JOURNAL OF Bio-X Research

REVIEW ARTICLE

http://dx.doi.org/10.1097/JBR.0000000000000003

OPEN

Single-cell sequencing leads a new era of profiling transcriptomic landscape

Huidan Zhang, PhD, MD[a,*], Naiwen Cui[a], Yamei Cai, PhD[a], Fengyang Lei, PhD, MD[a,b], David A. Weitz, PhD[a,c,*]

Abstract

Understanding the complexity of biological systems requires a comprehensive analysis of their cell populations. Ideally, this should be done at the single cell level, because bulk analysis of the full population obscured many critical details due to artifacts introduced by averaging. However, this has been technically challenging due to the cumbersome procedure, low throughput, and high costs of performing analysis on a single-cell basis. Excitingly, technical improvements in single-cell RNA sequencing are making it economically practical to profile the transcriptomics of large populations of cells at the single-cell level, and have yielded numerous results that address important biological and medical questions. Further development of the technology and data analysis will significantly benefit the biomedical field by unraveling the function of individual cells in their microenvironments and modeling their transcriptional dynamics.

Keywords: cell transcriptomics, CEL-Seq, Hi-SCL, MARS-Seq, SCRB-Seq, single-cell RNA sequencing, Smart-Seq2, sNuc-seq

Downloaded from https://journals.lww.com/jbioxresearch by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3OleNjk8BaNzRZRh5kRNAgPxwz16p83XaDGS5cgmwas= on 07/05/2018

Urgency of developing single-cell RNA sequencing technology

Traditional bulk RNA sequencing for studying cell transcriptomics measures the mean values from millions of cells. It is very difficult to accurately determine differentially expressed genes, except on an averaged basis. Nevertheless, this method has had great impact on biomedical research,[1] diagnostics,[2] and therapeutics.[3] However, even greater impact demands information at the level of individual cells, still collected over the full population of cells. Single-cell RNA sequencing (scRNA-seq) overcomes the averaging artifacts of bulk analysis by providing the expression profile of individual cells, and thus allows the analysis of the internal cell-to-cell differences and interactions within a biological system.[4,5] For example, through gene clustering analyses of scRNA-seq, different cell subpopulations can be resolved, thereby enabling characterization of a heterogeneous system.[6] Similarly, if scRNA-seq on a large number of single cells is possible, rare cell types that perform important functions can be identified,[7] providing valuable insights for diagnostics and treatment. These representative examples, and many others, pose the urgency of developing scRNA-seq technology.

The early implementations of scRNA-seq illustrated the power and potential of data at a single-cell level; however, they were technically challenging due to cumbersome procedures, low throughput and high costs. Excitingly, ongoing technical improvements of scRNA-seq have made it much more robust and economically practical to profile single-cell transcriptomics at the full population level. These studies have yielded numerous sets of data that address many biological and medical questions. In this review, we first describe the scRNA-seq methods and then discuss the technical platforms that have been developed to implement them. We discuss the advantages and limitations of the methods, with a focus on high-throughput scRNA-seq.

Fundamental development of scRNA-seq methodology

Smart-Seq and Smart-Seq2

Smart-Seq, Switching Mechanism at 5′ End of RNA Template, is a milestone in the development of scRNA-seq technology, contributing to the analysis of cell types existing in rare quantities in a heterogeneous population that are both biologically and clinically important.[8,9] As first developed, cells are individually isolated in distinct wells. Each cell is, and then reverse-transcribed into first-strand cDNA with tailed oligo(dT) using moloney murine leukemia virus reverse transcriptase. This also has a terminal transferase thereby adding a few nontemplated C nucleotides to the 3′ end of the cDNA when the reverse transcription (RT) reaction reaches the 5′ end of an RNA molecule. With the presence of a template switching oligo (TSO), tailed rGrGrG, moloney murine leukemia virus reverse transcriptase switches templates and continues synthesis of second-strand cDNA. The resulting full-length double-stranded cDNA is preamplified with the tail sequence on both ends, followed by

HZ and NC contributed equally to the writing of the article.

