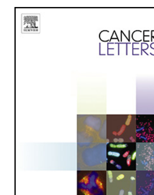
Cancer Letters 365 (2015) 1–10

Contents lists available at ScienceDirect

Cancer Letters

journal homepage: www.elsevier.com/locate/canlet



Mini-review

Recent advances and current issues in single-cell sequencing of tumors

Hai-jian Sun, Jian Chen, Bing Ni, Xia Yang *, Yu-zhang Wu **

Institute of Immunology, Third Military Medical University, Chongqing 400038, China

A R T I C L E   I N F O

Article history:
Received 31 January 2015
Received in revised form 19 April 2015
Accepted 20 April 2015

Keywords:
Single-cell sequencing
Tumor
Sequencing analysis
Intratumoral heterogeneity

A B S T R A C T

Intratumoral heterogeneity is a recently recognized but important feature of cancer that underlies the various biocharacteristics of cancer tissues. The advent of next-generation sequencing technologies has facilitated large scale capture of genomic data, while the recent development of single-cell sequencing has allowed for more in-depth studies into the complex molecular mechanisms of intratumoral heterogeneity. In this review, the recent advances and current challenges in single-cell sequencing methodologies are discussed, highlighting the potential power of these data to provide insights into oncological processes, from tumorigenesis through progression to metastasis and therapy resistance.

© 2015 The Authors. Published by Elsevier Ireland Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

Cancer remains as one of the world's most lethal and debilitating diseases, despite massive research efforts to understand the pathogenic processes and underlying molecular mechanisms. While these studies have identified many features of cancer development and progression, such as genetic mutations and pathogenic phenotypes, they have also uncovered perplexing complexity in the tumor microenvironment, even the individual cells within the tumor tissue showing distinctive phenotypes [1,2]. This newly recognized feature is known as intratumoral heterogeneity and may represent differing pathogenic potentials of tumor cells, where some cells may represent a greater threat to the organism by having a greater potential for metastasis or development of resistance to chemotherapeutic drugs.

Applying single-cell sequencing technology to study the phenomenon of intratumoral heterogeneity has uncovered its role in cancer diagnosis and prognosis, as well as identifying factors that may represent clinically useful diagnostic and/or prognostic biomarkers and targets of molecular therapies [2]. The single-cell sequencing technology is powerful and has the capability of capturing dynamic genomic data for a single cancer cell under specific cancer-related conditions, such as during the early stages of tumor development or the late stages of metastasis. The resultant data may help to define condition-related genetic profiles that underlie the mechanisms of tumorigenesis, tumor progression, and development of metastasis and resistance to various therapies [3–5].

One of the major technical challenges of single-cell sequencing is the limited amount (picogram levels) of DNA and RNA in a single tumor cell, which is below the threshold of sensitivity for even the most advanced sequencing platforms. To overcome this limitation, the nucleic acids from the isolated single cell must be amplified. Thus, the current protocol for single-cell sequencing involves four main steps: isolation of a single tumor cell, cell lysis, extraction of nucleic acids, and amplification. Each step, however, carries a potential for error that limits the efficacy and sensitivity of the technology. The use of micropipetting, albeit common throughout molecular biology experimental studies, is prone to operator error and mechanical failure [6]. Isolation of single tumor cell for analysis is often carried out by fluorescence activated cell sorting (FACS) or magnetic activated cell sorting (MACS), both of which may alter the transcriptional state of the cells [7], confounding the subsequent genomic analysis. The methods involved in the cell lysis and extraction steps carry the risk of causing DNA or RNA degradation, sample loss, or contamination [6,8,9]. Finally, the current methods used to amplify the low concentration of nucleic acids from a single cell can yield non-uniform levels of amplification products and inconsistent coverage of the original genomic information [10,11].

The sequencing operation itself is also imperfect and even the most advanced technological platforms have limited sensitivity, with a threshold of resolution, and require operator expertise and lengthy operation times, and the design and running of algorithms to interpret the sequencing data [12,13]. Certainly, the use of single-cell sequencing technology to describe the distinctive tumor-related genetic profiles for various oncological processes will benefit

* Corresponding author. Tel.: +86 23 68771896; fax: +86-23-68772230.
E-mail address: oceanyx@126.com (X. Yang).
** Corresponding author. Tel.: +86 23 68752235; fax: +86-23-68752230.
E-mail address: wuyuzhang@tmmu.edu.cn (Y. Wu).

http://dx.doi.org/10.1016/j.canlet.2015.04.022
0304-3835/© 2015 The Authors. Published by Elsevier Ireland Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

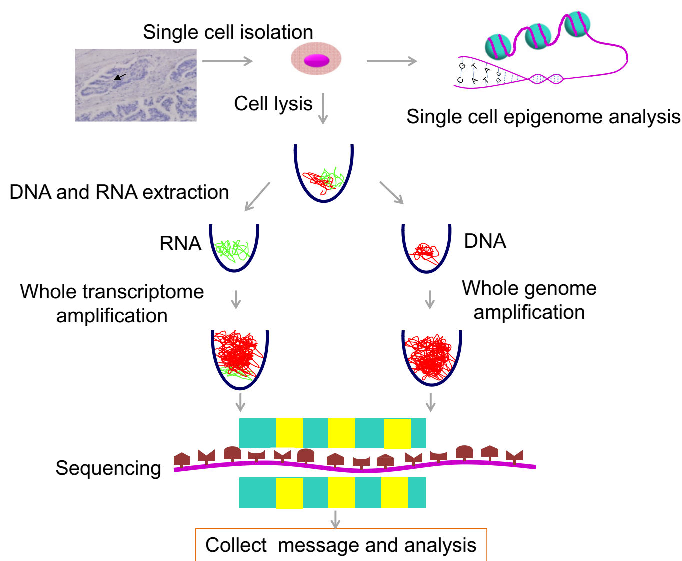


Fig. 1. Single-cell sequencing of a tumor cell. A tumor specimen is obtained by surgical excision and single cells are isolated by one of the several methods shown in Fig. 2. The individual cancer cell can be used for epigenome sequencing directly or lysed to extract the genetic material (DNA and RNA), which is in turn amplified by the methods shown in Fig. 3. Then, the amplified DNA and RNA are sequenced by single-cell sequencing technology and the result data are analyzed to provide insights into the molecular mechanisms underlying intratumor heterogeneity.

from research efforts to improve the sensitivity, efficiency and rapidity of this analytical approach.

Herein, we provide a review of the methods and procedures of single-cell sequencing (Fig. 1), including those that are more established and those that have been newly developed, and discuss their usefulness and limitations with the hope that this knowledge will support the ongoing research efforts to improve this technology and increase our knowledge in intratumor heterogeneity.

Isolation of single cells for sequencing

Isolating single cells from solid or liquid tumor tissues is the first and key step of single cancer cell sequencing. The following four methods are the most commonly used for single-cell isolation, and their primary characteristics are summarized in Table 1 [14].

Laser capture microdissection (LCM)

LCM is used to isolate individual cells from their native tissue directly and without any chemical or physical destruction. The basic procedure for LCM involves covering a section of the target tissue with a thermoplastic film and firing a laser to melt the film so that it adheres to a single tumor cell and facilitates mechanical isolation of that cell from its neighbors (Fig. 2A). Some researchers have proposed the use of immunofluorescence (i.e. fluorescent-tagged antibodies) to improve the speed and accuracy of this method [15–17]. However, the LCM laser's isolation of a single cell is still not capable of completely clearing all microenvironment materials' strong adhesion junction from a target cell's surface, which represents a source of contaminating materials from neighboring cells [18–20].

Micropipette isolation

The use of a micropipette to mechanically isolate a single cell by aspiration is a cheap and technically simple method since the only tools required are a basic micropipette and microscope with relatively lower resolution (at the individual cell level). The basic procedure for micropipette isolation involves enzymatic digestion of the tumor tissue to acquire a cell suspension, after which the suspension is diluted to approximately 10–20 cells per 1 μL solution and examined under the microscope's visual field to select a single cancer cell by micropipette-assisted aspiration (Fig. 2B). The limitations of this procedure are the laborious nature of the technique, resulting in low efficiency, and the dependence on the operator's expertise and low resolution visibility of the cells, making it highly prone to mistakes in identifying truly individual cells or contaminating non-cancer cells [6,21,22]. Some researchers have proposed the use of fluorescence *in situ* hybridization (FISH) to improve the efficiency of micropipette isolation [23].

Fluorescence-activated cell sorting (FACS)

The FACS platform is capable of both selecting and analyzing single cells from a heterogeneous tumor tissue, relying on simultaneous analysis of fluorescence signals and light scattering parameters. The operational procedure of FACS involves disruption of the cancer tissue (either by enzymatic digestion or mechanical disruption) to acquire a cell suspension, after which the target cells are immunoreacted with one or more fluorescence-tagged antibodies targeting specific antigens on the cell surface or intracellular markers; in this manner, specific cell types can be identified for isolation and analysis. After processing, the antibody-labeled cells are applied to a FACS machine in which they pass through a narrow stream of fluid with an intersecting laser beam that produces scattered light signals according to the labeled cell characteristics. The flow of the cell-containing liquid can be regulated to ensure that only one cell is present in a single droplet, and the light scattering pattern will help to identify that cell; in addition, during the FACS separation procedure, the cell will acquire temporary electrical properties that facilitate cell isolation (Fig. 2C) [24–26]. The FACS isolation approach has been modified to use a panel of isotope-tags for the immunoreactive antibodies [27].

