



An optimized FACS-free single-nucleus RNA sequencing (snRNA-seq) method for plant science research

Kaimeng Wang ^{a,b,1}, Caiyao Zhao ^{a,1}, Sunhuan Xiang ^{a,c}, Kunyu Duan ^{a,b}, Xiaoli Chen ^{a,b},
Xing Guo ^{a,*}, Sunil Kumar Sahu ^{a,*}

^a State Key Laboratory of Agricultural Genomics, BGI-Shenzhen, Shenzhen, 518083, China

^b BGI College & Henan Institute of Medical and Pharmaceutical Sciences, Zhengzhou University, Zhengzhou, China

^c College of Life Sciences, University of Chinese Academy of Sciences, Beijing, China

ARTICLE INFO

Keywords:

Arabidopsis thaliana
Mature leaves
ScRNA-seq
SnRNA-seq

ABSTRACT

Recently, single-cell RNA sequencing (scRNA-seq) provides unprecedented power for accurately understanding gene expression regulatory mechanisms. However, scRNA-seq studies have limitations in plants, due to difficulty in protoplast isolation that requires enzymatic digestion of the cell walls from various plant tissues. Therefore, to overcome this problem, we developed a nuclei isolation approach that does not rely on Fluorescence Activated Cell Sorting (FACS). We validated the robustness of the FACS-free single-nucleus RNA sequencing (snRNA-seq) methodology in mature *Arabidopsis* plant tissue by comparing it to scRNA-seq results based on protoplasts extracted from the same batch of leaf materials. Sequencing results demonstrated the high quality of snRNA-seq data, as well as its utility in cell type classification and marker gene identification. This approach also showed several advantages, including the ability to use frozen samples, taking less suspension preparation time, and reducing biased cellular coverage and dissociation-induced transcriptional artifacts. Surprisingly, snRNA-seq detected two epidermal pavement cell clusters, while scRNA-seq only had one. Furthermore, we hypothesized that these two epidermal cells represent the top and lower epidermis based on differences in expression patterns of cluster-specific expressed genes. In summary, this study has advanced the application of snRNA-seq in *Arabidopsis* leaves and confirmed the advantages of snRNA-seq in plant research.

1. Introduction

Higher plants and animals have complex functional organs consisting of diverse and specialized cell types, and, therefore, it is a key question to understand how the one cell of zygote with the same genomic blueprint develops and differentiates to hundreds of thousands of cells with distinct functions and morphologies. In the post-genome era, high-throughput RNA sequencing (RNA-seq) has fueled the discovery and innovation of research, promoting the identification of novel genes and transcription factors, which has contributed to the study of biological expression regulation mechanisms (Ayturk, 2019). However, bulk RNA-seq usually represents an average gene expression of the millions of cells (Olsen and Baryawno, 2018), ignoring the gene expression of the rare and subgroup cells. Therefore, many methods of single-cell RNA sequencing (scRNA-seq), especially the high-throughput microfluidic and microwell technologies (Picelli et al., 2013; Hashimshony et al., 2016; Gierahn et al.,

2017; Han et al., 2018; Macosko et al., 2015; Zheng et al., 2017; Klein et al., 2015), have been used to obtain hundreds of thousands of single-cell transcriptomes. With the advantages of high-resolution and high-throughput (Rostom et al., 2017), scRNA-seq has emerged as a powerful tool for further accelerating progress in dissecting cell types and functions, exploring the cooperative operation mode between cells, and studying tissue heterogeneity at single-cell resolution (Han et al. 2018; Han et al. 2020; Han et al. 2022).

Although genomics has enabled us to gain detailed insights into genetic architecture and evolution of the plants (Wang et al., 2022; Guo et al., 2021; Fan et al., 2021; Sahu et al., 2019; Fan et al., 2020; Yang et al., 2022), the technologies of high-throughput scRNA-seq have been successfully used to reveal the molecular mechanism for more than ten years and even several whole-cell-atlas of model animals have been completed (Han et al. 2018; Han et al. 2020; Han et al. 2022; Liu et al., 2022), the application of scRNA-seq in plant tissues is extremely limited. There are

* Corresponding authors.

E-mail addresses: guoxing@genomics.cn (X. Guo), sunilkumarsahu@genomics.cn (S.K. Sahu).

¹ These authors have contributed equally to this work.

currently a few scRNA-seq studies in model plant tissues, such as *Arabidopsis* root tip and shoot apical meristem, rice root, *maize* shoot apical meristem (SAM), tomato shoot apex, and *Physcomitrella* leaf cell (Jean-Baptiste et al., 2019; Denyer et al., 2019; L. Liu et al., 2019; Marand et al., 2021; Ryu et al., 2019; Zhang et al., 2019, 2021a; Marand et al., 2021; Zhang et al., 2021b; Tian et al., 2020; Kubo et al., 2019; Han et al., 2022). This is mainly due to the rigid plant cell wall which increases the difficulty of obtaining a single cell. Therefore, the protoplast isolation by hydrolyzing the cell wall is necessary for separating intact plant single cells, which imposes huge challenges. For example, the complex component of organic macromolecules in the cell wall between different species and tissues requires specific enzyme components and concentrations to hydrolyze the cell wall. Meantime, the contact area between the sample and the enzyme will also affect the separation efficiency. The second reason is that a wide range of plant cell sizes, coupled with the influence of osmotic pressure on cell size, may lead to experimental biases such as too large cells that cannot be captured during droplet capture, which will affect subsequent analyses (Bakken et al., 2018; Shaw et al., 2021; Sunaga-Franze et al., 2021). Apart from the reasons stated above, scRNA-seq cannot be used on frozen samples, which limits the application of samples when acquiring a fresh sample is problematic.

Besides scRNA-seq, single-nucleus RNA-sequencing (snRNA-seq) is an alternative option for the samples with difficulties in isolation of single cells. For example, Grindberg et al. used this method of snRNA-seq for neural progenitor cells and dentate gyrus tissue and generated comparable transcriptome with scRNA-seq of other tissues (Grindberg et al., 2013). This method also removes changes in genes expression caused by enzyme hydrolysis and could be widely used in eukaryotes (Bakken et al., 2018, 2021; Zeng et al., 2016). Considering the blocking of cell walls in scRNA-seq, the snRNA-seq could be a promising method to expand the range of RNA sequencing at a single cell level in a plant. Until now, snRNA-seq has been used for the *Populus* and the *Arabidopsis* root or tomato shoot apex (Tian et al., 2020b; Farmer et al., 2021; Conde et al., 2021), which not only provides considerable informative transcriptome as scRNA-seq, but also shows a bright expectation of transcripts in mononuclear cells research of plant biology.

Although plant scRNA-seq and snRNA-seq research have surged in recent years, the detailed comparison between these two methods in plants is obscure (Farmer et al., 2021). SnRNA-seq has been applied in plants, such as tomato (shoot apex), and poplar (shoot apices and lignified stems) (Marand et al., 2021; Tian et al., 2020b; Conde et al., 2021), collecting intact nuclei from plant tissues without RNA degradation remains a challenge for many species and tissues. At present, most snRNA-seq studies use a razor blade cutting method to prepare crude nucleus suspensions, and then obtain purified nucleus suspensions by FACS in a plant (Marand et al., 2021; Farmer et al., 2021; Conde et al., 2021; Thibivilliers et al., 2020). In the experiment, FACS increases the requirements for basic equipment and the cost, while the long-time sorting reduces the experimental efficiency. In this study, by using mature leaf materials of *Arabidopsis* as the source material, we present a nuclei isolation approach that does not rely on FACS. Compared with scRNA-seq data based on the protoplasts extracted from the same batch leaf materials, we validated the robustness of FACS-free snRNA-Seq approach in mature plant tissue. We confirmed that these two approaches yielded equivalent transcriptomic information and by examining the differences between the transcriptomes of scRNA-seq and snRNA-seq in detail we identified five major cell types in *Arabidopsis* leaves (epidermal cell, guard cell, vascular cell, mesophyll cell, and dividing cell). We also found that, compared to scRNA-seq of leaves, snRNA-seq can reduce bias and obtain a more reasonable ratio of cell types.

2. Materials and methods

2.1. Plant materials and growth condition

The wild-type *A. thaliana* (ecotypes Col-0) was used for all single

cells/nuclei RNA-seq experiments. *Arabidopsis* seeds were soaked in water in the dark for two days at 4 °C, and after being sterilized with 75 % alcohol and germination on vertical Murashige and Skoog (MS) plates at 21 °C in long-day conditions (16 h light and 8 h dark). After 10 days of growth, the plants were transferred from MS medium to soil and cultivated in the same temperature and light condition as above. The second and third pairs of mature leaves of 3-week-old *Arabidopsis thaliana* were harvested to isolate protoplasts and nuclei. Among them, part of the fresh sample obtained is used for scRNA-seq experiments, and the other part is stored in liquid nitrogen, and then used for snRNA-seq experiments.