[a] John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, [b] Department of Ophthalmology, Massachusetts Eye and Ear Infirmary, Harvard Medical School, Boston, [c] Department of Physics, Harvard University, Cambridge, MA, USA

[*] Corresponding author: Huidan Zhang, John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA, E-mail: hdzhang@seas.harvard.edu; and David A. Weitz, John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA. E-mail: weitz@seas.harvard.edu

Copyright © 2018 The Chinese Medical Association, Published by Wolters Kluwer Health, Inc. under the CCBY-NC-ND license. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

Journal of Bio-X Research (2018) 1:2–6

Received: 15 May 2018; Accepted: 15 May 2018

Published online 25 June 2018

http://dx.doi.org/10.1097/JBR.0000000000000003

construction of an Illumina sequencing library using either Illumina's shearing-ligation protocol or Epicentre's Nextera Tn5 transposon protocol. Smart-Seq is a full-length mRNA sequencing method, and is thus highly informative for identifying candidate biomarkers, single nucleotide polymorphisms and mutations.

With a few crucial technical modifications, Smart-Seq2 exhibits significant improvements.[10] Replacing the last guanylate at the TSO 3′ end with a locked nucleic acid (LNA) doubles the cDNA yield obtained with the TSO in Smart-Seq due to the increased thermal stability of LNA:DNA base pairs. The presence of the methyl group donor betaine in combination with higher $MgCl_2$ concentrations increases the cDNA yield by a factor of between 2 and 4. Adding deoxyribonucleoside triphosphates before the RNA denaturation rather than in the RT master mix increases the average length of the preamplified cDNA by 370 nt, presumably through mechanism that stabilizes the hybridization of RNA to the oligo(dT) primer. Removing bead purification step after first-stand cDNA synthesis combined with utilizing KAPA HiFi HotStart DNA polymerase during preamplification effectively avoids losing material with the promise of providing a good amplification efficacy, and more powerfully, the resulting average cDNA length is 450 nt greater. Thus, Smart-Seq2 has improved both the yield and the length of cDNA libraries generated from individual cells.

These methods to generate single-cell transcriptome libraries use off-the-shelf reagents, making them accessible for most labs. As a result they have been widely used.

CEL-Seq and CEL-Seq2

CEL-Seq, Cell Expression by Linear Amplification and Sequencing, relies on linear amplification.[11] The initial implementation was again done using cells in individual wells of a plate. After each cell is lysed, a tailed oligo(dT) is used to prime RT. From the 5′ end to the 3′ end, the sequence of the tailed oligo(dT) is T7 promoter, partial Illumina 5′ adapter, cell barcode, and polyT stretch. The second-strand cDNA is then synthesized to generate a double-stranded cDNA containing a T7 promoter. The cDNA are pooled and an in vitro transcription reaction is initiated to achieve linear amplification of cDNA. The amplified RNA (aRNA) generated are fragmented to a size distribution suitable for sequencing, and then the Illumina 3′ adaptor is added through ligation and reverse transcribing to DNA. Finally, the 3′-most fragments that contain both Illumina adaptors and a barcode are selectively amplified. The resultant amplicons undergo paired-end sequencing, where the first read recovers the barcode, and the second identifies the mRNA transcript. CEL-Seq gives more reproducible, linear, and sensitive results than a PCR-based amplification method.[12] Compared to Smart-Seq, CEL-Seq adds the barcode at an earlier stage, decreasing the hands-on work; however, it is only used for 3′-end sequencing.

A modified version of CEL-Seq, CEL-Seq2[13] adds a unique molecular identifier (UMI) upstream of the barcode to distinguish between PCR duplicates and transcript abundance in RNA scRNA-seq; this significantly improves the accuracy.[14–16] Then, the application of the Super-Script® II Double-Stranded cDNA Synthesis Kit combined with a shortening of the CEL-Seq primer provides great improvement in RT efficiency, thereby increasing the sensitivity yielding better detection of both transcripts and genes. Furthermore, the methods for dsDNA and aRNA clean-up are changed from a column to beads. Finally, instead of using low-efficiency ligation during the conversion of aRNA to a library compatible with Illumina sequencing, CEL-Seq2 inserts the Illumina adaptor directly at the RT step as a 5′-tail attached to a random hexamer. Collectively, twice as many transcripts per cell and 30% more genes are determined by CEL-Seq2 as compared to the original CEL-Seq protocol.

Similar to Smart-Seq, CEL-Seq uses off-the-shelf reagents to generate single-cell transcriptome libraries making it accessible to most labs in a 96-well plate platform.

