REVIEW

Single-cell genomics: An overview

Qichao WANG[1][*], Xianmin ZHU[1,2][*], Yun FENG[1], Zhigang XUE[1], Guoping FAN (✉)[2,3,4]

[1] *Stem Cell Research Center, Department of Regenerative Medicine, Tongji University School of Medicine, Tongji University, Shanghai 200092, China*
[2] *Advanced Institute of Translational Medicine, Tongji University, Shanghai 200092, China*
[3] *School of Life Sciences and Technology, Tongji University, Shanghai 200092, China*
[4] *Department of Human Genetics, UCLA, Los Angeles, CA 90095, USA*

© Higher Education Press and Springer-Verlag Berlin Heidelberg 2013

Abstract **The newly developed next-generation sequencing platforms, in combination with genome-scale amplification methods, provide a powerful tool to study genomics from a single cell. This mini-review summarizes the technologies of single cell genomics and their applications in several areas of biomedical research including stem cells, cancer biology and reproductive medicine. Particularly, it highlights recent advances in single cell exome sequencing, RNA-seq, and genome sequencing. The application of these powerful techniques will shed new light on the fundamental principles of gene transcription and genome organization at single-cell level and improve our understanding of cellular heterogeneity and diversity in multicellular organisms.**

Keywords single-cell genomics, next-generation sequencing, RNA-seq, single-nucleotide variation, copy-number variation, DNA methylation

Introduction

Single-cell genomics has become a prosperous research field because of its numerous advantages over conventional methods. Traditional sequencing technologies typically involve heterogeneous populations, thus masking the differences among individual cells (Geschwind and Konopka, 2009). In the field of cancer biology, for example, the differences among individual cells in a tumor tissue have already been confirmed decades ago (Farago and Chester, 1961), in particular with regard to genetic heterogeneity (Torres et al., 2007; Park et al., 2010). According to the cancer stem cell (CSC) theory (Kleinsmith and Pierce, 1964; Hamburger and Salmon, 1977; Reya et al., 2001), only a small fraction of tumor stem cells lead to recurrence of the tumor after chemotherapy and/or radiotherapy (Chaffer and Weinberg, 2011). Study of this small fraction of cells by single-cell analysis would be of great clinical importance. Several recent studies have been focused on the individual

differences among cells, including single nucleotide variant (SNV) (Clark et al., 2010), gene copy number variation (CNV) (Zong et al., 2012) and expression profiling (Tan et al., 2013). In addition to tumor research, single cell genomics becomes a valuable tool to study the biological differences between the neighboring cells during embryonic development (Xue et al., 2013; Zandi et al., 2012), cell proliferation (Lecault et al., 2011), differentiation (Galán et al., 2010), reprogramming (Buganim et al., 2012), and aging (Frontera et al., 2012). In short, single-cell genomics has the capability to uncover enormous cellular heterogeneity and yield novel biological mechanisms at a single cell level.

The amplification of nucleic acid is an essential step for single-cell genomics because DNA has to be amplified well over a thousand fold from a single cell that contains limited amount of genetic materials, e.g., a normal human somatic cell contains approximately 6 pg of genomic DNA and ~100 pg of total RNA of which ~1 pg is mRNA. Conventional DNA microarray and deep sequencing technologies generally require micrograms of DNA/RNA as a starting material. A significant problem is that the commonly used PCR amplification introduces high error rate due to low fidelity of DNA polymerase, so that the altered ratio of amplified PCR products generates biased sequencing results (Aird et al., 2011). To address the problem with respect to single-

[*]These authors contributed equally to this paper.

cell genomics, great progress has been made recently including isolation and amplification of DNA/RNA from a single or small number of cells followed by whole genome sequencing, RNA sequencing as well as epigenetic profiling. In this mini-review, we introduce the technique and applications of single-cell genomics in biomedical research. We also discuss how single cell analysis would revolutionize the field of cellular heterogeneity.

Isolation of single cell

The methods of single-cell separation and nucleotide extraction vary based on the characteristics of different cell types. Some cells, e.g., sperm and blood are in liquid form and are relatively easy to be isolated and collected. Traditional enzymatic digestion methods were developed to isolate cells of interest within compact tissues. However, these methods tend to lose the position information of a certain cell, which is important in research areas such as oncology. To overcome this problem, laser capture micro-dissection (LCM) was introduced since it does not alter or damage the morphological structure of the cell of interest as well as its surrounding cells. LCM cuts out and separates the cells from the adjacent tissues without contact using laser beam coupled into a microscope. Non-contact cutting ensures that the separated cells are free from instrument contamination and the accuracy reaches 1 μm (Simone et al., 1998). LCM is an ideal method to collect selected cells for DNA, RNA and protein analyses. It can be performed on many different tissue samples including cell cultures, solid tissues, and frozen and paraffin embedded archival tissues. For example, Guo et al. developed a protocol of LCM-mediated single nucleus purification for Class III gastric cancer tissue sample. The successful detection of a point mutation on the *PSCA* gene by PCR restriction fragment-length polymorphism (PCR-RFLP) demonstrated that this method can be effectively coupled with the downstream single cell polymorphism (SNP) analysis (Guo et al., 2012).

