Breaking Boundaries with Transformative Single-Cell Technologies in Plant Research

Somsole Bharath[1]* and Uttarayan Dasgupta[2]

[1,2]Ph.D. Research scholar, Division of Genetics, ICAR – Indian Agricultural Research Institute, New Delhi-110012

Corresponding Author
Somsole Bharath
Email: bharathsomsole@gmail.com

OPEN ACCESS

Keywords

Single-cell, Multiomics, Cell Technology

How to cite this article:

Bharath, S. and Dasgupta, U. 2024. Breaking Boundaries with Transformative Single-Cell Technologies in Plant Research. *Vigyan Varta* 5(7): 18-22.

ABSTRACT

Single-cell technologies are revolutionizing our understanding of cellular diversity by examining individual cell genomes, epigenomes, transcriptomes, and proteomes. These methods vary in accuracy and usability, with well-established techniques like single-cell RNA sequencing (scRNA-seq) and emerging multiomics approaches that combine different molecular layers. High-throughput cell isolation techniques, such as microfluidics, contrast with more labor-intensive methods like micromanipulation. Effective cell barcoding, especially combinatorial indexing, is key for sequencing efficiency. Single-cell genome amplification methods like MDA and LIANTI are being refined for better accuracy and coverage, while epigenome sequencing techniques face challenges due to limited DNA copies per cell. Transcriptome sequencing methods, including Smart-seq and Drop-seq, offer varying sensitivity and accuracy. Integrative approaches like scTrio-seq, which combine genomic, epigenetic, and transcriptional data, are enhancing our understanding of cellular behavior. Future advancements aim for "omni-omics," capturing a complete molecular and spatial picture of cells to fully understand their properties and lineage histories.

INTRODUCTION

A cell's characteristics are mainly influenced by the interaction of its genome, epigenome, transcriptome, and proteome. Single-cell technologies now allow researchers to examine these aspects at an individual level, revealing previously unnoticed differences between cells. These methods vary in accuracy, specificity, and

usability, and it's crucial to consider these factors when analyzing data. While some single-cell techniques, such as single-cell RNA sequencing (scRNA-seq), are well-known, others are still emerging. Multiomics approaches are developing, enabling the simultaneous capture of various molecular layers from a single cell, like combining genomics with transcriptomics or epigenomics with transcriptomics. This comprehensive method provides a better understanding of cellular functions and interactions. Analyzing different omic layers from the same cell can identify their relationships and offer insights into cell development and differentiation. For instance, tracking DNA mutations can outline a cell's history, while transcriptome data can highlight developmental pathways. Integrating this with epigenomic information reveals important regulatory processes. Additionally, the spatial arrangement of cells within tissues is essential for comprehending their roles and should be included as a significant data layer.

Key Techniques in Single-Cell Omics: Cell Isolation and Barcoding

Choosing a method for isolating single cells is crucial for single-cell omics studies. Isolation techniques differ in throughput, ease of use, and phenotypic data captured, as well as the proportion of cells retained. Low-throughput methods like manual micromanipulation, laser capture microdissection, raman tweezers, and patch clamp can isolate specific cells and keep spatial data but are labor-intensive and limited to tens or hundreds of cells per study.

Higher throughput can be achieved with fluorescence-activated cell sorting (FACS) from cell suspensions followed by plate-based processing, isolating thousands of cells into microliter volumes of lysis buffer. However, this method loses spatial information.

Microfluidic devices increase throughput to tens of thousands of cells using channels and chambers controlled by pressure valves, droplets in microfluidic chips, or nanowells. This scales down reaction volumes, reducing costs and making sequencing the main expense. These methods also require cell suspensions, losing spatial information. Though phenotypic data from FACS is lost, newer methods capture this information differently.

Cell barcoding is vital for sequencing multiple single-cell libraries together. Plate-based methods add barcodes during the final PCR step, creating a combined library pool. Microfluidic techniques add barcodes earlier, allowing earlier library preparation. Combinatorial indexing sorts fixed cells into wells, each getting a unique barcode. Cells are pooled, re-sorted, and barcoded again, uniquely tagging thousands of cells with minimal duplication, though many cells are lost. For example, in single-cell combinatorial indexing RNA sequencing, 150,000 *C. elegans* larvae produced six 96-well plates, resulting in 42,035 single-cell transcriptomes (Chappell *et al*., 2018).

