Review

Sequence-Based Platforms for Discovering Biomarkers in Liquid Biopsy of Non-Small-Cell Lung Cancer

Liam J. Brockley [1,*,†], Vanessa G. P. Souza [1,2,†], Aisling Forder [1], Michelle E. Pewarchuk [1], Melis Erkan [3,4,5], Nikita Telkar [1,6], Katya Benard [1], Jessica Trejo [1], Matt D. Stewart [1], Greg L. Stewart [1], Patricia P. Reis [2,7], Wan L. Lam [1] and Victor D. Martinez [3,4,5,*]

1   British Columbia Cancer Research Institute, Vancouver, BC V5Z 1L3, Canada; vg.souza@unesp.br (V.G.P.S.);
     aforder@bccrc.ca (A.F.); mpewarchuk@bccrc.ca (M.E.P.); ntelkar@bccrc.ca (N.T.); kbenard@bccrc.ca (K.B.);
     jtrejo@bccrc.ca (J.T.); mstewy@student.ubc.ca (M.D.S.); gstewart@bccrc.ca (G.L.S.); wanlam@bccrc.ca (W.L.L.)
2   Molecular Oncology Laboratory, Experimental Research Unit, School of Medicine, São Paulo State
     University (UNESP), Botucatu 18618-687, SP, Brazil; patricia.reis@unesp.br
3   Department of Pathology and Laboratory Medicine, IWK Health Centre, Halifax, NS B3K 6R8, Canada;
     merkan@dal.ca
4   Department of Pathology, Faculty of Medicine, Dalhousie University, Halifax, NS B3K 6R8, Canada
5   Beatrice Hunter Cancer Research Institute, Halifax, NS B3H 4R2, Canada
6   British Columbia Children's Hospital Research Institute, Vancouver, BC V5Z 4H4, Canada
7   Department of Surgery and Orthopedics, Faculty of Medicine, São Paulo State University (UNESP),
     Botucatu 18618-687, SP, Brazil
*   Correspondence: lbrockley@bccrc.ca (L.J.B.); victor.martinez@iwk.nshealth.ca (V.D.M.)
†   These authors contributed equally to this work.

Simple Summary: A lack of sensitive biomarkers hinders lung cancer detection and monitoring, resulting in late diagnosis and missed treatment response. Liquid biopsies have recently emerged as a promising method of detecting biomarkers in lung cancer patients without the need to conduct invasive procedures. New approaches for biomarker discovery have emerged due to advances in high-throughput sequencing technologies and bioinformatics tools. In this review, we comprehensively describe established and emerging methods for identifying lung cancer biomarkers using liquid biopsy. Furthermore, we highlight advanced bioinformatics tools and methods for processing NGS data, as well as recently developed software tailored for liquid biopsy biomarker detection.

Abstract: Lung cancer detection and monitoring are hampered by a lack of sensitive biomarkers, which results in diagnosis at late stages and difficulty in tracking response to treatment. Recent developments have established liquid biopsies as promising non-invasive methods for detecting biomarkers in lung cancer patients. With concurrent advances in high-throughput sequencing technologies and bioinformatics tools, new approaches for biomarker discovery have emerged. In this article, we survey established and emerging biomarker discovery methods using nucleic acid materials derived from bodily fluids in the context of lung cancer. We introduce nucleic acid biomarkers extracted from liquid biopsies and outline biological sources and methods of isolation. We discuss next-generation sequencing (NGS) platforms commonly used to identify novel biomarkers and describe how these have been applied to liquid biopsy. We highlight emerging biomarker discovery methods, including applications of long-read sequencing, fragmentomics, whole-genome amplification methods for single-cell analysis, and whole-genome methylation assays. Finally, we discuss advanced bioinformatics tools, describing methods for processing NGS data, as well as recently developed software tailored for liquid biopsy biomarker detection, which holds promise for early diagnosis of lung cancer.

Keywords: bioinformatics; CTCs; ctDNA; high throughput; liquid biopsy; lung cancer; NGS

1. Introduction

Lung cancer is one of the most commonly diagnosed types of cancer and is the leading cause of cancer deaths worldwide [1]. Liquid biopsy, which involves analyzing body fluids to detect cancer-specific biomarkers, has emerged as a promising tool for the diagnosis and management of lung cancer patients [2–5]. One of the main advantages of a liquid biopsy test is that it provides a minimally invasive method for detecting and monitoring the disease. Traditional tissue biopsies are invasive, associated with pain and carry the risk of complications, such as infections. In contrast, a liquid biopsy test only requires the collection of body fluid, such as blood, which is a much less invasive procedure [2]. Liquid biopsy can be serially repeated, with additional advantages, such as speed, low cost and safety of isolating body fluids compared to tumor tissue [6]. Liquid biopsies are useful to detect cancer-specific DNA and other biomarkers that are shed into the bloodstream from cancer cells, providing a way to monitor the disease even in its early stages [7]. In fact, liquid biopsies are clinically applicable to guiding treatment decisions, such as determining the best therapeutic regimen or assessing if a patient's cancer has changed or progressed [8]. Liquid biopsy tests are useful to monitoring response to treatment and detecting the emergence of drug-resistance mutations, allowing for treatment intervention and optimization. Liquid biopsy monitoring of drug response helps to improve overall outcomes and quality of life of lung cancer patients [2,9]. Therefore, liquid biopsy tests offer a number of important benefits for lung cancer patients, including improved accuracy and reliability of diagnosis, minimally invasive disease monitoring and the ability to guide and adjust treatment according to test results. The emergence of high-throughput sequencing technologies and bioinformatics tools has led to the development of new approaches for biomarker discovery using liquid biopsies [10–12]. High-throughput sequencing technologies allow for the simultaneous analysis of multiple biomarkers in a single sample, providing a comprehensive picture of the genomic landscape of a tumor. Bioinformatics tools applied to the analysis of large liquid biopsy datasets are key to identifying clinically actionable changes and predictive biomarkers to guide treatment decisions. Also, next-generation sequencing (NGS) technologies enable the simultaneous sequencing of millions of DNA fragments in a single test, providing a highly sensitive and specific method for detecting cancer-specific mutations, gene fusions and other genetic alterations associated with development and progression of lung cancer [13,14]. Such approaches are of utmost importance to improving the accuracy of lung cancer diagnosis and prognosis. The aim of this review is to survey the recent applications of NGS technologies to liquid biopsies in the context of non-small-cell lung cancer (NSCLC).

2. Sample Types and Materials for Biomarker Analysis

Liquid biopsy is a method that involves analyzing circulating tumor cells (CTCs) or molecules originating from the tumor in biological fluids [7]. These molecules include circulating tumor DNA (ctDNA) and RNA (ctRNA), proteins and extracellular vesicles (EVs) present in the biofluids of patients [7,15]. These biofluids include plasma, serum, sputum, bronchoalveolar lavage fluid (BAL), pleural effusion, urine, saliva and cerebral spinal fluid (CSF) [16] (Figure 1). Currently, the body fluid most commonly used for liquid biopsy in lung cancer is blood [16,17], which is used to identify specific biomarkers correlated with clinical outcomes in advanced NSCLC [18] and enables the monitoring of disease relapse and resistance of cancer [19]. Here, we briefly summarize materials used in liquid biopsy for biomarker analysis, as their features have been reviewed in depth elsewhere [7,15,16,20,21].
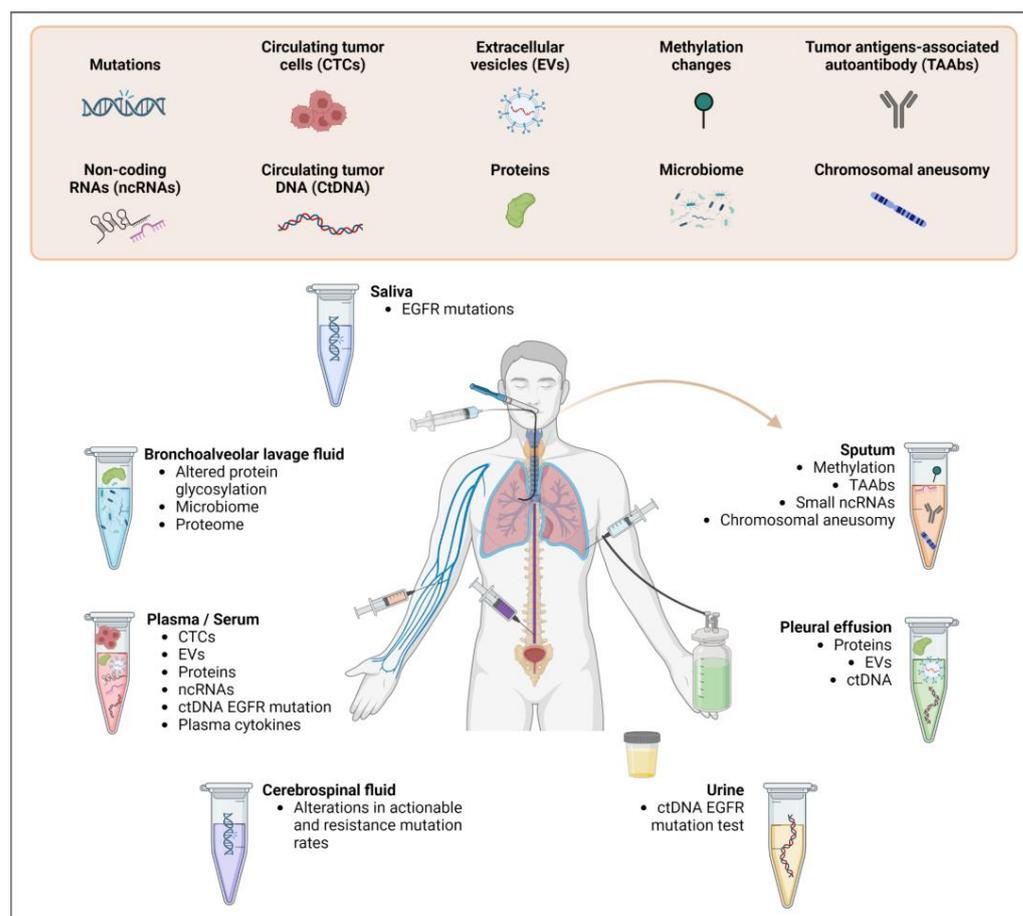
Figure 1. Types of biomarkers used to detect lung cancer in liquid biopsy and the types of bodily fluid samples from which they can be derived.

2.1. Circulating Tumor Cells (CTCs)

CTCs are isolated tumor cells or in clusters that are released by the primary tumor or metastatic cells that leak into the bloodstream towards the metastatic site [7]. CTCs are present only in extremely low concentrations in the blood, approximately one CTC per million leukocytes [22], but their presence can have diagnostic implications, for example, in lung cancer where high levels of CTCs have been associated with worse outcomes [23]. The presence of CTCs was also associated with low response rates, shorter progression-free survival and shorter overall survival in patients with advanced NSCLC treated with both targeted- and chemotherapy [24]. Different methods for enrichment, isolation and identification of CTCs have been developed according to the physical and biological characteristics of these cells [25–27]. In lung cancer, isolation by size of epithelial tumor cells (ISET) was the earliest size-based method used for CTC detection, showing high sensitivity and reproducibility [28]. Flow cytometry, fluorescence-activated cell sorting (FACS) and microfluidics are all commonly used techniques for isolating CTCs from liquid biopsy samples [29–32]. The most widely used immunomagnetic assay of epithelial cell adhesion molecule (EpCAM) enables the enumeration of CTCs of epithelial origin in whole blood, but it potentially misses a large cell population of "stem-like" characteristics [33,34]. Laser capture microdissection is another CTC sequestration method, which involves encapsulating a CTC on a hydrogel, extracting it with a laser and sequencing it [35].

2.2. Circulating Tumor DNA (ctDNA)

Circulating tumor DNA is part of the pool of total circulating cell-free DNA (cfDNA) and is released from tumor cells that have entered into apoptosis or necrosis [36]. In

addition, fragments of tumor DNA resulting from the phagocytosis of necrotic tumor cells by macrophages can be released by these cells in circulation. The level of ctDNA present can vary by cancer type, but overall increases in ctDNA correspond with tumor burden, disease progression and metastasis [36–39]. Circulating tumor DNA is commonly quantified in serum and plasma samples, which are obtained by removing cellular components from whole blood samples by centrifugation, but it has also been reported in other bodily fluids, such as cerebrospinal fluid. The three main categories of methods to isolate the ctDNA, irrespective of the sample of origin, are phase isolation, silicon membrane-based spin column and magnetic bead-based isolation [36]. Once isolated, ctDNA can be analyzed by a variety of methods, for example, methylation markers can be measured using bisulfite sequencing [40]. Specific sequences and mutations can be targeted using polymerase chain reaction (PCR)-based methods, including conventional PCR and digital droplet PCR (ddPCR) [41,42]. Untargeted sequencing methods, such as shotgun massively parallel sequencing, can be used to quantify copy number alterations, and NGS can be used to screen for mutations in large regions of the genome [43,44]. It is important to note that standardization for pre-analytical factors, such as clotting, number of freeze-thaws, DNAse activity in the blood and the time elapsed between the blood draw and analysis is still required [7]. DNA input is also an issue, with low input also potentially impairing sensitivity [36].

2.3. Circulating Tumor RNA (ctRNA)

Circulating tumor RNA includes protein-coding RNAs (mRNAs), long non-coding RNAs (lncRNAs), microRNAs (miRNAs) and others. Circulating tumor RNA is most commonly collected from peripheral blood, but can also be detected in BAL fluid, saliva, pleural effusion, urine and CSF (largely in the form of miRNAs, the most abundant ctRNA) [15,45,46]. Both ctDNA and ctRNA contain mutational information, but ctRNA can also provide information about the quantitative expression levels of genes of interest [47]. Similar to ctDNA, ctRNA is likely released from apoptotic or necrotic cells but may also be released in exosomes (discussed below), which protect them from degradation [46,48]. RNAs that are not encapsulated are also released into bodily fluids, but certain classes (e.g., mRNAs and lncRNAs) are more likely to be degraded by ribonucleases (RNases) [15,49]. miRNAs, however, are less susceptible to this due to their small size and their association with RNA-binding proteins [15,50]. Circulating tumor RNAs can be isolated from liquid biopsies by using commercial RNA extraction kits, phenol-chloroform methods, or guanidium thiocyanate methods [51]. Kits may introduce biases, but this can be controlled by using exogenous RNA spike-ins [51]. After or prior to extraction, samples should also be treated with a DNase to prevent DNA contamination [51]. Following collection, ctRNAs can be analyzed using ddPCR, reverse-transcription quantitative PCR (RT-qPCR), microarray or NGS methods, such as RNAseq or small RNAseq [52,53]. For biomarker discovery, RNAseq/small RNAseq are the gold standard, while PCR-based methods are more commonly used for the detection of specific, predetermined ctRNAs [51,54].

2.4. Extracellular Vesicles (EVs)

EVs are small, membrane-bound vesicles that include both exosomes, which are secreted from the endosomal system of the cell, and microvesicles, which bud off the plasma membrane. They contain cargo, such as DNA fragments, non-coding RNAs, mRNAs and proteins that facilitate extracellular communication and thus tumor progression and metastasis through a variety of downstream mechanisms [55]. EVs are present at high levels in bodily fluids including blood and are stable in circulation, but isolating tumor-specific EVs is an ongoing challenge in the field [56]. Differential expression of surface markers on tumor-derived EVs, as well as differential loading of disease-related cargo into EVs, mean that EVs can be used as a tool for liquid biopsy for detection, monitoring of disease progression and prognosis [57–59]. Differential ultracentrifugation is currently the gold standard for EV isolation [60]. Newer approaches use immunoaffinity [61], and other

physical properties or combinations of both [56]. Isolated EVs should be characterized with a Western blot for common EV-specific markers, such as CD63, CD81 and CD82 as well as at the single-exosome level with techniques, such as nanoparticle tracking or electron microscopy, to confirm their identity [62]. Isolated EVs can then be used for conventional protein analysis with Western blot, ELISA or novel methods, nucleic acid analysis with RT-qPCR or PCR-based technologies, such as digital droplet PCR and NGS [56].

3. Commonly Used Techniques for Biomarker Discovery in Lung Liquid Biopsy

There are several techniques that are commonly used for liquid biopsy, depending on the type of biomarkers being analyzed and the specific clinical application [15,63–65]. At present, ctDNA is the most extensively studied and widely used biomarker for liquid biopsy in NSCLC [66,67]. Both the European Medicines Agency (EMA) and the US Food and Drug Administration (FDA) have approved the use of ctDNA information for selecting NSCLC patients with EGFR-mutant cells for treatment with certain targeted therapies when a tumor sample cannot be evaluated [68–70]. Platforms and strategies currently available to detect ctDNA in lung cancer include both non–NGS- and NGS-based approaches [11,15,71].

3.1. PCR-Based Approaches

In general, the fraction of ctDNA in the blood of patients with cancer is thought to be relatively low, typically ranging from a few copies per milliliter to a few percent of the total DNA in the sample [72]. Consequently, traditional DNA analysis techniques (such as Sanger sequencing and pyrosequencing) are insufficient for detecting low amounts of DNA in blood samples [73]. PCR-based methods, such as real-time quantitative PCR (qPCR), digital PCR (dPCR) and mass-spectrometry-based methods, have been widely used as alternatives to traditional techniques since they have shown better sensitivities and are of low costs [2,11,15,74,75]. There are some variations of PCR-related methods including co-amplification at lower denaturation temperature-PCR (COLD-PCR) [76], refractory mutation system-PCR (ARMS-PCR) [77], locked nucleic acid (LNA)/DNA-PCR [78], peptide nucleic acid (PNA) clamp-PCR [79], beads, emulsions, amplification, magnetics (BEAMing) [80], ddPCR [81–85], intelligent multiplexed amplification for NGS applications (InPlex) [86,87] and Endpoint PCR [88]. The features of each method have been reviewed elsewhere [15,89]. Although the sensitivity of PCR-based methods is better when compared to traditional DNA analyses, it is still low, with the limit of detection of ctDNA ranging from < 0.0001% to 0.1% [15]. These techniques are also limited to the analysis of one or a very small number of genomic loci, even with multiplex analysis, and the specific mutation to be assayed is typically determined a priori [15,90,91]. While there have been promising uses of selective gene panels to screen ctDNA for oncogenic mutations in NSCLC patients, the sheer number of possible pathogenic mutations means that a predetermined panel is comparatively restricted in scope for biomarker discovery [92].