The FACS isolation approach has several advantages over the LCM and micropipette isolation procedures, including greater rapidity and higher throughput. However, FACS detection of fluorescent signals is relatively low, and low-expression markers are difficult or impossible to detect [24], so that there is a risk of missing some specific cell types.

Table 1
Methods of isolating single cells for sequencing.

Method	Sorting principle	Single cell suspension	Cost	Throughput	Automated?	Limitation
Micropipette isolation	Microexamination	Yes	Low	Low	No	Operational bias
LCM	Laser beam microdissection	No	High	Low	No	Contamination by surrounding material
FACS	Antigen–antibody hybrid	Yes	High	High	Yes	Needs large numbers of cells
MACS	Antigen–antibody hybrid	Yes	High	High	Yes	Operational complexity

Abbreviations: LCM, laser capture microdissection; FACS, fluorescence-activated cell sorting; MACS, magnetic-activated cell sorting.

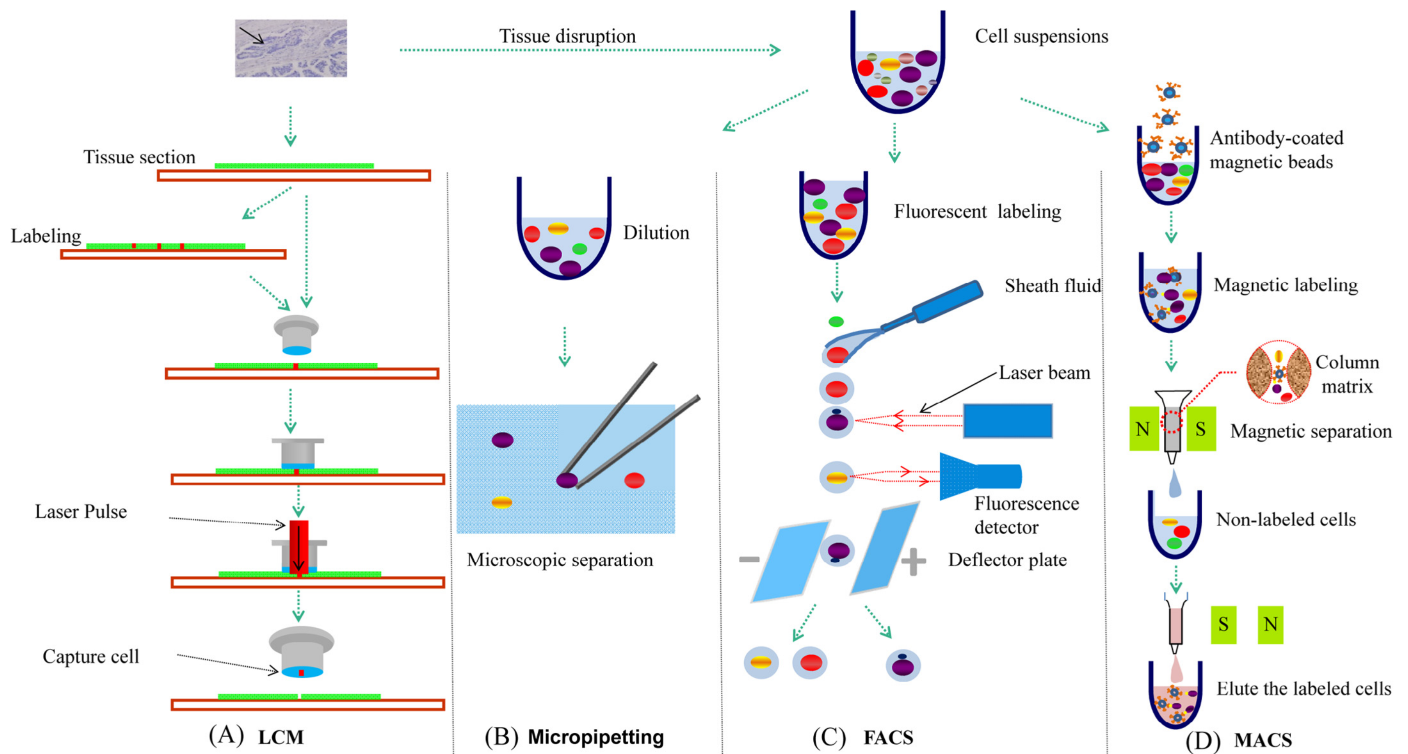


Fig. 2. Methods for single-cell isolation. (A) In LCM, the tumor tissue is placed on a microscope slide and covered with a thermoplastic film; a laser is beamed through the film, focused on the cell of interest, so that the film adheres to the cell for isolation by mechanical dislocation from the tissue. In the next three methods (B–D), the tumor tissue is first digested into a cell suspension. (B) Micropipette aspiration uses a standard micropipette to mechanically capture the single cell from a cell suspension on a microscope slide. (C) FACS detects the signal from fluorescent dyes that are bound to the surface of a specific target cell. (D) MACS relies on the immunoreactivity of a cell surface antigen with magnetic particles, so that the mixture can be sorted by a MACS Column placed in a MACS Separator. Abbreviations: LCM, laser capture microdissection; FACS, fluorescence-activated cell sorting; MACS, magnetic-activated cell sorting.

Magnetic-activated cell sorting (MACS)

MACS can capture and isolate single cells effectively and rapidly. The operational procedure of MACS involves immunoreactivity of antigens in the cell membrane and antibody conjunction with magnetic particles that allow for separation by magnetic fields (Fig. 2D) [28,29]. MACS has been successfully applied to isolate individual cells of different phenotypes from a variety of environments, and has been especially useful in separating circulating tumor cells (CTCs) from peripheral blood for use in cancer-related experimental research [28,30].

Lysis of isolated single cancer cells for sequencing

After isolation of the single cancer cell, lysis is necessary to obtain the genomic material for sequencing; an ideal lysis method will minimize substrate dissipation and potential contamination, and induce no changes to the genomic architecture that may influence downstream analyses. The currently available methods are either physics-dependent or reagents-dependent, and decision for the appropriate lysis method is made according to the downstream operation and sample state; furthermore, more than one method can be used on each sample [31–35].

The most common physics-dependent methods are optoelectronic tweezers [36], sonication, nanoknives, electroporation and freeze/thawing. The reagents-dependent methods include detergent-based lysis, alkaline lysis and enzymatic cell lysis [31,35]. Among all the methods, alkaline lysis and Proteinase K digestion are the most widely used in oncological studies for digestion of tumor tissues [37–43].

The alkaline lysis method is carried out by placing the isolated single cancer cell in phosphate-buffered saline, mixing in the alkaline lysis buffer, and incubating the solution for 10 minutes on ice. The lysis action is stopped upon addition of a neutralizing buffer [37,44,45]. In contrast, the proteinase K lysis protocol uses a solution consisting of proteinase K, 1× Tris–EDTA buffer and 1% Tween-20; the single cancer cell is incubated in this solution for 3 hours at 55 °C. Comparisons of these two reagent-dependent methods have shown that the proteinase K digestion is more efficient than the alkaline lysis method when the cell number is lower than 100, and that the former method produces less loss of nucleic acids due to its pH being closer to the normal physiological level [21,31,37].

Nucleic acid (NA) extraction from single cancer cells for sequencing

The currently available methods for NA extraction use either column-based or solution-based protocols, and most of these are available as convenient commercial kits [46]. The silica affinity binding column technology is the prototypical column-based method, and has been applied frequently for the extraction of DNA and RNA from single isolated tumor cells for subsequent sequencing [41,47–49]. The silica technology is based on the theory that high affinity exists between the positively-charged silica matrices and the negatively-charged DNA backbone in high ionic strength solutions; therefore, in this extraction approach DNA binds to silica and can be readily eluted by washing with a low ionic strength buffer [46,50]. This protocol has been made available as convenient commercial kits, such as the AllPrep DNA/RNA kit (Qiagen) and the Arcturus PicoPure RNA isolation kit (Applied Biosystems Inc.) [41,47].

Microfluidics devices have been widely used for NA extraction, and are especially useful with microquantities of DNA and RNA [51–53]. The microfluidics device is usually a laminar device bound by two opposing substrates; in the middle of the two substrates is a channel. A solution containing labeled target materials is passed through the channel and can be sorted during passage by application of an exterior power source, such as a magnetic field [52,54]. One of the most recently developed microfluidics methods is capable of extracting and purifying both mRNA and DNA from a single sample (1–10 cells) simultaneously; this method, known by the acronym SNARE (for "Selective Nucleic Acid Removal via Exclusion"), involves contiguous wells in a line connected by a trapezoid channel, consisting of an input well, followed by a middle well filled with silicon oil, and two side-by-side output wells for the DNA and mRNA outputs [9]. The input well is filled with a mixture of input materials (the cell lysis product), paramagnetic particles (PMPs; labeled with oligo$(dt)_{25}$) and lysis/binding buffer (which promotes the binding of PMPs and mRNA). An external magnet is applied to pull the mRNA- or DNA-bound PMPs (according to the charge of the magnet) toward the front of the input well and through the middle well until they enter the respective output well [9]. SNARE has been successfully used to extract and purify DNA and RNA from prostate cancer CTCs [9]. Moreover, the SNARE method is cost and time efficient, requires no hazardous chemicals, and has higher sensitivity than other extraction methods [9].

