Emerging Topics in Life Sciences (2021)
https://doi.org/10.1042/ETLS20200274

PORTLAND PRESS

Review Article

Space: the final frontier — achieving single-cell, spatially resolved transcriptomics in plants

Sai Guna Ranjan Gurazada[1,2,*], Kevin L. Cox, Jr[1,3,*], Kirk J. Czymmek[1,4] and Blake C. Meyers[1,5]

[1]Donald Danforth Plant Science Center, St. Louis, MO 63132, U.S.A.; [2]Delaware Biotechnology Institute, University of Delaware, Newark, DE 19711, U.S.A.; [3]Howard Hughes Medical Institute, Chevy Chase, MD 20815, U.S.A.; [4]Advanced Bioimaging Laboratory, Donald Danforth Plant Science Center, St. Louis, MO 63132, U.S.A.; [5]Division of Plant Sciences, University of Missouri, Columbia, MO 65201, U.S.A.

Correspondence: Blake C. Meyers (bmeyers@danforthcenter.org)

Single-cell RNA-seq is a tool that generates a high resolution of transcriptional data that can be used to understand regulatory networks in biological systems. In plants, several methods have been established for transcriptional analysis in tissue sections, cell types, and/or single cells. These methods typically require cell sorting, transgenic plants, protoplasting, or other damaging or laborious processes. Additionally, the majority of these technologies lose most or all spatial resolution during implementation. Those that offer a high spatial resolution for RNA lack breadth in the number of transcripts characterized. Here, we briefly review the evolution of spatial transcriptomics methods and we highlight recent advances and current challenges in sequencing, imaging, and computational aspects toward achieving 3D spatial transcriptomics of plant tissues with a resolution approaching single cells. We also provide a perspective on the potential opportunities to advance this novel methodology in plants.

Introduction

Identifying and quantifying differentially expressed genes is a primary activity of many projects focused on generating a molecular and cellular understanding of development and responses to internal or external factors. These studies utilize diverse experimental or environmental conditions and are fundamental to understanding phenotypic variation in organisms. A powerful, commonly used technology to characterize gene expression or abundance is a high-throughput, transcriptome profiling method known as RNA sequencing (RNA-seq) [1]. This has become a useful tool to answer fundamental questions in biology and its adoption has been accelerated because it is cost-effective and accurate. RNA-seq also provides a rich source of data because the technique can be adapted to many methods. For example, RNA-seq is being used to understand regulatory networks in development, responses to stresses, etc., often via computational modeling [2]. Modeled networks are used to predict new or additional functions for previously described genes or their products. Additionally, gene networks can predict potential interactions among genes and their protein products, while providing a systems-level view of the molecular mechanisms of a biological process [3]. However, a more complete understanding of these biological processes in eukaryotic organisms requires high-resolution spatial characterization of specialized tissue domains.

The advances in RNA-seq techniques have provided scientists the tools to conduct transcriptional analyses at a cellular resolution. This includes facilitating the study of transcriptional activity at a single cell level, known as single-cell RNA-seq (scRNA-seq). In addition, scRNA-seq typically generates datasets that contain thousands of transcripts per cell, and that identify cell types, including transcripts that are low in abundance or are previously undescribed. While this is a substantial increase in resolution compared with a whole tissue or 'bulk' RNA-seq approach, the drawback with most of these scRNA-seq technologies is that they are unable to spatially resolve the transcriptional profiles within the context of tissues — that is, where the cell types that are expressing specific transcripts are

*These authors contributed equally to this work.

Received: 15 November 2020
Revised: 5 January 2021
Accepted: 11 January 2021

Version of Record published: 1 February 2021

© 2021 The Author(s). Published by Portland Press Limited on behalf of the Biochemical Society and the Royal Society of Biology

located, broadly and relatively, is unclear. A method known as spatial transcriptomics accomplishes this by resolving the location of gene expression while maintaining relevant cellular adjacencies via tissue sections including specific cell types. The technique utilizes arrayed, barcoded capture probes to make thousands of positionally defined libraries from a single sample or tissue section [4]. Characterization of transcriptome profiles while retaining spatial information within the tissue permits more comprehensive and contextual studies for understanding regulatory networks. Spatial transcriptomics originated in mammalian systems and has substantially improved over the last few years. While this method has also been used in plants [5], the achieved resolution is not yet comparable to that in mammals — mainly, we argue, due to a lag in adoption rather than any inherent limitations specific to plants.

A variety of methods have been established in plants that allow transcriptional analysis at the level of individual tissue sections and/or cells. Yet, in order to generate datasets with a high spatial resolution of transcripts, the spatial context of the plant tissue is often destroyed. Transcriptional analysis experiments in plant roots are an example of this, as there are known marker genes that allow for single-cell transcriptomic analyses in Arabidopsis [6–9]. However, the data is mapped back to the roots through cell sorting instead of precise spatial data. Some other methods include fluorescence-activated cell sorting (FACS) [10–12], isolation of nuclei tagged in specific cell types (INTACT) [13], and laser capture microdissection (LCM) of tissue sections [14,15]. Alternatively, approaches like Drop-seq [16] or the Chromium system (10× Genomics) use large sets of isolated, protoplast cells. While these methods have been successful, they also have inherent limitations. For example, the FACS and INTACT methods require transgenic plants — challenging for some species. Additionally, FACS, Drop-seq, and the Chromium system require protoplasting of cells, a method that inherently reduces spatial data from the cells, while also potentially stressing the cells and causing artifacts due to handling. LCM is laborious, can be limited in quality by the cell type, and can damage RNA. Technologies that do offer high spatial resolution for RNA localization lack breadth in the number of transcripts characterized. That is, one of the current best ways to obtain precise spatial data in plants is to perform fluorescent *in situ* hybridization (FISH) microscopy, but this is typically performed one gene at a time, which greatly limits throughput [17,18].

