Open Access

CrossMark

Single cell sequencing: a distinct new field

Jian Wang and Yuanlin Song^{*}

Abstract

Single cell sequencing (SCS) has become a new approach to study biological heterogeneity. The advancement in technologies for single cell isolation, amplification of genome/transcriptome and next-generation sequencing enables SCS to reveal the inherent properties of a single cell from the large scale of the genome, transcriptome or epigenome at high resolution. Recently, SCS has been widely applied in various clinical and research fields, such as cancer biology and oncology, immunology, microbiology, neurobiology and prenatal diagnosis. In this review, we will discuss the development of SCS methods and focus on the latest clinical and research applications of SCS.

Introduction

The majority of experimental and clinical results from cell culture or tissues are based on the assumption that all of the cells in a culture or tissue are homogeneous. The thriving omics fields of study (genomics, proteomics, transcriptomics, etc.) analyze and mine biomarkers mainly based on the bulk of cells or tissue samples. However, this averaging of messages always misses the critical information because the heterogeneity of the samples is ignored, while the nature of biology is diverse. Heterogeneity is generally explained at three different levels in the biological universe: first, there is heterogeneity in different organisms; second, there is heterogeneity in different organs or tissues from an organism; third, cellular heterogeneity exists in the same organ or tissue. In fact, the concept of cellular heterogeneity was proposed as early as 1957 [1]. Each cell was considered as a unique unit with molecular coding across the DNA, RNA, and protein conversions [2]. Thus, it is necessary to conduct studies, especially omics studies, at the single cell level.

A single cell is the smallest structural and functional unit of an organism. The estimated number of single cells in the human body is 3.72×10^{13} [3]. The size or weight of a cell varies from different tissue backgrounds. The major components of a cell include water, inorganic ions, small organic molecules, proteins, RNA and DNA.

*Correspondence: song.yuanlin@zs-hospital.sh.cn Department of Pulmonary Medicine, Zhongshan Hospital, Fudan University. Shanghai 200030. China However, the minute numbers of copies of a gene (10-12 M) in a single cell are more than enough for conventional genomic analysis [2, 4]. In 2009, the first single cell whole transcriptome sequencing (WTA) protocol was applied to analyze transcriptome complexity in individual cells [5]. Subsequently, single cell whole genome sequencing (WGS) was created in 2011 [6], single cell whole exome sequencing (WES) was developed in 2012 [7, 8], and single cell epigenomic sequencing was developed in 2013 [9]. Currently, single cell sequencing (SCS) has been applied in various research and clinical fields, and the top five areas of SCS studies in order are cancer, embryonic development, microbiology, neurobiology and immunology, according to the reported statistics [10]. The number of SCS publications in these five areas has been increasing every year. Thus, this article will enable us to have a deep and broad view of SCS methods and to focus on the latest application of SCS in basic and clinical research.

Single cell isolation methods

Isolating single cells from a tissue mass or from cell culture is the first key step prior to SCS. Currently, the alternative methods used to isolate single cells from abundant populations include serial dilution, mechanical micromanipulation, laser capture microdissection (LCM), fluorescence activated cell sorting (FACS), and microfluidics [11, 12]. Although serial dilution is the simplest method to obtain a single cell in a single well via serial double



© The Author(s) 2017. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

dilution, it is a coarse and imprecise method that is rarely used in SCS (Fig. 1a). Our team has tried to use this method to isolate a single cell from primary lung cancer cells in cell suspension and found that it was hard to control the quality and quantity [13].

Mechanical micromanipulation is a classic method to isolate uncultivated microorganisms or early embryos, and it involves using a capillary pipette to suck up a single cell from a cell suspension with visual inspection of cellular morphology and coloring characteristics under a microscope [13, 14] (Fig. 1b). The drawback of mechanical micromanipulation is that it is low-throughput and time-consuming and can cause cellular injury from mechanical shearing during manipulation [15]. Additionally, it often leads to a failure for an unskilled manipulator or misidentification of the cellular morphology under the microscope.

FACS is the most efficient and economical method to isolate hundreds of thousands of individual cells per minute based on their size, granularity and fluorescence properties [4] (Fig. 1c). The high-throughput, time-saving and automatic properties are the main advantage of FACS. Additionally, it allows researchers to isolate

specific individual cells from heterogeneous cell samples by labeling the targeted cells with specific fluorescent antibodies [16], and it allows researchers to sort a single viral particle from a mixed viral assemblage for single viral genome sequencing [17]. BD Aria II/III (BD Biosciences, San Jose, CA, USA) and Beckman Coulter MO-FLO XDP cell sorter (Beckman Coulter, Brea, CA, USA) are two widely used commercial instruments for flow cytometry [11]. Our team has used the BD Aria III to sort individual living cells from lung cancer tissue single cell suspensions that were stained with carboxyfluorescein diacetate succinimidyl ester (CFSE) and sorted into 96well plates [18]. However, a bulk population of the cells (at least $5 \times 10^5 - 1 \times 10^6$ /ml) should be prepared as sorting material, which is greatly limited in accommodating low-abundance cell subpopulations. The high-speed fluid and fluorescent dye can damage the viability of cells.

Microfluidics is a newly developed and highly integrated system that sequentially processes or manipulates small volumes of fluids $(10^{-9}-10^{-18} \text{ l})$ in channels with dimensions of tens to hundreds of micrometers to achieve single cell culture and sequencing, that has been applied to single cell experiments [19, 20]



(Fig. 1e). Recently, various microfluidics platforms have emerged for single cell whole-genome, whole-transcriptome or epigenomics sequencing [21–23]. The advancement of microfluidics research has extended to separate nanoparticles, such as DNA isolation [24]. The advantages for microfluidics are the ability to input nanoliter-to-picoliter volumes of samples and to output accurate results with high resolution and sensitivity [19]. Additionally, microfluidics can provide parallel and timely analyses to make studies more efficient.

The main limitation of the above-mentioned methods is that the sample must be prepared in suspension and thus have lost the spatial location of the cells in the tissue. LCM overcomes this limitation and directly isolates single cells from tissue sections based on the cellular morphology (Fig. 1d). The targeted single cell can be stained with fluorescent or chromogenic antibodies for LCM [11]. The main drawbacks include lowthroughput, slicing the cells during the course of tissue sectioning, and UV damage to nuclei from the laser [12].

