REVIEW

Progress in Single Cell Sequencing Technology

Qicai Ma1,2 Wenli Wu1 Na Ye1,2 Xingdong Wang1,2 Ping Yan1,2 Heping Pan1*

1. College of Life Science and Engineering, Northwest University For Nationalities, Lanzhou, Gansu, 730030, China
2. Lanzhou Institute of Animal Husbandry and Veterinary Medicine, Chinese Academy of Agricultural Sciences, Lanzhou, Gansu, 730050; China

ARTICLE INFO

Article history
Received: 1 April 2019
Accepted: 1 April 2019
Published: 30 April 2019

Keywords:
Single cell isolation
Single cell sequencing
Whole genome
Transcriptome

ABSTRACT

Cells are the basic unit of life structure and life activities. Because of the complex micro-environment of cells, the content of components that play a key role is relatively small, so single-cell analysis is extremely challenging. In recent years, single-cell sequencing technology has been developed and matured. Single-cell sequencing can reveal the composition and physiological diversity of cells, and the existing single-cell separation technology, single-cell whole genome amplification technology, single cell transcriptome sequencing and the principles and applications of cell whole transcriptome amplification technology and single cell transcriptome sequencing are summarized and summarized.

1. Introduction

Cells are the basic unit of life activity, with complete vitality. Most living organisms are multicellular. However, single-cell living organisms can carry out life activities separately because cells have a complete set of self-regulating devices. In multicellular organisms, although the cells as a whole coordinate and control the function of each cell, each cell has a relatively independent set of metabolic systems, and each cell body performs specific functions in a cell-based unit. Only when the cells have a suitable growth environment, each isolated single cell can grow and multiply in vitro. The cell is the basic functional unit of life, and indirectly explains the laws of metabolism and regulation of the body. Therefore, research on single cells has been receiving public attention.

Single-cell sequencing technology is the analysis of single-cell DNA or RNA, revealing the genetic structure and expression of individual cells, changes in cell growth and development, genetic material status, and heterogeneity between cells and cells. Development is slower than other disciplines. It plays an irreplaceable role in tumor, developmental biology, stem cell biology, microbiology, neuroscience and so on. Single cell sequencing separates, sequences, etc. cells in a series of steps. However, there are certain difficulties in isolating single cells, single cell genomes, and amplification of transcriptomes. This paper mainly summarizes and analyzes the research progress of single cell genomic amplification technology, transcriptome amplification technology and single cell separation technology and their advantages and disadvantages in single cell sequencing technology.

*Corresponding Author:
Heping Pan,
College of Life Science and Engineering, Northwest University For Nationalities, Lanzhou, Gansu, 730030, China;
Email: 1397934532@qq.com.
2. Single Cell Separation and Technology

Single cell sequencing analysis is the sequencing of the genes of the cells of interest extracted from the tissue. In recent years, many researchers have continued to explore and improve the separation method of single cells. Currently, the techniques commonly used for single cell separation are: gradient dilution, micromanipulation, fluorescence activated cell sorting, Microfluidic technology, laser capture microdissection and so on.

2.1. Gradient Dilution Method

Serial dilution is used to clone stem cells or premature cells from different tissues in vitro. The cell suspension is diluted by a series of different gradients until a single cell is obtained. By gradient dilution calculation, single cells cannot be sorted intuitively. This method requires less equipment and operation, but the success rate of separating single cells is not high, and it is easy to make mistakes.

2.2. Micromanipulation

Micromanipulation is mainly applied under the condition that the target cells are sparse. The single cells are visually selected by micromanipulator or combined microscope and mouth pipette, and further research is carried out. This method has little effect on the activity and state of cells, and has low cost, but it has time-consuming shortcomings, which is easy to cause RNA degradation, low flux, and more manpower, which is not conducive to large-scale application.