Further development of scRNA-seq

SCRB-Seq

SCRB-Seq, Single Cell RNA Barcoding and Sequencing, is built on Smart-Seq, but only performs 3′ end sequencing, and is specifically optimized for profiling mRNA from a large number of cells using minimal reagents, reagent handling and sequencing reads per cell. It is particularly suitable for discovery of the major patterns of transcriptomics across heterogeneous populations.[17] In brief, single cells are first sorted into a 384-well plate with a fluorescence-activated cell sorter. The synthesis of cDNA is then primed using RT primers that are composed of partial Illumina adapters, well-specific barcodes, UMIs, and polyT stretches. The resultant cDNA from all cells is pooled, amplified, and prepared for sequencing using a modified transposon-based fragmentation approach that enriches for 3′ ends. In the first study, SCRB-Seq libraries from 44 microplates were sequenced yielding single-cell results from 12,832 cells.[17] SCRB-Seq is a complementary to protocols that are optimized for deep, full-length transcriptome coverage, such as Smart-Seq. The main limitation of SCRB-Seq is its potential for further scale up.

MARS-Seq

MARS-Seq, Massively Parallel RNA Single-Cell Sequencing, is built on CEL-Seq, but it is optimized for developing an automated workflow to analyze transcriptomics of thousands of cells, such that throughput and reproducibility are well promised.[18] In this method, the RT primer includes a T7 promoter, a partial Illumina adapter, a cell barcode, a UMI, and a polyT stretch. Single cells are first sorted into 384-well plates with a fluorescence-activated cell sorter, and subsequently, automated processing for library preparation is done mostly on pooled and labeled material, leading to a dramatic increase in throughput and reproducibility. In the first study, a total of 1536 cells were sequenced and 200 to 1500 distinct RNA molecules from each cell were unambiguously defined. With the progress of technology, MARS-Seq has been replaced by emerging high-throughput scRNA-seq methods.

Breakthrough of high-throughput scRNA-seq

The large number of early experiments and techniques clearly demonstrates that single-cell sequencing provides significant new information that is of great value in biology and medical applications. However, these methods are all cumbersome and difficult to perform, since they are all accomplished by isolating single cells in wells; this also limits the number of cells that can be probed, and hence limits the size of the population that can be explored. Further progress requires a means to increase the throughput of cells probed, while still exploiting the basic sample preparation techniques already developed.

This problem is ideally suited to the use of microfluidics technology to improve the throughput. The first commercial system was developed by Fluidigm. It produced the C1 Single Cell

Auto Prep system, which revolutionized single-cell sequencing.[19] The basic platform is based enables the capture of up to 96 cells in individual microfluidic wells, followed by an automated procedure to implement the Smart-Seq or CEL-Seq2 process on each cell captured in the chip. The only hands-on work is to prepare single-cell suspension and load into the C1 microfluidic chip. While the system can probe up to 96 cells, they are loaded randomly into the chip that typically yields considerably fewer cells than full capacity; moreover, only a small number of cells are actually captured, making the system inefficient in cell use, and not suitable for the study of limited populations of cells. Furthermore, the cost of the chip and operation of the system remains high. Subsequently, Fluidigm improved the performance of the C1 by introducing a chip that increases the throughput to up to 800 cells, however, further scaled up would be very challenging.[20,21]

The greatest increase in throughput for scRNA-seq has come from the development of drop-based microfluidic methods. The first of these is Hi-SCL (High-Throughput Single-Cell Labeling),[22] which was followed by inDrop (indexing Droplets) and Drop-Seq.[23,24] These methods are all based on compartmentalization of an individual cell, and a unique cell barcode, into a uniform, nanoliter water-in-oil droplet created by drop-based microfluidic techniques. These droplets significantly reduce the volume in which each cell in enclosed, particularly in comparison to the microliter volume commonly used with traditional well-based methods. The droplets reduce the volume by at least 3 orders of magnitude. Moreover, they can be generated at rates approaching thousands of droplets per second.[25] Furthermore, the inDrop technique very efficiently probes virtually all cells, and is hence suitable for populations of rare cells, where each cell must be utilized. Both inDrop and Drop-Seq have a very large capacity, and are capable of barcoding millions of cells with only microliter-level reagent in just a few hours. They are ultimately limited by the bandwidth of the sequencing used. In addition, the cost of droplet-based scRNA-seq is much lower than other scRNA-seq methods, facilitating studies of a large number of single-cell transcriptomics.[20,21,26]