In addition to LCM, a variety of manual or mechanical cutting techniques at the level of microstructure is widely used. In 1998, Lee et al. isolated tumor cells from adjacent tissue in paraffin embedded archival tissues without contamination, by developing a precise microdissection device that contains a disposable needle, a conventional microscope, and a micromanipulator (Lee et al., 1998). Many other methods have been particularly designed to adapt to specific tissues. When studying colon cancer, Dalerba et al. separated single tumor cells by flow cytometry using fluorescence-activated cell sorting (FACS) after immuno-staining against the cell surface marker (Dalerba et al., 2011). In analysis of cell-lineage commitment, blastomeres during human early embryonic development were separated by micromanipulation (Galán et al., 2010). With regard to single microbial cell separation, the conventional methods such as micro fluidics

and serial dilution are used (Schoenborn et al., 2004). Micro-fluidics method can be used for high-throughput isolation of single cell analysis as demonstrated by the recent Fluidigm C1 machine (Fluidigm Corporation) (Wu et al., 2013), which utilizes disposable 96-chamber collection disc to integrate all the upstream procedures for gene expression profiling, such as cell capture, wash, verification, cell lysis, reverse transcription and amplification. Fluidigm C1 is based on the patented NanoFlex valves. These valves are fabricated by sealing different layers of elastomeric rubber through Multi-layer Soft Lithography (MSL) (Unger et al., 2000). Pressurized gas in one layer results in rubber deflection on the other layer so that to control on and off of the fluid flow in the disc and thus enable separation of single cells.

New sequencing technology

The first generation of sequencing technology includes the Maxam-Gilbert (chemical cleavage) method and the Sanger (dideoxy) method, of which the Sanger method dominated the laboratories in the late 20th century for its convenience on benchtop. Since the new millennium, next-generation sequencing (NGS) technology, also known as deep sequencing, has been developed achieving millions of sequencing reads in parallel. NGS includes second-generation sequencing and now third-generation sequencing, indicating the evolution of NGS technology. Second-generation sequencing is referred to as sequencing by combination of optical detection and nucleoside triphosphates wash (Niedringhaus et al., 2011). It employs sequencing-by-synthesis or sequencing-by-ligation technology, and branches into three major commercial platforms i.e., Illumina HiSeq2000, Roche 454 and ABI SOLiD platform (Mardis, 2008; Morozova and Marra, 2008; Shendure and Ji, 2008; Metzker, 2010). Despite using different instruments and specific chemical reactions, all three platforms have the capability to achieve millions of sequencing reads in parallel efficiently. These three platforms have claimed their individual merit, e.g., HiSeq2000 is most cost-effective; 454 is capable of sequencing the longest fragments; SOLiD has claimed high precision. In practice, Illumina/HiSeq2000 is currently the most widely used platform among the three.

Third-generation sequencing introduces state-of-the-art technologies and aims to sequencing human genome under $1000 (Niedringhaus et al., 2011). Ion Torrent Personal Genome Machine (PGM) is a third-generation sequencing technology that sequences short-reads (about 100 bp in length) in a rapid and economical way (Rothberg et al., 2011). The principle of this method is to measure, as a pH meter, the proton release during dNTP incorporation in DNA synthesis. In DNA library construction step, DNA fragments with specific adapters are amplified onto the beads. Then DNA templates on the beads, along with sequencing primers and DNA polymerase are loaded into sensor wells to be

sequenced. In sequencing step, dNTP is provided sequentially so that the proton release from each nucleotide-incorporation reaction is recorded by an ion sensor, with the signal strength proportional to total number of incorporated nucleotides. The recent technical breakthrough, single-molecule real-time DNA sequencing (also known as SMRT), designed by Pacific Biosciences Company is also considered as a third-generation sequencing technology. SMRT is performed on a single DNA molecule and takes full advantage of the high catalytic ability and sustained synthesis ability of the DNA polymerase. Four types of nucleoside triphosphate are labeled with different fluorescent dyes on the γ-phosphate, so the reading signals are generated when certain nucleoside triphosphate is incorporated along the template. The fluorescent label is then cleaved off by the polymerase during DNA extension. SMRT has many important advantages. For example, it has low error rate and high sequencing speed and can sequence long DNA molecule. Besides, by analyzing the difference of DNA synthesis interval, it can determine the secondary structure of DNA molecule and also certain modifications such as DNA methylation (Feng et al., 2013). Overall, single-molecule sequencing technology such as SMRT promises great application prospects by avoiding the bias from PCR amplification (Eid et al., 2009). However, it still remains a major hurdle to reach a high sequencing speed while optimizing accuracy and reducing the cost for all the third-generation sequencing platforms.

DNA amplification

DNA library preparation for the major sequencing platforms shares similar procedures, including DNA fragmentation, end repair, 3′-end adenylation, ligation with specific adapters, ligation mediated-PCR, purification, quality control, etc. Single-cell genomics requires a specific step of whole-genome amplification (WGA) in DNA library preparation, which generates a sufficient amount of DNA from an individual cell for sequencing (Hanson and Ballantyne, 2005). PCR-based WGA is a popular approach nowadays. However, this method could generate amplification bias, leading to low genome coverage and skewed sequencing results.