2.3. Fluorescence Activated Cell Sorting Technology

In recent years, the development of immunological techniques and flow cytometry has rapidly achieved the goal of sorting cells by means of flow cytometry, which is debugged by a computer. The principle is that the cells to be tested are placed in a dye-stained sample tube, and the cells are in a single-row cell column under the action of air pressure, and then the fluorescence signal excited by the laser is captured by the optical system, and then detected according to the optical system. The fluorescence signal intensity sorts the cells. This technology is a technique for rapid quantitative analysis and is the most commonly used single cell separation technique. The advantages of fluorescence activated cell sorting (FACS) are high accuracy and throughput, uniform standards, mature experimental methods, and the ability to analyze individual cells for rapid detection of cellular physiology. Chemical parameters; the disadvantages of the process are complex, the demand for raw materials is large, and the spatial resolution is insufficient.

2.4. Micro Control Flow Technology

Microfluidics are the separation and capture of individual cells within a micron-scale channel. Due to the closedness of the operation space, this technology can greatly reduce the pollution. Due to the particularity of the operating device, the concentration of the sample is high, and the consumption of the detection reagent is small, and the operation error can be reduced. Low, this completes the data with higher reliability and better reaction efficiency, and has a good application prospect.

2.5. Laser Capture Microdissection Technology

Laser capture microdissection (LCM) is the precise separation of single cells from a smear of frozen or paraffin-embedded tissue-slurry thermoplastic films using a UV (320-400 nm) laser under a microscope. The biggest advantage is that the target cells to be preserved can be captured without destroying the tissue structure, and the degree of automation is high, and the application prospect is wide. However, the disadvantage of this technique is that the instrument is expensive and expensive, and for some fixed and uncovered tissue sections, the visual resolution is limited, resulting in inaccurate cutting, and may also lose nuclear genetic material due to the operation time.

3. Single Cell Whole Genome Sequencing

Single-cell gene sequencing refers to the process of separating single cells from tissues by tissue separation technology, then extracting the single-cell DNA and sequencing them to obtain a single-cell genome-wide map, but sometimes the single-cell genome DNA. As low as 6pg, and the copy number of each gene is only two, and often occurs due to lack of band, multi-band, etc. Therefore, it is not feasible to use a common method to amplify a single-cell genome, in order to meet the second-generation sequencing. The minimum DNA content is required for efficient amplification of genomic DNA.

3.1 Single-cell Whole Genome Amplification Technology

Whole genome amplification (WGA) is a technique for efficient amplification and sequencing of single-cell genomes. The principle is to efficiently amplify the micro-genomic DNA of isolated single cells to obtain high-coverage single-cell genome technology. Used to reveal cell population differences and cell evolutionary
relationships. So far, PCR-based whole genome amplification methods include: degenerate oligonucleotide primed PCR, DOP-PCR, ligation mediated PRC, LM-PCR, primer extension pre-amplification, PEP. There are multiple displacement amplification based on isothermal reaction, but not based on PCR, MDA, multiple annealing and looping-based amplification cycles, MALBAC [26-28].

3.1.1 PCR-based Approach

Among them, in the PCR-based method, due to the inevitable error or uncontrollable external factors in PCR, the fragment size or secondary structure and GC content of DNA are affected, and the coverage of the genome is insufficient. Complete, and the sequence of amplification results may be false positive, false negative and non-specific amplification products, resulting in impure product, amplification bias [29-31].

3.1.2 MDA and MALBAC Technology

The most commonly used techniques today are multiple displacement amplification techniques, amplification techniques for multiple annealing loops. MDA technology is a constant-temperature amplification method that relies on the principle of strand displacement amplification, highly amplified DNA, and utilizes the strong template binding ability of phi29 DNA polymerase and template for strand displacement amplification [32]. The advantages are simple operation, low error rate, high amplification coverage, and the amplified DNA fragments are 10~100 kb in length; the disadvantage is that non-specific amplification occurs during amplification, and a high initial template amount is required. At low times, the amplification bias is large and there is sequence deviation [33]. MALBAC technology uses the advantages of PCR amplification technology and MDA amplification technology. Through special primers, the end of the amplicon is complemented to form a loop, which avoids the exponential amplification of genomic DNA, and the amplification bias is low. The product increased coverage of more than 93% of the whole genome [33]. And the detection rate of MALPC in SNPs alleles can reach 70%, compared with MDA, the improvement effect is obvious [34-35].