3.2. NGS-Based Approaches

To overcome the aforementioned limitations of PCR-based techniques, recent advances in NGS technologies have made it possible to detect even very low levels of ctDNA in the blood, allowing the detection of different types of alterations (e.g., point mutations, gene fusions and translocations) in multiple genes in the same analysis [11,39,93–95]. Currently, NGS technology has the ability to detect a minor allele frequency (MAF) of less than 1% in lung cancer [96]. Furthermore, the use of unique molecular identifiers (UMIs) or unique barcodes can help to improve the accuracy and sensitivity of NGS-based liquid biopsy assays for lung cancer, especially for samples with low ctDNA concentrations [97,98]. Targeted NGS approaches with molecular barcoding have been used to detect ctDNA in early-stage lung cancer patients. The assay was able to detect ctDNA in both patients with stage I or II NSCLC, though the sensitivity of the assay increased with higher ctDNA concentrations and was highest for patients with stage II disease [39]. Another study used a targeted NGS approach with UMIs to detect ctDNA in NSCLC patients. The assay was able

to detect ctDNA in 59% of stage I or II NSCLC patients with a median MAF of 0.1% [98]. These studies demonstrated that the use of UMIs can reduce PCR and sequencing errors and improve the detection of low-frequency variants [98].

NGS uses probes to capture specific DNA fragments that are then sequenced in parallel and computationally aligned to a reference genome, enabling the sequencing of large numbers of gene targets and variant types in a single experiment. Furthermore, newer sequencing technologies provide options for sequencing portions of the genome at extremely high depths of coverage (i.e., the number of sequencing reads covering a specific position in the genome), which can reveal known mutations occurring at a low frequency or even uncover new driver mutations. Many methods are currently available to perform NGS analyses in liquid biopsy using either targeted or untargeted panels (Table 1). Targeted panels are designed to sequence specific genomic regions of interest, such as genes known to be frequently mutated in cancer. These panels can provide high coverage and depth of sequencing for the selected targets, which can increase the sensitivity of the assay for detecting low-frequency mutations. However, targeted panels may miss mutations outside of the selected regions, which can limit their ability to identify novel or unexpected mutations [8,11,15]. In contrast, untargeted panels, also known as whole-genome sequencing (WGS) or whole-exome sequencing (WES), can provide comprehensive and unbiased sequencing of the entire genome or exome, which can allow for the identification of novel or unexpected mutations [99]. However, untargeted panels may have lower coverage and depth of sequencing for specific regions, which can limit their sensitivity for detecting low-frequency mutations. Thus, both targeted and untargeted panels have advantages and limitations, and the choice of the panel may depend on the specific research or clinical question being addressed [8,11,15].

3.3. Clinically Validated Platforms for Biomarker Detection

In clinical practice, there are FDA-approved laboratory tests to analyze liquid biopsy biomarkers (typically ctDNA) for NSCLC. These include the Guardant360 CDx [100,101], FoundationOne Liquid CDx [102,103] and the cobas EGFR mutation Test v2 [104]. They use different panels to monitor a set of genes that may impact a patient's response to drug treatments. The Guardant360 CDx test can identify the *EGFR* exon 20 insertion mutation and thus indicate if patients are eligible for treatment with amivantamab-vmjw [105,106]. The FoundationOne NGS test indicates whether gefitinib, osimertinib or erlotinib are the appropriate treatment, while the real-time PCR cobas EGFR mutation test v2 identifies specific *EGFR* mutations for erlotinib [102–104].

Table 1. NGS platforms designed for biomarkers in liquid biopsy.

Technology	Brief Description	References
Non Targeted		
WES	Whole Exome Sequencing sequences all exons in ctDNA for mutation detection. Less expensive than WGS (lower coverage). Sample requirement not always feasible in liquid biopsy.	[107–109]
Digital Karyotyping	Uses WGS to sequence short DNA tags and then aligns these tags to the reference genome to identify genomic alterations, e.g., CNVs, SNVs and SVs. The short DNA tags are typically generated by restriction enzyme digestion. Requires high-quality genomic DNA.	[43,110–112]
FAST-SeqS	Fast Aneuploidy Screening Test-sequencing System uses individual primer pairs to amplify the repeat regions of interest. The WGS version, called mFAST-SeqS, identifies any somatic mutations in the tumor and then uses those mutations as unique markers for monitoring the disease.	[113–115]
PARE	Personalized Analysis of Rearranged Ends uses WGS data to identify rearranged ends in ctDNA. Detects structural variations, e.g., translocations and inversions.	[111,116,117]

Table 1. *Cont.*

Technology	Brief Description	References
Targeted panel		
Tam-seq	Tagged-Amplicon deep sequencing uses primers targeting regions of interest for a pre-amplification step. Templates undergo individual amplification, aiding purification.	[44]
Safe-SeqS	Safe-Sequencing System is a method for profiling low-frequency mutations. The method combines PCR amplification of targeted genomic regions with UMIs and deep sequencing to achieve high accuracy and sensitivity. The use of UMIs reduces errors introduced by PCR amplification and sequencing.	[118]
CAPP-Seq	Cancer Personalized Profiling by Deep Sequencing uses a library that generates many hybrid affinities captures of recurrently mutated genomic regions to create the selector, which is used to identify individual-specific mutations in the tumor DNA. The selector is then applied to ctDNA for quantification.	[119]
Ion AmpliSeq™	Customized multiplex PCR amplifies target regions for analysis with the Ion Torrent sequencing platform.	[120]
Guardant360®	Analyzes 73 genes commonly mutated in cancer. Digital sequencing technology for mutation detection with 99.5% sensitivity and 99.999% specificity. FDA approval for use in patients with advanced cancer without treatment options.	[100,101]
Foundation One®CDx	Analyzes 324 genes and selects genomic signatures, including MSI and TMB. Detects single nucleotide variants, small in/dels, copy number alterations and gene fusions. FDA-approved for use in patients with solid tumors, including NSCLC, to sort patients for specific targeted therapies.	[102,103]
iDES	In Integrated Digital Error Suppression, DNA is tagged with UIDs and tracked through library preparation and sequencing for error correction. By incorporating UIDs into NGS, iDES can improve the accuracy and sensitivity of NGS assays, particularly in low-frequency variant detection.	[121]
TEC-Seq	Targeted Error Correction Sequencing is a method for profiling low-frequency mutations in cfDNA. Utilizes molecular barcoding to distinguish true mutations from false positive variants. Before any amplification, DNA fragments are tagged with different "exogenous" DNA barcodes. Additionally, the start and end genome mapping positions of paired-end sequenced fragments are used as "endogenous barcodes" to differentiate between individual molecules. This combination of barcodes enables tracking each fragment, allowing for the detection of rare mutations with high accuracy and sensitivity.	[98]

Abbreviations: CNVs: Copy Number Variations; ctDNA: circulating tumor DNA; FDA: US Food and Drug Administration; mFAST-SeqS: Mutation-focused Assessment of Sequencing and Tracking by Sequencing; MSI: microsatellite instability; SNVs: Single Nucleotide Variations; SVs: Structural Variations; TMB: Tumor Mutational Burden; UIDs: Unique Identifiers; UMIs: Unique Molecular Identifiers; WGS: Whole-Genome Sequencing.

4. Emerging Methods for Liquid Biopsy Biomarker Discovery

Emerging methods for biomarker discovery in liquid biopsy have the potential to revolutionize the ability to diagnose and monitor diseases non-invasively and to identify novel therapeutic targets. In this section, we describe some of the emerging methods for biomarker discovery in lung cancer liquid biopsy, including long-read sequencing, DNA methylation, single-cell sequencing and fragmentomics analysis.

4.1. Long-Read Sequencing

Current NGS platforms implemented in diagnostic settings (mainly developed by Illumina, San Diego, CA, USA) produce as many as 6 billion highly accurate (greater than 99.9%) "short reads" (150–400 nucleotides) per run [122]. Accuracy, parallelization and relatively low cost of short-read sequencing have made this method the preferred option for the detection of SNVs (single nucleotide variants) and small insertion–deletion mutations (indels) in liquid biopsies. The large number of generated reads allows for an improvement of the signal-to-noise ratio with regard to the detection of mutations in liquid biopsies. However, the clinical performance of short-read NGS is limited in the detection of any cancer-related feature associated with large fragments and/or duplicated

regions of the genome. Newer long-read-based technologies such as Oxford Nanopore Technologies (ONT) (Oxford, UK) and PacBio (San Diego, CA, USA) have an advantage in detecting these features and avoiding amplification bias [123]. The most recent ONT devices use adaptors to attach phi29 DNA polymerase motor proteins to the ends of nucleic acid fragments, which feed the strands through α-hemolysin and MspA pores. Bases are detected as they pass through the pores and change the flow of an electric current, which is output as sequencing data in real-time [124,125]. In PacBio Single Molecule Real-Time (SMRT) sequencing, each free DNA strand is circularized with adapters. A DNA polymerase replicates the strand with labeled nucleotides, releasing fluorescent signals that are read as a sequence in real-time [126,127]. Long-read sequencing enables the analysis of structural properties and the fragmentation patterns of cfDNA (including fragment size, nucleosome footprinting and methylation patterns), which can give information about their tissues of origin. This is known as "fragmentomics" [128,128–130], which we discuss further below. Long-read sequencing technologies (PacBio, ONT) were recently compared for real-time detection of long cfDNA in plasma. PacBio SMRT sequencing generated data with higher percentages of long cfDNA compared with nanopore sequencing. Yet a higher number of long cfDNA fragments eligible for the tissue-of-origin analysis could be obtained from ONT sequencing due to its much higher throughput [131]. Newer methods aim to obtain long-range information through alternative methods, such as linked-read sequencing (LR-seq). LR-seq is a cost-effective library preparation technology that maintains the long-range structure of the original genomic material in the generated short reads, making it an advantageous option that can be integrated into existing clinical lab protocols without additional specialized equipment [122,123,131–133].

4.2. DNA Methylation Markers

Profiling of DNA methylation (DNAm), the epigenetic mechanism by which a methyl group is transferred to a cytosine making a 5-methylcytosine, has proven to be highly informative in the detection and prediction of several disorders [134]. CpG dinucleotides (a cytosine followed by a guanine) are the most frequent cytosines where DNAm occurs, and methylation signatures observed in diseased tumor tissue show significant overlap in liquid biopsy samples (most often using ctDNA) [135–137]. On average, tumors show global hypomethylation when compared to healthy tissue, but loci-specific CpG island hypermethylation, along with enrichment in repressive histone marks, such as H3K27me3 [138,139].

Analysis of plasma ctDNA has been successful in detecting the aberrant hypermethylation of promoters in genes associated with the development of lung cancer, in single genes, and in gene signatures, detectable even at early cancer stages [140,141]. A recent study involving 2800 participants compared the accuracy and limit of detection of different ctDNA analysis techniques for early cancer detection and found that whole-genome methylation analysis strongly outperformed WGS and targeted sequencing [142]. Sputum samples have proven to be highly effective in detecting lung cancer progression early [143,144], specifically in active and former smokers. While relatively few studies have investigated DNAm changes occurring in lung cancer using urine, saliva and pleural effusion, they have still demonstrated the utility of using these liquid biopsy samples for early detection, as well as their success in separating lung cancer subtypes [145]. The majority of DNAm studies use bulk tissue; however, the methylomes of individual cell types differ, as demonstrated by cell-specific gene expression profiles [146]. Single-cell lung methylome profiling is still in its infancy [147].

DNAm profiling can be performed using sequencing-based or array-based methods. Whole genome bisulfite sequencing (WGBS), in which unmethylated cytosines are converted to uracil, remains the gold standard for DNAm detection. However, the method performs best for shorter-length fragments, is prone to errors due to bisulfite treatment and remains expensive [148]. Methylated DNA immunoprecipitation (MeDIP), which employs antibody ligation to methylated CpGs, works well for detecting genome-wide methylation enrichment and performs well for limited sample material [149]. Illumina's Infinium Hu-

manMethylation450 BeadChip array assessing 450,000 CpG sites and its successive Human MethylationEPIC BeadChip at 850,000 CpG sites currently are the most cost-effective and widely used.

4.3. Single-Cell Sequencing

Single-cell analysis has emerged as a powerful tool in liquid biopsy research, particularly for sequencing CTCs and cancer-associated immune cells [150–154]. Capturing and profiling a tumor cell in the act of metastasizing is a good strategy to understand the current molecular status of cancer [32]. There are several techniques for single-cell analysis in liquid biopsy, including single-cell RNA sequencing, single-cell DNA sequencing, single-cell proteomics, single-cell secretomes and single-cell metabolomics [151]. These methods can be used to identify heterogeneous subpopulations of cells, track their behavior and evolution and identify novel biomarkers [150,155].

Single-cell DNA sequencing can be used to detect genomic features, such as single nucleotide variants, copy number variation and microsatellite instability within CTCs. Single-cell RNA sequencing, which requires reverse transcription to cDNA, provides additional information on gene expression signatures. Following cell isolation, library construction is performed. During this process, whole-genome or whole-transcriptome amplification (WGA/WTA) is necessary to generate sufficient genetic material for sequencing, which requires numerous quality control measurements [151]. Some current WGA/WTA methods include multiple displacement amplification (MDA), multiple annealing and looping-based amplification cycles (MALBAC), emulsion whole-genome amplification (eWGA) and Laser-induced Isolation of Microstructure On transferrable-chip and sequencing (LIMO-seq) [156]. MALBAC offers higher uniformity of genomic amplification than MDA by utilizing quasilinear rather than nonlinear amplification. MALBAC has been successfully used for sequencing CTCs derived from NSCLC patients [157,158]. LIMO-seq involves using a microfluidic chip and a laser pulse to isolate CTCs, followed by MDA [159]. eWGA entails separating DNA fragments into droplets where separate MDA reactions occur [160]. A group has recently compared WGA and WTA methods for CTC single-cell sequencing and found that methods using MDA had higher amplification efficiency than others [161].

Following library construction, molecules are often labeled with a UMI. Libraries may then be pooled together (multiplexed) during sequencing, which necessitates demultiplexing based on barcodes and UMIs if present [162]. Multiplex PCR with predetermined gene panels has been used to profile CTC transcriptomes from lung cancer patients at the single-cell level, which revealed expression patterns correlated with prognosis [163,164]. While the low concentration of starting material in CTCs can pose a challenge to accurate sequencing, developments in WGA and WTA as well as the use of targeted panels have made it possible to capture the genetic diversity of metastatic tumor cells.

4.4. Fragmentomics

Fragmentomics in liquid biopsy has emerged as a promising approach for the detection and monitoring of cancer [165–169]. It involves analyzing the fragmentation patterns of cfDNA using NGS technologies (such as WES or WGS) without performing the DNA sonication step or using long-read sequencing [129,170,171]. Different computational and experimental approaches have been applied to study fragmentomics at varying resolutions and scales [169–171]. Various aspects of fragmentation can be taken advantage of, for example, length analysis and variant analysis. In cancer, it has been observed that ctDNA has a shorter median length and greater variability in size compared to cfDNA from healthy control subjects [171–173]. These findings suggest that fragmentomics analysis could be a useful tool for cancer diagnosis and monitoring, as the presence of ctDNA with abnormal size profiles may indicate the presence of cancer. Fragment length analysis of cell-free DNA can also be a useful tool for the detection of cancer-specific mutations. In a study of melanoma patients, the *BRAF V600E* mutation was found to occur more commonly at

shorter fragment lengths (132–145 bp) than the wild-type allele (165 bp) [174]. In lung cancer patients, size-selecting for shorter cell-free DNA fragment lengths substantially increased the detection of the *EGFR T790M* mutant allele [174]. A study using WGS to analyze 344 plasma samples from 200 cancer patients showed that size-selected cfDNA identified clinically actionable mutations and copy number alterations that are otherwise not detected [175]. Additionally, the analysis of ctDNA fragments can be used to detect minimal residual disease (MRD) [176]. The combined analysis of both sequence variant and size fragmentation was reported to improve the stratification of patients into risk groups (low and high risk) for MRD in later-stage lung cancers (Stage II-IIIA) for guiding treatment decisions [177]. Fragmentomics data have also been used to detect early-stage lung cancer and to classify early-stage cancer patients [178–180].

5. Bioinformatics Pipelines for Analyzing Liquid Biopsy NGS Data

The process of generating and analyzing sequencing data using NGS technology involves three stages: primary, secondary and tertiary analysis [181,182]. The primary analysis involves the initial processing of raw data generated by the sequencing instrument, including base calling and quality control checks. The secondary analysis is focused on the pre-processed data, and it involves aligning the data to a reference genome and identifying genetic variants, such as single nucleotide polymorphisms (SNPs) or structural variations. In the tertiary analysis, the biological significance of the variants is interpreted and analyzed in the context of other available data sources [182].

