Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Biotechnology

Beyond bulk: a review of single cell transcriptomics methodologies and applications

Ashwinikumar Kulkarni, Ashley G Anderson, Devin P Merullo and Genevieve Konopka

Single-cell RNA sequencing (scRNA-seq) is a promising approach to study the transcriptomes of individual cells in the brain and the central nervous system (CNS). This technology acts as a bridge between neuroscience, computational biology, and systems biology, enabling an unbiased and novel understanding of the cellular composition of the brain and CNS. Gene expression at the single cell resolution is often noisy, sparse, and high-dimensional, creating challenges for computational analysis of such data. In this review, we overview fundamental sample preparation and data analysis processes of scRNA-seq and provide a comparative perspective for analyzing and visualizing these data.

Address
Department of Neuroscience, UT Southwestern Medical Center, Dallas, TX 75390, USA

Corresponding author:
Konopka, Genevieve (genevieve.konopka@utsouthwestern.edu)

Current Opinion in Biotechnology 2019, **58**:129–136

This review comes from a themed issue on **Systems biology**

Edited by **Maria Klapa** and **Ioannis P Androulakis**

For a complete overview see the Issue and the Editorial

Available online 10th April 2019

https://doi.org/10.1016/j.copbio.2019.03.001

Introduction

High-throughput single-cell transcriptomics has provided unprecedented insights into cellular diversity of tissues across diverse organisms. Transcriptomic studies using bulk tissue assume and represent all the cells as a homogeneous material, thereby ignoring the stochasticity of gene expression [1–3]. Single-cell transcriptomics, however, is able to address questions related to the stochastic nature of gene expression.

In this review, we will focus on a number of technical aspects of single-cell transcriptomic profiling, such as various methods of single-cell isolation and library preparation, and the incorporation of spatial information with single-cell gene expression profiling. We will also highlight commonalities and differences between single-cell and single-nuclei RNA-seq. Overall, we will highlight some of the outstanding questions in the field and provide some useful resources for the design and implementation of single-cell or single-nuclei RNA-seq studies.

Single-cell isolation and library preparation

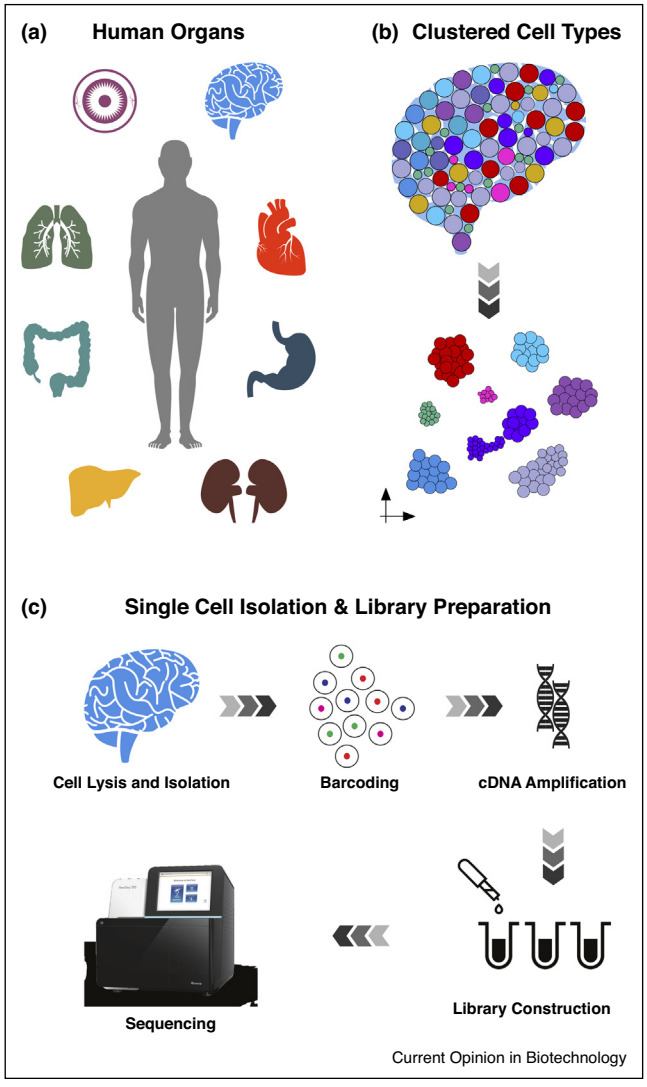
Methods to isolate single cells for RNA sequencing vary in the number of cells they isolate (high-throughput or low-throughput) and how they select cells (biased or unbiased) [4,5]. Droplet-based technologies are currently commonly used as a high-throughput, unbiased solution [6]. The three most widely employed platforms are 10X Genomics Chromium, DropSeq, and inDrop [7••,8•,9,10]. Each of these approaches uses microfluidics to tag individual cells with single beads containing a unique barcode. Each mRNA transcript is also linked with a unique molecular identifier (UMI). This approach results in a matrix containing the absolute number of counts for each transcript in each cell. Each platform has its strength and weaknesses: 10X Genomics Chromium is the most sensitive for detecting the greatest number of transcripts, although it can be cost-prohibitive for large numbers of samples; DropSeq generally is more affordable, but with the tradeoff of slight sensitivity losses by counting fewer transcripts; inDrop may be ideal to detect genes expressed at low levels because of its customizable parameters [7••].