The increasing number of studies on rare single cells (<1%) poses a challenge on the current methods for single cell isolation. Now, several new technologies have been developed to cover the shortcomings of the above-mentioned methods in rare single cancer cell isolation, such as CellSearch (Johnson & Johnson), MagSweeper (Illumina Inc.), DEP-Array (Silicon Biosciences), CellCelector (Automated Lab Solutions), and nanofabricated filters (CellSieve) [25]. The FDA-cleared CellSearch system is the most-advanced commercially available technology using anti-EpCAM ferrofluid and has been applied to the monitoring of patients with metastatic prostate, breast, or colorectal cancer in hospitals [26, 27] (Fig. 1f). Mag-Sweeper is an automated immunomagnetic separation technology for enrichment of rare cells in mixed populations with high purity [28]. DEP-Array uses a microfluidics chip with dielectrophoretic cages to isolate single cells by charge [12, 29]. The CellCelector uses a robotic arm carrying a module to retrieve single cells from microwells for micromanipulation [30]. The CellSieve system can capture a variety of circulating tumor cells based on size discrimination instead of specific cell surface markers [31].

Single cell sequencing methods

The advance in the next-generation sequencing (NGS) technologies has promoted the rapid development of SCS, including single cell whole-genome sequencing, single cell whole-transcriptomic sequencing and single cell epigenomic sequencing [32–34].

Single cell whole-genome/whole-exome sequencing

The amount of DNA (approximately 6 pg) in a single cell is insufficient to meet the demand for next-generation sequencing, and thus whole genome amplification (WGA) was developed to amplify the DNA by the hundreds of thousands [25, 35]. Recently, the alternative WGA technologies have polymerase chain reaction (PCR), multiple displacement amplification (MDA), or a combination of displacement pre-amplification and PCR amplification [36]. In PCR-based WGA methods, degenerate oligonucleotide-primed PCR (DOP-PCR) is the most widely used method to amplify the entire genome [37, 38]. The principle of DOP-PCR is to perform a low annealing degenerate primer extension on the DNA template and then to amplify the tagged sequences at a high annealing temperature [37, 39] (Fig. 2a). The main shortcoming for DOP-PCR is the low physical coverage (approximately 10%) of a single cell genome, which is prone to miss single-nucleotide polymorphisms (SNPs), but DOP-PCR is the optimal method for copy-number variations (CNVs) or aneuploidy detection because of the uniformity of amplification during WGS [12, 40, 41]. The established MDA technologies are based on the discovery of two specific DNA polymerases: Phi29 DNA polymerase isolated from the Bacillus subtilis, and Bst DNA polymerase isolated from Bacillus stearothermophilus [42–44]. The mechanism of MDA is to yield continuous strand displacement DNA amplification using Phi29 or Bst DNA polymerase and random primers under isothermal conditions [45] (Fig. 2b). Phi29 DNA polymerase has been considered the optimal choice for MDA because it shows higher efficiency, higher fidelity and a lower error rate compared with Bst DNA polymerase which has no proofreading activity [10, 46]. The advantages of MDA are that it has high single cell genome or exome coverage (>90%), which can accurately measure mutations at basepair resolution and that it yields adequate quantities of product (average length >10 kb) from single cell genomic DNA in a short time with high fidelity [47]. However, the main drawbacks of MDA are uneven genome coverage, chimeric sequences, and contamination issues [15]. Multiple annealing and looping based amplification cycles (MALBAC) is the newly applied WGS method that combines quasi-linear strand displacement pre-amplification by a polymerase and exponential amplification by PCR [33] (Fig. 2c). Remarkably, MALBAC has low amplification bias and can achieve 93% genome coverage $\geq 1 \times$ and $25 \times$ mean sequencing depth in a single cell during WGS. Moreover, MALBAC shows higher efficiency to detect CNVs and SNPs for its improved uniformity and a lower allele dropout rate, compared with MDA [36]. The pitfall of MALBAC is the extremely high false positive rate for SNV detection because of the low fidelity



of Bst DNA polymerase, and the loss of underamplified regions of the genome [48]. Another improved method, nuc-seq or single nucleus exome sequencing (SNES), has

been developed to reduce the bias, this method combines flow-sorting of single G1/0 or G2/M nuclei, time-limited isothermal MDA, exome capture, and NGS [49, 50]. The main advantage of this method is the high detection efficiencies for single-nucleotide variations (SNVs) and indels benefiting from the high physical coverage (96%) of the single cell genome and exome [50].

Single cell whole-transcriptomic sequencing

It is estimated that the amount of total RNA or mRNA is only approximately 10 pg or approximately 0.1 pg, respectively, in a single cell [10]. Thus, WTA is a necessary step to construct a cDNA library for single cell transcriptomic sequencing. WTA has been applied to amplify RNA from a single cell to obtain the gene expression profile in microarray prior to the advent of NGS [51, 52]. Tang et al. [5] improved the single cell whole-transcriptome amplification method and used NGS instead of microarray to identify more genes and previously unknown splice junctions in single cells. The principle of this method is to use oligo dT primers conjugated to adapter sequences for reverse transcription and selective amplification of polyadenylated mRNA by PCR [5, 10, 53] (Fig. 3a). However, this method generates 3'-end skew bias during reverse-transcription to miss proximal splicing events [34]. Another WTA method, called SMART-seq, was developed to use Moloney murine leukemia virus (MMLV) reverse transcriptase to construct full-length cDNA libraries [54]. The two key features, template-switching and terminal transferase activity, of the enzyme can lead to adding a few non-templated C nucleotides to the cDNA and switching templates to transcribe the other strand [55] (Fig. 3b). The advantage of SMART-seq is to generate and amplify full-length cDNA from single cell RNA, leading to the detection of alternatively spliced exons [56]. The low sensitivity of SMART is the main shortcoming that was improved in a subsequently developed method, called SMART-seq2 [57]. Similarly, single cell tagged reverse-transcription (STRT) is based on the template-switching property of the reverse transcriptase to tag the 5'-end of cDNA [58]. This method enables researchers to compare gene expression profile differences without bias in multiple samples, but it yields a strong 5'-end bias. Cell expression by linear amplification and sequencing (Cel-seq) labels cDNA with a barcode and pools these cDNA from multiple single cells for in vitro transcription (IVT) to linearly amplify cDNA [59]. The CEL-Seq generates more reproducible, linear, and sensitive results in comparison with the PCR-based amplification method, but it yields a high 3'-end skew bias and loses the full spectrum of transcript variant detection [60]. Additionally, the unique molecular identifiers (UMIs) labeling technique is applied in single cell WTA to achieve quantitative single cell RNA sequencing [61] (Fig. 3c). This method obviously increases the accuracy in single cell whole-transcriptome sequencing by eliminating amplification bias. Recently, two new droplet-based RNA-seq technologies, named as Drop-seq and inDrop (indexing droplets), has been exploited to sequence in parallel thousands of single cells from a tissue [62, 63]. Each nanoliter-scale aqueous droplet is a tiny reaction chamber that contains a single cell, barcoded and UMI-labeled primers, and reaction buffer.