Toward improving the capability of the traditional WGA, multiple displacement amplification (MDA) was newly developed by using strand-displacement Φ29 polymerase (Dean et al., 2002). The reaction can be carried out at a moderate isothermal condition near room temperature, therefore does not require a thermocycler. The hexamer primers are usually thiophosphate-modified at the 3′ end to convey resistance to the 3′–5′ exonuclease activity of Φ29 DNA polymerase (Luthra and Medeiros, 2004; Hutchison et al., 2005). During MDA, hyper-branched DNA structures are formed when primers continuously bind the newly formed DNA strands synthesized by strand displacement (Paez et al.,

2004). MDA has been widely used in analysis of SNV, CNV, sequence tagged sites (STS) and even methylation correlation detection (Tzvetkov et al., 2005; Corneveaux et al., 2007; Hughes and Jones, 2007; Ling et al., 2012). Although it has many advantages over the traditional PCR-based WGA, MDA still has insurmountable problems such as significant non-specific amplification and sequence deviation.

A new approach, multiple annealing and looping-based amplification cycles (MALBAC) was recently reported to be a potentially superior method over MDA (Lu et al., 2012; Zong et al., 2012). The amplification is based on a pool of random primers which share common 27-nucleotide sequences but with 8 variable nucleotides. This type of primers can be evenly distributed on the template, and ultimately results in the quasi-linear pre-amplification instead of exponential amplification by PCR. MALBAC achieves high genome coverage by quasi-linear pre-amplification of the genetic material from an individual cell. However, it remains to be verified whether amplification bias is significantly reduced compared to current WGA and MDA methods.

Single cell analysis of SNV and point mutation

Using the MALBAC method, Zong et al. detected CNVs of an individual cancer cell and identified SNVs within three kindred cells, which showed higher resolution than MDA (Zong et al., 2012). They also mapped recombination events in 99 sperm from an Asian male by using MALBAC (Lu et al., 2012). With SNPs as markers to map the positions of crossovers in each sperm, they found that the reduction of recombination rate near transcription start sites (TSSs) results from the intrinsic mechanism during meiosis rather than natural selection (Lu et al., 2012). The sequencing results of the sperm also showed that in 5% of the sperm aneuploidy causes serious congenital birth defects (Lu et al., 2012). Similarly, Wang et al. carried out microfluidic single-sperm WGA and studied the recombination hotspots in 91 human sperms (Wang et al., 2012). They found that there are on average 30 mutations per million base pairs between generations, which is slightly higher than reported previously (McVean et al., 2004).

The fast developing single-cell genomics has already provided guidance for the early diagnosis and personalized treatment for cancer. Since genomic variation is assumed to be the fundamental cause of a tumor, an accurate and in-depth analysis of tumor cell by single-cell genomics will allow researchers to locate cancer gene mutations and to identify the genotypes of certain cancer cell type. SNP, as the most common genomic variation, is considered as risk marker correlated with cancer. Single-cell genomics provides much more detailed information than conventional PCR and microarray data. By using single-nucleus sequencing, Navin

et al. analyzed 100 cells from breast cancer, categorizing three clonal sub-populations, and tracing the cell type that seeded a liver metastasis (Navin et al., 2011). In addition, exome sequencing can access information on single-nucleotide mutation at single-cell level. By single-cell exome-sequencing, for example, Xu et al. compared the genetic information of the individual tumor cells with their surrounding normal cells in the clear cell renal cell carcinoma (ccRCC) patient (Xu et al., 2012). They were able to pinpoint several potential causative genes for ccRCC and revealed different mutation mechanisms in these genes (Xu et al., 2012). Meanwhile, another exome-sequencing study on essential thrombocythemia (ET) was performed by Hou et al. (Hou et al., 2012), in which they identified the origin of the 58 single cells from the patients and characterized mutations on *SESN2* and *NTRK1* with combination of MDA, Hiseq2000 sequencing and SNP calling. Recently, Voet et al. developed single-cell paired-end sequencing method and analyzed the genomic variations in a variety of samples, including breast cancer cells, human zygote and blastomeres (Voet et al., 2013). In accordance with the previous findings, their data supported that MDA followed by sequencing is still a robust tool for SNP calling although bias is inevitably introduced by the DNA amplification step. Another important application for SNP analysis by single-cell genomics is to facilitate distinguishing transcripts that are paternally and maternally inherited, i.e. halotyping. Fan et al. developed a method called direct deterministic phasing (DDP) to separate single chromosome at metaphase from a single cell in a microfluidic disc (Fan et al., 2011). The authors performed the whole-genome halotyping in four individuals, implying its clinical application in autoimmune disease and transplantation by determination of human leukocyte antigen (HLA) haplotypes.

Transcriptome analysis

Individual cells have their unique transcriptomes that reflect expression of a subset of genes, although they share the same genotype. Theoretically, homogeneous cells with the same DNA may have very different transcriptomes even if they appear to be identical physiologically and morphologically (Li and Clevers, 2010). Cells and their neighboring cells have been shown to contain differentiated RNA transcripts, let alone the cells from different tissues, organs and systems.

