Wang *et al. Plant Methods* (2026) 22:12
https://doi.org/10.1186/s13007-025-01490-6

Plant Methods

RESEARCH Open Access

Integrated experimental and computational workflows for single-cell transcriptomics in plants

Jing Wang[1†], Shanqiao Zheng[2†], Bojie Lu[2†], Yuan Jiang[1], Yabing Zhu[2], Qun Liu[1], Song Gao[1], Peng Liu[2], Peng Yu[3], Sanjie Jiang[2*] and Liang Zong[1*]

Abstract

Background Single-cell transcriptomics is a powerful approach to resolve cellular heterogeneity, yet its application in plants is constrained by challenges in tissue preparation, nuclei isolation, and transcriptome quality. Optimized experimental and computational workflows are essential to achieve robust results in plant systems.

Results We systematically benchmarked bulk and single-cell transcriptomic workflows in maize and established an integrated, optimized framework. First, we developed an improved bulk RNA-seq protocol, providing higher consistency and serving as a reference for single-cell datasets. Second, we compared three input types, protoplasts, fresh nuclei, and frozen nuclei, across tissues, demonstrating overall comparability of their transcriptomic profiles and offering guidance for studies with limited material. Third, by leveraging bulk RNA-seq as a reference, these complementary data provide additional biological context that helps to interpret and validate findings derived from single-cell transcriptomic analyses. A combination of these strategies resulted in high transcriptome integrity and clear clustering resolution in the final dataset, supporting robust identification of plant cell types. While all experimental data are derived from maize, the principles and strategies described here provide practical guidance and inspiration for single-cell studies in other plant species.

Conclusions Our study establishes optimized experimental and computational workflows for plant single-cell transcriptomics. By validating input comparability and addressing the limitations of nuclear data, we provide methodological guidance that extends beyond maize and supports future single-cell investigations across diverse plant species.

Keywords Single-cell RNA sequencing, Plant transcriptomics, Workflow optimization

[†]Jing Wang, Shanqiao Zheng and Bojie Lu have contributed equally to this work.

*Correspondence:
Sanjie Jiang
jiangsanjie@bgi.com
Liang Zong
zongliang@bgi.com
[1]BGI Tech Solutions, BGI Genomics, Wuhan 430075, China
[2]BGI Tech Solutions, BGI Genomics, Shenzhen 518083, China
[3]Plant Genetics, TUM School of Life Sciences, Technical University of Munich, 85354 Freising, Germany

© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

Introduction

Understanding plant development and environmental responses requires high-resolution insights into cellular heterogeneity and gene regulation. Single-cell transcriptomic technologies, including single-cell RNA sequencing (scRNA-seq) and single-nucleus RNA sequencing (snRNA-seq), have transformed plant biology by enabling cell-type-specific gene expression profiling [1].

Protoplast-based scRNA-seq has been instrumental in generating cellular atlases, revealing developmental trajectories, and identifying key regulatory networks in *Arabidopsis thaliana*, *Oryza sativa* and *Zea mays* [2]. Despite its advantages in capturing full-picture of transcriptomes, protoplast isolation is often limited by stress-induced artifacts, low cell recovery, and reduced applicability to lignified tissues. In contrast, snRNA-seq provides access to recalcitrant tissues while preserving transcriptional states with minimal perturbation. However, it often yields fewer detected genes per nucleus and underrepresent cytoplasmic transcripts, affecting resolution for certain biological processes [3]. Despite their complementarity, direct comparisons of these methodologies in plants remain limited, leaving open questions about their respective advantages and biases in different tissues and conditions.

In mammalian tissues, scRNA-seq and snRNA-seq have shown systematic biases [4–7] and led to practical guidelines for selecting the appropriate method based on tissue type and experimental context [8–11]. As a major crop and genetic model, maize serves as an ideal system to refine and benchmark plant single-cell methodologies. Root development, which underlies nutrient uptake and stress resilience, and leaf architecture, essential for photosynthesis and gas exchange, have been extensively explored using single-cell approaches [12, 13]. In this study, we present a systematic comparison of protoplast- and nucleus-based single-cell transcriptomic methods in maize root and leaf tissues (Fig. 1). We characterize biases introduced during sample preparation, benchmark optimized bulk and pseudo-bulk sequencing strategies, and integrate data across platforms. Specifically, we (i) evaluate the advantages and limitations of each method, (ii) generate high-resolution transcriptional maps of maize leaf and root tissues, and (iii) provide practical guidance for optimizing single-cell transcriptomic workflows in plants. Together, these results offer insights into tissue-specific transcriptome profiling and establish the framework for expanding single-cell applications in plant biology.

Results

Preparation of single-cell protoplast and nucleus suspensions

Apical sections of primary roots and central portions of leaf 2 were collected from maize inbred line B73 (Fig. 2a). Protoplasts were isolated following previously established protocols [14]. For root sections, L-cysteine pretreatment was applied to enhance protoplast yield and improve recovery of meristematic cells. In leaf samples, the presence of chloroplasts enables direct visual monitoring of protoplast isolation. Under brightfield microscopy, viable protoplasts displayed a round, intact morphology with clearly defined plasma membranes (Fig. 2b, c). The viability can be evaluated under fluorescence microscopy using live/dead staining, and a stringent viability threshold could be critical for obtaining high-quality sequencing data (Fig. 2d, e).

Fresh nuclei were isolated using the 'chopping method' [15], in which unfrozen tissues were finely minced on ice with razor blades during lysis (Fig. 3a, b). For frozen nuclei, flash-frozen tissues were ground into fine powder in liquid nitrogen prior to lysis [16]. In leaf samples,

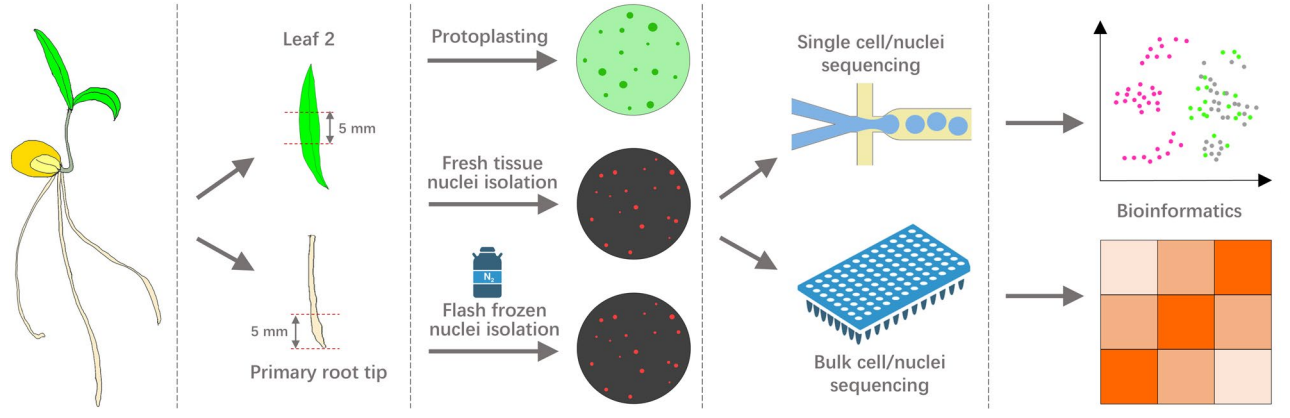

Fig. 1 Schematic overview of the experimental design. Samples were collected from maize B73 seedlings, then processed using three parallel strategies: protoplasting, fresh tissue nuclei isolation, and flash-frozen nuclei isolation. The resulting single cells or nuclei were subjected to either single-cell/nucleus RNA sequencing or bulk RNA sequencing. Bioinformatic analyses were performed to evaluate transcriptome profiles and assess technical variation introduced by different sample preparation methods

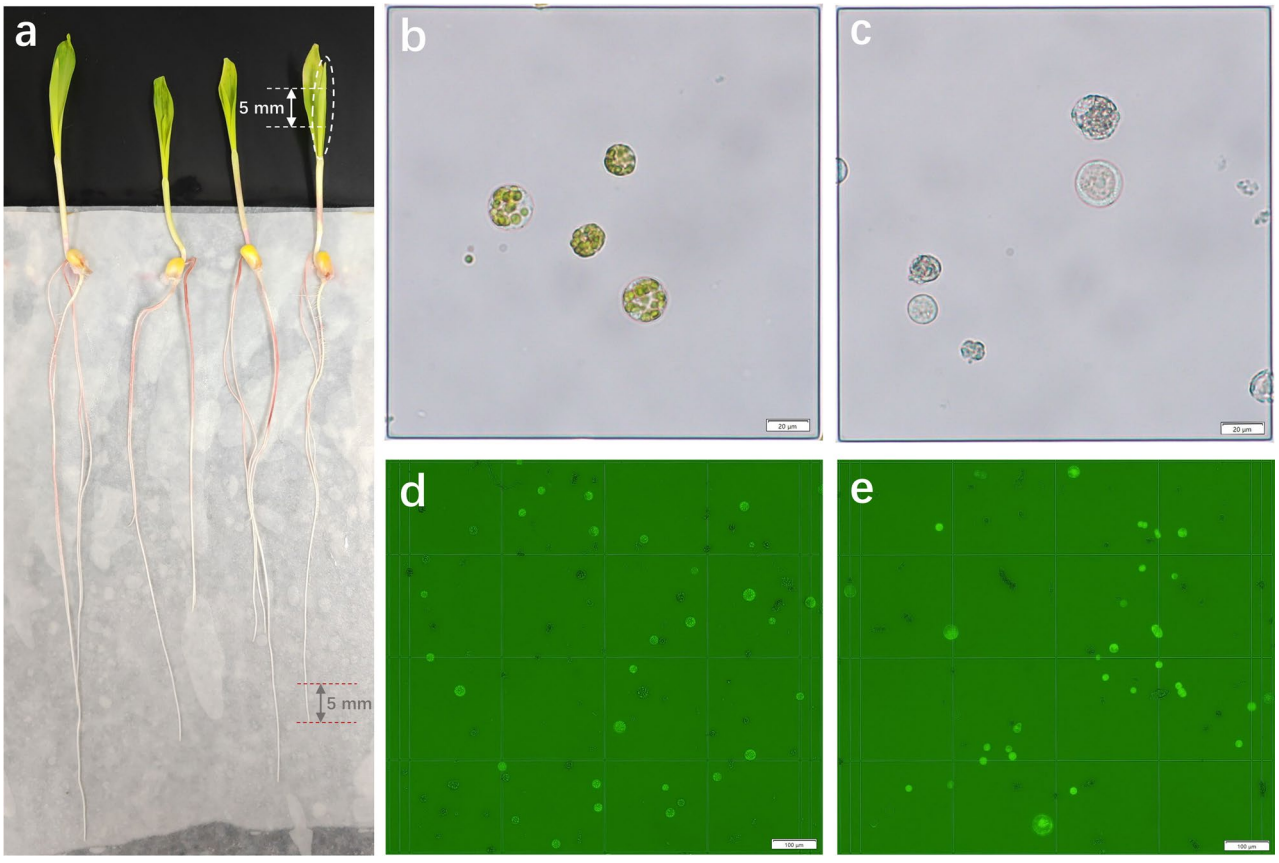

Fig. 2 Preparation of protoplast suspensions. **a** Maize seedlings were dissected to collect the central segments of leaf 2 (highlighted with a dashed outline) and the primary root tip. **b**, **c** Protoplasts of leaf (**b**) and root (**c**) observed under bright field. Scale bar = 20 µm. **d**, **e** Protoplasts of leaf (**d**) and root (**e**) counting under fluoresce field. Scale bar = 100 µm

