

# CfDNA Sequencing

Subjects: Genetics & Heredity

Contributor: Elodie Bohers

Cell free circulating DNA (cfDNA) refers to DNA fragments present outside of cells in body fluids such as plasma, urine, and cerebrospinal fluid (CSF). CfDNA was first identified in 1948 from plasma of healthy individuals. Afterward, studies showed that the quantity of this cfDNA in the blood was increased under pathological conditions such as auto-immune diseases but also cancers.

Keywords: cell-free DNA ; circulating tumoral DNA ; sequencing technologies ; bioinformatics

---

## 1. Introduction

Cell free circulating DNA (cfDNA) refers to DNA fragments present outside of cells in body fluids such as plasma, urine, and cerebrospinal fluid (CSF). In 1989, Philippe Anker and Maurice Stroun, from the University of Geneva, demonstrated that this cfDNA from cancer patients carries the characteristics of the DNA from tumoral cells <sup>[1]</sup>. Then, the research and identification of genomic anomalies specific of a cancer type in the circulating DNA, such as *NRAS* and *KRAS* mutations or *HER-2* amplifications <sup>[2][3][4]</sup>, started to expand, and for the first time, the term of circulating tumor DNA (ctDNA) appeared.

## 2. Detection of ctDNA by Sequencing Technologies

Sequencing technologies for detection and analysis of the ctDNA range from point mutations analyses using PCR-based methods to analyses of whole genome using NGS based methods. The choice of the method employed depends on the application and the sensitivity intended.

Targeted approaches can detect, with high sensitivity, specificity and at a fast and cost-effective rate, already known recurrent mutations. These hotspot mutations frequently occur in a specific type of tumor and can be, most of the time, targeted by a therapy. Thus, targeted approaches can be very useful for the follow up of minimal residual disease to early detect relapse or track resistant mutations. Contrariwise, untargeted approaches are less sensitive but are useful for the discovery of new DNA mutations and genome wide alterations such as copy number variations (CNV, or copy number alterations, CNA).

Several parameters in the sequencing processing can affect the sensitivity of detection. For example, as around 3000 copies of haploid genome are present in 10 ng of DNA, approximately 60 ng of cfDNA will be required for a sensitivity of 0.01% (one rare event in 10,000 molecules), which is often challenging, even more if we consider that more than one observation is necessary to determine a true variant. Amplification steps cannot replace low input of cfDNA because the polymerase will introduce errors, increasing the risk to have false positive variants. Another parameter that may improve sensitivity is to monitor multiple alterations simultaneously in order to increase the chances of detecting ctDNA.

### PCR-based methods

PCR-based methods, such as the derivatives of quantitative PCR (qPCR) and digital PCR, are fast, cost-effective, and relatively simple to carry out and analyze. They allow detection of single or few mutations at low variants allele frequency, up to 0.1% and less, with high specificity.

#### - Quantitative PCR

At first, the qPCR method, by measuring the fluorescence emitted by a labeled probe during amplification of a targeted gene, was used to estimate the concentration of cfDNA in plasma of patients with cancer <sup>[5]</sup>. Later, qPCR assays were developed to detect mutations in tumoral cfDNA and the sensitivity of detection was improved by promoting the specific amplification of the mutant allele.

ARMS-PCR (amplification-refractory mutation system) is a simple method for detecting point mutations or small deletions, in which DNA is amplified by allele specific primers. In this technique, the lack of 3' to 5' exonuclease proofreading activity of the Taq polymerase reduces dramatically the annealing and hence the amplification in case of mismatch at the 3' end of the primer. Although there are some improvements of the method, the false positive rate is still high with a limit of detection around 0.5 to 1% in plasma samples [6][7]. This limit can go down to 0.015% with ARMS-plus that includes a "Wild-type blocker" and in which amplicons were shortened to 50–80 bp, prohibiting the non-specific amplification and thus increasing the detection specificity [8].

PNA-LNA (peptide nucleic acid-locked nucleic acid) Clamp PCR uses a blocking synthetic nucleic acid analog complementary to wild type sequence to favor the amplification of the mutant allele. This method is particularly used in non-small cell lung cancer (NSCLC) to detect *EGFR* mutations, especially T790M mutations in tumor resistant to EGFR-TKIs (tyrosine kinase inhibitors), where cfDNA could be an alternative to the re-biopsy. Using smaller PCR products and by increasing the number of cycles, Watanabe and colleagues reached less than 0.1% detection rate [9].

COLD PCR (co-amplification at lower denaturation temperature-PCR) is an amplification method that selectively enriches low-abundance variant alleles from a mixture of wild-type and variation-containing DNA, irrespective of mutation type and position, by exploiting the critical denaturation temperature. The use of a lower denaturation temperature results in selective denaturation of molecules containing wild-type mutant heteroduplexes, which is followed by amplification. COLD-PCR has been used to improve the reliability of a number of different assays that traditionally use conventional PCR, such as Sanger sequencing, pyrosequencing or qPCR, greatly increasing their sensitivity. Thus, this method can detect mutant allele fraction down to 0.1% [10][11].

#### - Digital PCR

Digital PCR or droplet digital PCR (dPCR, ddPCR) allows clonal amplification of single DNA molecules, based on water-oil emulsion droplet technology. The PCR solution is partitioned into tens of thousands nano-liter sized droplet reactors where nucleic acid molecules are randomly distributed. A specific fluorescent amplification of the template (wild-type or mutated DNA) occurs in each individual droplet and positive droplets are subsequently counted to give an absolute target quantification using Poisson's statistics. The detection rate of hot-spot mutations with dPCR is 0.1% to 0.01% [12][13][14].

As an example in lymphoma, this technique has a potential clinical use in diffuse large B cell lymphoma (DLBCL), as co-occurring mutations in MYD88 and CD79B can predict response to Ibrutinib treatment, thus providing a predictive molecular tool for patient and therapy selection [15]. As well, in primary central nervous system lymphoma (PCNSL), mutation MYD88 L265P was identified by ddPCR in cerebrospinal fluid or vitreous fluid with a superior sensitivity when compared with qPCR [16][17]. Since this mutation is found in up to 85% of PCNSL cases and not in non-hematological brain tumors, this ddPCR assay may be a promising technique for minimally invasive confirmation of PCNSL diagnosis.

BEAMing (beads, emulsion, amplification, magnetics) is a highly sensitive digital PCR method that combines emulsion PCR and flow cytometry to identify and quantify specific somatic mutations present in DNA [18]. Diehl and coworkers used a BEAMing approach to detect mutations in cfDNA from patients with colorectal cancer, showing that ctDNA dynamics reflects tumor responses and progression, and that ctDNA detection after surgery represented a marker of residual disease [19]. This method, mainly used so far in solid tumors, such as colorectal [20], breast [21], and lung cancers [22], has a highly sensitive detection rate with variant allele fraction as low as 0.01%.

Although ddPCR allows for quantitative assessment of mutant frequencies in cfDNA, it is limited by the number of fluorescent probes that can be used in one assay (up to five) [23][24].

Copy number variations have also been investigated in cfDNA using ddPCR. Even if the number of targets is limited, it can be a useful tool for detecting, simply and rapidly, some gains or losses, which are associated with poor prognosis at diagnosis or during follow-up [25][26].