The newly developed techniques, e.g., single-cell cDNA microarray and RNA-seq, attempted to detect both known and unknown transcripts with relatively precise quantitation in a single cell. The cell fate commitment and embryo genome activation was analyzed by single-cell cDNA microarray using 5- to 8-cell human embryos (Galán et al., 2010). By coupling SOLiD sequencing system with improved amplification methods, mRNA-seq became a convenient platform to perform transcriptome analysis especially in the field of early embryonic development (Tang et al., 2009). However, it is well known that many RNA species, such as mRNA, rRNA,

tRNA and small nuclear RNA, undergo multiple variants due to alternative RNA processing (He, 2010). Thus, new single-cell transcriptome methods have been developed to improve the efficiency of detecting different transcript variants which are derived from different alleles and/or different RNA processing (Ramsköld et al., 2012). Single-cell tagged reverse transcription (STRT) (Islam et al., 2011) and switching mechanism at the 5′ end of the RNA template (SMART) (Ramsköld et al., 2012) take advantage of the "template switching" property of Moloney murine leukemia virus (MMLV) reverse transcriptase. After priming with oligo (dT), MMLV reverse transcriptase adds a tail of cytosines to the end of the first-strand cDNA. A helper oligonucleotide with guanine motif can anneal with the cytosine tail. MMLV reverse transcriptase then switches template and copies the helper oligo, which results in arbitrary ends on both ends allowing downstream DNA amplification that covers full length transcripts. An alternative approach is cell expression by linear amplification and sequencing (CEL-seq) (Hashimshony et al., 2012), which theoretically aims to reduce the bias in the amplification methods such as PCR by inserting an *in vitro* transcription step.

To generate single-cell transcriptome after single-cell isolation, the protocol by Tang et al. is to convert mRNAs into first-stand cDNA using a universal primer with oligo (dT). After reverse transcription, a poly(A) tail is added to the 3′ ends of the first-stand cDNA so that the second-strand cDNA is amplified by another universal primer with oligo (dT). Then the cDNA pool is amplified by PCR with both universal primers prior to sequencing and further analysis (Tang et al., 2011). However, the amplification steps often generate biased representation of mRNA, and lead to discriminated sequencing results. To solve this problem, Pan et al. developed two RNA amplification procedures to obtain full-length coverage of expressed mRNA from low quantities of cells or even single cell. These two methods demonstrated neither 3′ biased nor variable transcript representation in the previous methods (Pan et al., 2013). Briefly, one method utilized Φ29 DNA polymerase-based mRNA transcriptome amplification (PMA) to potentially capture all the end sequences of cDNA by intramolecular ligation, which is adapted from the previous whole DNA-pool amplification procedure (Pan et al., 2008). Meanwhile, the authors also developed semi-random primed PCR-based mRNA transcriptome amplification[SRP-transcriptome mRNA amplification (SMA)] by adapting from nano-ChIP-seq. The characteristic signature of SMA is to use the hairpin-containing semi-random primers for the initial library construction following cDNA reverse transcription. Both PMA and SMA were demonstrated to have high applicability in cDNA sequencing coupled with high-throughput sequencing technology. Although both methods are able to retrieve full-length coverage, SMA is considered to be superior to PMA in many respects. To date, a recent method Quartz-seq optimized the conditions for PCR based single-cell RNA-seq

by eliminating byproduct contamination and using a robust polymerase for one-tube reaction (Sasagawa et al., 2013).

Single-cell genomics holds great promise for analysis of gene expression variation at single-cell resolution in many research fields, such as immune response and preimplantation embryo development. By exploring the response of bone-marrow-derived dendritic cells to lipopolysaccharide (LPS), Shalek et al. recently elucidated the extensive bimodal expression of immune genes and differentiated mRNA splicing patterns via single-cell RNA-seq (Shalek et al., 2013). Our group recently performed a comprehensive analysis of the transcriptome during human and mouse preimplantation by taking advantage of single-cell RNA-seq (Xue et al., 2013), This work showed that the genetic programs in mammalian preimplantation development are evolutionarily conserved between the species, with the exception of developmental timing. By analyzing the allelic patterns of paternal and maternal gene transcripts, our analysis also revealed single nucleotide variants as well as synonymous and deleterious variants in a subset of protein coding genes. Our study paved the way for using single-cell RNA-seq in preimplantation genetic diagnosis and IVF technology.

DNA methylation

DNA methylation is referred to as the addition of a methyl group to the cytosine of DNA in vertebrates, which plays an important role in many cellular events. The base 5-hydroxymethycytosine (5hmC) and 5-methylcytosine (5mC) are both important modifications involved in many cellular functions especially during early development (Tahiliani et al., 2009; Bhutani et al., 2011). Site-specific methylation plays an important regulatory role during human embryonic development (Quenneville et al., 2012; Zhao et al., 2013). It is well known that cancer cells possess genome-wide hypo-methylation and site-specific hypermethylation (Esteller, 2007; Jones and Baylin, 2007). Analysis of these epigenetic markers including site-specific methylation will enhance our understanding of cellular regulation mechanisms. Song et al. developed a novel method that selectively labels 5-hmC with the biotin tag followed by SMRT. Based on the incorporation time change in modified nucleotides, different DNA methylation structures were directly detected by monitoring the interpulse duration (IPD) (Song et al., 2012). This method highlighted the bright prospects of SMRT as a third-generation sequencing technology, not only in genomic analysis but also in epigenetic analysis.

