Single-cell genomics in plants: current state, future directions, and hurdles to overcome

Josh T. Cuperus

1, *1, †

1 Department of Genome Sciences, University of Washington, Seattle, Washington, USA

*Author for communication: cuperusj@uw.edu
†Senior author.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (https://academic.oup.com/plphys/pages/general-instructions) is Josh T. Cuperus (cuperusj@uw.edu).

Abstract

Single-cell genomics has the potential to revolutionize the study of plant development and tissue-specific responses to environmental stimuli by revealing heretofore unknown players and gene regulatory processes. Here, I focus on the current state of single-cell genomics in plants, emerging technologies and applications, in addition to outlining possible future directions for experiments. I describe approaches to enable cheaper and larger experiments and technologies to measure multiple types of molecules to better model and understand cell types and their different states and trajectories throughout development. Lastly, I discuss the inherent limitations of single-cell studies and the technological hurdles that need to be overcome to widely apply single-cell genomics in crops to generate the greatest possible knowledge gain.

Introduction

In many ways, the promise of single-cell genomics—so well demonstrated in animal and human research—has finally arrived in plants. In 2019, five publications, appearing almost simultaneously, characterized gene expression in single Arabidopsis (Arabidopsis thaliana) root cells (Denyer et al., 2019; Jean-Baptiste et al., 2019; Ryu et al., 2019; Shulse et al., 2019; Zhang et al., 2019). The Arabidopsis root was the perfect system for these proof-of-principle studies. With few cell types and extensive existing knowledge on cell-type-specific expression and developmental trajectories (Brady et al., 2007), these studies were able to build on the known “ground truth” of cell-type proportions and spatial and temporal expression patterns. Since then, single-cell genomics has been applied to other plants, including crops like maize (Kim et al., 2021; Marand et al., 2021; Ortiz-Ramirez et al., 2021) and rice (Wang et al., 2020; Liu et al., 2021), and has contributed to the discovery of previously unknown biology, such as genes acting in lateral root development (Gala et al., 2021), or regulating root hair density in response to phosphate dosage (Wendrich et al., 2020). All studies published to date highlight the promise of single-cell genomics for fully...
understanding plant gene regulation; however, they also highlight the existing challenges before this promise can be fully realized. The first challenge is clear: we need to analyze many more cells to capture all the cell types and their various developmental states in large crops like maize.

We need more cells: some biological questions require more cells

The first droplet-based single-cell studies examined only a few thousand root cells (Denyer et al., 2019, 4,727 cells; Jean-Baptiste et al., 2019, 3,121 cells; Ryu et al., 2019, 7,522 cells; Shulse et al., 2019, 12,198 cells; Zhang et al., 2019, 7,695 cells). In stark contrast, single-cell studies in animal systems often rely on tens of thousands if not millions of cells to draw conclusions—for example, two million cells were assayed in a 2019 study of mouse embryos (Cao et al., 2019), 50,000 cells in a planarian single-cell transcriptome study (Fincher et al., 2018), and 80,000 cells in the Caenorhabditis elegans embryogenesis atlas (Packer et al., 2019). Although more does not always equate better, there is a strong argument for this to be true in single-cell genomics. Combining the 2019 Arabidopsis root single-cell datasets boosted cell numbers to about 50,000 and filled in developmental trajectories for some cell layers (McFaline-Figueroa et al., 2020). Following this approach, others subsequently integrated their newly generated, larger datasets with this existing data (for a total of approximately 96,000 cells), resulting in the largest, most complete analysis of Arabidopsis root development to date (Shahan et al., 2020).

Although the current cell numbers may have sufficed to study the exceedingly well-characterized Arabidopsis root, as the field moves to profiling more complex tissues of less well-characterized, larger plants, the ability to profile a greater number of cells will be critical for statistical rigor in interpreting the results. This is particularly true for capturing developmental transitions or specific cell states which will be present only in a few cells of a given cell type. It is worth noting that the ability to profile large numbers of cells negates the need for sampling different developmental stages because developmental trajectories of different cell layers, tissues, and organs can be readily reconstructed from high-quality data (Fincher et al., 2018).