Amplification of genetic material from single cancer cells for sequencing

Whole genome amplification (WGA)

A normal diploid human cell contains approximately 6~7 pg of DNA, which is inadequate for genomic sequencing using the currently available platforms [3,31]. Therefore, amplification methods are necessary and must be highly effective and low bias. The WGA methods currently available are either based on temperature cycling (routine PCR, PicoPLEX technology, multiple annealing and looping-based amplification cycles (MALBAC)) or isothermal conditions (multiple displacement amplification (MDA)) [10].

Routine PCR for WGA

Several PCR-based methods have been reported for WGA, including the commonly used degenerate oligonucleotide primed PCR (DOP-PCR) and primer extension preamplification-PCR (PEP-PCR).

The DOP-PCR reaction is carried out in two stages, each reliant upon a distinctive primer corresponding to specific sequences for amplification [55]. The first-stage primer is a semi-random priming sequence, so that both tail ends of the primer can bind to multiple sites in the genome and initiate amplification at a low annealing temperature [3,55,56]. The second-stage primer is a non-random priming sequence, so that only fragments tagged by the specific sequence of the former primer are bound; this facilitates exponential extension at a high annealing temperature (Fig. 3A) [55,56].

The PEP-PCR reaction is based on random priming and uses a collection of 15-base random oligonucleotides as primers, allowing for up to 4^{15} different primers (Fig. 3B). After 50 primer-extension cycles, no less than 30 copies of DNA should be produced, and the coverage of genomic should be >78% at the single cell level [57]. An improved version of this reaction, known as improved (I)-PEP-PCR, was recently developed, in which a proofreading DNA polymerase was added to the PEP-PCR reaction system in order to remove any mismatched nucleotides at the 3′ ends of the random primers. The I-PEP-PCR method has been shown to have remarkably better amplification efficiency than either DOP-PCR or PEP-PCR (40% vs. 15% or 3% at the single cell level) [58].

PicoPLEX technology for WGA

The PicoPLEX amplification technology was specifically designed to amplify DNA from a single cell by the use of a specific primer in the traditional PCR reaction. The methodology involves a DNA pre-amplification stage followed by DNA amplification (Fig. 3C). In the first stage, random primer extension is carried out with non-complementary primers to amplify genomic DNA fragments that are end-labeled with sequences that will facilitate hybridization to single-stranded (ss) DNA for the subsequent PCR-based amplification. This protocol has been made available as a convenient commercial kit, such as the PicoPLEX Single-Cell WGA kit (New England Biolabs). It has been reported that the PicoPLEX technology has a low level of inherent bias in amplification but, in comparison to the MDA approach (detailed below), is limited by a higher sensitivity to background contamination [10].

Multiple annealing and looping-based amplification cycles (MALBAC) for WGA

MALBAC is a quasi-linear amplification technology that combines linear displacement amplification methods with traditional PCR to amplify the DNA fragments for use in downstream sequencing applications [3]. The amplification process of MALBAC includes a single cycle of four isothermal conditions followed by 20 cycles of PCR amplification [59]. The MALBAC primer consists of a 27-mer common nucleotide sequence and eight variable nucleotides. In the single cycle, the four isothermal conditions include DNA polymerase hybridizing to the templates at 0 °C, DNA polymerase-mediated strand displacement generation of semi-amplicons at 65 °C, DNA double-strands uncoiling at 94 °C, and complementary sequences from both ends of the full amplicons hybridizing to each other to form looped DNAs at 58 °C, thereby preventing extra amplification. Afterwards, traditional PCR is carried out on the full amplicons for the 20 cycles (Fig. 3D) [60]. The MALBAC technique has been successfully applied in oncological studies to facilitate amplification of the whole genome of single CTCs obtained from lung cancer and to determine the genomic profiles of cells in metastatic cancer (providing data that distinguished primary cancer cells that are likely to become CTCs) [60,61]. Although MALBAC provides better uniformity of genome coverage than other amplification methods, it also has a high false-positive rate when detecting heterozygous loci and a high false-positive rate (40-fold higher than MDA) for genotyping single-nucleotide variants [11,59,60].

Multiple displacement amplification (MDA) for WGA

MDA is a non-specific amplification method that relies on two types of DNA polymerases, which not only bind to tandem DNA products generated by the single primer-initiated rolling circle amplification reaction but also can cause strand displacement; these types of DNA polymerases include the Phi29 DNA polymerase and the Bst large fragment DNA polymerase. The MDA methodology also relies on a random primer that binds to multiple binding sites in the displaced strand, thereby initiating complementary extension of the ssDNA. Use of these two types of primer-initiated reactions produces contiguous double-stranded DNA products and forms a network of hyperbranched DNA structures (Fig. 3E) [62,63]. Compared with traditional PCR-based WGA, MDA has several advantageous features, including the ability of Phi29 DNA polymerase to catalyze DNA extension reactions at low and constant temperatures and the ability of the process to generate lengthy amplicons (up to 12 kb) [56,63,64]. However, the MDA approach is not without limitations, and the procedure is disadvantaged by nonhomogeneity of amplification and high allele dropout rate (which can reach up to 65%) [6,31,65].

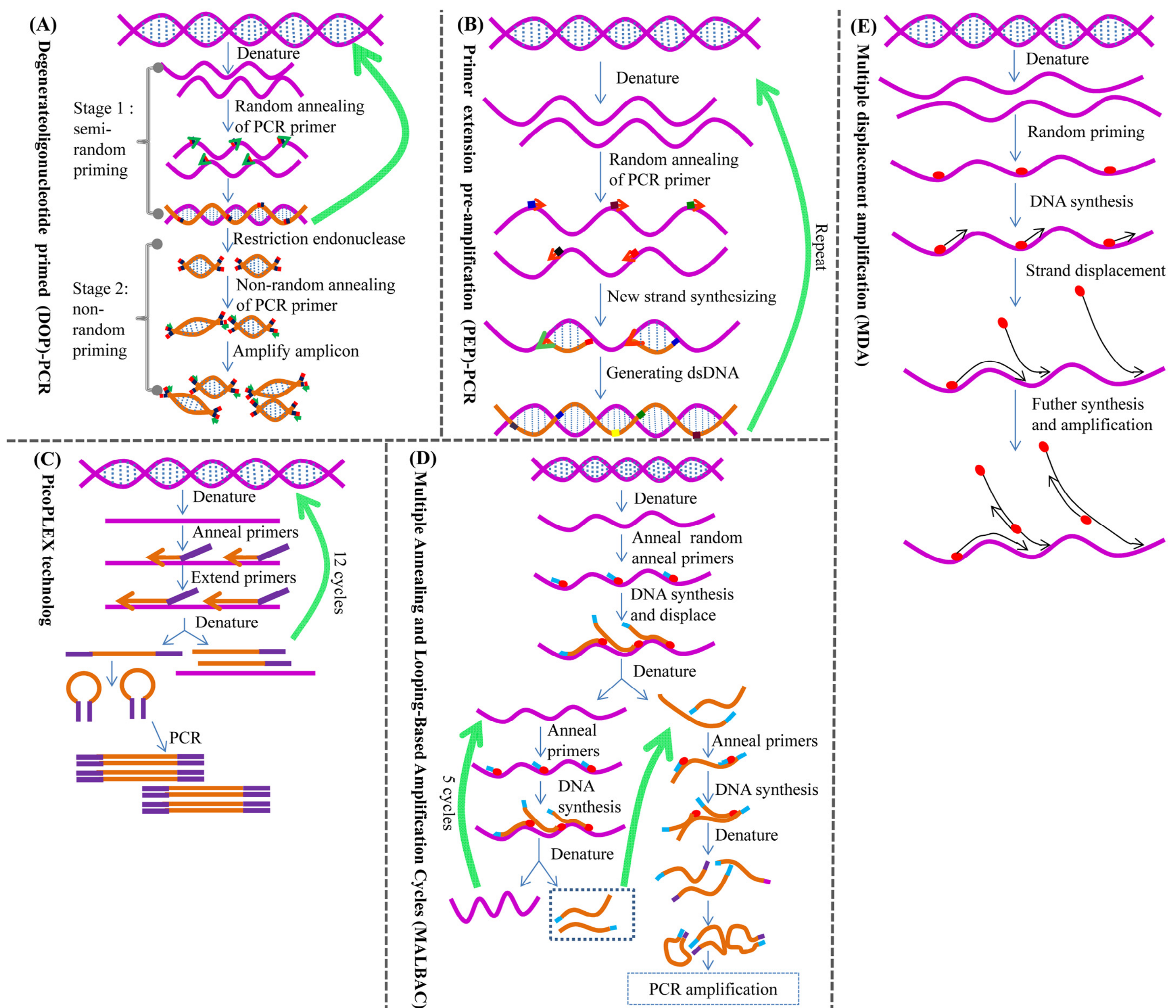


Fig. 3. Methods for whole-genome amplification. (A) DOP-PCR uses a random primer extension stage followed by an exponential amplicon replication stage. (B) PEP-PCR primers consist of 15-base random oligonucleotides, which are complementary to single strand template sequences and which can anneal at numerous positions in permissive thermocycling conditions. (C) PicoPLEX technology uses a specific primer that anneals to denatured DNA and initiates the extension reaction for 12 cycles, after which the amplification products are further amplified by PCR reaction for 14 cycles. (D) The MALBAC primer anneals to the single-strand DNA randomly, then synthesis of a new strand is initiated by a polymerase with displacement activity. In the next five amplification cycles, the DNA templates and semi-amplicons continue to extend, and the 5′ and 3′ sequences of the full amplicons are hybridized to generate a looped DNA fragment so that further pre-amplification is prevented. In the final step, the PCR reaction amplifies the DNA loop. (E) MDA is an isothermal reaction, in which random oligonucleotide primers anneal to denatured DNA for DNA synthesis in the 5′ to 3′ direction reliant on the strand displacement activity of DNA polymerases. When the polymerase reaches any newly synthesized double-stranded DNA, the polymerase is displaced to the newly synthesized strand to continue DNA polymerization.