Drug candidate screening and biomarker discovery

As stated previously in the review, cellular heterogeneity generates "noise" information when studying a cell popula-

tion. Before the era of single-cell genomics, all our results were obtained from mixed cell populations, which did not appreciate the unique property of an individual cell. Nowadays single-cell genomics has made the personalized medicine coming true. First of all, single-cell genomics requires very small sample size and thus is suitable for detection of rare cell-types and cells with very limited amount from the patients. Second, genomic DNA sequencing, transcriptome analysis and DNA modification analysis allow evaluating the efficacy of the drug candidates at different levels in the targeted cell compared to its neighbors. Last, the high-throughput advances allow single-cell genomics to retrieve more comprehensive information and to provide new biomarkers of human diseases. Pioneer investment on single-cell genomics has already been made in research on different cancer types (Dalerba et al., 2011; Navin et al., 2011; Hou et al., 2012; Xu et al., 2012; Heitzer et al., 2013; Voet et al., 2013) and preimplantation diagnosis (Xue et al., 2013; Yan et al., 2013). Its application has also been put on the agenda for other areas, such as neuronal diseases (Iourov et al., 2012). Interestingly, single-cell genomics is promptly adopted into studying uncultured pathogenic microbes, including bacteria (McLean et al., 2013) and virus (Allen et al., 2011; McWilliam Leitch and McLauchlan, 2013).

In summary, the single-cell analysis has been and will be widely used in all aspects of functional genomics, such as the profiling of mRNA and non-coding RNA expression, DNA methylation and histone modification patterns. It is possible to reveal the cellular and molecular differences between each individual cell. To date, single-cell analysis has already shown its promising potential in the translational medicine fields, such as cancer diagnosis and in vitro fertilization (IVF). The rapid development of technology will permit single cell genomic analysis with higher speed, wider coverage, higher accuracy, more convenience, and lower cost. It will certainly become a powerful tool for basic biomedical research, clinical diagnosis, and drug discovery in the near future.

Acknowledgements

Compliance with ethics guidelines

Qichao WANG, Xianmin ZHU, Yun FENG, Zhigang XUE, Guoping FAN declare that they have no conflict of interest.

this manuscript is a review article and does not involve a research protocol requiring approval by the relevant institutional review board or ethics committee.

References

Aird D, Ross M G, Chen W S, Danielsson M, Fennell T, Russ C, Jaffe D B, Nusbaum C, Gnirke A (2011). Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries. Genome Biol, 12 (2): R18

Allen L Z, Ishoey T, Novotny M A, McLean J S, Lasken R S, Williamson S J (2011). Single virus genomics: a new tool for virus discovery. PLoS ONE, 6(3): e17722

Bhutani N, Burns D M, Blau H M (2011). DNA demethylation dynamics. Cell, 146(6): 866–872

Buganim Y, Faddah D A, Cheng A W, Itskovich E, Markoulaki S, Ganz K, Klemm S L, van Oudenaarden A, Jaenisch R (2012). Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. Cell, 150(6): 1209–1222

Chaffer C L, Weinberg R A (2011). A perspective on cancer cell metastasis. Science, 331(6024): 1559–1564

Clark M J, Homer N, O'Connor B D, Chen Z, Eskin A, Lee H, Merriman B, Nelson S F (2010). U87MG decoded: the genomic sequence of a cytogenetically aberrant human cancer cell line. PLoS Genet, 6(1): e1000832

Corneveaux J J, Kruer M C, Hu-Lince D, Ramsey K E, Zismann V L, Stephan D A, Craig D W, Huentelman M J (2007). SNP-based chromosomal copy number ascertainment following multiple displacement whole-genome amplification. Biotechniques, 42(1): 77–83

Dalerba P, Kalisky T, Sahoo D, Rajendran P S, Rothenberg M E, Leyrat A A, Sim S, Okamoto J, Johnston D M, Qian D, Zabala M, Bueno J, Neff N F, Wang J, Shelton A A, Visser B, Hisamori S, Shimono Y, van de Wetering M, Clevers H, Clarke M F, Quake S R (2011). Single-cell dissection of transcriptional heterogeneity in human colon tumors. Nat Biotechnol, 29(12): 1120–1127

Dean F B, Hosono S, Fang L, Wu X, Faruqi A F, Bray-Ward P, Sun Z, Zong Q, Du Y, Du J, Driscoll M, Song W, Kingsmore S F, Egholm M, Lasken R S (2002). Comprehensive human genome amplification using multiple displacement amplification. Proc Natl Acad Sci USA, 99(8): 5261–5266

Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, Peluso P, Rank D, Baybayan P, Bettman B, Bibillo A, Bjornson K, Chaudhuri B, Christians F, Cicero R, Clark S, Dalal R, Dewinter A, Dixon J, Foquet M, Gaertner A, Hardenbol P, Heiner C, Hester K, Holden D, Kearns G, Kong X, Kuse R, Lacroix Y, Lin S, Lundquist P, Ma C, Marks P, Maxham M, Murphy D, Park I, Pham T, Phillips M, Roy J, Sebra R, Shen G, Sorenson J, Tomaney A, Travers K, Trulson M, Vieceli J, Wegener J, Wu D, Yang A, Zaccarin D, Zhao P, Zhong F, Korlach J, Turner S (2009). Real-time DNA sequencing from single polymerase molecules. Science, 323(5910): 133–138

Esteller M (2007). Cancer epigenomics: DNA methylomes and histone-modification maps. Nat Rev Genet, 8(4): 286–298

Fan H C, Wang J, Potanina A, Quake S R (2011). Whole-genome molecular haplotyping of single cells. Nat Biotechnol, 29(1): 51–57

Farago C, Chester I C (1961). Cancer in the Territory of Papua and New Guinea: a preliminary communication. Med J Aust, 48(2): 1033–1035