Large cell numbers and data integration are critical not only for comparing wild-type cell layers and tissues but also for detecting differences in cell-type proportions or cell-type-specific expression patterns in mutants (Ryu et al., 2019; Shahan et al., 2020). For example, by profiling mutants for the genes SCARECROW and SHORTROOT and comparing them to wild-type, Shahan et al. provide evidence for an alternative pathway that defines endodermal identity in the absence of SCARECROW but requires SHORTROOT function (Shahan et al., 2020). The increase in knowledge gain made possible by larger cell numbers is also well illustrated in a recent maize study that assayed more than 50,000 nuclei for chromatin accessibility (Marand et al., 2021).

Chromatin accessibility was captured from nuclei representing several tissues and stages of development, and tissue- and stage-specific accessible sites were integrated with other features such as genome organization and genetic diversity. As a result, the authors built a comprehensive framework in order to describe and understand gene regulation and its role in phenotypic diversity and domestication. One finding points toward several transcription factor families and their binding sites as acting in a fashion similar to the transcription factor CCCTC-binding factor (CTCF) and its binding sites in metazoans, which define the borders of chromatin loops with CTCF blocking the interaction between enhancers and promoters. Plant genomes do not contain CTCF homologs but surely its functional equivalent must exist to ensure the precision of plant gene expression and its dynamics in development and environmental responses. Marand et al. offer a first glimpse of possible functional equivalents, a finding that will undoubtedly stimulate further research.

Although more recent plant studies include increased numbers of cells, the current cost per nucleus is still very high, keeping the technology out of reach for many plant research groups. Commercial technologies such as the droplet-based method offered by 10× Genomics cost about $1.00 per cell, not including labor or downstream analysis. There are other, cheaper technologies developed in animal systems that could be applied to plants: single-cell combinatorial indexing, split-seq, and scifi-RNA-seq, among others (Cao et al., 2017; Rosenberg et al., 2018; Datlinger et al., 2019). Neither combinatorial indexing nor split-seq relies on a droplet-based approach; rather, they are based on several rounds of marking and mixing cells in a multiplexed format. As a result, each cell’s transcripts and/or accessible sites carry a unique combination of barcodes that can be attributed back to that cell. This approach has been applied extensively in mammalian cell culture experiments as well as to whole animals and bacterial cultures (Rosenberg et al., 2018; Cao et al., 2019; McFaline-Figueroa et al., 2019; Srivatsan et al., 2020; Kuchina et al., 2021).

Methods like "hashing" or MULTI-seq are multiplexing methods to mark cells (Stoeckius et al., 2017; Shin et al., 2019; McGinnis et al., 2019b; Gehring et al., 2020; Fang et al., 2021). In its simplest form, hashing relies on DNA oligos that mimic polyadenylated RNAs, but with well-specific barcodes, which enables large-scale assays of cells exposed to many different conditions, for example, drug concentrations (Srivatsan et al., 2020). Hashing could be used to increase the throughput of droplet-based methods by allowing to overload droplets with multiple cells or nuclei. The recently published scifi method uses a similar multiplexing approach, combining one-step combinatorial preindexing of single-cell transcriptomes with subsequent single-cell RNA-seq using commercially available droplet microfluidics. Similar to hashing, the preindexing step allows the loading of multiple cells per droplet, which increases the throughput of droplet-based single-cell RNA-seq up to 15-fold, enabling the multiplexing...
of many samples in a single scfi-RNA-seq experiment (Datlinger et al., 2019). Both hashing and scifi could make the 10× Genomics platform more affordable for plant researchers; however, developing plant-specific methods that do not rely on proprietary kits and instrumentation would further reduce costs.

There are drawbacks to increase the number of cells through combinations of barcodes: the higher the number of combinations (e.g. 2 versus 3 in combinatorial indexing), the fewer transcripts per cell are typically captured. The more cells are included in an experiment, the higher the incidence of barcode collision (i.e. two cells carrying the same set of barcodes) and thus “wasted” sequencing, which requires computational follow-up steps to remove suspect cells (Wolock et al., 2019; DePasquale et al., 2019; McGinnis et al., 2019a). These methods commonly rely on simulating doublets and identify real doublets by similarity, allowing for their removal from the analysis. Multiplexing based on natural variation is another strategy to increase throughput and simultaneously detect doublets effectively (Kang et al., 2018).

