

Cancer Detection from Plasma Cell-Free DNA

Li Chen¹

¹Department of Biomedical Engineering, Engineering College, Peking University, Beijing, China

Abstract. Liquid biopsy of cell-free DNA (cfDNA) has attracted much attention for its promise to realize pan-cancer detection in a non-invasive way. The whole-genome bisulfite sequencing (WGBS) is widely used in cfDNA sequencing analysis and lays the foundation for further study on cfDNA. The cfDNA released by multiple tissues carries genetic and epigenetic information. Methylation patterns, copy number variation (CNV) and fragmentation changes have been discovered in previous studies and achieved a promising accuracy. In this review, different cancer detection methods based on these three biomarkers are introduced. In addition, feature fusion is discussed for its potential in enhancing performance in clinical applications.

1 Introduction

Cancer-related mortality has increased in recent decades, and cancer is expected to become the leading cause of death. Despite much improvement in cancer treatment, earlier detection still offers the greatest chance of cure, supported by the fact that early diagnosed patients have a survival rate five to ten times higher than the late stage ones^[1]. Nowadays, the golden standard for cancer detection is tissue biopsy which is popular for its high accuracy but being invasive and inefficient. In fact, it may cause irreversible damage to patients and can only take one tissue sample for detection at a time. Researchers have been looking for a non-invasive way to achieve pan-cancer detection among people, whose realization may include cancer detection into our annual physical examination. Recently, cell-free DNA (cfDNA) in plasma has been proved to provide an ideal biomarker for pan-cancer detection through liquid biopsy.

The cfDNA refers to double-stranded DNA fragments that are released into human plasma from the apoptosis of normal cells of the hematopoietic lineage, with minimal contributions from other tissues. They exist in bodily fluids of humans and have a short half-life, thus can be detected easily and carry timely information of certain tissues. Discoveries of fetal-derived DNA in plasma of pregnant women and donor-derived DNA in transplantation patients have introduced cfDNA-based noninvasive prenatal testing^[2] and transplantation monitoring^[3]. Recent studies have demonstrated the clinical feasibility of cfDNA analysis for cancer screening via various ways, like genetic ones focusing on certain mutations and epigenetic ones that compare between DNA methylation or fragmentation patterns. With more works on the sensitivity and specificity of cfDNA liquid biopsy, it is believed to become a very promising way in early cancer detection.

2 Whole-genome bisulfite sequencing provides comprehensive signatures of cfDNA

Recent developments in next-generation sequencing technologies, like RNA-seq and BS-seq, have allowed for genome-wide profiling of DNA signatures at single-nucleotide or single-cell resolution^[4, 5]. As the published work has mentioned that DNA methylation patterns and DNA copy number variations can both suggest an abnormal physical state^[6], these technologies are adapted in liquid biopsy analysis to detect genetic alterations and epigenetic changes.

Different kinds of sequencing technologies vary in many aspects, such as experiment protocol and bioinformatics analysis. Although some sequencing technologies can capture specific cell-free DNA signatures, it is still puzzling how informative and high-quality biological signals can be acquired at a low cost. Previous works^[7-9] have discussed the superiority of a sequencing technology called whole-genome bisulfite sequencing (WGBS), which is capable of simultaneously providing multiple biological signals, like DNA methylation and fragmentation patterns, as well as copy number variations.

For WGBS, genomic DNA libraries are created and subsequently bisulfite converted, sequenced, and aligned to the reference genome. In bisulfite sequencing, denatured DNA is subjected to bisulfite treatment during which the unmodified cytosine is converted to uracil, but a methylated cytosine remains unchanged, thus allowing base resolution detection of cytosine methylation.

WGBS data preprocessing contains several steps. First, the sequencing FASTQ data should be checked for quality control using FATSQC^[10], during which the low quantity samples are deleted for a convincing analysis. Then, the adapter should be detected and removed. An

chenli2018@pku.edu.cn

adapter is added to the DNA fragment before sequencing. However, the adapter itself does not belong to the fragment, hence its deletion before alignment, otherwise it will be regarded as part of the original fragment and will further lead to whole fragment being discarded. This will result in the ignorance of the fragments which may come from cancerous cells[8]. Some bioinformatics toolkits like AdapterRemoval[11] and Trim Galore[12] can tackle this problem properly. Finally, the bisulfite sequencing reads should be aligned to human genome. It should be noted that the bisulfite sequencing reads are not complementary to the reference genome acid because of C-T conversion. These facts add complexity to bisulfite sequencing and render conventional alignment tools, such as Bowtie2[13] or Burrows-wheeler aligner[14], unsuitable. Several new computational tools have been developed to address this issue, like BSMAP[15] and Segemehl[16], which enumerate all C to T combinations, or BISMARKE[17] and BS-Seeker[18], which convert all C to T in both sequenced reads and genome reference before alignment.