Two distinct approaches for generating single-cell sequencing libraries are full-length and tag-based [11–13]. Full-length methods provide reads that are derived from across the entire length of genes, improving overall sensitivity. However, full-length library preparations tend to show a bias for longer genes because counts are often missed for shorter genes [14]. Tag-based methods incorporate UMIs, allowing for the identification and quantification of individual transcripts. These tags are added to either the 3′ or 5′ end of transcripts and therefore introduce a read bias to the end that receives the tag, although there is no length bias [14]. While full-length methods achieve more comprehensive read coverage, tag-based methods have become dominant for quantification purposes, especially for larger numbers of cells [11–13]. Figure 1 summarizes a generalized single-cell transcriptomic experiment.

Single-cell versus single nuclei

CNS cell types can be morphologically complex, making the isolation of intact cells challenging [15,16,17••,18].

Figure 1



Schematic diagram of a single-cell transcriptomic experiment. **(a)** Single-cell gene expression profiling can be carried out from any tissue from humans or model systems. **(b)** The human brain, as an example, is composed of diverse types of cells, each of which contains gene expression patterns that can be resolved into clusters of cells of similar types. **(c)** A generalized scRNA-seq pipeline from source tissue to sequencing.

Thus, unbiased neuronal profiling requires the preparation of RNA-seq libraries from freshly harvested, viable, and intact neurons. Most of the single-cell isolation or preparation methods applied to brain tissue make use of harsh enzymatic dissociation steps which can harm neurons and impair neuronal profiling [17••,19,20]. Several studies have indicated that some cell types are underrepresented or overrepresented in the final datasets due to artificial filtering from dissociation steps [21,22].

Similar issues arise for other tissues in the body where some cell types are more susceptible to the tissue dissociation process, such as the heart [23]. Additionally, isolating intact cells from post-mortem frozen tissue is challenging because cell membranes are easily damaged from mechanical and physical stress during the freeze–thaw process [23,24]. The isolation of nuclei offers an advantage over cells because nuclei are more resistant to mechanical and physical stress during freeze–thaw. As an alternative approach to scRNA-seq, single-nuclei RNA sequencing (snRNA-seq) has been successfully implemented to study cellular composition in post-mortem human brain [16,18,19,25], mouse hippocampus [26], and other tissues. Several studies have compared the matched single-cells versus single-nuclei, for example, in mouse 3T3 cells [19] and in mouse somatosensory cortex [17••,20]. snRNA-seq data have a high degree of similarity with whole-cell transcriptomes in terms of the number of detected genes and resolved cell types [17••,20], although snRNA-seq datasets showed a higher proportion of intronic reads than scRNA-seq, and the detection of more short-length genes [17••,20]. An important aspect of snRNA-seq is ensuring that reads that map both to either exons or introns are included, as counting intronic reads is required for high resolution cell-type identification from snRNA-seq data [20,27].

Some methods also attempt to cause less damage to cells during dissociation. These approaches are particularly relevant to the expression of immediate early genes, for which transcription changes quickly in response to stimuli. For example, Act-seq minimizes cellular disturbances during isolation and preparation by using the application of actinomycin D during dissociation to inhibit rapid transcriptional changes [28].

Performance evaluation of a single-cell RNA-seq experiment

The technical performance of scRNA-seq data should be evaluated for each experiment, including sensitivity, accuracy, and precision [27,29•,30]. *Sensitivity* refers to the lowest number of transcripts that can be reproducibly detected. This measurement is highly dependent on the depth of sequencing. It is crucial to understand this aspect of the dataset, as it directly estimates the proportion of RNA being measured in a single cell. *Accuracy* refers to the degree of agreement between the expression measurements and the actual transcript abundance [27,29•]. Accuracy can be affected by the factors specific to the protocol being used, such as exponential PCR amplification or sequencing bias [29•]. *Precision* is considered inversely proportional to the technical noise in the RNA-seq measurements. Technical noise is often considered as a coefficient of variation across replicate measurements. Precision currently remains the largest limitation for current scRNA-seq and snRNA-seq

protocols because the generated data are generally sparse and may contain many zeroes [27,29•].

Data preprocessing

Raw sequencing data, usually in the form of binary base calls (BCL files), are first de-multiplexed using oligonucleotide index sequences (such as chromium i7 index sets used during 10X Genomics library preparation) to generate FASTQ files. FASTQ files then undergo quality control (QC) to identify and remove low quality bases and sequencing adapters from the reads. Following filtering, reads are aligned to a reference genome. Uniquely mapped reads are then assigned to genes using reference annotation followed by counting the number of reads per gene per cell. This count matrix is then used to distinguish real cells from doublets and empty beads. The filtered count matrix ultimately enables downstream processing such as normalization, clustering, cell-type identification, and pseudo-time analysis.

Considering the limitations of using Cell Ranger pipelines, an alternative strategy that is more customizable is to use a combination of tools such as FastQC [FastQC, Babraham Bioinformatics, URL: https://www.bioinformatics.babraham.ac.uk/projects/fastqc], UMI tools [31], STAR [32], featureCounts [33], HTSeq [34], Samtools [35], zUMIs [36], and others. There are several customized pipelines [Analysis of single cell RNA-seq data, URL: http://hemberg-lab.github.io/scRNA.seq.course; Custom pipeline to preprocess the raw sequencing data for a single cell/nuclei transcriptomic profiling, URL: https://github.com/konopkalab/Pipeline-SingleCellRNASeq] available publicly (Figure 2) which take advantage of the above-mentioned tools to pre-process the data.