In these methods, the key technology to label of individual cell is to combine a barcoded droplet or a barcoded particle with each cell. In Hi-SCL, barcoded droplets are generated using a single microfluidic chip with 96 parallel drop-makers. Repeating this procedure multiple times can produce a larger number of different barcoded droplets. Then the droplets containing lysed cells are merged with the barcode drops and RT mix on a microfluidic-merging device.[22] The main limitation of Hi-SCL is that it is difficult to scale the size of the barcode library. To overcome this limitation, inDrop uses hydrogel microparticles as carriers of barcodes. Each hydrogel particle has 10^9 identical primers consisting of a T7 promoter, a partial Illumina adapter, drop- and therefore cell-specific barcodes, a synthesis adaptor, a UMI and a polyT stretch. The library construction basically follows that of CEL-Seq. A commonly used pool size of barcodes is 384×384 (147,456) barcodes; this can be readily increased by performing more split-and-pool cycles[23] to increase the length of the bead-specific barcode. By comparison, Drop-Seq uses a hard microparticle that contains 10^8 identical primers consisting of a partial Illumina adapter, a cell-specific barcode, a UMI, and a polyT stretch. The Smart-Seq protocol is used to prepare the library with the exception of sequencing the RNA only at the 3′ end. Its current pool size of barcodes is 4^{12} (16,777,216), which is generally sufficient for most applications.[24]

Another difference between these 2 methods is that the flow of hydrogel particles in Drop-Seq can be synchronized because their deformability allows close packing and regular loading into each drop as it is formed to achieve close to 100% droplet occupancy of hydrogels[26]; therefore, even when the cells are diluted to ensure that the vast majority of drops have at most one cell, every drop that contains a cell will also contain a barcoded particle. This feature is particularly valuable when the number of cells is limited, and each cell is highly valuable. By contrast, the hard particles used in Drop-Seq cannot be close-packed and therefore, to ensure at most a single barcode in each drop, the particles must be highly diluted so that the majority of drops do not have any particles. Thus, the percentage of drops containing both a cell and a barcoded particle is much less than inDrop due to the double-Poisson distribution, as shown in Fig. 1. Moreover, because of these Poisson statistics, most cells are in drops that do not contain a particle, meaning that they will not be included in the analysis. Thus, Drop-Seq is only suitable for experiments with an abundance of cells allowing most of them to not be analyzed. Ultimately, the choice of method for scRNA-seq depends on the requirement of a specific project.

Recently, more scRNA-seq methods have been developed using the droplet microfluidic platform. For example, sNuc-seq, Single Nucleus RNA-seq, was developed to massively profile transcriptomics of single nuclei for samples that cannot be dissociated into intact single cells.[27] Here, the sensitivity, efficiency and uniformity of classification of cell types were demonstrated by profiling 39,111 nuclei from mouse and human brain samples.[27] Another interesting area where droplet technology is being applied is to the study of the ultra-heterogeneity of immune system, including both profiling of transcriptomics of immune cells and their immune repertoires.[6,28–32]

Currently, there are several commercially available droplet-based scRNA-seq platforms. The most widely used is the Chromium Single Cell Gene Expression system from 10× genomics,[33] which is based on a drop loading system similar to that of inDrop, in that deformable hydrogel barcode beads are used. In addition, the inDrop™ System from 1CellBio[34] reproduces the inDrop method directly, while the Droplet encapsulation system from Dolomite Microfluidics[31] and the Nadia Instrument from Dolomite Bio[35] both use the Drop-Seq method. The availability of all these technologies and platforms ensures that scRNA-seq will have a large impact on biomedical research to revolutionize therapies for cancer and autoimmune diseases.

Summary and outlook

Single-cell RNA-seq has become a well established and widely used technology. Its sensitivity, accuracy, and throughput have improved significantly, while its costs have decreased substantially. Combining single-cell transcriptomic data with temporal information,[36] spatial information,[37,38] genomic sequencing data,[39,40] and epigenomic sequencing data[41] provides ever more precise transcriptional dynamics, helping elucidate the mechanisms underlying gene regulation, and key regulator genes, while exploring heterogeneous cell populations. These advances in single-cell sequencing technology enable systematic charting of the cell atlas, to define all human cell types in terms of distinctive molecular transcriptomic profiles and to correlate this information with conventional cell locations and morphology, thereby providing a framework for understanding cellular dysregulation in a variety of human diseases.[42–45]