The more recently developed microwell displacement amplification system (MIDAS) can reduce the amplification bias of MDA by using a reduced reaction volume [66,67]. In MIDAS, the genetic material from different cells (respectively placed in individual nanoliter wells) is amplified simultaneously by MDA in the respective wells as single reaction chambers. In order to overcome the potential problems of contamination (either exogenous or cross-well), a fluorescence label (e.g. SYBR Green I) has been included in the MDA master mix to facilitate visualization of the amplification process via fluorescence monitoring [44,63]. While MIDAS has been successfully used to study copy number variation, due to its excellent uniform genome coverage [44], it does not represent the perfect solution to overcome all limitations of MDA and carries its own limitations as well. The utilization rate of each microwell is only 10%, which is necessary to avoid the potential problems posed by more than one cell being present in the well, and the speed of amplification also remains suboptimal [44].

Whole transcriptome amplification (WTA)

The amount of RNA in a single cell, which is at pictogram level, cannot satisfy the requirements for analysis by the modern sequencing platforms, and WTA is necessary to perform single cell transcriptome sequencing. The currently available methods of WTA

Table 2
Methods for amplifying RNA for whole transcriptome analysis.

Method	Year	Principle	Coverage	Transcript lengths	Limitation	Ref
Tang's method	2009	Traditional PCR	64%	0.5~3 kb	3′ end bias	[6,14,70]
STRT-seq	2011	Modified PCR	5–25%	0.75~2.0 kb	Strong 3′ end bias	[3,69,73,79]
CEL-seq	2012	*In vitro* transcription	49%	Average 1.0 kb	Strong 3′ end bias, usually targets the last exons highly	[14,69,75,76,80]
Smart-seq	2012	Modified PCR	Nearly full-length	Average 1.5 kb	Cannot capture partially reverse-transcribed mRNA	[69,80–82]
PMA	2013	Phi29 DNA polymerase	Full-length	All sizes	Slight 5′ end bias	[69,78]
SMA	2013	Traditional PCR	Full-length	All sizes	Slight 5′ end bias	[69,78]
Quartz–seq	2013	*In vitro* transcription	81%	Average 2.5 kb	5′ end bias, cannot amplify cDNA with higher GC content	[69,76,83]
Smart-seq2	2014	Modified PCR	Nearly full-length	~2 kb	Cannot capture poly A^- RNA and reflects strand-specific mRNA	[6,69,84]

Abbreviations: STRT, single-cell tagged reverse transcription; CEL-seq, cell expression by linear amplification and sequencing; Smart-seq, switching mechanism at the 5′ end of the RNA template sequencing; PMA, Phi29 DNA polymerase-based mRNA transcriptome amplification; SMA, semi-random primed PCR-based mRNA transcriptome amplification procedure.

are based on traditional PCR, modified PCR, T7-*in vitro* transcription (IVT) and Phi29 DNA polymerase-mediated RNA amplification [68,69].

Traditional PCR for WTA

The traditional PCR-based methods begin with reverse transcription of the entire RNA of the cell, followed by rapid exponential amplification of the full-length cDNA [6,69]. The core principle of this methodology involves binding of the mRNA poly (A) tail to poly-(dT) primers with different anchor sequences; unfortunately, this feature precludes can not capture of mRNA without a poly (A) tail, such as histone mRNA, and causes 5′ end bias when the mRNA length is over 3 kb since the anchor sequence binds to the 3′ end [69–71]. In addition, the non-linear amplification will distort the initial representation of transcripts, especially for the last several PCR cycles [14].

Modified PCR for WTA

The modified PCR-based methods used to prepare the transcriptome for single-cell sequencing are also known as SMART techniques (an acronym for "Switching Mechanism At the 5′ end of the RNA Transcript") that exploit the intrinsic reverse transcription and template switching properties of the Moloney Murine Leukemia virus (MMLv) reverse transcriptase [72]. In SMART, a modified oligo (dT) primer is used to initiate the reverse transcription reaction and a non-template nucleotide sequence (consisting mainly of cytosines) corresponding to the 3′ end of the new single-stranded cDNA, known as the template-switching oligonucleotide, is annealed to the ssDNA generated upon template-extension to facilitate PCR amplification [72–74]. However, this methodological approach is incapable of amplifying mRNA without the 5′ end and fragments of mRNA that have only been partially reverse transcribed [69,74].

T7-IVT for WTA

The reverse transcription design of the primer used for the T7-IVT-based methods of WTA is different from those of all the other PCR-based methods. Several versions of this methodology have been developed. For the one known as CEL-seq, an acronym for "cell expression by linear amplification and sequencing", the primer includes an anchored poly-T, a unique barcode, the 5′ Illumina sequencing adaptor, and a T7 promoter [75]; for the one known as Quartz-Seq, the primer includes oligo-dT_{24}, the T7 promoter, and the specific sequences corresponding to the PCR target region [76]. The full procedure for the T7-IVT-based methods involves conversion of mRNA to DNA, to RNA, and again back to DNA [74]. Compared with the traditional PCR-based amplification method, the T7-IVT-based methods produce less accumulation of non-specific products, which supports their improved specificity and ratio fidelity [69,71,77]; however, the T7-IVT-based methods are more tedious and time-consuming [69,75].

Phi29 DNA polymerase for WTA

The Phi29 DNA polymerase has long been used in WGA, and has more recently been demonstrated as an effective tool for RNA amplification [69,78]. The Phi29-mRNA amplification (PMA) method carries out RNA reserve transcription with the use of a poly-dT primer, the subsequent cDNA amplification is carried out with the Phi29 DNA polymerase via the rolling circle amplification process [78]. A unique advantage of PMA is that it can achieve full-length amplification of all sizes of transcripts [69,78]; however, the genomic DNA must be removed before amplification to prevent its co-amplification [68,78].

A significant obstacle to all WTA methods is how to protect the intrinsic abundance of mRNA (without degradation) before and throughout the extraction and purification steps [6]. Moreover, the current amplification techniques for WTA (summarized in Table 2) should be improved to increase efficiency and reduce bias.

Sequencing technologies for single-cell sequencing

First-generation sequencing technology

The so-called 'first-generation' of sequencing technology is based on automation of dideoxy Sanger sequencing. The advent of this automated platform rocketed the progression of biological sciences and our approach to understanding normal physiological states and managing pathological conditions. The core concept of Sanger's method involves termination of the extension of dsDNA upon incorporation of a dideoxyribonucleoside triphosphate (ddNTP) into the extending DNA strand, with the termination point being determined by the relative concentration of available ddNTPs and dNTPs for this reaction. Unfortunately, the average length of Sanger's sequencing (up to 800 bases) is not adequate for WGS, and the laborious nature of the procedure (even when automated) has resulted in unacceptably high costs coupled with low throughput that limit its utility [85,86].

Next-generation sequencing (NGS)

Efforts to overcome the shortcomings of the first-generation sequencing technology culminated in the development of the NGS technologies in widespread use today. In NGS, the genomic DNA is sheared into millions of fragments, around 35–400 bp a piece, for *en masse* amplification [87]; the amplification approach varies among the different NGS platforms. The three most widely used NGS platforms are 454 Life Sciences/Roche, Illumina, and Applied Biosystems

SOLiD systems [88,89], although the first (454) has recently been taken off the market (including manufacturing (in 2016)).

The 454 and Illumina NGS platforms were based on the sequencing-by-synthesis (SBS) principle, while the SOLiD platform was based on the principle of sequencing-by-ligation (SBL). In SBS, light signals are generated by fluorescent-labeled single nucleotides that are incorporated into the extending strand. The 454 system uses pyrosequencing, a procedure that involves the release of a phosphate upon incorporation of a nucleotide which then activates firefly luciferase to produce light [88]. In contrast, the Illumina NGS system uses a unique cluster-amplification method in which adapter sequence-labeled DNA fragments are anchored to the surface of flow cell channels by binding with the corresponding sequences of the specific adapters and amplified by bridge amplification [87,90]. Sequencing of the amplified clusters involves incorporation of one of the four labeled reversible terminating nucleotides, and the imaging operation (i.e. reading of the sequence) is performed for each single nucleotide incorporation with the labeled reversible terminating nucleotide removed for the next incorporation by DNA polymerase [87,88]. In SBL, no DNA polymerase is used; instead, DNA ligase is used to match a known fluorescent-labeled nucleotide to an unknown nucleotide in a DNA fragment. In the SOLiD platform, DNA fragments are bound to flow cells, and fluorescent-labeled octamers are hybridized to the adaptor sequence for reading the sequence by fluorescence detection; while the first five labeled-nucleotides are integrated into the DNA sequence, only the first two can be identified and the other three (including the ones with a fluorescence signal) are removed for the next cycle [88,89].

Although NGS has many advantages over the traditional Sanger method and the first-generation technology, including lower cost, higher efficiency and lower background noise, it also has a substantially increased technical complexity, and requires a longer running time due to the advanced algorithms that analyze and process the enormous amounts of NGS data [13,91]. Furthermore, the NGS platforms require high-level bioinformatics expertise, such as that needed to develop and continually evolve effective databases to interconnect phenotypic information with DNA variants for clinical applications [86,92,93].