5.1. Sequence Data Processing

During the primary analysis, the sequencing instrument generates a raw binary file containing the nucleotide bases identified during the sequencing run. For Illumina sequencers, the primary output is a BCL (Base Call) file, which contains the raw signal intensities for each base call of each read generated during the run. BCL files are then converted into a text-based (ASCII) FASTQ format file using tools such as BCLConvert [183]. The FASTQ file is the standard format for sequence data and is used for downstream analysis [184]. Several quality control metrics are evaluated to assess the quality of the sequencing run and the data generated [182,185–187]. The most common metrics include yield (the total number of reads generated in a run), quality scores (a measure of the confidence in the base call accuracy for each position in the sequence read) and error rates (calculated as the mismatches occurring during read alignment to the phiX spike-in).

In the secondary analysis, the pre-processed data generated during the primary analysis is aligned to a reference genome or assembled de novo [182,188,189]. This process can be computationally expensive and time-consuming due to the need to realign the 150-nucleotide-long reads to their original position in the human genome, which spans around 3 billion nucleotides. To address this challenge, various algorithms have been developed to effectively map the reads back to the human genome, such as BWA [190], Bowtie2 [191] and others. While these aligners can be used for both DNA and RNA data, the latter may require specific settings or have specializations depending on the type of data being aligned. For example, RNA-Seq data can be aligned using algorithms, such as Bowtie and BWA, but splice events result in specific requirements for transcriptome alignments, such as accounting for splice junctions and allowing for accurate mapping of reads to the appropriate isoform, which requires using tools such as TopHat2 [192], HISAT2 [193,194], HISAT-3N [195] or STAR [196]. The input for the secondary analysis generally involves FASTQ files generated during the primary analysis. The output file generated during the alignment process is a Binary Alignment Map (BAM) or a Sequence Alignment Map (SAM) file [197]. BAM files are currently the standard because their storage footprint is considerably reduced compared to SAM files.

5.2. Sequence Data Interpretation

5.2.1. Biological Interpretation of the Sequencing Data

During the tertiary analysis stage, the processed data is annotated and interpreted, and several processes, such as variant calling, annotation and functional interpretation, take place to provide insights into the biological significance of the genetic variants identified [181,182]. Variant calling is the step where genetic variants, such as SNPs, indels and copy number variations (CNVs), are identified using tools such as GATK (Genome Analysis Toolkit), VarScan and SAMtools (Table 2) [198–200]. During the annotation process, publicly available databases, such as dbSNP, Ensembl, RefSeq, ClinVar, COSMIC and UCSC, are utilized to retrieve information regarding pathogenicity or significance from the identified genetic variants [201–206]. Tools, such as ANNOVAR, SnpEff and VEP (Variant Effect Predictor), are commonly used to position the genetic variants in a gene/transcript/genomic position context [207–209]. The functional interpretation step includes examining the functional relationships between genes and their involvement in biological pathways. Tools, such as IPA (Ingenuity Pathway Analysis), GSEA (Gene Set Enrichment Analysis), KEGG (Kyoto Encyclopedia of Genes and Genomes), Cytoscape, DAVID (Database for Annotation, Visualization and Integrated Discovery) and Enrichr are used to perform these tasks [210–215].

Table 2. Tools for computational analysis of NGS data.

Tool	Description	Resource	References
Variant calling			
GATK	Genome Analysis Toolkit has multiple applications, e.g., variant discovery, genotyping and mutation detection, quality control, coverage analysis and error correction.	https://gatk.broadinstitute.org/hc/en-us (accessed on 1 March 2023)	[198]
VarScan	Detects and characterizes variants, e.g., SNPs, indels and somatic mutations in tumor-normal pairs. Identifies low-frequency variants using Bayesian algorithms and statistical models for sensitivity and specificity.	https://varscan.sourceforge.net/ (accessed on 1 March 2023)	[199]
Vardict	Detects SNVs, indels and CNVs from tumors and tumor-normal pairs. Uses a combination of local realignment, base quality score recalibration and variant calling algorithms to identify variants. Handles data with high variability, e.g., low-coverage or high tumor heterogeneity.	https://github.com/AstraZeneca-NGS/VarDict (accessed on 1 March 2023)	[216]
Samtools	Analyzes alignment files in multiple formats, e.g., BAM, SAM and CRAM. Performs file conversion, sorting, indexing, filtering and merging. Quality control, coverage analysis and variant calling for reference genomes and alignment algorithms.	https://www.htslib.org/ (accessed on 1 March 2023)	[200]
Strelka2	Heuristic approach to detect SNVs, indels and structural variants. Employs a combination of probabilistic and machine learning methods to detect somatic mutations while minimizing false positives. Uses local assembly-based variant calling to improve variant detection sensitivity in regions with low read coverage or high levels of noise.	https://github.com/Illumina/strelka (accessed on 1 March 2023)	[217]
ANNOVAR	Enables genetic variants annotation in various genome builds, e.g., RefSeq, dbNSFP and gnomAD. Allows filtering and prioritization of variants based on the functional impact, population frequency, etc.	https://annovar.openbioinformatics.org/en/latest/ (accessed on 1 March 2023)	[207]
Variant annotation			
SnpEff	Annotation and functional analysis of genetic variants. Predict the effects of genetic variants on genes, transcripts and regulatory regions and classify variants based on their impact.	http://pcingola.github.io/SnpEff/ (accessed on 1 March 2023)	[208]

Table 2. *Cont.*

Tool	Description	Resource	References
VEP	Variant Effect Predictor performs analysis, annotation and prioritization of genomic variants in coding and non-coding regions	https://useast.ensembl.org/info/docs/tools/vep/index.html (accessed on 1 March 2023)	[209]
Functional interpretation			
GSEA	Gene Set Enrichment Analysis identifies enriched biological pathways, functions and processes based on the expression profiles of genes in a sample or dataset.	https://www.gsea-msigdb.org/gsea/index.jsp (accessed on 1 March 2023)	[210]
KEGG	Kyoto Encyclopedia of Genes and Genomes is a data and knowledge base of biological systems, e.g., metabolic pathways, regulatory networks and genetic information. A comprehensive set of reference genomes, gene annotations and pathway maps.	https://www.genome.jp/kegg/ (accessed on 1 March 2023)	[211]
Cytoscape	Open-source software for the visualization, analysis and interpretation of complex biological networks. Utilizes various data types, e.g., genetic, genomic, proteomic and metabolomic.	https://cytoscape.org/ (accessed on 1 March 2023)	[212]
DAVID	Resource for functional annotation and analysis of biological data. A comprehensive set of functional annotation tools, including gene ontology/pathway analysis and functional annotation clustering.	https://david.ncifcrf.gov/ (accessed on 1 March 2023)	[213]
Enrichr	Web-based analysis tool. Provides visualization summaries of collective functions of gene lists. Integrates public databases and annotations for identification and annotation of biological pathways, functions and processes associated with a set of genes or proteins.	https://maayanlab.cloud/Enrichr/ (accessed on 1 March 2023)	[214]
GeneMania	Identifies and analyzes functional gene networks. Uses combinations of functional genomics data sources, including protein-protein interactions, co-expression, genetic interactions and pathways to construct gene networks related to the biological function or disease.	http://genemania.org/ (accessed on 1 March 2023)	[215,218]
IPA	Ingenuity Pathway Analysis identifies key biological pathways, networks and functions associated with gene or protein sets. A range of visualization and reporting features. Supports various input and output file formats.	https://digitalinsights.qiagen.com/products-overview/discovery-insights-portfolio/analysis-and-visualization/qiagen-ipa/ (accessed on 1 March 2023)	Not applicable

Abbreviations: BAM: Binary Alignment Map; CNVs: Copy Number Variants; CRAM: Compressed Reference-oriented Alignment Map; indels: Insertion–deletion mutations; SAM: Sequence Alignment Map; SNPs: Single-nucleotide polymorphisms.

5.2.2. Bioinformatic Platforms for Analyzing Long-Read Sequence Data

Bioinformatic analysis of long-read sequencing data from Oxford Nanopore Technologies (ONT) and PacBio is a rapidly evolving field, and neural nets are often integrated into the newest tools. Long-read data analysis has been thoroughly reviewed elsewhere [125,219,220], and a comprehensive searchable and filterable catalog of long-read data analysis tools has been created (https://long-read-tools.org, accessed on 1 March 2023). Some important steps in long-read sequencing analysis include base calling, quality control, error correction, genome alignment and detection of genetic alterations. Base calling for long-read sequencing is somewhat more error-prone than for short-read sequencing, but physical updates to the sequencers and error correction programs have improved accuracy. The main PacBio base calling program is called CCS [221], and the most commonly used ONT base calling program is Guppy though others such as Flappie, Scrappie and Taiyaki are sometimes used [222]. More specialized programs for data analysis can be found in the long-read tools catalog linked above.

5.2.3. Analyzing DNA Whole-Genome Methylation Data

DNAm data numerically falls between 0 (no methylation) and 1 (complete methylation) per CpG, with intermediate methylation signifying CpGs with differential methylation. Methylation arrays (such as Illumina's 450 k HumanMethylation450 updated to the 850 k HumanMethylationEPIC) measure the intensity of methylation via two types of probes: the original Infinium 1 (Type 1) probes, which utilize two beads per CpG to read the methylated and unmethylated intensity (red or green fluorescence per bead depending upon the methylation status), and the Infinium II (Type 2) probes which use one bead per CpG emitting either a green fluorescence for methylation or red fluorescence for no methylation. Based on the intensity of these red and green fluorescence levels, methylation is measured by either beta values (methylated divided by methylated + unmethylated + 100) or MValue (log of methylated divided by unmethylated) [223]. Processing steps involve validation of internal control probes for quality control metrics, prediction of sample sex by sex-specific probes on the X and Y chromosomes and normalization. Many software packages exist for DNAm analysis depending on the method of profiling, with most written in the R Language for Statistical Programming [224,225].

The biggest caveat of DNAm profiling is the low global coverage of genome-wide CpGs [226,227], wherein the majority of CpGs assessed are within gene bodies, CpG Islands, CpG shores and CpG shelves, leaving intergenic regions largely unprofiled. The accuracy of profiling methods also depends upon the CpG density of a genome [228]. Efficient and accurate DNAm results require rigorous quality control metrics during sample collection, sample processing and sample distribution while methylation profiling, as well as accounting for technical variation during analysis [229].

5.2.4. Analyzing Single-Cell Sequence Data

Numerous platforms and software programs have been developed for processing single-cell sequencing data (Table 3), and a comprehensive searchable and filterable list is available online (https://www.scrna-tools.org/table, accessed on 1 March 2023). Frequently used platforms for quality control, visualization, and analysis of single-cell sequencing data include Seurat, Scanpy and Scater [230–232]. In most single-cell sequencing pipelines, to accomplish pre-processing and visualization, raw data are used to generate molecular barcodes for each cell. To remove reads from deceased and dying cells, quality control programs use molecular barcodes to measure covariates such as count depth, genes per barcode and mitochondrial genes per barcode, which can indicate unhealthy cells. Normalization is an important data processing step to account for artificial differences in expression observed due to sampling error during processing. Some normalization tools for single cells include MAST, SCDE and Basic [32]. The most frequently used formula for normalization is counts per million (CPM), which is the number of reads mapping to a particular feature divided by the total number of reads and multiplied by 10^6. Transcripts per million (TPM) is another common normalization formula for Single-cell RNA sequencing (scRNAseq) data, which normalizes for both read depth and gene length. Visualization of the output is often improved by dimensionality reduction methods, such as t-distributed stochastic neighbor embedding (t-SNE), which reduces the noise of differences between cells and better resolves similarities, as well as Uniform Manifold Approximation and Projection (UMAP) [233]. After these processing steps, downstream analysis is possible. This tends to focus on clustering cells by shared traits or cell trajectory inference, which entails mapping cell state changes over time. Clustering is often carried out using a k-means clustering algorithm. The popular Seurat platform uses the Louvain algorithm with K-nearest-neighbor graphs [162]. Annotation of the identified cell clusters can be aided with databases such as The Human Cell Atlas [234], CellMarker [235] and automated cluster annotation platforms, such as scmap, can be used [236]. Cell trajectory inference can be carried out with Monocle, Wanderlust or Slingshot [237–239]. For differential gene expression analysis between clusters, tools such as DESeq2 or EdgeR may be used [240,241].

Table 3. Platforms designed for bioinformatic analysis of single-cell sequencing data.

Tool	Brief Description	Resource	References
Seurat	R-based platform for raw data processing, paired sample analysis and visualizations. Uses machine learning and clustering algorithms to identify biological features. Assesses cellular heterogeneity via normalization, dimensionality reduction and integration tools.	http://satijalab.org/ (accessed on 1 March 2023)	[230,242]
Monocle	R-based scRNA-Seq analysis software. It uses algorithms and machine learning to determine cell developmental trajectories, identify molecular pathways and track changes in gene expression.	http://cole-trapnell-lab. github.io/monocle-release/ (accessed on 1 March 2023)	[237,243,244]
ChromVAR	R package for analyzing variations in chromatin accessibility in scATAC-Seq data to identify associated motifs or genomic annotations. It uses visualization techniques to detect and highlight changes in gene expression and provides users with powerful statistical methods. It is also capable of detecting and correlating molecular pathways.	https://greenleaflab.github. io/chromVAR/ (accessed on 1 March 2023)	[245]
DRAGEN Single-Cell RNA Pipeline	Cloud-based platform to analyze scRNA-Seq data: aligning and mapping reads, detecting features and biomarkers and generating visualizations. It processes multiplexed scRNA-Seq datasets from reads to a cell-by-gene UMI count gene expression matrix. Features splice-aware RNAseq alignment and matching to annotated genes for transcript reads, cell-barcode and UMI error correction and QC metrics.	http://illumina.com/ (accessed on 1 March 2023)	Not applicable
Tapestri	Pipeline to analyze scRNA-Seq data generated by the Tapestri platform. It Includes sequence import, data analysis and visualization capabilities. The software enables variant identification, including SNVs and CNVs, at clonal and subclonal levels.	https://support.missionbio. com/hc/en-us/categories/36 0002505454-Tapestri-Insights (accessed on 1 March 2023)	Not applicable
Scanorama	Integrates data from heterogenous scRNA-seq experiments via detecting common cell types among datasets. Identifies datasets, e.g., cells with similar transcriptional profiles, and leverages the matches for batch correction and integration. Can handle different dataset sizes and sources and does not require all datasets to share a cell population.	https://cb.csail.mit.edu/cb/ scanorama/ (accessed on 1 March 2023)	[246]
scmap v1.1.5	An R package that projects cells from a scRNA-Seq data set onto cell types or individual cells from various experiments. It is a widely applicable projection method, detecting the best-matching cell type or individual cell in the reference. It allows fast feature selection, centroid calculation and index creation.	https://scmap.sanger.ac.uk/ scmap/ (accessed on 1 March 2023)	[236]
Scrublet v0.1	Single-Cell Remover of Doublets, acronym Scrublet, is a framework for predicting the effect of multiplets in analysis and also identifies problematic multiplets. It can identify neotypic multiplets for an analyzed dataset. The Scrublet classifier can implement arbitrary functions for preprocessing and embedding of single-cell data.	https://github.com/ AllonKleinLab/scrublet (accessed on 1 March 2023)	[247]
CellRanger v2.2.0	A set of analysis pipelines that can process Chromium single-cell data to align reads, generate feature-barcode matrices, perform clustering, amongst other tasks. It contains five pipelines for the 3′ Single Cell Gene Expression Solutions and similar products.	https://support.10 xgenomics.com/single-cell-gene-expression/software (accessed on 1 March 2023)	Not applicable

Table 3. *Cont.*

Tool	Brief Description	Resource	References
CITE-seq-count v1.2	Python package that aids counting antibody tags from CITE-Seq or cell hashing experiments.	https://github.com/Hoohm/CITE-seq-Count (accessed on 1 March 2023)	[248]
Drop-seq tools v2.0.0	Java tool to analyze profiling of individual cells (UMI cells encapsulated in droplets for oligodT sequencing). Libraries produce paired-end reads (read 1: cell barcode and UMI; read 2: cDNA sequence) for identity and abundance of the transcripts in each cell.	https://github.com/broadinstitute/Drop-seq (accessed on 1 March 2023)	Not applicable
deepTools v3.1.2	Galaxy-based web server for processing and visualizing deeply sequenced data. Allows completion of bioinformatic workflows to integrative analyses. Supports four tasks: quality control, data processing and normalization, data integration and visualization.	Not applicable	[249]

Abbreviations: CITE-Seq: Cellular Indexing of Transcriptomes and Epitopes by Sequencing; CNV: Copy number variant; OligodT: Oligo-deoxythymidine; QC: Quality control; scATAC-Seq: Single-cell Assay for Transposase-Accessible Chromatin using sequencing; scRNA-Seq: Single-cell ribonucleic acid sequencing; SNV: Single nucleotide variant; UMI: Unique molecular identifier.

5.2.5. Other Software for Analyzing Liquid Biopsy Samples

Some tools have been developed specifically for long-read cfDNA sequencing data from ONT devices. A program called cfNano uses deconvolution to detect cancer-associated signatures in ctDNA, including nucleosome footprinting, copy number alterations, methylation changes and fragmentation patterns [250,251]. Another program called FrEIA uses variations in the sequences at the ends of cfDNA fragments to improve the sensitivity of cancer signal detection [252,253]. Nanomonsv is designed to detect somatic cancer-associated structural variants in paired tumor and normal samples [254,255]. Nanovar is a SV caller with the ability of detecting variants from low-depth, long-read sequencing (homozygous SVs can be detected using 4×, while heterozygous SVs are detected at a threshold of 8×) [256].

The development of bioinformatic tools to sequence single CTCs often includes quality control to account for errors introduced during whole genome or whole exome amplification. Several tools exist for different types of CTC studies [32]. Monovar is a variant caller for single-cell data [257]. OncoNEM and SCITE are designed to use single-cell data to trace the evolutionary trajectory of cancers [258,259]. RaceID2 and GiniClust are tools for identifying a cell's tissue of origin based on single-cell sequencing [260,261]. Ginkgo is a web platform that is used to identify copy number variants in single-cell sequencing data to construct a phylogenetic tree [262], which has been used to study the heterogeneity of CTCs [263].

cfDNApipe is an integrated tool for WGS or bisulphite cfDNA sequencing data that performs quality control and finds differentially methylated regions, copy number variation and alterations in fragment lengths, which facilitates tissue of origin analysis or fragmentomic analysis [264,265].