More recent single-cell genomics studies often profile nuclei rather than whole cells, presumably capturing a greater number of newly transcribed molecules. Although nuclei-based methods capture fewer molecules, these may be more biologically informative (Bakken et al., 2018; Farmer et al., 2021). Researchers are rightly concerned with the minimum requirements for cell numbers and number of molecules measured. At this point, however, it is challenging to formulate general guidelines, especially since these requirements critically depend on the experimental question asked. For example, if the scientific question concerns the relative size of a certain cell population among mutants, measuring a few, cell-type-specific transcripts can suffice. If the scientific question focuses on differential cell states or developmental transitions, many more cells and measured transcripts are needed for a statistically rigorous, meaningful analysis.

**Capturing low-abundance cells**

During development, heterogeneity in gene expression and ultimately different cell fates arise from a population of formerly identical cells. Many researchers are keen to capture the initial emergence of this regulatory heterogeneity in development, which likely occurs only in a small number of cells. Unfortunately, these cells will represent only a tiny fraction of the cells collected in a typical single-cell experiment, limiting statistical power. Because the costs per cell are still very high, simply expanding the number of cells is not the best answer. As an alternative, several groups have sought to enrich low-abundance cell types. Their strategies included careful tissue selection, for example, focusing on young seedlings, increased digestion of plant tissue to enrich for vasculature cells, tissue dissection, or physical removal of overrepresented cells (Bezrutczyk et al., 2021; Gala et al., 2021; Kim et al., 2021). These strategies met with varying success: mesophyll cells still accounted for ~75% of all assayed cells in Arabidopsis leaves (Bezrutczyk et al., 2021) and 98% in maize leaves (Bezrutczyk et al., 2021). With as few as approximately 50 cells, Bezrutczyk et al. still found many differentially expressed genes, subsequently validated with in situ hybridization. Similarly, Gala et al. relied on only 167 lateral root primordia (LRP) cells out of nearly 7,000 sequenced cells to identify approximately 800 differentially expressed, LRP-specific genes (Gala et al., 2021). Several of the selected top candidate genes influenced lateral root development when their expression was manipulated in validation experiments. Other groups have tackled low abundance cell types in even higher resolution, focusing on LRP in Arabidopsis (Serrano-Ron et al., 2021), shoot-borne-root meristems in tomato (Omary et al., 2020), inflorescence tissue in maize (Xu et al., 2021), and phloem root cells in Arabidopsis (Roszak et al., 2021).

Although these examples give us confidence that important knowledge can be gleaned even from a small number of cells, efficient enrichment methods are a high priority in the field. One proven strategy is fluorescence-activated cell sorting of either protoplasted cells or nuclei (Brady et al., 2007); a similar strategy relies on the capture of biotinylated nuclei (Deal and Henikoff, 2011). Both methods require prior knowledge of genes with cell-type or condition-specific expression; they also rely on transgenic lines that carry the respective constructs marking a particular cell type. Ultimately, the solution to capture low abundance cell types in crops will have to involve both: The development of innovative enrichment strategies coupled with the adaptation of sequencing methods that allow the analysis of millions of cells at a reasonable price.

**Cell-type annotation without the deep prior knowledge gathered from marker lines**

A major challenge of all single-cell genomics studies is the correct annotation of cell types. It is no coincidence that the first plant single-cell RNA-seq studies focused on the Arabidopsis root (Denyer et al., 2019; Jean-Baptiste et al., 2019; Ryu et al., 2019; Shulse et al., 2019; Zhang et al., 2019), and that in animals C. elegans with its exceedingly well-characterized invariant cell lineage (Sulston et al., 1983) served as a model for organismal single-cell genomics (Cao et al., 2017; Packer et al., 2019). Even in maize with its good annotation of certain tissue-specific expression patterns, cell type annotation remains a challenge. Marand et al. used extensive literature searches to identify the number of marker genes necessary to call a cluster reliably (Marand et al., 2021).

The initial Arabidopsis single-cell studies critically relied on root cell-type-specific gene expression data derived from the fluorescent marker lines that researchers had previously generated (Birnbaum et al., 2003). These resources are less available for shoot cell layers, especially throughout development. Such marker lines are even less common for most crops or really any other plant species on earth. A recent study ingeniously tried to overcome this limitation. Ortiz-Ramirez et al. relied on fluorescent dyes and an endodermis marker line to capture different maize cell types. Specifically, they used dyes...
with different abilities to penetrate cell layers to sort exterior and interior cell layers of maize root tips followed by RNA-seq. This approach facilitated the identification of cell clusters in their single-cell RNA-seq experiments, allowing them to confidently identify 16 distinct clusters at a sub-tissue level (Ortiz-Ramirez et al., 2021). The dye-penetrance method can easily be applied to other plants; however, is likely limited to root tips in its current form.