Feng Z, Fang G, Korlach J, Clark T, Luong K, Zhang X, Wong W, Schadt E (2013). Detecting DNA modifications from SMRT sequencing data by modeling sequence context dependence of polymerase kinetic. PLOS Comput Biol, 9(3): e1002935

Frontera W R, Zayas A R, Rodriguez N (2012). Aging of human muscle: understanding sarcopenia at the single muscle cell level. Phys Med Rehabil Clin N Am, 23(1): 201–207, xiii

Galán A, Montaner D, Póo M E, Valbuena D, Ruiz V, Aguilar C, Dopazo J, Simón C (2010). Functional genomics of 5- to 8-cell stage human embryos by blastomere single-cell cDNA analysis. PLoS ONE, 5 (10): e13615

Geschwind D H, Konopka G (2009). Neuroscience in the era of functional genomics and systems biology. Nature, 461(7266): 908–915

Guo Y, Yang Y, Zhou J, Czajkowsky D M, Liu B, Shao Z (2012). Microdissection of spatially identified single nuclei in a solid tumor for single cell whole genome sequencing. Biotechniques, 0(0): 1–3

Hamburger A W, Salmon S E (1977). Primary bioassay of human tumor stem cells. Science, 197(4302): 461–463

Hanson E K, Ballantyne J (2005). Whole genome amplification strategy for forensic genetic analysis using single or few cell equivalents of genomic DNA. Anal Biochem, 346(2): 246–257

Hashimshony T, Wagner F, Sher N, Yanai I (2012). CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. Cell Reports 2, 666–673

He C (2010). Grand challenge commentary: RNA epigenetics? Nat Chem Biol, 6(12): 863–865

Heitzer E, Auer M, Gasch C, Pichler M, Ulz P, Hoffmann E M, Lax S, Waldispuehl-Geigl J, Mauermann O, Lackner C, Höfler G, Eisner F, Sill H, Samonigg H, Pantel K, Riethdorf S, Bauernhofer T, Geigl J B, Speicher M R (2013). Complex tumor genomes inferred from single circulating tumor cells by array-CGH and next-generation sequencing. Cancer Res, 73(10): 2965–2975

Hou Y, Song L, Zhu P, Zhang B, Tao Y, Xu X, Li F, Wu K, Liang J, Shao D, Wu H, Ye X, Ye C, Wu R, Jian M, Chen Y, Xie W, Zhang R, Chen L, Liu X, Yao X, Zheng H, Yu C, Li Q, Gong Z, Mao M, Yang X, Yang L, Li J, Wang W, Lu Z, Gu N, Laurie G, Bolund L, Kristiansen K, Wang J, Yang H, Li Y, Zhang X, Wang J (2012). Single-cell exome sequencing and monoclonal evolution of a JAK2-negative myeloproliferative neoplasm. Cell, 148(5): 873–885

Hughes S, Jones J L (2007). The use of multiple displacement amplified DNA as a control for methylation specific PCR, pyrosequencing, bisulfite sequencing and methylation-sensitive restriction enzyme PCR. BMC Mol Biol, 8(1): 91

Hutchison C A 3rd, Smith H O, Pfannkoch C, Venter J C (2005). Cell-free cloning using phi29 DNA polymerase. Proc Natl Acad Sci USA, 102(48): 17332–17336

Iourov I Y, Vorsanova S G, Yurov Y B (2012). Single cell genomics of the brain: focus on neuronal diversity and neuropsychiatric diseases. Curr Genomics, 13(6): 477–488

Islam S, Kjällquist U, Moliner A, Zajac P, Fan J B, Lönnerberg P, Linnarsson S (2011). Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. Genome Res, 21(7): 1160–1167

Jones P A, Baylin S B (2007). The epigenomics of cancer. Cell, 128(4):

683–692

Kleinsmith L J, Pierce G B Jr (1964). Multipotentiality of Single Embryonal Carcinoma Cells. Cancer Res, 24: 1544–1551

Lecault V, Vaninsberghe M, Sekulovic S, Knapp D J, Wohrer S, Bowden W, Viel F, McLaughlin T, Jarandehei A, Miller M, Falconnet D, White A K, Kent D G, Copley M R, Taghipour F, Eaves C J, Humphries R K, Piret J M, Hansen C L (2011). High-throughput analysis of single hematopoietic stem cell proliferation in micro-fluidic cell culture arrays. Nat Methods, 8(7): 581–586

Lee J Y, Dong S M, Kim S Y, Yoo N J, Lee S H, Park W S (1998). A simple, precise and economical microdissection technique for analysis of genomic DNA from archival tissue sections. Virchows Archiv: an international journal of pathology 433, 305–309

Li L, Clevers H (2010). Coexistence of quiescent and active adult stem cells in mammals. Science, 327(5965): 542–545

Ling J, Deng Y, Long X, Liu J, Du H, Cao B, Xu K (2012). Single-nucleotide polymorphism array coupled with multiple displacement amplification: accuracy and spatial resolution for analysis of chromosome copy numbers in few cells. Biotechnol Appl Biochem, 59(1): 35–44

Lu S, Zong C, Fan W, Yang M, Li J, Chapman A R, Zhu P, Hu X, Xu L, Yan L, Bai F, Qiao J, Tang F, Li R, Xie X S (2012). Probing meiotic recombination and aneuploidy of single sperm cells by whole-genome sequencing. Science, 338(6114): 1627–1630