2.2. Protoplast isolation

Protoplast was isolated from mature leaves of *Arabidopsis thaliana* through the tape-sandwich method as described previously (Wu et al., 2009), with slight modifications. Briefly, 3-weeks-old *Arabidopsis thaliana* fresh leaves were adhered to the tape with the reverse side up, then a layer of tape was affixed to the upper layer, then the tape was peeled off. Both tapes with the leaves were placed in 10 ml of protoplasts enzymatic hydrolysis solution (0.4 M mannitol, 1.5 % CellulaseR-10, 0.4 % Macerozyme, 20 mM KCl, 20 mM MES pH=5.7, 10 mM CaCl₂, 0.1 % BSA, 1 mM Dithiothreitol (DTT)), and dark-incubated at 75 rpm on a platform shaker for 1.5 h at 28 °C. After cell wall digestion, the solution was filtered with a 40 µl cell sieve, then washed the remaining protoplasts on the tape with an equal volume of W5 solution (154.5 mM NaCl, 125 mM CaCl₂, 5 mM KCl, 2 mM MES PH=5.7, 5 mM D-Glucose), centrifuged at 100 g for 3 min, removed the supernatant, and repeatedly washed with 10 ml W5 solution for 3 times. Finally, the cells were resuspended in MMG solution (1 M mannitol, 1 M MgCl₂, 1 M MES PH=5.7) and counted under the microscope.

2.3. Nuclei isolation

The nucleus isolation was done as previously (Dorrity et al 2021), with slight adjustments. The 10 ml buffer A solution (0.8 M sucrose, 10 mM MgCl₂, 25 mM Tris-HCl pH = 8.0, 0.1 mM DTT, 0.4 U/µl RNase inhibitor, 0.1 mM phenylmethylsulfonyl fluoride (PMSF)) was prepared under ice-cold condition. About 8 mature leaves of *A. thaliana* were taken out from -80 °C, placed on a cold petri dish, and 2 ml buffer A was added for infiltration, and then used a razor to gently chop for 5 min until there is no obvious visible block of tissue. Using a 40 µl cell sieve for filtration, the remaining nuclei were washed with the remaining buffer A, then secondary filtrated with a 30 µl sieve, and then incubated on ice for 10 min, and then centrifuged at 50g for 5 min at 4 °C to remove bulk impurities, and then centrifuged the supernatant at 2000g for 10 min at 4 °C. After the supernatant was removed, the pellet was suspended with 1 ml buffer B (0.4 M sucrose, 10 mM MgCl₂, 25 mM Tris-HCl pH = 8.0, 0.2 U/µl RNase inhibitor, 0.1 mM PMSF, 1 % Triton X-100) and added 25 %/75 % Percoll gradient upper layer (first add 25 % Percoll, then add 75 % Percoll to the bottom of 25 % Percoll), centrifuged at 3500g for 15 min at 4 °C. Then the upper layer was discarded and taken the middle layer (with white material on the 25 % Percoll layer) into a new centrifuge tube. The solution obtained in the previous step was washed with 5 vol buffer B, centrifuged at 1800g for 10 min at 4 °C, discarded the supernatant, and gently pipetted the cell suspension buffer (CRB) to resuspend the pellet. 5 µl nucleus suspension and 5 µl DAPI solution were mixed for count under a fluorescence microscope. The intact nuclei were stained by DAPI and counted under a fluorescent inverted microscope. However, when the nuclei membrane was broken and the nuclei appear to be clustered, the nuclei were not counted. Nuclei with intact nuclei and high fluorescent signal were counted, and about 200 µl of nuclei suspension was obtained in this study at a concentration of ~3875 nuclei/µl.

2.4. snRNA-seq/scRNA-seq library construction

We used DNBelab C Series Single-Cell seq Library Prep Set (v1.0, MGI, #1000021878) (L. Han et al., 2020) as previously described (C. Liu et al., 2019) to generate 14 RNA-seq libraries. The library was constructed according to the method provided by the manufacturer. Generally speaking, single-nuclei/cells suspension is converted to bar-coded single-cell/nucleus RNA-seq libraries through procedures including droplet encapsulation, emulsion breakage, capture beads collection, reverse transcription, DNA amplification, and purification. Indexed sequencing libraries were prepared according to the user guide. The concentration of the libraries was determined using Qubit 3.0, while snRNA-seq/scRNA-seq library sequencing was performed using DNBSEQ-T1&T5 platform at China National GeneBank.

2.5. Generation of single-cell/nuclei expression matrices

snRNA-seq and scRNA-seq data for the seven wild-type replicates from this study, (Supp. Table S1), and other scRNA-seq data directly for the two replicates of previously published data (Kim et al., 2021) were used for the analyses. Our data were processed individually using the BGI cell ranger count and these pipelines include processing of raw scRNA-seq data to align reads and generate gene-cell matrices. The genome and GTF annotation files of *A. thaliana* were downloaded from the TAIR10 website (<https://www.arabidopsis.org/>). Reads were aligned to the TAIR10 reference genome by the aligner STAR (v. 2.7.4a). The Bam file and gene-cell matrices were formed by PISA (v.1.1.0) for downstream analysis.

2.6. Data filter, dimensionality reduction, UMAP visualization, and identification of cell types

Unless specifically mentioned, all downstream snRNA-seq/scRNA-seq analyses were performed using the Seurat R package (v.4.0.3) (Hao et al., 2021). When filtering data, we set the parameter “min. cell” to 3, and “min. feature” to 200, while nFeature_RNA is between 200 and 3000, nCount_RNA is greater than 1000, percent.atm and percent.atc less than 5. Then, each dataset was log-normalized and integrated with the Seurat R package, and the top 2000 variable genes were used for feature selection. Principal component analysis (PCA) was performed to reduce the dimensionality on the log-transformed gene-barcode matrices of the most variable genes. The top 30 PCs were selected as input for UMAP visualized (McInnes et al., 2018), clustering analysis, and performed using Seurat’s FindClusters method using a resolution value of 0.5. Cluster marker genes were identified using “FindAll-Markers” with parameters min.pct = 0.25, logfc.threshold = 0.25. Cells in our dataset were assigned a cell type based on difference express genes(DEGs) and cluster-specific marker genes.

2.7. Correlation analyses

Seurat’s AverageExpression function was used to separately obtain expression values of each gene in the subsets of cells and nuclei belonging to each cluster. Then the pseudo-bulk transcriptomes were created by further calculating the mean expression of each gene in all cells by weight of cell numbers of clusters. Then, we obtained DEGs of the bulk transcriptome and the non-bulk transcriptome of the mature leaves of *Arabidopsis thaliana* from the published article (Kim et al., 2021). Then, we calculated the Spearman correlation coefficient between the log2 detection components for the gene expression, and used the ggplot2 package to make a correlation scatter diagram.

2.8. Heat map and gene ontology (GO) enrichment analysis

Differentially expressed genes ($|log_2FC| > 1$) between candidate upper and lower epidermis clusters (C10 and C3) were used for the

heatmap analysis. The expression levels of these DEGs in shoot apex part reporter system were downloaded from http://bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi?dataSource=Shoot_Apex. The resulting dendrogram and heat map were plotted using hiplot (<https://hiplot.com.cn/basic>). The data of row were scale by default parameters. At the same time, set the parameter cluster methods to "ward.D", and the row distance measure of the row or column to "euclidean".

The enrichment of gene ontology (GO) terms was analyzed using TopGO Package (Alexa and Rahnenführer, 2009).

3. Results

3.1. snRNA-seq and scRNA-seq generate comparable transcriptome datasets

The first critical step is single-cell/nucleus isolation in transcriptome studies on individual cells. In this study, we can quickly obtain purified nuclei in a short time by optimizing the existing single-nucleus isolation protocol using density gradient centrifugation instead of relying on flow cytometry sorting technology (Dorrity et al., 2021) (detail in methods). For the comparison, protoplast was also obtained digesting the cell wall as described by Zhang et al. (Zhang et al., 2019). Protoplasts and nuclei were isolated from *Arabidopsis thaliana* mature leaves (Fig. 1A, Supp. Fig S1), to obtain single-cell/nucleus suspension by the microfluidic system of DNBelab C4 (Han et al., 2022). We constructed seven libraries by following similar steps for scRNA-seq and snRNA-seq assays, and used the same clustering parameters for data processing to compare snRNA-seq and scRNA-seq transcriptomes.

Subsequently, we assessed the quality of the sequencing data for each library through several parameters, such as the total number of reads, the ratio of clean reads in raw reads, the Q30 bases in RNA reads and the barcode. A total of about 2931 million and 2633 million raw reads were obtained in the nucleus and protoplasts respectively (Supp. Table S1). After the quality control, the clean reads accounted for about 90.5 % and 86.3 % of the raw reads, respectively. The average Q30 bases on RNAs and barcodes were 94.79 % and 90 % in the nucleus, respectively, while in protoplasts they account for 95.25 % and 90 %, respectively. After filtering low-quality data, snRNA-seq obtained 10,516 nuclei containing 21,668 genes and scRNA-seq obtained 9015 cells containing 19,955 genes (Supp. Table S1). Genes of snRNA-seq and scRNA-seq transcriptomes accounted for 79 % and 73 % of the 27,420 predicted protein-coding genes in *Arabidopsis*, respectively.

To identify distinct cell populations, an unbiased, unsupervised clustering was performed on the 10,516 nucleus and 9015 cells transcriptomes using the Seurat software package (Satija et al., 2015). The plotting of the transcriptomes from the seven replicates in two dimensions using Uniform Manifold Approximation and Projection (UMAP) (McInnes et al., 2018; Becht et al., 2019) yielded largely overlapping distributions of cells (Fig. 1B-C) and a similar proportion of cells in different clusters (Fig. 1A-B). The results illustrated the high degree of reproducibility of different replicates.