STAMPs (Single-cell Transcriptomes Attached to Microparticles) is PCR amplified for sequencing in Drop-seq, while Cell-seq is used by inDrop for sequencing. The advantages of these methods are to differentiate the cellof-origin of each mRNA which helps to develop single cell analysis in a complicated tissue, and the low technical noise that allows the analysis of thousands of different cells in parallel. The latest commercial platform—ChromiumTM System from 10× Genomics—integrates the Gemcode platform, which separates long pieces of DNA into droplets to create barcoded sequencing libraries [64, 65]. The high efficiency and flexible throughput of this method allows researchers to dynamically detect transcriptional profiles of single cells at scale [66].

Single cell epigenomic sequencing

Epigenomics is defined as a phenomenon that changes the final outcome of a chromosome without changing the underlying DNA sequence, including DNA methylation, histone modifications, chromatin packaging, small RNA, etc. [67]. Recently, single cell epigenomic sequencing studies are on the rise with the application of new single cell epigenomic sequencing methods. Single cell reduced representation bisulfite sequencing (scRRBS) integrates all the experimental steps before PCR amplification into a single-tube reaction to avoid unnecessary DNA loss and enables the detection of approximately 40% of the CpG sites in comparison with standard RRBS using thousands of cells [68, 69]. Another method, single cell bisulfite sequencing (scBS-seq), modifies the Post-Bisulfite Adaptor Tagging (PBAT) to perform bisulfite conversion prior to successive primer extension with oligo1 and oligo2 tagged random primers to generate amplicons [70]. The drawbacks of these methods are DNA loss, purification, and disability to discriminate 5mC from 5hmC for bisulfite conversion [34]. Moreover, single cell chromatin immunoprecipitation followed by sequencing (scChIPseq) combines microfluidics, DNA barcoding and sequencing to collect low coverage maps of the chromatin state at single cell resolution [71]. Additionally, other methods have been developed for single cell epigenomic sequencing, such as Hi-C methods that characterize chromatin interactions in the genome of single cells [9], and single chromatin molecule analysis at the nanoscale (SCAN) that extracts single chromatin with fluorescent antibodies through fluidic channels [72].

Application of single cell sequencing Cancer

Cancer heterogeneity comes from clone diversity and mutational evolution, which promote cancer cell survival and metastasis and confound the cancer diagnosis and treatment [10, 73]. A deep understanding of cancer

heterogeneity can contribute to therapeutic decisions. Thus, SCS as an ideal tool has been increasingly applied to reveal intratumor heterogeneity in various primary tumors, such as breast cancer [6, 49, 74], lung cancer [75], brain cancer [76], colon cancer [33, 77], bladder cancer [78], acute myeloid leukemia [79, 80] and melanoma [81].

Navin et al. [6] first applied single nucleus sequencing (SNS) to study tumor population structure and evolution in two breast cancer cases by analysis of genome copy number variation. The results found punctuated clonal evolution in tumors and confirmed that metastatic cells emerged from a main advanced expansion. Another study used nuc-seq to find a difference in the pattern of occurrence for aneuploid rearrangements and point mutations in breast tumor evolution [49]. Furthermore, Eirew et al. [74] studied the dynamics of genomic clones in breast cancer patient xenografts at single cell resolution to indicate that genomic aberrations can be reproducible determinants of evolutionary trajectories. Interestingly, a single cell whole genome sequencing study for colon cancer identified an abundant amount of mutated gene SLC12A5 at the individual level, which was sparse at the bulk cells level, and discovered that colon cancer had a biclonal origin [77]. However, another study using single cell exome sequencing to reveal the evolutionary process in bladder cancer indicated that 66 individual bladder cancer cells were derived from a single ancestral cell, but they developed into two distinct tumor cell subpopulations with subsequent evolution [78]. Single cell exome sequencing was also applied to elucidate the intratumoral genetic characteristics at a single cell level in a kidney cancer [8]. Additionally, the clonal evolution has been studied in hematopoietic tumors. Hughes et al. [79]. sequenced single cells from three myelodysplastic syndrome (MDS)-derived secondary acute myeloid leukemias (sAMLs) to confirm the clonal architecture that was identified from the bulk sample analysis. Single cell exome sequencing revealed a monoclonal evolution in a JAK2-negative myeloproliferative neoplasm and further identified candidate gene mutations for neoplasm progression [7].

In addition to single cell DNA and exome sequencing applications, single cell RNA-seq has also been widely used to study clonal evolution in different cancers. Single cell RNA-seq demonstrated subclonal heterogeneity in xenograft tumor cells and found a candidate tumor cell subpopulation associated with anti-cancer drug resistance in lung adenocarcinoma (ADC) [75]. Equally, intertumor and intratumor heterogeneity was elucidated in melanoma by single cell RNA-seq [81]. Tirosh et al. [76] used single cell RNA-seq to find a new subpopulation marked with stem or progenitor cell-like characteristics, which supported developmental programs in oligodendroglioma. Furthermore, another study identified several rare tumor-related genes in squamous cell carcinoma of urinary bladder using single cell RNA-seq [81].

