

Potentials of single-cell genomics in deciphering cellular phenotypes

Short Title:

Single cell genomics deciphers cell phenotypes

Authors: Abbas Shojaee[1], Michelle Saavedra[1], Shao-shan Carol Huang[1]*

Affiliation:

1. Center for Genomics and Systems Biology
   Department of Biology
   New York University
   12 Waverly Pl, New York, New York 10003

***Author for Correspondence:**

Shao-shan Carol Huang,

12 Waverly Pl, New York, New York 10003.

Phone: 212-998-8286

Email: s.c.huang@nyu.edu

Keywords:

Single cell, RNA sequencing, expression waves, pseudotime, spatial transcriptomics, chromatin accessibility, multi-omics, causal inference, plant biology,

Abstract

Single-cell genomics and particularly single-cell transcriptome profiling by RNA sequencing (scRNA-seq) have transformed the possibilities to relate genes to functions, structures, and eventually phenotypes. We can now observe changes in each cell's transcriptome and among its neighborhoods, interrogate the sequence of transcriptional events, and assess their influence on subsequent events. This paradigm shift in biology enables us to infer causal relationships in these events with high accuracy. Here we review the latest single-cell studies in plants that uncover how cellular phenotypes emerge as a result of the transcriptome process such as waves of expression, trajectories of development and responses to the environment, and spatial information. With an eye on the advances made in animal and human studies, we further highlight some of the needed areas for future research and development, including computational methods.

Introduction

Every study of genotype at heart is motivated by a question about the effects of the genotype on the specific phenotypes. However, as we can more accurately quantify genomic and phenotypic variations, we have started to learn that many phenotypes are a continuum of incremental changes in functionalities and structures rather than discrete entities. Single-cell genomics has been instrumental in creating this paradigm shift by revealing the blurry boundaries of the conventional definition of cell phenotypes in biology.

Omics-scale data acquisition started from analyzing bulk tissues, producing average measures of all the cells collected. The analysis of individual cells began in 2009 by Tang et al. by sequencing RNA in four cell-stage murine blastomeres [1]. Over the past decade, improvements in techniques for isolating and tagging cells and the advancement of next-generation sequencing methods have made it possible to profile individual cells at a large scale and in a highly reproducible way. Unlike bulk measurements, these methods can analyze a collection of cells individually or as subgroups defined by desired criteria. Single-cell profiling of samples taken at multiple time points and spatial locations creates snapshots of the cell states, akin to a series of photos captured from a group dance which could reveal individual players, subgroups, and interplays between them. These high-resolution pictures of transcriptional programs allow characterization and investigation of the continuum of transcriptomic events across cell types along time and space. We can now induce *in vitro, in vivo*, and soon *in silico* perturbations to both the cellular transcriptional machinery and to the environment. By tracing the effects of perturbations, we can ask causal questions on what, why, and how the induced changes in regulatory programs affected particular aspects of cell phenotypes. This is far more powerful than bulk tissue analysis, which is inherently limited to answer questions about co-expression and correlation among vaguely defined groups, often with insufficient accuracy.

Single-cell analysis technologies have advanced our understanding of complex and dynamic regulatory networks underlying cellular phenotypes. In response to cellular signals, chromatin accessibility changes dynamically to shape the subsequent transcriptional programs [2]. Chromatin accessibility can be mapped genome-wide at the single cell level by assay for transposase-accessible chromatin using sequencing (scATAC-seq) [3] and matched to transcriptional profiles captured by single cell RNA-seq (scRNA-seq). Further integration of chromatin accessibility maps, transcriptome profiles and genomic and epigenomic annotation can reveal the upstream regulatory histone modifications, DNA chemical modifications such as methylation, and transcription factor (TF) binding sites, delineating the transcriptional programs that drive cellular fates. Technologies of *in situ* spatial analysis further allow us to investigate transcriptional changes due to interactions with neighboring cells. Beyond the genome and transcriptome, continuous innovations are enabling single-cell proteomics and metabolomics [4] [5] [6] [7]. These advancements provide a comprehensive and high-resolution picture of cellular identity and behavior (Figure 1), which will bring us closer to answer fundamental questions on what, why and how particular cell fates, responses and diseases are decided.