Luthra R, Medeiros L J (2004). Isothermal multiple displacement amplification: a highly reliable approach for generating unlimited high molecular weight genomic DNA from clinical specimens. J Mol Diagn, 6(3): 236–242

Mardis E R (2008). The impact of next-generation sequencing technology on genetics. Trends Genet, 24(3): 133–141

McLean J S, Lombardo M J, Ziegler M G, Novotny M, Yee-Greenbaum J, Badger J H, Tesler G, Nurk S, Lesin V, Brami D, Hall A P, Edlund A, Allen L Z, Durkin S, Reed S, Torriani F, Nealson K H, Pevzner P A, Friedman R, Venter J C, Lasken R S (2013). Genome of the pathogen Porphyromonas gingivalis recovered from a biofilm in a hospital sink using a high-throughput single-cell genomics platform. Genome Res, 23(5): 867–877

McVean G A, Myers S R, Hunt S, Deloukas P, Bentley D R, Donnelly P (2004). The fine-scale structure of recombination rate variation in the human genome. Science, 304(5670): 581–584

McWilliam Leitch E C, McLauchlan J (2013). Determining the cellular diversity of hepatitis C virus quasispecies by single-cell viral sequencing. J Virol, 87(23): 12648–12655

Metzker M L (2010). Sequencing technologies- the next generation. Nat Rev Genet, 11(1): 31–46

Morozova O, Marra M A (2008). Applications of next-generation sequencing technologies in functional genomics. Genomics, 92(5): 255–264

Navin N, Kendall J, Troge J, Andrews P, Rodgers L, McIndoo J, Cook K, Stepansky A, Levy D, Esposito D, Muthuswamy L, Krasnitz A, McCombie W R, Hicks J, Wigler M (2011). Tumour evolution inferred by single-cell sequencing. Nature, 472(7341): 90–94

Niedringhaus T P, Milanova D, Kerby M B, Snyder M P, Barron A E (2011). Landscape of next-generation sequencing technologies. Anal Chem, 83(12): 4327–4341

Paez J G, Lin M, Beroukhim R, Lee J C, Zhao X, Richter D J, Gabriel S, Herman P, Sasaki H, Altshuler D, Li C, Meyerson M, Sellers W R

(2004). Genome coverage and sequence fidelity of phi29 polymerase-based multiple strand displacement whole genome amplification. Nucleic Acids Res, 32(9): e71

Pan X, Durrett R E, Zhu H, Tanaka Y, Li Y, Zi X, Marjani S L, Euskirchen G, Ma C, Lamotte R H, Park I H, Snyder M P, Mason C E, Weissman S M (2013). Two methods for full-length RNA sequencing for low quantities of cells and single cells. Proc Natl Acad Sci USA, 110(2): 594–599

Pan X, Urban A E, Palejev D, Schulz V, Grubert F, Hu Y, Snyder M, Weissman S M (2008). A procedure for highly specific, sensitive, and unbiased whole-genome amplification. Proc Natl Acad Sci USA, 105 (40): 15499–15504

Park S Y, Gönen M, Kim H J, Michor F, Polyak K (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

Quenneville S, Turelli P, Bojkowska K, Raclot C, Offner S, Kapopoulou A, Trono D (2012). The KRAB-ZFP/KAP1 system contributes to the early embryonic establishment of site-specific DNA methylation patterns maintained during development. Cell Reports 2: 766–773

Ramsköld D, Luo S, Wang Y C, Li R, Deng Q, Faridani O R, Daniels G A, Khrebtukova I, Loring J F, Laurent L C, Schroth G P, Sandberg R (2012). Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. Nat Biotechnol, 30(8): 777–782

Reya T, Morrison S J, Clarke M F, Weissman I L (2001). Stem cells, cancer, and cancer stem cells. Nature, 414(6859): 105–111

Rothberg J M, Hinz W, Rearick T M, Schultz J, Mileski W, Davey M, Leamon J H, Johnson K, Milgrew M J, Edwards M, Hoon J, Simons J F, Marran D, Myers J W, Davidson J F, Branting A, Nobile J R, Puc B P, Light D, Clark T A, Huber M, Branciforte J T, Stoner I B, Cawley S E, Lyons M, Fu Y, Homer N, Sedova M, Miao X, Reed B, Sabina J, Feierstein E, Schorn M, Alanjary M, Dimalanta E, Dressman D, Kasinskas R, Sokolsky T, Fidanza J A, Namsaraev E, McKernan K J, Williams A, Roth G T, Bustillo J (2011). An integrated semiconductor device enabling non-optical genome sequencing. Nature, 475(7356): 348–352

Sasagawa Y, Nikaido I, Hayashi T, Danno H, Uno K D, Imai T, Ueda H R (2013). Quartz-Seq: a highly reproducible and sensitive single-cell RNA sequencing method, reveals non-genetic gene-expression heterogeneity. Genome Biol, 14(4): R31

Schoenborn L, Yates P S, Grinton B E, Hugenholtz P, Janssen P H (2004). Liquid serial dilution is inferior to solid media for isolation of cultures representative of the phylum-level diversity of soil bacteria. Appl Environ Microbiol, 70(7): 4363–4366