Recently, an array of studies that used SCS to understand the necessary knowledge of different rare circulating cancer cells have been published. Ni et al. [82] combined MALBAC with NGS to elucidate the CNV patterns for metastasis of cancer in circulating tumor cells (CTCs) from lung ADC. Lohr et al. [83] sequenced entire exomes of CTCs from two prostate cancer patients and observed 73% CTC mutations that were identified in bulk tissue. The results were consistent with another study that compared CTCs with tissue using WGS in prostate cancer [84]. Additionally, one study built a new system to assess the genomic heterogeneity of single CTCs from metastatic breast cancer patients and found a cell subpopulation related to drug resistance [85]. However, a recent study indicated that a targeted mutation detection rate is approximately 27.7% in CTCs from pancreatic cancer compared with bulk cells but is negative in white blood cells [86]. Furthermore, single CTCs studies based on whole RNA-seq have also been published. Lohr et al. [87] classified multiple myeloma (MM) and quantitatively assessed prognosis related genes using single CTC RNA-seq. Another CTC RNA-seq study revealed that noncanonical Wnt signaling took part in antiandrogen resistance in prostate cancer [88].

Immunology

The heterogeneity of the immune system contributes to an efficient defense against a multitude of different pathogens [89]. The SCS technologies can help to define new classifications and differentiation trajectories of immune cells. CD4+ T helper cell, which play a key role in adaptive immune responses, are further investigated to unravel the heterogeneity of this celluar population at the single cell level. Mahata et al. [90] used single cell RNAseq to reveal the extensive heterogeneity within the Th2 population and to identify a new Th2 cell subpopulation marked with Cyp11a1 that modulated the steroid synthesis pathway. Additionally, functional and structural studies of the T cell receptor repertoire have also benefited from SCS approaches. Dash et al. [91] developed a new method to sequence the TCR α and TCR β chains from single CD8+ T cells. The data showed a characterized expression of TCR α for an influenza epitope. Another study combined TCRa and TCRB sequencing with phenotypic analysis to reveal the clonal structure of T cells at the single cell level [92]. In addition to T cells, Shalek et al. [93] examined the mouse bone-marrow-derived dendritic cells (BMDCs), which is an important antigenpresenting cell subpopulation in the adaptive immune system, using single cell RNA-seq. The results indicated that hundreds of immune related genes displayed bimodal expression in single cells. Further study demonstrated that paracrine signaling from early-induced dendritic cells plays an important role in inflammatory program [94]. Although the application of SCS to study the immune system is limited at present, SCS has shown robust potential for defining immune cell subpopulations and for examining gene expression variability, differential splicing and gene-regulatory networks [89].

Microbiology

The vast majority of microorganisms are uncultivated with current culturing methods which has extremely limited our ability to understand the biological diversity of the microbiome [95]. Recently, the difficulty in microbial research has been overcome with the development of SCS. The first study combined FACS with MDA to sequence single TM7 bacterial cells from the soil and gained a deep insight into the evolution and metabolism of these cells [96]. The member of TM7 phylum from the human mouth was also investigated in a similar method [97]. The subsequent study conducted the single cell genomic sequencing in other candidate uncultured phyla from different environments, including anoxic springderived OP11 [98], human microbiota-derived SR1 [99], hospital sink biofilm-derived TM6 [100] and hot spring sediments-derived OP9 [101]. In addition to sequencing the genome of various bacterial phyla, SCS can reveal the lifestyle and metabolism of uncultivated microorganisms, supporting the potential development of cultivation approaches and commercial applications. Marc et al. [102] sequenced over 70% of the genome of *Beggiatoa* from the surface of marine sediment and confirmed the chemolithoautotrophic physiology via investigating the pathway for sulfur oxidation, oxygen and nitrate respiration, and carbon metabolism. The findings supported the establishment of a particulate cultivating environment in which there was coexistence of different members of the microbial community and some missing supplementary materials [95]. In another study, Mason et al. [103] used MDA to sequence and assemble the single cell genome of Oceanospirillales from seawater after the Deepwater Horizon oil spill and identified enzymes that can degrade crude oil.

Prenatal diagnosis

The application of SCS to prenatal diagnosis, including pre-implantation genetic diagnosis (PGD) and noninvasive prenatal diagnosis (NIPD) has greatly increased the opportunities for healthy birth [32]. Recently, SCS has been widely used to detect aneuploidy and SNPs in prenatal diagnosis. Well et al. [104] used a rapid WDA protocol to diagnosis of aneuploidy in embryo biopsy with high accuracy and cost-efficiency. In another study, Fiorentino et al. [105] confirmed the validation and accuracy of a single cell NGS-based method for aneuploidy screening in single blastomeres. In the subsequent study, they compared this protocol with array comparative genomic hybridization (array-CGH) and demonstrated that a single cell NGS-based method improved the aneuploidy detection with high-throughput, automation and reliability [106]. Furthermore, Vera-Rodríguez et al. [107] used single cell NGS to investigate the distribution patterns of segmental aneuploidies in trophectoderm biopsy. The efficiency of NGS in the detection of pure and mosaic segmental aneuploidies equated with that of CGH. Lu et al. [108] used MALBAC to sequence 99 sperm from an Asian male to detect aneuploidy and single nucleotide polymorphisms. The same method was used to accurately detect aneuploidy and SNPs in a single oocyte [109]. Additionally, using NIPD as a safe and reliable method to identify affected fetuses before birth is becoming increasingly popular for clinical and research applications in combination with NGS technologies. Zhang et al. [110] used low-coverage massively parallel sequencing to detect CNVs in four single cells from peripheral blood. The sensitivity and specificity for CNVs and aneuploidies were 99.63 and 97.71%, respectively. Hua et al. [111] used WGA and Illumina MiSeq to sequence single fetal nucleated red blood cells from placental villi and to diagnose aneuploidy in 5 cases in 10 single cells.