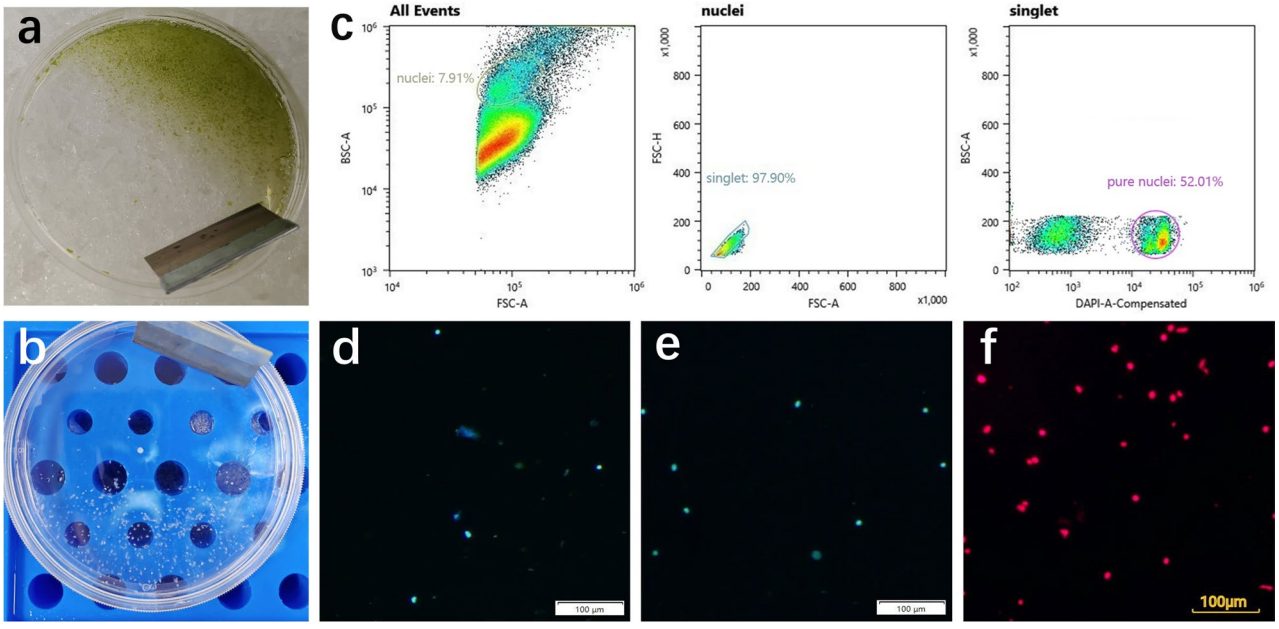

Fig. 3 Preparation of nucleus suspensions. **a**, **b** The chopping procedure of leaf (**a**) and root (**b**) samples on ice, respectively. **c** The gating strategy for purification of leaf fresh nuclei. **d**, **e** Leaf fresh nuclei observed under fluoresce field before (**d**) and after (**e**) FANS, respectively. Scale bar = 100 µm. **f** Nuclei counting using a fluoresce cell counter

a considerable amount of debris was observed in the nuclear suspensions under fluorescence microscopy, possibly due to a higher degree of lignification. To obtain high-purity nuclei with intact morphology, fluorescence-activated nuclei sorting (FANS) was applied to eliminate debris, incomplete nuclei, and aggregates (Fig. 3c, e). For root nuclei, density gradient centrifugation was used instead [17, 18]. All final suspensions were evaluated using an automated cell counter to assess concentration and quality (Fig. 3f).

The optimization of bulk sequencing workflows

Bulk RNA sequencing plays a critical role in single-cell transcriptomics by serving as a benchmark for quality control, normalization, and data validation. Because scRNA-seq suffers from technical challenges such as dropout effects and incomplete transcript capture, bulk RNA-seq provides a comprehensive reference to distinguish true biological signals from artifacts [19]. Additionally, bulk sequencing helps identify cell-type-specific biases, distinguishing between true biological variation and technical noise [20]. Integrating bulk RNA-seq improves normalization and batch correction, yet several input sources can be used to generate bulk RNA-seq datasets, including: (i) RNA extracted from the same tissue used in scRNA-seq, (ii) RNA extracted from the same tissue but specific regions enriched or removed, and (iii) RNA extracted from protoplast suspensions [21]. Library construction can proceed via poly(A)-primed full-length cDNA synthesis or random-primed fragmented cDNA generation, followed by amplification and sequencing. However, each of these approaches has limitations: (i) RNA extraction through mechanical grinding may introduce stress-related transcriptional artifacts not present in scRNA workflows, (ii) differences in library preparation chemistry between bulk and scRNA-seq may introduce systematic biases, and (iii) workflows utilizing isolated nuclei as RNA sources remain underexplored, representing a gap in nuclei-based bulk transcriptomics.

In this study, we established an optimized bulk RNA-seq workflow compatible based on SMARTer-seq3 [22] and FLASH-seq [23] protocols (Supplemental File). Bulk libraries were generated directly from input of protoplasts, fresh nuclei, or frozen nuclei, matched to their respective single-cell datasets. The protocol was optimized with the following key modifications: (i) Triton X-100 in the lysis buffer was replaced with guanidine hydrochloride (GuHCl) to inhibit nucleases and reduce interference from plant-derived secondary metabolites during reverse transcription, (ii) reverse transcription and cDNA amplification were combined into a single-step reaction to minimize RNA degradation and template loss, resulting in improved cDNA yield (Table S1), and (iii) the quality of cDNA was assessed using Bioanalyzer traces to evaluate full-length transcript coverage, providing a more accurate measure than RNA integrity number (RIN), which is not applicable to nuclei samples lacking ribosomal RNA (Figure S1). The optimized bulk protocol eliminates the need for a separate RNA extraction step and closely mirrors the chemistry of common scRNA-seq protocols [24]. By aligning preparation steps, it minimizes technical discrepancies between bulk and single-cell datasets and enhances the reliability of downstream comparisons.

The statistics of bulk sequencing results

Approximately 200 cells or nuclei were used for bulk sequencing, the cDNA profiles are shown in Figure S1. cDNA generated from protoplasts closely resembles the high-integrity control sample, Universal Human Reference RNA (UHRR), featuring a smooth baseline below 300 bp and a distinct peak around 1500 bp, indicating minimal RNA degradation during protoplast isolation and cDNA synthesis. In contrast, cDNA derived from nuclei exhibits a major peak below 1000 bp, reflecting partial RNA degradation, along with visible amplification products between 200 and 600 bp. These shorter fragments suggest residual ribosomal RNA contamination and amplification artifacts, including incomplete reverse transcription and byproducts containing the template-switching oligo (TSO). These features are indicative of inefficient mRNA template utilization, which is already 10- to 100-fold lower in nuclei compared to cytoplasm [3]. The rRNA read proportions from bulk RNA-seq are summarized in Table S2. Consistent with the Bioanalyzer traces, nuclei-derived reads exhibit rRNA contamination ranging from 5.44% to 19.58%, which is significantly higher than expected, even in samples from leaf nuclei purified via FANS. However, FANS appears to outperform gradient centrifugation, as root nuclei purified by the latter exhibit higher rRNA levels despite minimal visible debris under microscopy. In leaf nuclei data, chloroplast rRNA accounts for approximately 30% of total rRNA reads, while it is absent in root samples. This observation aligns with the anticipated organelle contamination during nuclei isolation that RNA-containing organelles or cell debris may adhere to intact nuclei through molecular interactions, as extracellular DNA release has been reported to facilitate aggregation of nuclei, potentially via electrostatic interactions or binding with nuclear proteins, and further supports the reliability of the statistical pipeline. These aggregates likely exhibit stronger signals on flow cytometers, making them susceptible to be retained and recovered during FANS. Alternative purification strategies, such as antibody-based staining against nuclear pore complex (NPC) proteins in FANS, have been reported to yield average mitochondrial read percentages of 6.74% in nuclei

datasets, indicating incomplete removal of cytoplasmic RNA and offering no substantial advantage over standard dead/live staining methods [5].