6. Conclusions

The use of liquid biopsy and NGS technology has revolutionized the diagnosis and management of lung cancer, providing a non-invasive alternative to traditional solid tissue biopsies. Liquid biopsy enables the detection of cancer-specific mutations and other biomarkers in a minimally invasive manner, allowing for serial testing and monitoring of lung cancer progression or response to treatment.

While the use of liquid biopsy and NGS technology has brought numerous benefits to the field of lung cancer diagnosis and management, there are also some limitations and challenges. One of the major limitations compared to solid tissue biopsies is the signal-to-

noise ratio. The low abundance of circulating tumor DNA/RNA, circulating tumor cells and extracellular vesicles in biological fluids, detecting cancer-specific mutations and other biomarkers, can be problematic, as these signals can easily be drowned out by processing protocols, background noise and/or sequencing errors. As a result, the sensitivity and specificity of liquid biopsy and NGS-based analysis can be affected, leading to false-positive or false-negative results. Consequently, one of the main challenges is the need for increased sensitivity and specificity of liquid biopsy analysis, particularly in detecting early-stage lung cancer or minimal residual disease. This is especially important for patients who are not eligible for surgery or who have a high risk of recurrence.

NGS technology has enabled a comprehensive and highly detailed analysis of liquid biopsy samples, providing valuable information for cancer diagnosis, monitoring and treatment selection. With its ability to analyze multiple types of genetic alterations (including SNVs, indels and copy number/structural variations) and biomarkers simultaneously, NGS has opened up new avenues for identifying potential therapeutic targets and developing precision medicine strategies. Moreover, emerging sequencing technologies, such as long reads sequencing and single-cell sequencing, are expected to expand the nature of biomarkers that can be detected in liquid biopsy. Long reads sequencing has the potential to detect complex genomic rearrangements at an unprecedented level of detail, while single-cell sequencing has the potential to provide a more comprehensive analysis of the heterogeneity of cancer cells in liquid biopsy, which can play a significant role in the development of resistance to treatment. The combination of current NGS approaches and emerging sequencing technologies is expected to further enhance the precision and efficacy of lung cancer diagnosis and treatment.

In conclusion, the use of NGS in the analysis of liquid biopsy has revolutionized lung cancer diagnostics and management; however, limitations and challenges remain to fully realize its potential. Further research will improve the ability of NGS to further unlock the biological information contained in liquid biopsies.

Author Contributions: L.J.B. and V.G.P.S. contributed equally to this work. L.J.B., V.G.P.S., P.P.R., W.L.L. and V.D.M. conceptualized and designed the article. All authors contributed to data collection, curation, analysis, interpretation and writing. L.J.B., V.G.P.S., G.L.S., P.P.R., W.L.L., A.F., M.E.P., N.T., K.B. and V.D.M. were responsible for editing. J.T., M.E., V.G.P.S., M.D.S. and G.L.S. developed the display items. P.P.R., W.L.L. and V.D.M. are the principal investigators. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Siegel, R.L.; Miller, K.D.; Fuchs, H.E.; Jemal, A. Cancer Statistics, 2022. *CA Cancer J. Clin.* **2022**, *72*, 7–33. [CrossRef] [PubMed]
2. Li, W.; Liu, J.-B.; Hou, L.-K.; Yu, F.; Zhang, J.; Wu, W.; Tang, X.-M.; Sun, F.; Lu, H.-M.; Deng, J.; et al. Liquid Biopsy in Lung Cancer: Significance in Diagnostics, Prediction, and Treatment Monitoring. *Mol. Cancer* **2022**, *21*, 25. [CrossRef]

3. Shen, H.; Jin, Y.; Zhao, H.; Wu, M.; Zhang, K.; Wei, Z.; Wang, X.; Wang, Z.; Li, Y.; Yang, F.; et al. Potential Clinical Utility of Liquid Biopsy in Early-Stage Non-Small Cell Lung Cancer. *BMC Med.* **2022**, *20*, 480. [CrossRef]

4. Gristina, V.; Barraco, N.; La Mantia, M.; Castellana, L.; Insalaco, L.; Bono, M.; Perez, A.; Sardo, D.; Inguglia, S.; Iacono, F.; et al. Clinical Potential of Circulating Cell-Free DNA (CfDNA) for Longitudinally Monitoring Clinical Outcomes in the First-Line Setting of Non-Small-Cell Lung Cancer (NSCLC): A Real-World Prospective Study. *Cancers* **2022**, *14*, 6013. [CrossRef] [PubMed]

5. Guibert, N.; Pradines, A.; Favre, G.; Mazieres, J. Current and Future Applications of Liquid Biopsy in Nonsmall Cell Lung Cancer from Early to Advanced Stages. *Eur. Respir. Rev. Off. J. Eur. Respir. Soc.* **2020**, *29*, 190052. [CrossRef]

6. Bracht, J.W.P.; Mayo-de-Las-Casas, C.; Berenguer, J.; Karachaliou, N.; Rosell, R. The Present and Future of Liquid Biopsies in Non-Small Cell Lung Cancer: Combining Four Biosources for Diagnosis, Prognosis, Prediction, and Disease Monitoring. *Curr. Oncol. Rep.* **2018**, *20*, 70. [CrossRef] [PubMed]

7. Lone, S.N.; Nisar, S.; Masoodi, T.; Singh, M.; Rizwan, A.; Hashem, S.; El-Rifai, W.; Bedognetti, D.; Batra, S.K.; Haris, M.; et al. Liquid Biopsy: A Step Closer to Transform Diagnosis, Prognosis and Future of Cancer Treatments. *Mol. Cancer* **2022**, *21*, 79. [CrossRef]

8. Hasenleithner, S.O.; Speicher, M.R. A Clinician's Handbook for Using CtDNA throughout the Patient Journey. *Mol. Cancer* **2022**, *21*, 81. [CrossRef]

9. Saarenheimo, J.; Eigeliene, N.; Andersen, H.; Tiirola, M.; Jekunen, A. The Value of Liquid Biopsies for Guiding Therapy Decisions in Non-Small Cell Lung Cancer. *Front. Oncol.* **2019**, *9*, 129. [CrossRef]

10. Lightbody, G.; Haberland, V.; Browne, F.; Taggart, L.; Zheng, H.; Parkes, E.; Blayney, J.K. Review of Applications of High-Throughput Sequencing in Personalized Medicine: Barriers and Facilitators of Future Progress in Research and Clinical Application. *Brief. Bioinform.* **2019**, *20*, 1795–1811. [CrossRef] [PubMed]

11. Chen, M.; Zhao, H. Next-Generation Sequencing in Liquid Biopsy: Cancer Screening and Early Detection. *Hum. Genom.* **2019**, *13*, 34. [CrossRef]

12. Pantel, K.; Alix-Panabières, C. Liquid Biopsy and Minimal Residual Disease—Latest Advances and Implications for Cure. *Nat. Rev. Clin. Oncol.* **2019**, *16*, 409–424. [CrossRef]

13. Lin, C.; Liu, X.; Zheng, B.; Ke, R.; Tzeng, C.-M. Liquid Biopsy, CtDNA Diagnosis through NGS. *Life* **2021**, *11*, 890. [CrossRef] [PubMed]

14. Hasegawa, N.; Kohsaka, S.; Kurokawa, K.; Shinno, Y.; Takeda Nakamura, I.; Ueno, T.; Kojima, S.; Kawazu, M.; Suehara, Y.; Ishijima, M.; et al. Highly Sensitive Fusion Detection Using Plasma Cell-Free RNA in Non-Small-Cell Lung Cancers. *Cancer Sci.* **2021**, *112*, 4393–4403. [CrossRef] [PubMed]

15. Siravegna, G.; Marsoni, S.; Siena, S.; Bardelli, A. Integrating Liquid Biopsies into the Management of Cancer. *Nat. Rev. Clin. Oncol.* **2017**, *14*, 531–548. [CrossRef] [PubMed]

16. Palmieri, M.; Frullanti, E. Different Liquid Biopsies for the Management of Non-Small Cell Lung Cancer in the Mutational Oncology Era. *Med. Sci.* **2023**, *11*, 8. [CrossRef] [PubMed]

17. Lianidou, E.; Pantel, K. Liquid Biopsies. *Genes. Chromosomes Cancer* **2019**, *58*, 219–232. [CrossRef]

18. Soyano, A.E.; Dholaria, B.; Marin-Acevedo, J.A.; Diehl, N.; Hodge, D.; Luo, Y.; Manochakian, R.; Chumsri, S.; Adjei, A.; Knutson, K.L.; et al. Peripheral Blood Biomarkers Correlate with Outcomes in Advanced Non-Small Cell Lung Cancer Patients Treated with Anti-PD-1 Antibodies. *J. Immunother. Cancer* **2018**, *6*, 129. [CrossRef]

19. Underwood, J.J.; Quadri, R.S.; Kalva, S.P.; Shah, H.; Sanjeevaiah, A.R.; Beg, M.S.; Sutphin, P.D. Liquid Biopsy for Cancer: Review and Implications for the Radiologist. *Radiology* **2020**, *294*, 5–17. [CrossRef]

20. Michela, B. Liquid Biopsy: A Family of Possible Diagnostic Tools. *Diagnostics* **2021**, *11*, 1391. [CrossRef]

21. Fernández-Lázaro, D.; García Hernández, J.L.; García, A.C.; Córdova Martínez, A.; Mielgo-Ayuso, J.; Cruz-Hernández, J.J. Liquid Biopsy as Novel Tool in Precision Medicine: Origins, Properties, Identification and Clinical Perspective of Cancer's Biomarkers. *Diagnostics* **2020**, *10*, 215. [CrossRef]

22. Young, R.; Pailler, E.; Billiot, F.; Drusch, F.; Barthelemy, A.; Oulhen, M.; Besse, B.; Soria, J.-C.; Farace, F.; Vielh, P. Circulating Tumor Cells in Lung Cancer. *Acta Cytol.* **2012**, *56*, 655–660. [CrossRef]

23. Lindsay, C.R.; Faugeroux, V.; Michiels, S.; Pailler, E.; Facchinetti, F.; Ou, D.; Bluthgen, M.V.; Pannet, C.; Ngo-Camus, M.; Bescher, G.; et al. A Prospective Examination of Circulating Tumor Cell Profiles in Non-Small-Cell Lung Cancer Molecular Subgroups. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* **2017**, *28*, 1523–1531. [CrossRef]

24. Tamminga, M.; de Wit, S.; Schuuring, E.; Timens, W.; Terstappen, L.W.M.M.; Hiltermann, T.J.N.; Groen, H.J.M. Circulating Tumor Cells in Lung Cancer Are Prognostic and Predictive for Worse Tumor Response in Both Targeted- and Chemotherapy. *Transl. Lung Cancer Res.* **2019**, *8*, 854–861. [CrossRef] [PubMed]

25. Li, X.; Li, Y.; Shao, W.; Li, Z.; Zhao, R.; Ye, Z. Strategies for Enrichment of Circulating Tumor Cells. *Transl. Cancer Res.* **2020**, *9*, 2012–2025. [CrossRef]

26. Sharma, S.; Zhuang, R.; Long, M.; Pavlovic, M.; Kang, Y.; Ilyas, A.; Asghar, W. Circulating Tumor Cell Isolation, Culture, and Downstream Molecular Analysis. *Biotechnol. Adv.* **2018**, *36*, 1063–1078. [CrossRef] [PubMed]

27. Ju, S.; Chen, C.; Zhang, J.; Xu, L.; Zhang, X.; Li, Z.; Chen, Y.; Zhou, J.; Ji, F.; Wang, L. Detection of Circulating Tumor Cells: Opportunities and Challenges. *Biomark. Res.* **2022**, *10*, 58. [CrossRef] [PubMed]

28. Tong, B.; Wang, M. Circulating Tumor Cells in Patients with Lung Cancer: Developments and Applications for Precision Medicine. *Future Oncol. Lond. Engl.* **2019**, *15*, 2531–2542. [CrossRef]

29. Bhagwat, N.; Dulmage, K.; Pletcher, C.H.; Wang, L.; DeMuth, W.; Sen, M.; Balli, D.; Yee, S.S.; Sa, S.; Tong, F.; et al. An Integrated Flow Cytometry-Based Platform for Isolation and Molecular Characterization of Circulating Tumor Single Cells and Clusters. *Sci. Rep.* **2018**, *8*, 5035. [CrossRef]

30. Basu, S.; Campbell, H.M.; Dittel, B.N.; Ray, A. Purification of Specific Cell Population by Fluorescence Activated Cell Sorting (FACS). *J. Vis. Exp.* **2010**, *41*, e1546. [CrossRef]

31. Volovetskiy, A.B.; Malinina, P.A.; Kapitannikova, A.Y.; Smetanina, S.V.; Kruglova, I.A.; Maslennikova, A.V. Isolation of Circulating Tumor Cells from Peripheral Blood Samples of Cancer Patients Using Microfluidic Technology. *Sovrem. Tekhnol. Med.* **2020**, *12*, 62–68. [CrossRef] [PubMed]

32. Rossi, E.; Zamarchi, R. Single-Cell Analysis of Circulating Tumor Cells: How Far Have We Come in the -Omics Era? *Front. Genet.* **2019**, *10*, 958. [CrossRef] [PubMed]

33. Ramirez, J.-M.; Fehm, T.; Orsini, M.; Cayrefourcq, L.; Maudelonde, T.; Pantel, K.; Alix-Panabières, C. Prognostic Relevance of Viable Circulating Tumor Cells Detected by EPISPOT in Metastatic Breast Cancer Patients. *Clin. Chem.* **2014**, *60*, 214–221. [CrossRef]

34. Eslami-S, Z.; Cortés-Hernández, L.E.; Thomas, F.; Pantel, K.; Alix-Panabières, C. Functional Analysis of Circulating Tumour Cells: The KEY to Understand the Biology of the Metastatic Cascade. *Br. J. Cancer* **2022**, *127*, 800–810. [CrossRef]

35. Park, E.S.; Yan, J.P.; Ang, R.A.; Lee, J.H.; Deng, X.; Duffy, S.P.; Beja, K.; Annala, M.; Black, P.C.; Chi, K.N.; et al. Isolation and Genome Sequencing of Individual Circulating Tumor Cells Using Hydrogel Encapsulation and Laser Capture Microdissection. *Lab. Chip* **2018**, *18*, 1736–1749. [CrossRef] [PubMed]

36. Sánchez-Herrero, E.; Serna-Blasco, R.; Robado de Lope, L.; González-Rumayor, V.; Romero, A.; Provencio, M. Circulating Tumor DNA as a Cancer Biomarker: An Overview of Biological Features and Factors That May Impact on CtDNA Analysis. *Front. Oncol.* **2022**, *12*, 943253. [CrossRef]

37. Parkinson, C.A.; Gale, D.; Piskorz, A.M.; Biggs, H.; Hodgkin, C.; Addley, H.; Freeman, S.; Moyle, P.; Sala, E.; Sayal, K.; et al. Exploratory Analysis of TP53 Mutations in Circulating Tumour DNA as Biomarkers of Treatment Response for Patients with Relapsed High-Grade Serous Ovarian Carcinoma: A Retrospective Study. *PLoS Med.* **2016**, *13*, e1002198. [CrossRef]

38. Chabon, J.J.; Hamilton, E.G.; Kurtz, D.M.; Esfahani, M.S.; Moding, E.J.; Stehr, H.; Schroers-Martin, J.; Nabet, B.Y.; Chen, B.; Chaudhuri, A.A.; et al. Integrating Genomic Features for Non-Invasive Early Lung Cancer Detection. *Nature* **2020**, *580*, 245–251. [CrossRef]

39. Abbosh, C.; Birkbak, N.J.; Wilson, G.A.; Jamal-Hanjani, M.; Constantin, T.; Salari, R.; Le Quesne, J.; Moore, D.A.; Veeriah, S.; Rosenthal, R.; et al. Phylogenetic CtDNA Analysis Depicts Early-Stage Lung Cancer Evolution. *Nature* **2017**, *545*, 446–451. [CrossRef]

40. Hai, L.; Li, L.; Liu, Z.; Tong, Z.; Sun, Y. Whole-Genome Circulating Tumor DNA Methylation Landscape Reveals Sensitive Biomarkers of Breast Cancer. *MedComm* **2022**, *3*, e134. [CrossRef]

41. Diehl, F.; Li, M.; Dressman, D.; He, Y.; Shen, D.; Szabo, S.; Diaz, L.A.; Goodman, S.N.; David, K.A.; Juhl, H.; et al. Detection and Quantification of Mutations in the Plasma of Patients with Colorectal Tumors. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 16368–16373. [CrossRef] [PubMed]

42. de Kock, R.; van den Borne, B.; Youssef-El Soud, M.; Belderbos, H.; Brunsveld, L.; Scharnhorst, V.; Deiman, B. Therapy Monitoring of EGFR-Positive Non-Small-Cell Lung Cancer Patients Using DdPCR Multiplex Assays. *J. Mol. Diagn.* **2021**, *23*, 495–505. [CrossRef] [PubMed]

43. Chan, K.C.A.; Jiang, P.; Zheng, Y.W.L.; Liao, G.J.W.; Sun, H.; Wong, J.; Siu, S.S.N.; Chan, W.C.; Chan, S.L.; Chan, A.T.C.; et al. Cancer Genome Scanning in Plasma: Detection of Tumor-Associated Copy Number Aberrations, Single-Nucleotide Variants, and Tumoral Heterogeneity by Massively Parallel Sequencing. *Clin. Chem.* **2013**, *59*, 211–224. [CrossRef] [PubMed]

44. Forshew, T.; Murtaza, M.; Parkinson, C.; Gale, D.; Tsui, D.W.Y.; Kaper, F.; Dawson, S.-J.; Piskorz, A.M.; Jimenez-Linan, M.; Bentley, D.; et al. Noninvasive Identification and Monitoring of Cancer Mutations by Targeted Deep Sequencing of Plasma DNA. *Sci. Transl. Med.* **2012**, *4*, 136ra68. [CrossRef]

45. Wu, Z.; Yang, Z.; Dai, Y.; Zhu, Q.; Chen, L.-A. Update on Liquid Biopsy in Clinical Management of Non-Small Cell Lung Cancer. *OncoTargets Ther.* **2019**, *12*, 5097–5109. [CrossRef]

46. Tzimagiorgis, G.; Michailidou, E.Z.; Kritis, A.; Markopoulos, A.K.; Kouidou, S. Recovering Circulating Extracellular or Cell-Free RNA from Bodily Fluids. *Cancer Epidemiol.* **2011**, *35*, 580–589. [CrossRef]

47. Ishiba, T.; Hoffmann, A.-C.; Usher, J.; Elshimali, Y.; Sturdevant, T.; Dang, M.; Jaimes, Y.; Tyagi, R.; Gonzales, R.; Grino, M.; et al. Frequencies and Expression Levels of Programmed Death Ligand 1 (PD-L1) in Circulating Tumor RNA (CtRNA) in Various Cancer Types. *Biochem. Biophys. Res. Commun.* **2018**, *500*, 621–625. [CrossRef]

48. Liu, Z.; Wang, T.; Yang, X.; Zhou, Q.; Zhu, S.; Zeng, J.; Chen, H.; Sun, J.; Li, L.; Xu, J.; et al. Polyadenylation Ligation-Mediated Sequencing (PALM-Seq) Characterizes Cell-Free Coding and Non-Coding RNAs in Human Biofluids. *Clin. Transl. Med.* **2022**, *12*, e987. [CrossRef]

49. Kopreski, M.S.; Benko, F.A.; Kwak, L.W.; Gocke, C.D. Detection of Tumor Messenger RNA in the Serum of Patients with Malignant Melanoma. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **1999**, *5*, 1961–1965.