In this review, we highlight the current advances and challenges, experimentally and computationally, for achieving 3-D spatial transcriptomics in plants, with a particular focus on how this method might approach single-cell resolution. We also identify and discuss potentially novel opportunities to further advance spatial transcriptomics as a final frontier in plant transcriptomics: the opportunity to acquire spatially resolved data at single-cell resolution.

Experimental challenges

While spatial transcriptomics methods have developed and expanded in recent years, these methods come with many challenges. These range from sample and template preparation, and RNA transfer strategies, to downstream computational analyses (discussed later) — all of which may require special consideration when working with plants. First, we will describe the process of performing 'wet lab' aspects of this technology and its resulting challenges, before offering potential strategies and solutions to overcome these hurdles.

Resolution and space

While the newest methods of spatial transcriptomics have received attention recently, approaches to spatially resolve gene expression in tissue sections and cells have been around for decades. Techniques such as LCM and FISH are commonly used in plant and mammalian studies and can achieve single-cell and subcellular resolution. However, the area of tissue that can be reasonably accessed, probed, and analyzed in an experiment is relatively limited. When the microarray-based system was first introduced in 2016 for mammalian systems, the objective was to capture larger spatial content in the tissue section, at sufficiently high resolution, to capture data about the distribution of gene expression [4,19]. Indeed, this approach allowed numerous regions containing different tissues or cell layers to be analyzed in an experiment without prior knowledge of the genes expressed in each represented cell type. The primary drawback of the 2016 advance was that the resolution was limited to ~100 μm, as the arrays were composed of 100 μm circular spots with a center-to-center distance of 200 μm, making this far from single-cell resolution. While the platforms (focused primarily on mammalian tissues) have advanced to produce significantly higher resolutions, the approach in plants has yet to be applied beyond the original 100 μm platform [5]. One reason that the technology has lagged for plant applications compared with mammals is related to the cell wall. The majority of high resolution, spatially resolved methods,

© 2021 The Author(s). Published by Portland Press Limited on behalf of the Biochemical Society and the Royal Society of Biology

https://doi.org/10.1042/ETLS20200274

including Slide-seq [20] and High Definition Spatial Transcriptomics [21], would likely require protoplasting, an approach not readily achieved in some plant systems, but is required for both capture methods which use beads instead of probes.

To achieve single-cell, spatial transcriptomics in plants, a potential approach is to continue to increase the resolution of the microarray-based approach (i.e. smaller spots, more densely spaced). The microarray approach is being commercialized by 10× Genomics using smaller printed spots, called Visium Spatial Gene Expression, with a spot size of 55 μm, and a 100 μm center-to-center distance between capture probes [22]. While Visium has yet to be used on plant tissues (no data are reported, as this is written), it is likely to be applicable as the protocols are similar to the existing 100 μm resolution technology [23]. Conceivably, the development of an instrument that can synthesize more densely packed microarrays with high fidelity oligo probes would be key to production of a platform for highly spatially resolved transcriptomics in plants.

Preparation of tissue sections

Plant cells have some unique challenges for transcriptional analysis, including the nature of certain cellular structures, such as the cell wall, vacuole, and chloroplasts [24]. Additionally, plant tissues often contain apoplastic pockets of air that can lead to problems in cryosectioning. Furthermore, if the target tissue is leaves, inherent curvature results in the requirement of more sections to completely capture the tissue volume (i.e. adaxial surface to abaxial surface). Lastly, the thickness of the section needs to be considered. Most spatial transcriptomics methods, both in plant and mammalian systems, use tissue sections with a thickness ranging between 10 and 20 μm [20,23]. Since plants have walled cells with complex structures, sections that are too thick could potentially lead to reduced mRNA release or capture while sections too thin are prone to sectioning artifacts or simply low transcript levels. Moreover, the large central vacuole of plant cells restricts cytoplasmic RNA to a thin, peripheral cytoplasmic strip that may be problematic depending on its orientation relative to the array surface.

There are some possible solutions to these challenges. For example, the approach typically used for leaf or root tissues is to apply a cryo-compatible embedding medium to encapsulate the plant structure [23]. The embedding medium provides a solid matrix to provide support and improve cutting properties. Therefore, applying an additional thin layer of this embedding medium onto the specimen before cryosectioning will better maintain the morphology and improve section quality [23]. Enzymatic improvements could also facilitate removal of the cell wall [23]. Additionally, modifications to imaging and RNA capture procedures could serve as solutions to the challenges of tissue preparations as outlined in the following sections.

Imaging

Imaging is a key part of spatial transcriptomics. The function of barcoding individual spots in a large array for transcriptional analysis is to match the captured mRNA to corresponding and specific plant tissue structures (i.e. meristem, epidermis, vascular tissue). In the simplest form, plant cell walls and boundaries can be delineated in unstained samples using standard brightfield light optics, or further enhanced with stains such as toluidine blue [23] or hematoxylin and eosin [19]. In some instances, autofluorescence (i.e. chloroplasts or cell walls) could be leveraged to provide positional information or indicate cellular type or status. Moreover, imaging can be effectively used to assess the quality of sections and the retention of targets. The release of plant material can be assessed by inclusion of fluorescent nucleotides during cDNA synthesis [19,23]. Imaging can also be used to visualize array spots that are labeled by hybridization of fluorescently-labeled probes. Additional spatial targeting could be achieved by the selective expression of genetically encoded fluorescent proteins in plants and/or associated microbes. Regardless of the contrasting method, larger tissue sections will require multi-image montage maps and fiducial reference points that serve to register to the fluorescently-labeled spot arrays [19] and their respective cells/tissues and accordant transcripts. Ideally, for 3D transcriptomics, each imaged serial section must be registered to its corresponding address. Optimization of the quality of the upstream sample and sectioning will aid significantly in the downstream data processing and visualization. However, moderate section-to-section distortion and compression artifacts are practically not avoidable but can be addressed with non-linear transformation and image alignment tools [25].