In order to overcome the challenge of cell-type annotation, animal researchers have developed computational approaches to predict cell types in single-cell space based on previous experiments. Although this approach has not yet been applied outside of mammalian systems, programs like Garnet, clustifyR, scClassify, scNym, and scCATCH can predict cell types for the same types of data from the same organism with high reproducibility (Pliner et al., 2019; Fu et al., 2020; Kimmel and Kelley, 2020; Lin et al., 2020; Shao et al., 2020). I envision a pan-classifier that builds on existing single-cell studies across several plant species, and includes homolog information to predict cell types in single-cell experiments in additional plant species without the reliance on extensive marker genes and tissue level bulk RNA-seq. There is evidence from animal systems for this to work: for Garnet, cell-type annotation across species was remarkably accurate when using mouse data to classify human single cells (Pliner et al., 2019). Using single-cell data from mouse lung cells, over 92% of alveolar, B cells, T cells, epithelial (ciliated) cells, endothelial cells, and fibroblasts were accurately assigned in a human lung tumor.

Similar approaches in plants are faced with the challenge of deep divergence in the plant kingdom. Monocots and dicots are far more diverged than mouse and human, and even within dicots, many tissues do not share cell-type-specific expression patterns (Kajala et al., 2021). Nevertheless, other aspects of gene regulation such as transcription factors and their binding motifs are conserved across plants. The potential for classifying cell types in less-well characterized plants on the basis of existing single-cell data from Arabidopsis is well illustrated by a recent study that tested several of the 111 identified marker genes for trichoblast cells in five other plant species (Yan et al., 2020). As the efficacy of machine learning relies critically on data volume, the results will only improve as more and more datasets become available.

Using the power of long reads to gain high-resolution knowledge

Long-read sequencing is getting more affordable and more accurate (Wenger et al., 2019), making it possible to apply it to single cells. There are already several groups that have successfully applied long-read sequencing to different systems (Gupta et al., 2018; Singh et al., 2019; Lebrigand et al., 2020; Volden and Vollmers, 2020; Hård et al., 2021; Long et al., 2021). Long et al. used a combination of Illumina and Nanopore sequencing to assay transcriptomes in Arabidopsis roots at single-cell resolution, enabling them to identify cell-type-specific splice forms (Long et al., 2021). This is clearly an important step forward to understanding plant development and plant responses to environmental stimuli because alternative splicing is so prevalent in plants.

There are also new technologies that rely on long-read technology to assay chromatin occupancy by nucleosomes and transcription factors (Abdulhay et al., 2020; Stergachis et al., 2020). Because both Nanopore and PacBio can detect DNA methylation, these methods rely on methylating accessible DNA and reading out the resulting patterns on 10 kb or larger DNA molecules. In this way, accessible regions, transcription factor footprints, and nucleosomes can be assayed in context with one another, which could help to overcome the sparseness of single-cell ATAC data. Although these methods have not yet been applied to single cells or even just to plants, I predict that it is only a matter of time when this will be attempted.

Co-assays for maximal knowledge gain

Several software packages allow researchers to computationally combine different single-cell measurements. For example, Dorrity et al. co-embedded existing single-cell RNA-seq from Arabidopsis roots with their newly generated ATAC-seq data from young roots (Dorrity et al., 2020). This approach allowed the authors to annotate their cells with the much richer features that can be gathered from gene expression data, including scores for cell cycle state, endoreplication, and developmental progression. They were also able to associate the increase of chromatin accessibility at sites containing a binding motif for a certain transcription factor family with the increase in expression of a specific member of this transcription factor family. In short, they were able to make inferences about specific regulatory events at individual loci, something the field has struggled to achieve due to the generally poor association of chromatin accessibility and gene expression.