50. Schwarzenbach, H.; Hoon, D.S.B.; Pantel, K. Cell-Free Nucleic Acids as Biomarkers in Cancer Patients. *Nat. Rev. Cancer* **2011**, *11*, 426–437. [CrossRef]

51. Cabús, L.; Lagarde, J.; Curado, J.; Lizano, E.; Pérez-Boza, J. Current Challenges and Best Practices for Cell-Free Long RNA Biomarker Discovery. *Biomark. Res.* **2022**, *10*, 62. [CrossRef] [PubMed]

52. Bruno, R.; Fontanini, G. Next Generation Sequencing for Gene Fusion Analysis in Lung Cancer: A Literature Review. *Diagnostics* **2020**, *10*, 521. [CrossRef] [PubMed]

53. Mensah, M.; Borzi, C.; Verri, C.; Suatoni, P.; Conte, D.; Pastorino, U.; Orazio, F.; Sozzi, G.; Boeri, M. MicroRNA Based Liquid Biopsy: The Experience of the Plasma MiRNA Signature Classifier (MSC) for Lung Cancer Screening. *J. Vis. Exp.* **2017**, *128*, 56326. [CrossRef]

54. Costa, C.; Giménez-Capitán, A.; Karachaliou, N.; Rosell, R. Comprehensive Molecular Screening: From the RT-PCR to the RNA-Seq. *Transl. Lung Cancer Res.* **2013**, *2*, 87–91. [CrossRef]

55. Forder, A.; Hsing, C.-Y.; Trejo Vazquez, J.; Garnis, C. Emerging Role of Extracellular Vesicles and Cellular Communication in Metastasis. *Cells* **2021**, *10*, 3429. [CrossRef] [PubMed]

56. Yu, D.; Li, Y.; Wang, M.; Gu, J.; Xu, W.; Cai, H.; Fang, X.; Zhang, X. Exosomes as a New Frontier of Cancer Liquid Biopsy. *Mol. Cancer* **2022**, *21*, 56. [CrossRef]

57. Melo, S.A.; Luecke, L.B.; Kahlert, C.; Fernandez, A.F.; Gammon, S.T.; Kaye, J.; LeBleu, V.S.; Mittendorf, E.A.; Weitz, J.; Rahbari, N.; et al. Glypican-1 Identifies Cancer Exosomes and Detects Early Pancreatic Cancer. *Nature* **2015**, *523*, 177–182. [CrossRef] [PubMed]

58. Barile, L.; Vassalli, G. Exosomes: Therapy Delivery Tools and Biomarkers of Diseases. *Pharmacol. Ther.* **2017**, *174*, 63–78. [CrossRef]

59. Fang, T.; Lv, H.; Lv, G.; Li, T.; Wang, C.; Han, Q.; Yu, L.; Su, B.; Guo, L.; Huang, S.; et al. Tumor-Derived Exosomal MiR-1247-3p Induces Cancer-Associated Fibroblast Activation to Foster Lung Metastasis of Liver Cancer. *Nat. Commun.* **2018**, *9*, 191. [CrossRef]

60. Johnstone, R.M.; Adam, M.; Hammond, J.R.; Orr, L.; Turbide, C. Vesicle Formation during Reticulocyte Maturation. Association of Plasma Membrane Activities with Released Vesicles (Exosomes). *J. Biol. Chem.* **1987**, *262*, 9412–9420. [CrossRef]

61. Kanwar, S.S.; Dunlay, C.J.; Simeone, D.M.; Nagrath, S. Microfluidic Device (ExoChip) for on-Chip Isolation, Quantification and Characterization of Circulating Exosomes. *Lab Chip* **2014**, *14*, 1891–1900. [CrossRef]

62. Théry, C.; Witwer, K.W.; Aikawa, E.; Alcaraz, M.J.; Anderson, J.D.; Andriantsitohaina, R.; Antoniou, A.; Arab, T.; Archer, F.; Atkin-Smith, G.K.; et al. Minimal Information for Studies of Extracellular Vesicles 2018 (MISEV2018): A Position Statement of the International Society for Extracellular Vesicles and Update of the MISEV2014 Guidelines. *J. Extracell. Vesicles* **2018**, *7*, 1535750. [CrossRef]

63. Palmirotta, R.; Lovero, D.; Cafforio, P.; Felici, C.; Mannavola, F.; Pellè, E.; Quaresmini, D.; Tucci, M.; Silvestris, F. Liquid Biopsy of Cancer: A Multimodal Diagnostic Tool in Clinical Oncology. *Ther. Adv. Med. Oncol.* **2018**, *10*, 1758835918794630. [CrossRef] [PubMed]

64. Ma, S.; Zhou, M.; Xu, Y.; Gu, X.; Zou, M.; Abudushalamu, G.; Yao, Y.; Fan, X.; Wu, G. Clinical Application and Detection Techniques of Liquid Biopsy in Gastric Cancer. *Mol. Cancer* **2023**, *22*, 7. [CrossRef] [PubMed]

65. Nikanjam, M.; Kato, S.; Kurzrock, R. Liquid Biopsy: Current Technology and Clinical Applications. *J. Hematol. Oncol.* **2022**, *15*, 131. [CrossRef]

66. Paweletz, C.P.; Sacher, A.G.; Raymond, C.K.; Alden, R.S.; O'Connell, A.; Mach, S.L.; Kuang, Y.; Gandhi, L.; Kirschmeier, P.; English, J.M.; et al. Bias-Corrected Targeted Next-Generation Sequencing for Rapid, Multiplexed Detection of Actionable Alterations in Cell-Free DNA from Advanced Lung Cancer Patients. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **2016**, *22*, 915–922. [CrossRef]

67. Cescon, D.W.; Bratman, S.V.; Chan, S.M.; Siu, L.L. Circulating Tumor DNA and Liquid Biopsy in Oncology. *Nat. Cancer* **2020**, *1*, 276–290. [CrossRef] [PubMed]

68. Premarket Approval (PMA). Available online: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpma/pma.cfm?id=P150 044 (accessed on 1 March 2023).

69. EMA Iressa. Available online: https://www.ema.europa.eu/en/medicines/human/EPAR/iressa (accessed on 5 March 2023).

70. EMA Tagrisso. Available online: https://www.ema.europa.eu/en/medicines/human/EPAR/tagrisso (accessed on 5 March 2023).

71. Rolfo, C.; Mack, P.; Scagliotti, G.V.; Aggarwal, C.; Arcila, M.E.; Barlesi, F.; Bivona, T.; Diehn, M.; Dive, C.; Dziadziuszko, R.; et al. Liquid Biopsy for Advanced NSCLC: A Consensus Statement From the International Association for the Study of Lung Cancer. *J. Thorac. Oncol. Off. Publ. Int. Assoc. Study Lung Cancer* **2021**, *16*, 1647–1662. [CrossRef]

72. Fleischhacker, M.; Schmidt, B. Circulating Nucleic Acids (CNAs) and Cancer—A Survey. *Biochim. Biophys. Acta* **2007**, *1775*, 181–232. [CrossRef]

73. Ma, M.; Zhu, H.; Zhang, C.; Sun, X.; Gao, X.; Chen, G. "Liquid Biopsy"-CtDNA Detection with Great Potential and Challenges. *Ann. Transl. Med.* **2015**, *3*, 235. [CrossRef] [PubMed]

74. Köhn, L.; Johansson, M.; Grankvist, K.; Nilsson, J. Liquid Biopsies in Lung Cancer-Time to Implement Research Technologies in Routine Care? *Ann. Transl. Med.* **2017**, *5*, 278. [CrossRef]

75. Mosko, M.J.; Nakorchevsky, A.A.; Flores, E.; Metzler, H.; Ehrich, M.; van den Boom, D.J.; Sherwood, J.L.; Nygren, A.O.H. Ultrasensitive Detection of Multiplexed Somatic Mutations Using MALDI-TOF Mass Spectrometry. *J. Mol. Diagn.* **2016**, *18*, 23–31. [CrossRef]

76. Milbury, C.A.; Chen, C.C.; Mamon, H.; Liu, P.; Santagata, S.; Makrigiorgos, G.M. Multiplex Amplification Coupled with COLD-PCR and High Resolution Melting Enables Identification of Low-Abundance Mutations in Cancer Samples with Low DNA Content. *J. Mol. Diagn.* **2011**, *13*, 220–232. [CrossRef] [PubMed]

77. Spindler, K.-L.G.; Pallisgaard, N.; Vogelius, I.; Jakobsen, A. Quantitative Cell-Free DNA, KRAS, and BRAF Mutations in Plasma from Patients with Metastatic Colorectal Cancer during Treatment with Cetuximab and Irinotecan. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **2012**, *18*, 1177–1185. [CrossRef]

78. Watanabe, K.; Fukuhara, T.; Tsukita, Y.; Morita, M.; Suzuki, A.; Tanaka, N.; Terasaki, H.; Nukiwa, T.; Maemondo, M. EGFR Mutation Analysis of Circulating Tumor DNA Using an Improved PNA-LNA PCR Clamp Method. *Can. Respir. J.* **2016**, *2016*, 5297329. [CrossRef]

79. Oh, J.E.; Lim, H.S.; An, C.H.; Jeong, E.G.; Han, J.Y.; Lee, S.H.; Yoo, N.J. Detection of Low-Level KRAS Mutations Using PNA-Mediated Asymmetric PCR Clamping and Melting Curve Analysis with Unlabeled Probes. *J. Mol. Diagn.* **2010**, *12*, 418–424. [CrossRef]

80. Diehl, F.; Li, M.; He, Y.; Kinzler, K.W.; Vogelstein, B.; Dressman, D. BEAMing: Single-Molecule PCR on Microparticles in Water-in-Oil Emulsions. *Nat. Methods* **2006**, *3*, 551–559. [CrossRef] [PubMed]

81. Li, C.; He, Q.; Liang, H.; Cheng, B.; Li, J.; Xiong, S.; Zhao, Y.; Guo, M.; Liu, Z.; He, J.; et al. Diagnostic Accuracy of Droplet Digital PCR and Amplification Refractory Mutation System PCR for Detecting EGFR Mutation in Cell-Free DNA of Lung Cancer: A Meta-Analysis. *Front. Oncol.* **2020**, *10*, 290. [CrossRef] [PubMed]

82. Yang, K.; Li, J.; Zhao, J.; Ren, P.; Wang, Z.; Wei, B.; Dong, B.; Sun, R.; Wang, X.; Groen, H.J.M.; et al. Developing Ultrasensitive Library-Aliquot-Based Droplet Digital PCR for Detecting T790M in Plasma-Circulating Tumor DNA of Non-Small-Cell-Lung-Cancer Patients. *Anal. Chem.* **2018**, *90*, 11203–11209. [CrossRef]

83. Dressman, D.; Yan, H.; Traverso, G.; Kinzler, K.W.; Vogelstein, B. Transforming Single DNA Molecules into Fluorescent Magnetic Particles for Detection and Enumeration of Genetic Variations. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 8817–8822. [CrossRef]

84. Hindson, B.J.; Ness, K.D.; Masquelier, D.A.; Belgrader, P.; Heredia, N.J.; Makarewicz, A.J.; Bright, I.J.; Lucero, M.Y.; Hiddessen, A.L.; Legler, T.C.; et al. High-Throughput Droplet Digital PCR System for Absolute Quantitation of DNA Copy Number. *Anal. Chem.* **2011**, *83*, 8604–8610. [CrossRef]

85. Taly, V.; Pekin, D.; Benhaim, L.; Kotsopoulos, S.K.; Le Corre, D.; Li, X.; Atochin, I.; Link, D.R.; Griffiths, A.D.; Pallier, K.; et al. Multiplex Picodroplet Digital PCR to Detect KRAS Mutations in Circulating DNA from the Plasma of Colorectal Cancer Patients. *Clin. Chem.* **2013**, *59*, 1722–1731. [CrossRef]

86. Mouliere, F.; El Messaoudi, S.; Gongora, C.; Guedj, A.-S.; Robert, B.; Del Rio, M.; Molina, F.; Lamy, P.-J.; Lopez-Crapez, E.; Mathonnet, M.; et al. Circulating Cell-Free DNA from Colorectal Cancer Patients May Reveal High KRAS or BRAF Mutation Load. *Transl. Oncol.* **2013**, *6*, 319–328. [CrossRef]

87. Mouliere, F.; El Messaoudi, S.; Pang, D.; Dritschilo, A.; Thierry, A.R. Multi-Marker Analysis of Circulating Cell-Free DNA toward Personalized Medicine for Colorectal Cancer. *Mol. Oncol.* **2014**, *8*, 927–941. [CrossRef]

88. Bardelli, A.; Corso, S.; Bertotti, A.; Hobor, S.; Valtorta, E.; Siravegna, G.; Sartore-Bianchi, A.; Scala, E.; Cassingena, A.; Zecchin, D.; et al. Amplification of the MET Receptor Drives Resistance to Anti-EGFR Therapies in Colorectal Cancer. *Cancer Discov.* **2013**, *3*, 658–673. [CrossRef] [PubMed]

89. Wen, X.; Pu, H.; Liu, Q.; Guo, Z.; Luo, D. Circulating Tumor DNA-A Novel Biomarker of Tumor Progression and Its Favorable Detection Techniques. *Cancers* **2022**, *14*, 6025. [CrossRef]

90. Wang, W.; Song, Z.; Zhang, Y. A Comparison of DdPCR and ARMS for Detecting EGFR T790M Status in CtDNA from Advanced NSCLC Patients with Acquired EGFR-TKI Resistance. *Cancer Med.* **2017**, *6*, 154–162. [CrossRef] [PubMed]

91. FDA Approves Blood Tests That Can Help Guide Cancer Treatment—NCI. Available online: https://www.cancer.gov/news-events/cancer-currents-blog/2020/fda-guardant-360-foundation-one-cancer-liquid-biopsy (accessed on 1 March 2023).