RNA capture

An advantage of spatial transcriptomics technologies is the ability to use barcoded oligo probes to capture RNA from fresh-frozen tissue. To briefly explain this process, the section is placed on a microarray coated with

© 2021 The Author(s). Published by Portland Press Limited on behalf of the Biochemical Society and the Royal Society of Biology

Emerging Topics in Life Sciences (2021)
https://doi.org/10.1042/ETLS20200274

reverse transcription oligo(dT) primers, with each spot having a unique barcode (Figure 1). After fixing and imaging the sample on the array, the mRNA is released from the cells, captured by the oligo probe via its polyT tail, and reverse transcribed into cDNA before subsequent elution and use for library preparation and RNA-seq. The resulting gene expression data are mapped back to specific locations within the tissue section via the barcodes embedded within the array. As previously mentioned, releasing mRNA from plant cells has the added difficulty due to a physical barrier — their cell walls. Therefore, a cocktail of enzymes that specifically hydrolyzes the components of cell walls must be used and has been demonstrated to work [5,23]. However, plants also possess secondary metabolites that can form complexes with the released mRNA [26]. This, combined with the need for harsh enzymes to break the cell wall, potentially limits the efficiency of RNA capture. While this can be overcome by adding chemicals (i.e. poly 1-vinylpyrrolidone-2) [23], another potential solution is to force the released transcripts immediately down towards the slide surface where they can be captured by the oligo probes, perhaps via electrophoresis, to increase the RNA capture efficiency. This method could operate similarly to tissue printing methods in plants [27]. Nevertheless, additional modifications to the protocol would be needed for this theoretical approach to be successful in practice.

Computational challenges

The study of transcriptional activity and gene regulatory networks in plants using RNA-seq technologies has heavily relied on computational algorithms and bioinformatics tools. Typically, millions of short RNA reads are processed to remove sequencing adapters and low quality reads before being mapped to known reference gene transcripts to quantify individual gene expression. Depending on the experiment, an appropriate statistical

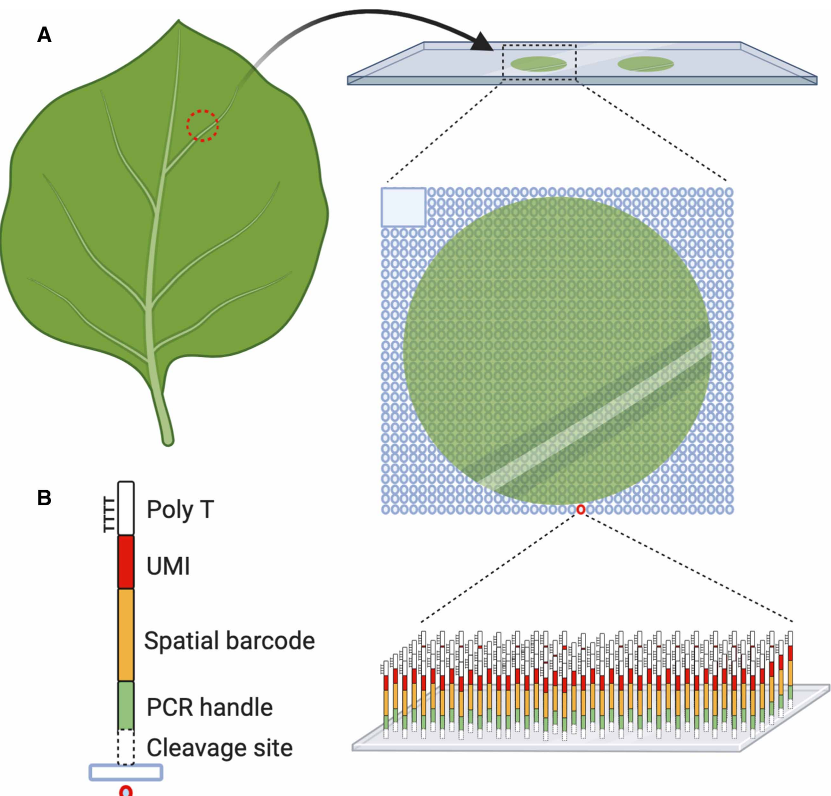

Figure 1. Microarray and oligo design.
(**A**) Tissue section from a leaf sample of *N. benthamiana* is fixed and placed on a microarray with gridded, barcoded spots. The section is processed to permeabilize the cellular material, and convert the mRNA to cDNA (not shown). (**B**) Millions of oligonucleotide probes are found at each spot on the microarray, and used to capture cDNA from the tissue section. Each probe sequence consists of a spot-specific barcode followed by a unique molecular identifier (UMI); the PCR handle includes the sequences necessary to amplify and sequence the RNA-seq library. Created with BioRender.com.

© 2021 The Author(s). Published by Portland Press Limited on behalf of the Biochemical Society and the Royal Society of Biology

https://doi.org/10.1042/ETLS20200274

method is applied to normalize gene expression counts for effective comparison across samples so that any variation in sequencing depth, sample size and batch effects are neutralized. Finally, differentially expressed genes are identified using computational methods [28–32] by comparing samples across developmental stages, stress response or disease progression over time, etc. The core computational and analytical strategies to study whole transcriptomes of both mammalian and plant species using bulk RNA-seq have been well established over the past several years [33,34].

The recent advent of scRNA-seq technologies has allowed the study of transcriptional activity at a much higher resolution, including at the cellular or even subcellular levels when combined with *in situ* hybridization or sequencing. While this provides exciting new opportunities, it also presents new computational challenges [35]. The goal of spatial transcriptomics is to combine spatial context with cellular transcriptomics to enable the visualization of the variation of gene expression across cells in a tissue section, to see where the identified cell types are spatially located within the section, and to gain a better visual understanding of the regulatory networks at play in a region of interest based on the gene expression of interacting cells [36]. Given the high resolution of scRNA-seq and spatial transcriptomics, several computational challenges of scRNA-seq extend to spatial transcriptomics, with the additional aspect of visualizing and correlating expression trends across spots to the tissue images in the latter. Once the resolution of the spots is at or below the size of plant cells, the method may generate scRNA-seq data that are spatially resolved. In this section, we outline some of the computational challenges that apply to both scRNA-seq and spatial transcriptomics, and highlight potential solutions.