Clustering and visualization of cellular clusters

Interpreting single-cell transcriptomic data requires a few additional steps after the raw count matrix is ready, which are discussed in publicly available user tutorials for R-packages and python-based tools [10,37–40]. It remains fundamentally unclear how to ideally cluster and visualize heterogeneous cellular population into groups of homogeneous so-called "cell types."

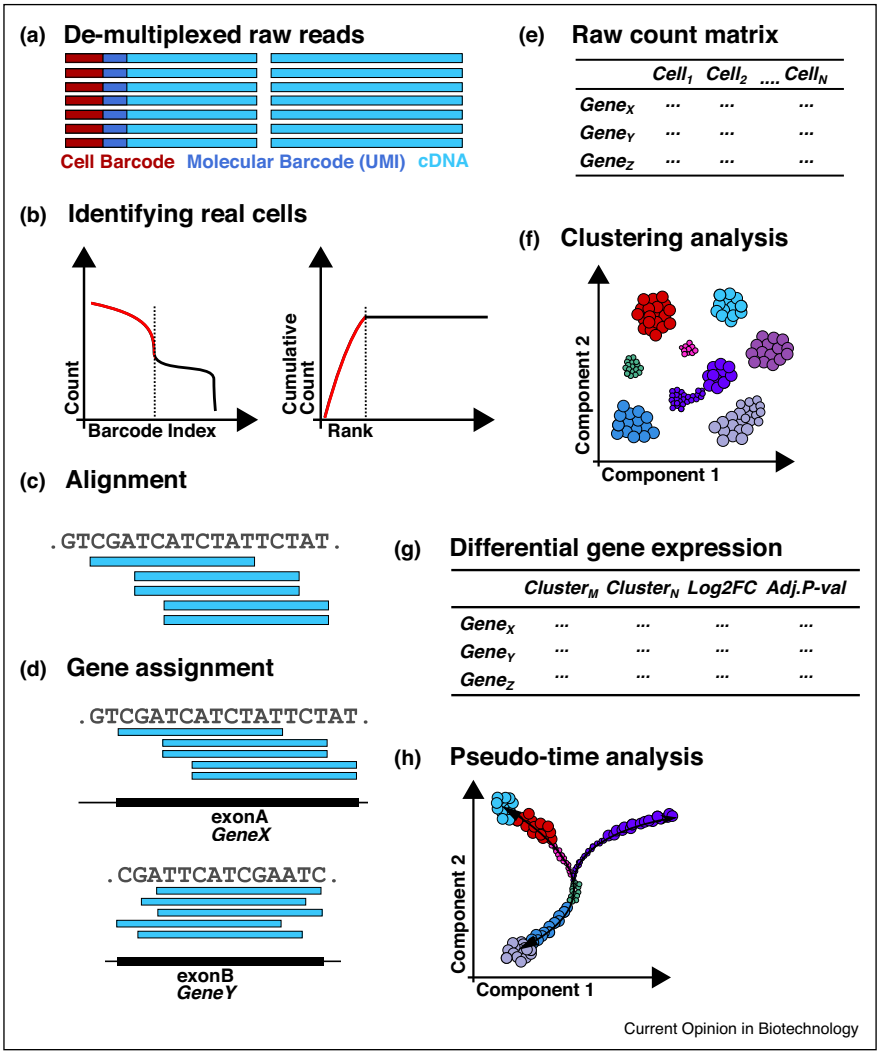
Dimensionality reduction is used in scRNA-seq data processing to allow for downstream analyses such as clustering, visualization, and further interpretation. Common algorithms for this include principle component analysis (PCA) [41], multi-dimensional scaling (MDS) [42], Sammon mapping [43], t-distributed stochastic neighbor embedding (t-SNE) [44–47], and Isomap [48]. Linear dimensionality reduction methods such as PCA [41] often cannot represent high-dimensional data into low-dimensional data. Among many non-linear methods, t-SNE [44–47] is a very popular choice for cluster visualization. Though most of the single-cell RNA-seq studies published to date rely on t-SNE [44–47] for visualization, several limitations have now become apparent, for example, when compared to a more recent method of visualization [49,50,51•]. Uniform manifold approximation and projection (UMAP) is a novel algorithm for non-linear dimensionality reduction based on Riemannian geometry and algebraic topology [49]. UMAP offers promising advantages over t-SNE for visualization of single-cell RNA-seq data [51•]. Furthermore, UMAP also retains and projects the information related to multi-branched trajectories to facilitate pseudo-time analysis [40,51•,52••]. In our unpublished dataset from mouse striatal neurons, UMAP not only facilitated annotating the cells more accurately, but also helped visualize the clusters in accordance with their developmental trajectories (Figure 3).