DdPCR can also be suitable to detect chromosomal rearrangements, especially in hematological malignancies. Among others, assays have been developed for translocation t(11;14) deregulating the *CCND1* gene and translocation t(14;18) deregulating the *BCL2* gene, which are frequently observed in Mantle cell lymphoma (MCL) and follicular lymphoma (FL), respectively [27][28]. The sensitivity of these techniques can go down to 0.01%.

#### - PCR coupled with mass spectrometry

The major limitation of the previous PCR-based approaches is their very limited multiplexing ability. Mass spectrometry-based methods such as surface-enhanced Raman spectroscopy (SERS) and UltraSEEK are adaptation of the conventional PCR method with a unique advantage in multiplexing to detect ctDNA mutations at low frequency with low

input amount of cfDNA and fast turnaround time.

SERS is a surface-sensitive technique that enhances Raman scattering by molecules adsorbed on rough metal surfaces or by nanostructures such as plasmonic-magnetic silica nanotubes [29]. The detection of target specific DNA is based on the use of labeled nanotags (Raman reporters) and the measurement of the Shift in the spectrum of Raman reporter that can provide information about low-frequency transitions in molecules. The status of mutations is then analyzed with SERS spectrum where unique spectral peaks demonstrated the presence of targeted mutations. Multiplex PCR/SERS identifying three hotspot mutations has been developed in melanoma and colorectal cancer with a limit of detection as few as 0.1% [30][31].

The UltraSEEK chemistry is able to interrogate multiple informative variants within a single reaction. In this method, the mutant allele is specifically targeted by a primer extension step that omits the wild type allele. In NSCLC, the limit of detection of the UltraSEEK Lung Panel, consists of 73 variants, was 0.125–1% with low input of specific tumoral cfDNA fragments beforehand measured with the LiquidIQ Panel [32]. Of note, this study showed the importance of preanalytical cfDNA quality control and input amount for the accuracy of liquid biopsy testing.

Overall, these PCR-based assays are very effective tools for detecting mutations at a relatively low-cost, which make them feasible in routine clinical practices. The main limitation is the limited multiplexing ability, which restricts the possibility of targets and can lead to a greater consumption of material. Furthermore, the alterations detected must be previously known such as hotspot mutations, which is more suitable for a minimal residual disease but less as a diagnostic tool.

### **Targeted NGS based methods**

Targeted deep sequencing techniques are still limited to a certain number of regions but can cover entire genes or entire coding regions of genes. Thus, they are suitable for genes without hotspot mutations, which is often the case for loss of function mutations in tumor suppressor genes.

Targeted enrichment in library construction can be achieved by direct amplification (amplicon or multiplex PCR) or hybridization capture (hybrid capture) of the DNA regions of interest. Techniques using multiplex PCR-based methods are more dependent on the length of the fragments and may require several simultaneous reactions for target enrichment to cover a large region of a gene, consuming more DNA. Hybrid capture methods employ custom RNA probes complementary to targeted regions and are able to detect both single nucleotide variants (SNV) and structural variants [33]. In this method of enrichment, the fragmentation of cfDNA can lead to a heterogeneous coverage across targeted exons with a lower fragment depth in the edge regions of exons, which must be taken into consideration when designing the panels for ctDNA sequencing [34].

Although they have high sensitivity and specificity, NGS platforms show a random error rate between 0.1 and 1.5% per base call, but library preparation protocols have been upgraded to improve the detection of rare variants [35][36]. In targeted DNA sequencing, the use of few DNA molecules combined with ultra-deep sequencing increases the risk to read several times the same molecule where polymerase errors are introduced at any step during the NGS process, leading to the inability to confidently call rare variants. One of the major recent technological advances is the use of molecular barcodes, which are random sequences introduced before any amplification step. They allow the counting of original DNA molecules instead of PCR duplicates, thereby enabling digital sequencing and resulting in unbiased and accurate mutation profiles with an increased sensitivity [37][38][39][40].

#### **- Tagged-amplicon deep sequencing (Tam-Seq)**

Tam-Seq is an amplicon method using a target enrichment array with barcoded primers to prepare the amplicon library for NGS. First, an initial targeted preamplification step is carried out, followed by a selective amplification of the regions of interest in single-plex reactions. Then, sequencing adaptors and sample-specific barcodes are attached to the amplicons in a further PCR. It was first able to detect mutations in circulating DNA with high sensitivity and specificity (>97%) at allele frequencies as low as 2% [41]. The technique has been recently improved (enhanced Tam-Seq, eTam-Seq) with a primer design strategy, allowing for amplification of highly fragmented DNA, a workflow reducing the background error rate, and a more efficient calling algorithm with better detection of SNV and indels (insertions/deletions), and also CNV [42]. This assay, using an optimal amount of DNA, detected 94% mutations at 0.25–0.33% allele fraction (AF) with a limit of detection down to 0.02% AF with high per-base specificity (99.9997%). In this study comparison of eTam-Seq with dPCR showed a good concordance between the two techniques, demonstrating the quantitative accuracy of eTam-Seq technology for reliable detection of mutations at low allele frequency [42].

#### **- Safe-Sequencing System (Safe-SeqS)**

This amplicon method was originally described by the group of Bert Vogelstein [37]. It was the first approach using molecular barcodes in DNA sequencing, to increase sensitivity of massively parallel sequencing. Thus, this method allows a correction of amplification and sequencing errors and can quantify rare mutations with a sensitivity of 0.05% of allele fraction. Safe-SeqS showed high performance in detecting mutations in cfDNA from patients with solid tumors, for molecular profiling as well as real-time monitoring of minimal residual disease [43]. A recent study on three independent cohort of nonmetastatic colorectal cancer, showed a median mutant allele frequency of 0.046% with a minimum of 0.01% [44].

#### **- Duplex sequencing**

Duplex sequencing is an improvement of the Safe-SeqS technique [45][46]. In this method, a semi-degenerated double stranded unique barcoded adapter is ligated to a target double stranded DNA. Thus, in addition to get rid of PCR and sequencing errors, the advantage of this technique is to identify artifacts due to sample alterations [47] because it can examine both strands individually and the damage to them is usually not identical (error correction by double-stranded consensus sequence). The theoretical sensitivity of this approach to discovering mutants is one molecule among  $10^7$  which is much higher in accuracy than conventional next-generation sequencing methods [45][46].

Several studies, in various types of cancers, applied this method on plasma cfDNA. In combination with target enrichment using hybrid capture, this approach allowed detection of tumoral fraction at 0.1% and below with high sensitivity and specificity, providing a powerful tool for diagnosis as well as longitudinal monitoring of disease [48][49][50].

#### **- Targeted error correction sequencing (TEC-Seq)**

In this technique, molecular barcoding is also used to facilitate the discrimination between true mutations and false positive variants. This combination of barcodes allows keeping track of each fragment as they are sequenced around 30,000 times [38]. The analytical sensitivity was 100% and 89% for detecting mutations present at 0.2% and 0.1%, respectively, using minimum thresholds of 0.05% in hot-spot positions and 0.1% at all other locations, resulting in a sensitivity of 97.4% overall, and without detection of false positives (less than one error in three million bases sequenced).