In animals, combinatorial indexing has been used to co-capture chromatin accessibility and gene expression; however, this co-assay reduces power for both measures because many fewer molecules are captured (Cao et al., 2018). Although I am not aware of any publications or preprints applying them yet, commercial platforms have begun to offer droplet-based co-assays. Notably, the published co-assays of accessibility and gene expression continue to observe the weak correlation between chromatin accessibility and gene expression that we and others have attributed to regulatory complexity (Brady et al., 2011; Gaudinier et al., 2011; MacNeil et al., 2015; Fuxman Bass et al., 2016; Alexandre et al., 2017; Cao et al., 2018; Chen et al., 2019; Dorrity et al., 2020). A weak correlation is even observed when associating only the accessibility of promoters of differentially expressed genes with their expression levels in mouse cerebral cortex clusters ($r = 0.34$ on average; Chen et al., 2019). Co-assays of accessibility and gene expression in plants with their more plastic development may help facilitate the resolution of the ongoing debate whether to attribute the low correlation
OUTSTANDING QUESTIONS

• Will single-cell genomics help us resolve the relationship between accessibility and gene expression?
• How will we harness the underlying information we gather from single-cell experiments to manipulate crops for better outcomes?
• What emerging technologies can we best apply to plant systems with meaningful outcomes?

between accessibility and expression to regulatory complexity, technical limitations, or a combination of both.

Plants derive much of their unsurpassed capacity for phenotypic plasticity from their vast universe of small RNAs. Although there have been a few small-scale attempts to measure both small RNAs and messenger RNA in mammalian cell culture (Wang et al., 2019), this type of co-assay really needs to be applied in plants to gain important biological insights. Capturing the regulatory small RNAs in plants in conjunction with the mRNAs and genes they regulate will enable us to better understand gene regulation at a single-cell level. In stark contrast to animal systems, small RNAs in plants rely on full or near full sequence alignment to target sites, which readily enables target predictions. This technology will likely be pioneered in Arabidopsis or maize due to the wealth of existing tools.

Another, only somewhat futuristic, plant-specific application of single-cell genomics would be the potential to multiplex experiments by exploring the effects of hormones, herbicides, temperature, or fertilizers on seedlings, possibly including mutant collections. I envision these experiments with very small plant species, such as duckweed, or with young seedlings. Although comparable human single-cell research can rely on multiplexing cell lines, plant researchers are able to work within the organismal context that matters for responses to genetic or environmental perturbations.

To what end?

The hope and promise of single-cell genomics are that someday we will fully understand gene expression in order to predict and manipulate gene expression in future crops, using either natural variants or targeted genome editing. As traditional breeding becomes too slow to accommodate the rapid changes in the environmental conditions around us, this knowledge is critical for obtaining food and energy security.

Acknowledgments

I thank Christine Queitsch for valuable discussions and support.

Funding

This work was supported by the National Science Foundation (RESEARCH-PGR grant 17488843). This work was also supported by NIH grant R01-GM079712 and R35-GM139532.

Conflict of interest statement. none declared.

References

Abdulhajy NJ, McNally CP, Hsieh LJ, Kasinathan S, Keith A, Estes LS, Karimzadeh M, Underwood JG, Goodarzi H, Narlikar GJ, et al. (2020) Massively multiplex single-molecule oligonucleosome footprinting. eLife 9: e59404

Alexandre CM, Urton JR, Jean-Baptiste K, Huddleston JL, Dorrity MW, Cuperus JT, Sullivan AM, Benm F, Jolic D, Arsovski AA, et al. (2017) Complex relationships between chromatin accessibility, sequence divergence, and gene expression in A. thaliana. Mol Biol Evol 35: 837–854

Bakken TE, Hodge RD, Miller JA, Yao Z, Nguyen TN, Aevermann B, Barkan E, Bertagnolli D, Casper T, Dee N, et al. (2018) Single-nucleus and single-cell transcriptomes compared in matched cortical cell types. PLoS One 13: e0209648

Bezrutzczyk M, Zöllner NR, Kruse CPS, Hartwig T, Lautwein T, Köhrer K, Frommer WB, Kim JY (2021) Evidence for phloem loading via the abaxial bundle sheath cells in maize leaves. Plant Cell 33: 531–547

Birnbaum K, Shasha DE, Wang JY, Jung JW, Lambert GM, Galbraith DW, Benfey PN (2003) A gene expression map of the Arabidopsis root. Science 302: 1956–1960

Brady SM, Orlando DA, Lee J-Y, Wang JY, Koch J, Dinneny JR, Mace D, Ohler U, Benfey PN (2007) A high-resolution root spatiotemporal map reveals dominant expression patterns. Science 318: 801–806