Incorporating spatial information

A major drawback to scRNA-seq methods is the loss of spatial information [53•,54–56]. The first attempts to identify the spatial locations of genes identified via scRNA-seq that were co-expressed in clusters involved visually labeling individual gene markers through fluorescence in situ hybridization (FISH). In single-molecule fish (smFISH), multiple fluorescent probes are used to characterize distinct cell groups based on the resulting combinations [57]. However, this technique is limited by the number of colors that can be visualized at once. Sequential fluorescence *in situ* hybridization (seqFISH) resolves this problem by sequentially labeling tissue for different RNA markers [58–60]. Fluorescence is washed out after each round of labeling, thereby binarily labelling cell populations for the presence or absence of a particular gene. A unique barcode is then assigned to each cell population. Such combinatorial barcodes generated by sequential labeling enable cell populations to be identified with much greater efficiency and robustness than smFISH. The development of multiplexed error-robust FISH (MERFISH) allows even more RNA species to be labeled with a reduced error rate [61••,62–64].

Several other methods to observe spatial gene expression also exist [65]. Transcriptomic scRNA-seq data can be computationally integrated with spatial seqFISH information [66]. Expansion microscopy, which physically increases the space between molecules while preserving spatial integrity, has been used in conjunction with the fluorescent labelling of RNA molecules to visualize gene expression [67,68]. osmFISH is a recently developed method for labelling gene expression without the use of barcodes, and a chief advantage of it is that the detection of highly expressed genes does not obscure the detection of lowly expressed genes [69]. Together, these efforts to preserve the spatial locations of specific cell types will offer invaluable insight into the organization and functioning of diverse tissues.

Figure 2

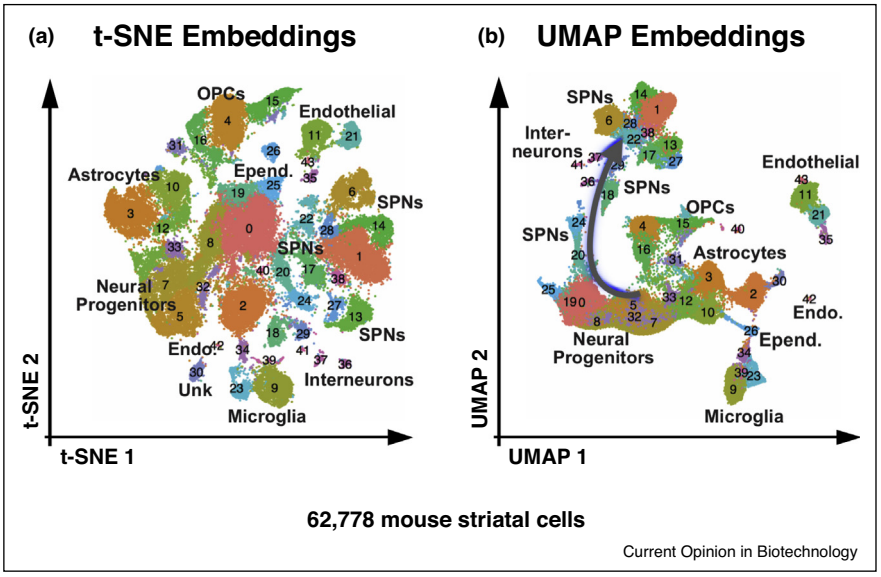


General workflow of scRNA-seq data analysis.
(a) Raw BCL files containing reads are quality-filtered and de-multiplexed. These reads contain information for the cell barcode, a molecular barcode or UMI, and the cDNA sequence. **(b)** Using UMI tools [31], one can distinguish real cells from potential background noise. UMI tools estimate these cell barcodes from the data using the knee method. This whitelist of cell barcodes is then used to extract cDNA reads corresponding to the estimated real cells. **(c)** The next step is to align the reads to the reference genome, for example using STAR, [32] and **(d)** assign reads to the reference annotation gene model using a tool such as featureCounts [33]. **(e)** UMI tools can further collapse the read counts corresponding to each gene in each cell, creating a raw digital expression count matrix. **(f)** Using the raw count table, clustering pipelines such as Seurat [37,38] separate the cells into different clusters based on the cell types and **(g)** facilitate differential gene expression analysis between any given clusters. One can additionally use tools like SCDE [78,79] which are designed specifically for single-cell transcriptional data to calculate differential expression. **(h)** Additional analyses could include a pseudo-time trajectory of cells using tools such as Velocyto [52••] and Monocle [80].

Application of single-cell transcriptomics in neuroscience

Single-cell transcriptomics has uncovered novel cell-types and elucidated neuronal and glial diversity within various brain regions and across species [70–73]. This technology has also enabled the molecular examination of "cell states" as well as cell types [71,74]. Novel methodologies to implement single-cell transcriptomics on both fixed and frozen postmortem human brain tissue using nuclei are providing insight into human brain evolution, development, and function [19]. The application of single-cell transcriptomics within experimental systems modelling human brain development or disease states, such as human brain organoids [75] or animal

Figure 3



Comparison of non-linear dimensionality reduction techniques.
Using the same dataset of 62,778 mouse striatal neurons (unpublished dataset), plotting the cells using t-SNE embeddings **(a)** does not spatially capture the trajectory of distinct cell types like **(b)** UMAP. Arrow shows the progression of neural progenitor clusters to fully mature spiny projection neuron clusters that is visible in UMAP; however, no such trajectory can be visually discerned using t-SNE.

models [76••], can ultimately help illuminate conserved and unique aspects of human brain development and function.