Previous reports have shown that the process of enzyme digestion of plant cell walls introduce transcriptional artifacts during enzyme digestion, limiting the applicability of scRNA-seq using isolated protoplasts (Sunaga-Franze et al., 2021). Therefore, in order to confirm the effect of enzymatic protoplasts on scRNA-seq, differentially expressed genes (DEGs) between the non-protoplast bulk RNA-seq and protoplast bulk RNA-seq ($|log_2FC| > 8$) were selected to compare their average expression difference between scRNA-seq and snRNA-seq (Kim et al., 2021) (Supp. Table S2). We found that most of the highly expressed genes in bulk RNA-seq ($log_2FC > 8$) are also highly expressed in our snRNA-seq (Fig. 1D indicated in red), while highly expressed genes in protoplast bulk RNA-seq ($log_2FC < -8$) mostly had higher expression levels in scRNA-seq (Fig. 1D indicated in green). This result suggest that scRNA-seq can indeed introduce artifacts at enzymatic digestion of plant cell walls as proved in protoplast bulk RNA-seq. Next, to evaluate the

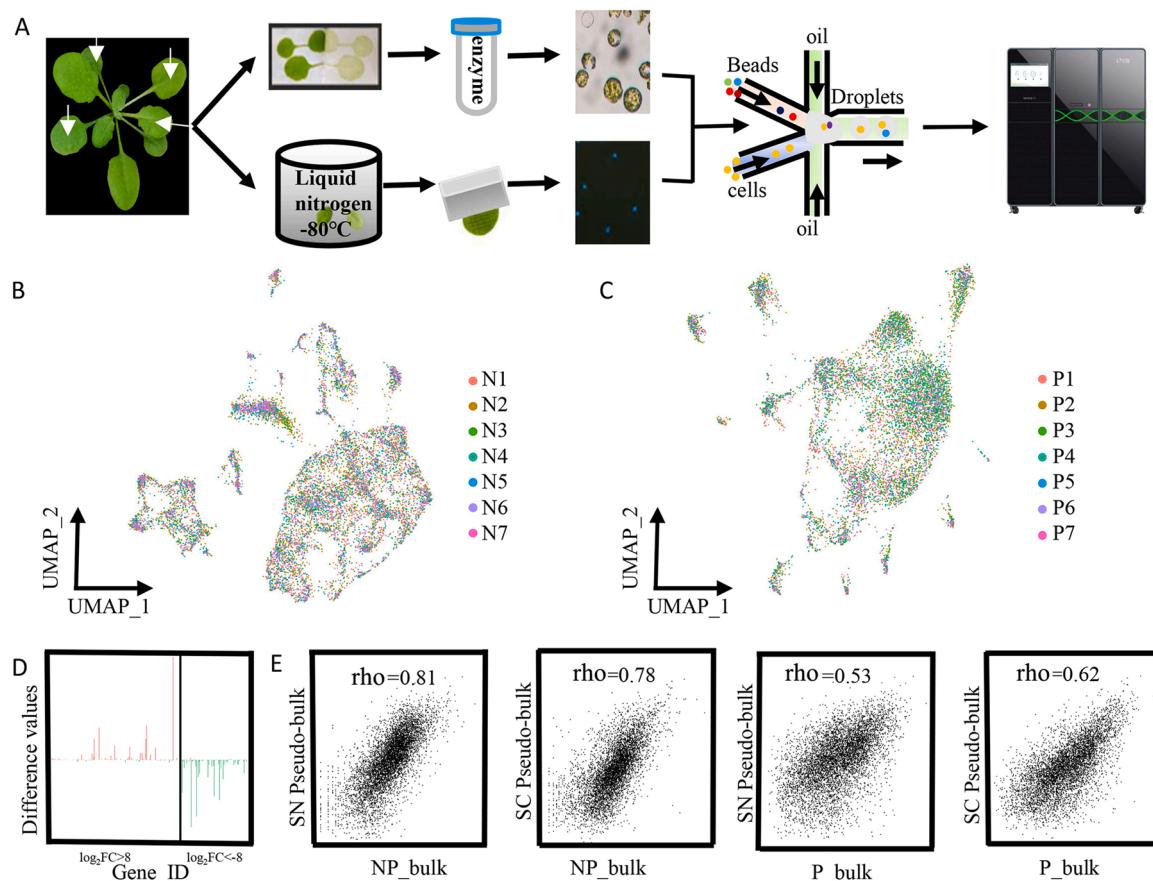


Fig. 1. Isolation and Uniform Manifold Approximation and Projection (UMAP) cluster analysis of single-cell and single-nucleus transcriptomes from the *Arabidopsis* leaf. (A) Workflow used for snRNA-seq and scRNA-seq to obtain transcriptomes from individual *Arabidopsis* leaf cells. (B) and (C) UMAP projection plots showing dimensional reduction of snRNA-seq and scRNA-seq transcriptomes from seven biological replicates. (D) Barplot showing the expression up-regulation (green) and down-regulation (red) of previously identified protoplasting-responsive genes in scRNA-seq comparing snRNA-seq. x-axis: the selected differentially expressed genes in protoplasting bulk RNA-seq (P_{bulk}) vs. NP_{bulk}). y-axis: The differential expression values of these genes in scRNA-seq compared with snRNA-seq. (E) snRNA-seq and scRNA-seq pseudo-bulked expression data (SN Pseudo-bulk and SC Pseudo-bulk) are compared to bulk RNA-seq of non-protoplasted (NP_{bulk}) and protoplasted (P_{bulk}) leaf samples, respectively.

relation of sc/snRNA-seq with bulk transcriptomes, we evaluated the extent to which snRNA-seq and scRNA-seq combined pseudo-bulk transcriptomes resemble bulk protoplast transcriptomes or bulk non-protoplast transcriptomes of *A. thaliana* leaf (Kim et al., 2021) (Supp. Table S2). We observed strong correlations between single-cell/nucleus pseudo-bulked and non-protoplasts bulk transcriptomes (snRNA-seq: Spearman's rho= 0.81; scRNA-seq: Spearman's rho= 0.78) (Fig. 1E), but lower correlation with the protoplast bulk transcriptome (snRNA-seq: Spearman's rho= 0.53; scRNA-seq: Spearman's rho= 0.62). This result suggests that the process of protoplast isolation indeed changes the original transcriptome of the plant sample. As a result, snRNA-seq pseudo-bulk transcriptome has a higher relationship with bulk transcriptome than scRNA-seq pseudo-bulk transcriptome, and scRNA-seq pseudo-bulk transcriptome has a higher association with bulk protoplast transcriptome than snRNA-seq pseudo-bulk transcriptome. In brief, snRNA-seq and scRNA-seq both generate biologically meaningful transcriptome datasets.

3.2. The major cell types of mature leaves can be identified in snRNA-seq transcriptome

For further analyses, first, we tried to annotate the 17 clusters of snRNA-seq by identifying cluster-specific marker genes and the expression patterns of well-known marker genes in different clusters (Fig. 2A-C, Supp. Table S3). Each cluster showed the accumulation of highly specific gene expression (Fig. 2B, Fig. 2A) which, in general,

enrich specific gene ortholog (GO) corresponding to the functions of one cell type (Supp. Table S4). We identified epidermal cells (EC), mesophyll cells (MC), guard cells (GC) of stomata, vasculature cells (VC), and dividing cells (DC) based on the accumulation of marker genes (Fig. 2A). As mentioned in previous reports, compared with t-distributed stochastic neighborhood embedding (t-SNE), UMAP technology generates the topology of clusters revealing the meaningful organization of cells and nuclei (Jean-Baptiste et al., 2019; Zhang et al., 2021a; Farmer et al., 2021; Satterlee and Strable, 2020; Zhu et al., 2021). In agreement with this observation, we also found that the cells of similar types such as vascular cells (clusters 5, 13, 14, 15) were gathered together on the UMAP plot, which were separated in the opposite orientation on the t-SNE plot (Fig. 2A, B). Therefore, we only used the UMAP visualization for the rest of this study.

Epidermal cells consist of two clusters (cluster 3 and 10, Fig. 2A), in which epidermal-specific genes involved cuticular wax and suberin biosynthesis such as *FDH*, *LTP1*, *LTPG1*, *CUT1*, and *ATML1* were predominantly expressed (Fig. 2B-C) (Efremova et al., 2004; Clark and Bohnert, 1999; Lee et al., 2009; Kunst et al., 2000; Takada et al., 2013). Although the transcriptome of C3 is highly correlated to that of C10 (Fig. 2C), the enriched top GO terms of the differentially expressed genes (DEGs) of C3 and C10 are slightly different. While DEGs of C10 included a relatively higher number of genes involved in the wax, cuticle metabolic process, the DEGs of C3 included more genes involved in the cellular response to internal and external signals (such as oxygen levels, bacterium, chitin, a nitrogen compound, jasmonic acid, etc.) and