92. Wan, J.C.M.; Massie, C.; Garcia-Corbacho, J.; Mouliere, F.; Brenton, J.D.; Caldas, C.; Pacey, S.; Baird, R.; Rosenfeld, N. Liquid Biopsies Come of Age: Towards Implementation of Circulating Tumour DNA. *Nat. Rev. Cancer* **2017**, *17*, 223–238. [CrossRef]

93. Belic, J.; Koch, M.; Ulz, P.; Auer, M.; Gerhalter, T.; Mohan, S.; Fischereder, K.; Petru, E.; Bauernhofer, T.; Geigl, J.B.; et al. Rapid Identification of Plasma DNA Samples with Increased CtDNA Levels by a Modified FAST-SeqS Approach. *Clin. Chem.* **2015**, *61*, 838–849. [CrossRef]

94. Castro-Giner, F.; Gkountela, S.; Donato, C.; Alborelli, I.; Quagliata, L.; Ng, C.K.Y.; Piscuoglio, S.; Aceto, N. Cancer Diagnosis Using a Liquid Biopsy: Challenges and Expectations. *Diagnostics* **2018**, *8*, 31. [CrossRef] [PubMed]

95. McCoach, C.E.; Blakely, C.M.; Banks, K.C.; Levy, B.; Chue, B.M.; Raymond, V.M.; Le, A.T.; Lee, C.E.; Diaz, J.; Waqar, S.N.; et al. Clinical Utility of Cell-Free DNA for the Detection of ALK Fusions and Genomic Mechanisms of ALK Inhibitor Resistance in Non-Small Cell Lung Cancer. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **2018**, *24*, 2758–2770. [CrossRef]

96. Baer, C.; Kern, W.; Koch, S.; Nadarajah, N.; Schindela, S.; Meggendorfer, M.; Haferlach, C.; Haferlach, T. Ultra-Deep Sequencing Leads to Earlier and More Sensitive Detection of the Tyrosine Kinase Inhibitor Resistance Mutation T315I in Chronic Myeloid Leukemia. *Haematologica* **2016**, *101*, 830–838. [CrossRef]

97. Goldberg, S.B.; Narayan, A.; Kole, A.J.; Decker, R.H.; Teysir, J.; Carriero, N.J.; Lee, A.; Nemati, R.; Nath, S.K.; Mane, S.M.; et al. Early Assessment of Lung Cancer Immunotherapy Response via Circulating Tumor DNA. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **2018**, *24*, 1872–1880. [CrossRef] [PubMed]

98. Phallen, J.; Sausen, M.; Adleff, V.; Leal, A.; Hruban, C.; White, J.; Anagnostou, V.; Fiksel, J.; Cristiano, S.; Papp, E.; et al. Direct Detection of Early-Stage Cancers Using Circulating Tumor DNA. *Sci. Transl. Med.* **2017**, *9*, eaan2415. [CrossRef]

99. Imperial, R.; Nazer, M.; Ahmed, Z.; Kam, A.E.; Pluard, T.J.; Bahaj, W.; Levy, M.; Kuzel, T.M.; Hayden, D.M.; Pappas, S.G.; et al. Matched Whole-Genome Sequencing (WGS) and Whole-Exome Sequencing (WES) of Tumor Tissue with Circulating Tumor DNA (CtDNA) Analysis: Complementary Modalities in Clinical Practice. *Cancers* **2019**, *11*, 1399. [CrossRef] [PubMed]

100. Lanman, R.B.; Mortimer, S.A.; Zill, O.A.; Sebisanovic, D.; Lopez, R.; Blau, S.; Collisson, E.A.; Divers, S.G.; Hoon, D.S.B.; Kopetz, E.S.; et al. Analytical and Clinical Validation of a Digital Sequencing Panel for Quantitative, Highly Accurate Evaluation of Cell-Free Circulating Tumor DNA. *PLoS ONE* **2015**, *10*, e0140712. [CrossRef] [PubMed]

101. Solutions—Guardant360. Available online: https://www.guardanthealthamea.com/solutions/ (accessed on 1 March 2023).

102. Milbury, C.A.; Creeden, J.; Yip, W.-K.; Smith, D.L.; Pattani, V.; Maxwell, K.; Sawchyn, B.; Gjoerup, O.; Meng, W.; Skoletsky, J.; et al. Clinical and Analytical Validation of FoundationOne®CDx, a Comprehensive Genomic Profiling Assay for Solid Tumors. *PLoS ONE* **2022**, *17*, e0264138. [CrossRef] [PubMed]

103. FoundationOne CDx | Foundation Medicine. Available online: https://www.foundationmedicine.com/test/foundationone-cdx (accessed on 1 March 2023).

104. Torres, S.; González, Á.; Cunquero Tomas, A.J.; Calabuig Fariñas, S.; Ferrero, M.; Mirda, D.; Sirera, R.; Jantus-Lewintre, E.; Camps, C. A Profile on Cobas® EGFR Mutation Test v2 as Companion Diagnostic for First-Line Treatment of Patients with Non-Small Cell Lung Cancer. *Expert Rev. Mol. Diagn.* **2020**, *20*, 575–582. [CrossRef]

105. Remon, J.; Hendriks, L.E.L.; Cardona, A.F.; Besse, B. EGFR Exon 20 Insertions in Advanced Non-Small Cell Lung Cancer: A New History Begins. *Cancer Treat. Rev.* **2020**, *90*, 102105. [CrossRef]

106. FDA Grants Accelerated Approval to Amivantamab-Vmjw for Metastatic Non-Small Cell Lung Cancer | FDA. Available online: https://www.fda.gov/drugs/resources-information-approved-drugs/fda-grants-accelerated-approval-amivantamab-vmjw-metastatic-non-small-cell-lung-cancer (accessed on 6 April 2023).

107. Koeppel, F.; Blanchard, S.; Jovelet, C.; Genin, B.; Marcaillou, C.; Martin, E.; Rouleau, E.; Solary, E.; Soria, J.-C.; André, F.; et al. Whole Exome Sequencing for Determination of Tumor Mutation Load in Liquid Biopsy from Advanced Cancer Patients. *PLoS ONE* **2017**, *12*, e0188174. [CrossRef]

108. Murtaza, M.; Dawson, S.-J.; Tsui, D.W.Y.; Gale, D.; Forshew, T.; Piskorz, A.M.; Parkinson, C.; Chin, S.-F.; Kingsbury, Z.; Wong, A.S.C.; et al. Non-Invasive Analysis of Acquired Resistance to Cancer Therapy by Sequencing of Plasma DNA. *Nature* **2013**, *497*, 108–112. [CrossRef] [PubMed]

109. Bos, M.K.; Angus, L.; Nasserinejad, K.; Jager, A.; Jansen, M.P.H.M.; Martens, J.W.M.; Sleijfer, S. Whole Exome Sequencing of Cell-Free DNA—A Systematic Review and Bayesian Individual Patient Data Meta-Analysis. *Cancer Treat. Rev.* **2020**, *83*, 101951. [CrossRef]

110. Wang, T.-L.; Maierhofer, C.; Speicher, M.R.; Lengauer, C.; Vogelstein, B.; Kinzler, K.W.; Velculescu, V.E. Digital Karyotyping. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 16156–16161. [CrossRef] [PubMed]

111. Leary, R.J.; Sausen, M.; Kinde, I.; Papadopoulos, N.; Carpten, J.D.; Craig, D.; O'Shaughnessy, J.; Kinzler, K.W.; Parmigiani, G.; Vogelstein, B.; et al. Detection of Chromosomal Alterations in the Circulation of Cancer Patients with Whole-Genome Sequencing. *Sci. Transl. Med.* **2012**, *4*, 162ra154. [CrossRef]

112. Diaz, L.A.; Sausen, M.; Fisher, G.A.; Velculescu, V.E. Insights into Therapeutic Resistance from Whole-Genome Analyses of Circulating Tumor DNA. *Oncotarget* **2013**, *4*, 1856–1857. [CrossRef]

113. Kinde, I.; Papadopoulos, N.; Kinzler, K.W.; Vogelstein, B. FAST-SeqS: A Simple and Efficient Method for the Detection of Aneuploidy by Massively Parallel Sequencing. *PLoS ONE* **2012**, *7*, e41162. [CrossRef] [PubMed]

114. Mendelaar, P.A.J.; Robbrecht, D.G.J.; Rijnders, M.; de Wit, R.; de Weerd, V.; Deger, T.; Westgeest, H.M.; Aarts, M.J.B.; Voortman, J.; Martens, J.W.M.; et al. Genome-Wide Aneuploidy Detected by MFast-SeqS in Circulating Cell-Free DNA Is Associated with Poor Response to Pembrolizumab in Patients with Advanced Urothelial Cancer. *Mol. Oncol.* **2022**, *16*, 2086–2097. [CrossRef] [PubMed]

115. Belic, J.; Koch, M.; Ulz, P.; Auer, M.; Gerhalter, T.; Mohan, S.; Fischereder, K.; Petru, E.; Bauernhofer, T.; Geigl, J.B.; et al. MFast-SeqS as a Monitoring and Pre-Screening Tool for Tumor-Specific Aneuploidy in Plasma DNA. *Adv. Exp. Med. Biol.* **2016**, *924*, 147–155. [CrossRef]

116. Leary, R.J.; Kinde, I.; Diehl, F.; Schmidt, K.; Clouser, C.; Duncan, C.; Antipova, A.; Lee, C.; McKernan, K.; De La Vega, F.M.; et al. Development of Personalized Tumor Biomarkers Using Massively Parallel Sequencing. *Sci. Transl. Med.* **2010**, *2*, 20ra14. [CrossRef]

117. McBride, D.J.; Orpana, A.K.; Sotiriou, C.; Joensuu, H.; Stephens, P.J.; Mudie, L.J.; Hämäläinen, E.; Stebbings, L.A.; Andersson, L.C.; Flanagan, A.M.; et al. Use of Cancer-Specific Genomic Rearrangements to Quantify Disease Burden in Plasma from Patients with Solid Tumors. *Genes. Chromosomes Cancer* **2010**, *49*, 1062–1069. [CrossRef] [PubMed]

118. Kinde, I.; Wu, J.; Papadopoulos, N.; Kinzler, K.W.; Vogelstein, B. Detection and Quantification of Rare Mutations with Massively Parallel Sequencing. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 9530–9535. [CrossRef]

119. Newman, A.M.; Bratman, S.V.; To, J.; Wynne, J.F.; Eclov, N.C.W.; Modlin, L.A.; Liu, C.L.; Neal, J.W.; Wakelee, H.A.; Merritt, R.E.; et al. An Ultrasensitive Method for Quantitating Circulating Tumor DNA with Broad Patient Coverage. *Nat. Med.* **2014**, *20*, 548–554. [CrossRef] [PubMed]

120. Rothé, F.; Laes, J.-F.; Lambrechts, D.; Smeets, D.; Vincent, D.; Maetens, M.; Fumagalli, D.; Michiels, S.; Drisis, S.; Moerman, C.; et al. Plasma Circulating Tumor DNA as an Alternative to Metastatic Biopsies for Mutational Analysis in Breast Cancer. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* **2014**, *25*, 1959–1965. [CrossRef] [PubMed]

121. Newman, A.M.; Lovejoy, A.F.; Klass, D.M.; Kurtz, D.M.; Chabon, J.J.; Scherer, F.; Stehr, H.; Liu, C.L.; Bratman, S.V.; Say, C.; et al. Integrated Digital Error Suppression for Improved Detection of Circulating Tumor DNA. *Nat. Biotechnol.* **2016**, *34*, 547–555. [CrossRef] [PubMed]

122. Logsdon, G.A.; Vollger, M.R.; Eichler, E.E. Long-Read Human Genome Sequencing and Its Applications. *Nat. Rev. Genet.* **2020**, *21*, 597–614. [CrossRef]

123. Pollard, M.O.; Gurdasani, D.; Mentzer, A.J.; Porter, T.; Sandhu, M.S. Long Reads: Their Purpose and Place. *Hum. Mol. Genet.* **2018**, *27*, R234–R241. [CrossRef]

124. Jain, M.; Olsen, H.E.; Paten, B.; Akeson, M. The Oxford Nanopore MinION: Delivery of Nanopore Sequencing to the Genomics Community. *Genome Biol.* **2016**, *17*, 239. [CrossRef]

125. Wang, Y.; Zhao, Y.; Bollas, A.; Wang, Y.; Au, K.F. Nanopore Sequencing Technology, Bioinformatics and Applications. *Nat. Biotechnol.* **2021**, *39*, 1348–1365. [CrossRef]

126. Eid, J.; Fehr, A.; Gray, J.; Luong, K.; Lyle, J.; Otto, G.; Peluso, P.; Rank, D.; Baybayan, P.; Bettman, B.; et al. Real-Time DNA Sequencing from Single Polymerase Molecules. *Science* **2009**, *323*, 133–138. [CrossRef]

127. Rhoads, A.; Au, K.F. PacBio Sequencing and Its Applications. *Genom. Proteom. Bioinform.* **2015**, *13*, 278–289. [CrossRef]

128. Lo, Y.M.D.; Han, D.S.C.; Jiang, P.; Chiu, R.W.K. Epigenetics, Fragmentomics, and Topology of Cell-Free DNA in Liquid Biopsies. *Science* **2021**, *372*, eaaw3616. [CrossRef] [PubMed]

129. Yu, S.C.Y.; Jiang, P.; Peng, W.; Cheng, S.H.; Cheung, Y.T.T.; Tse, O.Y.O.; Shang, H.; Poon, L.C.; Leung, T.Y.; Chan, K.C.A.; et al. Single-Molecule Sequencing Reveals a Large Population of Long Cell-Free DNA Molecules in Maternal Plasma. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2114937118. [CrossRef]

130. Choy, L.Y.L.; Peng, W.; Jiang, P.; Cheng, S.H.; Yu, S.C.Y.; Shang, H.; Olivia Tse, O.Y.; Wong, J.; Wong, V.W.S.; Wong, G.L.H.; et al. Single-Molecule Sequencing Enables Long Cell-Free DNA Detection and Direct Methylation Analysis for Cancer Patients. *Clin. Chem.* **2022**, *68*, 1151–1163. [CrossRef] [PubMed]

131. Yu, S.C.Y.; Deng, J.; Qiao, R.; Cheng, S.H.; Peng, W.; Lau, S.L.; Choy, L.Y.L.; Leung, T.Y.; Wong, J.; Wong, V.W.-S.; et al. Comparison of Single Molecule, Real-Time Sequencing and Nanopore Sequencing for Analysis of the Size, End-Motif, and Tissue-of-Origin of Long Cell-Free DNA in Plasma. *Clin. Chem.* **2023**, *69*, 168–179. [CrossRef]

132. Peters, B.A.; Kermani, B.G.; Sparks, A.B.; Alferov, O.; Hong, P.; Alexeev, A.; Jiang, Y.; Dahl, F.; Tang, Y.T.; Haas, J.; et al. Accurate Whole-Genome Sequencing and Haplotyping from 10 to 20 Human Cells. *Nature* **2012**, *487*, 190–195. [CrossRef] [PubMed]

133. Zheng, G.X.Y.; Lau, B.T.; Schnall-Levin, M.; Jarosz, M.; Bell, J.M.; Hindson, C.M.; Kyriazopoulou-Panagiotopoulou, S.; Masquelier, D.A.; Merrill, L.; Terry, J.M.; et al. Haplotyping Germline and Cancer Genomes with High-Throughput Linked-Read Sequencing. *Nat. Biotechnol.* **2016**, *34*, 303–311. [CrossRef]

134. Zhao, L.; Wu, X.; Zheng, J.; Dong, D. DNA Methylome Profiling of Circulating Tumor Cells in Lung Cancer at Single Base-Pair Resolution. *Oncogene* **2021**, *40*, 1884–1895. [CrossRef]

135. Gonçalves, E.; Gonçalves-Reis, M.; Pereira-Leal, J.B.; Cardoso, J. DNA Methylation Fingerprint of Hepatocellular Carcinoma from Tissue and Liquid Biopsies. *Sci. Rep.* **2022**, *12*, 11512. [CrossRef]

136. Luo, H.; Wei, W.; Ye, Z.; Zheng, J.; Xu, R.-H. Liquid Biopsy of Methylation Biomarkers in Cell-Free DNA. *Trends Mol. Med.* **2021**, *27*, 482–500. [CrossRef]

137. Zhang, Y.; Bewerunge-Hudler, M.; Schick, M.; Burwinkel, B.; Herpel, E.; Hoffmeister, M.; Brenner, H. Blood-Derived DNA Methylation Predictors of Mortality Discriminate Tumor and Healthy Tissue in Multiple Organs. *Mol. Oncol.* **2020**, *14*, 2111–2123. [CrossRef]

138. Sproul, D.; Meehan, R.R. Genomic Insights into Cancer-Associated Aberrant CpG Island Hypermethylation. *Brief. Funct. Genom.* **2013**, *12*, 174–190. [CrossRef]

139. Nishiyama, A.; Nakanishi, M. Navigating the DNA Methylation Landscape of Cancer. *Trends Genet. TIG* **2021**, *37*, 1012–1027. [CrossRef] [PubMed]

140. Balgkouranidou, I.; Chimonidou, M.; Milaki, G.; Tsaroucha, E.; Kakolyris, S.; Georgoulias, V.; Lianidou, E. SOX17 Promoter Methylation in Plasma Circulating Tumor DNA of Patients with Non-Small Cell Lung Cancer. *Clin. Chem. Lab. Med.* **2016**, *54*, 1385–1393. [CrossRef] [PubMed]

141. Liu, B.; Ricarte Filho, J.; Mallisetty, A.; Villani, C.; Kottorou, A.; Rodgers, K.; Chen, C.; Ito, T.; Holmes, K.; Gastala, N.; et al. Detection of Promoter DNA Methylation in Urine and Plasma Aids the Detection of Non-Small Cell Lung Cancer. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **2020**, *26*, 4339–4348. [CrossRef] [PubMed]

142. Jamshidi, A.; Liu, M.C.; Klein, E.A.; Venn, O.; Hubbell, E.; Beausang, J.F.; Gross, S.; Melton, C.; Fields, A.P.; Liu, Q.; et al. Evaluation of Cell-Free DNA Approaches for Multi-Cancer Early Detection. *Cancer Cell* **2022**, *40*, 1537–1549.e12. [CrossRef]

143. Jamieson, E.; Korologou-Linden, R.; Wootton, R.E.; Guyatt, A.L.; Battram, T.; Burrows, K.; Gaunt, T.R.; Tobin, M.D.; Munafò, M.; Davey Smith, G.; et al. Smoking, DNA Methylation, and Lung Function: A Mendelian Randomization Analysis to Investigate Causal Pathways. *Am. J. Hum. Genet.* **2020**, *106*, 315–326. [CrossRef]

144. Lissa, D.; Robles, A.I. Sputum-Based DNA Methylation Biomarkers to Guide Lung Cancer Screening Decisions. *J. Thorac. Dis.* **2017**, *9*, 4308–4310. [CrossRef]

145. Li, P.; Liu, S.; Du, L.; Mohseni, G.; Zhang, Y.; Wang, C. Liquid Biopsies Based on DNA Methylation as Biomarkers for the Detection and Prognosis of Lung Cancer. *Clin. Epigenet.* **2022**, *14*, 118. [CrossRef]