Individual Cell Genome Sequencing

Techniques for single-cell whole genome amplification (scWGA), such as degenerated oligonucleotide primer (DOP)-PCR and multiple displacement amplification (MDA), have been used extensively. Recently, novel methods like Multiple Annealing and Looping Based Amplification Cycles (MALBAC), emulsion multiple displacement amplification (eMDA), Linear amplification via Transposon Insertion (LIANTI), single-stranded sequencing using microfluidic reactors (SISSOR), and Microfluidics-based Enrichment of Single-Target Amplified Regions with an ultra-high throughput approach (META-CS) have emerged. MDA, employing Bacteriophage phi29 DNA polymerase for isothermal amplification, boasts high fidelity, detecting single nucleotide

variants (SNVs) accurately, albeit with exponential amplification leading to errors and reduced coverage uniformity. LIANTI employs T4 RNA polymerase for linear amplification, enhancing uniformity and accuracy.

META-CS and SISSOR improve accuracy by utilizing information from both DNA strands, a feature also found in MDA through single-nucleotide polymorphism (SNP) data, albeit at the cost of fewer informative SNVs. Research is ongoing to refine single-cell genome sequencing for lineage tracing.

Uniformity is gauged by the coefficient of variation of sequence bias. LIANTI exhibits the highest uniformity at a bin size of 1,000 kb, while eMDA, MALBAC, META-CS, and DOP-PCR range between 0.1 and 0.15, and MDA exceeds 0.21. Accuracy, measured by false positive rate (FPR), is lowest in META-CS and SISSOR ($<2.4 \times 10^{-8}$), followed by LIANTI (5.4×10^{-6}), MDA and eMDA (1.3×10^{-4}), MALBAC (3.8×10^{-4}), and DOP-PCR (9.6×10^{-4}). LIANTI achieves the highest genome coverage (95% with 83 Gb data), with META-CS, MDA, and eMDA following, while SISSOR and DOP-PCR offer lower coverage. Experimental ease varies, with DOP-PCR, MDA, MALBAC, and META-CS simpler than eMDA, SISSOR, and LIANTI (Wen & Tang, 2022).

Individual Cell Epigenome Sequencing

Analyzing the epigenome of a single cell is challenging because each diploid cell contains only two DNA copies. This differs from single-cell transcriptome analysis, which has the advantage of multiple mRNA copies per gene. High sensitivity in enzymes or chemicals is essential for successful epigenome sequencing. To minimize DNA loss, the experimental process should be simple, and early cell barcoding enables pooling hundreds or thousands of cells, increasing throughput. Automation is facilitated by methods using droplets and microchips, and combinatorial indexing further boosts throughput. Tn5 transposase is crucial due to its sensitivity, simplicity, early barcoding capability, and high throughput, making it central to recent advances. Tools like chromVAR and ArchR address the challenges of sparse epigenomic data.

DNA methylation is another key epigenetic layer with patterns specific to cell types. Techniques for single-cell DNA methylome sequencing, such as reduced representation bisulfite sequencing (RRBS) and post-bisulfite adaptor tagging (PBAT), have been established. PBAT's mapping efficiency and throughput are enhanced with 3’ tagging methods like adaptase and TdT tailing, but coverage decreases due to a single round of random amplification. The sci-MET method increases throughput with combinatorial indexing, adding barcodes through Tn5 transposon and random priming. Traditional RRBS targets CpG regions between MspI (CCGG) sites, covering much of the genome. The single-cell extended-representation bisulfite sequencing (scXRBS) method uses adaptors ligated to single MspI sites, balancing coverage and enrichment of important genomic regions (Chappell *et al*., 2018).

Individual Cell Transcriptome Sequencing

The transcriptome plays a crucial role in cell identity and function, defining cell types and states. Analyzing single-cell transcriptomes reveals transcriptional noise and regulated heterogeneity. This is essential for understanding gene coexpression and exclusive expression. Transcripts vary greatly, with many at low levels (5–20 copies per cell), and human cells contain less than 1 pg of mRNA.