Third-generation sequencing (TGS)

Although NGS has proven incredibly useful in obtaining massive amounts of genomic data from tissues, there still exist several limitations when it is applied to single-cell sequencing, including insufficient resolution and read length, and unacceptable operation time and bias [12,94]. The next generation of sequencing technology, known as TGS, is being designed to investigate single molecules more accurately, such as nucleic acid sequencing of single cancer cells. Three TGS systems are currently available and have shown promising results from application; several TGS methods are available, such as the Single-Molecule Real-Time (SMRT) sequencing technology, the Nanopore technology and the DNA transistor technology [94].

SMRT sequencing is based on visualization of the second-strand of DNA as it is synthesized and is based on the zero-mode wave guide measurement as it is affected by the delayed incorporation of a florescent-labeled nucleotide by an anchored polymerase [94,95]. In Nanopore sequencing, a polymerase is used to unzip the dsDNA and feed the ssDNA through a charged nanopore; as a single nucleotide goes through the pore, it disrupts the current in a characteristic manner so that the molecule (and order of those molecules, such as nucleotides in a sequence) can be identified by its pattern of current disruption [94,96]. The DNA transistor technology also uses nanometer-sized pores and reads the individual nucleotides of ssDNA nucleotides according to their unique electronic signatures when passing through the pores with an appropriate modulation [94].

Each of these TGS technologies has advantages over the NGS technologies, but none is perfect. The primary advantage of all these technologies is low bias in genome coverage, even in high-GC regions, because they can interrogate the single nucleotides without the bias induced by amplification [12,93,97]. Unfortunately, SMRT suffers from low throughput and high raw read error rates [94], while the Nanopore technology struggles with regulating the polymerase-mediated passage of DNA through the pore [94,98]. The DNA transistor technology has the advantage of rapidity, but it cannot yet accurately regulate the miscibility of signals from neighboring nucleotides [94,99,100].

Single-cell epigenomic detection

The most widely used methods to detect epigenetic profiles are chromatin immunoprecipitation-sequencing (ChIP-Seq) and bisulfite sequencing (BS-seq) [101]. ChIP, an enrichment-based approach, is a well-established immunoreactive-based method that has been widely used to detect (and measure, when coupled with quantitative real-time PCR) epigenetic changes, and when coupled with sequencing can detect genome-wide histone modifications [101,102]. BS-seq can capture the DNA methylation status of cytosine residues (5-MeC) throughout the genome by performing double-stranded DNA degeneration followed by bisulfate-induced conversion of unmethylated cytosine to uracil and then performing PCR amplification and data analysis [6,103,104]. Currently, BS-seq is considered the gold standard for methylation analysis [102], but both approaches are disadvantaged by their inherent limitation of only being able to capture one feature at a time (and having high monetary and labor costs for not being high-throughput) [101].

Several other novel methods for detecting epigenetic marks have been reported but are not yet well established. Among them, SCAN (the acronym for "Single Chromatin molecule Analysis in Nanochannels") is a nanofluidic-based approach that can capture more than one epigenetic feature of ssDNA and associated chromatin molecules [105], while iChmo (the acronym for "Imaging of a Combination of Histone Modifications") uses a dual hybridization reaction of antigen–antibodies to capture two histone modification marks simultaneously [106].

Applications of single cell sequencing

Discovery of cancer-related genetic aberrations

Oncogenesis is a disease process manifesting from multigene aberrations, which ultimately causes the transition of a normal cell phenotype to an abnormal cell phenotype, such as that of a cancer cell. Identifying the specific genetic aberrations underlying different cancer types and different oncological stages (i.e. tumorigenesis, metastasis, chemotherapy resistance) may help us to better understand the mechanisms of tumor development and persistence as well as to provide molecular targets for novel strategies of cancer prevention and therapy.

Recently, a whole-exome sequencing project involving 22 gastric cancer samples was performed [107], and 7036 somatic mutations were measured, including 4653 located in coding regions or essential splice sites. All except 59 of the affected genes harbored protein-altering somatic mutations, and 2831 new genes were detected as mutated in the gastric cancer condition. Moreover, 20 new genes were identified as having potential for promoting gastric cancer (gastric cancer-driver genes), with low false discovery rate (FDR ≤ 0.2). The authors also performed Sanger sequencing of one of these new genes, ARID1A, in specimens from 109 gastric cancer cases and found that 39 of these carried either a single or double mutation; in addition, they determined that the mutation rate of ARID1A was

different among the different molecular subtypes of gastric cancers examined [107].

Transcriptome and epigenome sequencing data are also important for uncovering cancer-related aberrations. A transcriptome analysis of colorectal cancer CTCs revealed that the CTCs may represent a state of dormancy since the cancer-related down-regulated transcripts were associated with genes that play a role in “stemness”, such as CD166 (ALCAM) and CD26 (DPP4). In addition, some of the genes that showed lower expression in the colorectal cancer CTCs are known to play roles in cell–cell contact and motility. The only upregulated gene found in these colorectal cancer CTCs was CD47, which has been proposed to help cancer cells escape immune-killing [41]. In addition, several genes that play important roles in normal cellular functions, such as Septin 9 (SEPT9), POU4F1 and NOL4 (nucleolar protein 4), have been demonstrated as having cancer-related epigenetic modifications, such as those at the CpG islands, which may contribute to the tumorigenic mechanism [108–111].

Reconstructing cancer cell lineage trees

Cancer evolution usually involves the accumulation of genetic aberrations, which act to disrupt the function of normal cells and promote the transition to a carcinoma phenotype [112]. Capturing these genetic aberrations, both as spatial and temporal data, will allow for the reconstruction of cancer cell lineage trees so that the evolutionary process of cancer can be analyzed and can provide insights into the diagnosis and treatment of this disease.

Various models have been developed to construct such cancer cell lineage trees using sequencing data, including that from single-cell sequencing. The minimal spanning tree method can accommodate sequencing errors to order all the mutations sites from the sequencing data. In the mutation tree itself, the early mutation sites are placed in the root, while other mutations are placed in different layers, with the branching architecture depending on the relative distances. In this approach, a novel algorithm model is used to estimate the time it took for a proportion of mutations to have accumulated, from the first mutation to the most recent common ancestor mutation. Unfortunately, this bioinformatics method is limited by the precision of the currently available single-cell sequencing technologies [113].

The PhyloSub statistical model divides cancer cells into several major subclonal lineages and estimates their ancestry according to the presence of single-nucleotide variants (SNVs) and allele frequencies, as determined from sequencing data. When the clonal frequencies are insufficient to reconstruct the phylogeny, the Markov Chain Monte Carlo (MCMC) method can be used to capture the “partial order plot” for inferring the phylogenetic relationships. The PhyloSub model is also applicable to research of the various somatic mutations, such as small insertions and deletions; however, the hindrance to constructing cancer cell lineage trees is the unknown of how many cells are involved in the research, so it can cover all the single nucleotide mutations [6].

Although the research based on tumor cell lineage tree construction has provided some very important insights into the processes of cancer development and progression, the power of this approach remains limited by the currently available algorithms for analyzing sequencing data and the high error rate generated during the sequencing process.

Cancer diagnosis and therapy

The accumulated genetic aberrations that underline tumorigenesis can also mediate the tumor cell’s response to therapy [114]. Single-cell sequencing technology may identify markers related to cancer diagnosis as well as molecules that will act as targets for individualized therapy [115,116].

Various cancer-related mutations are already in clinical used as biomarkers for the detection of cancer in general and diagnosis of specific cancer types, including the BRAF mutations of melanoma, the KRAS mutations of colon cancer, and the epidermal growth factor receptor (EGFR) mutations of non-small-cell lung cancer [116]. However, these biomarkers are only sensitive enough for use as indicators and definitive diagnoses require more detailed and comprehensive analyses, as no single mutation to date has been absolutely correlated with a specific cancer type and stage or prognosis. For example, while the NOL4 (Nucleolar protein 4) gene and the LHFPL4 (lipoma HMGIC fusion partner-like protein 4) gene have been observed as having high-frequency methylation in many cervical cancer cases [105,110,111], the smaller percentage of cases that do not show this profile would be missed by a diagnostic approach relying solely on these markers.

It has been proposed that the cancer phenotype may be restored to a normal phenotype by modifying the cancer-related genetic aberrations. Various targeting drugs capable of inhibiting the function of the abnormal gene are available and some have been used in clinic trials as individualized cancer treatments, such as the histone deacetylase inhibitors for hematological malignancies, the BRAF inhibitors for melanoma, and the PARP inhibitors for breast cancer [116–119]. These targeted drugs do not appear to be a general answer to all cancer types, as they have been shown to exert different effects in different or even the same kinds of cancers, because the cancers can hardly be the result of the same genetic aberrations [118]. Single-cell sequencing to uncover the distinctive panels of genetic aberrations in various cancer types and stages will likely help improve this type of therapeutic approach.

Conclusion and future perspectives

The cell is the functional unit of an organism; in multi-celled organisms, such as humans, the heterogeneity that inherently exists among its individual cells involves both the phenotype and genotype levels. Single-cell sequencing technology provides an unprecedented methodology for cancer research, allowing for the analysis that is present among heterogeneous cancer cell populations at the single cell level. The data from single-cell sequencing will advance our understanding of cancer pathogenesis, so that improved methods of diagnosis and treatment may be developed; in addition, these data will act as foundational information for further research studies, such as those to reconstruct the cancer cell evolution lineage.