Cell identities are points on a continuum

Using scRNA-seq to study cells during the development or response to environmental conditions reveals that discrete cell identities are indeed points on a continuum of cell state changes. The existence of such continuum has been verified with various independent methods such as marker gene expression [8], DNA barcode editing [9] [10], and comparison with cell identity mutants [11]. Arranging the continuous cell states based on their progression in the underlying process uncovers distinct trajectories along the progression time (pseudotime) that carry the cell to particular phenotypes. Assisted by these methods, scRNA-seq studies in plants recapitulated and refined the cell states and known developmental trajectories, and identified new cell types and differentiation branches. Lopez *et al*. used scRNA-seq to study *Arabidopsis thaliana* stomata lineage and found overlapping and flexible stomatal lineage cell states along a continuum [12]. Similarly, Liu *et al.* identified TF regulatory networks in different developmental stages of stomatal lineage cells using pseudotime analysis and found several novel marker genes, interactions and mutual regulation among key marker genes at different developmental stages [13]. The transcriptional programs that drive the development of specific cell types are shaped by differential chromatin remodeling that results in distinct chromatin accessibility profiles [14, 15] and cell-type specific networks [16]. Recent studies of *A. thaliana* roots using scRNA-seq and scATAC-seq showed that the transcriptome profiles and the accessible chromatin profiles for many cell types were correlated [17] [14] [18], and scATAC-seq can complement scRNA-seq results in identifying the continuum of cell states and types [14]. Using scATAC-seq in maize, Marand and colleagues identified 92 chromatin accessibility states across 165,913 putative cis-regulatory elements in 52 known cell types and showed that the combinatorial accessibility of TFs and their binding motif defines various cell states [18]. The study of maize pluripotent stem cells in shoot apical meristems (SAMs) and its differentiating cellular descendants by Satterlee *et al*. by scRNA-seq showed that cell differentiation followed a continuum of transcriptional states [19]. In *A. thaliana*, Gala *et al*. used CRISPR/dCas9 to repress three histone deacetylases chromatin regulators and showed that the induced chromatin remodeling created a substantially different phenotype by increasing the density and frequency of lateral root development [11]. Technologies that combine large-scale genetic perturbations and scRNA-seq, such as Perturb-seq [20, 21], will enable unbiased characterization and validation of key regulators that drive the cell state trajectories.

Transcriptional response to the environment is heterogeneous across cell types

Plant species' plasticity in adapting to diverse environments emerges from genotype-environment interactions. Single cell analysis allows us to characterize the heterogeneity of transcriptional changes in such genotype-environment interactions at the resolution of the individual cell and cell types. Jean-Baptiste and colleague applied heat stress to *A. thaliana* roots and used scRNA-seq to identify large gene expression changes, particularly in the outer layers such as the epidermis and cortex [22]. Wendrich *et al*. used scRNA-seq to examine *A. thaliana* roots in low phosphate conditions and discovered that increases in the TMO5/LHW TF activity in young xylem cells induced cytokinin biosynthesis, which then diffused through the vasculature and directed both the length and identity of the outer trichoblast cells to support roots to forage the soil for phosphate [23]. Through comparison of scRNAseq profiles of *A. thaliana* roots grown in sucrose versus others, Shulse *et al*. found that the addition of sucrose as a carbon source altered proportions of cell populations and induced gene expression changes highly specific to distinct tissues or cell populations [24]. Studying the effect of low nitrogen and high salinity condition in rice using scRNA-seq, Wang *et al.* found that the induced transcriptional changes were specific to cell types. Interestingly, they discovered that the same set of genes expressed under the low nitrogen condition tended to express differentially during the high salinity treatment, suggesting a rebalancing of common regulatory networks in response to various abiotic stresses [25].