3.2 Single-cell Whole Genome Sequencing Technology

After efficient amplification of single-cell whole-genome DNA, a high-coverage genome is obtained, and single-cell genomic DNA is sequenced. So far, under the premise of ensuring the accuracy of sequencing, high-throughput sequencing technology is constantly optimizing the operating procedures, which is thousands of times higher than the traditional methods, which can greatly reduce the cost of gene sequencing. Single-cell genome sequencing has been used to analyze the recombination patterns and aneuploidy of human single germ cells, blastocysts and polar bodies pre-implantation screening, and heterogeneity of genomes in tumor cells and circulating tumor cells. The current high-throughput sequencing next-generation technologies include: Roche 454 pyrosequencing, Illumina Solexa sequencing, and Ion torrent ion semiconductor sequencing.

3.2.1 Roche 454 Pyrosequencing

Roche454 pyrosequencing technology uses a magnetic bead emulsification polymerase chain reaction technique based on pyrosequencing, which produces pyrophosphate when a base is paired with a template, and fluorescein and fluorescein under the action of ATP sulfated enzyme. The enzyme oxidizes to fluorescein, and the release of the optical signal is determined by the high-sensitivity charge-coupled device CCD to determine the base sequence of the template [36]. The technology has an accuracy of 99.96%, a coverage rate of 96%, and a long read length of up to 400bp. The disadvantage of this technique is that the continuous incorporation of bases cannot terminate the penetration of the same base due to the lack of termination elements. The length of the same base can only be inferred from the strength of the signal. In this process, errors of insertion or deletion may occur, and the cost is high, and the flux is low [37, 38].

3.2.2 Illumina Solexa Sequencing

Illumina Solexa sequencing technology utilizes "DNA clusters" and "reversible end terminations" [39]. By recovering 100 bp-200 bp DNA fragments randomly interrupted by physical methods, a linker is added at both ends to dilute at a certain concentration, and then the single-stranded DNA is complementary to the single-stranded primer on the surface of the chip, and after 30 rounds of amplification, a single sheet is formed. Clone "cluster". After linearization of the amplicon, four fluorescently labeled dNTPs (these dNTP "reversible terminators") were used for sequencing while sequencing, and the 3'-OH with a chemical cleavage site allowed only one base to be added. Then, after fluorescence is obtained by pyrosequencing, the 3'-OH terminal group is chemically cleaved, and the 3' viscosity is lowered, and the next round of reaction sequencing is performed [40]. The advantage of Illumina Solexa sequencing technology is that it can obtain higher throughput, requires less sample volume, high precision, avoids insertion or missing errors,
less manual operation, disadvantages of long running time, and sequence reading due to optical signals and other reasons. The length is short and the error rate accumulates as the chain lengthens [37, 41].

3.2.3 Ion Torrent Ion Semiconductor Sequencing

Ion torrent ion semiconductor sequencing technology is based on synthetic sequencing, using chemical and digital information technology, combined with nucleic acid chemistry and semiconductor technology, to directly complete non-optical application model DNA sequencing in integrated circuits [42-43]. The advantages of Ion torrent ion semiconductor sequencing technology are that DNA synthesis and detection can be performed under natural conditions, and the cost is low, the time is short, the flexibility is fast, and the flux measurement is fast, and the disadvantage is that misreading of polybases is easy.

4. Single Cell Whole Transcriptome Analysis

The transcriptome was first proposed by Vecalesuc and Kinzler et al [44], in 1997. It is a collection of genes expressed in the genome, which refers to the sum of all RNAs processed and transcribed by a living unit or tissue under certain conditions (including messenger RNA and non-coding RNA). In a single cell, mRNA has thousands of copies [45], and sequencing analysis requires that mRNA is reverse transcribed into cDNA and then amplified by PCR. Therefore, efficient and unbiased amplification is reverse transcription. Sequencing is a key factor [46].

4.1 Single Cell Whole Transcriptome Amplification Technology

At present, the methods for amplifying single cell cDNA mainly include: (1) PCR index amplification; (2) linear amplification in vitro; (3) Phi29 polymerase amplification.