Read alignment statistics for bulk RNA-seq data are summarized in Table 1. After removing rRNA reads, all samples exhibited high genome mapping rates, exceeding 90%. When assessing mapping ratio to genes and exonic regions, protoplast-derived samples showed consistently higher mapping ratios than nuclei-derived samples, while frozen nuclei outperformed fresh nuclei. Across tissue types, root samples exhibited significantly lower gene and exon mapping rates than leaf samples. This discrepancy may reflect the increased abundance of secondary metabolites and active enzymatic processes in root tips due to ongoing differentiation and development, which can interfere with RNA integrity and library construction [25, 26]. Protoplasts detected the lowest gene numbers; however, when assessing gene abundance, protoplasts identified a higher number of genes with RPKM > 1, whereas nuclei captured more low-abundance genes with RPKM ≤ 1 (Figure S2), consistent with observations reported in mammalian cells [10]. Both root and leaf fresh nuclei exhibited the greatest number of low-abundance transcripts and the fewest highly expressed genes (RPKM ≥ 10), suggesting a reduction of highly abundant genes during nuclei isolation and an overall decrease in gene expression specifically in the fresh nuclei isolation process. Alignment quality and transcript coverage were visualized in Fig. 4 and Figure S3. Consistent with the cDNA Bioanalyzer profiles, protoplast-derived libraries exhibited even read distribution across gene bodies, with most transcripts recovering exceeding 85% of their full length (Fig. 4a, b). In contrast, nuclei-derived data recovered less than 50% of the full length for most transcripts and required twice the number of reads to reach saturation, due to divergences in transcript abundance and complexity between nucleus and cytoplasm (Fig. 4c). The pronounced bias toward the 3' end of transcripts further confirmed RNA degradation during the nuclei isolation process, which would negatively affect the studies focused on alternative splices or isoform-specific expressions [19].

Principal component analysis (PCA) of gene expression profiles revealed a clear separation between root and leaf samples, while datasets from different preparation methods within the same tissue clustered closely together (Fig. 4d). Spearman correlation heatmaps confirmed the overall comparability of datasets across preparation methods (Fig. 4e, f). Comparisons between protoplast and nuclei data showed higher similarity in root samples than in leaf samples. Among all single-cell isolation strategies, libraries from fresh nuclei consistently showed the lowest intra-group correlation, suggesting that this method may lead to reduced technical reproducibility.

Gene Ontology (GO) enrichment analysis revealed that GO terms related to chloroplast thylakoids and photosystems were enriched when comparing leaf protoplasts and nuclei, while terms associated with osmotic stress response and response to chemicals were enriched in the comparison between root protoplasts and nuclei, findings that align with the expected tissue-specific physiological roles and the experimental design [2]. Meanwhile, differentially expressed genes (DEGs) between fresh and frozen nuclei were enriched not only in stress response but also in DNA damage stimulus response and chromosome organization and maintenance, suggesting that the strategy for nuclei isolation significantly affected expression profiles and should be carefully considered when addressing specific biological questions (Figure S4).

When transcription factor (TF) expression profiles were extracted and analyzed via heatmap clustering, surprisingly, the datasets grouped primarily by experimental method of protoplasts, fresh nuclei, and frozen nuclei, rather than by tissue origin (Figure S5). Although attempts to identify overrepresented TF families yielded no meaningful enrichment, this clustering demonstrated the strong influence of cell isolation method on TF expression profiles. While the exact mechanisms remain unclear, these findings highlight the critical importance

Table 1 Summary statistics of bulk RNA-seq data

Samples	Number of clean reads		Genome mapping ratio %		Gene mapping ratio %		Exon mapping ratio %		Number of genes detected	
	Leaf	Root	Leaf	Root	Leaf	Root	Leaf	Root	Leaf	Root
fresh_nuclei_b1	104,404,106	103,972,401	94.5	90.36	49.87	41.73	34.87	27.3	36,277	36,414
fresh_nuclei_b2	104,083,710	104,293,662	92.51	89.79	50.95	42.98	35.95	27.37	36,547	36,116
fresh_nuclei_b3	104,407,006	103,173,971	92.81	90.45	51.44	42.08	37.43	27.8	36,170	36,051
frozen_nuclei_b1	103,905,516	105,401,927	92.05	93.58	62.27	66.59	46.1	51.25	33,967	31,796
frozen_nuclei_b2	102,616,680	102,832,843	93.16	93.96	66.30	62.26	49.67	47.8	34,233	30,556
frozen_nuclei_b3	102,716,799	103,405,605	92.34	93.56	63.99	65.11	47.22	49.58	31,882	30,414
protoplasts_b1	101,191,437	106,843,448	95.29	94.94	68.52	78.72	53.47	67.94	28,271	28,744
protoplasts_b2	104,814,821	104,992,919	96.86	96.03	74.02	77.60	59.07	66.13	27,891	28,130
protoplasts_b3	103,811,053	105,582,367	96.69	95.62	73.52	77.31	59.25	66.62	27,504	28,887

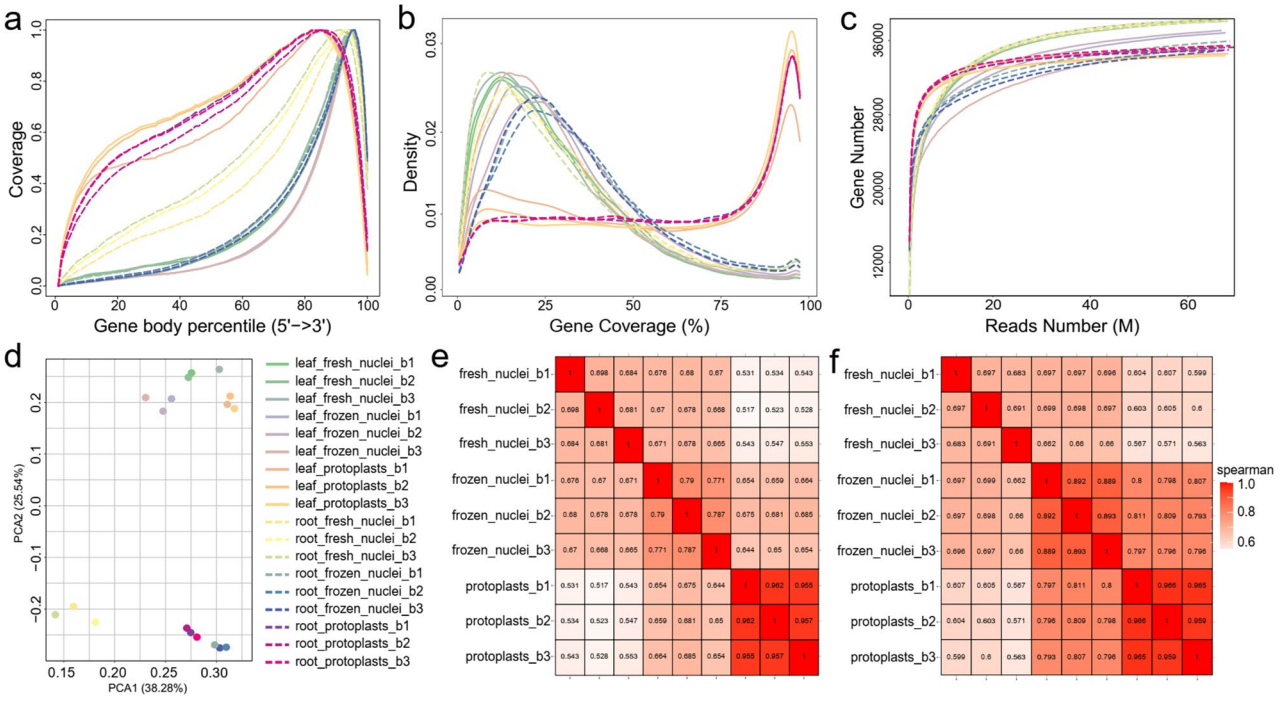

Fig. 4 Assessment of bulk RNA-seq libraries across input types. **a** Gene body coverage curves from 5′ to 3′. **b** Distribution of gene coverage density across input types. **c** Saturation curves showing the number of genes detected versus sequencing depth. **d** PCA projection based on gene expression by input type and tissue origin. **e**, **f** Spearman correlation heatmaps based on gene expression for leaf (**e**) and root (**f**) samples, respectively

Table 2 Summary statistics of single-cell RNA-seq data

Samples	Number of clean reads	Number of cells	Fraction of reads in cells %	Reads mapped to transcriptome %	Mean UMI count per cell	Mean genes per cell	Total genes detected
leaf_fresh_nuclei_b1	624,710,926	4,607	64.03	41.80	1,013	751	35,469
leaf_fresh_nuclei_b2	609,103,491	3,490	63.92	41.40	1,083	790	35,163
leaf_frozen_nuclei_b1	590,418,043	3,166	56.67	59.10	2,403	1,659	28,712
leaf_frozen_nuclei_b2	702,503,944	3,286	59.95	59.50	2,177	1,566	28,834
leaf_protoplasts_b1	673,310,873	5,092	73.48	71.70	8,493	2,967	31,799
leaf_protoplasts_b2	598,000,357	5,177	73.78	72.10	7,966	2,904	30,993
root_fresh_nuclei_b1	713,079,681	2,982	43.93	55.30	3,231	1,838	29,862
root_fresh_nuclei_b2	876,663,236	2,231	43.20	57.30	2,938	1,724	28,958
root_frozen_nuclei_b1	571,989,237	3,523	39.15	58.90	5,397	3,051	32,697
root_frozen_nuclei_b2	742,919,146	3,245	41.16	58.40	6,726	3,553	33,380
root_protoplasts_b1	613,294,046	3,301	80.67	78.30	20,372	4,022	31,134
root_protoplasts_b2	657,382,825	3,120	82.73	78.80	23,682	4,260	31,623

of methodological consistency and careful interpretation when evaluating transcriptional regulation.

The construction of the single-cell atlas

The alignment statistics of the scRNA-seq data are summarized in Table 2. Although approximately 12,000 protoplasts or nuclei were loaded during library preparation, the recovery rate varied across groups, ranging from 18.6% to 43.1%. Protoplasts consistently outperformed nuclei across most quality metrics, with the exception of the total number of genes detected, a pattern also observed in the bulk RNA-seq results. Key indicators of data quality, such as the fraction of reads mapped to cells and the fraction of reads aligned to the transcriptome, further highlighted that nuclei-derived libraries exhibited higher proportions of uninformative reads, contributing greater background noise and potential contaminants [12]. Fresh nuclei samples yielded a substantially lower average number of genes detected per cell, which aligns with values reported in some previous studies [15], this reduction in sensitivity may lead to the loss of important transcriptomic information.

Matrix construction, integration, dimensionality reduction, and cell-type annotation were performed separately for leaf and root datasets. Protoplasts tend to clusters distinct from nuclei, indicating that input material has a substantial impact on gene expression profiles, and that cross-sample normalization and careful handling of batch effects are essential when comparing data generated from different input types (Figure S6a, b).