Data processing

A major step in single-cell methods is data processing, which involves mapping reads back to cells and determining the nature of the cell based on those reads. In the Drop-seq method [16] for scRNA-seq, DNA-barcoded microbead particles or oligo probes consisting of millions of primers coded with a unique cellular barcode followed by a UMI (unique molecular identifier) are used to capture the mRNA molecules from dissociated single cells. For spatial transcriptomics, a microarray-based method [4] adapted this approach to capture mRNA molecules from 100 μm 'spots' (instead of dissociated single cells) gridded with 200 μm center-to-center spacing on an array that is fixed with tissue sections, by replacing the cellular barcode in the oligo primers with a spatial barcode that uniquely maps back to the X-Y locations of the spots (Figure 1). More recently, the Slide-seq method [20] achieved higher resolution (10 μm tissue regions) by using spatially barcoded microbeads packed across a slide (termed 'puck') instead of the microarray. The rationale in these spatial transcriptomics methods is that the mRNA molecules can be traced back to each gene in each spot (or spatial bead in case of Slide-seq) in the tissue by using (a) the spatial barcode to distinguish the location of different spots and (b) the UMI to differentiate molecules within the same spot. UMIs are helpful in removing sequencing biases during the PCR amplification step. Typically, for data generation, paired-end sequencing is done with the first read of the pair coinciding with the spatial barcode and UMI part of the primer, while the second read captures the cDNA of the gene transcript (Figure 2A).

After obtaining paired-end reads of the array-captured cDNA, one can estimate transcript abundances for each gene, comparing across spots in a specific region of the sectioned tissue. To determine these abundances, the workflow involves the following steps: (i) standard RNA-seq cleanup to remove sequencing adapters and low quality reads, (ii) deconvolution of the reads and grouping by spatial barcode, (iii) map the cDNA reads to reference gene transcripts and estimate gene expression by counting UMIs per spot per gene, (iv) normalize gene expression by total reads per spot, and finally, specific to spatial transcriptomics (v) map the spatial barcode and associated set of genes to a cell or location based on the microscopy image of the tissue section to assign it a 'real' physical location. The final output from steps (i–iv) is a gene expression table, with columns representing spatially oriented spots and rows representing genes (Figure 2A). This can be achieved using the Cellranger pipeline by 10x Genomics, with the exception of assuming a spatial barcode in place of a cell barcode, or it can be achieved by the ST pipeline [37], as previously used in plants [5]. However, for step (v), which involves visualization of these spots and gene expression on a tissue image, there are limited tools currently available, particularly to create a *Z*-stack of 3D abundances by overlaying the 2D slices and assigning X-Y-Z coordinates within the plant tissue. We discuss this issue in further detail below.

Dealing with inherent noise

The next step after obtaining the gene expression table is to cluster the data based on gene expression profiles across cells or spatial positions. This identifies subpopulations of cell types or regions with similar expression

© 2021 The Author(s). Published by Portland Press Limited on behalf of the Biochemical Society and the Royal Society of Biology

Emerging Topics in Life Sciences (2021)
https://doi.org/10.1042/ETLS20200274

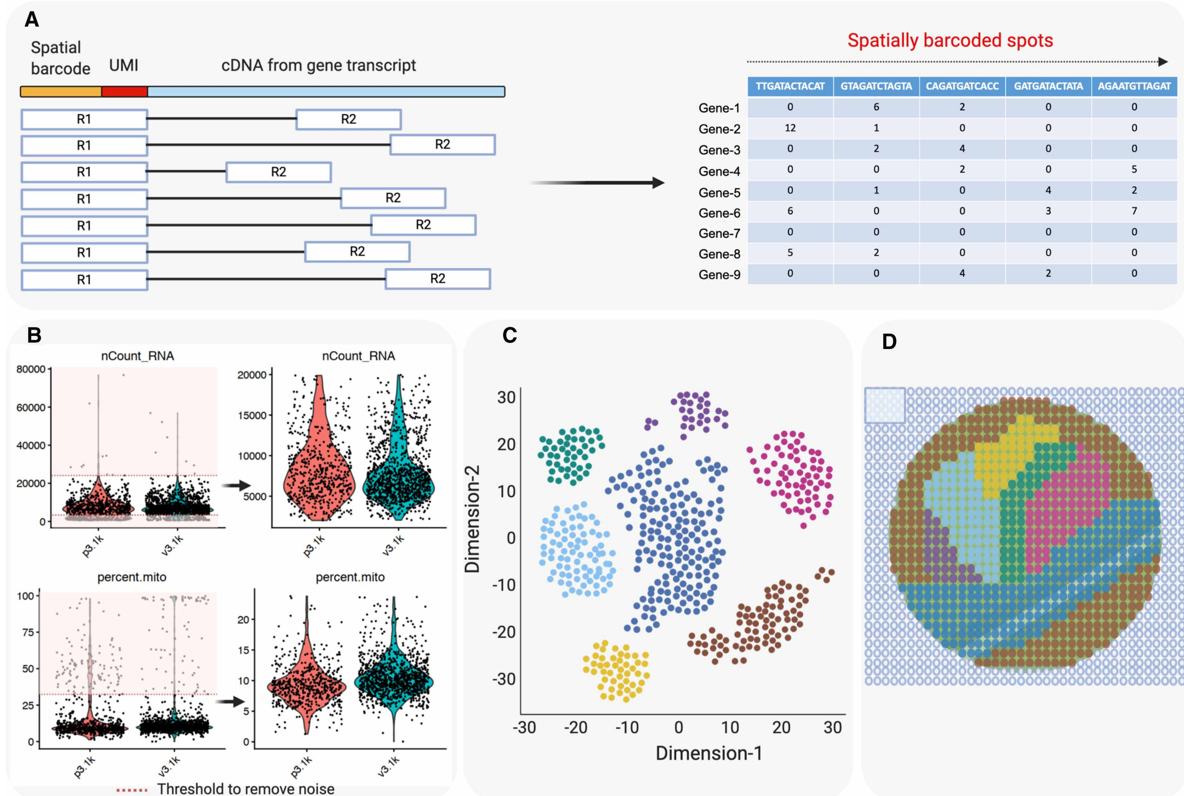

Figure 2. Summary of computational challenges.