Distinct expression waves underlie the continuity of cell states

In closer examination, the continuous cell state transitions revealed by scRNA-seq is not smooth but a choppy surface made of mostly distinct TF and gene expression waves. These expression waves of highly interconnected TFs and target genes shift the cell state along the developmental trajectory or environmental response. Studying male and female reproductive development in balsam poplar, Cronk and colleagues found successive waves of gene expression corresponding to distinct alterations to the chromatin landscape [26]. Denyer et al. performed scRNA-seq of *A. thaliana* root cells and identified 239 TF corresponding to highly precise gene expression waves that occurred during differentiation and maturation of meristematic cells in unidimensional growth and root hair elongation [27]. Similarly, Liu *et al*. studied over 12,844 individual cells from the cotyledons of five-day-old *A. thaliana* seedlings and found that TF activity and gene expression were very high at the early stage of development of stomatal cells but quickly decreased along the pseudotime [13]. Examining maize male meiosis by scRNA-seq, Nelms *et al*. observed that periods of rapid gene expression occurred during the meiotic prophase, where significant cellular physiology changes took place [28]. They defined the concept of pseudotime velocity to measure the rate of gene expression change along the pseudotime of developmental trajectories. It is conceivable that the continuum of changes that take place during environmental response is also made of explicit expression waves. Gould and colleagues studied *A. thaliana* seedling's response to day and night cycles and observed robust transcription oscillations [29]. Interestingly, they found that the waves of clock gene expression that spread across the plant were seemingly coordinated through cell-cell communication.

Spatial techniques characterize transcriptional programs *in situ*

The spatial structure of cells in tissues and the ultimate tissue functions are governed by interactions and coregulations of transcriptional programs with neighboring cells and the environment. The development of *in situ* spatial transcriptomics has enabled us to characterize these programs in their original tissue context. Experimenting with barley leaf, Solanki and colleagues optimized the sample preparation and tissue pretreatment steps of RNAscope, a multiplex fluorescent RNA *in situ* hybridization method, for plant tissues and achieved tissue section integrity, RNA stability and high-quality signal [30]. Giacomello *et al*. optimized the spatial transcriptomics platform of barcoded oligo-DT microarrays for multiple plant species and for *A. thaliana* detected significant differences in the expression levels of 141 genes between eight inflorescence tissue domains at 100 µm resolution [31]. To study the cellular response to heat stress in *A. thaliana* and *Nicotiana benthamiana* leaf, Alamos and colleagues used PP7 and MS2 technologies for labeling nascent RNA together with quantitative imaging to measure individual cell’s transcriptional changes and discovered tissue level patterns of mRNA accumulation [32]. Many new developments have been made in spatial transcriptomics in animal studies, for instance, the method seqFISH+ could detect 10,000 genes in a single experiment in mouse brain [33]. Implementing these methods in plant biology hold great potential to further our understanding of how a plant is made from its genetic instructions.

Limitations of current methods

Like any technology, scRNA-seq comes with limitations that may impact the conclusions we derive from our studies. Some limitations are the result of the current state of technologies, while others are specific to plant studies. For example, the composition and thickness of the cell wall and large variations of cell size present unique challenges [34]. Widely-used protocols of preparing single cell suspensions for microfluidic-based scRNA-seq platforms usually involve protoplasting, which induces changes in the transcriptome [35] [36] [37]. We invite interested readers to find more detailed accounts on technological challenges in these recent reviews and papers [34, 38] [39] [40] [41] [42]. Addressing this limitation, several recent studies showed that single nucleus RNA-seq is a promising alternative to protoplasting [14, 43] [44]. Even with appropriate preparation, some inherent limitations of single cell experiments must also be overcome. For example, sufficient and unbiased sampling along the real development time and adequate representation of the diverse cell states in a sample are necessary to preserve the fingerprint of rare cell states [45] [46] or rare cells that lead essential changes, such as quiescent center cells in the root.