Brady SM, Zhang L, Megraw M, Martinez NJ, Jiang E, Yi CS, Liu W, Zeng A, Taylor-Teeple M, Kim D, et al. (2011) A sterile-enriched gene regulatory network in the Arabidopsis root. Mol Syst Biol 7: 459

Cao J, CusanoVich DA, Ramani V, Aghamirzadeh D, Pliner HA, Hill AJ, Daza RM, McFalone-Figueroa JL, Packer JS, Christiansen L, et al. (2018) Joint profiling of chromatin accessibility and gene expression in thousands of single cells. Science 361: 1380–1385

Cao J, Packer JS, Ramani V, CusanoVich DA, Huyhnh C, Daza R, Qiu X, Lee C, Furlan SN, Steemers FJ, et al. (2017) Comprehensive single-cell transcriptional profiling of a multicellular organism. Science 357: 661–667

Cao J, Spielmann M, Qiu X, Huang X, Ibrahim DM, Hill AJ, Zhang F, Mundlos S, Christiansen L, Steemers FJ, et al. (2019) The single-cell transcriptional landscape of mammalian organogenesis. Nature 566: 496–502

Chen S, Lake BB, Zhang K (2019) High-throughput sequencing of the transcriptome and chromatin accessibility in the same cell. Nat Biotechnol 37: 1452–1457

Datlinger P, Rendeiro AF, Boenke T, Krausgruber T, Barreca D, Bock C (2019) Ultra-high throughput single-cell RNA sequencing by combinatorial fluidic indexing. bioRxiv doi: 10.1101/2019.12.17.879304

Deal RB, Henikoff S (2011) The INTACT method for cell type-specific gene expression and chromatin profiling in Arabidopsis thaliana. Nat Protocol 6: 56–68

Denyer T, Ma X, Kleen S, Scacchi E, Niesel K, Timmermans MCP (2019) Spatiotemporal developmental trajectories in the Arabidopsis root revealed using high-throughput single-cell RNA sequencing. Dev Cell 48: 840–852.e5