Neurobiology

Defining neuronal heterogeneity is an enormous task in nervous research [112]. SCS has been increasingly used to understand neural cell diversity and to classify neurons. Many studies have been reported to use single cell RNA-seq to classify the type of neurons in various regions of the mouse nervous system, including monoaminergic systems, dorsal root ganglia, cortex and retina [112]. Okaty et al. [113] used single cell RNA-seq to distinguish serotonergic neurons from five hindbrain rhombomeres and confirmed the subpopulation grouped from the population-scale transcriptomes. Additionally, the subtype-specific behavioral function-related genes were identified at the single cell level. In another study, Zeisel et al. [114] used large-scale single cell RNA-seq to sequence the neuronal cells from the somatosensory cortex and hippocampal CA1 region, and they identified an interneuron and a postmitotic oligodendrocyte labeled with Pax6 and ltpr2, respectively. Similarly, Tasic et al. [115] defined 49 transcriptomic cortical cell types, including 23 GABAergic, 19 glutamatergic and 7 non-neuronal types based on single cell RNA sequencing. In humans, the structure and function of brain is more complex. Johnson et al. [116] combined FACS and single cell RNA-seq to detect the heterogeneity in evolution of human outer radial glia (ORG). In another study, single cell RNA-seq was used to identify the transcriptome diversity in adult and fetal brains. The results indicated that there was differential gene expression between adult and fetal neurons, and the gradient patterns of gene expression contributed to the understanding of the evolution of neurons in the brain [117]. The latest study used single nucleus RNA-seq to sequence the single neuron from six distinct regions of the human cerebral cortex and identified 16 neuronal subtypes with subtype-specific transcriptome profiles [118]. Additionally, single cell DNA-seq was used for CNV detection in brain diseases. Using single cell DNA-seq, McConnell et al. [119] demonstrated that there were abundant mosaic CNVs in human neurons, especially in hiPSCderived neurons. In another study, Cai et al. [120] used WGS to find a somatic CNV of chromosome 1q in more than 20% of neurons in a hemimegalencephaly (HMG) patient.

Conclusions

Biological heterogeneity must be considered in clinical and basic studies. With the advancement of next-generation sequencing, SCS, including single cell genomic, transcriptomic and epigenomic sequencing, has been become the major tool to unlock the secrets of biological diversity [41]. Recently, the application of SCS has been widespread in various research fields, such as cancer, immunology, microbiology, neurobiology and embryogenesis, and many successful commercial kits have emerged in the market [34]. Most exciting is the transformation of the use of SCS from bench to bedside. For example, SCS has been applied to the assessment of human embryos prior to implantation, non-invasive prenatal diagnosis and cancer diagnosis and prognosis [32]. However, there are still several shortcomings of SCS [10, 25]. It is hard to comprehensively and simultaneously sequence the genome, transcriptome and epigenome in a single cell. The high cost of SCS impedes its clinical application, and thus it will be a great challenge for researchers and engineers to provide highly efficient and low-cost technologies in SCS. Furthermore, in situ, real-time and in vivo sequencing and analysis of the DNA and RNA from single cells will be a new field that obtains deep insight into the spatial and temporal measurement of the molecular profiles of single cells. Lastly, new analysis models for the enormous data obtained from SCS should be built to unbiasedly mine the inherent properties of a single cell. Although still evolving, new SCS technologies have become powerful approaches for us to unravel the complexities of nature.

Authors' contributions

JW was involved in the conception, design and drafting of the manuscript. YLS was involved in the revision and final acceptance of the manuscript. Both authors read and approved the final manuscript.

Acknowledgements

This research was supported by the National Natural Science Foundation of China (81500026,81490533,81570028,81400018) and the Shanghai Science and Technology Committee (15DZ1930602). Dr. YL Song was supported by the State Key Basic Research Program (973) project (2015CB553404) and by the Doctoral Fund of the Ministry of Education of China (20130071110044).

Competing interests

The authors declare that they have no competing interests.

Received: 28 December 2016 Accepted: 11 February 2017 Published online: 20 February 2017

References

- 1. Novick A, Weiner M (1957) Enzyme induction as an all-or-none phenomenon. Proc Natl Acad Sci USA 43(7):553–566
- Coskun AF, Eser U, Islam S (2016) Cellular identity at the single-cell level. Mol Bio Syst 12(10):2965–2979
- 3. Bianconi E, Piovesan A, Facchin F et al (2013) An estimation of the number of cells in the human body. Ann Hum Biol 40(6):463–471
- Lindstrom S, Andersson-Svahn H (2010) Overview of single-cell analyses: microdevices and applications. Lab Chip 10(24):3363–3372
- Tang F, Barbacioru C, Wang Y et al (2009) mRNA-Seq whole-transcriptome analysis of a single cell. Nat Methods 6(5):377–382
- Navin N, Kendall J, Troge J et al (2011) Tumour evolution inferred by single-cell sequencing. Nature 472(7341):90–94
- Hou Y, Song L, Zhu P et al (2012) Single-cell exome sequencing and monoclonal evolution of a JAK2-negative myeloproliferative neoplasm. Cell 148(5):873–885
- Xu X, Hou Y, Yin X et al (2012) Single-cell exome sequencing reveals single-nucleotide mutation characteristics of a kidney tumor. Cell 148(5):886–895
- Nagano T, Lubling Y, Stevens TJ et al (2013) Single-cell Hi-C reveals cellto-cell variability in chromosome structure. Nature 502(7469):59–64
- 10. Wang Y, Navin NE (2015) Advances and applications of single-cell sequencing technologies. Mol Cell 58(4):598–609
- 11. Navin N, Hicks J (2011) Future medical applications of single-cell sequencing in cancer. Genome Med 3(5):31
- 12. Navin NE (2014) Cancer genomics: one cell at a time. Genome Biol 15(8):452
- Zhang D, Wang X (2015) A simple protocol for single lung cancer cell isolation-making the single cell based lung cancer research feasible for individual investigator. In Single cell sequencing and systems immunology. Springer, Berlin
- 14. Brehm-Stecher BF, Johnson EA (2004) Single-cell microbiology: tools, technologies, and applications. Microbiol Mol Biol Rev 68(3):538–559
- 15. Yilmaz S, Singh AK (2012) Single cell genome sequencing. Curr Opin Biotechnol 23(3):437–443
- von Boehmer L, Liu C, Ackerman S et al (2016) Sequencing and cloning of antigen-specific antibodies from mouse memory B cells. Nat Protoc 11(10):1908–1923
- 17. Allen LZ, Ishoey T, Novotny MA et al (2011) Single virus genomics: a new tool for virus discovery. PLoS ONE 6(3):e17722
- Wang J, Min Z, Jin M, et al (2015) Protocol for single cell isolation by flow cytometry. In Single cell sequencing and systems immunology. Springer, Berlin
- Whitesides GM (2006) The origins and the future of microfluidics. Nature 442(7101):368–373
- 20. Le Gac S, Nordhoff V (2016) Microfluidics for mammalian embryo culture and selection: where do we stand now? Mol Hum Reprod 27:61
- Streets AM, Zhang X, Cao C et al (2014) Microfluidic single-cell wholetranscriptome sequencing. Proc Natl Acad Sci USA 111(19):7048–7053