The tools that support single-cell sequencing, including reagents, equipment and bioinformatics algorithms, are continually being refined to increase the power of this platform. The recent technological advances have allowed for both DNA and RNA to be extracted simultaneously from a single sample [9] and have decreased the amplification bias (such as with the newly introduced MIDAS approach by reducing the reaction volume of MDA [44]).The ongoing efforts of the sequencing research community at large will almost certainly culminate in innovative approaches for the third generation of sequencing technology, especially for the single-molecule sequencing approach [94]. Moreover, these efforts will include generation of more sophisticated deficient algorithms to process and interpret the single-cell sequencing data [113,120–122].

The creation of an optimal single-cell sequencing technology must overcome the numerous limitations that currently exist for the various tools involved. Significant challenges are still posed by contamination [21], amplification bias (for both DNA and RNA) [11,14], efficiency of sequencing [94], and algorithms advanced enough to link the sequencing data [113,120]. Improved single-cell sequencing technologies will decrease or even eliminate these challenges so that more precise sequencing data will be obtained and the relationship between sequencing data and pathological features will be convenient and accurate.

Acknowledgements

This research was supported by grants from the National High Technology Research and Development Program of China (863 program) (No. 2012AA02A407) and the National Basic Research Program of China (973 program) (No. 2013CB531500). We would like to thank Dr. Jennifer C van Velkinburgh (van Velkinburgh Initiative for Collaboratory BioMedical Research, Santa Fe, NM, USA) for helpful discussions and for polishing the grammatical presentation of the manuscript.

Conflict of interest

The authors declare that they have no conflicts of interests.

References

[1] T. Shibata, Cancer genomics and pathology: all together now, Pathol. Int. 62 (2012) 647–659.
[2] C. Hiley, E.C. de Bruin, N. McGranahan, C. Swanton, Deciphering intratumor heterogeneity and temporal acquisition of driver events to refine precision medicine, Genome Biol. 15 (2014) 453.
[3] P. Van Loo, T. Voet, Single cell analysis of cancer genomes, Curr. Opin. Genet. Dev. 24 (2014) 82–91.
[4] S.C. Ren, M. Qu, Y.H. Sun, Investigating intratumour heterogeneity by single-cell sequencing, Asian J. Androl. 15 (2013) 729–734.
[5] L.R. Patel, M. Nykter, K. Chen, W. Zhang, Cancer genome sequencing: understanding malignancy as a disease of the genome, its conformation, and its evolution, Cancer Lett. 340 (2013) 152–160.
[6] E. Shapiro, T. Biezuner, S. Linnarsson, Single-cell sequencing-based technologies will revolutionize whole-organism science, Nat. Rev. Genet. 14 (2013) 618–630.
[7] J. Eberwine, J.Y. Sul, T. Bartfai, J. Kim, The promise of single-cell sequencing, Nat. Methods 11 (2014) 25–27.
[8] T. Nawy, Single-cell sequencing, Nat. Methods 11 (2013) 18.
[9] L. Strotman, R. O'Connell, B.P. Casavant, S.M. Berry, J.M. Sperger, J.M. Lang, et al., Selective nucleic acid removal via exclusion (SNARE): capturing mRNA and DNA from a single sample, Anal. Chem. 85 (2013) 9764–9770.
[10] C.F. de Bourcy, I. De Vlaminck, J.N. Kanbar, J. Wang, C. Gawad, S.R. Quake, A quantitative comparison of single-cell whole genome amplification methods, PLoS ONE 9 (2014) e105585.
[11] L. Ning, G. Liu, G. Li, Y. Hou, Y. Tong, J. He, Current challenges in the bioinformatics of single cell genomics, Front. Oncol. 4 (2014) 7.
[12] S. Koren, G.P. Harhay, T.P. Smith, J.L. Bono, D.M. Harhay, S.D. McVey, et al., Reducing assembly complexity of microbial genomes with single-molecule sequencing, Genome Biol. 14 (2013) R101.
[13] M. Kim, K.H. Lee, S.W. Yoon, B.S. Kim, J. Chun, H. Yi, Analytical tools and databases for metagenomics in the next-generation sequencing era, Genomics Inform. 11 (2013) 102–113.
[14] I.C. Macaulay, T. Voet, Single cell genomics: advances and future perspectives, PLoS Genet. 10 (2014) e1004126.
[15] L.V. Hooper, Laser microdissection: exploring host-bacterial encounters at the front lines, Curr. Opin. Microbiol. 7 (2004) 290–295.
[16] C.P. Satori, V. Kostal, E.A. Arriaga, Review on recent advances in the analysis of isolated organelles, Anal. Chim. Acta 753 (2012) 8–18.
[17] P. Sluka, L. O'Donnell, R.I. McLachlan, P.G. Stanton, Application of laser-capture microdissection to analysis of gene expression in the testis, Prog. Histochem. Cytochem. 42 (2008) 173–201.
[18] L. Cheng, S. Zhang, G.T. MacLennan, S.R. Williamson, D.D. Davidson, M. Wang, et al., Laser-assisted microdissection in translational research: theory, technical considerations, and future applications, Appl. Immunohistochem. Mol. Morphol. 21 (2013) 31–47.
[19] M.R. Emmert-Buck, R.F. Bonner, P.D. Smith, R.F. Chuaqui, Z. Zhuang, S.R. Goldstein, et al., Laser capture microdissection, Science 274 (1996) 998–1001.
[20] M. Vandewoestyne, K. Goossens, C. Burvenich, A. Van Soom, L. Peelman, D. Deforce, Laser capture microdissection: should an ultraviolet or infrared laser be used?, Anal. Biochem. 439 (2013) 88–98.
[21] P.C. Blainey, The future is now: single-cell genomics of bacteria and archaea, FEMS Microbiol. Rev. 37 (2013) 407–427.
[22] J. Ankarklev, S.G. Svard, M. Lebbad, Allelic sequence heterozygosity in single Giardia parasites, BMC Microbiol. 12 (2012) 65.
[23] T. Ishoey, T. Woyke, R. Stepanauskas, M. Novotny, R.S. Lasken, Genomic sequencing of single microbial cells from environmental samples, Curr. Opin. Microbiol. 11 (2008) 198–204.
[24] T. Kalisky, P. Blainey, S.R. Quake, Genomic analysis at the single-cell level, Annu. Rev. Genet. 45 (2011) 431–445.
[25] K.R. Willison, D.R. Klug, Quantitative single cell and single molecule proteomics for clinical studies, Curr. Opin. Biotechnol. 24 (2013) 745–751.
[26] M.J. Smith, A.C. Culhane, M. Donovan, J.C. Coffey, B.D. Barry, M.A. Kelly, et al., Analysis of differential gene expression in colorectal cancer and stroma using fluorescence-activated cell sorting purification, Br. J. Cancer 100 (2009) 1452–1464.
[27] Z.B. Bjornson, G.P. Nolan, W.J. Fantl, Single-cell mass cytometry for analysis of immune system functional states, Curr. Opin. Immunol. 25 (2013) 484–494.
[28] J. Autebert, B. Coudert, F.C. Bidard, J.Y. Pierga, S. Descroix, L. Malaquin, et al., Microfluidic: an innovative tool for efficient cell sorting, Methods 57 (2012) 297–307.
[29] J. Chen, J. Li, Y. Sun, Microfluidic approaches for cancer cell detection, characterization, and separation, Lab Chip 12 (2012) 1753–1767.
[30] Z.Y. Zhang, H.Y. Ge, Micrometastasis in gastric cancer, Cancer Lett. 336 (2013) 34–45.
[31] S. Yilmaz, A.K. Singh, Single cell genome sequencing, Curr. Opin. Biotechnol. 23 (2012) 437–443.
[32] E. Sharif, J. Kiely, R. Luxton, Novel immunoassay technique for rapid measurement of intracellular proteins using paramagnetic particles, J. Immunol. Methods 388 (2013) 78–85.
[33] A. Stahlberg, M. Bengtsson, Single-cell gene expression profiling using reverse transcription quantitative real-time PCR, Methods 50 (2010) 282–288.
[34] R. Stepanauskas, Single cell genomics: an individual look at microbes, Curr. Opin. Microbiol. 15 (2012) 613–620.
[35] R.B. Brown, J. Audet, Current techniques for single-cell lysis, J. R. Soc. Interface 5 (Suppl. 2) (2008) S131–S138.
[36] P.Y. Chiou, A.T. Ohta, M.C. Wu, Massively parallel manipulation of single cells and microparticles using optical images, Nature 436 (2005) 370–372.
[37] M.S. Rook, S.M. Delach, G. Deyneko, A. Worlock, J.L. Wolfe, Whole genome amplification of DNA from laser capture-microdissected tissue for high-throughput single nucleotide polymorphism and short tandem repeat genotyping, Am. J. Pathol. 164 (2004) 23–33.
[38] T. Baslan, J. Kendall, L. Rodgers, H. Cox, M. Riggs, A. Stepansky, et al., Genome-wide copy number analysis of single cells, Nat. Protoc. 7 (2012) 1024–1041.
[39] C. Zhang, C. Zhang, S. Chen, X. Yin, X. Pan, G. Lin, et al., A single cell level based method for copy number variation analysis by low coverage massively parallel sequencing, PLoS ONE 8 (2013) e54236.
[40] A. Imle, B. Polzer, S. Alexander, C.A. Klein, P. Friedl, Genomic instability of micronucleated cells revealed by single-cell comparative genomic hybridization, Cytometry A. 75 (2009) 562–568.
[41] G. Steinert, S. Scholch, T. Niemietz, N. Iwata, S.A. Garcia, B. Behrens, et al., Immune escape and survival mechanisms in circulating tumor cells of colorectal cancer, Cancer Res. 74 (2014) 1694–1704.
[42] J.F. Swennenhuis, J. Reumers, K. Thys, J. Aerssens, L.W. Terstappen, Efficiency of whole genome amplification of single circulating tumor cells enriched by CellSearch and sorted by FACS, Genome Med. 5 (2013) 106.
[43] N.E. Potter, L. Ermini, E. Papaemmanuil, G. Cazzaniga, G. Vijayaraghavan, I. Titley, et al., Single-cell mutational profiling and clonal phylogeny in cancer, Genome Res. 23 (2013) 2115–2125.
[44] J. Gole, A. Gore, A. Richards, Y.J. Chiu, H.L. Fung, D. Bushman, et al., Massively parallel polymerase cloning and genome sequencing of single cells using nanoliter microwells, Nat. Biotechnol. 31 (2013) 1126–1132.
[45] Y. Yang, H.R. Hebron, J. Hang, A method for preparing DNA sequencing templates using a DNA-binding microplate, J. Biomol. Tech. 20 (2009) 165–171.
[46] S.C. Tan, B.C. Yiap, DNA, RNA, and protein extraction: the past and the present, J. Biomed. Biotechnol. 2009 (2009) 574398.
[47] L. Melchor, A. Brioli, C.P. Wardell, A. Murison, N.E. Potter, M.F. Kaiser, et al., Single-cell genetic analysis reveals the composition of initiating clones and phylogenetic patterns of branching and parallel evolution in myeloma, Leukemia 28 (2014) 1705–1715.
[48] J. Gunawardana, F.C. Chan, A. Telenius, B. Woolcock, R. Kridel, K.L. Tan, et al., Recurrent somatic mutations of PTPN1 in primary mediastinal B cell lymphoma and Hodgkin lymphoma, Nat. Genet. 46 (2014) 329–335.
[49] M.N. Fletcher, M.A. Castro, X. Wang, I. de Santiago, M. O'Reilly, S.F. Chin, et al., Master regulators of FGFR2 signalling and breast cancer risk, Nat. Commun. 4 (2013) 2464.
[50] J. Kim, M. Johnson, P. Hill, B.K. Gale, Microfluidic sample preparation: cell lysis and nucleic acid purification, Integr. Biol. 1 (2009) 574–586.
[51] X. Wang, S. Takebayashi, E. Bernardin, D.M. Gilbert, R. Chella, J. Guan, Microfluidic extraction and stretching of chromosomal DNA from single cell nuclei for DNA fluorescence in situ hybridization, Biomed. Microdevices 14 (2012) 443–451.
[52] A.M. Streets, Y. Huang, Microfluidics for biological measurements with single-molecule resolution, Curr. Opin. Biotechnol. 25 (2014) 69–77.
[53] R. Zhang, H.Q. Gong, X. Zeng, C. Lou, C. Sze, A microfluidic liquid phase nucleic acid purification chip to selectively isolate DNA or RNA from low copy/single bacterial cells in minute sample volume followed by direct on-chip quantitative PCR assay, Anal. Chem. 85 (2013) 1484–1491.
[54] X.T. Sun, M. Liu, Z.R. Xu, Microfluidic fabrication of multifunctional particles and their analytical applications, Talanta 121 (2014) 163–177.
[55] H. Telenius, N.P. Carter, C.E. Bebb, M. Nordenskjold, B.A. Ponder, A. Tunnacliffe, Degenerate oligonucleotide-primed PCR: general amplification of target DNA by a single degenerate primer, Genomics 13 (1992) 718–725.
[56] W. Peng, H. Takabayashi, K. Ikawa, Whole genome amplification from single cells in preimplantation genetic diagnosis and prenatal diagnosis, Eur. J. Obstet. Gynecol. Reprod. Biol. 131 (2007) 13–20.
[57] L. Zhang, X. Cui, K. Schmitt, R. Hubert, W. Navidi, N. Arnheim, Whole genome amplification from a single cell: implications for genetic analysis, Proc. Natl. Acad. Sci. U.S.A. 89 (1992) 5847–5851.
[58] W. Dietmaier, A. Hartmann, S. Wallinger, E. Heinmoller, T. Kerner, E. Endl, et al., Multiple mutation analyses in single tumor cells with improved whole genome amplification, Am. J. Pathol. 154 (1999) 83–95.