#### **- Single primer extension (SPE) with unique molecular barcode**

This approach uses only one gene specific primer (GSP) for amplification of each genomic region, which makes it less dependent on the size of DNA fragments than PCR using two primers and offers a uniform coverage. As for capture, the first step is a fragmentation step in which the buffer used inhibits fragmentation of the high length fragments of DNA such as contaminating. These adapters will be used for amplification of targeted region (together with GSP) and contain the degenerated molecular barcodes (UMI, Unique Molecular Index). Moreover, given this UMI contains 12 base pairs, it allows a large number of combinations and a very little risk for redundancy [39].

This technique of deep sequencing, using molecular barcodes to improve accuracy in variant detection, has been used at diagnosis in order to identify actionable genetic alteration with targeted therapies available for treatment or hotspot mutations to be tracked with ddPCR during follow up, with a detection of variant allele frequency down to 1–5% [51][52]. Further investigations are needed to find the real limit of detection of this technology, which may be below 1% as other techniques using molecular barcoding.

This approach also allowed detection of CNV. In PCR-based library construction, amplification introduces biases in further reads count because the amplification factor is dependent on many parameters such as library size, GC content, region length or competition between primers overlapping the same locus. Thus, the use of UMI via the mCNA tool allows the direct count of targeted DNA molecules before any amplification and the detection of CNV in a robust and sensitive way [53].

#### **- Cancer Personalized Profiling by Deep Sequencing (CAPP-Seq)**

In this technique, the first important step is to query cancer databases to identify known recurrent mutations for a particular cancer type. Then, biotinylated oligonucleotide probes, named “Selector”, are designed to target large segments of the concerned regions. The sensitivity is also improved by its ability to detect simultaneously various types of alterations: single nucleotide variants, rearrangements, insertions/deletions, and copy number alterations. It was originally described to detect and monitor lung cancer but was successfully adapted to a broad range of cancers, including different types of solid tumors as well as hematological malignancies such as DLBCL, LF, and HL [54][55][56][57][58].

With this method, ctDNA was detected in blood of NSCLC patients with 96% specificity for mutant allele fraction down to 0.02%. This iDES-enhanced CAPP-Seq combines CAPP-Seq with duplex barcoding sequencing technology and with a computational algorithm that removes stereotypical errors associated with the CAPP-Seq hybridization step. This

improved version of CAPP-Seq has shown a high sensitivity in the detection of *EGFR* mutations in cfDNA of NSCLC patients, with variant allele frequency as low as 0.004% with >99.99% specificity. Moreover, using duplex sequencing and covering a large number of mutations ( $\geq 200$ ), the authors outperformed iDES and managed to detect ctDNA down to 0.00025%, with an input of only 32ng of cfDNA [59].

#### - Immunoglobulin high-throughput sequencing (Ig-HTS)

This test was specifically developed for MRD in hematological malignancies. In this method, ultra-deep sequencing of genomic DNA, with a set of locus-specific multiplex PCR covering all possible rearranged IgH, IgK, and IgL receptor gene sequences, firstly identifies the tumor-specific clonotype. Then, this clonotype can be tracked as a specific fingerprint to quantify ctDNA in lymphoma disease monitoring with a sensitivity of approximately  $10^{-6}$  [60][61]. This technique presents some technical limitations, including the need of tissue biopsy to identify clonotype and difficulties to identify clonotype sequences in some lymphoma types such as DLBCL of the germinal center type and FL because of somatic hypermutation (SHM). Nevertheless, this method has shown high performance in surveillance ctDNA, after complete remission, to identify risk of recurrence before any clinical evidence of disease in most patients (with a median of 3.5 months)

This approach was also used for MRD monitoring in DLBCL patients after CAR-T cell therapy, showing correlation with clinical and radiologic outcomes for all the patients tested [62].

#### Untargeted NGS based methods :

As mentioned previously, untargeted approaches, namely whole exome and whole genome sequencing (WES, WGS), are less sensitive than targeted approaches. The sensitivity of these techniques on cfDNA is estimated around 5–10%, as compared to less than 0.1% for a targeted sequencing approach [63], making it difficult to detect rare events, especially in situations of early detection or minimal residual disease. However, these approaches may be necessary for the discovery of new alterations in the context of initial profiling at diagnosis, to provide information for the use of more sensitive targeted techniques during disease monitoring. Even if they are not suitable to detect subclonal events, they may be useful, considering intra tumoral heterogeneity, to highlight new drug targets or to track drug resistance clones [64].

WES is, most of the time, limited to coding regions and splicing sites of genes. The technical feasibility of whole-exome sequencing (WES) on cfDNA has been demonstrated in various solid tumors and some hematological malignancies [64]. Low coverage and sensitivity, compared to targeted NGS technologies does not allow for the detection of rare variants but WES of cfDNA is suitable for mutational analysis of patients with advanced tumors and increased ctDNA fractions (>5% mutant allele fraction). In metastatic patients, WES could serve as a surrogate for tumor genome analysis, considering the difficulties of doing multiple biopsies and the high ctDNA allele frequencies [65].

Additionally, given intra tumoral heterogeneity, analysis comparing mutational profile between tumor and cfDNA mostly identified more mutations in cfDNA with a high prevalence of targetable genes. Beyond SNV detection, WES of cfDNA also allowed analysis of mutational signatures, copy number variations, fusion genes, rearrangements, predicted neoantigens, and tumor mutational burden [64].

Contrariwise to WES, WGS technologie is more suitable to detect ctDNA by identifying structural and non-coding variations such as genome-wide copy number aberrations, methylation profiles, and fragmentation patterns.

To override the cost and analysis time limitations caused by WGS, Heitzer and colleagues developed a shallow genome-wide sequencing approach called Plasma-Seq [66]. This technique does not have a sufficient sequencing resolution to identify SNV but is able to detect CNV in cfDNA at a depth of 0.1x, with a specificity >80% when ctDNA fraction is  $\geq 10\%$ . Recently, this approach of shallow WGS has been successfully used in cfDNA of DLBCL and HL patients to identify copy number patterns that can differentiate the two diseases at diagnosis [67]. These copy number aberrations were also correlated with clinical parameters, and longitudinal analyses showed correlation with disease status. than in tumor, as for mutation detection [54][68][67].

Aneuploidy has also been explored with WGS derived techniques such as Fast-SeqS (Fast Aneuploidy Screening Test-Sequencing System) and (Within Sample Aneuploidy Detection), using a single specific primer pair to amplify dispersed retrotransposon regions throughout the genome (long interspersed nuclear elements (LINEs)) [69][70]. By simulations with synthetic DNA, the bioinformatic tool WALDO showed high performance to detect individual chromosome arm gain or loss with a fraction of ctDNA >5%, and up to 1% of tumoral fraction with a sensitivity of 78%. However, due to their mechanism of detection, these techniques are limited to cancers presenting aneuploidy.