Leaf cells were integrated and grouped into 16 clusters (Figure S6c), which were annotated into four major cell types: pavement cells (PC), mesophyll cells (MC), bundle sheath cells (BSC), and companion cells (CC) (Fig. 5a), based on well-established and experimentally validated marker genes (Table S3). Cell-type proportions varied markedly across input methods. MCs dominated in protoplast-derived data (91.2%), consistent with previous studies. In contrast, nuclei-derived samples showed reduced MC proportions (62.1% in fresh and 52.0% in frozen nuclei) and a notable increase in BSCs (20.3% in fresh and 32.5% in frozen nuclei, compared to 0.77% in protoplasts) (Fig. 5b, c). The underrepresentation of BSCs in protoplast datasets is likely due to their larger size (~ 32 μm width, 44 μm length) [27], which reduces droplet recover efficiency. Cryosectioning of the corresponding region of leaf 2 confirmed that nuclei data more accurately reflected in vivo cell composition (Fig. 5e). Additionally, BSCs are tightly encased by MCs and possess thickened cell walls, which may hinder protoplast recovery more than nuclei isolation. Given the critical roles of BSCs in C4 photosynthesis and other biological

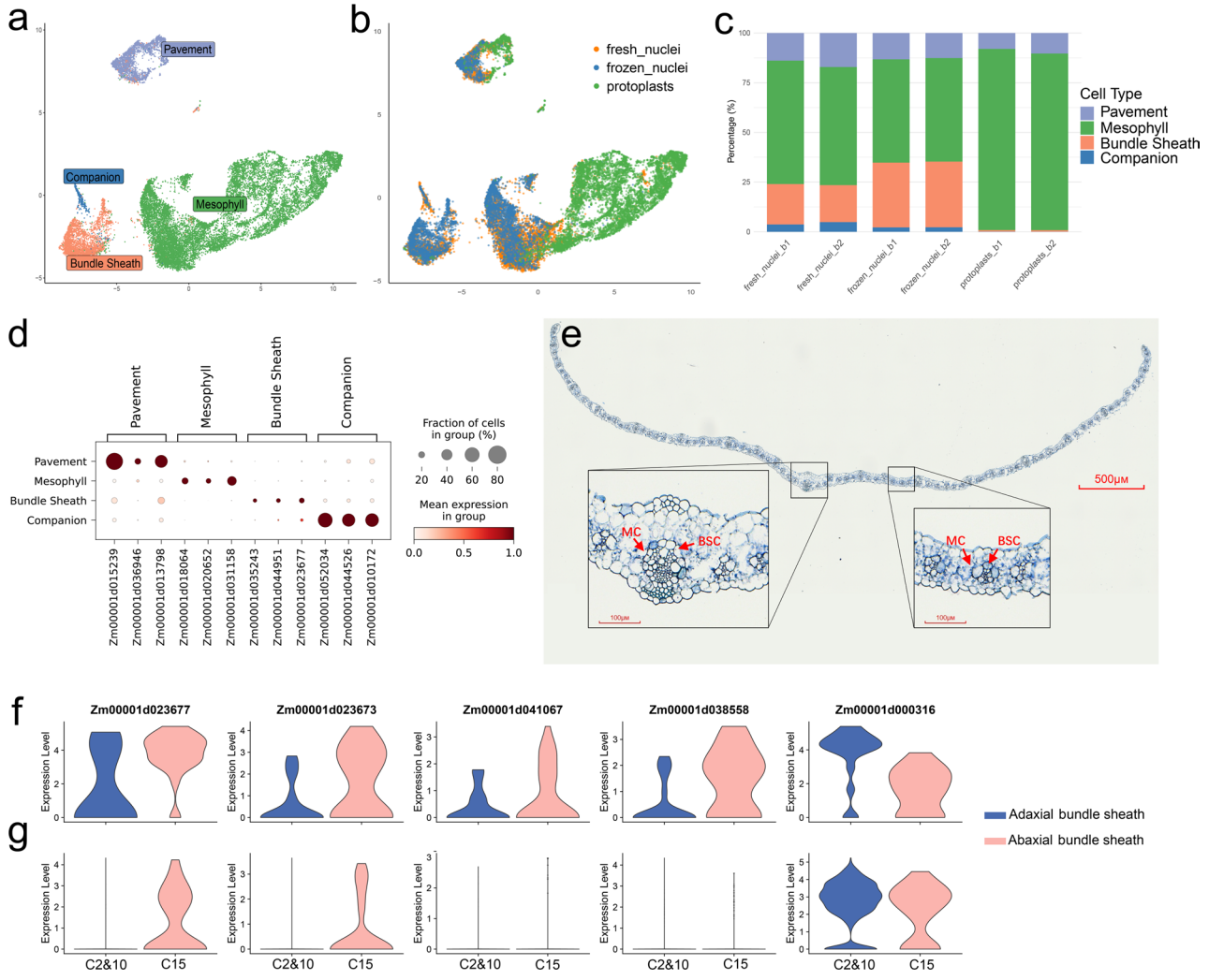

Fig. 5 Cell type annotation and distribution of leaf single-cell transcriptomes. **a** UMAP visualization of leaf cells classified into four major cell types based on known marker genes: Pavement, Mesophyll, Bundle Sheath and Companion. **b** UMAP projection colored by input type of fresh nuclei, frozen nuclei, and protoplasts. **c** Bar plots showing the proportion of each cell type across biological replicates and input types. **d** Dot plot showing the expression of representative marker genes across cell types, where dot size indicates the percentage of cells expressing the gene and color intensity represents the average expression level. **e** Bright field microscopy image of leaf 2 sections. **f**, **g** Violin plots comparing the expression levels of five marker genes between adaxial (blue) and abaxial (pink) bundle sheath cells, derived from protoplast (**f**) and nuclei (**g**) datasets, respectively

processes, the inclusion of snRNA-seq data is essential for comprehensive characterization of leaf samples.

A subset of mesophyll cells (cluster 6, 8 and 9) was uniquely detected in protoplast datasets and expressed high levels of mesophyll markers along with three additional genes, *Zm00001d014116*, *Zm00001d034562*, and *Zm00001d028692*, previously reported with protoderm identity [28]. Given the rarity of protoderm cells in the central section of leaf 2 and the stem-like expression characteristics of these genes, this subgroup was annotated as an early developmental subset of mesophyll cells. Both adaxial and abaxial BSC clusters were identified across protoplast and nuclei datasets, marked by expression of *RBCS1* (*Zm00001d052595*), with abaxial BSCs (abBSC) further distinguished by the presence of *SWEET13* (*Zm00001d023677*, *Zm00001d023673*, *Zm00001d041067*) and *CC3* (*Zm00001d038558*) [29]. However, both the proportion of cells expressing these markers and their expression levels were markedly decreased in nuclei-derived datasets (Fig. 5f, g). These findings emphasize the importance of method selection based on target cell type and required sensitivity for downstream applications.

Root cells were integrated and grouped into 23 clusters (Figure S6d), which were annotated into nine major cell types: Epidermis cells, Cortex cells, Endodermis cells, Pericycle cells, Stele cells, Phloem cells, Xylem cells, Meristematic cells and Root cap cells (Fig. 6a). In protoplast-derived data, outer root layers were more efficiently recovered (Fig. 6b, c). A subset of cells (cluster 3 and 10, predominately from nuclei libraries) was labeled as "undetermined" due to the lack of clear and reliable marker gene expression associated to known cell identities (Fig. 6d). This observation aligns with findings from published single-nucleus atlases of maize [30], rice [31] and *Arabidopsis thaliana* [32], highlighting the limited sensitivity of snRNA-seq in plants samples. In previous studies, such ambiguous clusters have been referred to as "mixing droplets", as they often display gene expression profiles suggestive of multiple or conflicting cell identities. The elevated proportion of undetermined cells in frozen nuclei suggests that mechanical disruption during sample processing may compromise nuclear integrity and contribute to transcriptomic ambiguity. Root cap cells were exclusively recovered in nuclei-derived datasets. Based on known columella marker genes (*Zm00001d004089*, *Zm00001d008268*, *Zm00001d025773*, *Zm00001d029546* and *Zm00001d034644*) [33], root cap cells could be further classified into lateral root cap and columella subgroups (Fig. 6e). These results demonstrate the importance of integrating both protoplast- and nuclei-based