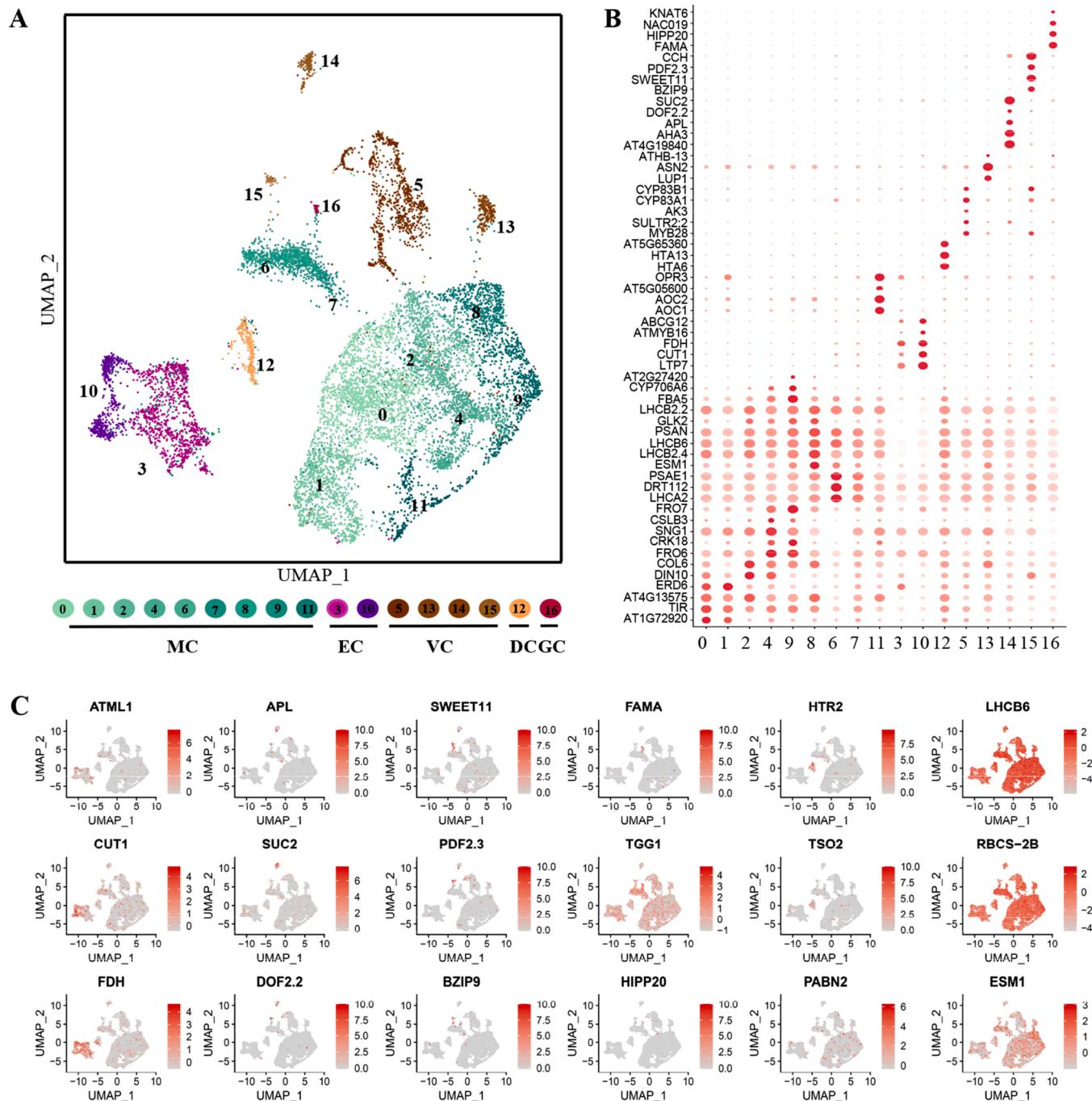


Fig. 2. Functional annotation of the *Arabidopsis* leaf cell types based on snRNA-seq data. (A) Assignment of *Arabidopsis* leaf cell types based on the characterization of the expression profile of cell-type marker genes. (B) Expression pattern of representative cluster-specific marker genes. Dot diameter indicates the proportion of cluster cells expressing a given gene. (C) UMAP plot showing the selected cell-type marker genes for different cell types.

regulation of defense response, suggesting that those cells of C3 and C10 may represent two different sub-cell types of the epidermis (Supp. Table S4). Guard cell marker genes, *FAMA* and *TGG1* (Shirakawa et al., 2014; Islam et al., 2009) were specifically highly expressed in C16 (Fig. 2B-C). C16 was assigned to guard cells after the enriched GO terms of DEGs of C16 were combined with the biological processes of response to water and stomatal movement (Supp. Table S4).

We also identified clusters of vascular tissues based on the enrichment of known marker genes. C14 was defined as companion cells (CC), in which several CC marker genes were detected such as *APL* (Bonke et al., 2003), an indispensable regulator gene for phloem differentiation, *SUC2* encoding a transporter required for apoplastic phloem sucrose

loading in source tissues to transport it to sink tissues (Gottwald et al., 2000) and *PHLOEM PROTEIN2-A1* (PP2-A1). The proton ATPase gene *AHA3*, only detected in the phloem, was also specially expressed in cluster 14 (Fig. 2B). *SWEET11* is highly expressed in C15, which had previously been reported to be expressed in the phloem parenchyma for providing sucrose to Sucrose transporters (SUTs) for phloem loading (Kim et al., 2021; Eom et al., 2015). *CCH* was also enriched in C15 (Fig. 2B), which is mainly located in the vascular bundles and accumulates in stem sieve elements (Mira et al., 2001; Lara et al., 2003; Weltmeier et al., 2009). *ASN2*, one of the three genes encoding asparagine synthetase (Gaufichon et al., 2013; Lam et al., 1998), which has been reported to be expressed in the vascular region adjacent to the leaf

mesophyll cells. In this study, *ASN2* was highly enriched in C13 (Fig. 2B), indicating that C13 represents vascular tissues. *SULTR2;2* had previously been reported to be localized specifically in the leaf vascular bundle sheath cells (Takahashi et al., 2000), and was found to be highly expressed in cluster C5 (Fig. 2B).

The mesophyll group comprised eight clusters (C0, C1, C2, C4, C6, C7, C8, C9, C11) in which photosynthesis-related genes, such as *RBCS1A*, *LHCA2*, *BCA1*, etc., were highly expressed. Meanwhile, we have also identified a large number of enriched gene functions and processes related to photosynthesis in DEGs of these clusters (Supp. Table S4). These mesophyll cells accounted for 71.7 % of the total

number of the leaf cells identified in snRNA-seq dataset (Supp. Table S5), which is consistent with the previous results (Kim et al., 2021). It was found that the abaxial marker gene *FIL* (*YAB1*) (Uemoto et al., 1830) was slightly highly expressed in C8, suggesting that this cluster may represent spongy parenchyma which consisted of the mesophyll cells located in the abaxial side of *Arabidopsis* leaf (Supp. Fig S2). We have also identified a group of cells (C12), in which GO terms of DEGs involved in mitosis, such as DNA modification, histone binding, nucleosome positioning, chromosome segregation, etc., are significantly enriched. Except for mitosis-related genes (such as DNA synthesis genes *HIS4* and *HTR2*), there were no marker genes of any specifically