Identification of cell states, measuring the similarity between them, and aligning them along a pseudotime or pseudospace [35, 47] is at the core of deciphering transcriptional programs using scRNA-seq. However, the assumption that similarity in the transcriptome profile corresponds to developmental similarity should be examined with independent methods [47]. The measure of similarity can also be influenced by highly expressed genes or dominant expression patterns, such as cell cycle progression [47] or gene expression oscillation, masking the key differences of lowly expressed genes [48]. Furthermore, assumptions made during analysis about data distribution, structure and input features, such as predefined marker genes, could strongly influence the number of the identified clusters [49] and the similarity between them and may reduce the chance of revealing new or underrepresented information. Therefore, the design of scRNA-seq experiments and analysis choices should be informed by the particular biological process that we want to identify.

The representation of cell state transitions as a tree or limited graph, a fundamental assumption for constructing pseudotime trajectories, could be a source of bias [50]. As more data accumulates, we learn that groups of cells may take an alternative shortcut in differentiation processes to another part of the lineage hierarchy [47]. These shortcuts take the shape of a directed acyclic or cyclic graph, which the pseudotime method cannot represent. Even using successful graph-based presentations, such as partition-based graph abstraction [51], could yield biased lineage graphs due to the algorithm's assumptions [52]. As a result, caution must be taken in conducting trajectories analyses and interpreting their results.

Conclusion and future directions

In this review, we highlighted several key recent developments and applications of single-cell technologies in plant studies. We envision that the waves of expression, observed in multiple processes, will emerge as a key concept in representing units of transformation in cell states. Each expression wave is the output of a multi-layer, coordinated regulation within the cell and is temporally related to an inward or outward interaction with the environment. Developing and adding functionalities to existing analysis pipelines to identify and resolve the superimposition of expression waves, and mapping them to the underlying regulatory changes and the developmental and environmental signals will be necessary for disentangling the regulatory networks that connect genotype to phenotypes.

Beyond the transcriptome, multimodal single cell technologies could simultaneously capture, in each cell, the state of two or more layers of gene expression machinery such as DNA methylation, chromatin accessibility, RNA expression, protein profile and metabolites [53] [54] [55]. These technologies provide a more comprehensive characterization of cell identity [56] [18], and the interactions between regulatory layers that determine the developmental and environmental responses. Although not yet widely adopted in plant studies, we note that multimodal technologies could enhance the usability of the data produced by unimodal single-cell methods toward an unbiased view of regulatory networks. Unbiased results reduce the uncertainty of computational methods that perform data integration, ultimately leading to more reliable predictions and reducing the cost of validating these predictions.

In single-cell studies, we face a widening gap between the data production rate and our methods' ability to analyze it effectively. Bulk tissue analysis is often constrained to finding correlations between co-expression clusters in a limited number of data matrices. However, integrative single-cell studies have presented challenges and opportunities for inferring regulatory circuits from millions of data points across many connected data matrices. These challenges so far has been met by extending and combining the existing computational methods for parallel data types [57, 58] to learn both the cross-modality variations and cross-cell correlations [57]. Several difficulties faced by these extended methods have been discussed in the recent literature [48] [57-62]. We argue that even when these difficulties are resolved, there is still a fundamental challenge that calls for theoretical expansion. The challenge is that the analysis of multi-omics single-cell data requires simultaneous inference of local interactions, such as the influence of a TF on a particular gene, and global interactions, such as the metabolic effects of that expressed gene on neighboring cells and other tissues. Keeping both the local and global connections in perspective requires improving the tradeoff made by existing methods that prefer one over the other. Effective concurrent inference of local and global connections requires novel theoretical insights beyond correlations and conditional probabilities. We suggest that such novel frameworks would come from innovations in inferring causality in uncontrolled and complex data. Such methods would be able to draw more certain connections between many changing parameters to weed out billions of coincidental proximal and distal correlations that emerge from multi-omics single-cell data. The novel theoretical frameworks for causality would significantly facilitate our endeavors to answer what, how, and why questions that link the intricate regulatory circuits to phenotypes.

Acknowledgments

This work was supported by grants from the National Science Foundation (IOS-1916804) and the National Institutes of Health (R35 GM138143) to S.C.H.