(**A**) Data processing: involves deconvolution of paired-end reads by the spatial barcode, mapping the cDNA reads to reference gene transcripts, and finally creating a gene expression table with columns representing spatial spots and rows representing genes. (**B**) Dealing with inherent noise: Thresholds are established to remove spots with too low and too high RNA, given by nCount_RNA that reflects number of UMIs per spot (top panel) and spots with a very high percentage of mitochondrial reads, given by percent.mito (bottom panel). Violin plots showing distribution of spots (black dots) in two samples (different by color) before and after removing noisy data. (**C**) Dimensionality reduction and clustering: After removing noise, spots are clustered based on gene expression levels using dimensionality reduction and clustering techniques such as t-SNE or UMAP. (**D**) Visualization of spatial expression maps: Using the spatial barcodes the clustered spots are mapped back to the physical spots on the tissue. Not shown: *Z*-stacks of tissue sections would be reassembled to create 3D spatial maps of the transcriptome. Created with BioRender.com.

trends. However, there are challenges unique to scRNA-seq or spatial data that can seriously compromise clustering and downstream functional characterization. One such challenge is inherent noise. Noise in scRNA-seq or spatial data is both technical and biological. From a computational perspective, technical noise in spatial data results from 'spatial dropouts' (spots with low capture efficiency that increase at higher resolution), intrinsic variation between spots covering partial cells versus spots covering multiple cells, spots with highly expressing ribosomal genes or mitochondrial RNA, variation in sequencing depth, and experimental batch effects. Biological noise arises from transient cell states, cell-cycle variation, and cell heterogeneity. For accurate clustering of the cells, it is important to remove noise and reduce the feature set to the most informative genes [38]. One way to address this is to model the technical variation in the data and establish thresholds to remove cells or spots with either very low RNA or very high RNA, as well as those with a high percentage of mitochondrial/ribosomal genes (Figure 2B). Another way to reduce noise is to smooth the gene expression values of each spot by averaging these with those of its adjacent spots [39,40]. Expression of spike-in control molecules can be used for normalizing read counts and addressing variation in sequencing depth [41]. PCA-regression-like approaches can be used to remove confounding factors such as batch effects, and it has been shown that variability due to cell-cycle phase can be accounted for by using a Gaussian-based linear regression approach [42].

© 2021 The Author(s). Published by Portland Press Limited on behalf of the Biochemical Society and the Royal Society of Biology

R or python packages such as Seurat [43], Scanpy [44], and Scater [45] have been developed particularly to analyze large-scale scRNA-seq data and to provide tools to remove the inherent noise, and they can also be applied to high-resolution spatial transcriptomics data.

Dimensionality reduction and clustering

ScRNA-seq generates high-dimensional data due to the large number of genes assayed across individual cells. After removing technical noise and reducing the feature set to the most informative genes, computational strategies may be applied to reduce the data to a lower dimensional space, at which point unsupervised clustering (grouping cells based on gene expression profiles) is performed (Figure 2C). With complementary biological annotation gathered from cell atlases and cell ontologies, one can identify and characterize new cell types from the clustering results. Several clustering strategies have been well documented and are widely adopted to analyze scRNA-seq data [35,38] including in plant studies [46]. t-SNE [47] and UMAP [48] are the two most commonly used techniques for nonlinear dimensionality reduction and clustering of scRNA-seq data. In the case of high-resolution spatial RNA-seq data, one could similarly carry out unsupervised clustering of spatial features (as dissociated single cells) and later correlate clusters with spatial regions in the tissue. Some of these computational methods for dimensionality reduction and clustering are available in the Seurat v3 R package [49] with support for spatial transcriptomics data as well.

Visualization of spatial expression maps

The primary advantage of spatial transcriptomics is the ability to visualize gene expression dynamics within the 2D spatial contours of a section of plant tissue. Visualization in 3D space can be achieved by stacking multiple 2D tissue sections. This then entails the challenge of correlating gene expression data from array spots defined by X-Y-Z coordinates (derived from the spatial basis of the 2D approach and then overlaid to create a *Z*-stack of 3D abundances) with microscopy-based 3D array imaging of the plant tissue. To achieve this correlation, it is first important to establish virtual or physical 'fiducials' on the spatially barcoded arrays used to physically orient and align the array to the microscopy image. A simple example of a physical fiducial could be an empty area in the top left corner of the array [4]. Secondly, the correlation itself can either be done manually using imaging softwares such as Adobe Photoshop [5] or using ST Spot Detector, a computational tool developed specifically to analyze spatial transcriptomics imaging data [50]. ST Spot Detector can overlay gene expression data, defined by spatial coordinates, on a high-quality tissue image. Once the 'real' spatial context of gene