DePasquale EAK, Schnell DJ, Van Camp P-J, Valiente-Alandi ñ, Blaxall BC, Grimes HL, Singh H, Salomonis N (2019) Doublet
decon: deconvoluting doublets from single-cell RNA-sequencing data. Cell Rep 29: 1718–1727.e8
Dority MW, Alexandre CM, Hamm M, Vigil A-L, Fields S, Queitsch C, Cuperus J (2020) The regulatory landscape of Arabidopsis thaliana roots at single-cell resolution. bioRxiv doi: 10.1101/2020.07.12.204792
Fang L, Li G, Sun Z, Zhu Q, Cui H, Li Y, Zhang J, Liang W, Wei W, Hu Y, et al. (2021) CASB: a concanavalin A-based sample barcoding strategy for single-cell sequencing. Mol Syst Biol 17: e10060
Farmer A, Thibivilliers S, Ryu KH, Schiefelein J, Libault M (2021) 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: 372383
Fincher CT, Wurtzel O, de Hoog T, Kravarik KM, Reddien PW (2018) Cell type transcriptome atlas for the planar Schmidtea mediterranea. Science 360: eaau1736
Fu R, Gillen AE, Sheridan RM, Tian C, Daya M, Hao Y, Hesselberth JR, Riemondy KA (2020) clustifyr: an R package for automated single-cell RNA sequencing cluster classification. F1000Res 9: 223
Fuxman Bass JI, Pons C, Kozlowski L, Reeece-Hoyes JS, Shrestha S, Holdorf AD, Mori A, Myers CL, Walhout AJ (2016) A gene-centered C. elegans protein-DNA interaction network provides a framework for functional predictions. Mol Syst Biol 12: 884
Gala HP, Lancot A, Jean-Baptiste K, Giuzio C, Chu JC, Zhai J, George W, Queitsch C, Cuperus JT, Nemhauser JL (2021) A single cell view of the transcriptome during lateral root initiation in Arabidopsis thaliana. Plant Cell 33: 2197–2220
Gaudinier A, Zhang L, Reeece-Hoyes JS, Taylor-Teepeles M, Pu L, Liu Z, Breton G, Pruneda-Paz JL, Kim D, Kay SA, et al. (2021) Enhanced Y1H assays for Arabidopsis. Nat Methods 8: 1053–1055
Dueck H, Stefanik D, Tan K, Trapnell C, Kim J (2019) A pooled single-cell genetic screen identifies regulatory checkpoints in the continuum of the epithelial-to-mesenchymal transition. Nat Genet 51: 1389–1398
McFaline-Figueroa JL, Trapnell C, Cuperus JT (2020) The promise of single-cell genomics in plants. Cur Opin Plant Biol 54: 114–121
McGinnis CS, Murrow LM, Gartner JR (2019a) DoubletFinder: Doublet detection in single-cell RNA sequencing data using artificial neural networks. Cell Syst 8: 329–337.e4
McGinnis CS, Patterson DM, Winkler J, Conrad DN, Heim MY, Heim SY, Serrano-Ron L, Perez-Garcia P, Sanchez-Corrionero A, Gude I, Jia J, Mo W, Fang L, Lu D, Liu B, Zhang H, Chen W, M, Miras M, Zollner N, Hartwig T, Wudick MM et al. (2019b) MULTI-seq: sample multiplexing for single-cell RNA sequencing using lipid-tagged indices. Nat Methods 16: 619–626
Ohler U, Shao X, Liao J, Du Z, Li R, Hu G, Zhang P, Ma Y, et al. (2021) Transcriptional landscape of rice roots at the single-cell resolution. Mol Plant 14: 384–394
Ouyang C, Liao J, Xu R, Ai N, Fan X (2020) scCATCH: automatic annotation on cell types of clusters from single-cell RNA sequencing data. iScience 3: 100882
Shanahan R, Hsieh C, Weng C, Tan H, Shao L, Lin D, Sung SH, Chien Y, et al. (2021) A con-regulatory atlas in maize at single-cell resolution. Cell 184: 3041–3055.e21
Liu Z, Breton G, Pruneda-Paz JL, Kim D, Kay SA, et al. (2018) A single-cell profiling of the developing mouse brain and spinal cord reveals developmental trajectories in wild type and cell identity mutants. bioRxiv doi: 10.1101/2020.06.29.178863
Rajewsky N, Benfey PN, Ohler U (2018) Distinct identities of leaf phloem cells revealed by single cell transcriptomics. Plant Cell 31: 114–121
Serrano-Ron L, Perez-Garcia P, Sanchez-Corrionero A, Gude I, Jia J, Mo W, Fang L, Lu D, Liu B, Zhang H, Chen W, Zhai J (2021) FlsnRNA-seq: protoplasting-full length single-nucleus RNA profiling in plants. Genome Biol 22: 66
Shao X, Liu J, Zhou Y, Yu D, Zang X, Wang G, Zhang J, Zhou Z (2018) DeepSEA: deep learning-based regulatory element prediction. Mol Syst Biol 14: 993–1011
Yan Z, Xu X, Rahni R, Gingeras TR, Jackson D, Gallagher KL, et al. (2021) Ground tissue circuitry regulates organ complexity in cereal roots. bioRxiv doi: 10.1101/2021.04.28.441892
Yao Z, Graybuck LT, Peeler DJ, Mukherjee S, Chen W, et al. (2018) SUPERLOCUS: a framework for functional predictions. Mol Syst Biol 360: 511–530
Zhai J, Dueck H, Stefanik D, Tan K, Trapnell C, Kim J, et al. (2019) A lineage-resolved molecular atlas of C. elegans embryogenesis at single-cell resolution. Science 365: eaaq1736
Pliner HA, Shendure J, Trapnell C (2019) Supervised classification enables rapid annotation of cell atlases. Nat Methods 16: 983–986
Mansour AB, Roco CM, Muscat RA, Kuchina A, Sample P, Yao Z, Graybuck LT, Peeler DJ, Mukherjee S, Chen W, et al. (2018) Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. Science 360: 176–182
Rosenberg AB, Roco CM, Muscat RA, Kuchina A, Sample P, Yao Z, Graybuck LT, Peeler DJ, Mukherjee S, Chen W, et al. (2018) Single-cell RNA sequencing reveals molecular relationships among individual plant cells. Plant Physiol 179: 1444–1456
Serrano-Ron L, Perez-Garcia P, Sanchez-Corrionero A, Uede I, Cabrera J, Ip P-L, Birnbaum KD, Moreno-Risueno MA (2020) Reconstructed of lateral root formation through single-cell RNA sequencing resolves order of tissue initiation. Mol Plant 13: 1362–1378
Shaman R, Hsu CW, Nolan TM, Cole BJ, Taylor IW, Vlot AHC, Benfey PN, Ohler U (2020) A single cell Arabidopsis ploidal atlas reveals developmental trajectories in wild type and cell identity mutants. bioRxiv doi: 10.1101/2020.06.29.178863
Shao X, Liao J, Lu X, Xue R, Ai N, Fan X (2020) scCATCH: automatic annotation on cell types of clusters from single-cell RNA sequencing data. iScience 3: 100882
Shin D, Lee W, Lee JH, Bang D (2019) Multiplexed single-cell RNA-seq via transient barcoding for simultaneous expression profiling of various drug perturbations. Sci Adv 5: eaa2249
Shulse CN, Cole BJ, Ciobanu D, Lin J, Yoshinaga Y, Gouran M, Turco GM, Zhu Y, O’Malley RC, Brady SM, et al. (2019) High-throughput single-cell transcriptome profiling of plant cell types. Cell Rep 27: 2241–2247.e4