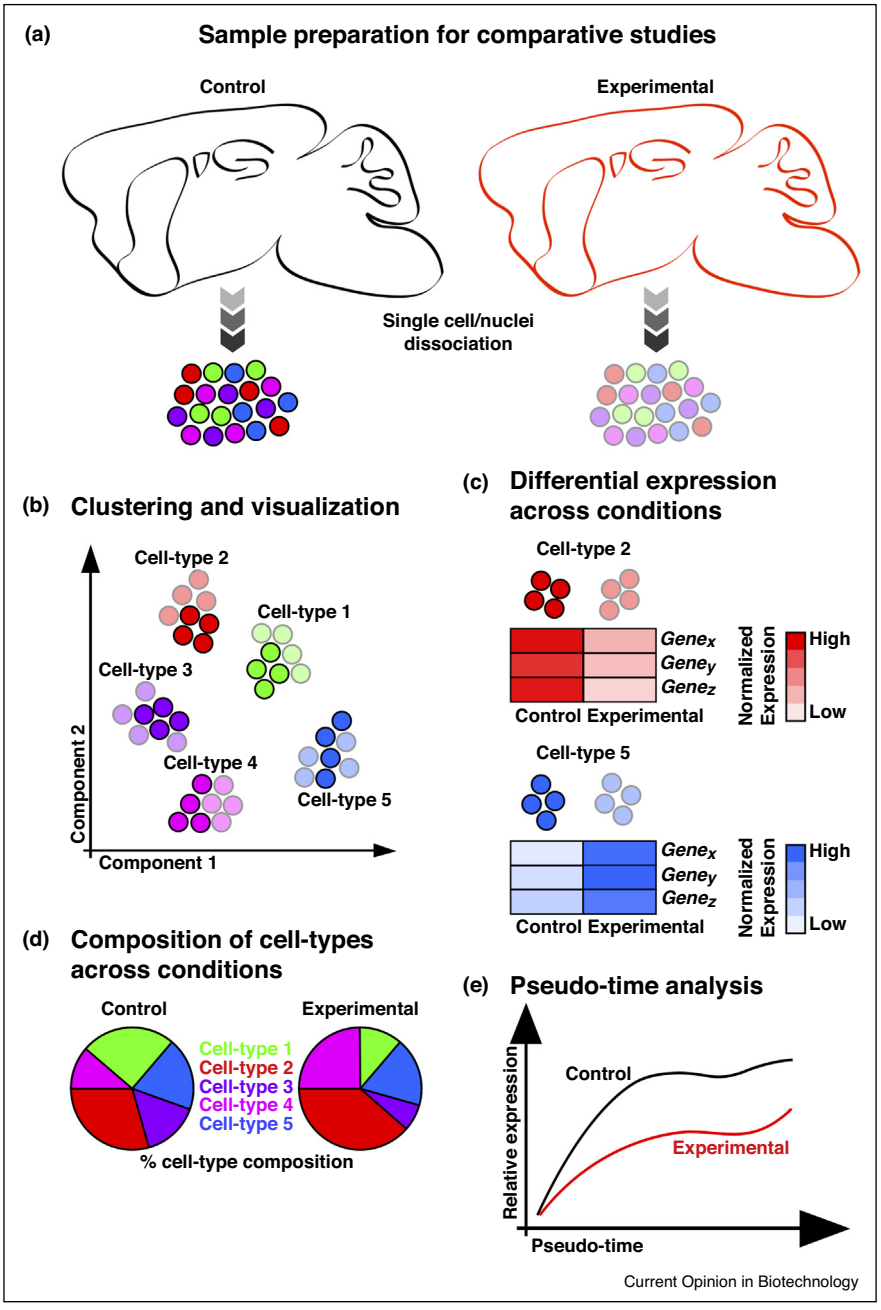
Using scRNA-seq to study neurological disorders has been a major technical advancement given the enormous diversity of cell types within the brain. Single-cell transcriptomics has helped to characterize Rett syndrome, an X-linked neurodevelopmental disorder, where random X-chromosome inactivation leads to a mixture of mutant and normal cells within the same patient [76••]. A new methodology that identifies single-cell single-nucleotide polymorphism sequencing (SNP-seq) could reliably detect both mutant and normal cells in postmortem human brain tissue [76••]. This single-cell transcriptomic tool could be broadly applied to other X-linked neurodevelopmental disorders, such as Fragile-X syndrome or X-linked intellectual disability.

Studies using scRNA-seq in control vs. knockout studies in mice have analyzed changes in cellular composition, differentially expressed genes, and pseudo-time or pseudo-differentiation tools at cellular resolution with deletion or manipulation of a given gene [77••] (Figure 4). Given the immense cellular heterogeneity of any given brain region, single-cell transcriptomics allows for the identification of cell-types particularly vulnerable in disease states and for addressing the question of why these cell types are particularly vulnerable. Another benefit of using single-cell transcriptomics within these comparative approaches is the ability to examine both cell-autonomous and non-cell-autonomous changes. For example, in studies using models to delete a specific gene in distinct cell types, downstream analyses could examine how disruption of one cell type alters the function or number of other cell types.