Table 1 Challenges, solutions and outcomes of spatial transcriptomics in plants

	Challenge	Possible solution	Outcome
Experimental	Current resolution in plants is limited at 100 μm	Increase the resolution of the microarray-based approach by adding smaller spots that are densely spaced	A platform that provides single-cell resolution in plants
	Nature of plant cells and tissue makes it difficult prep the tissue for sectioning	Apply a cryo-compatible embedding medium to encapsulate the plant structure	Better support and improved cutting properties; morphology of the tissue sections is maintained
	Components of cell wall limits the efficiency of RNA capture	Push the released mRNA down towards the probes through electrophoresis	Increased RNA capture efficiency
Computational	Biological/Technical noise when handling high resolution data	Model the technical variation in the data and establish thresholds to remove cells or spots with either very low/high RNA content	Reduced noise that allows for large scale, high resolution spatial data to be analyzed
	Correlating gene expression data from array spots with microscopy-based 3D imaging of the plant tissue	Establishing virtual/physical fiducials on arrays; manually correlate gene expression or using established analysis tools (i.e. ST Spot Detector)	Ability to analyze single-cell, spatial transcriptomics in 3D context

© 2021 The Author(s). Published by Portland Press Limited on behalf of the Biochemical Society and the Royal Society of Biology

expression across cells is resolved, one can proceed to visualize spatial expression maps of particular genes of interest, view the differentially expressed genes in a specific region and visually locate the cell types on the tissue that are clustered together (Figure 2D). Computational tools such as ST viewer [51] can also help with these visualizations. Trendsceek [52] will identify spatial expression trends of significant genes including hot-spots, step or linear gradients, and streaks; this is based on a statistical method using marked point processes, in which points represent the spatial location of cells and marks on each point represent expression levels. It is also worth noting that computational methods such as NovoSparc [53] attempt to reconstruct *de novo* spatial positioning from dissociated scRNA-seq data, based on assumptions that cells in proximity show higher similarity in expression than cells farther apart. However, most of the approaches discussed above are currently known to work with 2D spatial data and will need to be extended and improved to analyze spatial transcriptomics data in 3D space at a single-cell resolution.

Conclusion

The future is promising for advances in transcriptomics, particularly with the development of a new generation of methods that map gene expression on tissues and organs. These advances, collectively representing the method known as spatial transcriptomics, will allow the 3D reconstruction of potentially cellular-level variation in gene expression in multicellular organisms. There are a number of technical challenges yet to be overcome, including for both experimental and computational aspects of the approach (Table 1). Validation of the resulting data will require methods of equally high spatial resolution; the most promising are *in situ* hybridization techniques for RNA, as these are now optimized for plants [18]. *In situ* methods have both sequence specificity and, using super-resolution microscopy, the power to localize and quantify RNAs within individual cells [18]. Therefore, in our opinion, spatial analysis of gene expression, at a single-cell resolution and reconstructing 3D organs or tissues, is the final frontier in RNA-seq analysis — and one that we are on the cusp of achieving.

Summary

- Spatial transcriptomics is an emerging technology that allows for gene expression data to be spatially resolved in plant systems.
- Advancing spatial transcriptomics in plants to a single-cell resolution platform will require overcoming several obstacles within wet lab techniques and computational analysis.
- Implementing improvements in the methodology for spatial transcriptomics in plants will help achieve the ultimate goal of having plant systems with complete, spatially resolved gene expression at single cell resolution.

Competing Interests

The authors declare that there are no competing interests associated with the manuscript.

Funding

This work was supported by a Howard Hughes Medical Institute Hanna H. Gray Fellowship to K.L.C., and NSF MCB award 1945854 to B.C.M. and K.J.C. Additional support for the work comes from the Donald Danforth Plant Science Center, including the Advanced Bioimaging Laboratory (RRID:SCR_018951).

Author Contributions

S.G.R.G. and K.L.C. made the outline, drafted the manuscript and figures, and all authors revised and finalized the manuscript.

Acknowledgements

We would like to thank Joanna Friesner for critical reading of the manuscript.

© 2021 The Author(s). Published by Portland Press Limited on behalf of the Biochemical Society and the Royal Society of Biology

Abbreviations

FACS, fluorescence-activated cell sorting; FISH, fluorescent *in situ* hybridization; INTACT, isolation of nuclei tagged in specific cell types; LCM, laser capture microdissection; UMI, unique molecular identifier.