In order to detect genomic rearrangements, Leary et al. developed a technique called PARE (personalized analysis of rearranged ends), which uses WGS mate-paired analysis of the tumoral DNA to identify patient specific genomic rearrangement. Analyses, in breast and colorectal cancers, suggest that ctDNA concentrations at levels >0.75% could be detected in the cfDNA of patients with a sensitivity >90% and a specificity >99%, and that even a single copy of rearrangement from ctDNA can be detected without false positives<sup>[71]</sup>. In a recent study, PARE was employed to detect rearrangements in gastric tumor, which were used to design a quantitative PCR assay targeting rearranged loci for quantitative monitoring in cfDNA<sup>[72]</sup>.

WGS, combined with artificial intelligence, can also identify genome-wide fragmentation patterns in cfDNA. Several studies in different cancer types have shown that these patterns can be used to detect ctDNA in body fluids and with very low plasma ctDNA fraction <sup>[73][74]</sup>. Indeed, the cfDNA fragmentation landscape represents a nucleosome footprint reflecting the cell and tissue of origin, potentially enabling non-invasive diagnosis of cancer type <sup>[73]</sup>. This nucleosome footprinting firstly identified by WGS represents nucleosome depletion at transcription start sites of highly expressed genes and the capture of this chromatin accessibility profile was used by CAPP-Seq technology to define gene expression differences and thus determine the cell-of-origin in DLBCL subjects from cfDNA <sup>[75]</sup>.

Among epigenetic alterations, aberrant DNA methylation events can also represent an ideal biomarker for detection and classification of early stage cancer, as they occur early in cancer development, sometimes before the acquisition of SNVs. Multiple liquid biopsy studies have been performed utilizing DNA methylation markers in various cancer types <sup>[76]</sup>. These strategies are either very targeted, as methylation events of interest occur at known, stereotyped positions <sup>[77]</sup>, or larger to identify methylation patterns, which have been shown to enable accurate determination of cell-of-origin from cfDNA and non-invasive cancer classification. In lymphoma, aberrant promoter methylation patterns detected in cfDNA have been shown to be an independent and significant poor prognostic factor for 5-year overall survival in DLBCL, outperforming existing clinical risk parameters an independent <sup>[78][79]</sup>.

Moreover, as healthy cells also participate to epigenetic changes, it may need to be distinguished from these of cancers cells <sup>[76]</sup>. Thus, it could be of major interest to combine epigenetic analysis of the entire cfDNA pool with mutational analysis of ctDNA molecules.

### **3. Bioinformatical Methods**

While common bioinformatics strategies allow variant identification down to 2–5% allele frequency, in most cases, ctDNA accounts for a small fraction of total cfDNA since most of cfDNA is derived from non-cancer cells and especially blood cells. ctDNA fraction can be lower than 0.1%, leading to the detection of somatic mutations at the same level as the sequencing noise. It implies the use of in silico strategies to distinguish true positive variant calls from sequencing noise.

It has been reported from healthy controls that under an allele fraction of 0.02%, more than 50% of sequenced genomic positions had sequencing artifacts <sup>[59]</sup>. These errors are particularly due to library preparation, the error rates of NGS technologies, and the physical characteristic of the cfDNA fragments.

In addition, there are many tools and therefore many bioinformatics parameters that need to be optimized when analyzing cfDNA samples. While major progress has been made in the harmonization of tumor analyzes with the GATK4 Best Practices Workflows <sup>[80]</sup>, there is not yet an international consensus for bioinformatic cfDNA analysis and research in this area remains very active.

#### **Adapter contamination**

The quality of cfDNA analysis is particularly impacted by adapter contamination. fragments could be shorter than usual which may result in the sequencing of adapters due to too many sequencing cycles compared to their lengths. Consequently, these reads could be either unmappable to the reference genome or could have a lower alignment score. These alignment scores are considered by a large number of bioinformatic tools and could finally affect the results of variant caller algorithms.

Many softwares were developed to find and trim adapters, like Cutadapt <sup>[81]</sup>, TagCleaner <sup>[82]</sup>, Trim Galore <sup>[83]</sup>, or Trimmomatic <sup>[84]</sup>. In general, these algorithms also integrate the trimming of low quality nucleotides and the extraction of molecular barcodes.

## Library biases and molecular barcoding

The amplification of the libraries by PCR includes many biases for counting mutated reads because the number of aligned reads is no longer directly proportional to the number of initial unique targeted DNA fragments. The amplification factor of each region is unknown and depends on many parameters such as library size, GC content, or fragment length. Recent advances in library preparation allow the addition of Unique Molecular Identifiers (UMI) to each read. UMIs are especially useful to correct library amplification biases by making each DNA molecule in a population of reads distinct.

There are two main bioinformatic approaches to use UMI for cfDNA analysis.

The first one consists in grouping PCR duplicates prior to any downstream analysis by merging sequences harboring the same UMI tag. The advantage of this approach is that it allows the use of classic bioinformatic pipelines after deduplicating the reads. It erases amplification biases due to cfDNA characteristics. However, it no longer provides access to essential information such as the amplification factor of each UMI or the discordant mutation calls of reads having the same UMI.

More recent approaches consist in using new bioinformatic algorithms for variant and CNV calling which are able to take into account the information carried by the UMIs after alignment, i.e., at the end of data processing.

For example, the UMI-VarCal algorithm <sup>[85]</sup> tries to quantify the number of concordant and discordant UMIs for each candidate variant during the variant calling process. Concordant UMIs were defined as number of unique UMIs for which all the reads carrying these UMIs validate the presence of the variant. Conversely, discordant UMIs quantify the number of abnormal substitutions like sequencing or PCR errors. SmCounter uses a barcode level allele probability and UMI counts to reject candidate mutations lacking enough barcodes with good read evidence.

Many biases due to the amplification step while preparing sequencing libraries prevent the direct quantification of loci copy-number <sup>[86]</sup>. cfDNA fragments are often shorter than DNA extracted from tissue and make it impossible to use conventional approaches for the detection of CNV such as read-depth algorithms. Recent approaches, like mCNA <sup>[53]</sup>, use the UMI counts instead of read counts to improve high-resolution copy number variation of genes.

## Bioinformatics processing

There is not yet an international consensus for bioinformatic cfDNA analysis pipeline. The bioinformatics tools and parameters must be adapted to the nature of the sequenced samples (quantity of DNA, quality of extraction, integrity of the cfDNA, etc.), to the kits used to prepare libraries, to the presence of UMI in library construction or not, and finally to the sequencing depth. In addition, sequencing biases are often sample specific which requires an objective assessment of sequencing noise at sample level.

Some first tools, like UMI-Gen <sup>[87]</sup>, allow to create in silico alignment datasets to evaluate the performance of variant calling and filtration tools. UMI-Gen is a UMI-based read simulator, which reproduces targeted sequencing paired-end alignment files (BAM) by estimating sequencing noise from a set of reference BAM files. It is particularly useful for evaluating the performance of variant calling tools because it allows to vary many parameters (sequencing depth, number of initial UMI, etc.) and to insert variants at frequencies of interest during the simulation. It thus makes it possible to optimize bioinformatic pipelines according to the targeted panels or the sequencing technology.