146. Loyfer, N.; Magenheim, J.; Peretz, A.; Cann, G.; Bredno, J.; Klochendler, A.; Fox-Fisher, I.; Shabi-Porat, S.; Hecht, M.; Pelet, T.; et al. A DNA Methylation Atlas of Normal Human Cell Types. *Nature* **2023**, *613*, 355–364. [CrossRef]
147. Chen, Q.-F.; Gao, H.; Pan, Q.-Y.; Wang, Y.-J.; Zhong, X.-N. Analysis at the Single-Cell Level Indicates an Important Role of Heterogeneous Global DNA Methylation Status on the Progression of Lung Adenocarcinoma. *Sci. Rep.* **2021**, *11*, 23337. [CrossRef]
148. Yong, W.-S.; Hsu, F.-M.; Chen, P.-Y. Profiling Genome-Wide DNA Methylation. *Epigenet. Chromatin* **2016**, *9*, 26. [CrossRef]
149. Taiwo, O.; Wilson, G.A.; Morris, T.; Seisenberger, S.; Reik, W.; Pearce, D.; Beck, S.; Butcher, L.M. Methylome Analysis Using MeDIP-Seq with Low DNA Concentrations. *Nat. Protoc.* **2012**, *7*, 617–636. [CrossRef] [PubMed]
150. Xu, J.; Liao, K.; Yang, X.; Wu, C.; Wu, W. Using Single-Cell Sequencing Technology to Detect Circulating Tumor Cells in Solid Tumors. *Mol. Cancer* **2021**, *20*, 104. [CrossRef]
151. Lim, S.B.; Di Lee, W.; Vasudevan, J.; Lim, W.-T.; Lim, C.T. Liquid Biopsy: One Cell at a Time. *NPJ Precis. Oncol.* **2019**, *3*, 23. [CrossRef] [PubMed]
152. Dong, Y.; Wang, Z.; Shi, Q. Liquid Biopsy Based Single-Cell Transcriptome Profiling Characterizes Heterogeneity of Disseminated Tumor Cells from Lung Adenocarcinoma. *Proteomics* **2020**, *20*, e1900224. [CrossRef]
153. Guo, X.; Zhang, Y.; Zheng, L.; Zheng, C.; Song, J.; Zhang, Q.; Kang, B.; Liu, Z.; Jin, L.; Xing, R.; et al. Global Characterization of T Cells in Non-Small-Cell Lung Cancer by Single-Cell Sequencing. *Nat. Med.* **2018**, *24*, 978–985. [CrossRef] [PubMed]
154. Ballesteros, I.; Cerezo-Wallis, D.; Hidalgo, A. Understanding NSCLC, One Cell at a Time. *Cancer Cell* **2022**, *40*, 1459–1461. [CrossRef]
155. Keller, L.; Pantel, K. Unravelling Tumour Heterogeneity by Single-Cell Profiling of Circulating Tumour Cells. *Nat. Rev. Cancer* **2019**, *19*, 553–567. [CrossRef]
156. Pei, H.; Li, L.; Han, Z.; Wang, Y.; Tang, B. Recent Advances in Microfluidic Technologies for Circulating Tumor Cells: Enrichment, Single-Cell Analysis, and Liquid Biopsy for Clinical Applications. *Lab Chip* **2020**, *20*, 3854–3875. [CrossRef]
157. Zong, C.; Lu, S.; Chapman, A.R.; Xie, X.S. Genome-Wide Detection of Single-Nucleotide and Copy-Number Variations of a Single Human Cell. *Science* **2012**, *338*, 1622–1626. [CrossRef]
158. Ni, X.; Zhuo, M.; Su, Z.; Duan, J.; Gao, Y.; Wang, Z.; Zong, C.; Bai, H.; Chapman, A.R.; Zhao, J.; et al. Reproducible Copy Number Variation Patterns among Single Circulating Tumor Cells of Lung Cancer Patients. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 21083–21088. [CrossRef]
159. Kim, O.; Lee, D.; Chungwon Lee, A.; Lee, Y.; Bae, H.J.; Lee, H.-B.; Kim, R.N.; Han, W.; Kwon, S. Whole Genome Sequencing of Single Circulating Tumor Cells Isolated by Applying a Pulsed Laser to Cell-Capturing Microstructures. *Small Weinh. Bergstr. Ger.* **2019**, *15*, e1902607. [CrossRef] [PubMed]
160. Fu, Y.; Li, C.; Lu, S.; Zhou, W.; Tang, F.; Xie, X.S.; Huang, Y. Uniform and Accurate Single-Cell Sequencing Based on Emulsion Whole-Genome Amplification. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 11923–11928. [CrossRef]
161. Kojima, M.; Harada, T.; Fukazawa, T.; Kurihara, S.; Saeki, I.; Takahashi, S.; Hiyama, E. Single-Cell DNA and RNA Sequencing of Circulating Tumor Cells. *Sci. Rep.* **2021**, *11*, 22864. [CrossRef]
162. Luecken, M.D.; Theis, F.J. Current Best Practices in Single-Cell RNA-Seq Analysis: A Tutorial. *Mol. Syst. Biol.* **2019**, *15*, e8746. [CrossRef]
163. Park, S.-M.; Wong, D.J.; Ooi, C.C.; Kurtz, D.M.; Vermesh, O.; Aalipour, A.; Suh, S.; Pian, K.L.; Chabon, J.J.; Lee, S.H.; et al. Molecular Profiling of Single Circulating Tumor Cells from Lung Cancer Patients. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E8379–E8386. [CrossRef]
164. Lim, S.B.; Yeo, T.; Lee, W.D.; Bhagat, A.A.S.; Tan, S.J.; Tan, D.S.W.; Lim, W.-T.; Lim, C.T. Addressing Cellular Heterogeneity in Tumor and Circulation for Refined Prognostication. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 17957–17962. [CrossRef]
165. Zhitnyuk, Y.V.; Koval, A.P.; Alferov, A.A.; Shtykova, Y.A.; Mamedov, I.Z.; Kushlinskii, N.E.; Chudakov, D.M.; Shcherbo, D.S. Deep CfDNA Fragment End Profiling Enables Cancer Detection. *Mol. Cancer* **2022**, *21*, 26. [CrossRef]
166. Gianni, C.; Palleschi, M.; Merloni, F.; Di Menna, G.; Sirico, M.; Sarti, S.; Virga, A.; Ulivi, P.; Cecconetto, L.; Mariotti, M.; et al. Cell-Free DNA Fragmentomics: A Promising Biomarker for Diagnosis, Prognosis and Prediction of Response in Breast Cancer. *Int. J. Mol. Sci.* **2022**, *23*, 14197. [CrossRef] [PubMed]
167. Thierry, A.R. Circulating DNA Fragmentomics and Cancer Screening. *Cell Genom.* **2023**, *3*, 100242. [CrossRef] [PubMed]
168. Zhang, X.; Wang, Z.; Tang, W.; Wang, X.; Liu, R.; Bao, H.; Chen, X.; Wei, Y.; Wu, S.; Bao, H.; et al. Ultrasensitive and Affordable Assay for Early Detection of Primary Liver Cancer Using Plasma Cell-Free DNA Fragmentomics. *Hepatol. Baltim. Md* **2022**, *76*, 317–329. [CrossRef] [PubMed]
169. Ding, S.C.; Lo, Y.M.D. Cell-Free DNA Fragmentomics in Liquid Biopsy. *Diagnostics* **2022**, *12*, 978. [CrossRef]
170. Ivanov, M.; Baranova, A.; Butler, T.; Spellman, P.; Mileyko, V. Non-Random Fragmentation Patterns in Circulating Cell-Free DNA Reflect Epigenetic Regulation. *BMC Genom.* **2015**, *16* (Suppl. S13), S1. [CrossRef]
171. Liu, Y. At the Dawn: Cell-Free DNA Fragmentomics and Gene Regulation. *Br. J. Cancer* **2022**, *126*, 379–390. [CrossRef]
172. Stroun, M.; Anker, P.; Lyautey, J.; Lederrey, C.; Maurice, P.A. Isolation and Characterization of DNA from the Plasma of Cancer Patients. *Eur. J. Cancer Clin. Oncol.* **1987**, *23*, 707–712. [CrossRef]
173. Cristiano, S.; Leal, A.; Phallen, J.; Fiksel, J.; Adleff, V.; Bruhm, D.C.; Jensen, S.Ø.; Medina, J.E.; Hruban, C.; White, J.R.; et al. Genome-Wide Cell-Free DNA Fragmentation in Patients with Cancer. *Nature* **2019**, *570*, 385–389. [CrossRef]
174. Underhill, H.R.; Kitzman, J.O.; Hellwig, S.; Welker, N.C.; Daza, R.; Baker, D.N.; Gligorich, K.M.; Rostomily, R.C.; Bronner, M.P.; Shendure, J. Fragment Length of Circulating Tumor DNA. *PLoS Genet.* **2016**, *12*, e1006162. [CrossRef] [PubMed]

175. Mouliere, F.; Chandrananda, D.; Piskorz, A.M.; Moore, E.K.; Morris, J.; Ahlborn, L.B.; Mair, R.; Goranova, T.; Marass, F.; Heider, K.; et al. Enhanced Detection of Circulating Tumor DNA by Fragment Size Analysis. *Sci. Transl. Med.* **2018**, *10*, eaat4921. [CrossRef] [PubMed]

176. Peng, Y.; Mei, W.; Ma, K.; Zeng, C. Circulating Tumor DNA and Minimal Residual Disease (MRD) in Solid Tumors: Current Horizons and Future Perspectives. *Front. Oncol.* **2021**, *11*, 763790. [CrossRef]

177. Vessies, D.C.L.; Schuurbiers, M.M.F.; van der Noort, V.; Schouten, I.; Linders, T.C.; Lanfermeijer, M.; Ramkisoensing, K.L.; Hartemink, K.J.; Monkhorst, K.; van den Heuvel, M.M.; et al. Combining Variant Detection and Fragment Length Analysis Improves Detection of Minimal Residual Disease in Postsurgery Circulating Tumour DNA of Stage II-IIIA NSCLC Patients. *Mol. Oncol.* **2022**, *16*, 2719–2732. [CrossRef] [PubMed]

178. Bao, H.; Wang, Z.; Ma, X.; Guo, W.; Zhang, X.; Tang, W.; Chen, X.; Wang, X.; Chen, Y.; Mo, S.; et al. Letter to the Editor: An Ultra-Sensitive Assay Using Cell-Free DNA Fragmentomics for Multi-Cancer Early Detection. *Mol. Cancer* **2022**, *21*, 129. [CrossRef] [PubMed]

179. Wang, S.; Meng, F.; Li, M.; Bao, H.; Chen, X.; Zhu, M.; Liu, R.; Xu, X.; Yang, S.; Wu, X.; et al. Multi-Dimensional Cell-Free DNA Fragmentomic Assay for Detection of Early-Stage Lung Cancer. *Am. J. Respir. Crit. Care Med.* **2022**. [CrossRef]

180. Guo, W.; Chen, X.; Liu, R.; Liang, N.; Ma, Q.; Bao, H.; Xu, X.; Wu, X.; Yang, S.; Shao, Y.; et al. Sensitive Detection of Stage I Lung Adenocarcinoma Using Plasma Cell-Free DNA Breakpoint Motif Profiling. *EBioMedicine* **2022**, *81*, 104131. [CrossRef]

181. Moorthie, S.; Hall, A.; Wright, C.F. Informatics and Clinical Genome Sequencing: Opening the Black Box. *Genet. Med. Off. J. Am. Coll. Med. Genet.* **2013**, *15*, 165–171. [CrossRef]

182. Pereira, R.; Oliveira, J.; Sousa, M. Bioinformatics and Computational Tools for Next-Generation Sequencing Analysis in Clinical Genetics. *J. Clin. Med.* **2020**, *9*, 132. [CrossRef]

183. BCL Convert Support. Available online: https://support.illumina.com/sequencing/sequencing_software/bcl-convert.html (accessed on 1 March 2023).

184. Cock, P.J.A.; Fields, C.J.; Goto, N.; Heuer, M.L.; Rice, P.M. The Sanger FASTQ File Format for Sequences with Quality Scores, and the Solexa/Illumina FASTQ Variants. *Nucleic Acids Res.* **2010**, *38*, 1767–1771. [CrossRef]

185. de Sena Brandine, G.; Smith, A.D. Falco: High-Speed FastQC Emulation for Quality Control of Sequencing Data. *F1000Research* **2019**, *8*, 1874. [CrossRef] [PubMed]

186. Patel, R.K.; Jain, M. NGS QC Toolkit: A Toolkit for Quality Control of next Generation Sequencing Data. *PLoS ONE* **2012**, *7*, e30619. [CrossRef]

187. Babraham Bioinformatics—FastQC A Quality Control Tool for High Throughput Sequence Data. Available online: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (accessed on 1 March 2023).

188. Reinert, K.; Langmead, B.; Weese, D.; Evers, D.J. Alignment of Next-Generation Sequencing Reads. *Annu. Rev. Genom. Hum. Genet.* **2015**, *16*, 133–151. [CrossRef] [PubMed]

189. Trapnell, C.; Salzberg, S.L. How to Map Billions of Short Reads onto Genomes. *Nat. Biotechnol.* **2009**, *27*, 455–457. [CrossRef]

190. Li, H.; Durbin, R. Fast and Accurate Short Read Alignment with Burrows-Wheeler Transform. *Bioinformatics* **2009**, *25*, 1754–1760. [CrossRef] [PubMed]

191. Langmead, B.; Salzberg, S.L. Fast Gapped-Read Alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359. [CrossRef] [PubMed]

192. Kim, D.; Pertea, G.; Trapnell, C.; Pimentel, H.; Kelley, R.; Salzberg, S.L. TopHat2: Accurate Alignment of Transcriptomes in the Presence of Insertions, Deletions and Gene Fusions. *Genom. Biol.* **2013**, *14*, R36. [CrossRef]

193. Kim, D.; Langmead, B.; Salzberg, S.L. HISAT: A Fast Spliced Aligner with Low Memory Requirements. *Nat. Methods* **2015**, *12*, 357–360. [CrossRef] [PubMed]

194. Kim, D.; Paggi, J.M.; Park, C.; Bennett, C.; Salzberg, S.L. Graph-Based Genome Alignment and Genotyping with HISAT2 and HISAT-Genotype. *Nat. Biotechnol.* **2019**, *37*, 907–915. [CrossRef] [PubMed]

195. Zhang, Y.; Park, C.; Bennett, C.; Thornton, M.; Kim, D. Rapid and Accurate Alignment of Nucleotide Conversion Sequencing Reads with HISAT-3N. *Genom. Res.* **2021**, *31*, 1290–1295. [CrossRef] [PubMed]

196. Dobin, A.; Davis, C.A.; Schlesinger, F.; Drenkow, J.; Zaleski, C.; Jha, S.; Batut, P.; Chaisson, M.; Gingeras, T.R. STAR: Ultrafast Universal RNA-Seq Aligner. *Bioinformatics* **2013**, *29*, 15–21. [CrossRef]

197. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R. 1000 Genome Project Data Processing Subgroup The Sequence Alignment/Map Format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079. [CrossRef] [PubMed]

198. McKenna, A.; Hanna, M.; Banks, E.; Sivachenko, A.; Cibulskis, K.; Kernytsky, A.; Garimella, K.; Altshuler, D.; Gabriel, S.; Daly, M.; et al. The Genome Analysis Toolkit: A MapReduce Framework for Analyzing next-Generation DNA Sequencing Data. *Genom. Res.* **2010**, *20*, 1297–1303. [CrossRef]

199. Koboldt, D.C.; Chen, K.; Wylie, T.; Larson, D.E.; McLellan, M.D.; Mardis, E.R.; Weinstock, G.M.; Wilson, R.K.; Ding, L. VarScan: Variant Detection in Massively Parallel Sequencing of Individual and Pooled Samples. *Bioinformatics* **2009**, *25*, 2283–2285. [CrossRef] [PubMed]

200. Danecek, P.; Bonfield, J.K.; Liddle, J.; Marshall, J.; Ohan, V.; Pollard, M.O.; Whitwham, A.; Keane, T.; McCarthy, S.A.; Davies, R.M.; et al. Twelve Years of SAMtools and BCFtools. *GigaScience* **2021**, *10*, giab008. [CrossRef] [PubMed]

201. Sherry, S.T.; Ward, M.H.; Kholodov, M.; Baker, J.; Phan, L.; Smigielski, E.M.; Sirotkin, K. DbSNP: The NCBI Database of Genetic Variation. *Nucleic Acids Res.* **2001**, *29*, 308–311. [CrossRef]

202. Cunningham, F.; Allen, J.E.; Allen, J.; Alvarez-Jarreta, J.; Amode, M.R.; Armean, I.M.; Austine-Orimoloye, O.; Azov, A.G.; Barnes, I.; Bennett, R.; et al. Ensembl 2022. *Nucleic Acids Res.* **2022**, *50*, D988–D995. [CrossRef] [PubMed]

203. O'Leary, N.A.; Wright, M.W.; Brister, J.R.; Ciufo, S.; Haddad, D.; McVeigh, R.; Rajput, B.; Robbertse, B.; Smith-White, B.; Ako-Adjei, D.; et al. Reference Sequence (RefSeq) Database at NCBI: Current Status, Taxonomic Expansion, and Functional Annotation. *Nucleic Acids Res.* **2016**, *44*, D733–D745. [CrossRef]

204. Landrum, M.J.; Chitipiralla, S.; Brown, G.R.; Chen, C.; Gu, B.; Hart, J.; Hoffman, D.; Jang, W.; Kaur, K.; Liu, C.; et al. ClinVar: Improvements to Accessing Data. *Nucleic Acids Res.* **2020**, *48*, D835–D844. [CrossRef]