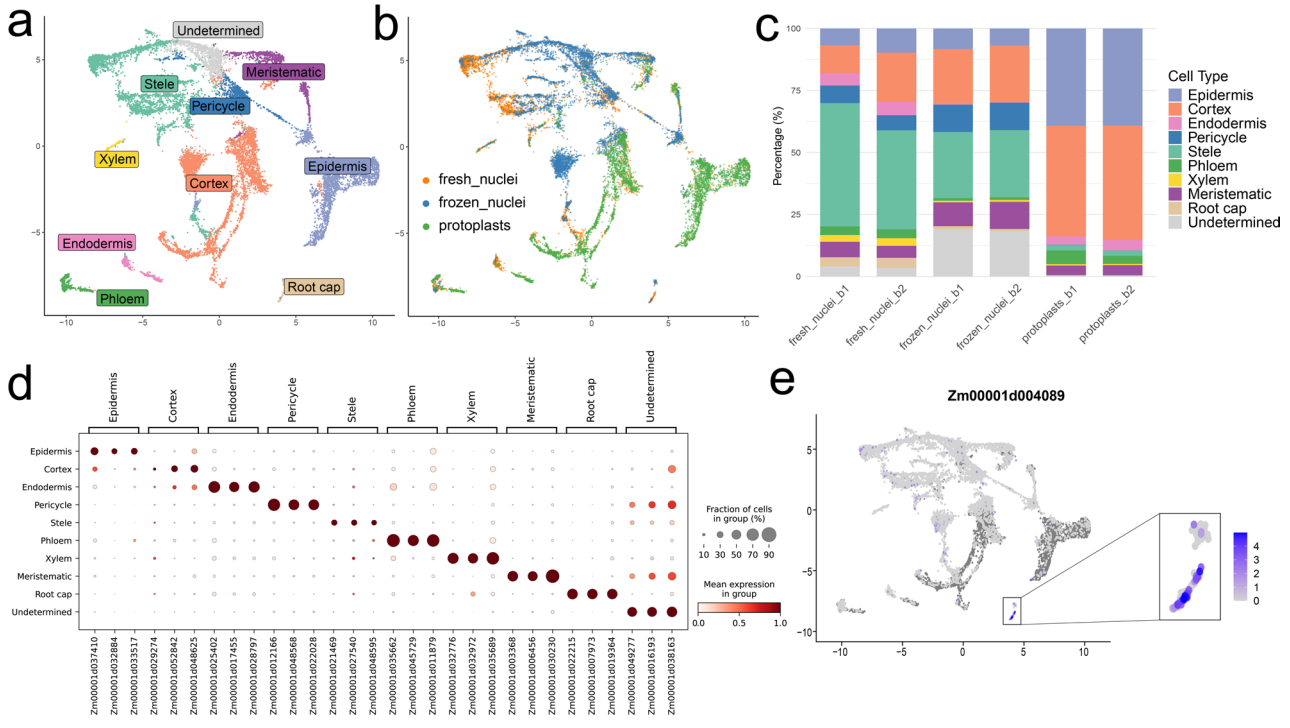

Fig. 6 Cell type annotation and distribution of root single-cell transcriptomes. **a** UMAP visualization of root cells classified into nine major cell types based on known marker genes: Epidermis, Cortex, Endodermis, Pericycle, Stele, Phloem, Xylem, Meristematic and Root cap. **b** UMAP projection colored by input type of fresh nuclei, frozen nuclei, and protoplasts. **c** Bar plots showing the proportion of each cell type across biological replicates and input types. **d** Dot plot showing the expression of representative marker genes across cell types. **e** UMAP plot showing the expression of *Zm00001d004089*, specifically enriched in columella cells

approaches to achieve a comprehensive and accurate representation of plant cell diversity in single-cell studies.

The optimization, assessment and benchmarking of single nuclei data

TSO artifacts represent a major source of low genome alignment rates in challenging samples, such as the nuclei-derived inputs of plant tissues [34]. A high proportion of TSO-containing reads may reflect cDNA degradation or the presence of abnormally short cDNA fragments during reverse transcription, consistent with observations in nuclei-derived bulk RNA-seq data. Although optimization of the reverse transcription step, such as redesigning the TSO oligo can reduce artifact formation, such adjustments are typically not feasible for researchers relying on commercial kits with fixed chemistries. On the bioinformatics side, two commonly used strategies for mitigating TSO interference include: (i) discarding entire reads containing TSO sequences, which is complicated by difficulties in defining TSO positions and overlap lengths; and (ii) filtering out reads below a predefined length threshold, which fails to remove more complex structures such as TSO concatemers [35].

Cutadapt offers a flexible and effective solution to remove sequencing artifacts [36]. It can be employed to detect TSO sequences in a defined orientation and repeat number, removes reads shorter than a user-specified threshold, and retains the remaining clean reads of varying lengths for downstream analysis. As shown in Figure S7, TSO removal was performed on a random subset of raw 100-nt reads from fresh leaf nuclei using Cutadapt, configured with a 25-nt TSO sequence, a maximum of four repeats, and a minimum 3-nt overlap (see Methods). Notably, 11.8% of reads containing TSO concatemers were successfully identified and discarded, while the remaining reads were retained and usable. This strategy significantly improved data quality and provided a cost-effective way to minimize unnecessary loss of sequencing resources.

Background noise is an inherent challenge in droplet-based scRNA-seq platforms and can obscure true biological signals. Notably, contamination levels have been reported to be significantly higher in snRNA-seq compared to scRNA-seq [8]. A key contributing factor is the markedly lower sensitivity of snRNA-seq, as reflected by the reduced number of detected features per cell (Fig. 7a, b), consistent with trends observed in bulk transcriptomic data. SoupX was applied to quantify background contamination [37]. Protoplast samples exhibited relatively stable global contamination rates, ranging from 5% to 7%, whereas nuclei samples showed broader and substantially higher contamination levels, ranging from 14% to 52% (Fig. 7c), consistent with trends reported in mammalian snRNA-seq studies [38].

To address this challenge, CellBender was used for ambient RNA correction [39]. Using a marker gene list sourced from previous studies, CellBender effectively suppressed background noise in nuclei-derived data, reducing contamination from highly abundant genes such as those specific to root cap cells (Figure S8). However, this improvement came at the cost of reduced overall sensitivity across cell types, which in turn complicated cell-type annotation. Importantly, CellBender failed to recover cell-type-specific features that were uniquely captured in protoplast datasets, such as markers of the meristematic cells. These observations emphasized the need for caution when applying background correction tools, and suggest that, when feasible, integrating both protoplast- and nuclei-derived datasets remains the most effective strategy for achieving a comprehensive and reliable view of plant cellular diversity.

An alternative strategy to improve the quality of nuclei sequencing involves first isolating protoplasts from fresh tissues, followed by enzymatic digestion of the plasma membrane to release intact nuclei. This method minimizes the risk of mechanical damage to nuclei often introduced by chopping or grinding. It is particularly suitable for tissues or organs composed of discrete structures, or for studies focusing exclusively on outer cell layers, such as investigations of the epidermis of cotton ovules [40, 41]. However, this approach has not been applied to maize in previous studies, primarily due to its relatively low yield compared to the protocols used in this study. More importantly, it combines two potential sources of systematic bias: incomplete recovery of internal cell layers during protoplast isolation and reduced transcriptomic sensitivity caused by cytoplasmic removal.

Maintaining consistency in input material is also essential for accurate interpretation of snRNA-seq and bulk RNA-seq data. In this study, pseudo-bulk expression profiles were generated from single-cell datasets and compared with bulk profiles using the optimized sequencing workflow. Spearman correlation coefficients between pseudo-bulk and bulk protoplast-derived profiles ranged from 0.854 to 0.887 (Fig. 7d, e, Figure S9), outperforming previous studies that employed conventional protocols [20, 21, 42]. For nuclei-derived data, the correlation between pseudo-bulk and bulk expression profiles was notably higher when input material was matched. These findings highlight the importance of aligning input sources when designing bulk transcriptome experiments, especially in plant snRNA-seq studies.

Conclusions

In this study, we systematically evaluated single-cell transcriptomic profiling strategies in maize using both protoplast- and nucleus-based approaches across different

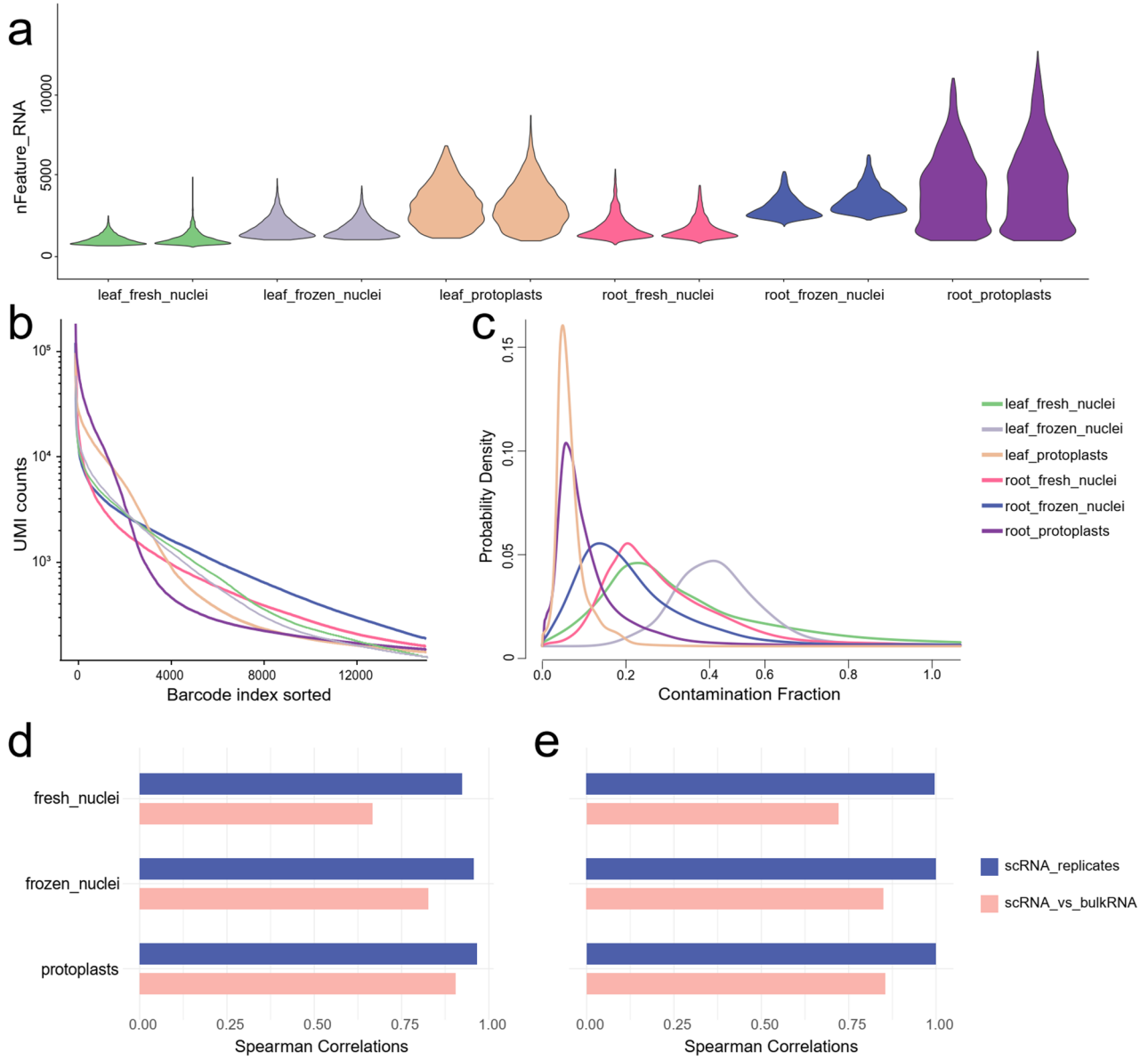

Fig. 7 Quality assessment and comparative transcriptomic features across input types. **a** Violin plots showing the number of detected genes per cell across six input conditions: fresh nuclei, frozen nuclei, and protoplasts from both leaf and root tissues. **b** UMI count distributions per barcode ranked by abundance. **c** Distribution of contamination fractions estimated by SoupX, indicating ambient RNA contamination across conditions. **d**, **e** Spearman correlation coefficients comparing replicates within each input type (scRNA_replicates, blue) and each input type with matched bulk RNA-seq profiles (scRNA_vs_bulkRNA, pink) for leaf (**d**) and root (**e**), respectively