- 22. Han L, Zi X, Garmire LX et al (2014) Co-detection and sequencing of genes and transcripts from the same single cells facilitated by a micro-fluidics platform. Sci Rep 4:6485
- 23. Wu AR, Kawahara TL, Rapicavoli NA et al (2012) High throughput automated chromatin immunoprecipitation as a platform for drug screening and antibody validation. Lab Chip 12(12):2190–2198
- 24. Salafi T, Zeming KK, Zhang Y (2016) Advancements in microfluidics for nanoparticle separation. Lab Chip 17:11–33
- 25. Zhang X, Marjani SL, Hu Z et al (2016) Single-cell sequencing for precise cancer research: progress and prospects. Cancer Res 76(6):1305–1312
- 26. Swennenhuis JF, van Dalum G, Zeune LL et al (2016) Improving the cell search(R) system. Expert Rev Mol Diagn 16(12):1291–1305
- 27. Pantel K, Alix-Panabieres C, Riethdorf S (2009) Cancer micrometastases. Nat Rev Clin Oncol 6(6):339–351
- Talasaz AH, Powell AA, Huber DE et al (2009) Isolating highly enriched populations of circulating epithelial cells and other rare cells from blood using a magnetic sweeper device. Proc Natl Acad Sci USA 106(10):3970–3975
- 29. Altomare L, Borgatti M, Medoro G et al (2003) Levitation and movement of human tumor cells using a printed circuit board device based on software-controlled dielectrophoresis. Biotechnol Bioeng 82(4):474–479
- Choi JH, Ogunniyi AO, Du M et al (2010) Development and optimization of a process for automated recovery of single cells identified by microengraving. Biotechnol Prog 26(3):888–895
- Adams DL, Adams DK, Alpaugh RK et al (2016) Circulating cancerassociated macrophage-like cells differentiate malignant breast cancer and benign breast conditions. Cancer Epidemiol Biomark Prev 25(7):1037–1042
- Zhu W, Zhang XY, Marjani SL et al (2016) Next-generation molecular diagnosis: single-cell sequencing from bench to bedside. Cell Mol Life Sci 13:1
- Zong C, Lu S, Chapman AR et al (2012) Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. Science 338(6114):1622–1626
- 34. Liang J, Cai W, Sun Z (2014) Single-cell sequencing technologies: current and future. J Genet Genom 41(10):513–528
- 35. Van Loo P, Voet T (2014) Single cell analysis of cancer genomes. Curr Opin Genet Dev 24:82–91
- 36. Grun D, van Oudenaarden A (2015) Design and analysis of single-cell sequencing experiments. Cell 163(4):799–810
- Telenius H, Carter NP, Bebb CE et al (1992) Degenerate oligonucleotideprimed PCR: general amplification of target DNA by a single degenerate primer. Genomics 13(3):718–725
- Huang L, Ma F, Chapman A et al (2015) Single-cell whole-genome amplification and sequencing: methodology and applications. Annu Rev Genom Hum Genet 16:79–102
- Arneson N, Hughes S, Houlston R et al (2008) Whole-genome amplification by degenerate oligonucleotide primed PCR (DOP-PCR). CSH Protoc 2008;t4919
- Hou Y, Wu K, Shi X et al (2015) Comparison of variations detection between whole-genome amplification methods used in single-cell resequencing. Gigascience 4:37
- 41. Baslan T, Hicks J (2014) Single cell sequencing approaches for complex biological systems. Curr Opin Genet Dev 26:59–65
- 42. Zhang DY, Zhang W, Li X et al (2001) Detection of rare DNA targets by isothermal ramification amplification. Gene 274(1–2):209–216
- Aliotta JM, Pelletier JJ, Ware JL et al (1996) Thermostable Bst DNA polymerase I lacks a 3' -> 5' proofreading exonuclease activity. Genet Anal 12(5–6):185–195
- 44. Baner J, Nilsson M, Mendel-Hartvig M et al (1998) Signal amplification of padlock probes by rolling circle replication. Nucleic Acids Res 26(22):5073–5078
- 45. Dean FB, Nelson JR, Giesler TL et al (2001) Rapid amplification of plasmid and phage DNA using Phi 29 DNA polymerase and multiplyprimed rolling circle amplification. Genome Res 11(6):1095–1099
- Spits C, Le Caignec C, De Rycke M et al (2006) Optimization and evaluation of single-cell whole-genome multiple displacement amplification. Hum Mutat 27(5):496–503
- Dean FB, Hosono S, Fang L et al (2002) Comprehensive human genome amplification using multiple displacement amplification. Proc Natl Acad Sci USA 99(8):5261–5266