There are several issues to be noted in WGBS. First, the timing of bisulfite treatment in WGBS matters because the process of bisulfite conversion includes a high temperature, high salinity, acid, and alkaline environment, during which more than 90% of the DNA templates are damaged. Therefore, traditional WGBS needs a large amount of input data and outputs low effective reads. One way to improve the efficiency of bisulfite sequencing is to construct the library after the bisulfite conversion is done (Post-Bisulfite Construction, PBC). This strategy can greatly increase template utilization and richness, since damaged DNA fragments can still be included into the library. Notably, PBC strategy changes the original DNA fragment length, thus WGBS data with PBC strategy cannot be used in DNA fragment size analysis. Second, there is a conversion rate of each bisulfite conversion, and it changes according to experimental protocols and specific operations. Generally, over 90% of the unmodified cytosine is converted to uracil, and the rest will be regarded as methylated cytosine. Accordingly, the conversion rate has a non-negligible impact on the DNA methylation patterns in the sequencing result. For example, if the actual conversion rate is lower than expected, the methylation rate will be falsely high because of the unexpected cytosine that is not methylated but remains unchanged.

3 Signatures in WGBS for cancer detection

Recent studies have shown that many DNA changes founded in tumor tissue could also be identified in cell-free DNA, like DNA methylation and copy number variations. Furthermore, specific cfDNA features are also founded like DNA fragmentation profile changes. In the following scenarios, features commonly used in cancer detection of cfDNA sequencing data will be discussed.

3.1 Methylation pattern changes in cfDNA of cancer patients

DNA methylation is a ubiquitous phenomenon and linked to the developmental state and cellular differentiation processes in mammalian organisms. It is a kind of DNA modification through which genetic alterations can happen without sequence changes. It involves the addition of a methyl group to position 5 of the DNA cytosine rings by DNA methyltransferase enzymes and usually occurs at CpG dinucleotide sequences in mammals[19]. Generally speaking, CpG sites scattered on the genome are usually hypermethylated, while those areas with higher than average density of CpG (also known as CpG island) tend to be relatively hypomethylated. Research has found that erroneous DNA methylation may change gene expression patterns through different mechanisms, some of which are likely to be a driver of certain types of cancer. In fact, DNA hypermethylation is a widespread mechanism for the silencing of tumor suppressor genes in human cancers, and it is estimated to be as common as mutation. Also, DNA hypomethylation, which has been proved to cause chromosome instability, can be a driver factor to cancer formation[20]. Therefore, the DNA methylation pattern has great potential as a biomarker for cancer detection. DNA methylation signatures can be found in blood from cfDNA, which is released during cell apoptosis, and they are highly consistent between cfDNA and genomic DNA from its tissue origins. Many existing works have investigated specific methods of liquid biopsy for cancer detection.

Sun, et al. developed a statistical approach for studying the major tissue contributors to the circulating DNA pool[21]. The work identified 5820 specific methylation regions from 14 human tissues. These well-defined marker loci are highly variable between different tissues. Therefore, plasma DNA signal can be deconvoluted into different tissue components. The method was tested on pregnant women, patients with hepatocellular carcinoma and subjects following bone marrow and liver transplant. All these scenarios showed significant results to reveal the strong correlation between the methylation deconvolution and single tissue markers. In addition, Moss, et al. reported a similar approach for unbiased determination of the tissue origins of cfDNA[22]. Not like the method proposed by Sun, et al., Moss, et al. took a single CpG site as the deconvolution marker based on Illumina Infinium Human Methylation 450K or EPIC BeadChip arrays. What's more, this work made a comparison between cell-type and whole-tissue reference methylomes and found the former had a better performance, which highlighted the importance of using a comprehensive and cell type-specific DNA methylation atlas for sensitive detection of rare contributors to mixed methylomes.