---

## References

1. Stroun, M.; Anker, P.; Maurice, P.; Lyautey, J.; Lederrey, C.; Beljanski, M. Neoplastic Characteristics of the DNA Found in the Plasma of Cancer Patients. *Oncology* 1989, 46, 318–322.
2. Vasioukhin, V.; Anker, P.; Maurice, P.; Lyautey, J.; Lederrey, C.; Stroun, M. Point mutations of the N-ras gene in the blood plasma DNA of patients with myelodysplastic syndrome or acute myelogenous leukaemia. *Br. J. Haematol.* 1994, 86, 774–779.
3. Anker, P.; Lefort, F.; Vasioukhin, V.; Lyautey, J.; Lederrey, C.; Chen, X.Q.; Stroun, M.; Mulcahy, H.E.; Farthing, M.J. K-ras mutations are found in DNA extracted from the plasma of patients with colorectal cancer. *Gastroenterology* 1997, 112, 1114–1120.
4. Austrup, F.; Uciechowski, P.; Eder, C.; Böckmann, B.; Suchy, B.; Driesel, G.; Jäckel, S.; Kusiak, I.; Grill, H.J.; Giesing, M. Prognostic value of genomic alterations in minimal residual cancer cells purified from the blood of breast cancer patients. *Br. J. Cancer* 2000, 83, 1664–1673.

5. Hohaus, S.; Giachella, M.; Massini, G.; Mansueto, G.; Vannata, B.; Bozzoli, V.; Criscuolo, M.; D'Alò, F.; Martini, M.; Larocca, L.M.; et al. Cell-free circulating DNA in Hodgkin's and non-Hodgkin's lymphomas. *Ann. Oncol.* 2009, 20, 1408–1413.
6. Aung, K.L.; Donald, E.; Ellison, G.; Bujac, S.; Fletcher, L.; Cantarini, M.; Brady, G.; Orr, M.; Clack, G.; Ranson, M.; et al. Analytical Validation of BRAF Mutation Testing from Circulating Free DNA Using the Amplification Refractory Mutation Testing System. *J. Mol. Diagn.* 2014, 16, 343–349.
7. Siggillino, A.; Ulivi, P.; Pasini, L.; Reda, M.S.; Chiadini, E.; Tofanetti, F.R.; Baglivo, S.; Metro, G.; Crinó, L.; Delmonte, A.; et al. Detection of EGFR Mutations in Plasma Cell-Free Tumor DNA of TKI-Treated Advanced-NSCLC Patients by Three Methodologies: Scorpion-ARMS, PNA-Clamp, and Digital PCR. *Diagnostics* 2020, 10, 1062.
8. Zhang, X.; Chang, N.; Yang, G.; Zhang, Y.; Ye, M.; Cao, J.; Xiong, J.; Han, Z.; Wu, S.; Shang, L.; et al. A comparison of ARMS-Plus and droplet digital PCR for detecting EGFR activating mutations in plasma. *Oncotarget* 2017, 8, 112014–112023.
9. 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, 1–7.
10. Milbury, C.A.; Li, J.; Liu, P.; Makrigiorgos, G.M. COLD-PCR: Improving the sensitivity of molecular diagnostics assays. *Expert Rev. Mol. Diagn.* 2011, 11, 159–169.
11. Galbiati, S.; Damin, F.; Burgio, V.; Brisci, A.; Soriani, N.; Belcastro, B.; Di Resta, C.; Gianni, L.; Chiari, M.; Ronzoni, M.; et al. Evaluation of three advanced methodologies, COLD-PCR, microarray and ddPCR, for identifying the mutational status by liquid biopsies in metastatic colorectal cancer patients. *Clin. Chim. Acta* 2019, 489, 136–143.
12. Vincent Camus; Nasrin Sarafan-Vasseur; Elodie Bohers; Sydney Dubois; Sylvain Mareschal; Philippe Bertrand; Pierre-Julien Viailly; Philippe Ruminy; Catherine Maingonnat; Emilie Lemasle; et al. Digital PCR for quantification of recurrent and potentially actionable somatic mutations in circulating free DNA from patients with diffuse large B-cell lymphoma. *L Leukemia & Lymphoma* 2015, 57, 2171-2179, [10.3109/10428194.2016.1139703](https://doi.org/10.3109/10428194.2016.1139703).
13. Joost H. van Ginkel; Manon M. H. Huibers; Robert J. J. van Es; Remco de Bree; Stefan M. Willems; Droplet digital PCR for detection and quantification of circulating tumor DNA in plasma of head and neck cancer patients. *BMC Cancer* 2017, 17, 428, [10.1186/s12885-017-3424-0](https://doi.org/10.1186/s12885-017-3424-0).
14. Laura S. Hiemcke-Jiwa; Monique C. Minnema; Joyce H. Radersma-Van Loon; N. Mehdi Jiwa; Mirthe De Boer; Roos J. Leguit; Roel A. De Weger; Manon M.H. Huibers; The use of droplet digital PCR in liquid biopsies: A highly sensitive technique for MYD88 p.(L265P) detection in cerebrospinal fluid. *Hematological Oncology* 2017, 36, 429-435, [10.1002/hon.2489](https://doi.org/10.1002/hon.2489).
15. Wilson, W.H.; Young, R.M.; Schmitz, R.; Yang, Y.; Pittaluga, S.; Wright, G.; Lih, C.-J.; Williams, P.M.; Shaffer, A.L.; Gerecitano, J.; et al. Targeting B cell receptor signaling with ibrutinib in diffuse large B cell lymphoma. *Nat. Med.* 2015, 21, 922–926.
16. Hiemcke-Jiwa, L.S.; Minnema, M.C.; Loon, J.H.R.-V.; Jiwa, N.M.; De Boer, M.; Leguit, R.J.; De Weger, R.A.; Huibers, M.M. The use of droplet digital PCR in liquid biopsies: A highly sensitive technique for MYD88 p.(L265P) detection in cerebrospinal fluid. *Hematol. Oncol.* 2018, 36, 429–435.
17. Chen, K.; Ma, Y.; Ding, T.; Zhang, X.; Chen, B.; Guan, M. Effectiveness of digital PCR for MYD88L265P detection in vitreous fluid for primary central nervous system lymphoma diagnosis. *Exp. Ther. Med.* 2020, 20, 301–308.
18. 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.
19. Diehl, F.; Schmidt, K.; Choti, M.A.; Romans, K.; Goodman, S.; Li, M.; Thornton, K.; Agrawal, N.; Sokoll, L.; Szabo, S.A.; et al. Circulating mutant DNA to assess tumor dynamics. *Nat. Med.* 2008, 14, 985–990.
20. Holdhoff, M.; Schmidt, K.; Donehower, R.; Diaz, L.A. Analysis of Circulating Tumor DNA to Confirm Somatic KRAS Mutations. *J. Natl. Cancer Inst.* 2009, 101, 1284–1285.
21. O'Leary, B.; Hrebien, S.; Beaney, M.; Fribbens, C.; Garcia-Murillas, I.; Jiang, J.; Li, Y.; Bartlett, C.H.; Andre, F.; Loibl, S.; et al. Comparison of BEAMing and Droplet Digital PCR for Circulating Tumor DNA Analysis. *Clin. Chem.* 2019, 65, 1405–1413.
22. Garcia, J.; Gauthier, A.; Lescuyer, G.; Barthelemy, D.; Geiguer, F.; Balandier, J.; Edelstein, D.L.; Jones, F.S.; Holtrup, F.; Duruisseau, M.; et al. Routine Molecular Screening of Patients with Advanced Non-SmallCell Lung Cancer in Circulating Cell-Free DNA at Diagnosis and During Progression Using OncoBEAM™ EGFR V2 and NGS Technologies. *Mol. Diagn. Ther.* 2021, 25, 239–250.