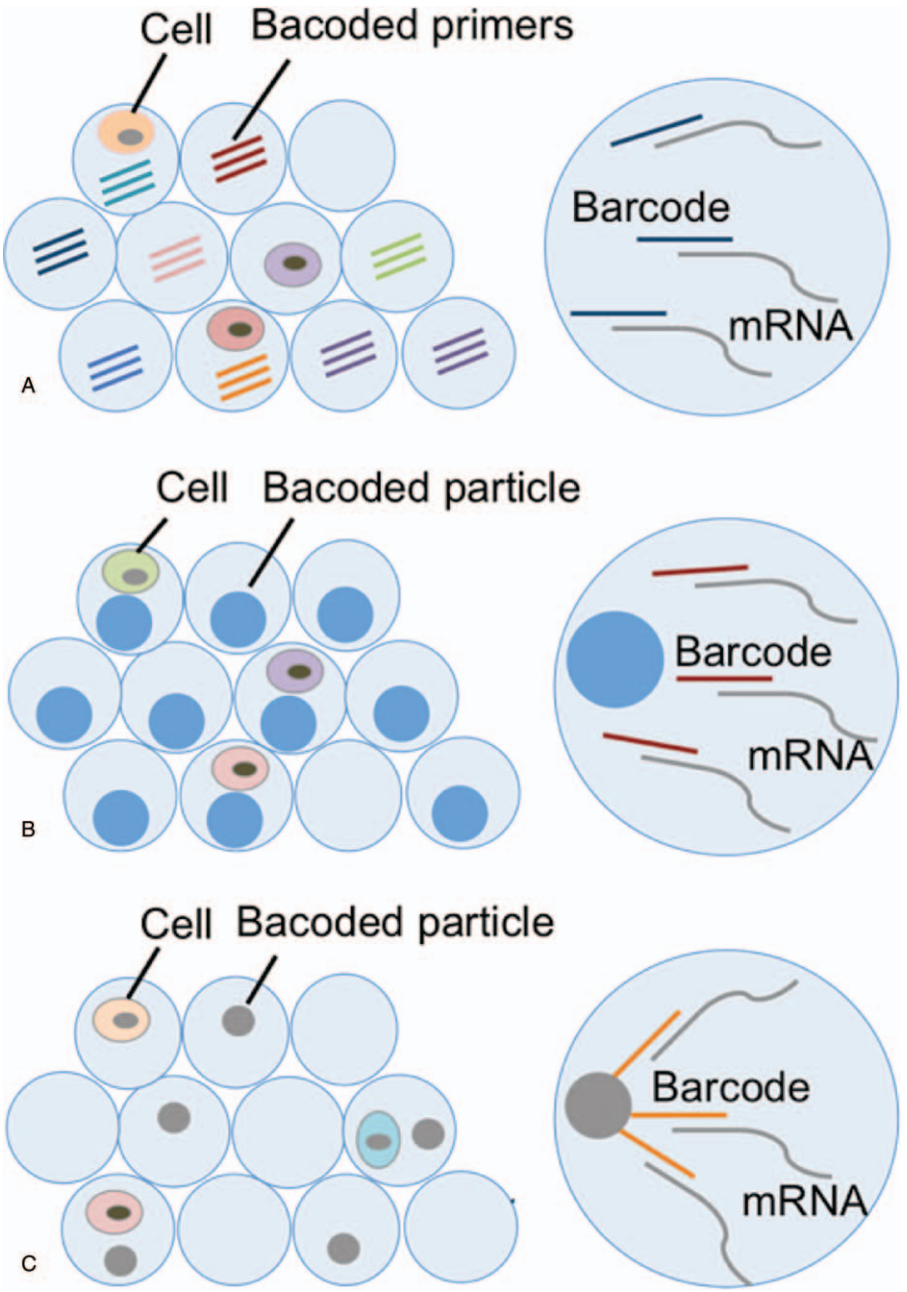

Figure 1. Different approaches of barcoding mRNA molecules in single cells using droplet-based microfluidics. (A) High-throughput single-cell labeling (Hi-SCL).[22] Droplets containing cells are merged with droplets containing free barcoded primers. Then, the barcoded primers pair with mRNA molecules released from cells and start priming reverse transcription. (B) Indexing Droplets (inDrop).[23] Cells are encapsulated into drops with deformable hydrogel microparticles, followed by photo-releasing of barcodes into solution and perform reverse transcription inside. (C) Drop-Seq.[24] Cells are encapsulated into drops with hard microparticles, where the mRNA molecules are captured onto microparticles.