Protocols for scRNA-seq of eukaryotic polyadenylated mRNA vary in sensitivity and

accuracy, with synthetic spike-in RNAs like ERCCs aiding comparison. Methods include oligo(dT) primers and MMLV reverse transcriptases for selective amplification, and in vitro transcription with T7 promoters for linear amplification. Plate-based methods like Smart-seq and CEL-seq, and microfluidics-based approaches like Drop-seq, InDrop, and Seq-well process many cells. Combinatorial indexing also handles large cell numbers but requires fixation.

Other techniques, like SINC-seq, separate and quantify nuclear and cytoplasmic mRNA, offering insights into posttranscriptional regulation. Noncoding RNAs are less explored, though some long noncoding RNAs and small RNAs can be captured at the single-cell level (Fu *et al*., 2015).

Unified Single-Cell Genome and Transcriptome Sequencing

Methods that analyze DNA and RNA from the same single cell can clearly link genetic changes to transcriptional variations, aiding in understanding cell development and diseases. One method, DNA–RNA sequencing (DR-seq), amplifies gDNA and mRNA together before separating them for sequencing, though RNA amplification can contaminate DNA sequencing data. Another method, genome and transcriptome sequencing (G&T-seq), separates mRNA from gDNA using oligo(dT)-coated beads, allowing independent processing and automation, though potential nucleic acid loss is a concern. A third approach separates the nucleus and cytoplasm, analyzing nuclear DNA and cytoplasmic RNA separately. This method can result in some RNA loss but offers flexibility in protocol choice and potential for automation.

Unified Single-Cell Epigenome and Transcriptome Sequencing

The role of the epigenome in cell function is crucial, impacting how genes are read and transcribed. Understanding the link between epigenetic and transcriptional changes in single cells remains challenging but is vital for deciphering cell behavior during differentiation.

scTrio-seq analyzes genomic variations, DNA methylation, and transcriptomes in individual cells. It gently lyses cells, separates mRNA for RNA-seq, and analyzes DNA methylation and copy number. scMT-seq, like scTrio-seq, isolates nuclei for DNA methylation and amplifies cytoplasmic mRNA separately. scM&T-seq separates DNA and RNA and profiles CpG methylation using scBS-seq. scGEM isolates cells, converts mRNA to cDNA, and removes chromatin proteins for gene expression and methylation analysis. scNOMe-seq and scCOOL-seq evaluate chromatin accessibility and DNA methylation. scNMT-seq combines these with transcriptome sequencing for a holistic view of cellular epigenetics and gene expression.

Forward-looking Strategies

Emergence of new methods characterizes the rapid progression of single-cell multiomics. Expansion of large datasets from single-cell RNA sequencing to multiomics is anticipated. Enhancements in sensitivity and precision of scRNA-seq and scDNA-seq could enhance multiomics. Long-read technologies may enable simultaneous detection of DNA sequence, methylation, and structural variants. Spatial information capture might broaden to include diverse molecules. Challenges persist in single-cell metabolomics and cost constraints. Crucial is the development of computational methods for integrating multiple omics layers. Future multiomics endeavors aim for thorough cellular characterization, covering all molecules, traits, and lineage history. The ultimate goal is achieving "omni-omics," comprehensively capturing cellular molecules, traits, and lineage history (Mo & Jiao, 2022).

REFERENCES

Chappell, L., Russell, A. J. C., & Voet, T. (2018). Single-Cell (Multi)omics Technologies. *Annual review of genomics and human genetics*, *19*, 15–41. https://doi.org/10.1146/annurev-genom-091416-035324

Fu, Y., Li, C., Lu, S., Zhou, W., Tang, F., Xie, X. S., & Huang, Y. (2015). Uniform and accurate single-cell sequencing based on emulsion whole-genome amplification. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(38), 11923–11928. https://doi.org/10.1073/pnas.1513988112

Mo, Y., & Jiao, Y. (2022). Advances and applications of single-cell omics technologies in plant research. *The Plant journal: for cell and molecular biology*, *110*(6), 1551–1563. https://doi.org/10.1111/tpj.15772

Wen, L., & Tang, F. (2022). Recent advances in single-cell sequencing technologies. *Precision clinical medicine*, *5*(1), pbac002. https://doi.org/10.1093/pcmedi/pbac002