Li, et al. proposed a cancer detection method termed "CancerDetector"[23]. It probabilistically modeled the joint methylation states of multiple adjacent CpG sites on an individual sequencing read and achieved the ultrasensitive resolution which can identify a trace amount of circular tumor DNA in plasma. It had a promising performance in both simulation experiments and real data tests. However, the work depends on an iteration step to

delete ambiguous markers which may cause a serious bias in circular tumor DNA burden estimation.

In comparison with other biomarkers for liquid biopsy cancer detection, DNA methylation patterns have two superior features that promise its great potential. First, DNA methylation signature is highly tissue-specific and has low individual differences. Second, as an epigenetic change, DNA methylation signature is more consistent in cancer. These two features allow for a systematic diagnosis method with high applicability.

3.2 Copy number variations can be detected in cfDNA

The copy number variation (CNV) refers to the phenomenon that sections of the genome are erroneously repeated or deleted during DNA replication. It affects a great number of genes, some of which are key functional regions that affect gene expression. Therefore, CNV can lead to various pathological states. Down syndrome and Huntington's disease are typical representatives of this type of disease. Existing research has proved the correlation between CNV and a certain type of cancers. Hence, CNV is a prognostic biomarker for cancer detection.

Chen, et al. made a comparison of whole-genome sequencing (WGS) at different sequencing coverage and demonstrated the feasibility of using low-pass whole-genome sequencing (LP-WGS) to detect CNVs in cfDNA and estimate tumor fraction in the plasma of patients[24]. This work shed light on numerous applications, since LP-WGS offers an unbiased and high-throughput way to study genome-wide CNVs that are cancer-related. It is also valuable for identifying novel CNVs as prognostic or predictive biomarkers. Because LP-WGS is relatively cost-effective, CNV analysis is also suitable for initial signal screening and continuous monitoring after treatment. Further, Jiang, et al. used larger scale CNV signals to distinguish HCC patients from healthy person[8]. This work focused on the deletion of chromosomes 1p and 8p and the amplification of chromosomes 1q and 8q, which are commonly observed in the HCC tissues. They quantified copy number gains or losses by chromosome arm-level z-score analysis (CAZA). The analysis outputs a z-score, whose value identifies whether the DNA fragment shows a certain CNV. In this study, the CNV was regarded as a tool for HCC detection, and it partially reflected that the CNV as a biomarker was relatively mature in comparison with other features like DNA fragmentation profiles.

CNV as a biomarker for cancer detection is a relatively well-studied one, and efforts have been made in the recent years to achieve greater specificity and sensitivity. Allen Chan, et al. demonstrated the upgrade in specificity and sensitivity of hypomethylation analysis after taking the CNV patterns into consideration. In the future, CNV is still a very potential biomarker for cancer detection.

3.3 Unique cfDNA fragmentation profile reveals cancer identity

The cfDNA released into plasma through cell apoptosis has different lengths. Research has shown that these fragments have nonrandom fragmentation patterns. For instance, peaks at around 147bp and 167bp correspond to nucleosomes and chromatosomes (nucleosome + linker) respectively[25]. This kind of feature shows difference between cancer patients and healthy person. Therefore, many efforts have been paid into this area.

Jiang, et al. designed an effective way (described in section 3.2) to utilize DNA CNV patterns to distinguish tumor-derived DNA fragments from non-tumor-derived ones, and massive parallel sequencing enabled them to measure the length of every individual plasma DNA molecules. It turned out that plasma DNA molecules bearing tumor-associated CNAs were shorter than those not carrying such signatures, and the length got even shorter when the fractional concentrations of tumor DNA in plasma increased[8]. However, patients with low fractional concentrations of tumor DNA plasma showed longer DNA fragmentation length in comparison with healthy controls. Jiang, et al supposed that it was because more non-tumor-derived DNA fragments were released from necrosis instead of apoptosis, which was consistent with the former work that reported longer DNA fragments from tissue necrosis than normal apoptosis. Further, a study was proposed to enhance the sensitivity for detecting the presence of circulating tumor DNA (ctDNA) by utilizing differences in fragment lengths of circulating DNA[26]. Notably, this work proposed an alternative deeper sequencing of cfDNA by combining LP-WGS (0.4X) and selective sequencing of specific fragment sizes. Plus, with predictive models integrating fragment length and CNV, the area under the curve (AUC) achieved more than 0.99, in comparison with that less than 0.80 without fragmentation features.