Here, we have addressed main scRNA-seq methods and their developmental history. Given the requirement of analyzing more and more cells, the cost of instruments, reagents, consumables, labor, and sequencing remains high. Therefore, scRNA-seq methods must be carefully chosen based on the specific application. Optimization of the protocol, including the efficiency of capturing and converting mRNA transcripts into cDNA molecules, the precision of quantification of mRNA, and various features across different types of cells is essential. Further development of technology and data analytics will benefit the biomedical field and help unravel the function of individual cells in their individual microenvironments and model their transcriptional dynamics. Finally, new discoveries from scRNA-seq must still be validated due to potential bias in the profiling of transcriptomics and due to possible computational errors. However, the rapid development of the field ensures that it will continue to advance and have major impact on biology and medical applications.

Acknowledgements

None.

Author contributions

DAW and HZ conceived the manuscript. All authors participated in the discussions and writing of the manuscript. All authors read and approved the final manuscript.

Financial support

This work was supported by the Harvard Materials Research Science and Engineering Center (NSF DMR-1420570), National Science Foundation (DMR-1708729), National Institutes of Health (P01HL120839), Harvard-Suzhou Industrial Park Research and engineering innovation initiative grant, and National Natural Science Foundation of China (81372496).

Conflicts of interest

DAW is a co-founder of 1-Cell Bio. The other authors declare no conflicts of interest.