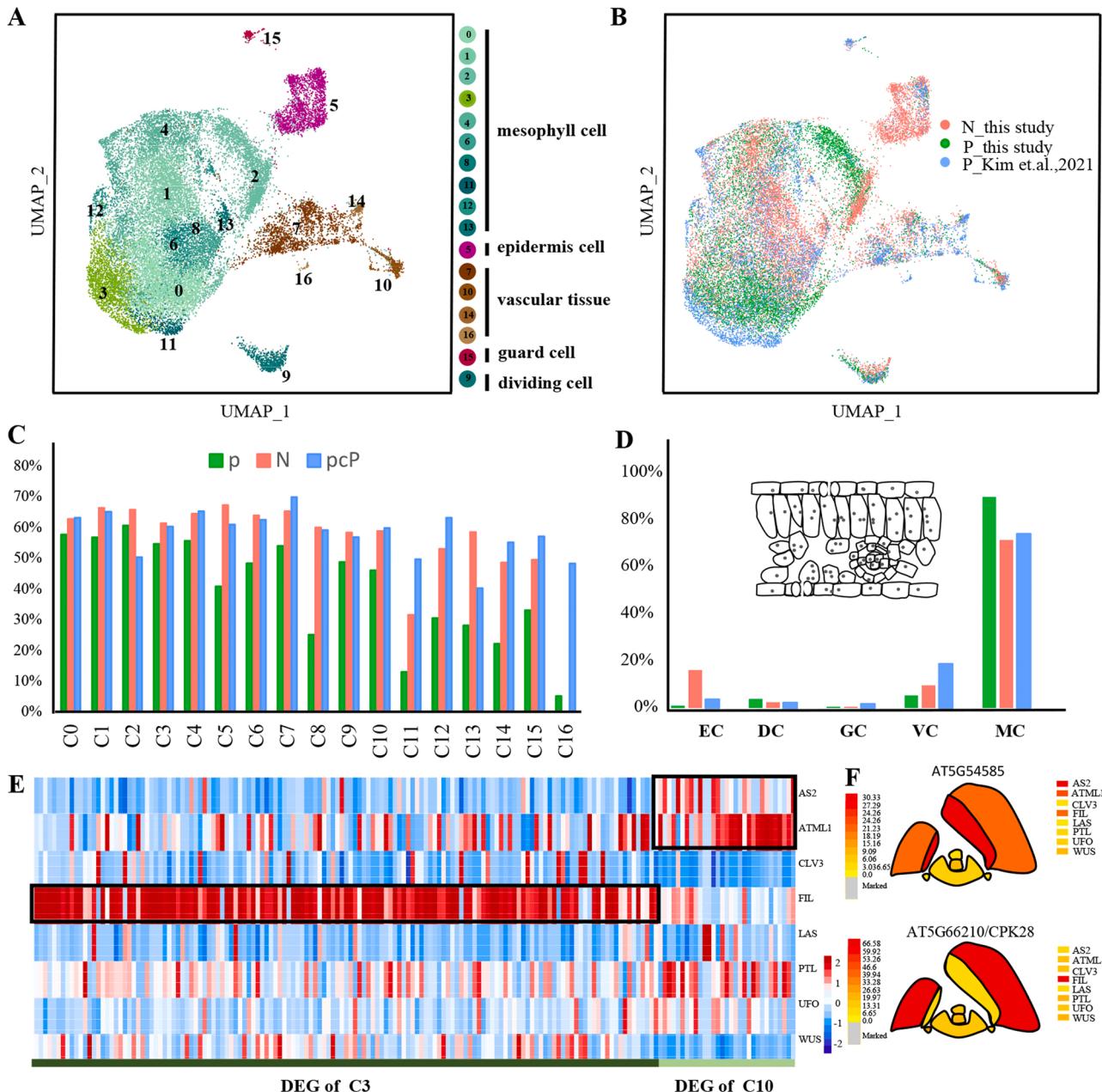


Fig. 3. Comparison of snRNA-seq and scRNA-seq transcriptomes of *Arabidopsis* leaf. (A-B) (A) Visualization of 17 cell clusters using UMAP by co-clustering three datasets (snRNA-seq and scRNA-seq of this study, and scRNA-seq of Kim et al., 2021) as shown in (B) indicated by red, green and blue, respectively. (C) Percentage of *Arabidopsis* genes found expressed upon analysis of snRNA-seq (red), scRNA-seq of this study (green), and scRNA-seq of Kim et al., 2021 (blue) for each 17 clusters. (D) Percentage of different cell types identified in three datasets as in (B). (E) Expression heatmap of DEGs between C3 and C10 of the snRNA-seq in the transcriptome of different cell types of the *Arabidopsis* shoot apex (data from http://bar.utoronto.ca/efp_arabidopsis). x-axis: DEGs of C3 and C10: differentially expressed genes of cluster3 and cluster10; y-axis: gene expression value in different cell types marked by the eight marker genes: AS2, leaf adaxial; ATML, epidermal L1 layer; CLV3, central zone; FIL, leaf abaxial; LAS, organ boundary; PTL, leaf margin; UFO, enlarged peripheral zone; WUS, rib meristem. (F) The expression of two genes of DEGs in (E) which were preferentially highly expressed in adaxial and abaxial tissues are shown in the cartoon of *Arabidopsis* shoot apex.

differentiated cells, indicating that this cluster represents proliferating cells. In short, these 17 clusters of snRNA-seq contain the main cell types of mature leaves of *A. thaliana*.

3.3. Comparison between transcriptomes of snRNA-seq and scRNA-seq

Using the same method as snRNA-seq, scRNA-seq generated 16 clusters through the same unsupervised clustering parameters (Fig. 3A), of which the transcriptome of each cluster also showed the accumulation of highly specific gene expression (Suppl. Fig. 4A). According to the specific expression patterns of the same marker genes used in the annotation of clusters in snRNA-seq, these clusters of scRNA-seq transcriptome also were assigned to five major groups of cells: mesophyll cell (C0, C1, C2, C3, C5, C7, C8, C9, C11, C13, C15), epidermal cell (C12), guard cell (C14), vascular tissue (C6, C10) and dividing cell (C4) (Fig. 3A–C). At the same time, there is a strong correlation between the subgroups of mesophyll cells (Suppl. Fig. 4B).

Next, to further assess the differences between the two approaches, we co-clustered snRNA-seq and scRNA-seq data of this study, along with the scRNA-seq transcriptome data from a recent study (Kim et al., 2021), resulting in a total of 17 clusters (Fig. 3A). These three datasets produced a good distribution overlap on the UMAP plot, implying that our data is comparable with previous research (Fig. 3B).

To further investigate the biological significance of cluster in the merging data, according to the specificity of transcripts of the above-mentioned marker genes, we defined the cell types of the co-clusters. Likewise, the merged data identified five major cell types in *Arabidopsis* leaves: mesophyll cell, epidermal cell, guard cell, vascular tissue and dividing cell. (Fig. 3A, Suppl. Fig. 5). Taken together, the principal cell types of *A. thaliana* leaves could be derived using both snRNA-seq and scRNA-seq data (Fig. 2A, Fig. 3A, A).

Next, we calculated the percentages of genes captured in each cluster of these three datasets in the co-clustering analysis of the total genes encoding proteins in *A. thaliana*. The ratio of genes obtained by scRNA-seq (Kim et al., 2021) and snRNA-seq (this study) was found to be similar in each cluster (Fig. 3C). Then, we separately counted the proportion of cells of each cluster to the total cells identified in each of the three datasets (Fig. 3D, Supp. Table S5). Although the cells of three methods in each cluster are variable, the cell ratio of the data in each type of cell is comparable. The results revealed that mesophyll cells were found in the greatest number, while guard cells were found in the smallest number, reflecting the actual ratio of each cell type in an *Arabidopsis* leaf.

While the cell ratio of most cell types identified by the two methods of snRNA-seq and scRNA-seq is similar, the proportion of epidermal cells identified in the snRNA-seq dataset was significantly higher than that in scRNA-seq. It's worth noting that there were two clusters of epidermal cells in the snRNA-seq dataset, but only one in the scRNA-seq dataset of this work or previous research (Kim et al., 2021). Considering differences among the top enriched GO terms in DEGs of the two epidermal cell clusters (C3 and C10) identified by snRNA-seq, we proposed that these two clusters may represent the upper and lower epidermis, respectively. To test this hypothesis, we further analyzed the DEGs ($| \log_2FC | > 1$) between these two clusters of the epidermis (Supp. Table S6). We obtained the expression information of these DEGs in other transcriptomic datasets which included transcriptomes of *FIL* marked cells and *AS2* marked cells representing the leaf tissues of abaxial and adaxial areas, respectively, and other specifically expressed genes marked cells in leaf (such as *ATML1* gene marked epidermis, etc., data collected and normalized from <http://bar.utoronto.ca/efp/>). The results showed that, while many highly expressed genes in C10 were enriched in the transcriptomes of *AS2* and *ATML1* marked cells, the most highly and specifically expressed genes in C3 were enriched in *FIL* marked cells (Fig. 3E). For example, *AT5G54585* of the DEGs in C10 was highly expressed in the adaxial region of the leaf, while *AT5G66210/CPK28* in C3 was highly expressed in the abaxial side of the leaf (Fig. 3F). Although more functional and expression evidence is

necessary, these results suggest that C3 and C10 in snRNA-seq represent two very similar, but different sub-cell types of the epidermis, which may be the lower and upper epidermis, respectively.

4. Discussion

Compared with scRNA-seq, snRNA-seq has become a powerful tool for transcriptomics due to its unique advantages, such as less biased cellular coverage, reduced dissociation-induced transcriptional artifacts, and applicability to frozen specimens (Bakken et al., 2018; Ma et al., 2021). In this study, we introduce a simple method of isolating plant single-nucleus for the snRNA-seq on the droplet platform of DNBelab C4 of BGI (Han et al., 2022), that does not rely on flow cytometry sorting selection and is easier and faster than existing snRNA-seq nuclei and scRNA-seq protoplast isolation methods. The ability to obtain high-quality nuclei by centrifugation makes single-cell research accessible to regular laboratories, which contributes to accelerating the single-cell research. By employing the new method, we accomplished snRNA-seq of *A. thaliana* leaves and obtained biologically informative transcriptome, of which the most of cell types of leaf were identified, such as epidermal cell, mesophyll cell, guard cell, vascular tissue, and a group of proliferating cells. By comparing the transcriptome of snRNA-seq with scRNA-seq in detail, we found that snRNA-seq could produce comparable transcriptomic information (for example, the numbers of cell types and genes captured), and even distinguish the cell-types of lower and upper leaf epidermis.

The advantages of snRNA-seq outweigh the limitations of plant scRNA-seq, which requires the time-consuming but compulsory usage of protoplasts for single-cell isolation. Although both scRNA-seq and snRNA-seq can obtain almost the same cell groups of *Arabidopsis* leaves, snRNA-seq seems to have a greater advantage in avoiding the bias of capturing proportions of different cell types. While in the scRNA-seq data of this study or previous research, the proportion of epidermal cells captured was obviously lower than the real cell ratio of *Arabidopsis* leaves, snRNA-seq can obtain relatively more epidermal cells, which may be closer to the actual condition (Kalve et al., 2014). When isolating protoplasts, prolonged enzymatic hydrolysis and changes in osmotic pressure will cause epidermal cells to be more easily broken due to their location in the outermost layer of leaf tissue. However, the isolation of single-nucleus by physically breaking the cell wall and cytomembrane can extremely reduce the bias of cell capture. This method of nucleus isolation can also be used to capture the transcriptome of some cell types with too hard and rigid secondary cell walls to be hydrolyzed such as *Arabidopsis* root endodermal cells with suberized casparyan strip (Farmer et al., 2021). On the other hand, although scRNA-seq can theoretically obtain more transcripts included in the nucleus and cytoplasm than snRNA-seq (Farmer et al., 2021), the transcriptome of snRNA-seq has a higher correlation with non-protoplast bulk RNA-seq than scRNA-seq (Kim et al., 2021). The discrepancy in the results may be due to the longer time spent in the process of protoplast isolation, which could have probably changed the transcriptome (Shaw et al., 2021; Sunaga-Franze et al., 2021). Consistent with that, the transcriptome of scRNA-seq was found closer to the bulk RNA-seq of protoplasts than snRNA-seq.