Snyder, et al. proposed a brand-new mechanism to distinguish between contributing tissues[25]. Unlike traditional methods that used genetic differences as a standard of distinction, Snyder, et al. made a hypothesis that the fragmentation patterns observed in an individual's cfDNA might contain evidence of their tissue-of-origin. The work built a map of nucleosome occupancy in human cells genome-wide with deeply sequenced cfDNA data. Then, comparisons of gene expression and regulatory site profiles were carried out, through which the epigenetic signatures of hematopoietic lineages contributing to cfDNA in healthy individuals were identified, with plausible additional contributions from one or more non-hematopoietic tissues in a small panel of individuals with advanced cancers. Although the work was not necessarily superior to mutation-based monitoring of ctDNA in sensitivity, but it envisioned a unique application of non-invasively classifying cancers at the time of diagnosis, simply by matching the epigenetic signature of cfDNA fragmentation patterns with reference datasets corresponding to diverse cancer types. It may also help the research of non-malignant conditions like myocardial infarction and stroke.

Sun et al. further explored the clinical potential of cfDNA fragmentation patterns[9]. The work reported that in open chromatin regions, cfDNA molecules showed characteristic fragmentation patterns reflected by sequencing coverage imbalance and differentially phased fragment end signals that preferentially occurred in tissue-specific open chromatin regions where the corresponding tissues contributed DNA into the plasma. Quantitative analyses of such signals allowed measurement of the relative contributions of various tissues toward the plasma DNA pool and thus, proposed a new method for nucleosome positioning profiling and quantitative determination of the relative contributions of various tissues in plasma DNA by fragmentation pattern analyses. These findings were validated by plasma DNA obtained from pregnant women, organ transplantation recipients, and cancer patients.

DNA fragmentation length has shown its great potential as a biomarker for liquid biopsy of cancer detection; however, this field is still in its infancy and further developments are necessary before clinical implementations.

4 Feature fusion in cancer detection

The aforementioned works primarily focus on single feature to distinguish cancer patients and the healthy. However, due to the heterogeneity of cancers, it is hard to detect all patients accurately from only one feature. Different features combining can provide more information, so it is more likely for models considering more features to achieve a better performance. In addition, paired end WGBS with pre-BS strategy can acquire DNA methylation, CNV, and fragmentation information simultaneously, so from a cost perspective, it is acceptable to comprehensively utilize all the available information. Some works have pioneered in this direction.

Jiang's work[8] mentioned before proposed a creative way to combine DNA CNV and fragmentation information. DNA CNV information is relatively well-studied, therefore, the work used it as a standard to differentiate tumor-derived DNA fragments from non-tumor-derived ones. This classification laid the foundation for follow-up research on DNA fragmentation length, which is still controversial. With such combination, this work provided a rather convincing result about DNA fragmentation patterns in cancer patients, and it is a good example for feature fusion in cfDNA research.

Another existing case for feature fusion is Mouliere's work[26]. Unlike Jiang's work mentioned above, Mouliere used DNA fragmentation patterns as a biomarker for initial screening. In fact, the work established a tumor-guided personalized deep sequencing based on size distribution of ctDNA, and identified clinically actionable mutations and copy number variations that were otherwise not detected with low-pass WGS. Then a predictive model integrating fragmentation length and CNV information improved the identification of patients with advanced cancer, with the area under the curve (AUC) more than 0.99 compared to that less than 0.80 without fragmentation features. Increased

identification of cfDNA from patients with glioma, renal, and pancreatic cancer was achieved with AUC more than 0.91 compared to that less than 0.5 without fragmentation features.

Xu Z, et al. built a model named GUseek for genitourinary (GU) cancer detection[27]. Notably, GUseek integrated DNA methylation and CNA features together to construct binary classifiers and multiclass classifiers. It can distinguish patients with GU cancer from healthy people, as well as tell apart UC, PRAD, and KIRC patients. GUseek took open-access database information as well as shallow whole-genome bisulfite sequencing data to evaluate CNA and DNA methylation profile changes and select features with multiple principles. The binary classifiers were able to diagnose GU cancer with an average accuracy of more than 95%, and the multiclass classifier could further distinguish different kinds of GU cancers with a total accuracy of 90.57%. What's more, the performances of GUseek between early- and late-stage cancer were similar, which indicated its potential in early-stage cancer detection and location.