Singh M, Al-Eryani G, Carswell S, Ferguson JM, Blackburn J, Barton K, Roden D, Luciani F, Giang Phan T, Junankar S, et al. (2019) High-throughput targeted long-read single cell sequencing reveals the clonal and transcriptional landscape of lymphocytes. Nat Commun 10: 3120

Srivatsan SR, McFalone-Figueroa JL, Ramani V, Saunders L, Cao J, Packer J, Pliner HA, Jackson DL, Daza RM, Christiansen L, et al. (2020) Massively multiplex chemical transcriptomics at single-cell resolution. Science 367: 45–51

Stergachis AB, Debo BM, Haugen E, Churchman LS, Stamatoyannopoulos JA (2020) Single-molecule regulatory architectures captured by chromatin fiber sequencing. Science 368: 1449–1454

Stoeckius M, Hafemeister C, Stephenson W, Houck-Loomis B, Chattopadhyay PK, Sweedlow J, Satija R, Smibert P (2017) Simultaneous epitope and transcriptome measurement in single cells. Nat Methods 14: 865–868

Sulston JE, Schierenberg E, White JG, Thomson JN (1983) The embryonic cell lineage of the nematode Caenorhabditis elegans. Dev Biol 100: 64–119

Volden R, Vollmers C (2020) Highly multiplexed single-cell full-length cDNA sequencing of human immune cells with 10X genomics and R2C2. bioRxiv doi: 10.1101/2020.01.01.902361

Wang N, Zheng J, Chen Z, Liu Y, Dura B, Kwak M, Xavier-Ferrucio J, Lu Y-C, Zhang M, Roden C, et al. (2019) Single-cell microRNA-mRNA co-sequencing reveals non-genetic heterogeneity and mechanisms of microRNA regulation. Nat Commun 10: 95

Wang Y, Huan Q, Chu X, Li K, Qian W (2020) Single-cell transcriptome analyses recapitulate the cellular and developmental responses to abiotic stresses in rice. bioRxiv. doi: 10.1101/2020.01.30.926329

Wendrich JR, Yang B, Vandamme N, Verstaen K, Smet W, Van de Velde C, Minne M, Wybouw B, Mor E, Arents HE, et al. (2020) Vascular transcription factors guide plant epidermal responses to limiting phosphate conditions. Science 370: eaay4970

Wenger AM, Peluso P, Rowell WJ, Chang P-C, Hall RJ, Concepcion GT, Ebler J, Fungtammasan A, Kolesnikov A, Olson ND, et al. (2019) Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome. Nat Biotechnol 37: 1155–1162

Wolock SL, Lopez R, Klein AM (2019) Scrublet: computational identification of cell doublets in single-cell transcriptomic data. Cell Syst 8: 281–291.e9

Xu X, Crow M, Rice BR, Li F, Harris B, Liu L, Demesa-Arevalo E, Lu Z, Wang L, Fox N, et al. (2021) Single-cell RNA sequencing of developing maize ears facilitates functional analysis and trait candidate gene discovery. Dev Cell 56: 557–568.e6

Yan H, Song Q, Lee J, Schiefelbein J, Li S (2020) Identification of cell-type marker genes from plant single-cell RNA-seq data using machine learning. bioRxiv doi: 10.1101/2020.11.22.393165

Zhang TQ, Xu ZG, Shang GD, Wang JW (2019) A single-cell RNA sequencing profiles the developmental landscape of Arabidopsis root. Mol Plant 12: 648–660