tissue types. By optimizing sample preparation and sequencing protocols, we generated high-quality single-cell datasets that captured distinct cellular compositions and expression patterns in maize tissues. By introducing the bulk RNA-seq protocol designed to match droplet-based single-cell chemistry, we were able to improve input-matched bulk and pseudo-bulk correlations, thereby enhancing the interpretability and reproducibility of single-cell datasets, offering practical guidance for experimental design and data interpretation. Protoplasts are generally the preferred input for plant single-cell sequencing due to their consistently higher cell recovery, improved read alignment, broader transcriptome coverage, and enhanced sensitivity. In cases where tissue architecture is complex or target cells are deeply embedded, nuclei isolated from flash-frozen samples are also recommended as a complementary approach, while stress- or damage-related transcriptional biases may also be introduced. When access to liquid nitrogen is limited, fresh nuclei can serve as an alternative, delivering slightly reduced performance while maintaining overall comparable data quality. However, maintaining all procedures

on ice is essential and non-negotiable to preserve sample integrity. Our results provide a comprehensive framework for constructing high-resolution single-cell atlases in complex plant systems and provide practical guidance for selecting appropriate experimental strategies. As our study is based on a single maize inbred line (B73), two organs (leaf 2 and primary root tip), direct generalization to other genotypes, tissue types, or species remains to be validated. These findings highlight the necessity of optimization of experimental design to specific biological questions, particularly in comparative studies or when investigating subtle regulatory differences.

Methods and materials

Seedling germination and growth

Dry seeds of B73 were sterilized with 1% sodium hypochlorite (NaClO) for 30 min and then rinsed three times with distilled water. The sterilized seeds were placed between two layers of wetted germination paper until the roots emerged and elongated to approximately 10 mm. The germination papers were then rolled and positioned vertically in a water-filled glass beaker, incubated at room temperature under light-shielded conditions for 3 days, and subsequently exposed to an 8 h light/16 h dark cycle until the leaf 2 emerged but not fully expanded.

The apical sections (approximately 5 mm) of primary roots and the central sections (approximately 5 mm) of the leaf 2 were collected. Root and leaf samples were divided into three portions: one was used immediately for protoplast isolation, another for fresh nucleus isolation, and the third was flash-frozen in liquid nitrogen for frozen nucleus isolation.

Protoplast isolation

Root protoplasts were isolated following the protocol previously published with modifications [14]. Root tips were collected and promptly transferred into 50 mL centrifuge tubes (BGI Tech, #CH-CT50) containing 5 mL of freshly prepared pre-treatment solution [0.75 g sorbitol (Biosharp, #BS115) and 0.0075 g L-cysteine (Biosharp, #BS181) dissolved in 25 mL double-distilled water]. The samples were incubated at room temperature for 45 min. After incubation, most of the pre-treatment solution was removed. The samples were rinsed with distilled water, then finely chopped using a double-edged razor blade and transferred into 5 mL of freshly prepared enzyme mix solution [1.82 g D-mannitol (Sigma-Aldrich, #M4125), 0.097 g MES (Sigma-Aldrich, #M3671), 0.5 mL 1 M KCl (BBI, #A610440), 0.3 g Cellulase "Onozuka" RS (Yakult Pharmaceutical Industry Co.), 0.3 g Cellulase "Onozuka" R10 (Yakult Pharmaceutical Industry Co.), 0.1 g Macerozyme R10 (Yakult), and 0.09 g Pectolyase Y-23 (MP Biomedicals, #9033-35-6) dissolved in 25 mL solution (pH 5.7), then 0.25 mL 1 M $CaCl_2$ (Sino Pharm, #10005861) and 0.025 g bovine serum albumin (BBI, #A600332) were added]. The mixture was incubated at room temperature with gentle shaking (80 rpm) for 150 min. The solution was then resuspended by pipetting several times until it appeared "cloudy." The released protoplasts were then filtered through a 70 μm cell strainer (Jetbiofil, #CSS013070), washed, and resuspended in wash solution (recipe same as enzyme mix solution, with no enzymes added). Leaf protoplasts were isolated using the commercially available Protoplast Preparation Kit (Coolaber, #PPT111) according to the manufacturer's instructions. The resuspended root and leaf protoplasts were purified by centrifugation, diluted and incubated with FDA staining solution (Invitrogen, # F1303). The stained protoplasts were loaded into a hemocytometer (INCYTO, #DHC-F01-5) and observed under a fluorescence microscope (Olympus, #IX73) using 488 nm excitation and 530 nm emission filters. Viability was determined by calculating the ratio of fluorescently stained intact protoplasts to the total number of intact protoplasts observed.

Nucleus isolation

Nuclei were isolated using the commercially available Nuclei Preparation Kit (Shanghai Biotechnology Corporation, #52305-10) with modifications. To preserve nuclear RNA integrity and minimize degradation by endogenous and exogenous RNases, RNase inhibitors (ABCLONAL, #RK21401) and DTT (Invitrogen, #D1532) were added to all buffers. Tissues were either freshly minced or ground in liquid nitrogen. For fresh samples, tissues were roughly chopped and transferred into a 2.0 mL centrifuge tube containing 1 mL of PL buffer. The tissue was further minced with a blade until no visible debris remained, followed by incubation on ice for 5 min. For frozen samples, tissues were ground into fine powder in liquid nitrogen, transferred into a pre-chilled 2.0 mL centrifuge tube containing 1 mL of PL buffer, and incubated on ice for 5 min. The lysate was then filtered, washed, and resuspended in 500 μL of PA buffer by gentle pipetting. For root nuclei, a density gradient purification was applied by slowly pipetting 500 μL of PB buffer to the bottom of the tube, forming two distinct layers. Another 500 μL of PC buffer was then added, resulting in three layers. The sample was centrifuged at 3,000 g for 20 min at 4 °C, and the nuclei layer (~ 200 μL at the PB-PC interface) was carefully collected. For leaf nuclei, the lysate suspension was stained with DAPI (Invitrogen, #04-24-2006) and sorted using a SONY SH800S cell sorter (SONY Biotechnology), where only DAPI-positive events were gated and collected into the resuspension buffer. All purified nuclei were stained with AOPI (Beyotime, #C2019A) and analyzed using a Countstar Mira FL Fluorescence Cell Analyzer (Alit Biotech) for automated counting.

Library preparation and sequencing

The optimized bulk library preparation protocol is provided in Supplemental File. Bulk libraries that passed quality control were sequenced on a DNBSEQ-T7 sequencer (MGI) using a paired-end 100 + 100 bp strategy. Due to the shorter insert size distribution of nuclei libraries, all bulk sequencing reads were trimmed to PE50 for downstream analysis. scRNA-seq libraries were prepared using the DNBelab C4 Single-Cell RNA Library Preparation Set V3.0 (MGI, #940-001818-00) following the manufacturer's instructions. scRNA-seq libraries that passed quality control were sequenced on DNBSEQ-T7 using a paired-end 47 + 100 bp strategy.

Bulk reads processing and analysis

The reference genome Zm-B73-REFERENCE-GRAMENE-4.0 and the corresponding gene annotation file Zm-B73-REFERENCE-GRAMENE-4.0_Zm00001d.2.gff3 were downloaded from MaizeGDB (https://maizegdb.org). Due to relatively incomplete rRNA annotations, Zea_mays.B73_RefGen_v4.42.gff3 was obtained from Ensembl to supplement rRNA gene information. rRNA index was then built using Bowtie2 v2.4.5 with default parameters.

rRNAs were first removed from the raw reads by aligning against the known rRNAs using Bowtie2 with parameters: --very-sensitive-local --no-unal -I 1 -X 1000, followed by adapters and low-quality reads filtering using SOAPnuke v1.5.6 with parameters: -l 15 -q 0.2 -n 0.01 -Q 2 -f AAGTCGGAGGCCAAGCGGTCTTAGGAAGACAA -r AAGTCGGATCGTAGCCATGTCGTTCTGTGAGCCAAGGAGTTG. Clean reads were then aligned to the reference genome using HISAT2 with the following parameters: --phred64 --sensitive --no-discordant --no-mixed -I 1 -X 1000, and to the reference transcriptome using Bowtie2 with parameters: -q –phred64 –sensitive –dpad 0 –gbar 99 –mp 1,1 –np 1 –score-min L,0,-0.1 -I 1 -X 1000 –no-mixed –no-discordant -k 200. Gene and transcript abundances were quantified with RSEM v1.3.1 with parameters: --forward-prob 0.5 --paired-end. Differential expression analysis was performed using R package DESeq2, genes with fold changes > = 2 and adjusted pvalue < = 0.05 were regarded as differentially expressed. GO enrichment analysis was performed using R package topGO with default parameters.

Single cell reads processing, data integration, clustering and annotation

rRNA reads were first removed using Bowtie2, following the same filtering strategy as applied to bulk RNA-seq data. Reads containing TSO artifacts were trimmed using Cutadapt (https://cutadapt.readthedocs.io) with the following parameters: -n 4 -m 50 --error-rate 0.3 -g AAGCAGTGGTATCAACGCAGAGGGG -o. Clean reads were then demultiplexed and aligned to the maize reference genome using DNBtools v2.1.3 (https://github.com/MGI-tech-bioinformatics/DNBelab_C_Series_HT_scRNA-analysis-software). Background RNA contamination was assessed using SoupX v1.6.2 using the default settings, where contamination fraction was estimated automatically using autoEstCont with the parameters: tfidfMin = 0.01, soupQuantile = 0.8, and forceAccept = TRUE. The raw and filtered gene-barcode matrices were used as inputs to estimate the expression profiles of ambient mRNAs. In parallel, the same dataset was either processed with CellBender v0.2.0 with default recommended settings (--fpr 0.1) to generate a corrected UMI count matrix or directly used for downstream analysis in Seurat v5.3.0.