Shalek A K, Satija R, Adiconis X, Gertner R S, Gaublomme J T, Raychowdhury R, Schwartz S, Yosef N, Malboeuf C, Lu D, Trombetta J J, Gennert D, Gnirke A, Goren A, Hacohen N, Levin J Z, Park H, Regev A (2013). Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. Nature, 498 (7453): 236–240

Shendure J, Ji H (2008). Next-generation DNA sequencing. Nat Biotechnol, 26(10): 1135–1145

Simone N L, Bonner R F, Gillespie J W, Emmert-Buck M R, Liotta L A (1998). Laser-capture microdissection: opening the microscopic frontier to molecular analysis. Trends Genet, 14(7): 272–276

Song C X, Clark T A, Lu X Y, Kislyuk A, Dai Q, Turner S W, He C, Korlach J (2012). Sensitive and specific single-molecule sequencing

of 5-hydroxymethylcytosine. Nat Methods, 9(1): 75–77

Tahiliani M, Koh K P, Shen Y, Pastor W A, Bandukwala H, Brudno Y, Agarwal S, Iyer L M, Liu D R, Aravind L, Rao A (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science, 324(5929): 930–935

Tan D W, Jensen K B, Trotter M W, Connelly J T, Broad S, Watt F M (2013). Single-cell gene expression profiling reveals functional heterogeneity of undifferentiated human epidermal cells. Development, 140(7): 1433–1444

Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, Wang X, Bodeau J, Tuch B B, Siddiqui A, Lao K, Surani M A (2009). mRNA-Seq whole-transcriptome analysis of a single cell. Nat Methods, 6(5): 377–382

Tang F, Lao K, Surani M A (2011). Development and applications of single-cell transcriptome analysis. Nat Methods, 8(4 Suppl): S6–S11

Torres L, Ribeiro F R, Pandis N, Andersen J A, Heim S, Teixeira M R (2007). Intratumor genomic heterogeneity in breast cancer with clonal divergence between primary carcinomas and lymph node metastases. Breast Cancer Res Treat, 102(2): 143–155

Tzvetkov M V, Becker C, Kulle B, Nürnberg P, Brockmöller J, Wojnowski L (2005). Genome-wide single-nucleotide polymorphism arrays demonstrate high fidelity of multiple displacement-based whole-genome amplification. Electrophoresis, 26(3): 710–715

Unger M A, Chou H P, Thorsen T, Scherer A, Quake S R (2000). Monolithic microfabricated valves and pumps by multilayer soft lithography. Science, 288(5463): 113–116

Voet T, Kumar P, Van Loo P, Cooke S L, Marshall J, Lin M L, Zamani Esteki M, Van der Aa N, Mateiu L, McBride D J, Bignell G R, McLaren S, Teague J, Butler A, Raine K, Stebbings L A, Quail M A, D'Hooghe T, Moreau Y, Futreal P A, Stratton M R, Vermeesch J R, Campbell P J (2013). Single-cell paired-end genome sequencing reveals structural variation per cell cycle. Nucleic Acids Res, 41(12): 6119–6138

Wang J, Fan H C, Behr B, Quake S R (2012). Genome-wide single-cell analysis of recombination activity and de novo mutation rates in human sperm. Cell, 150(2): 402–412

Wu A R, Neff N F, Kalisky T, Dalerba P, Treutlein B, Rothenberg M E, Mburu F M, Mantalas G L, Sim S, Clarke M F, Quake S R (2013). Quantitative assessment of single-cell RNA-sequencing methods. Nat Methods,

Xu X, Hou Y, Yin X, Bao L, Tang A, Song L, Li F, Tsang S, Wu K, Wu H, He W, Zeng L, Xing M, Wu R, Jiang H, Liu X, Cao D, Guo G, Hu X, Gui Y, Li Z, Xie W, Sun X, Shi M, Cai Z, Wang B, Zhong M, Li J, Lu Z, Gu N, Zhang X, Goodman L, Bolund L, Wang J, Yang H, Kristiansen K, Dean M, Li Y, Wang J (2012). Single-cell exome sequencing reveals single-nucleotide mutation characteristics of a kidney tumor. Cell, 148(5): 886–895

Xue Z, Huang K, Cai C, Cai L, Jiang C Y, Feng Y, Liu Z, Zeng Q, Cheng L, Sun Y E, Liu J Y, Horvath S, Fan G (2013). Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. Nature, 500(7464): 593–597

Yan L, Yang M, Guo H, Yang L, Wu J, Li R, Liu P, Lian Y, Zheng X, Yan J, Huang J, Li M, Wu X, Wen L, Lao K, Li R, Qiao J, Tang F (2013). Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. Nat Struct Mol Biol, 20(9): 1131–1139

Zandi S, Ahsberg J, Tsapogas P, Stjernberg J, Qian H, Sigvardsson M (2012). Single-cell analysis of early B-lymphocyte development suggests independent regulation of lineage specification and commitment in vivo. Proc Natl Acad Sci USA, 109(39): 15871–15876

Zhao Y, Gong X, Chen L, Li L, Liang Y, Chen S, Zhang Y (2013). Site-specific methylation of placental HSD11B2 gene promoter is related to intrauterine growth restriction. Eur J Hum Genet, doi: 10.1038/ejhg.2013.226

Zong C, Lu S, Chapman A R, Xie X S (2012). Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. Science, 338(6114): 1622–1626