Conclusions and future challenges

Currently, annotating clustered cells using a handful of known marker genes or using enrichment for other published dataset is a common approach, but this poses a real issue to identify novel cell types hidden in the scRNA-seq datasets. Machine learning approaches combined with literature-based knowledge can potentially resolve this problem. Many collaborative approaches across the globe, such as the Human Cell Atlas [https://www.humancellatlas.org], are creating an enormous catalog to account for every single type of cell in human body. As these data are rapidly increasing, the present challenge is a careful curation, analysis, and annotation of these datasets. Efforts are also needed to overcome the limitations due to the sparse and noisy nature of the dataset, low starting material, biases introduced due to amplification and sequencing, technical variations, and cost. Nevertheless, single cell transcriptomics technology has great

Figure 4



Example schematic of a comparative single-cell transcriptomic study.
(a)–(b) Control versus experimental cells can be clustered together and visualized using a dimensionality reduction tool. **(c)–(d)** Gene expression differences within distinct cell types between control and experimental cells can be uncovered and cell-type composition within the tissue can be compared across samples. **(e)** Other scRNA-seq tools, such as pseudo-time analyses, can be examined between control and experimental systems.

potential to contribute to the fields of spatial transcriptomics and personalized medicine.

Conflict of interest statement

Nothing declared.

Funding

G.K. is a Jon Heighten Scholar in Autism Research at UT Southwestern. This work was supported by the N.I.H. (T32MH076690 to A.A.; DC014702, DC016340, MH102603, MH103517, and NS106447 to G.K.), the James S. McDonnell Foundation 21st Century Science

Initiative in Understanding Human Cognition (Scholar Award 220020467), the Simons Foundation (SFARI awards 401220 and 573689), and the Chan Zuckerberg Initiative, an advised fund of Silicon Valley Community Foundation (HCA-A-1704-10747) to G.K.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Raj A, van Oudenaarden A: **Nature, nurture, or chance: stochastic gene expression and its consequences**. *Cell* 2008, **135**:216-226.

2. Munsky B, Neuert G, van Oudenaarden A: **Using gene expression noise to understand gene regulation**. *Science* 2012, **336**:183-187.

3. Stegle O, Teichmann SA, Marioni JC: **Computational and analytical challenges in single-cell transcriptomics**. *Nat Rev Genet* 2015, **16**:133.

4. Cristinelli S, Ciuffi A: **The use of single-cell RNA-Seq to understand virus–host interactions**. *Curr Opin Virol* 2018, **29**:39-50.

5. Gross A *et al.*: **Technologies for single-cell isolation**. *Int J Mol Sci* 2015, **16**.

6. Prakadan SM, Shalek AK, Weitz DA: **Scaling by shrinking: empowering single-cell 'omics' with microfluidic devices**. *Nat Rev Genet* 2017, **18**:345.

7. •• Zhang X *et al.*: **Comparative analysis of droplet-based ultra-high-throughput single-cell RNA-seq systems**. *Mol Cell* 2019, **73(1)**:130-142.e5 https://doi.org/10.1016/j.molcel.2018.10.020.
This study provides a comparative analysis of results achieved by three different microfluidics platforms (10X Genomics Chromium, DropSeq, and inDrop), highlighting the strengths and limitations of each approach.

8. • Zheng GXY *et al.*: **Massively parallel digital transcriptional profiling of single cells**. *Nat Commun* 2017, **8**:1-12.
The 10X Genomics Chromium platform and Cell Ranger pipeline are used to characterize single cells within the immune system. The authors provide an overview of this method and the utility of its applications.

9. Klein AM *et al.*: **Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells**. *Cell* 2015, **161**:1187-1201.

10. Macosko EZ *et al.*: **Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets**. *Cell* 2015, **161**:1202-1214.

11. Hedlund E, Deng Q: **Single-cell RNA sequencing: technical advancements and biological applications**. *Mol Aspects Med* 2018, **59**:36-46.

12. See P *et al.*: **A single-cell sequencing guide for immunologists**. *Front Immunol* 2018, **9**:2425.

13. Ziegenhain C *et al.*: **Comparative analysis of single-cell RNA sequencing methods**. *Mol Cell* 2017, **65**:631-643.e4.

14. Phipson B, Zappia L, Oshlack A: **Gene length and detection bias in single cell RNA sequencing protocols**. *F1000Res* 2017, **6**:595.

15. Grindberg RV *et al.*: **RNA-sequencing from single nuclei**. *Proc Natl Acad Sci U S A* 2013, **110**:19802-19807.

16. Krishnaswami SR *et al.*: **Using single nuclei for RNA-seq to capture the transcriptome of postmortem neurons**. *Nat Protoc* 2016, **11**:499-524.

17. •• Lake BB *et al.*: **A comparative strategy for single-nucleus and single-cell transcriptomes confirms accuracy in predicted cell-type expression from nuclear RNA**. *Sci Rep* 2017, **7**:6031.
Authors provide a very insightful comparison for using single-nuclei RNA-seq over single-cell to study neuronal complexity. This is especially very useful in situations when access to fresh surgical tissue material is limited.

18. Lake BB *et al.*: **Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain**. *Science* 2016, **352**:1586.

19. Habib N *et al.*: **Massively parallel single-nucleus RNA-seq with DroNc-seq**. *Nat Methods* 2017, **14**:955-958.

20. Bakken TE, Bakken TE, Hodge RD, Miller JA, Yao Z, Nguyen TN: **Single-nucleus and single-cell transcriptomes compared in matched cortical cell types**. *PLOS One* 2018, **13**:e0209648 https://doi.org/10.1371/journal.pone.0209648.

21. Darmanis S *et al.*: **A survey of human brain transcriptome diversity at the single cell level**. *Proc Natl Acad Sci U S A* 2015, **112**:7285-7290.