23. Butler, T.; Spellman, P.T.; Gray, J. Circulating-tumor DNA as an early detection and diagnostic tool. *Curr. Opin. Genet. D ev.* 2017, 42, 14–21.
24. Milbury, C.A.; Zhong, Q.; Lin, J.; Williams, M.; Olson, J.; Link, D.R.; Hutchison, B. Determining lower limits of detection of digital PCR assays for cancer-related gene mutations. *Biomol. Detect. Quantif.* 2014, 1, 8–22.
25. Shoda, K.; Ichikawa, D.; Fujita, Y.; Masuda, K.; Hiramoto, H.; Hamada, J.; Arita, T.; Konishi, H.; Komatsu, S.; Shiozaki, A.; et al. Monitoring the HER2 copy number status in circulating tumor DNA by droplet digital PCR in patients with gastric cancer. *Gastric Cancer* 2016, 20, 126–135.
26. Lee, K.S.; Nam, S.K.; Seo, S.H.; Park, K.U.; Oh, H.-K.; Kim, D.-W.; Kang, S.-B.; Kim, W.H.; Lee, H.S. Digital polymerase chain reaction for detecting c-MYC copy number gain in tissue and cell-free plasma samples of colorectal cancer patients. *Sci. Rep.* 2019, 9, 1–9.
27. Delfau-Larue, M.-H.; Van Der Gucht, A.; Dupuis, J.; Jais, J.-P.; Nel, I.; Beldi-Ferchiou, A.; Hamdane, S.; Benmaad, I.; Labouere, G.; Verret, B.; et al. Total metabolic tumor volume, circulating tumor cells, cell-free DNA: Distinct prognostic value in follicular lymphoma. *Blood Adv.* 2018, 2, 807–816.
28. Pott, C.; Brüggemann, M.; Ritgen, M.; Van Der Velden, V.H.J.; Van Dongen, J.J.M.; Kneba, M. MRD Detection in B-Cell Non-Hodgkin Lymphomas Using Ig Gene Rearrangements and Chromosomal Translocations as Targets for Real-Time Quantitative PCR. *Meth. Mol. Biol.* 2019, 199–228.
29. Pyrak, E.; Krajczewski, J.; Kowalik, A.; Kudelski, A.; Jaworska, A. Surface Enhanced Raman Spectroscopy for DNA Biosensors—How Far Are We? *Molecular* 2019, 24, 4423.
30. Wee, E.J.; Wang, Y.; Tsao, S.C.-H.; Trau, M. Simple, Sensitive and Accurate Multiplex Detection of Clinically Important Melanoma DNA Mutations in Circulating Tumour DNA with SERS Nanotags. *Theranostics* 2016, 6, 1506–1513.
31. Lyu, N.; Rajendran, V.K.; Diefenbach, R.; Charles, K.; Clarke, S.J.; Engel, A.; Rizos, H.; Molloy, M.P.; Wang, Y. Sydney 1000 Colorectal Cancer Study Investigators\* Multiplex detection of ctDNA mutations in plasma of colorectal cancer patients by PCR/SERS assay. *NanoTheranostics* 2020, 4, 224–232.
32. Lamy, P.-J.; Van Der Leest, P.; Lozano, N.; Becht, C.; Duboeuf, F.; Groen, H.J.M.; Hilgers, W.; Pourel, N.; Rifaela, N.; Schuur, E.; et al. Mass Spectrometry as a Highly Sensitive Method for Specific Circulating Tumor DNA Analysis in NSCLC: A Comparison Study. *Cancers* 2020, 12, 3002.
33. Mamanova, L.; Coffey, A.J.; Scott, C.E.; Kozarewa, I.; Turner, E.; Kumar, A.; Howard, E.; Shendure, J.; Turner, D.J. Target-enrichment strategies for next-generation sequencing. *Nat. Methods* 2010, 7, 111–118.
34. Deveson, I.W.; Gong, B.; Lai, K.; LoCoco, J.S.; Richmond, T.A.; Schageman, J.; Zhang, Z.; Novorodovskaya, N.; Wille, J.C.; Jones, W.; et al. Evaluating the analytical validity of circulating tumor DNA sequencing assays for precision oncology. *Nat. Biotechnol.* 2021, 1–14.
35. Glenn, T.C. Field guide to next-generation DNA sequencers. *Mol. Ecol. Resour.* 2011, 11, 759–769.
36. Loman, N.J.; Misra, R.V.; Dallman, T.J.; Constantinidou, C.; Gharbia, S.E.; Wain, J.; Pallen, M.J. Performance comparison of benchtop high-throughput sequencing platforms. *Nat. Biotechnol.* 2012, 30, 434–439.
37. 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.
38. 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.
39. Xu, C.; Ranjbar, M.R.N.; Wu, Z.; Dicarlo, J.; Wang, Y. Detecting very low allele fraction variants using targeted DNA sequencing and a novel molecular barcode-aware variant caller. *BMC Genom.* 2017, 18, 5.
40. Salk, J.J.; Schmitt, M.W.; Loeb, L.A. Enhancing the accuracy of next-generation sequencing for detecting rare and subclonal mutations. *Nat. Rev. Genet.* 2018, 19, 269–285.
41. 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.
42. Gale, D.; Lawson, A.R.J.; Howarth, K.; Madi, M.; Durham, B.; Smalley, S.; Calaway, J.; Blais, S.; Jones, G.; Clark, J.; et al. Development of a highly sensitive liquid biopsy platform to detect clinically-relevant cancer mutations at low allele fractions in cell-free DNA. *PLoS ONE* 2018, 13, e0194630.
43. Fostira, F.; Oikonomopoulou, P.; Kladi, A.; Edelstein, D.; Stieler, K.; Heim, D.; Gkatzamanidou, M.; Anastasiou, M.; Kotsantis, I.; Kavourakis, G.; et al. Blood-based testing of mutations in patients with head and neck squamous cell carcinoma (HNSCC) using highly sensitive SafeSEQ technology. *Ann. Oncol.* 2019, 30, v469.