- Lasken RS (2013) Single-cell sequencing in its prime. Nat Biotechnol 31(3):211–212
- Wang Y, Waters J, Leung ML et al (2014) Clonal evolution in breast cancer revealed by single nucleus genome sequencing. Nature 512(7513):155–160
- 50. Leung ML, Wang Y, Waters J et al (2015) SNES: single nucleus exome sequencing. Genome Biol 16:55
- Kurimoto K, Yabuta Y, Ohinata Y et al (2006) An improved single-cell cDNA amplification method for efficient high-density oligonucleotide microarray analysis. Nucleic Acids Res 34(5):e42
- Iscove NN, Barbara M, Gu M et al (2002) Representation is faithfully preserved in global cDNA amplified exponentially from sub-picogram quantities of mRNA. Nat Biotechnol 20(9):940–943
- Tang F, Barbacioru C, Nordman E et al (2010) RNA-Seq analysis to capture the transcriptome landscape of a single cell. Nat Protoc 5(3):516–535
- Ramskold D, Luo S, Wang YC et al (2012) Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. Nat Biotechnol 30(8):777–782
- Zhu YY, Machleder EM, Chenchik A et al (2001) Reverse transcriptase template switching: a SMART approach for full-length cDNA library construction. Biotechniques 30(4):892–897
- Goetz JJ, Trimarchi JM (2012) Transcriptome sequencing of single cells with Smart-Seq. Nat Biotechnol 30(8):763–765
- 57. Picelli S, Bjorklund AK, Faridani OR et al (2013) Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Nat Methods 10(11):1096–1098
- Islam S, Kjallquist U, Moliner A et al (2011) Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. Genome Res 21(7):1160–1167
- Hashimshony T, Wagner F, Sher N et al (2012) CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. Cell Rep 2(3):666–673
- 60. Shapiro E, Biezuner T, Linnarsson S (2013) Single-cell sequencing-based technologies will revolutionize whole-organism science. Nat Rev Genet 14(9):618–630
- 61. Islam S, Zeisel A, Joost S et al (2014) Quantitative single-cell RNA-seq with unique molecular identifiers. Nat Methods 11(2):163–166
- Macosko EZ, Basu A, Satija R et al (2015) Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. Cell 161(5):1202–1214
- Klein AM, Mazutis L, Akartuna I et al (2015) Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. Cell 161(5):1187–1201
- 64. Eisenstein M (2015) Startups use short-read data to expand long-read sequencing market. Nat Biotechnol 33(5):433–435
- 65. Coombe L, Warren RL, Jackman SD et al (2016) Assembly of the complete Sitka Spruce chloroplast genome using 10× genomics' GemCode sequencing data. PLoS ONE 11(9):e163059
- 66. Zheng GX, Terry JM, Belgrader P et al (2017) Massively parallel digital transcriptional profiling of single cells. Nat Commun 8:14049
- 67. Goldberg AD, Allis CD, Bernstein E (2007) Epigenetics: a landscape takes shape. Cell 128(4):635–638
- Guo H, Zhu P, Guo F et al (2015) Profiling DNA methylome landscapes of mammalian cells with single-cell reduced-representation bisulfite sequencing. Nat Protoc 10(5):645–659
- 69. Guo H, Zhu P, Wu X et al (2013) Single-cell methylome landscapes of mouse embryonic stem cells and early embryos analyzed using reduced representation bisulfite sequencing. Genome Res 23(12):2126–2135
- Smallwood SA, Lee HJ, Angermueller C et al (2014) Single-cell genomewide bisulfite sequencing for assessing epigenetic heterogeneity. Nat Methods 11(8):817–820
- Rotem A, Ram O, Shoresh N et al (2015) Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. Nat Biotechnol 33(11):1165–1172
- 72. Cipriany BR, Zhao R, Murphy PJ et al (2010) Single molecule epigenetic analysis in a nanofluidic channel. Anal Chem 82(6):2480–2487
- 73. Park SY, Gonen M, Kim HJ et al (2010) Cellular and genetic diversity in the progression of in situ human breast carcinomas to an invasive phenotype. J Clin Invest 120(2):636–644

- Eirew P, Steif A, Khattra J et al (2015) Dynamics of genomic clones in breast cancer patient xenografts at single-cell resolution. Nature 518(7539):422–426
- 75. Kim KT, Lee HW, Lee HO et al (2015) Single-cell mRNA sequencing identifies subclonal heterogeneity in anti-cancer drug responses of lung adenocarcinoma cells. Genome Biol 16:127
- Tirosh I, Venteicher AS, Hebert C et al (2016) Single-cell RNA-seq supports a developmental hierarchy in human oligodendroglioma. Nature 539(7628):309–313
- 77. Yu C, Yu J, Yao X et al (2014) Discovery of biclonal origin and a novel oncogene SLC12A5 in colon cancer by single-cell sequencing. Cell Res 24(6):701–712
- Li Y, Xu X, Song L et al (2012) Single-cell sequencing analysis characterizes common and cell-lineage-specific mutations in a muscle-invasive bladder cancer. Gigascience 1(1):12
- Hughes AE, Magrini V, Demeter R et al (2014) Clonal architecture of secondary acute myeloid leukemia defined by single-cell sequencing. PLoS Genet 10(7):e1004462
- Paguirigan AL, Smith J, Meshinchi S et al (2015) Single-cell genotyping demonstrates complex clonal diversity in acute myeloid leukemia. Sci Transl Med 7(281):281r–282r
- Gerber T, Willscher E, Loeffler-Wirth H et al (2016) Mapping heterogeneity in patient-derived melanoma cultures by single-cell RNA-seq. Oncotarget 26:8
- Ni X, Zhuo M, Su Z et al (2013) Reproducible copy number variation patterns among single circulating tumor cells of lung cancer patients. Proc Natl Acad Sci USA 110(52):21083–21088
- Lohr JG, Adalsteinsson VA, Cibulskis K et al (2014) Whole-exome sequencing of circulating tumor cells provides a window into metastatic prostate cancer. Nat Biotechnol 32(5):479–484
- Jiang R, Lu YT, Ho H et al (2015) A comparison of isolated circulating tumor cells and tissue biopsies using whole-genome sequencing in prostate cancer. Oncotarget 6(42):44781–44793
- Polzer B, Medoro G, Pasch S et al (2014) Molecular profiling of single circulating tumor cells with diagnostic intention. EMBO Mol Med 6(11):1371–1386
- Court CM, Ankeny JS, Sho S et al (2016) Reality of single circulating tumor cell sequencing for molecular diagnostics in pancreatic cancer. J Mol Diagn 18(5):688–696
- Lohr JG, Kim S, Gould J et al (2016) Genetic interrogation of circulating multiple myeloma cells at single-cell resolution. Sci Transl Med 8(363):147r–363r
- Miyamoto DT, Zheng Y, Wittner BS et al (2015) RNA-Seq of single prostate CTCs implicates noncanonical Wnt signaling in antiandrogen resistance. Science 349(6254):1351–1356
- Proserpio V, Mahata B (2016) Single-cell technologies to study the immune system. Immunology 147(2):133–140
- 90. Mahata B, Zhang X, Kolodziejczyk AA et al (2014) Single-cell RNA sequencing reveals T helper cells synthesizing steroids de novo to contribute to immune homeostasis. Cell Rep 7(4):1130–1142
- Douek DC, Betts MR, Brenchley JM et al (2002) A novel approach to the analysis of specificity, clonality, and frequency of HIV-specific T cell responses reveals a potential mechanism for control of viral escape. J Immunol 168(6):3099–3104
- 92. Han A, Glanville J, Hansmann L et al (2014) Linking T-cell receptor sequence to functional phenotype at the single-cell level. Nat Biotechnol 32(7):684–692
- 93. Shalek AK, Satija R, Adiconis X et al (2013) Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. Nature 498(7453):236–240
- 94. Shalek AK, Satija R, Shuga J et al (2014) Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. Nature 510(7505):363–369
- Lasken RS, McLean JS (2014) Recent advances in genomic DNA sequencing of microbial species from single cells. Nat Rev Genet 15(9):577–584
- 96. Podar M, Abulencia CB, Walcher M et al (2007) Targeted access to the genomes of low-abundance organisms in complex microbial communities. Appl Environ Microbiol 73(10):3205–3214
- 97. Marcy Y, Ouverney C, Bik EM et al (2007) Dissecting biological "dark matter" with single-cell genetic analysis of rare and uncultivated