However, apart from the advantages of capturing cells, the disadvantages also need to be stated. In theory, genetic information in the cytoplasm was lost in snRNA-seq compared to protoplast-based scRNA-seq. So, the proportion of mature mRNA with poly-A tails in the nucleus is lower which leads to a risk that the genes detected in each nucleus in snRNA-seq may be under-reported, which is not conducive to the identification of cell subtypes. Although snRNA-seq obtains fewer gene numbers than scRNA-seq, this difference seems to have a very tiny effect on the clustering of cells in this study. In this study, scRNA-seq obtained 16 cell subsets, and snRNA-seq produced 17 cell subsets, both of which produce clusters of the main cell types of *Arabidopsis* leaves. The study of snRNA-seq of *Arabidopsis* roots reported a similar conclusion about cell

clustering (Tian et al., 2020). Moreover, snRNA-seq cannot be used for the cells without a nucleus, such as adult sieve element cells of plants and red blood cells of mammals. Both scRNA-seq and snRNA-seq have their own limitations, and the overlap between scRNA-seq and snRNA-seq datasets are issues that remain to be addressed. This study shows that both scRNA-seq and snRNA-seq can produce biologically significant data, but snRNA-seq has more advantages in the preparation of single cells samples, so it is necessary to choose the appropriate method according to the purpose of the study.

Unsupervised clustering snRNA-seq transcriptome of *Arabidopsis* leaves produced two very similar, but clearly different clusters of the epidermis. Considering that DEGs of C3 were predominantly expressed in *FIL1*-marked *Arabidopsis* abaxial cells, whereas DEGs of C10 were predominantly expressed in *AS2* and *ATML*-marked cells, we hypothesized that the distinct differences between C3 and C10 could be due to their occurrence in the lower and upper epidermis, respectively. Corresponding with this hypothesis, there were several genes involved in wax, cuticle metabolic, or biosynthetic process that were highly expressed in C10, indicating that the harder cell wall of the upper epidermis could utilize them to adapt against direct sunlight. However, these identifications and analyses for these two epidermal clusters need to be verified by subsequent detailed gene expression and functional experiments. Similar studies on somewhat different cell types, such as the upper and lower epidermis, are still difficult to identify based on scRNA/snRNA-seq, and there is greater uncertainty. In recent years, new technologies such as spatiotemporal RNA-seq have made it possible to sequence total RNA *in situ* with a very high-throughput. Especially, the resolution of BGI's technology has now reached single cell and even sub-cellular levels (Xia et al., 2022; X. Liu et al., 2022), making it easier to discern the transcriptome differences between very similar sub-cell types and accelerating life science research.

Funding

This work was supported by the National Key R&D Program of China (No. 2019YFC1711000). This work is also supported by China National GeneBank (CNGB; <https://www.cngb.org/>).

CRediT authorship contribution statement

KW conceived the idea with the assistance of CZ, XG and SKS. SKS and XG supervised the study. KW performed experiments and generated data. KW analyzed the data with the assistance of CZ and SX. KW wrote the original manuscript. SKS revised and edited the manuscript. KD, XG and XC provided helpful comments on this study. All authors contributed to the article and approved the submitted version.

Data Availability Statement

The data that support the findings of this study have been deposited into CNGB Sequence Archive (CNSA) (Guo et al., 2020) of China National GeneBank DataBase (CNGBdb) (Chen et al., 2020) with accession number CNP0002469 (<https://db.cngb.org/cnsa/download/>).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

The data that support the findings of this study have been deposited into CNGB sequence Archive (CNSA) (Chen et al., 2020) of China National GeneBank DataBase (CNGBdb) with accession number CNP0002469.

Acknowledgments

We are grateful to Cong Tan, Shuqiang Chen, Shuai Yang, Tianying Zhang, Sisi Chen, Yuting Yang, Fang Wang, Dongming Fang, Feng Yang, Liyuan Zhong for general technical assistance or discussion.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:[10.1016/j.plantsci.2022.111535](https://doi.org/10.1016/j.plantsci.2022.111535).

References

- A. Alexa, J. Rahnenführer, Gene set enrichment analysis with topGO, Bioconductor Improv 27 (2009) 1–26.
- U. Ayturk, RNA-seq in skeletal biology, Curr. Osteoporos. Rep. 17 (2019) 178–185.
- T.E. Bakken, R.D. Hodge, J.A. Miller, Z. Yao, T.N. Nguyen, B. Aeversmann, E. Barkan, D. Bertagnolli, T. Casper, N. Dee, Single-nucleus and single-cell transcriptomes compared in matched cortical cell types, PLoS One 13 (2018), e0209648.
- T.E. Bakken, C.T. van Velthoven, V. Menon, R.D. Hodge, Z. Yao, T.N. Nguyen, L. T. Graybuck, G.D. Horwitz, D. Bertagnolli, J.J.E. Goldy, Single-cell and single nucleus RNA-seq uncovers shared and distinct axes of variation in dorsal LGN neurons in mice, non-human primates, and humans, Elife 10 (2021), e64875.
- E. Becht, L. McInnes, J. Healy, C.-A. Dutertre, I.W.H. Kwok, L.G. Ng, F. Ginhoux, E. W. Newell, Dimensionality reduction for visualizing single-cell data using UMAP, Nat. Biotechnol. 37 (2019) 38–44.
- M. Bonke, S. Thitamadee, A.P. Mähönen, M.T. Hauser, Y. Helariutta, APL regulates vascular tissue identity in *Arabidopsis*, Nature 426 (2003) 181–186.
- F.Z. Chen, L.J. You, F. Yang, L.N. Wang, X.Q. Guo, F. Gao, C. Hua, C. Tan, L. Fang, R. Q. Shan, W.J. Zeng, B. Wang, R. Wang, X. Xu, X.F. Wei, CNGBdb: china national GeneBank DataBase, Yi chuan = Hered. 42 (2020) 799–809.
- A.M. Clark, H.J. Bohnert, Cell-specific expression of genes of the lipid transfer protein family from *Arabidopsis thaliana*, Plant Cell Physiol. 40 (1999) 69–76.
- D. Conde, P.M. Triozzi, K.M. Balmant, A.L. Doty, M. Miranda, A. Boullosa, H.W. Schmidt, W.J. Pereira, C. Dervinis, M. Kirst, A robust method of nuclei isolation for single-cell RNA sequencing of solid tissues from the plant genus *Populus*, PLoS One 16 (2021), e0251149.
- T. Denyer, X. Ma, S. Klesen, E. Scacchi, K. Nieselt, M.C.P. Timmermans, Spatiotemporal developmental trajectories in the *Arabidopsis* root revealed using high-throughput single-cell RNA sequencing, Dev. Cell 48 (2019), 840–852.e845.
- M.W. Dorrrity, C.M. Alexandre, M.O. Hamm, A.-L. Vigil, S. Fields, C. Queitsch, J.T.J. Nc Cuperus, The regulatory landscape of *Arabidopsis thaliana* roots at single-cell resolution, Nat. Commun. 12 (2021) 1–12.
- N. Efremova, L. Schreiber, S. Bär, I. Heidmann, P. Huijser, K. Wellesen, Z. Schwarz-Sommer, H. Saedler, A. Yephremov, Functional conservation and maintenance of expression pattern of FIDDLEHEAD-like genes in *Arabidopsis* and *Antirrhinum*, Plant Mol. Biol. 56 (2004) 821–837.
- J.S. Eom, L.-Q. Chen, D. Sosso, B.T. Julius, I. Lin, X.-Q. Qu, D.M. Braun, W.B. Frommer, SWEETs, transporters for intracellular and intercellular sugar translocation, Curr. Opin. Plant Biol. 25 (2015) 53–62.
- Y. Fan, S.K. Sahu, T. Yang, W. Mu, J. Wei, L. Cheng, J. Yang, J. Liu, Y. Zhao, M. Lisby, H. Liu, The *Clausena lansium* (Wampee) genome reveal new insights into the carbazole alkaloids biosynthesis pathway, Genomics 113 (2021) 3696–3704.
- Y. Fan, S.K. Sahu, T. Yang, W. Mu, J. Wei, L. Cheng, J. Yang, R. Mu, J. Liu, J. Zhao, Y. Zhao, X. Xu, X. Liu, H. Liu, Dissecting the genome of star fruit (*Averrhoa carambola* L.), Hortic. Res. 7 (2020) 1–10.
- A. Farmer, S. Thibivilliers, K.H. Ryu, J. Schiefelbein, M. Libault, Single-nucleus RNA and ATAC sequencing reveals the impact of chromatin accessibility on gene expression in *Arabidopsis* roots at the single-cell level, Mol. Plant 14 (2021) 372–621.
- L. Gaufichon, C. Masclaux-Daubresse, G. Tcherkez, M. Reisdorf-Cren, Y. Sakakibara, T. Hase, G. Clément, J.C. Avice, O. Grandjean, A. Marmagne, S. Boutet-Mercey, M. Azzopardi, F. Soulay, A. Suzuki, *Arabidopsis thaliana* ASN2 encoding asparagine synthetase is involved in the control of nitrogen assimilation and export during vegetative growth, Plant Cell Environ. 36 (2013) 328–342.
- T.M. Gierahn, M.H. Wadsworth 2nd, T.K. Hughes, B.D. Bryson, A. Butler, R. Satija, S. Fortune, J.C. Love, A.K. Shalek, Seq-Well: portable, low-cost RNA sequencing of single cells at high throughput, Nat. Methods 14 (2017) 395–398.
- J.R. Gottwald, P.J. Krysan, J.C. Young, R.F. Evert, M.R. Sussman, Genetic evidence for the in planta role of phloem-specific plasma membrane sucrose transporters, Proc. Natl. Acad. Sci. USA 97 (2000) 13979–13984.
- R.V. Grindberg, J.L. Yee-Greenbaum, M.J. McConnell, M. Novotny, A.L. O'Shaughnessy, G.M. Lambert, M.J. Araúzo-Bravo, J. Lee, M. Fishman, G.E. Robbins, X. Lin, P. Venepally, J.H. Badger, D.W. Galbraith, F.H. Gage, R.S. Lasken, RNA-Seq. Single Nucl. 110 (2013) 19802–19807.
- X. Guo, D. Fang, S.K. Sahu, S. Yang, X. Guang, R. Folk, S.A. Smith, A.S. Chanderbali, S. Chen, M. Liu, *Chloranthus* genome provides insights into the early diversification of angiosperms, Nature, Communications 12 (2021) 1–14.
- X. Guo, F. Chen, F. Gao, L. Li, K. Liu, L. You, C. Hua, F. Yang, W. Liu, C. Peng, L. Wang, X. Yang, F. Zhou, J. Tong, J. Cai, Z. Li, B. Wan, L. Zhang, T. Yang, M. Zhang, L. Yang, Y. Yang, W. Zeng, B. Wang, X. Wei, X. Xu, CNSA: a data repository for archiving omics data, Database: J. Biol. Database. Curation (2020).