References

1. Tang, F., et al., *mRNA-Seq whole-transcriptome analysis of a single cell.* Nature methods, 2009. **6**(5): p. 377-382.
2. Klemm, S.L., Z. Shipony, and W.J. Greenleaf, *Chromatin accessibility and the regulatory epigenome.* Nature Reviews Genetics, 2019. **20**(4): p. 207-220.
3. Buenrostro, J.D., et al., *Single-cell chromatin accessibility reveals principles of regulatory variation.* Nature, 2015. **523**(7561): p. 486-490.
4. Mora-Castilla, S., et al., *Miniaturization Technologies for Efficient Single-Cell Library Preparation for Next-Generation Sequencing.* Journal of Laboratory Automation, 2016. **21**(4): p. 557-567.
5. Shahi, P., et al., *Abseq: Ultrahigh-throughput single cell protein profiling with droplet microfluidic barcoding.* Sci Rep, 2017. **7**: p. 44447.
6. Kelly, R.T., *Single-cell Proteomics: Progress and Prospects.* Molecular & Cellular Proteomics, 2020. **19**(11): p. 1739-1748.
7. Hansen, R.L. and Y.J. Lee, *High-Spatial Resolution Mass Spectrometry Imaging: Toward Single Cell Metabolomics in Plant Tissues.* The Chemical Record, 2018. **18**(1): p. 65-77.
8. Zhang, T.-Q., et al., *A single-cell RNA sequencing profiles the developmental landscape of Arabidopsis root.* Molecular plant, 2019. **12**(5): p. 648-660.
9. Masuyama, N., H. Mori, and N. Yachie, *DNA barcodes evolve for high-resolution cell lineage tracing.* Current Opinion in Chemical Biology, 2019. **52**: p. 63-71.

10. Spanjaard, B., et al., *Simultaneous lineage tracing and cell-type identification using CRISPR-Cas9-induced genetic scars.* Nature biotechnology, 2018. **36**(5): p. 469-473.
11. Gala, H.P., et al., *A single cell view of the transcriptome during lateral root initiation in Arabidopsis thaliana.* bioRxiv, 2020: p. 2020.10.02.324327.
12. Lopez-Anido, C.B., et al., 2020.
13. Liu, Z., et al., *Global Dynamic Molecular Profiling of Stomatal Lineage Cell Development by Single-Cell RNA Sequencing.* Mol Plant, 2020. **13**(8): p. 1178-1193.
14. Farmer, A., et al., *Single-nucleus RNA and ATAC sequencing reveals the impact of chromatin accessibility on gene expression in Arabidopsis roots at the single-cell level.* Molecular Plant, 2021. **14**(3): p. 372-383.
15. Alexandre, C.M., et al., *Complex Relationships between Chromatin Accessibility, Sequence Divergence, and Gene Expression in Arabidopsis thaliana.* Molecular Biology and Evolution, 2017. **35**(4): p. 837-854.
16. Xu, X., et al., *Single-cell RNA sequencing of developing maize ears facilitates functional analysis and trait candidate gene discovery.* Developmental Cell, 2021. **56**(4): p. 557-568.e6.
17. Dorrity, M.W., et al., *The regulatory landscape of Arabidopsis thaliana roots at single-cell resolution.* bioRxiv, 2020: p. 2020.07.17.204792.
18. Marand, A.P., et al., *A cis-regulatory atlas in maize at single-cell resolution.* bioRxiv, 2020: p. 2020.09.27.315499.
19. Satterlee, J.W., J. Strable, and M.J. Scanlon, *Plant stem-cell organization and differentiation at single-cell resolution.* Proceedings of the National Academy of Sciences, 2020. **117**(52): p. 33689-33699.
20. Jin, X., et al., *In vivo Perturb-Seq reveals neuronal and glial abnormalities associated with autism risk genes.* Science, 2020. **370**(6520).
21. Dixit, A., et al., *Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens.* Cell, 2016. **167**(7): p. 1853-1866 e17.
22. Jean-Baptiste, K., et al., *Dynamics of Gene Expression in Single Root Cells of Arabidopsis thaliana.* Plant Cell, 2019. **31**(5): p. 993-1011.
23. Wendrich, J.R., et al., *Vascular transcription factors guide plant epidermal responses to limiting phosphate conditions.* Science, 2020. **370**(6518): p. eaay4970.
24. Shulse, C.N., et al., *High-Throughput Single-Cell Transcriptome Profiling of Plant Cell Types.* Cell Reports, 2019. **27**(7): p. 2241-2247.e4.
25. Wang, Y., et al., *Single-cell transcriptome analyses recapitulate the cellular and developmental responses to abiotic stresses in rice.* bioRxiv, 2020.
26. Cronk, Q., S. Raju, and K. Bräutigam, *Gene expression trajectories during male and female reproductive development in balsam poplar (<i>Populus balsamifera</i> L.).* Scientific Reports (Nature Publisher Group), 2020. **10**(1).
27. Denyer, T., et al., *Spatiotemporal Developmental Trajectories in the Arabidopsis Root Revealed Using High-Throughput Single-Cell RNA Sequencing.* Dev Cell, 2019. **48**(6): p. 840-852 e5.
28. Nelms, B. and V. Walbot, *Defining the developmental program leading to meiosis in maize.* Science, 2019. **364**(6435): p. 52-56.
29. Gould, P.D., et al., *Coordination of robust single cell rhythms in the Arabidopsis circadian clock via spatial waves of gene expression.* Elife, 2018. **7**: p. e31700.
30. Solanki, S., et al., *Visualization of spatial gene expression in plants by modified RNAscope fluorescent in situ hybridization.* Plant Methods, 2020. **16**(1): p. 71.
31. Giacomello, S. and J. Lundeberg, *Preparation of plant tissue to enable Spatial Transcriptomics profiling using barcoded microarrays.* Nature Protocols, 2018. **13**(11): p. 2425-2446.