44. Tie, J.; Cohen, J.D.; Lo, S.N.; Wang, Y.; Li, L.; Christie, M.; Lee, M.; Wong, R.; Kosmider, S.; Skinner, I.; et al. Prognostic significance of postsurgery circulating tumor DNA in nonmetastatic colorectal cancer: Individual patient pooled analysis of three cohort studies. *Int. J. Cancer* 2021, 148, 1014–1026.
45. Schmitt, M.W.; Kennedy, S.R.; Salk, J.J.; Fox, E.; Hiatt, J.B.; Loeb, L.A. Detection of ultra-rare mutations by next-generation sequencing. *Proc. Natl. Acad. Sci. USA* 2012, 109, 14508–14513.
46. Kennedy, S.R.; Schmitt, M.W.; Fox, E.J.; Kohn, B.; Salk, J.J.; Ahn, E.H.; Prindle, M.J.; Kuong, K.J.; Shen, J.-C.; Risques, R.-A.; et al. Detecting ultralow-frequency mutations by Duplex Sequencing. *Nat. Protoc.* 2014, 9, 2586–2606.
47. Costello, M.; Pugh, T.J.; Fennell, T.J.; Stewart, C.; Lichtenstein, L.; Meldrim, J.C.; Fostel, J.L.; Friedrich, D.C.; Perrin, D.; Dionne, D.; et al. Discovery and characterization of artifactual mutations in deep coverage targeted capture sequencing data due to oxidative DNA damage during sample preparation. *Nucleic Acids Res.* 2013, 41, e67.
48. Alcaide, M.; Yu, S.; Davidson, J.; Albuquerque, M.; Bushell, K.; Fornika, D.; Arthur, S.; Grande, B.M.; McNamara, S.; du Tertre, M.C.; et al. Targeted error-suppressed quantification of circulating tumor DNA using semi-degenerate barcoded adapters and biotinylated baits. *Sci. Rep.* 2017, 7, 10574.
49. Ren, Y.; Zhang, Y.; Wang, D.; Liu, F.; Fu, Y.; Xiang, S.; Su, L.; Li, J.; Dai, H.; Huang, B. SinoDuplex: An Improved Duplex Sequencing Approach to Detect Low-frequency Variants in Plasma cfDNA Samples. *Genom. Proteom. Bioinform.* 2020, 18, 81–90.
50. Mallampati, S.; Zalles, S.; Duose, D.Y.; Hu, P.C.; Medeiros, L.J.; Wistuba, I.I.; Kopetz, S.; Luthra, R. Development and Application of Duplex Sequencing Strategy for Cell-Free DNA-Based Longitudinal Monitoring of Stage IV Colorectal Cancer. *J. Mol. Diagn.* 2019, 21, 994–1009.
51. Sim, W.C.; Loh, C.H.; Toh, G.L.-X.; Lim, C.W.; Chopra, A.; Chang, A.Y.C.; Goh, L.L. Non-invasive detection of actionable mutations in advanced non-small-cell lung cancer using targeted sequencing of circulating tumor DNA. *Lung Cancer* 2018, 124, 154–159.
52. Zou, D.; Day, R.; Cocadiz, J.A.; Parackal, S.; Mitchell, W.; Black, M.A.; Lawrence, B.; Fitzgerald, S.; Print, C.; Jackson, C.; et al. Circulating tumor DNA is a sensitive marker for routine monitoring of treatment response in advanced colorectal cancer. *Carcinogenesis* 2020, 41, 1507–1517.
53. Viailly, P.-J.; Sater, V.; Viennot, M.; Bohers, E.; Vergne, N.; Berard, C.; Dauchel, H.; Lecroq, T.; Celebi, A.; Ruminy, P.; et al. Improving high-resolution copy number variation analysis from next generation sequencing using unique molecular identifiers. *BMC Bioinform.* 2021, 22, 1–15.
54. Spina, V.; Brusca, A.; Cuccaro, A.; Martini, M.; Di Trani, M.; Forestieri, G.; Manzoni, M.; Condoluci, A.; Arribas, A.; Terzi-Di-Bergamo, L.; et al. Circulating tumor DNA reveals genetics, clonal evolution, and residual disease in classical Hodgkin lymphoma. *Blood* 2018, 131, 2413–2425.
55. Kurtz, D.M.; Scherer, F.; Jin, M.C.; Soo, J.; Craig, A.F.; Esfahani, M.S.; Chabon, J.J.; Stehr, H.; Liu, C.L.; Tibshirani, R.; et al. Circulating Tumor DNA Measurements As Early Outcome Predictors in Diffuse Large B-Cell Lymphoma. *J. Clin. Oncol.* 2018, 36, 2845–2853.
56. Klass, D.; Newman, A.; Lovejoy, A.; Zhou, L.; Stehr, H.; Xu, T.; He, J.; Komaki, R.; Liao, Z.; Maru, D.; et al. Analysis of Circulating Tumor DNA in Esophageal Carcinoma Patients Treated with Chemoradiation Therapy. *Int. J. Radiat. Oncol.* 2015, 93, S104–S105.
57. Scherer, F.; Kurtz, D.M.; Newman, A.; Stehr, H.; Craig, A.F.M.; Esfahani, M.S.; Lovejoy, A.F.; Chabon, J.J.; Klass, D.M.; Liu, C.L.; et al. Distinct biological subtypes and patterns of genome evolution in lymphoma revealed by circulating tumor DNA. *Sci. Transl. Med.* 2016, 8, 364ra155.
58. Dudley, J.C.; Schroers-Martin, J.; Lazzareschi, D.V.; Shi, W.Y.; Chen, S.B.; Esfahani, M.S.; Trivedi, D.; Chabon, J.J.; Chaudhuri, A.A.; Stehr, H.; et al. Detection and Surveillance of Bladder Cancer Using Urine Tumor DNA. *Cancer Discov.* 2019, 9, 500–509.
59. Newman, A.; 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.
60. Mark Roschewski; Kieron Dunleavy; Stefania Pittaluga; Martin Moorhead; Francois Pepin; Katherine Kong; Margaret Shovlin; Elaine S Jaffe; Louis M Staudt; Catherine Lai; et al. Circulating tumour DNA and CT monitoring in patients with untreated diffuse large B-cell lymphoma: a correlative biomarker study. *The Lancet Oncology* 2015, 16, 541-549, [10.1016/S1470-2045\(15\)70106-3](https://doi.org/10.1016/S1470-2045(15)70106-3).
61. David M. Kurtz; Michael R. Green; Scott V. Bratman; Florian Scherer; Chih Long Liu; Christian Kunder; Kazuhiro Takahashi; Cynthia Glover; Colm Keane; Shingo Kihira; et al. Noninvasive monitoring of diffuse large B-cell lymphoma by immunoglobulin high-throughput sequencing. *Blood* 2015, 125, 3679-3687, [10.1182/blood-2015-03-635169](https://doi.org/10.1182/blood-2015-03-635169).