TM7 microbes from the human mouth. Proc Natl Acad Sci USA 104(29):11889–11894

- Youssef NH, Blainey PC, Quake SR et al (2011) Partial genome assembly for a candidate division OP11 single cell from an anoxic spring (Zodletone Spring, Oklahoma). Appl Environ Microbiol 77(21):7804–7814
- Campbell JH, O'Donoghue P, Campbell AG et al (2013) UGA is an additional glycine codon in uncultured SR1 bacteria from the human microbiota. Proc Natl Acad Sci USA 110(14):5540–5545
- McLean JS, Lombardo MJ, Badger JH et al (2013) Candidate phylum TM6 genome recovered from a hospital sink biofilm provides genomic insights into this uncultivated phylum. Proc Natl Acad Sci USA 110(26):E2390–E2399
- 101. Dodsworth JA, Blainey PC, Murugapiran SK et al (1854) Single-cell and metagenomic analyses indicate a fermentative and saccharolytic lifestyle for members of the OP9 lineage. Nat Commun 2013:4
- Mussmann M, Hu FZ, Richter M et al (2007) Insights into the genome of large sulfur bacteria revealed by analysis of single filaments. PLoS Biol 5(9):e230
- Mason OU, Hazen TC, Borglin S et al (2012) Metagenome, metatranscriptome and single-cell sequencing reveal microbial response to deepwater horizon oil spill. ISME J 6(9):1715–1727
- 104. Wells D, Kaur K, Grifo J et al (2014) Clinical utilisation of a rapid low-pass whole genome sequencing technique for the diagnosis of aneuploidy in human embryos prior to implantation. J Med Genet 51(8):553–562
- 105. Fiorentino F, Biricik A, Bono S et al (2014) Development and validation of a next-generation sequencing-based protocol for 24-chromosome aneuploidy screening of embryos. Fertil Steril 101(5):1375–1382
- Fiorentino F, Bono S, Biricik A et al (2014) Application of next-generation sequencing technology for comprehensive aneuploidy screening of blastocysts in clinical preimplantation genetic screening cycles. Hum Reprod 29(12):2802–2813
- 107. Vera-Rodriguez M, Michel CE, Mercader A et al (2016) Distribution patterns of segmental aneuploidies in human blastocysts identified by next-generation sequencing. Fertil Steril 105(4):1047–1055
- Lu S, Zong C, Fan W et al (2012) Probing meiotic recombination and aneuploidy of single sperm cells by whole-genome sequencing. Science 338(6114):1627–1630

- Hou Y, Fan W, Yan L et al (2013) Genome analyses of single human oocvtes. Cell 155(7):1492–1506
- Zhang C, Zhang C, Chen S et al (2013) A single cell level based method for copy number variation analysis by low coverage massively parallel sequencing. PLoS ONE 8(1):e54236
- 111. Hua R, Barrett AN, Tan TZ et al (2015) Detection of aneuploidy from single fetal nucleated red blood cells using whole genome sequencing. Prenat Diagn 35(7):637–644
- 112. Poulin JF, Tasic B, Hjerling-Leffler J et al (2016) Disentangling neural cell diversity using single-cell transcriptomics. Nat Neurosci 19(9):1131–1141
- 113. Okaty BW, Freret ME, Rood BD et al (2015) Multi-scale molecular deconstruction of the serotonin neuron system. Neuron 88(4):774–791
- Zeisel A, Munoz-Manchado AB, Codeluppi S et al (2015) Brain structure. Cell types in the mouse cortex and hippocampus revealed by singlecell RNA-seq. Science 347(6226):1138–1142
- Tasic B, Menon V, Nguyen TN et al (2016) Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. Nat Neurosci 19(2):335–346
- 116. Johnson MB, Walsh CA (2016) Cerebral cortical neuron diversity and development at single-cell resolution. Curr Opin Neurobiol 42:9–16
- 117. Darmanis S, Sloan SA, Zhang Y et al (2015) A survey of human brain transcriptome diversity at the single cell level. Proc Natl Acad Sci USA 112(23):7285–7290
- Lake BB, Ai R, Kaeser GE et al (2016) Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. Science 352(6293):1586–1590
- 119. McConnell MJ, Lindberg MR, Brennand KJ et al (2013) Mosaic copy number variation in human neurons. Science 342(6158):632–637
- 120. Cai X, Evrony GD, Lehmann HS et al (2014) Single-cell, genome-wide sequencing identifies clonal somatic copy-number variation in the human brain. Cell Rep 8(5):1280–1289

Submit your manuscript to a SpringerOpen[™] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- ► High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at > springeropen.com