22. Tasic B *et al.*: **Shared and distinct transcriptomic cell types across neocortical areas**. *Nature* 2018, **563**:72-78.

23. Farrant J: **Mechanism of cell damage during freezing and thawing and its prevention**. *Nature* 1965, **205**:1284.

24. McGann LE, Yang H, Walterson M: **Manifestations of cell damage after freezing and thawing**. *Cryobiology* 1988, **25**:178-185.

25. Lake BB *et al.*: **Integrative single-cell analysis of transcriptional and epigenetic states in the human adult brain**. *Nat Biotechnol* 2018, **36**:70-80.

26. Habib N *et al.*: **Div-Seq: single-nucleus RNA-Seq reveals dynamics of rare adult newborn neurons**. *Science* 2016, **353**:925.

27. Ziegenhain C *et al.*: **Comparative analysis of single-cell RNA sequencing methods**. *Mol Cell* 2017, **65**:631-643 e4.

28. Wu YE *et al.*: **Detecting activated cell populations using single-cell RNA-seq**. *Neuron* 2017, **96**:313-329.e6.

29. • Wu AR *et al.*: **Single-cell transcriptional analysis**. *Annu Rev Anal Chem (Palo Alto Calif)* 2017, **10**:439-462.
This review focuses on almost every aspect of a single-cell RNA-seq experiment. The most interesting fact discussed is regarding the performance evaluation of a single-cell RNA-seq experiment.

30. Saliba A-E *et al.*: **Single-cell RNA-seq: advances and future challenges**. *Nucleic Acids Res* 2014, **42**:8845-8860.

31. Smith Tom, Heger Andreas, Sudbery Ian: **UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy**. *Genome Res* 2017, **27**:491-499 http://dx.doi.org/10.1101/gr.209601.116.

32. Dobin A *et al.*: **STAR: ultrafast universal RNA-seq aligner**. *Bioinformatics* 2013, **29**:15-21.

33. Liao Y, Smyth GK, Shi W: **Feature counts: an efficient general purpose program for assigning sequence reads to genomic features**. *Bioinformatics* 2014, **30**:923-930.

34. Anders S, Pyl PT, Huber W: **HTSeq—a python framework to work with high-throughput sequencing data**. *Bioinformatics* 2015, **31**:166-169.

35. Li H *et al.*: **The sequence alignment/map format and SAMtools**. *Bioinformatics* 2009, **25**:2078-2079.

36. Parekh S *et al.*: **zUMIs - a fast and flexible pipeline to process RNA sequencing data with UMIs**. *GigaScience* 2018, **7** giy059-giy059.

37. Butler A *et al.*: **Integrating single-cell transcriptomic data across different conditions, technologies, and species**. *Nat Biotechnol* 2018, **36**:411.

38. Satija R *et al.*: **Spatial reconstruction of single-cell gene expression data**. *Nat Biotechnol* 2015, **33**:495.

39. Qiu X *et al.*: **Single-cell mRNA quantification and differential analysis with census**. *Nat Methods* 2017, **14**:309.

40. Wolf FA, Angerer P, Theis FJ: **SCANPY: large-scale single-cell gene expression data analysis**. *Genome Biol* 2018, **19**:15.

41. Jolliffe IT: **Principal component analysis and factor analysis**. *Springer Series in Statistics*. New York, NY: Springer; 1986.

42. Cox MAACF: **Multidimensional scaling**. *Handbook of Data Visualization*. Berlin, Heidelberg: Springer; 2008.

43. Sammon JW: **A nonlinear mapping for data structure analysis**. *IEEE Trans Comput* 1969, **C-18**:401-409.

44. van der Maaten L: **Accelerating t-SNE using tree-based algorithms**. *J Mach Learn Res* 2014, **15**.

45. van der Maaten L, Hinton G: **Visualizing non-metric similarities in multiple maps**. *Mach Learn* 2012, **87**:33-55.

46. van der Maaten L: **Learning a parametric embedding by preserving local structure**. *PMLR* 2009:384-391.

47. van der Maaten L, Geoffrey H: **Visualizing data using t-SNE**. *J Mach Learn Res* 2008, **9**.

48. Tenenbaum Joshua B, de Silva Vin, Langford John C: **A global geometric framework for nonlinear dimensionality reduction**. *Science* 2000, **290**:2319.

49. McInnes Leland, Healy John, Saul Nathaniel, Großberger Lukas: **UMAP: uniform manifold approximation and projection**. *J Open Source Software* 2018, **3**:861.

50. Ding J, Condon A, Shah SP: **Interpretable dimensionality reduction of single cell transcriptome data with deep generative models**. *Nat Commun* 2018, **9**:2002.

51. • Becht E *et al.*: **Evaluation of UMAP as an alternative to t-SNE for single-cell data**. *bioRxiv* 2018:298430.
Authors have identified UMAP as a better alternative to t-SNE for visualizing single-cell transcriptomic data in order to identify clusters. Use of UMAP is also crucial to study developmental trajectories using single-cell data.

52. •• La Manno G *et al.*: **RNA velocity of single cells**. *Nature* 2018, **560**:494-498.
Studying developmental trajectory is a very unique feature of single cell transcriptomic data. Authors have very systematically made use of the data to extract trajectory information.