Low-quality cells and genes were filtered based on the following criteria: (i) outliers in gene count (nFeature_RNA) were identified using the median absolute deviation (MAD); (ii) cells with abnormal mitochondrial transcript proportions (percent.mt) were flagged using the interquartile range (IQR) [43]; (iii) doublets were detected using the DoubletFinder package with default parameters (https://github.com/chris-mcginnis-ucsf/DoubletFinder). This strategy was chosen over fixed thresholds (e.g., 200 < nFeature_RNA < 5,000 and percent.mt < 25%), which may be not suitable for plant cells, particularly nuclei-derived data. Expression matrices were normalized using the NormalizeData function (method = "LogNormalize", scale factor = 10,000). Highly variable genes (n = 2,000) were identified with FindVariableFeatures (method = "vst"), followed by data scaling using ScaleData, regressing out percent.mt. Principal component analysis was performed with RunPCA, and significant PCs were determined using the JackStraw method. Cell clustering was carried out with the Louvain algorithm via FindNeighbors and FindClusters. UMAP visualization was performed using RunUMAP. For datasets with two biological replicates, batch effects were corrected using RunHarmony, and Harmony embeddings were used for downstream clustering and visualization. Cluster-specific marker genes were identified using FindAllMarkers (Wilcoxon rank-sum test; logfc.threshold = 0.5, min.pct = 0.1). Cell types were annotated based on canonical and experimentally validated marker genes (Table S3).

Cryosection

The central sections (approximately 5 mm) of the leaf 2 were collected and fixed in 10% sucrose (BBI, #A610498) solution for 2 h, followed by fixation in 20% sucrose solution for an additional 4 h. The samples were rinsed in PBS and gently blotted dry with lint-free paper to remove residual moisture. After fixation, samples were embedded in OCT compound (Sakura, #4583) and stored at – 80 °C.

Prior to sectioning, embedded tissues were equilibrated at – 20 °C for 30 min, then cryosectioned at 10 μm thickness using a cryostat (Leica, #CT520). Sections were mounted onto adhesive microscope slides (Biosharp, #AR1065), fixed in methanol (Sigma-Aldrich, #34860) at room temperature for 10 min, and stained with 0.1% toluidine blue (Sigma-Aldrich, #T3260) for 1 min. After rinsing under running water, stained sections were observed under the scanning microscope (Motic Scientific, #PA53-FS6).

Data visualization

The graphs were generated using the R package ggplot2, Adobe Illustrator and Microsoft PowerPoint.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13007-025-01490-6.

Supplementary material 1.

Supplementary material 2.

Supplementary material 3.

Acknowledgements

We would like to thank Dr. Yu Liu of China Agricultural University for generously providing the B73 seeds. We would like to thank Xiaohui Li, Lina Wang and BGI sequencing team for supporting this project. We would also like to thank Lele Zong of Wuhan New City Experimental School for the hand-drawn draft of maize seedlings and Biorender.com for the other visual elements in Figure 1. ChatGPT was used to improve the clarity and fluency of the text.

Author contributions

L.Z. and J.W. conceived the study. J.W., Y.J., Q.L., S.G. and P.L. performed sample preparation and sequencing experiments. S.Z., B.L. and Y.Z. performed bioinformatic analysis. L.Z. wrote the initial draft. S.J. and P.Y. oversaw the project and reviewed the paper. All authors read and approved the final manuscript.

Funding

Not applicable.

Data availability

All sequencing data have been deposited in the CNGB Nucleotide Sequence Archive (CNSA) under accession number CNP0007662 ([https://db.cngb.org/cnsa](https:/db.cngb.org/cnsa)). Processed single-cell count matrices are publicly available on Figshare ([https://doi.org/10.6084/m9.figshare.30741611](https:/doi.org/10.6084/m9.figshare.30741611)).

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 20 August 2025 / Accepted: 17 December 2025
Published online: 03 January 2026

References

1. Seyfferth C, Renema J, Wendrich JR, Eekhout T, Seurinck R, Vandamme N, Blob B, Saeys Y, Helariutta Y, Birnbaum KD, De Rybel B. Advances and opportunities in single-cell transcriptomics for plant research. Ann Rev Plant Biol. 2021;72:847–66. https://doi.org/10.1146/annurev-arplant-081720-010120.
2. Mo Y, Jiao Y. Advances and applications of single-cell omics technologies in plant research. Plant J. 2022;110(6):1551–63. https://doi.org/10.1111/tpj.15772.
3. Kuo A, Hansen KD, Hicks SC. Quantification and statistical modeling of droplet-based single-nucleus RNA-sequencing data. Biostatistics. 2024;25(3):801–17. https://doi.org/10.1093/biostatistics/kxad010.
4. Santo B, Fink EE, Krylova AE, Lin YC, Eltemamy M, Wee A, Wessely O, Lee BH, Ting AH. Exploring the utility of snRNA-seq in profiling human bladder tissue: a comprehensive comparison with scRNA-seq. iScience. 2024;28(1):111628. https://doi.org/10.1016/j.isci.2024.111628.
5. Lai Y, Ramírez-Pardo I, Isern J, An J, Perdiguero E, Serrano AL, Li J, García-Domínguez E, Segalés J, Guo P, Lukesova V, Andrés E, Zuo J, Yuan Y, Liu C, Viña J, Doménech-Fernández J, Gómez-Cabrera MC, Song Y, Liu L, Esteban MA. Multimodal cell atlas of the ageing human skeletal muscle. Nature. 2024;629(8010):154–64. https://doi.org/10.1038/s41586-024-07348-6.
6. Heuston EF, Doumatey AP, Naz F, Islam S, Anderson S, Kirby MR, Wincovitch S, Dell'Orso S, Rotimi CN, Adeyemo AA. Optimized methods for scRNA-seq and snRNA-seq of skeletal muscle stored in nucleic acid stabilizing preservative. Commun Biol. 2025;8(1):10. https://doi.org/10.1038/s42003-024-07445-2.
7. Thrupp N, Frigerio S, Wolfs C, Skene L, Fattorelli NG, Poovathingal N, Fourne S, Matthews Y, Theys PM, Mancuso T, de Strooper R, B., Fiers M. Single-nucleus RNA-Seq is not suitable for detection of microglial activation genes in humans. Cell Rep. 2020;32(13):108189. https://doi.org/10.1016/j.celrep.2020.108189.
8. Ding J, Adiconis X, Simmons SK, Kowalczyk MS, Hession CC, Marjanovic ND, Hughes TK, Wadsworth MH, Burks T, Nguyen LT, Kwon JYH, Barak B, Ge W, Kedaigle AJ, Carroll S, Li S, Hacohen N, Rozenblatt-Rosen O, Shalek AK, Villani AC, Levin JZ. Systematic comparison of single-cell and single-nucleus RNA-sequencing methods. Nat Biotechnol. 2020;38(6):737–46. https://doi.org/10.1038/s41587-020-0465-8.
9. Slyper M, Porter CBM, Ashenberg O, Waldman J, Drokhlyansky E, Wakiro I, Smillie C, Smith-Rosario G, Wu J, Dionne D, Vigneau S, Jané-Valbuena J, Tickle TL, Napolitano S, Su MJ, Patel AG, Karlstrom A, Gritsch S, Nomura M, Waghray A, Regev A. A single-cell and single-nucleus RNA-Seq toolbox for fresh and frozen human tumors. Nat Med. 2020;26(5):792–802. https://doi.org/10.1038/s41591-020-0844-1.
10. Denisenko E, Guo BB, Jones M, Hou R, de Kock L, Lassmann T, Poppe D, Clément O, Simmons RK, Lister R, Forrest ARR. Systematic assessment of tissue dissociation and storage biases in single-cell and single-nucleus RNA-seq workflows. Genome Biol. 2020;21(1):130. https://doi.org/10.1186/s13059-020-02048-6.
11. Tosti L, Hang Y, Debnath O, Tiesmeyer S, Trefzer T, Steiger K, Ten FW, Lukassen S, Ballke S, Kühl AA, Spieckermann S, Bottino R, Ishaque N, Weichert W, Kim SK, Eils R, Conrad C. Single-Nucleus and in situ rna-sequencing reveal cell topographies in the human pancreas. Gastroenterol. 2021;160(4):1330–e134411. https://doi.org/10.1053/j.gastro.2020.11.010.
12. Grones C, Eekhout T, Shi D, Neumann M, Berg LS, Ke Y, Shahan R, Cox KL Jr, Gomez-Cano F, Nelissen H, Lohmann JU, Giacomello S, Martin OC, Cole B, Wang JW, Kaufmann K, Raissig MT, Palfalvi G, Greb T, Libault M, De Rybel B. Best practices for the execution, analysis, and data storage of plant single-cell/nucleus transcriptomics. Plant Cell. 2024;36(4):812–28. https://doi.org/10.1093/plcell/koae003.
13. Cole B, Bergmann D, Blaby-Haas CE, Blaby IK, Bouchard KE, Brady SM, Ciobanu D, Coleman-Derr D, Leiboff S, Mortimer JC, Nobori T, Rhee SY, Schmutz J, Simmons BA, Singh AK, Sinha N, Vogel JP, O'Malley RC, Visel A, Dickel DE. Plant single-cell solutions for energy and the environment. Commun Biol. 2021;4(1):962. https://doi.org/10.1038/s42003-021-02477-4.
14. Ortiz-Ramírez C, Guillotin B, Xu X, Rahni R, Zhang S, Yan Z, Coqueiro Dias Araujo P, Demesa-Arevalo E, Lee L, Van Eck J, Gingeras TR, Jackson D, Gallagher KL, Birnbaum KD. Ground tissue circuitry regulates organ complexity in maize and setaria. Science. 2021;374(6572):1247–52. https://doi.org/10.1126/science.abj2327.
15. Sun G, Xia M, Li J, Ma W, Li Q, Xie J, Bai S, Fang S, Sun T, Feng X, Guo G, Niu Y, Hou J, Ye W, Ma J, Guo S, Wang H, Long Y, Zhang X, Zhang J, Song CP. The maize single-nucleus transcriptome comprehensively describes signaling networks governing movement and development of grass stomata. Plant Cell. 2022;34(5):1890–911. https://doi.org/10.1093/plcell/koac047.