[59] R.S. Lasken, Single-cell sequencing in its prime, Nat. Biotechnol. 31 (2013) 211–212.
[60] C. Zong, S. Lu, A.R. Chapman, X.S. Xie, Genome-wide detection of single-nucleotide and copy-number variations of a single human cell, Science 338 (2012) 1622–1626.
[61] X. Ni, M. Zhuo, Z. Su, J. Duan, Y. Gao, Z. Wang, et al., Reproducible copy number variation patterns among single circulating tumor cells of lung cancer patients, Proc. Natl. Acad. Sci. U.S.A. 110 (2013) 21083–21088.
[62] P.M. Lizardi, X. Huang, Z. Zhu, P. Bray-Ward, D.C. Thomas, D.C. Ward, Mutation detection and single-molecule counting using isothermal rolling-circle amplification, Nat. Genet. 19 (1998) 225–232.
[63] C. Spits, C. Le Caignec, M. De Rycke, L. Van Haute, A. Van Steirteghem, I. Liebaers, et al., Whole-genome multiple displacement amplification from single cells, Nat. Protoc. 1 (2006) 1965–1970.
[64] J. Martin, A. Cervero, P. Mir, J.A. Martinez-Conejero, A. Pellicer, C. Simon, The impact of next-generation sequencing technology on preimplantation genetic diagnosis and screening, Fertil. Steril. 99 (2013) 1054–1061.e3.
[65] L. Ning, G. Liu, G. Li, Y. Hou, Y. Tong, J. He, Current challenges in the bioinformatics of single cell, Front. Oncol. 4 (2014) 7.
[66] C.A. Hutchison 3rd, H.O. Smith, C. Pfannkoch, J.C. Venter, Cell-free cloning using phi29 DNA polymerase, Proc. Natl. Acad. Sci. U.S.A. 102 (2005) 17332–17336.
[67] Y. Marcy, T. Ishoey, R.S. Lasken, T.B. Stockwell, B.P. Walenz, A.L. Halpern, et al., Nanoliter reactors improve multiple displacement amplification of genomes from single cells, PLoS Genet. 3 (2007) 1702–1708.
[68] X. Pan, Single cell analysis: from technology to biology and medicine, Single Cell Biol. 3 (2014).
[69] N. Liu, L. Liu, X. Pan, Single-cell analysis of the transcriptome and its application in the characterization of stem cells and early embryos, Cell. Mol. Life Sci. 71 (2014) 2707–2715.
[70] F. Tang, C. Barbacioru, Y. Wang, E. Nordman, C. Lee, N. Xu, et al., mRNA-Seq whole-transcriptome analysis of a single cell, Nat. Methods 6 (2009) 377–382.
[71] F. Tang, K. Lao, M.A. Surani, Development and applications of single-cell transcriptome analysis, Nat. Methods 8 (2011) S6–S11.
[72] Y.Y. Zhu, E.M. Machleder, A. Chenchik, R. Li, P.D. Siebert, Reverse transcriptase template switching: a SMART approach for full-length cDNA library construction, Biotechniques 30 (2001) 892–897.
[73] S. Islam, U. Kjallquist, A. Moliner, P. Zajac, J.B. Fan, P. Lonnerberg, et al., Highly multiplexed and strand-specific single-cell RNA 5' end sequencing, Nat. Protoc. 7 (2012) 813–828.
[74] D. Hebenstreit, Methods, challenges and potentials of single cell RNA-seq, Biology 1 (2012) 658–667.
[75] T. Hashimshony, F. Wagner, N. Sher, I. Yanai, CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification, Cell Rep. 2 (2012) 666–673.
[76] Y. Sasagawa, I. Nikaido, T. Hayashi, H. Danno, K.D. Uno, T. Imai, et al., Quartz-Seq: a highly reproducible and sensitive single-cell RNA sequencing method, reveals non-genetic gene-expression heterogeneity, Genome Biol. 14 (2013) R31.
[77] J. Eberwine, H. Yeh, K. Miyashiro, Y. Cao, S. Nair, R. Finnell, et al., Analysis of gene expression in single live neurons, Proc. Natl. Acad. Sci. U.S.A. 89 (1992) 3010–3014.
[78] X. Pan, R.E. Durrett, H. Zhu, Y. Tanaka, Y. Li, X. Zi, et al., Two methods for full-length RNA sequencing for low quantities of cells and single cells, Proc. Natl. Acad. Sci. U.S.A. 110 (2013) 594–599.
[79] S. Islam, U. Kjallquist, A. Moliner, P. Zajac, J.B. Fan, P. Lonnerberg, et al., Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq, Genome Res. 21 (2011) 1160–1167.
[80] V. Bhargava, S.R. Head, P. Ordoukhanian, M. Mercola, S. Subramaniam, Technical variations in low-input RNA-seq methodologies, Sci. Rep. 4 (2014) 3678.
[81] D. Ramskold, S. Luo, Y.C. Wang, R. Li, Q. Deng, O.R. Faridani, et al., Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells, Nat. Biotechnol. 30 (2012) 777–782.
[82] D. Ramskold, E.T. Wang, C.B. Burge, R. Sandberg, An abundance of ubiquitously expressed genes revealed by tissue transcriptome sequence data, PLoS Comput. Biol. 5 (2009) e1000598.
[83] Z. Zhang, W.E. Theurkauf, Z. Weng, P.D. Zamore, Strand-specific libraries for high throughput RNA sequencing (RNA-Seq) prepared without poly(A) selection, Silence 3 (2012) 9.
[84] S. Picelli, O.R. Faridani, A.K. Bjorklund, G. Winberg, S. Sagasser, R. Sandberg, Full-length RNA-seq from single cells using Smart-seq2, Nat. Protoc. 9 (2014) 171–181.
[85] K.M. Kim, J.H. Park, D. Bhattacharya, H.S. Yoon, Applications of next-generation sequencing to unravelling the evolutionary history of algae, Int. J. Syst. Evol. Microbiol. 64 (2014) 333–345.
[86] M. Jessri, C.S. Farah, Next generation sequencing and its application in deciphering head and neck cancer, Oral Oncol. 50 (2014) 247–253.
[87] D. Ulahannan, M.B. Kovac, P.J. Mulholland, J.B. Cazier, I. Tomlinson, Technical and implementation issues in using next-generation sequencing of cancers in clinical practice, Br. J. Cancer 109 (2013) 827–835.
[88] E.R. Mardis, Next-generation DNA sequencing methods, Annu. Rev. Genomics Hum. Genet. 9 (2008) 387–402.
[89] M. Jessri, C.S. Farah, Harnessing massively parallel sequencing in personalized head and neck oncology, J. Dent. Res. 93 (2014) 437–444.
[90] J.S. Ross, M. Cronin, Whole cancer genome sequencing by next-generation methods, Am. J. Clin. Pathol. 136 (2011) 527–539.
[91] S. Wang, J. Xing, A primer for disease gene prioritization using next-generation sequencing data, Genomics Inform. 11 (2013) 191–199.