62. Hossain, N.M.; Dahiya, S.; Le, R.; Abramian, A.M.; Kong, K.A.; Muffly, L.S.; Miklos, D.B. Circulating tumor DNA assessment in patients with diffuse large B-cell lymphoma following CAR T-cell therapy. *Leuk. Lymphoma* 2019, 60, 503–506.
63. Corcoran, R.B.; Chabner, B.A. Application of Cell-free DNA Analysis to Cancer Treatment. *N. Engl. J. Med.* 2018, 379, 1754–1765.
64. Bos, M.K.; Angus, L.; Nasserinejad, K.; Jager, A.; Jansen, M.P.; Martens, J.W.; 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.
65. Muhammed Murtaza; Sarah-Jane Dawson; Dana W. Y. Tsui; Davina Gale; Tim Forshew; Anna M. Piskorz; Christine Parkinson; Suet-Feung Chin; Zoya Kingsbury; Alvin S. C. Wong; et al. Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. *Nature* 2013, 497, 108–112, [10.1038/nature12065](https://doi.org/10.1038/nature12065).
66. Ellen Heitzer; Peter Ulz; Jelena Belic; Stefan Gutsch; Franz Quehenberger; Katja Fischereeder; Theresa Benezeder; Martina Auer; Carina Pischler; Sebastian Mannweiler; et al. Tumor-associated copy number changes in the circulation of patients with prostate cancer identified through whole-genome sequencing. *Genome Medicine* 2013, 5, 30–30, [10.1186/gm434](https://doi.org/10.1186/gm434).
67. Raman, L.; Van Der Linden, M.; De Vriendt, C.; Broeck, B.V.D.; Muylle, K.; Deeren, D.; Dedeurwaerdere, F.; Verbeke, S.; Dendooven, A.; De Grove, K.; et al. Shallow-depth sequencing of cell-free DNA for Hodgkin and diffuse large B-cell lymphoma (differential) diagnosis: A standardized approach with underappreciated potential. *Haematologica* 2020.
68. Camus, V.; Viennot, M.; LeQuesne, J.; Vially, P.-J.; Bohers, E.; Bessi, L.; Marcq, B.; Etancelin, P.; Dubois, S.; Picquenoit, J.-M.; et al. Targeted genotyping of circulating tumor DNA for classical Hodgkin lymphoma monitoring: A prospective study. *Haematologica* 2020, 106, 154–162.
69. 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.
70. Douville, C.; Springer, S.; Kinde, I.; Cohen, J.D.; Hruban, R.H.; Lennon, A.M.; Papadopoulos, N.; Kinzler, K.W.; Vogelstein, B.; Karchin, R. Detection of aneuploidy in patients with cancer through amplification of long interspersed nucleotide elements (LINEs). *Proc. Natl. Acad. Sci. USA* 2018, 115, 1871–1876.
71. 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.
72. Young-Woo Kim; Yura Song; Han-Seong Kim; Hye Won Sim; Shiv Poojan; Bang Wool Eom; Myeong-Cherl Kook; Jungnam Joo; Kyeong-Man Hong; Young-Ho Kim; et al. Monitoring circulating tumor DNA by analyzing personalized cancer-specific rearrangements to detect recurrence in gastric cancer. *Experimental & Molecular Medicine* 2019, 51, 1–10, [10.1038/s12276-019-0292-5](https://doi.org/10.1038/s12276-019-0292-5).
73. Snyder, M.W.; Kircher, M.; Hill, A.J.; Daza, R.M.; Shendure, J. Cell-free DNA Comprises an In Vivo Nucleosome Footprint that Informs Its Tissues-Of-Origin. *Cell* 2016, 164, 57–68.
74. 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.
75. Mehrmohamadi, M.; Esfahani, M.S.; Soo, J.; Scherer, F.; Schroers-Martin, J.G.; Chen, B.; Kurtz, D.M.; Hamilton, E.; Liu, C.L.; Diehn, M.; et al. Distinct Chromatin Accessibility Profiles of Lymphoma Subtypes Revealed By Targeted Cell Free DNA Profiling. *Blood* 2018, 132, 672.
76. Van der Pol, Y.; Mouliere, F. Toward the Early Detection of Cancer by Decoding the Epigenetic and Environmental Fingerprints of Cell-Free DNA. *Cancer Cell* 2019, 36, 350–368.
77. Cirillo, M.; Craig, A.F.; Borchmann, S.; Kurtz, D.M. Liquid biopsy in lymphoma: Molecular methods and clinical applications. *Cancer Treat. Rev.* 2020, 91, 102106.
78. Kristensen, L.S.; Hansen, J.W.; Kristensen, S.S.; Tholstrup, D.; Harsløf, L.B.S.; Pedersen, O.B.; Brown, P.D.N.; Grønbaek, K. Aberrant methylation of cell-free circulating DNA in plasma predicts poor outcome in diffuse large B cell lymphoma. *Clin. Epigenet.* 2016, 8, 1–11.
79. Wedge, E.; Hansen, J.W.; Garde, C.; Asmar, F.; Tholstrup, D.; Kristensen, S.S.; Munch-Petersen, H.D.; Ralfkiaer, E.; Brown, P.; Grønbaek, K.; et al. Global hypomethylation is an independent prognostic factor in diffuse large B cell lymphoma. *Am. J. Hematol.* 2017, 92, 689–694.
80. Van der Auwera, G.A.; O’Connor, B.D. *Genomics in the Cloud: Using Docker, GATK, and WDL in Terra*, 1st ed.; O’Reilly Media: Cambridge, MA, USA, 2020.

81. Kechin, A.; Boyarskikh, U.; Kel, A.; Filipenko, M. cutPrimers: A New Tool for Accurate Cutting of Primers from Reads of Targeted Next Generation Sequencing. *J. Comput. Biol.* 2017, 24, 1138–1143.
82. Schmieder, R.; Lim, Y.W.; Rohwer, F.; Edwards, R. TagCleaner: Identification and removal of tag sequences from genomic and metagenomic datasets. *BMC Bioinform.* 2010, 11, 341.
83. Babraham Bioinformatics—Trim Galore! Available online: (accessed on 11 May 2021).
84. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 2014, 30, 2114–2120.
85. Sater, V.; Vially, P.-J.; Lecroq, T.; Prieur-Gaston, É.; Bohers, É.; Viennot, M.; Ruminy, P.; Dauchel, H.; Vera, P.; Jardin, F. UMI-VarCal: A new UMI-based variant caller that efficiently improves low-frequency variant detection in paired-end sequencing NGS libraries. *Bioinform. Oxf. Engl.* 2020, 36, 2718–2724.
86. Boeva, V.; Popova, T.; Lienard, M.; Toffoli, S.; Kamal, M.; Le Tourneau, C.; Gentien, D.; Servant, N.; Gestraud, P.; Frio, T.R.; et al. Multi-factor data normalization enables the detection of copy number aberrations in amplicon sequencing data. *Bioinform. Oxf. Engl.* 2014, 30, 3443–3450.
87. Sater, V.; Vially, P.-J.; Lecroq, T.; Ruminy, P.; Bérard, C.; Prieur-Gaston, É.; Jardin, F. UMI-Gen: A UMI-based read simulator for variant calling evaluation in paired-end sequencing NGS libraries. *Comput. Struct. Biotechnol. J.* 2020, 18, 2270–2280.
88. David M. Kurtz; Michael R. Green; Scott V. Bratman; Florian Scherer; Chih Long Liu; Christian Kunder; Kazuhiro Takahashi; Cynthia Glover; Colm Keane; Shingo Kihira; et al. Noninvasive monitoring of diffuse large B-cell lymphoma by immunoglobulin high-throughput sequencing. *Blood* **2015**, 125, 3679-3687, [10.1182/blood-2015-03-635169](https://doi.org/10.1182/blood-2015-03-635169).

---

Retrieved from <https://encyclopedia.pub/entry/history/show/27633>