16. Cui Y, Su Y, Bian J, Han X, Guo H, Yang Z, Chen Y, Li L, Li T, Deng XW, Liu X. Single-nucleus RNA and ATAC sequencing analyses provide molecular insights into early pod development of peanut fruit. Plant Commun. 2024;5(8):100979. https://doi.org/10.1016/j.xplc.2024.100979.
17. Lu Z, Hofmeister BT, Vollmers C, DuBois RM, Schmitz RJ. Combining ATAC-seq with nuclei sorting for discovery of cis-regulatory regions in plant genomes. Nucleic Acids Res. 2017;45(6):e41. https://doi.org/10.1093/nar/gkw1179.
18. Wang K, Zhao C, Xiang S, Duan K, Chen X, Guo X, Sahu SK. An optimized FACS-free single-nucleus RNA sequencing (snRNA-seq) method for plant science research. Plant Sci. 2023;326:111535. https://doi.org/10.1016/j.plantsci.2022.111535.
19. Xuhui L, Weiwei C, Siqi L, Junteng F, Hang Z, Xiangbo Z, Yongwen Q. Full-length transcriptome analysis of maize root tips reveals the molecular mechanism of cold stress during the seedling stage. BMC Plant Biol. 2022;22(1):398. https://doi.org/10.1186/s12870-022-03787-3.
20. Guillotin B, Rahni R, Passalacqua M, Mohammed MA, Xu X, Raju SK, Ramírez CO, Jackson D, Groen SC, Gillis J, Birnbaum KD. A pan-grass transcriptome reveals patterns of cellular divergence in crops. Nature. 2023;617(7962):785–91. https://doi.org/10.1038/s41586-023-06053-0.
21. Cao Y, Ma J, Han S, Hou M, Wei X, Zhang X, Zhang ZJ, Sun S, Ku L, Tang J, Dong Z, Zhu Z, Wang X, Zhou X, Zhang L, Li X, Long Y, Wan X, Duan C. Single-cell RNA sequencing profiles reveal cell type-specific transcriptional regulation networks conditioning fungal invasion in maize roots. Plant Biotechnol J. 2023;21(9):1839–59. https://doi.org/10.1111/pbi.14097.
22. Hagemann-Jensen M, Ziegenhain C, Chen P, Ramsköld D, Hendriks GJ, Larsson AJM, Faridani OR, Sandberg R. Single-cell RNA counting at allele and isoform resolution using Smart-seq3. Nat Biotechnol. 2020;38(6):708–14. https://doi.org/10.1038/s41587-020-0497-0.
23. Hahaut V, Picelli S. Full-Length Single-Cell RNA-Sequencing with FLASH-seq. Methods Mol Biol. 2023;2584:123–64. https://doi.org/10.1007/978-1-0716-2756-3_5.
24. Guillaumet-Adkins A, Rodríguez-Esteban G, Mereu E, Mendez-Lago M, Jaitin DA, Villanueva A, Vidal A, Martinez-Marti A, Felip E, Vivancos A, Keren-Shaul H, Heath S, Gut M, Amit I, Gut I, Heyn H. Single-cell transcriptome conservation in cryopreserved cells and tissues. Genome Biol. 2017;18(1):45. https://doi.org/10.1186/s13059-017-1171-9.
25. Liu Z, Kong X, Long Y, Liu S, Zhang H, Jia J, Cui W, Zhang Z, Song X, Qiu L, Zhai J, Yan Z. Integrated single-nucleus and Spatial transcriptomics captures transitional States in soybean nodule maturation. Nat Plants. 2023;9(4):515–24. https://doi.org/10.1038/s41477-023-01387-z.
26. Yang Q, Li Z, Guan K, Wang Z, Tang X, Hong Y, Liu Z, Zhai J, Huang A, Long Y, Song Y. Comparative single-nucleus RNA-seq analysis revealed localized and cell type-specific pathways governing root-microbiome interactions. Nat Commun. 2025;16(1):3169. https://doi.org/10.1038/s41467-025-58395-0.
27. Billakurthi K, Hibberd JM. A rapid and robust leaf ablation method to visualize bundle sheath cells and chloroplasts in C_3 and C_4 grasses. Plant Methods. 2023;19(1):69. https://doi.org/10.1186/s13007-023-01041-x.
28. Yan XC, Liu Q, Yang Q, Wang KL, Zhai XZ, Kou MY, Liu JL, Li ST, Deng SH, Li MM, Duan HJ. Single-cell transcriptomic profiling of maize cell heterogeneity and systemic immune responses against puccinia Polysora underw. Plant Biotechnol J. 2025;23(2):549–63. https://doi.org/10.1111/pbi.14519.
29. Bezrutczyk M, Zöllner NR, Kruse CPS, Hartwig T, Lautwein T, Köhrer K, Frommer WB, Kim JY. Evidence for phloem loading via the abaxial bundle sheath cells in maize leaves. Plant Cell. 2021;33(3):531–47. https://doi.org/10.1093/plcell/koaa055.
30. Ke Y, Pujol V, Staut J, Pollaris L, Seurinck R, Eekhout T, Grones C, Saura-Sanchez M, Van Bel M, Vuylsteke M, Ariani A, Liseron-Monfils C, Vandepoele K, Saeys Y, De Rybel B. A single-cell and Spatial wheat root atlas with cross-species annotations delineates conserved tissue-specific marker genes and regulators. Cell Rep. 2025;44(2):115240. https://doi.org/10.1016/j.celrep.2025.115240.
31. Wang W, Zhang X, Zhang Y, Zhang Z, Yang C, Cao W, Liang Y, Zhou Q, Hu Q, Zhang Y, Wang Y, Xing Y, Qian W, Yao N, Xu N, Liu J. Single-Cell and Spatial transcriptomics reveals a stereoscopic response of rice leaf cells to Magnaporthe oryzae infection. Adv Sci. 2025;12(19):e2416846. https://doi.org/10.1002/advs.202416846.
32. Denyer T, Ma X, Klesen S, Scacchi E, Nieselt K, Timmermans MCP. Spatiotemporal developmental trajectories in the Arabidopsis root revealed using High-Throughput Single-Cell RNA sequencing. Dev Cell. 2019;48(6):840–e8525. https://doi.org/10.1016/j.devcel.2019.02.022.
33. Wang T, Wang F, Deng S, Wang K, Feng D, Xu F, Guo W, Yu J, Wu Y, Wuriyanghan H, Li ST, Gu X, Le L, Pu L. Single-cell transcriptomes reveal Spatiotemporal heat stress response in maize roots. Nat Commun. 2025;16(1):177. https://doi.org/10.1038/s41467-024-55485-3.
34. Kumari P, Kaur M, Dindhoria K, Ashford B, Amarasinghe SL, Thind AS. Advances in long-read single-cell transcriptomics. Hum Genet. 2024;143(9–10):1005–20. https://doi.org/10.1007/s00439-024-02678-x.
35. James T, Webber DA, Bartlett G, Al-Eryani A, Khorgade A, Shin C, Georgescu H, Yu, Brian J, Haas. Victoria Popic, Aziz Al'Khafaji. Multi-platform evaluation and optimization of single-cell RNA isoform sequencing. bioRxiv 2025.02.22.639662; https://doi.org/10.1101/2025.02.22.639662
36. Kechin A, Boyarskikh U, Kel A, Filipenko M. CutPrimers: A new tool for accurate cutting of primers from reads of targeted next generation sequencing. J Comput Biol. 2017;24(11):1138–43. https://doi.org/10.1089/cmb.2017.0096.
37. Young MD, Behjati S. SoupX removes ambient RNA contamination from droplet-based single-cell RNA sequencing data. GigaScience. 2020;9(12):giaa151. https://doi.org/10.1093/gigascience/giaa151.
38. Janssen P, Kliesmete Z, Vieth B, Adiconis X, Simmons S, Marshall J, McCabe C, Heyn H, Levin JZ, Enard W, Hellmann I. The effect of background noise and its removal on the analysis of single-cell expression data. Genome Biol. 2023;24(1):140. https://doi.org/10.1186/s13059-023-02978-x.
39. Fleming SJ, Chaffin MD, Arduini A, Akkad AD, Banks E, Marioni JC, Philippakis AA, Ellinor PT, Babadi M. Unsupervised removal of systematic background noise from droplet-based single-cell experiments using cellbender. Nat Method. 2023;20(9):1323–35. https://doi.org/10.1038/s41592-023-01943-7.
40. Wang D, Hu X, Ye H, Wang Y, Yang Q, Liang X, Wang Z, Zhou Y, Wen M, Yuan X, Zheng X, Ye W, Guo B, Yusuyin M, Russinova E, Zhou Y, Wang K. Cell-specific clock-controlled gene expression program regulates rhythmic fiber cell growth in cotton. Genome Biol. 2023;24(1):49. https://doi.org/10.1186/s13059-023-02886-0.
41. Zhu X, Xu Z, Wang G, Cong Y, Yu L, Jia R, Qin Y, Zhang G, Li B, Yuan D, Tu L, Yang X, Lindsey K, Zhang X, Jin S. Single-cell resolution analysis reveals the Preparation for reprogramming the fate of stem cell niche in cotton lateral meristem. Genome Biol. 2023;24(1):194. https://doi.org/10.1186/s13059-023-03032-6.
42. Chen X, Yao R, Hua X, Du K, Liu B, Yuan Y, Wang P, Yan Q, Dong L, Groen SC, Jiang S, Zhou T. Identification of maize genes that condition early systemic infection of sugarcane mosaic virus through single-cell transcriptomics. Plant Commun. 2025;6(5):101297. https://doi.org/10.1016/j.xplc.2025.101297.
43. Germain PL, Sonrel A, Robinson MD, pipeComp. A general framework for the evaluation of computational pipelines, reveals performant single cell RNA-seq preprocessing tools. Genome Biol. 2020;21:227. https://doi.org/10.1186/s13059-020-02136-7.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.