- L. Han, X. Wei, C. Liu, G. Volpe, Z. Zhuang, X. Zou, Z. Wang, T. Pan, Y. Yuan, X. Zhang, Cell transcriptomic atlas of the non-human primate *Macaca fascicularis*, *Nature* 604 (2022) 723–731.
- X. Han, R. Wang, Y. Zhou, L. Fei, H. Sun, S. Lai, A. Saadatpour, Z. Zhou, H. Chen, F. Ye, D. Huang, Y. Xu, W. Huang, M. Jiang, X. Jiang, J. Mao, Y. Chen, C. Lu, J. Xie, Q. Fang, Y. Wang, R. Yue, T. Li, H. Huang, S.H. Orkin, G.C. Yuan, M. Chen, G. Guo, Mapping the mouse cell atlas by microwell-Seq, *Cell* 172 (2018), 1091–1107.e1017.
- X. Han, Z. Zhou, L. Fei, H. Sun, R. Wang, Y. Chen, H. Chen, J. Wang, H. Tang, W. Ge, Y. Zhou, F. Ye, M. Jiang, J. Wu, Y. Xiao, X. Jia, T. Zhang, X. Ma, Q. Zhang, X. Bai, S. Lai, C. Yu, L. Zhu, R. Lin, Y. Gao, M. Wang, Y. Wu, J. Zhang, R. Zhan, S. Zhu, H. Hu, C. Wang, M. Chen, H. Huang, T. Liang, J. Chen, W. Wang, D. Zhang, G. Guo, Construction of a human cell landscape at single-cell level, *Nature* 581 (2020) 303–309.
- Y. Hao, S. Hao, E. Andersen-Nissen, S. Mauck 3rd, S. Zheng, A. Butler, M.J. Lee, A. J. Wilk, C. Darby, M. Zager, P. Hoffman, M. Stoeckius, E. Papalexi, E.P. Mimitou, J. Jain, A. Srivastava, T. Stuart, L.M. Fleming, B. Yeung, A.J. Rogers, J.M. McElrath, C.A. Blish, R. Gottardo, P. Smibert, R. Satija, Integrated analysis of multimodal single-cell data, *Cell*, 184 (2021) 3573–3587.e3529., Integrated analysis of multimodal single-cell data, *Cell* 184 (2021) 3573–3587, e3529.
- T. Hashimshony, N. Senderovich, G. Avital, A. Klochendler, Y. de Leeuw, L. Anavy, D. Gennert, S. Li, K.J. Livak, O. Rozenblatt-Rosen, Y. Dor, A. Regev, I. Yanai, CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq, *Genome Biol.* 17 (2016) 77.
- M.M. Islam, C. Tani, M. Watanabe-Sugimoto, M. Uraji, M.S. Jahan, C. Masuda, Y. Nakamura, I.C. Mori, Y. Murata, Myrosinases, TGG1 and TGG2, redundantly function in ABA and MeJA signaling in *Arabidopsis* guard cells, *Plant Cell Physiol.* 50 (2009) 1171–1175.
- K. Jean-Baptiste, J.L. McFaline-Figueroa, C.M. Alexandre, M.W. Dorrrity, L. Saunders, K. L. Bubb, C. Trapnell, S. Fields, C. Queitsch, J.T. Cuperus, Dynamics of gene expression in single root cells of *Arabidopsis thaliana*, *Plant Cell* 31 (2019) 993–1011.
- S. Kalve, J. Fotschki, T. Beeckman, K. Vissenberg, G.T. Beemster, Three dimensional patterns of cell division and expansion throughout the development of *Arabidopsis thaliana* leaves, *J. Exp. Bot.* 65 (2014) 6385–6397.
- J.-Y. Kim, E. Symeonidi, T.Y. Pang, T. Denyer, D. Weidauer, M. Bezrutczyk, M. Miras, N. Zöllner, T. Hartwig, M.M. Wudick, M. Lercher, L.-Q. Chen, M.C.P. Timmermans, W.B. Frommer, Distinct identities of leaf phloem cells revealed by single cell transcriptomics, *Plant Cell* 33 (2021) 511–530.
- A.M. Klein, L. Mazutis, I. Akartuna, N. Tallapragada, A. Veres, V. Li, L. Peshkin, D. A. Weitz, M.W. Kirschner, Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells, *Cell* 161 (2015) 1187–1201.
- M. Kubo, T. Nishiyama, Y. Tamada, R. Sano, M. Ishikawa, T. Murata, A. Imai, D. Lang, T. Demura, R. Reski, M. Hasebe, Single-cell transcriptome analysis of Physcomitrella leaf cells during reprogramming using microcapillary manipulation, *Nucleic Acids Res.* 47 (2019) 4539–4553.
- L. Kunst, S. Clemens, T. Hooker, Expression of the wax-specific condensing enzyme CUT1 in *Arabidopsis*, *Biochem. Soc. Trans.* 28 (2000) 651–654.
- H.M. Lam, M.H. Hsieh, G. Coruzzi, Reciprocal regulation of distinct asparagine synthetase genes by light and metabolites in *Arabidopsis thaliana*, *Plant J.* 16 (1998) 345–353.
- P. Lara, B. Oñate-Sánchez, Z. Abraham, C. Ferrández, I. Díaz, P. Carbonero, J. Vicente-Carbajosa, Synergistic activation of seed storage protein gene expression in *Arabidopsis* by ABI3 and two bZIPs related to OPAQUE2, *J. Biol. Chem.* 278 (2003) 21003–21011.
- S.B. Lee, Y.S. Go, H.J. Bae, J.H. Park, S.H. Cho, H.J. Cho, D.S. Lee, O.K. Park, I. Hwang, M.C. Suh, Disruption of glycosylphosphatidylinositol-anchored lipid transfer protein gene altered cuticular lipid composition, increased plastoglobules, and enhanced susceptibility to infection by the fungal pathogen *Alternaria brassicicola*, *Plant Physiol.* 150 (2009) 42–45.
- C. Liu, R. Li, Y. Li, X. Lin, K. Zhao, Q. Liu, S. Wang, X. Yang, X. Shi, Y. Ma, Spatiotemporal mapping of gene expression landscapes and developmental trajectories during zebrafish embryogenesis, *Dev. Cell* 57 (2022a), 1284–1298. e1285.
- C. Liu, T. Wu, F. Fan, Y. Liu, L. Wu, M. Junkin, Z. Wang, Y. Yu, W. Wang, W. Wei, Y. Yuan, M. Wang, M. Cheng, X. Wei, J. Xu, Q. Shi, S. Liu, A. Chen, O. Wang, M. Ni, W. Zhang, Z. Shang, Y. Lai, P. Guo, C. Ward, G. Volpe, L. Wang, H. Zheng, Y. Liu, B. A. Peters, J. Beecher, Y. Zhang, M.A. Esteban, Y. Hou, X. Xu, I.-J. Chen, L. Liu, A portable and cost-effective microfluidic system for massively parallel single-cell transcriptome profiling, *bioRxiv* (2019b), 818450.
- L. Liu, C. Liu, A. Quintero, L. Wu, Y. Yuan, M. Wang, M. Cheng, L. Leng, L. Xu, G.J. Nc Dong, Deconvolution of single-cell multi-omics layers reveals regulatory heterogeneity, *Nat. Commun.* 10 (2019a) 1–10.
- X. Liu, Y. Jiang, D. Song, L. Zhang, G. Xu, R. Hou, Y. Zhang, J. Chen, Y. Cheng, L.J.C. Liu, T. medicine, *Clin. Chall. Tissue Prep. Spat. Transcr.* 12 (2022b), e669.
- S. Ma, S. Sun, J. Li, Y. Fan, J. Qu, L. Sun, S. Wang, Y. Zhang, S. Yang, Z.J.C. Liu, Single-cell transcriptomic atlas of primate cardiopulmonary aging, *Cell Res.* 31 (2021) 415–432.
- E.Z. Macosko, A. Basu, R. Satija, J. Nemesh, K. Shekhar, M. Goldman, I. Tirosh, A. R. Bialas, N. Khamtaki, E.M. Martersteck, J.J. Trombetta, D.A. Weitz, J.R. Sanes, A. K. Shalek, A. Regev, S.A. McCarroll, Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets, *Cell* 161 (2015) 1202–1214.
- A.P. Marand, Z. Chen, A. Gallavotti, R.J. Schmitz, A cis-regulatory atlas in maize at single-cell resolution, *Cell* 184 (2021), 3041–3055. e3021.
- L. McInnes, J. Healy, J.J.A.P.A. Melville, Umap: Uniform manifold approximation and projection for dimension reduction, *arXiv* (2018).
- H. Mira, F. Martínez-García, L. Peñarrubia, Evidence for the plant specific intercellular transport of the *Arabidopsis* copper chaperone CCH, *Plant J.* 25 (2001) 521–528.