References

[1] Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 2009;10:57–63.
[2] Byron SA, Van Keuren-Jensen KR, Engelthaler DM, et al. Translating RNA sequencing into clinical diagnostics: opportunities and challenges. Nat Rev Genet 2016;17:257–271.
[3] Kim KT, Lee HW, Lee HO, et al. Application of single-cell RNA sequencing in optimizing a combinatorial therapeutic strategy in metastatic renal cell carcinoma. Genome Biol 2016;17:80.
[4] Villani AC, Satija R, Reynolds G, et al. Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. Science 2017;356:eaah4573.
[5] Tirosh I, Izar B, Prakadan SM, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science 2016;352:189–196.
[6] Papalexi E, Satija R. Single-cell RNA sequencing to explore immune cell heterogeneity. Nat Rev Immunol 2018;18:35–45.
[7] Grun D, Lyubimova A, Kester L, et al. Single-cell messenger RNA sequencing reveals rare intestinal cell types. Nature 2015;525:251–255.
[8] Ramskold D, Luo S, Wang YC, et al. Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. Nat Biotechnol 2012;30:777–782.
[9] Goetz JJ, Trimarchi JM. Transcriptome sequencing of single cells with Smart-Seq. Nat Biotechnol 2012;30:763–765.
[10] Picelli S, Bjorklund AK, Faridani OR, et al. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Nat Methods 2013;10:1096–1098.
[11] Hashimshony T, Wagner F, Sher N, et al. CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. Cell Rep 2012;2:666–673.
[12] Islam S, Kjallquist U, Moliner A, et al. Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. Genome Res 2011;21:1160–1167.
[13] Hashimshony T, Senderovich N, Avital G, et al. CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq. Genome Biol 2016;17:77.
[14] Kivioja T, Vähärautio A, Karlsson K, et al. Counting absolute numbers of molecules using unique molecular identifiers. Nat Methods 2011;9:72–74.
[15] Shiroguchi K, Jia TZ, Sims PA, et al. Digital RNA sequencing minimizes sequence-dependent bias and amplification noise with optimized single-molecule barcodes. Proc Natl Acad Sci U S A 2012;109:1347–1352.
[16] Stegle O, Teichmann SA, Marioni JC. Computational and analytical challenges in single-cell transcriptomics. Nat Rev Genet 2015;16:133–145.
[17] Soumillon M, Cacchiarelli D, Semrau S, et al. Characterization of directed differentiation by high-throughput single-cell RNA-Seq. bioRxiv 2014;003236.
[18] Jaitin DA, Kenigsberg E, Keren-Shaul H, et al. Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. Science 2014;343:776–779.
[19] Wu AR, Neff NF, Kalisky T, et al. Quantitative assessment of single-cell RNA-sequencing methods. Nat Methods 2014;11:41–46.
[20] Han X, Wang R, Zhou Y, et al. Mapping the mouse cell atlas by Microwell-Seq. Cell 2018;172:1091.e17–1107.e17.
[21] Ziegenhain C, Vieth B, Parekh S, et al. Comparative analysis of single-cell RNA sequencing methods. Mol Cell 2017;65:631.e4–643.e4.
[22] Rotem A, Ram O, Shoresh N, et al. High-throughput single-cell labeling (Hi-SCL) for RNA-Seq using drop-based microfluidics. PLoS ONE 2015;10:e0116328.
[23] Klein AM, Mazutis L, Akartuna I, et al. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. Cell 2015;161:1187–1201.
[24] Macosko EZ, Basu A, Satija R, et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. Cell 2015;161:1202–1214.
[25] Guo MT, Rotem A, Heyman JA, et al. Droplet microfluidics for high-throughput biological assays. Lab Chip 2012;12:2146–2155.
[26] Abate AR, Chen CH, Agresti JJ, et al. Beating Poisson encapsulation statistics using close-packed ordering. Lab Chip 2009;9:2628–2631.
[27] Habib N, Avraham-Davidi I, Basu A, et al. Massively parallel single-nucleus RNA-seq with DroNc-seq. Nat Methods 2017;14:955–958.
[28] Stubbington MJT, Rozenblatt-Rosen O, Regev A, et al. Single-cell transcriptomics to explore the immune system in health and disease. Science 2017;358:58–63.
[29] Medaglia C, Giladi A, Stoler-Barak L, et al. Spatial reconstruction of immune niches by combining photoactivatable reporters and scRNA-seq. Science 2017;358:1622–1626.
[30] Zheng GX, Terry JM, Belgrader P, et al. Massively parallel digital transcriptional profiling of single cells. Nat Commun 2017;8:14049.
[31] Briggs A, Goldfless S, Timberlake S, et al. Tumor-infiltrating immune repertoires captured by single-cell barcoding in emulsion. bioRxiv 2017;134841.
[32] DeKosky BJ, Kojima T, Rodin A, et al. In-depth determination and analysis of the human paired heavy- and light-chain antibody repertoire. Nat Med 2015;21:86–91.
[33] Skelly DA, Squiers GT, McLellan MA, et al. Single-cell transcriptional profiling reveals cellular diversity and intercommunication in the mouse heart. Cell Rep 2018;22:600–610.
[34] Baron M, Veres A, Wolock SL, et al. A single-cell transcriptomic map of the human and mouse pancreas reveals inter- and intra-cell population structure. Cell Syst 2016;3:346.e4–360.e4.
[35] Croset V, Treiber CD, Waddell S. Cellular diversity in the Drosophila midbrain revealed by single-cell transcriptomics. Elife 2018;7:e34550.
[36] Bendall SC, Davis KL, Amir el AD, et al. Single-cell trajectory detection uncovers progression and regulatory coordination in human B cell development. Cell 2014;157:714–725.
[37] Satija R, Farrell JA, Gennert D, et al. Spatial reconstruction of single-cell gene expression data. Nat Biotechnol 2015;33:495–502.
[38] Lee JH, Daugharthy ER, Scheiman J, et al. Highly multiplexed subcellular RNA sequencing in situ. Science 2014;343:1360–1363.
[39] Dey SS, Kester L, Spanjaard B, et al. Integrated genome and transcriptome sequencing of the same cell. Nat Biotechnol 2015;33:285–289.
[40] Macaulay IC, Haerty W, Kumar P, et al. G&T-seq: parallel sequencing of single-cell genomes and transcriptomes. Nat Methods 2015;12:519–522.
[41] Smallwood SA, Lee HJ, Angermueller C, et al. Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. Nat Methods 2014;11:817–820.
[42] Regev A, Teichmann SA, Lander ES, et al. The human cell atlas. Elife 2017;6:e27041.
[43] Hutter C, Zenklusen JC. The cancer genome atlas: creating lasting value beyond its data. Cell 2018;173:283–285.
[44] Sanchez-Vega F, Mina M, Armenia J, et al. Oncogenic signaling pathways in the cancer genome atlas. Cell 2018;173:321.e10–337.e10.
[45] Thorsson V, Gibbs DL, Brown SD, et al. The immune landscape of cancer. Immunity 2018;48:812.e14–830.e14.