32. Alamos, S., et al., *Quantitative imaging of RNA polymerase II activity in plants reveals the single-cell basis of tissue-wide transcriptional dynamics.* bioRxiv, 2020: p. 2020.08.30.274621.
33. Eng, C.-H.L., et al., *Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH+.* Nature, 2019. **568**(7751): p. 235-239.
34. Shaw, R., X. Tian, and J. Xu, *Single-Cell Transcriptome Analysis in Plants: Advances and Challenges.* Molecular Plant, 2021. **14**(1): p. 115-126.
35. McFaline-Figueroa, J.L., C. Trapnell, and J.T. Cuperus, *The promise of single-cell genomics in plants.* Current Opinion in Plant Biology, 2020. **54**: p. 114-121.
36. Birnbaum, K., et al., *A Gene Expression Map of the Arabidopsis Root.* Science, 2003. **302**(5652): p. 1956-1960.
37. Jean-Baptiste, K., et al., *Dynamics of gene expression in single root cells of Arabidopsis thaliana.* The plant cell, 2019. **31**(5): p. 993-1011.
38. Luo, C., A.R. Fernie, and J. Yan, *Single-Cell Genomics and Epigenomics: Technologies and Applications in Plants.* Trends Plant Sci, 2020. **25**(10): p. 1030-1040.
39. Iqbal, M.M., et al., *Status and Potential of Single-Cell Transcriptomics for Understanding Plant Development and Functional Biology.* Cytometry Part A, 2020. **97**(10): p. 997-1006.
40. Rich-Griffin, C., et al., *Single-cell transcriptomics: a high-resolution avenue for plant functional genomics.* Trends in plant science, 2020. **25**(2): p. 186-197.
41. Muhammad, II, et al., *RNA-seq and ChIP-seq as Complementary Approaches for Comprehension of Plant Transcriptional Regulatory Mechanism.* Int J Mol Sci, 2019. **21**(1).
42. Bai, D., J. Peng, and C. Yi, *Advances in single-cell multi-omics profiling.* RSC Chemical Biology, 2021.
43. Tian, C., et al., *Single-nucleus RNA-seq resolves spatiotemporal developmental trajectories in the tomato shoot apex.* bioRxiv, 2020: p. 2020.09.20.305029.
44. Long, Y., et al., *FlsnRNA-seq: protoplasting-free full-length single-nucleus RNA profiling in plants.* Genome biology, 2021. **22**(1): p. 1-14.
45. Wagner, D.E., et al., *Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo.* Science, 2018. **360**(6392): p. 981-987.
46. Fischer, D.S., et al., *Inferring population dynamics from single-cell RNA-sequencing time series data.* Nature Biotechnology, 2019. **37**(4): p. 461-468.
47. Tritschler, S., et al., *Concepts and limitations for learning developmental trajectories from single cell genomics.* Development, 2019. **146**(12): p. dev170506.
48. Weinreb, C., et al., *Fundamental limits on dynamic inference from single-cell snapshots.* Proceedings of the National Academy of Sciences, 2018. **115**(10): p. E2467-E2476.
49. Camara, P.G., *Methods and challenges in the analysis of single-cell RNA-sequencing data.* Current Opinion in Systems Biology, 2018. **7**: p. 47-53.
50. Saelens, W., et al., *A comparison of single-cell trajectory inference methods.* Nature Biotechnology, 2019. **37**(5): p. 547-554.
51. Wolf, F.A., et al., *PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells.* Genome Biology, 2019. **20**(1): p. 59.
52. Setty, M., et al., *Characterization of cell fate probabilities in single-cell data with Palantir.* Nature Biotechnology, 2019. **37**(4): p. 451-460.
53. Ma, A., et al., *Integrative methods and practical challenges for single-cell multi-omics.* Trends in Biotechnology, 2020.
54. Chen, G., B. Ning, and T. Shi, *Single-Cell RNA-Seq Technologies and Related Computational Data Analysis.* Front Genet, 2019. **10**: p. 317.
55. Zhu, C., S. Preissl, and B. Ren, *Single-cell multimodal omics: the power of many.* Nat Methods, 2020. **17**(1): p. 11-14.