[92] F. Chang, M.M. Li, Clinical application of amplicon-based next-generation sequencing in cancer, Cancer Genet. 206 (2013) 413–419.
[93] S.C. Shin, H. Ahn do, S.J. Kim, H. Lee, T.J. Oh, J.E. Lee, et al., Advantages of single-molecule real-time sequencing in high-GC content genomes, PLoS ONE 8 (2013) e68824.
[94] E.E. Schadt, S. Turner, A. Kasarskis, A window into third-generation sequencing, Hum. Mol. Genet. 19 (2010) R227–R240.
[95] J. Eid, A. Fehr, J. Gray, K. Luong, J. Lyle, G. Otto, et al., Real-time DNA sequencing from single polymerase molecules, Science 323 (2009) 133–138.
[96] I.M. Derringtona, T.Z. Butler, M.D. Collins, E. Manrao, M. Pavlenok, M. Niederweis, et al., Nanopore DNA sequencing with MspA, Proc. Natl Acad. Sci. U.S.A. 107 (2010) 16060–16065.
[97] A. Hendel, E.J. Kildebeck, E.J. Fine, J.T. Clark, N. Punjya, V. Sebastiano, et al., Quantifying genome-editing outcomes at endogenous loci with SMRT sequencing, Cell Rep. 7 (2014) 293–305.
[98] R.D. Maitra, J. Kim, W.B. Dunbar, Recent advances in nanopore sequencing, Electrophoresis 33 (2012) 3418–3428.
[99] M. Krems, M. Zwolak, Y.V. Pershin, M. Di Ventra, Effect of noise on DNA sequencing via transverse electronic transport, Biophys. J. 97 (2009) 1990–1996.
[100] B.Q. Luan, H.B. Peng, S. Polonsky, S. Rossnagel, G. Stolovitzky, G. Martyna, Base-by-base ratcheting of single stranded DNA through a solid-state nanopore, Phys. Rev. Lett. 104 (2010) 238103.
[101] J.A. Hagarman, P.D. Soloway, Epigenomics: one molecule at a time, Cell Cycle 12 (2013) 3451–3452.
[102] K. Mensaert, S. Denil, G. Trooskens, W. Van Criekinge, O. Thas, T. De Meyer, Next-generation technologies and data analytical approaches for epigenomics, Environ. Mol. Mutagen. 55 (2014) 155–170.
[103] K. Patterson, L. Molloy, W. Qu, S. Clark, DNA methylation: bisulphite modification and analysis, J. Vis. Exp. e3170 (2011) 1–10.
[104] B.R. Cipriany, R. Zhao, P.J. Murphy, S.L. Levy, C.P. Tan, H.G. Craighead, et al., Single molecule epigenetic analysis in a nanofluidic channel, Anal. Chem. 82 (2010) 2480–2487.
[105] P.J. Murphy, B.R. Cipriany, C.B. Wallin, C.Y. Ju, K. Szeto, J.A. Hagarman, et al., Single-molecule analysis of combinatorial epigenomic states in normal and tumor cells, Proc. Natl. Acad. Sci. U.S.A. 110 (2013) 7772–7777.
[106] N. Hattori, T. Niwa, K. Kimura, K. Helin, T. Ushijima, Visualization of multivalent histone modification in a single cell reveals highly concerted epigenetic changes on differentiation of embryonic stem cells, Nucleic Acids Res. 41 (2013) 7231–7239.
[107] K. Wang, J. Kan, S.T. Yuen, S.T. Shi, K.M. Chu, S. Law, et al., Exome sequencing identifies frequent mutation of ARID1A in molecular subtypes of gastric cancer, Nat. Genet. 43 (2011) 1219–1223.
[108] F. Gao, G. Ji, Z. Gao, X. Han, M. Ye, Z. Yuan, et al., Direct ChIP-bisulfite sequencing reveals a role of H3K27me3 mediating aberrant hypermethylation of promoter CpG islands in cancer cells, Genomics 103 (2014) 204–210.
[109] R. Wasserkort, A. Kalmar, G. Valcz, S. Spisak, M. Krispin, K. Toth, et al., Aberrant septin 9 DNA methylation in colorectal cancer is restricted to a single CpG island, BMC Cancer 13 (2013) 398.
[110] M. Faryna, C. Konermann, S. Aulmann, J.L. Bermejo, M. Brugger, S. Diederichs, et al., Genome-wide methylation screen in low-grade breast cancer identifies novel epigenetically altered genes as potential biomarkers for tumor diagnosis, FASEB J. 26 (2012) 4937–4950.
[111] S.S. Wang, D.J. Smiraglia, Y.Z. Wu, S. Ghosh, J.S. Rader, K.R. Cho, et al., Identification of novel methylation markers in cervical cancer using restriction landmark genomic scanning, Cancer Res. 68 (2008) 2489–2497.
[112] D.E. Newburger, D. Kashef-Haghighi, Z. Weng, R. Salari, R.T. Sweeney, A.L. Brunner, et al., Genome evolution during progression to breast cancer, Genome Res. 23 (2013) 1097–1108.
[113] K.I. Kim, R. Simon, Using single cell sequencing data to model the evolutionary history of a tumor, BMC Bioinformatics 15 (2014) 27.
[114] J.T. Dudley, R. Chen, A.J. Butte, Matching cancer genomes to established cell lines for personalized oncology, Pac. Symp. Biocomput. (2011) 243–252.
[115] K. Mabert, M. Cojoc, C. Peitzsch, I. Kurth, S. Souchelnytskyi, A. Dubrovska, Cancer biomarker discovery: current status and future perspectives, Int. J. Radiat. Biol. 90 (2014) 659–677.
[116] N. Normanno, A.M. Rachiglio, C. Roma, F. Fenizia, C. Esposito, R. Pasquale, et al., Molecular diagnostics and personalized medicine in oncology: challenges and opportunities, J. Cell. Biochem. 114 (2013) 514–524.
[117] L. Zhang, Y. Han, Q. Jiang, C. Wang, X. Chen, X. Li, et al., Trend of histone deacetylase inhibitors in cancer therapy: isoform selectivity or multitargeted strategy, Med. Res. Rev. 35 (2015) 63–84.
[118] A. Prahallad, C. Sun, S. Huang, F. Di Nicolantonio, R. Salazar, D. Zecchin, et al., Unresponsiveness of colon cancer to BRAF(V600E) inhibition through feedback activation of EGFR, Nature 483 (2012) 100–103.
[119] S. Vinayak, J.M. Ford, PARP inhibitors for the treatment and prevention of breast cancer, Curr. Breast Cancer Rep. 2 (2010) 190–197.
[120] W. Jiao, S. Vembu, A.G. Deshwar, L. Stein, Q. Morris, Inferring clonal evolution of tumors from single nucleotide somatic mutations, BMC Bioinformatics 15 (2014) 35.
[121] Y. Hou, L. Song, P. Zhu, B. Zhang, Y. Tao, X. Xu, et al., Single-cell exome sequencing and monoclonal evolution of a JAK2-negative myeloproliferative neoplasm, Cell 148 (2012) 873–885.
[122] Y. Li, X. Xu, L. Song, Y. Hou, Z. Li, S. Tsang, et al., Single-cell sequencing analysis characterizes common and cell-lineage-specific mutations in a muscle-invasive bladder cancer, GigaScience 1 (2012) 12.