4.1.1 PCR Index Amplification

The first is PCR amplification using cDNA with anchor sequences (left and right primers). The enzymes used in the synthesis of the second cDNA strand can be classified into mRNA-Seq [47], Smart-Seq [48], Smart-Seq2 [50], STRT-Seq [50], and SMA [51]. The anchor sequence at the 3' end of the transcript can be introduced by oligo (dT) primers during reverse transcription. Methods such as Smart-Seq, Smart-Seq2, and STRT-Seq utilize reverse-transferase terminal transferase and template-switching activities to synthesize complementary cDNA strands by template mapping. This method can avoid the interference of ribosomal RNA (rRNA) and transport RNA (tRNA), but can not detect various RNAs without poly(A) tail, the transcriptome with low level expression may be lost and the instability of Smart technology Wait.

The SMA semirandom primed PCR-based mRNA transcriptome amplification technique uses oligoribonucleotides with a hairpin structure as primers [52]. The full-length sequence of mRNA can be obtained, but the starting amount, sensitivity, and stability are different, and the range of use is also different.

4.1.2 In Vitro Transcription Linear Amplification

The second is to sequence the entire transcript, and the in vitro transcription (IVT) method [53] for linear amplification. IVT amplification avoids PCR bias but requires multiple rounds of reverse transcription to achieve the initial amount. The advantage of this technique is that the exponential amplification of cDNA is faster, but slower than the PCR process, the result is more accurate, and it is easy to cause Primer dimer and byproduct accumulation [54].

4.1.3 Phi29 Polymerase Amplification

The third is Phi29-mRNA amplification (PMA) [55] using the high fidelity and high amplification rate of Phi29 DNA polymerase used in whole genome amplification technology, with excellent strand displacement and continuous The characteristics of the synthesis, the cyclized cDNA is subjected to multiple generations of continuous amplification, and the enzyme also has 3-, 5-exonuclease reading activity. This reaction can be amplified at room temperature, which not only strongly avoids the influence of DNA degradation on the quality of the amplified product at high temperature, but also reduces the amplification advantage caused by the difference in GC content. However, the amplification uniformity of this method is not high.

4.2 Single Cell Transcriptome Sequencing Technology

Single-cell RNA-sequencing method, scRNA-seq can obtain a single-cell genome-wide expression profile, which can effectively avoid the cellular heterogeneity of conventional transcriptome analysis technology in biological systems and tissues, and can directly distinguish different organisms. The difference in gene expression and transcriptional characteristics between cell types can reflect the expression and regulation of all genes from the overall level of cells, revealing the molecular composition in cells and tissues, and can be clear about the randomness of gene expression [56].

General transcriptome sequencing is the choice to utilize high-throughput sequencing or microarray analysis [57-59]. The advantage of microarray analysis is that it can
provide transcriptome information and regulate network information, and high-throughput sequencing can reveal dynamic information more accurately from multiple angles [10]. In both methods, the ultimate goal is to first obtain the original RNA molecule, and then effectively perform consistent and stable amplification, and finally analyze the RNA sequence according to the experimental purpose. In addition, a new generation of sequencing methods is also constantly exploring and developing, single molecule fluorescence in situ hybridization, fluorescence in situ hybridization RNA sequencing and intracellular transcript analysis [26].

5. Outlook

So far, single-cell separation technology is still immature, whole-genome amplification bias, non-specific amplification, pollution during database construction, high technical noise, operational errors, low reproducibility, poor non-coding RNA detection The lack of barriers to splicing software still exists. Perfecting and improving sequencing analysis methods understand cell heterogeneity, randomness and synergy at the cellular level and play an important role in grasping the body’s growth and development process and biological key events. With the continuous advancement of sequencing technology, single-cell sequencing technology is widely used in the field of tumors, which can analyze the heterogeneity between solid tumors and circulating tumor cells from a large scale, and use this technology to predict and monitor diseases. Detection of drug resistance, etc. In the process of growth and development of tissues and organs, search for new cell types and marker genes; in the nerve and immune system, it can clearly understand the advantages of cell heterogeneity; it also has important applications in the fields of artificial insemination and stem cell transplantation. In order to achieve true proteomics and metabolomics analysis, with the continuous optimization of technical means, these problems are solved just around the corner. Single-cell sequencing technology will become an important technical means for humans to explore the growth and development of the body and various clinical fields.

References