56. Leonavicius, K., et al., *Multi-omics at single-cell resolution: comparison of experimental and data fusion approaches.* Current Opinion in Biotechnology, 2019. **55**: p. 159-166.
57. Ma, A., et al., *Integrative Methods and Practical Challenges for Single-Cell Multi-omics.* Trends in Biotechnology, 2020. **38**(9): p. 1007-1022.
58. Dugourd, A., et al., *Causal integration of multi-omics data with prior knowledge to generate mechanistic hypotheses.* bioRxiv, 2020: p. 2020.04.23.057893.
59. Jamil, I.N., et al., *Systematic Multi-Omics Integration (MOI) Approach in Plant Systems Biology.* Front Plant Sci, 2020. **11**: p. 944.
60. Todorov, H., et al., *Network Inference from Single-Cell Transcriptomic Data*, in *Gene Regulatory Networks: Methods and Protocols*, G. Sanguinetti and V.A. Huynh-Thu, Editors. 2019, Springer New York: New York, NY. p. 235-249.
61. Gross, T. and N. Blüthgen, *Identifiability and experimental design in perturbation studies.* Bioinformatics, 2020. **36**(Supplement_1): p. i482-i489.
62. Colomé-Tatché, M. and F.J. Theis, *Statistical single cell multi-omics integration.* Current Opinion in Systems Biology, 2018. **7**: p. 54-59.

