-Griffin, C., Stechemesser, A., Finch, J., Lucas, E., Ott, S., Schafer, P., 2020. Single-cell transcriptomics: a high-resolution avenue for plant functional genomics. Trends Plant Sci. 25, 186–197.

Ryu, K.H., Huang, L., Kang, H.M., Schiefelbein, J., 2019. Single-cell RNA sequencing resolves molecular relationships among individual plant cells. Plant Physiol. 179, 1444–1456.

Shulse, C.N., Cole, B.J., Ciobanu, D., Lin, J., Yoshinaga, Y., Gouran, M., Turco, G.M., Zhu, Y., O'Malley, R.C., Brady, S.M., et al., 2019. High-throughput single-cell transcriptome profiling of plant cell types. Cell Rep. 27, 2241–2247. e2244.

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck 3rd, W.M., Hao, Y., Stoeckius, M., Smibert, P., Satija, R., 2019. Comprehensive integration of single-cell data. Cell 177, 1888–1902. e1821.

Takehisa, H., Sato, Y., Igarashi, M., Abiko, T., Antonio, B.A., Kamatsuki, K., Minami, H., Namiki, N., Inukai, Y., Nakazono, M., et al., 2012. Genome-wide transcriptome dissection of the rice root system: implications for developmental and physiological functions. Plant J. 69, 126–140.

Tian, C., Du, Q., Xu, M., Du, F., Jiao, Y., 2020. Single-nucleus RNA-seq resolves spatiotemporal developmental trajectories in the tomato shoot apex. bioRxiv, 2020.2009.2020.305029.

Tian, T., Liu, Y., Yan, H., You, Q., Yi, X., Du, Z., Xu, W., Su, Z., 2017. agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. Nucleic Acids Res. 45, W122–W129.

Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N.J., Livak, K.J., Mikkelsen, T.S., Rinn, J.L., 2014. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat. Biotechnol. 32, 381–386.

Tsutsumi, K., Taniguchi, Y., Kawasaki, M., Taniguchi, M., Miyake, H., 2006. Expression of photosynthesis-related genes during the leaf development of a C3 plant rice as visualized by *in situ* hybridization. Plant Prod. Sci. 9, 232–241.

Turco, G.M., Rodriguez-Medina, J., Siebert, S., Han, D., Valderrama-Gomez, M.A., Vahldick, H., Shulse, C.N., Cole, B.J., Juliano, C.E., Dickel, D.E., et al., 2019. Molecular mechanisms driving switch behavior in xylem cell differentiation. Cell Rep. 28, 342–351 e344.

Umeda, M., Umeda-Hara, C., Yamaguchi, M., Hashimoto, J., Uchimiya, H., 1999. Differential expression of genes for cyclin-dependent protein kinases in rice plants. Plant Physiol. 119, 31–40.

Vahisalu, T., Kollist, H., Wang, Y.F., Nishimura, N., Chan, W.Y., Valerio, G., Lamminmaki, A., Brosche, M., Moldau, H., Desikan, R., et al., 2008. SLAC1 is required for plant guard cell S-type anion channel function in stomatal signalling. Nature 452, 487–491.

Wang, Y., Song, F., Zhu, J., Zhang, S., Yang, Y., Chen, T., Tang, B., Dong, L., Ding, N., Zhang, Q., et al., 2017. GSA: Genome Sequence Archive. Genom. Proteom. Bioinform. 15, 14–18.

Wendrich, J.R., Yang, B., Vandamme, N., Verstaen, K., Smet, W., Van de Velde, C., Minne, M., Wybouw, B., Mor, E., Arents, H.E., et al., 2020. Vascular transcription factors guide plant epidermal responses to limiting phosphate conditions. Science 370.

Zeng, M., Hu, B., Li, J., Zhang, G., Ruan, Y., Huang, H., Wang, H., Xu, L., 2016. Stem cell lineage in body layer specialization and vascular patterning of rice root and leaf. Sci. Bull. 61, 847–858.

Zhang, T.Q., Chen, Y., Liu, Y., Lin, W.H., Wang, J.W., 2021a. Single-cell transcriptome atlas and chromatin accessibility landscape reveal differentiation trajectories in the rice root. Nat. Commun. 12, 2053.

Zhang, T.Q., Chen, Y., Wang, J.W., 2021b. A single-cell analysis of the *Arabidopsis* vegetative shoot apex. Dev. 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. Mol. Plant 12, 648–660.

Zheng, G.X., Terry, J.M., Belgrader, P., Ryvkin, P., Bent, Z.W., Wilson, R., Ziraldo, S.B., Wheeler, T.D., McDermott, G.P., Zhu, J., et al., 2017. Massively parallel digital transcriptional profiling of single cells. Nat. Commun. 8, 14049.

Zhou, L.J., Xiao, L.T., Xue, H.W., 2017. Dynamic cytology and transcriptional regulation of rice lamina joint development. Plant Physiol. 174, 1728–1746.