53. • Lein E, Borm LE, Linnarsson S: **The promise of spatial transcriptomics for neuroscience in the era of molecular cell typing**. *Science* 2017, **358**:64-69.
The authors provide an in-depth review of the history of, and current progress in, attempts to unify spatial information and transcriptomic data.

54. Crosetto N, Bienko M, van Oudenaarden A: **Spatially resolved transcriptomics and beyond**. *Nat Rev Genet* 2015, **16**:57.

55. Zeng H, Sanes JR: **Neuronal cell-type classification: challenges, opportunities and the path forward**. *Nat Rev Neurosci* 2017, **18**:530.

56. Yuan G-C *et al.*: **Challenges and emerging directions in single-cell analysis**. *Genome Biol* 2017, **18**:84.

57. Raj A *et al.*: **Imaging individual mRNA molecules using multiple singly labeled probes**. *Nat Methods* 2008, **5**:877.

58. Shah S *et al.*: **In situ transcription profiling of single cells reveals spatial organization of cells in the mouse hippocampus**. *Neuron* 2016, **92**:342-357.

59. Lubeck E *et al.*: **Single-cell in situ RNA profiling by sequential hybridization**. *Nat Methods* 2014, **11**:360.

60. Lubeck E, Cai L: **Single-cell systems biology by super-resolution imaging and combinatorial labeling**. *Nat Methods* 2012, **9**:743.

61. •• Moffitt JR *et al.*: **Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region**. *Science* 2018, **362**.
This study uses MERFISH to create a spatially informed transcriptomic map of single cells in the mouse hypothalamic preoptic region. Furthermore, distinct cell populations are associated with specific social behaviors.

62. Chen KH *et al.*: **Spatially resolved, highly multiplexed RNA profiling in single cells**. *Science* 2015, **348**.

63. Moffitt JR, Zhuang X: **RNA imaging with multiplexed error-robust fluorescence in situ hybridization (MERFISH)**. In *Methods in Enzymology*. Edited by Filonov GS, Jaffrey SR. Academic Press; 2016:1-49.

64. Moffitt JR *et al.*: **High-throughput single-cell gene-expression profiling with multiplexed error-robust fluorescence in situ hybridization**. *Proc Natl Acad Sci U S A* 2016, **113**:11046-11051.

65. Wang X *et al.*: **Three-dimensional intact-tissue sequencing of single-cell transcriptional states**. *Science* 2018:eaat5691.

66. Zhu Q *et al.*: **Identification of spatially associated subpopulations by combining scRNAseq and sequential fluorescence in situ hybridization data**. *Nat Biotechnol* 2018, **36**:1183-1190 http://dx.doi.org/10.1038/nbt.4260.

67. Karagiannis ED, Boyden ES: **Expansion microscopy: development and neuroscience applications**. *Curr Opin Neurobiol* 2018, **50**:56-63.

68. Alon S, Huynh GH, Boyden ES: **Expansion microscopy: enabling single cell analysis in intact biological systems**. *FEBS J* 2018 http://dx.doi.org/10.1111/febs.14597.

69. Codeluppi S *et al.*: **Spatial organization of the somatosensory cortex revealed by osmFISH**. *Nat Methods* 2018, **15**:932-935.

70. Nowakowski TJ *et al.*: **Spatiotemporal gene expression trajectories reveal developmental hierarchies of the human cortex**. *Science* 2017, **358**:1318.

71. Saunders A *et al.*: **Molecular diversity and specializations among the cells of the adult mouse brain**. *Cell* 2018, **174**:1015-1030.e16.

72. Zeisel A *et al.*: **Molecular architecture of the mouse nervous system**. *Cell* 2018, **174**:999-1014.e22.

73. Zhong S *et al.*: **A single-cell RNA-seq survey of the developmental landscape of the human prefrontal cortex**. *Nature* 2018, **555**:524.

74. Hrvatin S *et al.*: **Single-cell analysis of experience-dependent transcriptomic states in the mouse visual cortex**. *Nat Neurosci* 2018, **21**:120-129.

75. Quadrato G *et al.*: **Cell diversity and network dynamics in photosensitive human brain organoids**. *Nature* 2017, **545**:48.

76. •• Renthal W *et al.*: **Characterization of human mosaic Rett syndrome brain tissue by single-nucleus RNA sequencing**. *Nat Neurosci* 2018, **21**:1670-1679.
These authors develop a new application of single-cell transcriptomics (SNP-seq) for overcoming challenges for characterizing X-linked neurodevelopmental disorders.

77. •• Zywitza V *et al.*: **Single-cell transcriptomics characterizes cell types in the subventricular zone and uncovers molecular defects impairing adult neurogenesis**. *Cell Rep* 2018, **25**:2457-2469.e8.
One of the first studies to use scRNA-seq in comparative analysis; in this case, LRP2 knockout versus littermate control cells were compared to better understand the cell-types affected by LRP2 ablation and adult neurogenesis.

78. Kharchenko PV, Silberstein L, Scadden DT: **Bayesian approach to single-cell differential expression analysis**. *Nat Methods* 2014, **11**:740.

79. Fan J *et al.*: **Characterizing transcriptional heterogeneity through pathway and gene set overdispersion analysis**. *Nat Methods* 2016, **13**:241.

80. Trapnell C *et al.*: **The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells**. *Nat Biotechnol* 2014, **32**:381.