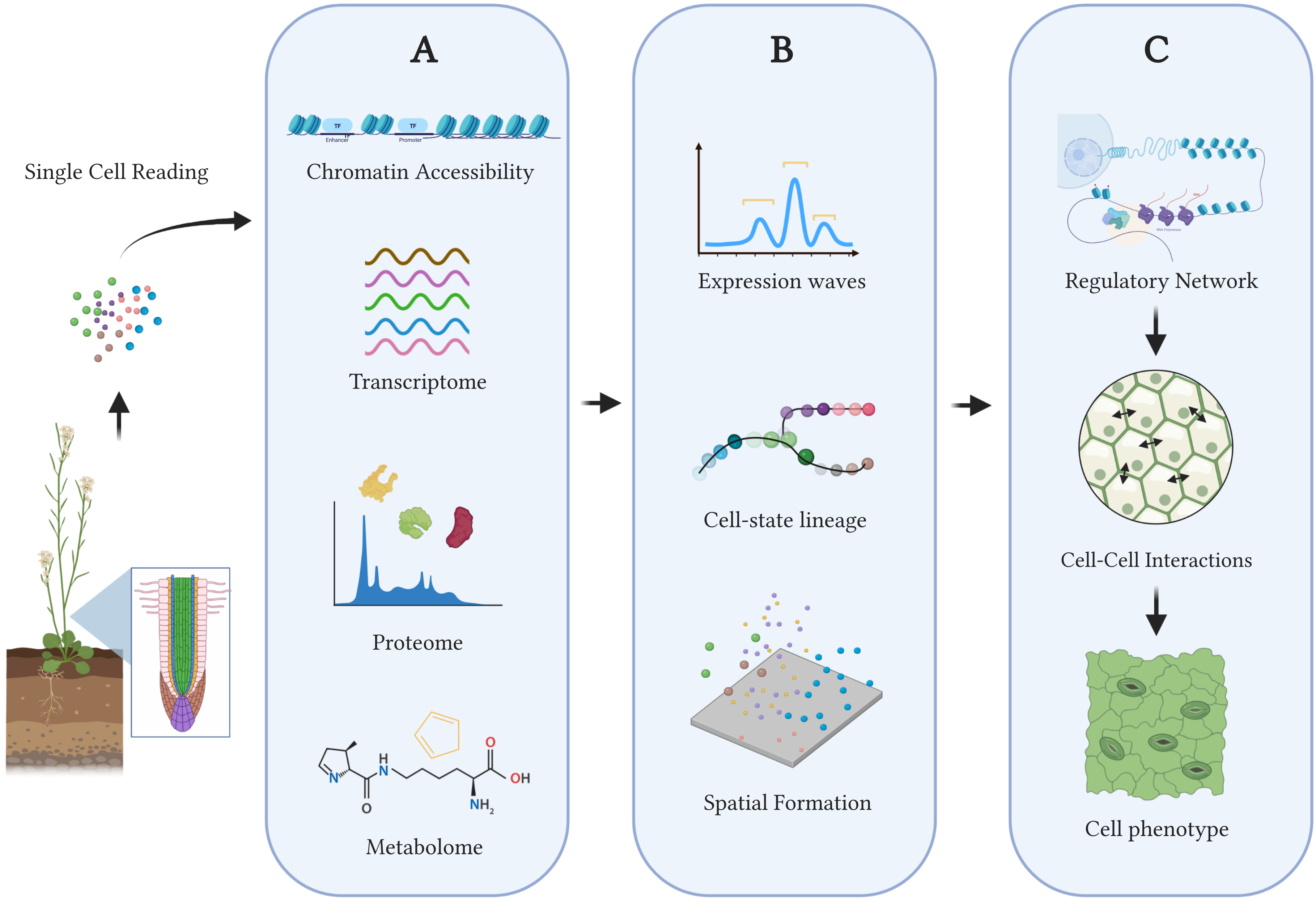
Single Cell Reading
A
Chromatin Accessibility
Transcriptome
Proteome
Metabolome
B
Expression waves
Cell-state lineage
Spatial Formation
C
Regulatory Network
Cell-Cell Interactions
Cell phenotype

Figure 1: Overview of the role of single cell technologies in understanding transcriptome machinery and cellular phenotypes

Single-cell data acquisition technologies (A), Conceptual frameworks and methods for deciphering genotype to phenotype translation (B), cellular phenotype emerges from the interaction of its regulatory network with developmental and environmental signals(C). The figure is created with BioRender.com.