- T.K. Olsen, N. Baryawno, Introduction to single-cell RNA sequencing, *Curr. Protoc. Mol. Biol.* 122 (2018), e57.
- S. Picelli, Å.K. Björklund, O.R. Faridani, S. Sagasser, G. Winberg, R. Sandberg, Smart-seq2 for sensitive full-length transcriptome profiling in single cells, *Nat. Methods* 10 (2013) 1096–1098.
- R. Rostom, V. Svensson, S.A. Teichmann, G. Kar, Computational approaches for interpreting scRNA-seq data, *FEBS Lett.* 591 (2017) 2213–2225.
- K.H. Ryu, L. Huang, H.M. Kang, J. Schiefelbein, Single-cell RNA sequencing resolves molecular relationships among individual plant cells, *Plant Physiol.* 179 (2019) 1444–1456.
- S.K. Sahu, M. Liu, A. Yssel, R. Kariba, S. Muthemba, S. Jiang, B. Song, P.S. Hendre, A. Muchugi, R. Jamnadass, S.M. Kao, J. Featherston, N.J.C. Zerega, X. Xu, H. Yang, A.V. Deynze, Y.V. Peer, X. Liu, H. Liu, Draft genomes of two artocarpus plants, jackfruit (*A. heterophyllus*) and Breadfruit (*A. altilis*), *Genes* 11 (2019) 1–17.
- R. Satija, J.A. Farrell, D. Gennert, A.F. Schier, A.J.N. Regev, Spatial reconstruction of single-cell gene expression data 33 (2015) 495–502.
- J.W. Satterlee, J. Strable, M.J. Scanlon, Plant stem-cell organization and differentiation at single-cell resolution, *Proc Natl Acad Sci* 117 (2020) 33689–33699. U.S.A.
- R. Shaw, X. Tian, J. Xu, Single-cell transcriptome analysis in plants: advances and challenges, *Mol. Plant* 14 (2021) 115–126.
- M. Shirakawa, H. Ueda, A.J. Nagano, T. Shimada, T. Kohchi, I. Hara-Nishimura, FAMA is an essential component for the differentiation of two distinct cell types, myrosin cells and guard cells, in *Arabidopsis*, *Plant Cell* 26 (2014) 4039–4052.
- D.Y. Sunaga-Franze, J.M. Muino, C. Braeuning, X. Xu, M. Zong, C. Smacznak, W. Yan, C. Fischer, R. Vidal, M. Klemi, Single-nucleus RNA sequencing of plant tissues using a nanowell-based system, *Plant J.* (2021).
- S. Takada, N. Takada, A. Yoshida, ATML1 promotes epidermal cell differentiation in *Arabidopsis* shoots,, *Development* 140 (2013) 1919–1923 (Cambridge, England).
- H. Takahashi, A. Watanabe-Takahashi, F.W. Smith, M. Blake-Kalff, M.J. Hawkesford, K. Saito, The roles of three functional sulphate transporters involved in uptake and translocation of sulphate in uptake and translocation of sulphate in *Arabidopsis thaliana*, *Plant J.* 23 (2000) 171–182.
- S. Thibivilliers, D. Anderson, M. Libault, Isolation of Plant Root Nuclei for Single Cell RNA Sequencing, *Curr. Protoc. Plant Biol.* 5 (2020), e20120.
- C. Tian, Q. Du, M. Xu, F. Du, Y. Jiao, Single-nucleus RNA-seq resolves spatiotemporal developmental trajectories in the *tomato* shoot apex, *bioRxiv* (2020), e20120.
- K. Uemoto, T. Araki, M. Endo, Isolation of *arabidopsis* palisade and spongy mesophyll cells, *Methods Mol. Biol.* 2018 (1830) 141–148.
- S. Wang, H. Liang, H. Wang, L. Li, Y. Xu, Y. Liu, M. Liu, J. Wei, T. Ma, C. Le, J. Yang, C. He, J. Liu, J. Zhao, Y. Zhao, M. Lisby, S. K. Sahu, H. Liu, The chromosome scale genomes of Dipterocarpus turbinatus and Hopea hainanensis (Dipterocarpaceae) provide insights into fragrant oleoresin biosynthesis and hardwood formation, *Plant Biotechnol. J.* 20 (2022) 538–553.
- F. Weltmeier, F. Rahmani, A. Ehrlert, K. Dietrich, K. Schütze, X. Wang, C. Chaban, J. Hansson, M. Teige, K. Harter, J. Vicente-Carbojosa, S. Smeekens, W. Dröge-Laser, Expression patterns within the *Arabidopsis* C/S1 bZIP transcription factor network: availability of heterodimerization partners controls gene expression during stress response and development, *Plant Mol. Biol.* 69 (2009) 107–119.
- F.H. Wu, S.C. Shen, L.Y. Lee, S.H. Lee, M.T. Chan, C.S. Lin, Tape-Arabidopsis Sandwich - a simpler *Arabidopsis* protoplast isolation method, *Plant Methods* 5 (2009) 16.
- K. Xia, H.-X. Sun, J. Li, J. Li, Y. Zhao, L. Chen, C. Qin, R. Chen, Z. Chen, G. Liu, R. Yin, B. Mu, X. Wang, M. Xu, X. Li, P. Yuan, Y. Qiao, S. Hao, J. Wang, Q. Xie, J. Xu, S. Liu, Y. Li, A. Chen, L. Liu, Y. Yin, H. Yang, J. Wang, Y. Gu, X. Xu, The single-cell stereoseq reveals region-specific cell subtypes and transcriptome profiling in *Arabidopsis* leaves, *Dev. Cell* 57 (2022), 1299–1310.e1294.
- T. Yang, S.K. Sahu, L. Yang, Y. Liu, W. Mu, X. Liu, M.L. Strube, H. Liu, B. Zhong, Comparative analyses of 3,654 plastid genomes unravel insights into evolutionary dynamics and phylogenetic discordance of green plants, *Front. Plant Sci.* 13 (2022).
- W. Zeng, S. Jiang, X. Kong, N. El-Ali, A.R. Ball Jr, C.I.-H. Ma, N. Hashimoto, K. Yokomori, A. Mortazavi, Single-nucleus RNA-seq of differentiating human myoblasts reveals the extent of fate heterogeneity, *Nucleic Acids Res.* 44 (2016) e158–e158.
- T.Q. Zhang, Z.G. Xu, G.D. Shang, J.W. Wang, A. Single-Cell, RNA Sequencing Profiles the Developmental Landscape of *Arabidopsis* Root, *Mol. Plant* 12 (2019) 648–660.
- T.Q. Zhang, Y. Chen, Y. Liu, W.H. Lin, J.W. Wang, Single-cell transcriptome atlas and chromatin accessibility landscape reveal differentiation trajectories in the rice root, *Nat. Commun.* 12 (2021b) 2053.
- T.-Q. Zhang, Y. Chen, J.-W. Wang, A single-cell analysis of the *Arabidopsis* vegetative shoot apex, *Dev. Cell* 56 (2021a), 1056–1074. e1058.
- G.X. Zheng, J.M. Terry, P. Belgrader, P. Ryvkin, Z.W. Bent, R. Wilson, S.B. Ziraldo, T. D. Wheeler, G.P. McDermott, J. Zhu, M.T. Gregory, J. Shuga, L. Montesclaros, J. G. Underwood, D.A. Masquelier, S.Y. Nishimura, M. Schnall-Levin, P.W. Wyatt, C. M. Hindson, R. Bharadwaj, A. Wong, K.D. Ness, L.W. Beppu, H.J. Deeg, C. McFarland, K.R. Loeb, W.J. Valente, N.G. Ericson, E.A. Stevens, J.P. Radich, T. S. Mikkelsen, B.J. Hindson, J.H. Bielas, Massively parallel digital transcriptional profiling of single cells, *Nat. Commun.* 8 (2017) 14049.
- J. Zhu, F. Chen, L. Luo, W. Wu, J. Dai, J. Zhong, X. Lin, C. Chai, P. Ding, L. Liang, S. Wang, X. Ding, Y. Chen, H. Wang, J. Qiu, F. Wang, C. Sun, Y. Zeng, J. Fang, X. Jiang, P. Liu, G. Tang, X. Qiu, X. Zhang, Y. Ruan, S. Jiang, J. Li, S. Zhu, X. Xu, F. Li, Z. Liu, G. Cao, D. Chen, Single-cell atlas of domestic pig cerebral cortex and hypothalamus, *Sci. Bull.* 66 (2021) 1448–1461.