References

1 Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L. and Wold, B. (2008) Mapping and quantifying mammalian transcriptomes by RNA-seq. *Nat. Methods* **5**, 621–628 https://doi.org/10.1038/nmeth.1226
2 Windram, O. and Denby, K.J. (2015) Modelling signaling networks underlying plant defence. *Curr. Opin. Plant Biol.* **27**, 165–171 https://doi.org/10.1016/j.pbi.2015.07.007
3 Li, Y., Pearl, S.A. and Jackson, S.A. (2015) Gene networks in plant biology: approaches in reconstruction and analysis. *Trends Plant Sci.* **20**, 664–675 https://doi.org/10.1016/j.tplants.2015.06.013
4 Ståhl, P.L., Salmén, F., Vickovic, S., Lundmark, A., Navarro, J.F., Magnusson, J. et al. (2016) Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* **353**, 78–82 https://doi.org/10.1126/science.aaf2403
5 Giacomello, S., Salmén, F., Terebieniec, B.K., Vickovic, S., Navarro, J.F., Alexeyenko, A. et al. (2017) Spatially resolved transcriptome profiling in model plant species. *Nat. Plants* **3**, 17061 https://doi.org/10.1038/nplants.2017.61
6 Shulse, C.N., Cole, B.J., Ciobanu, D., Lin, J., Yoshinaga, Y., Gouran, M. et al. (2019) High-throughput single-cell transcriptome profiling of plant cell types. *Cell Rep.* **27**, 2241–2247.e4 https://doi.org/10.1016/j.celrep.2019.04.054
7 Zhang, T.-Q., Xu, Z.-G., Shang, G.-D. and Wang, J.-W. (2019) A single-cell RNA sequencing profiles the developmental landscape of *Arabidopsis* root. *Mol. Plant* **12**, 648–660 https://doi.org/10.1016/j.molp.2019.04.004
8 Jean-Baptiste, K., McFaline-Figueroa, J.L., Alexandre, C.M., Dorrity, M.W., Saunders, L., Bubb, K.L. et al. (2019) Dynamics of gene expression in single root cells of *Arabidopsis thaliana*. *Plant Cell* **31**, 993–1011 https://doi.org/10.1105/tpc.18.00785
9 Ryu, K.H., Huang, L., Kang, H.M. and Schiefelbein, J. (2019) Single-cell RNA sequencing resolves molecular relationships among individual plant cells. *Plant Physiol.* **179**, 1444–1456 https://doi.org/10.1104/pp.18.01482
10 Birnbaum, K., Shasha, D.E., Wang, J.Y., Jung, J.W., Lambert, G.M., Galbraith, D.W. et al. (2003) A gene expression map of the Arabidopsis root. *Science* **302**, 1956–1960 https://doi.org/10.1126/science.1090022
11 Brady, S.M., Orlando, D.A., Lee, J.-Y., Wang, J.Y., Koch, J., Dinneny, J.R. et al. (2007) A high-resolution root spatiotemporal map reveals dominant expression patterns. *Science* **318**, 801–806 https://doi.org/10.1126/science.1146265
12 Yadav, R.K., Tavakkoli, M., Xie, M., Girke, T. and Reddy, G.V. (2014) A high-resolution gene expression map of the Arabidopsis shoot meristem stem cell niche. *Development* **141**, 2735–2744 https://doi.org/10.1242/dev.106104
13 Deal, R.B. and Henikoff, S. (2010) The INTACT method for cell type–specific gene expression and chromatin profiling in *Arabidopsis thaliana*. *Nat. Protoc.* **6**, 56 https://doi.org/10.1038/nprot.2010.175
14 Nelson, T., Tausta, S.L., Gandotra, N. and Liu, T. (2006) Laser microdissection of plant tissue: what you see is what you get. *Annu. Rev. Plant Biol.* **57**, 181–201 https://doi.org/10.1146/annurev.arplant.56.032604.144138
15 Gautam, V., Singh, A., Singh, S. and Sarkar, A.K. (2016) An efficient LCM-based method for tissue specific expression analysis of genes and miRNAs. *Sci. Rep.* **6**, 21577 https://doi.org/10.1038/srep21577
16 Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M. et al. (2015) Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 https://doi.org/10.1016/j.cell.2015.05.002
17 Tirichine, L., Andrey, P., Biot, E., Maurin, Y. and Gaudin, V. (2009) 3D fluorescent in situ hybridization using arabidopsis leaf cryosections and isolated nuclei. *Plant Methods* **5**, 11 https://doi.org/10.1186/1746-4811-5-11
18 Huang, K., Batish, M., Teng, C., Harkess, A., Meyers, B.C. and Caplan, J.L. (2020) Quantitative fluorescence in situ hybridization detection of plant mRNAs with single-molecule resolution. *Methods Mol. Biol.* **2166**, 23–33 https://doi.org/10.1007/978-1-0716-0712-1_2
19 Salmén, F., Ståhl, P.L., Mollbrink, A., Navarro, J.F., Vickovic, S., Frisén, J. et al. (2018) Barcoded solid-phase RNA capture for spatial transcriptomics profiling in mammalian tissue sections. *Nat. Protoc.* **13**, 2501–2534 https://doi.org/10.1038/s41596-018-0045-2
20 Rodriques, S.G., Stickels, R.R., Goeva, A., Martin, C.A., Murray, E., Vanderburg, C.R. et al. (2019) Slide-seq: a scalable technology for measuring genome-wide expression at high spatial resolution. *Science* **363**, 1463–1467 https://doi.org/10.1126/science.aaw1219
21 Vickovic, S., Eraslan, G., Salmén, F., Klughammer, J., Stenbeck, L., Schapiro, D. et al. (2019) High-definition spatial transcriptomics for in situ tissue profiling. *Nat. Methods* **16**, 987–990 https://doi.org/10.1038/s41592-019-0548-y
22 Rao, N., Clark, S. and Habern, O. (2020) Bridging genomics and tissue pathology: 10x genomics explores new frontiers with the visium spatial gene expression solution. *Genet. Eng. Biotechnol. News* **40**, 50–51 https://doi.org/10.1089/gen.40.02.16
23 Giacomello, S. and Lundeberg, J. (2018) Preparation of plant tissue to enable spatial transcriptomics profiling using barcoded microarrays. *Nat. Protoc.* **13**, 2425–2446 https://doi.org/10.1038/s41596-018-0046-1
24 Cosgrove, D.J. (2005) Growth of the plant cell wall. *Nat. Rev. Mol. Cell. Biol.* **6**, 850–861 https://doi.org/10.1038/nrm1746
25 Wang, C.-W., Budiman Gosno, E. and Li, Y.-S. (2015) Fully automatic and robust 3D registration of serial-section microscopic images. *Sci. Rep.* **5**, 15051 https://doi.org/10.1038/srep15051
26 Bourgaud, F., Gravot, A., Milesi, S. and Gontier, E. (2001) Production of plant secondary metabolites: a historical perspective. *Plant Sci.* **161**, 839–851 https://doi.org/10.1016/S0168-9452(01)00490-3
27 Taylor, R., Inamine, G. and Anderson, J.D. (1993) Tissue printing as a tool for observing immunological and protein profiles in young and mature celery petioles. *Plant Physiol.* **102**, 1027–1031 https://doi.org/10.1104/pp.102.3.1027
28 Love, M.I., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 https://doi.org/10.1186/s13059-014-0550-8
29 Anders, S. and Huber, W. (2010) Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 https://doi.org/10.1186/gb-2010-11-10-r106

© 2021 The Author(s). Published by Portland Press Limited on behalf of the Biochemical Society and the Royal Society of Biology