These existing works have proved the feasibility of feature fusion in cfDNA cancer detection. However, it is still in the starting stage. Some work, like Jiang's and Mouliere's mentioned above, utilized two features in the whole designing process, but only one was actually taken as the biomarker during detection. GUseek was a good example for feature fusion. Similarly, Chan, et al. has enhanced the performance of multiple diagnostic models of DNA methylation patterns with CNV information[24]. It turns out that it is not only a more accurate way for cancer screening, but is more cost-effective as well.

As our knowledge about different cfDNA biomarkers becoming more comprehensive, it is reasonable to pay more attention to feature fusion of these biomarkers. It would be very tempting to utilize modern technologies like deep learning in cfDNA detection, which are expected to make revolutionary advancement in the field.

5 Discussion

Liquid biopsy of cfDNA has shown its potential in non-invasive pan-cancer detection, which is expected to offer earlier diagnose and thus, a better chance for cure. WGBS provides us with multiple biomarkers information to interpret cfDNA signals. Efforts have been made in each biomarker to achieve better performance, and this article summarizes some of the representative works to provide readers with an overview of the field. The DNA methylation pattern is a relatively traditional and well-studied biomarker. It is highly specific among different types of tumors and is conservative among different individuals, thus being very ideal for clinical cancer screening. DNA methylation patterns are now being analyzed primarily under two methods, namely the linear regression-based deconvolution model and the statistical model-based DNA fragment classifier, and many attempts have been made to improve the accuracy of models. Similarly, CNV change is also mature and has already been applicable in real situation, like prenatal testing and transplantation monitoring. Researchers have tried from

different angles to interpret CNV data so as to extract more information from a minimal amount of cfDNA. For example, unlike traditional bin-level models that focus on a single DNA base as the evidence of CNV, the arm-level model integrates more information and pushes the accuracy to a new level. Besides, DNA fragmentation patterns is a novel biomarker that seems to have a promising future. It contains new independent information from CNV or methylation patterns, providing a brand-new insight into cfDNA.

Until now, liquid biopsy of cfDNA has been used in many clinical scenarios, like prenatal and transplantation assessment. In comparison, the application in cancer detection is not yet popular. The major reason is that liquid biopsy of cfDNA for cancer detection still suffers from accuracy anxiety. To be more specific, as an early screening tool for cancer, it is acceptable to be imperfect in specificity, which at most requires for more rounds of the test, but insufficient sensitivity will greatly lower its application value. This dilemma is because our research on biomarkers of cfDNA is still in a correlation level. For example, it is still hard to pinpoint whether methylation changes are the driving factor in tumorigenesis, or just being the consequence. Lack of understanding in biological mechanism to some extent sets the ceiling for improvement of accuracy. With the ceiling existed, it is important for us to take full advantage of the processed information. In other words, how to realize feature fusion of cfDNA biomarkers is likely to be an important breakthrough in the field. It would be very interesting and of great potential in early diagnosis of cancer to explore more methods to combine the information and show a more comprehensive picture of cfDNA.