205. Tate, J.G.; Bamford, S.; Jubb, H.C.; Sondka, Z.; Beare, D.M.; Bindal, N.; Boutselakis, H.; Cole, C.G.; Creatore, C.; Dawson, E.; et al. COSMIC: The Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Res.* **2019**, *47*, D941–D947. [CrossRef]

206. Kent, W.J.; Sugnet, C.W.; Furey, T.S.; Roskin, K.M.; Pringle, T.H.; Zahler, A.M.; Haussler, D. The Human Genome Browser at UCSC. *Genom. Res.* **2002**, *12*, 996–1006. [CrossRef] [PubMed]

207. Wang, K.; Li, M.; Hakonarson, H. ANNOVAR: Functional Annotation of Genetic Variants from High-Throughput Sequencing Data. *Nucleic Acids Res.* **2010**, *38*, e164. [CrossRef]

208. Cingolani, P.; Platts, A.; Wang, L.L.; Coon, M.; Nguyen, T.; Wang, L.; Land, S.J.; Lu, X.; Ruden, D.M. A Program for Annotating and Predicting the Effects of Single Nucleotide Polymorphisms, SnpEff: SNPs in the Genome of Drosophila Melanogaster Strain W1118; Iso-2; Iso-3. *Fly* **2012**, *6*, 80–92. [CrossRef]

209. McLaren, W.; Gil, L.; Hunt, S.E.; Riat, H.S.; Ritchie, G.R.S.; Thormann, A.; Flicek, P.; Cunningham, F. The Ensembl Variant Effect Predictor. *Genom. Biol.* **2016**, *17*, 122. [CrossRef] [PubMed]

210. Subramanian, A.; Tamayo, P.; Mootha, V.K.; Mukherjee, S.; Ebert, B.L.; Gillette, M.A.; Paulovich, A.; Pomeroy, S.L.; Golub, T.R.; Lander, E.S.; et al. Gene Set Enrichment Analysis: A Knowledge-Based Approach for Interpreting Genome-Wide Expression Profiles. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 15545–15550. [CrossRef] [PubMed]

211. Kanehisa, M.; Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **2000**, *28*, 27–30. [CrossRef]

212. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genom. Res.* **2003**, *13*, 2498–2504. [CrossRef]

213. Huang, D.W.; Sherman, B.T.; Tan, Q.; Collins, J.R.; Alvord, W.G.; Roayaei, J.; Stephens, R.; Baseler, M.W.; Lane, H.C.; Lempicki, R.A. The DAVID Gene Functional Classification Tool: A Novel Biological Module-Centric Algorithm to Functionally Analyze Large Gene Lists. *Genom. Biol.* **2007**, *8*, R183. [CrossRef] [PubMed]

214. Chen, E.Y.; Tan, C.M.; Kou, Y.; Duan, Q.; Wang, Z.; Meirelles, G.V.; Clark, N.R.; Ma'ayan, A. Enrichr: Interactive and Collaborative HTML5 Gene List Enrichment Analysis Tool. *BMC Bioinform.* **2013**, *14*, 128. [CrossRef]

215. Franz, M.; Rodriguez, H.; Lopes, C.; Zuberi, K.; Montojo, J.; Bader, G.D.; Morris, Q. GeneMANIA Update 2018. *Nucleic Acids Res.* **2018**, *46*, W60–W64. [CrossRef] [PubMed]

216. Lai, Z.; Markovets, A.; Ahdesmaki, M.; Chapman, B.; Hofmann, O.; McEwen, R.; Johnson, J.; Dougherty, B.; Barrett, J.C.; Dry, J.R. VarDict: A Novel and Versatile Variant Caller for next-Generation Sequencing in Cancer Research. *Nucleic Acids Res.* **2016**, *44*, e108. [CrossRef]

217. Kim, S.; Scheffler, K.; Halpern, A.L.; Bekritsky, M.A.; Noh, E.; Källberg, M.; Chen, X.; Kim, Y.; Beyter, D.; Krusche, P.; et al. Strelka2: Fast and Accurate Calling of Germline and Somatic Variants. *Nat. Methods* **2018**, *15*, 591–594. [CrossRef] [PubMed]

218. Warde-Farley, D.; Donaldson, S.L.; Comes, O.; Zuberi, K.; Badrawi, R.; Chao, P.; Franz, M.; Grouios, C.; Kazi, F.; Lopes, C.T.; et al. The GeneMANIA Prediction Server: Biological Network Integration for Gene Prioritization and Predicting Gene Function. *Nucleic Acids Res.* **2010**, *38*, W214–W220. [CrossRef]

219. Amarasinghe, S.L.; Su, S.; Dong, X.; Zappia, L.; Ritchie, M.E.; Gouil, Q. Opportunities and Challenges in Long-Read Sequencing Data Analysis. *Genom. Biol.* **2020**, *21*, 30. [CrossRef]

220. Alser, M.; Rotman, J.; Deshpande, D.; Taraszka, K.; Shi, H.; Baykal, P.I.; Yang, H.T.; Xue, V.; Knyazev, S.; Singer, B.D.; et al. Technology Dictates Algorithms: Recent Developments in Read Alignment. *Genom. Biol.* **2021**, *22*, 249. [CrossRef] [PubMed]

221. CCS Home | CCS Docs. Available online: https://ccs.how/ (accessed on 5 March 2023).

222. Oxford Nanopore Technologies. Available online: https://github.com/nanoporetech (accessed on 1 March 2023).

223. Du, P.; Zhang, X.; Huang, C.-C.; Jafari, N.; Kibbe, W.A.; Hou, L.; Lin, S.M. Comparison of Beta-Value and M-Value Methods for Quantifying Methylation Levels by Microarray Analysis. *BMC Bioinform.* **2010**, *11*, 587. [CrossRef]

224. Huber, W.; Carey, V.J.; Gentleman, R.; Anders, S.; Carlson, M.; Carvalho, B.S.; Bravo, H.C.; Davis, S.; Gatto, L.; Girke, T.; et al. Orchestrating High-Throughput Genomic Analysis with Bioconductor. *Nat. Methods* **2015**, *12*, 115–121. [CrossRef]

225. Bioconductor—BiocViews. Available online: http://www.bioconductor.org/packages/release/BiocViews.html#___DNAMethylation (accessed on 1 March 2023).

226. Kurdyukov, S.; Bullock, M. DNA Methylation Analysis: Choosing the Right Method. *Biology* **2016**, *5*, 3. [CrossRef]

227. Harris, R.A.; Wang, T.; Coarfa, C.; Nagarajan, R.P.; Hong, C.; Downey, S.L.; Johnson, B.E.; Fouse, S.D.; Delaney, A.; Zhao, Y.; et al. Comparison of Sequencing-Based Methods to Profile DNA Methylation and Identification of Monoallelic Epigenetic Modifications. *Nat. Biotechnol.* **2010**, *28*, 1097–1105. [CrossRef]

228. Beck, D.; Ben Maamar, M.; Skinner, M.K. Genome-Wide CpG Density and DNA Methylation Analysis Method (MeDIP, RRBS, and WGBS) Comparisons. *Epigenetics* **2022**, *17*, 518–530. [CrossRef]

229. Konwar, C.; Del Gobbo, G.; Yuan, V.; Robinson, W.P. Considerations When Processing and Interpreting Genomics Data of the Placenta. *Placenta* **2019**, *84*, 57–62. [CrossRef] [PubMed]

230. Hao, Y.; Hao, S.; Andersen-Nissen, E.; Mauck, W.M.; Zheng, S.; Butler, A.; Lee, M.J.; Wilk, A.J.; Darby, C.; Zager, M.; et al. Integrated Analysis of Multimodal Single-Cell Data. *Cell* **2021**, *184*, 3573–3587.e29. [CrossRef] [PubMed]

231. Wolf, F.A.; Angerer, P.; Theis, F.J. SCANPY: Large-Scale Single-Cell Gene Expression Data Analysis. *Genom. Biol.* **2018**, *19*, 15. [CrossRef]

232. McCarthy, D.J.; Campbell, K.R.; Lun, A.T.L.; Wills, Q.F. Scater: Pre-Processing, Quality Control, Normalization and Visualization of Single-Cell RNA-Seq Data in R. *Bioinformatics* **2017**, *33*, 1179–1186. [CrossRef] [PubMed]

233. GitHub–Lmcinnes/Umap: Uniform Manifold Approximation and Projection. Available online: https://github.com/lmcinnes/umap (accessed on 1 March 2023).

234. Regev, A.; Teichmann, S.A.; Lander, E.S.; Amit, I.; Benoist, C.; Birney, E.; Bodenmiller, B.; Campbell, P.; Carninci, P.; Clatworthy, M.; et al. The Human Cell Atlas. *eLife* **2017**, *6*, e27041. [CrossRef]

235. Zhang, X.; Lan, Y.; Xu, J.; Quan, F.; Zhao, E.; Deng, C.; Luo, T.; Xu, L.; Liao, G.; Yan, M.; et al. CellMarker: A Manually Curated Resource of Cell Markers in Human and Mouse. *Nucleic Acids Res.* **2019**, *47*, D721–D728. [CrossRef]

236. Kiselev, V.Y.; Yiu, A.; Hemberg, M. Scmap: Projection of Single-Cell RNA-Seq Data across Data Sets. *Nat. Methods* **2018**, *15*, 359–362. [CrossRef] [PubMed]

237. Trapnell, C.; Cacchiarelli, D.; Grimsby, J.; Pokharel, P.; Li, S.; Morse, M.; Lennon, N.J.; Livak, K.J.; Mikkelsen, T.S.; Rinn, J.L. The Dynamics and Regulators of Cell Fate Decisions Are Revealed by Pseudotemporal Ordering of Single Cells. *Nat. Biotechnol.* **2014**, *32*, 381–386. [CrossRef]

238. Bendall, S.C.; Davis, K.L.; Amir, E.-A.D.; Tadmor, M.D.; Simonds, E.F.; Chen, T.J.; Shenfeld, D.K.; Nolan, G.P.; Pe'er, D. Single-Cell Trajectory Detection Uncovers Progression and Regulatory Coordination in Human B Cell Development. *Cell* **2014**, *157*, 714–725. [CrossRef] [PubMed]

239. Street, K.; Risso, D.; Fletcher, R.B.; Das, D.; Ngai, J.; Yosef, N.; Purdom, E.; Dudoit, S. Slingshot: Cell Lineage and Pseudotime Inference for Single-Cell Transcriptomics. *BMC Genom.* **2018**, *19*, 477. [CrossRef] [PubMed]

240. Love, M.I.; Huber, W.; Anders, S. Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2. *Genom. Biol.* **2014**, *15*, 550. [CrossRef]

241. Robinson, M.D.; McCarthy, D.J.; Smyth, G.K. EdgeR: A Bioconductor Package for Differential Expression Analysis of Digital Gene Expression Data. *Bioinform. Oxf. Engl.* **2010**, *26*, 139–140. [CrossRef] [PubMed]

242. Stuart, T.; Butler, A.; Hoffman, P.; Hafemeister, C.; Papalexi, E.; Mauck, W.M.; Hao, Y.; Stoeckius, M.; Smibert, P.; Satija, R. Comprehensive Integration of Single-Cell Data. *Cell* **2019**, *177*, 1888–1902.e21. [CrossRef]

243. Qiu, X.; Mao, Q.; Tang, Y.; Wang, L.; Chawla, R.; Pliner, H.A.; Trapnell, C. Reversed Graph Embedding Resolves Complex Single-Cell Trajectories. *Nat. Methods* **2017**, *14*, 979–982. [CrossRef]

244. Qiu, X.; Hill, A.; Packer, J.; Lin, D.; Ma, Y.-A.; Trapnell, C. Single-Cell MRNA Quantification and Differential Analysis with Census. *Nat. Methods* **2017**, *14*, 309–315. [CrossRef] [PubMed]

245. Schep, A.N.; Wu, B.; Buenrostro, J.D.; Greenleaf, W.J. ChromVAR: Inferring Transcription-Factor-Associated Accessibility from Single-Cell Epigenomic Data. *Nat. Methods* **2017**, *14*, 975–978. [CrossRef]

246. Hie, B.; Bryson, B.; Berger, B. Efficient Integration of Heterogeneous Single-Cell Transcriptomes Using Scanorama. *Nat. Biotechnol.* **2019**, *37*, 685–691. [CrossRef]

247. Wolock, S.L.; Lopez, R.; Klein, A.M. Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. *Cell Syst.* **2019**, *8*, 281–291.e9. [CrossRef]

248. Stoeckius, M.; Hafemeister, C.; Stephenson, W.; Houck-Loomis, B.; Chattopadhyay, P.K.; Swerdlow, H.; Satija, R.; Smibert, P. Simultaneous Epitope and Transcriptome Measurement in Single Cells. *Nat. Methods* **2017**, *14*, 865–868. [CrossRef] [PubMed]

249. Ramírez, F.; Ryan, D.P.; Grüning, B.; Bhardwaj, V.; Kilpert, F.; Richter, A.S.; Heyne, S.; Dündar, F.; Manke, T. DeepTools2: A next Generation Web Server for Deep-Sequencing Data Analysis. *Nucleic Acids Res.* **2016**, *44*, W160–W165. [CrossRef] [PubMed]

250. Katsman, E.; Orlanski, S.; Martignano, F.; Fox-Fisher, I.; Shemer, R.; Dor, Y.; Zick, A.; Eden, A.; Petrini, I.; Conticello, S.G.; et al. Detecting Cell-of-Origin and Cancer-Specific Methylation Features of Cell-Free DNA from Nanopore Sequencing. *Genom. Biol.* **2022**, *23*, 158. [CrossRef]

251. GitHub—Methylgrammarlab/Cfdna-Ont. Available online: https://github.com/methylgrammarlab/cfdna-ont (accessed on 5 March 2023).

252. GitHub—Mouliere-Lab/FrEIA: Fragment End Integrated Analysis Tool. Available online: https://github.com/mouliere-lab/FrEIA (accessed on 5 March 2023).

253. Moldovan, N.; van der Pol, Y.; van den Ende, T.; Boers, D.; Verkuijlen, S.; Creemers, A.; Ramaker, J.; Vu, T.; Fransen, M.F.; Pegtel, M.; et al. Genome-Wide Cell-Free DNA Termini in Patients with Cancer. *medRxiv* **2021**. [CrossRef]

254. GitHub—Friend1ws/Nanomonsv: SV Detection Tool for Nanopore Sequence Reads. Available online: https://github.com/friend1ws/nanomonsv (accessed on 5 March 2023).

255. Shiraishi, Y.; Koya, J.; Chiba, K.; Okada, A.; Arai, Y.; Saito, Y.; Shibata, T.; Kataoka, K. Precise Characterization of Somatic Complex Structural Variations from Paired Long-Read Sequencing Data with Nanomonsv. *bioRxiv* **2023**. [CrossRef]

256. Tham, C.Y.; Tirado-Magallanes, R.; Goh, Y.; Fullwood, M.J.; Koh, B.T.H.; Wang, W.; Ng, C.H.; Chng, W.J.; Thiery, A.; Tenen, D.G.; et al. NanoVar: Accurate Characterization of Patients' Genomic Structural Variants Using Low-Depth Nanopore Sequencing. *Genom. Biol.* **2020**, *21*, 56. [CrossRef]

257. Zafar, H.; Wang, Y.; Nakhleh, L.; Navin, N.; Chen, K. Monovar: Single-Nucleotide Variant Detection in Single Cells. *Nat. Methods* **2016**, *13*, 505–507. [CrossRef]

258. Ross, E.M.; Markowetz, F. OncoNEM: Inferring Tumor Evolution from Single-Cell Sequencing Data. *Genom. Biol.* **2016**, *17*, 69. [CrossRef]

259. Jahn, K.; Kuipers, J.; Beerenwinkel, N. Tree Inference for Single-Cell Data. *Genom. Biol.* **2016**, *17*, 86. [CrossRef] [PubMed]

260. Grün, D.; Muraro, M.J.; Boisset, J.-C.; Wiebrands, K.; Lyubimova, A.; Dharmadhikari, G.; van den Born, M.; van Es, J.; Jansen, E.; Clevers, H.; et al. De Novo Prediction of Stem Cell Identity Using Single-Cell Transcriptome Data. *Cell Stem Cell* **2016**, *19*, 266–277. [CrossRef] [PubMed]

261. Tsoucas, D.; Yuan, G.-C. Recent Progress in Single-Cell Cancer Genomics. *Curr. Opin. Genet. Dev.* **2017**, *42*, 22–32. [CrossRef]

262. Garvin, T.; Aboukhalil, R.; Kendall, J.; Baslan, T.; Atwal, G.S.; Hicks, J.; Wigler, M.; Schatz, M.C. Interactive Analysis and Assessment of Single-Cell Copy-Number Variations. *Nat. Methods* **2015**, *12*, 1058–1060. [CrossRef]

263. Deger, T.; Mendelaar, P.A.J.; Kraan, J.; Prager-van der Smissen, W.J.C.; van der Vlugt-Daane, M.; Bindels, E.M.J.; Sieuwerts, A.M.; Sleijfer, S.; Wilting, S.M.; Hollestelle, A.; et al. A Pipeline for Copy Number Profiling of Single Circulating Tumour Cells to Assess Intrapatient Tumour Heterogeneity. *Mol. Oncol.* **2022**, *16*, 2981–3000. [CrossRef] [PubMed]

264. Zhang, W.; Wei, L.; Huang, J.; Zhong, B.; Li, J.; Xu, H.; He, S.; Liu, Y.; Liu, J.; Lv, H.; et al. CfDNApipe: A Comprehensive Quality Control and Analysis Pipeline for Cell-Free DNA High-Throughput Sequencing Data. *Bioinformtics* **2021**, *37*, 4251–4252. [CrossRef]

265. CfDNApipe. Available online: https://xwanglabthu.github.io/cfDNApipe/ (accessed on 1 March 2023).