30 Robinson, M.D., McCarthy, D.J. and Smyth, G.K. (2010) Edger: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 https://doi.org/10.1093/bioinformatics/btp616
31 Zhou, Y.-H., Xia, K. and Wright, F.A. (2011) A powerful and flexible approach to the analysis of RNA sequence count data. *Bioinformatics* **27**, 2672–2678 https://doi.org/10.1093/bioinformatics/btr449
32 Hardcastle, T.J. and Kelly, K.A. (2010) Bayseq: empirical Bayesian methods for identifying differential expression in sequence count data. *BMC Bioinformatics* **11**, 422 https://doi.org/10.1186/1471-2105-11-422
33 Tong, L., Wu, P.-Y., Phan, J.H., Hassazadeh, H.R., SEQC Consortium, Tong, W. et al. (2020) Impact of RNA-seq data analysis algorithms on gene expression estimation and downstream prediction. *Sci. Rep.* **10**, 17925 https://doi.org/10.1038/s41598-020-74567-y
34 Garber, M., Grabherr, M.G., Guttman, M. and Trapnell, C. (2011) Computational methods for transcriptome annotation and quantification using RNA-seq. *Nat. Methods* **8**, 469–477 https://doi.org/10.1038/nmeth.1613
35 Stegle, O., Teichmann, S.A. and Marioni, J.C. (2015) Computational and analytical challenges in single-cell transcriptomics. *Nat. Rev. Genet.* **16**, 133–145 https://doi.org/10.1038/nrg3833
36 Faridani, O.R. and Sandberg, R. (2015) Putting cells in their place. *Nat. Biotechnol.* **33**, 490–491 https://doi.org/10.1038/nbt.3219
37 Navarro, J.F., Sjöstrand, J., Salmén, F., Lundeberg, J. and Ståhl, P.L. (2017) ST pipeline: an automated pipeline for spatial mapping of unique transcripts. *Bioinformatics* **33**, 2591–2593 https://doi.org/10.1093/bioinformatics/btx211
38 Kiselev, V.Y., Andrews, T.S. and Hemberg, M. (2019) Publisher correction: challenges in unsupervised clustering of single-cell RNA-seq data. *Nat. Rev. Genet.* **20**, 310 https://doi.org/10.1038/s41576-019-0095-5
39 He, B., Bergenstråhle, L., Stenbeck, L., Abid, A., Andersson, A., Borg, Å. et al. (2020) Integrating spatial gene expression and breast tumour morphology via deep learning. *Nat. Biomed. Eng.* **4**, 827–834 https://doi.org/10.1038/s41551-020-0578-x
40 Zhao, E., Stone, M.R., Ren, X., Pulliam, T., Nghiem, P., Bielas, J.H. et al.) Bayesspace enables the robust characterization of spatial gene expression architecture in tissue sections at increased resolution. *bioRxiv* https://doi.org/10.1101/2020.09.04.283812
41 Brennecke, P., Anders, S., Kim, J.K., Kołodziejczyk, A.A., Zhang, X., Proserpio, V. et al. (2013) Accounting for technical noise in single-cell RNA-seq experiments. *Nat. Methods* **10**, 1093–1095 https://doi.org/10.1038/nmeth.2645
42 Buettner, F., Natarajan, K.N., Casale, F.P., Proserpio, V., Scialdone, A., Theis, F.J. et al. (2015) Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. *Nat. Biotechnol.* **33**, 155–160 https://doi.org/10.1038/nbt.3102
43 Butler, A., Hoffman, P., Smibert, P., Papalexi, E. and Satija, R. (2018) Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 https://doi.org/10.1038/nbt.4096
44 Wolf, F.A., Angerer, P. and Theis, F.J. (2018) SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 https://doi.org/10.1186/s13059-017-1382-0
45 McCarthy, D.J., Campbell, K.R., Lun, A.T.L. and Wills, Q.F. (2017) Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. *Bioinformatics* **33**, 1179–1186 https://doi.org/10.1093/bioinformatics/btw777
46 Rich-Griffin, C., Stechemesser, A., Finch, J., Lucas, E., Ott, S. and Schäfer, P. (2020) Single-cell transcriptomics: a high-resolution avenue for plant functional genomics. *Trends Plant Sci.* **25**, 186–197 https://doi.org/10.1016/j.tplants.2019.10.008
47 van der Maaten, L. and Hinton, G. (2008) Visualizing data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605
48 McInnes, L., Healy, J. and Melville, J. (2018) UMAP: Uniform manifold approximation and projection for dimension reduction. *arXiv* https://doi.org/1802.03426
49 Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck WM, I.I.I. et al. (2019) Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902.e21 https://doi.org/10.1016/j.cell.2019.05.031
50 Wong, K., Navarro, J.F., Bergenstråhle, L., Ståhl, P.L. and Lundeberg, J. (2018) ST spot detector: a web-based application for automatic spot and tissue detection for spatial transcriptomics image datasets. *Bioinformatics* **34**, 1966–1968 https://doi.org/10.1093/bioinformatics/bty030
51 Fernández Navarro, J., Lundeberg, J. and Ståhl, P.L. (2019) ST viewer: a tool for analysis and visualization of spatial transcriptomics datasets. *Bioinformatics* **35**, 1058–1060 https://doi.org/10.1093/bioinformatics/bty714
52 Edsgärd, D., Johnsson, P. and Sandberg, R. (2018) Identification of spatial expression trends in single-cell gene expression data. *Nat. Methods* **15**, 339–342 https://doi.org/10.1038/nmeth.4634
53 Nitzan, M., Karaiskos, N., Friedman, N. and Rajewsky, N. (2019) Gene expression cartography. *Nature* **576**, 132–137 https://doi.org/10.1038/s41586-019-1773-3

© 2021 The Author(s). Published by Portland Press Limited on behalf of the Biochemical Society and the Royal Society of Biology