REFERENCES

1. Cho H, Mariotto AB, Schwartz LM, et al. When do changes in cancer survival mean progress? The insight from population incidence and mortality[J]. *Journal of the National Cancer Institute Monographs*, 2014, 2014(49):187-197.
2. Van Opstal D, Van Maarle MC, Lichtenbelt K, et al. Origin and clinical relevance of chromosomal aberrations other than the common trisomies detected by genome-wide NIPS: results of the TRIDENT study[J]. *Genetics in Medicine*, 2018, 20(5):480-485.
3. Schütz E, Fischer A, Beck J, et al. Graft-derived cell-free DNA, a noninvasive early rejection and graft damage marker in liver transplantation: A prospective, observational, multicenter cohort study[J]. *PLoS medicine*, 2017, 14(4):e1002286.
4. Chen Y-R, Yu S, Zhong S (2018) Profiling DNA methylation using bisulfite sequencing (BS-Seq), pp. 31-43, Springer.
5. Chu Y, Corey DR. RNA sequencing: platform selection, experimental design, and data interpretation[J]. *Nucleic acid therapeutics*, 2012, 22(4):271-274.
6. Jiang P, Chan KA, Lo YD. Liver-derived cell-free nucleic acids in plasma: Biology and applications in liquid biopsies[J]. *Journal of hepatology*, 2019, 71(2):409-421.
7. Cristiano S, Leal A, Phallen J, et al. Genome-wide cell-free DNA fragmentation in patients with cancer[J]. *Nature*, 2019, 570(7761):385-389.
8. Jiang P, Chan CW, Chan KA, et al. Lengthening and shortening of plasma DNA in hepatocellular carcinoma patients[J]. *Proceedings of the National Academy of Sciences*, 2015, 112(11):E1317-E1325.
9. Sun K, Jiang P, Cheng SH, et al. Orientation-aware plasma cell-free DNA fragmentation analysis in open chromatin regions informs tissue of origin[J]. *Genome research*, 2019, 29(3):418-427.
10. Andrews S (2010) FastQC: a quality control tool for high throughput sequence data, Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom.
11. Schubert M, Lindgreen S, Orlando L. AdapterRemoval v2: rapid adapter trimming, identification, and read merging[J]. *BMC Res Notes*, 2016, 9:88.
12. Krueger F. Trim galore[J]. A wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files, 2015, 516:517.
13. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2[J]. *Nat Methods*, 2012, 9(4):357-359.
14. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform[J]. *bioinformatics*, 2009, 25(14):1754-1760.
15. Xi Y, Li W. BSMAP: whole genome bisulfite sequence MAPping program[J]. *BMC bioinformatics*, 2009, 10(1):1-9.
16. Hoffmann S, Otto C, Kurtz S, et al. Fast mapping of short sequences with mismatches, insertions and deletions using index structures[J]. *PLoS Comput Biol*, 2009, 5(9):e1000502.
17. Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications[J]. *bioinformatics*, 2011, 27(11):1571-1572.
18. Guo W, Fizev P, Yan W, et al. BS-Seeker2: a versatile aligning pipeline for bisulfite sequencing data[J]. *BMC genomics*, 2013, 14(1):774.
19. Pfeifer GP. Defining driver DNA methylation changes in human cancer[J]. *International journal of molecular sciences*, 2018, 19(4):1166. [1] Cho H, Mariotto AB, Schwartz LM, et al. When do changes in cancer survival mean progress? The insight from population incidence and mortality[J]. *Journal of the National Cancer Institute Monographs*, 2014, 2014(49):187-197.
20. Pfeifer GP. Defining driver DNA methylation changes in human cancer[J]. *International journal of molecular sciences*, 2018, 19(4):1166.
21. Hentze JL, Høgdall CK, Høgdall EV. Methylation and ovarian cancer: Can DNA methylation be of diagnostic use?[J]. *Molecular and clinical*

- oncology,2019, 10(3):323-330.
22. Sun K, Jiang P, Chan KC, et al. Plasma DNA tissue mapping by genome-wide methylation sequencing for noninvasive prenatal, cancer, and transplantation assessments[J]. *Proc Natl Acad Sci U S A*,2015, 112(40):E5503-5512.
 23. Liu X, Ren J, Luo N, et al. Comprehensive DNA methylation analysis of tissue of origin of plasma cell-free DNA by methylated CpG tandem amplification and sequencing (MCTA-Seq)[J]. *Clinical Epigenetics*,2019, 11(1)
 24. Li W, Li Q, Kang S, et al. CancerDetector: ultrasensitive and non-invasive cancer detection at the resolution of individual reads using cell-free DNA methylation sequencing data[J]. *Nucleic Acids Res*,2018, 46(15):e89.
 25. Chen X, Chang C-W, Spoerke JM, et al. Low-pass Whole-genome Sequencing of Circulating Cell-free DNA Demonstrates Dynamic Changes in Genomic Copy Number in a Squamous Lung Cancer Clinical Cohort[J]. *Clinical Cancer Research*,2019, 25(7):2254-2263.
 26. Snyder MW, Kircher M, Hill AJ, et al. Cell-free DNA Comprises an In Vivo Nucleosome Footprint that Informs Its Tissues-Of-Origin[J]. *Cell*,2016, 164(1-2):57-68.
 27. Mouliere F, Chandrananda D, Piskorz AM, et al. Enhanced detection of circulating tumor DNA by fragment size analysis[J]. *Sci Transl Med*,2018, 10(466)
 28. Xu Z, Ge G, Guan B, et al. Noninvasive Detection and Localization of Genitourinary Cancers Using Urinary Sediment DNA Methylomes and Copy Number Profiles[J]. *European Urology*,2020, 77(2):288-290.