



## The emerging role of cell-free DNA as a molecular marker for cancer management

Abel Jacobus Bronkhorst, Vida Ungerer, Stefan Holdenrieder\*

Institute for Laboratory Medicine, German Heart Centre, Technical University Munich, Lazarettstraße. 36, D-80636, Munich, Germany



### ARTICLE INFO

Handled by Chennai Guest Editor

#### Keywords:

Cell-free DNA  
Circulating tumor DNA  
Biomarker  
Liquid biopsy  
Cancer  
Precision oncology

### ABSTRACT

An increasing number of studies demonstrate the potential use of cell-free DNA (cfDNA) as a surrogate marker for multiple indications in cancer, including diagnosis, prognosis, and monitoring. However, harnessing the full potential of cfDNA requires (i) the optimization and standardization of preanalytical steps, (ii) refinement of current analysis strategies, and, perhaps most importantly, (iii) significant improvements in our understanding of its origin, physical properties, and dynamics in circulation. The latter knowledge is crucial for interpreting the associations between changes in the baseline characteristics of cfDNA and the clinical manifestations of cancer. In this review we explore recent advancements and highlight the current gaps in our knowledge concerning each point of contact between cfDNA analysis and the different stages of cancer management.

### 1. Introduction

Fragmented cell-free DNA (cfDNA) molecules were discovered in the human circulatory system in 1948 [1]. While this phenomenon seemed to be trivial at first, its clinical importance was recognized when researchers observed differences between the characteristics of cfDNA from healthy and diseased individuals. Raised concentrations of cfDNA were first reported for patients with autoimmune disease and leukemia [2,3]. In the subsequent decades numerous studies have intermittently demonstrated that cancer patients generally have high levels of cfDNA vs healthy subjects [4]. In 1989, Stroun and colleagues demonstrated that a fraction of the cfDNA present in the plasma of cancer patients is derived from cancer cells [5], and shortly thereafter another group detected TP53 mutations in the DNA of urinary sediments from patients with invasive bladder cancer [6]. Follow-up studies not only confirmed that cancer cells do release detectable amounts of cfDNA fragments into circulation and other biofluids, but also revealed that these fragments bear the unique genetic and epigenetic alterations that are characteristic of the tumor from which they originate (reviewed in [7]).

These proof-of-principle studies indicated that kinetic assessment and molecular profiling of cfDNA may serve a potentially useful role in non-invasive cancer management. This, in concurrence with the advent of ultrasensitive technologies (e.g., next-generation sequencing (NGS), BEAMing (beads, emulsions, amplification and magnetics), and droplet digital PCR (ddPCR)) and concomitant improvements in most analytical techniques acted as a catalyst for a surge in studies investigating the correlation between

the characteristics of cfDNA and a wide range of pathological signatures in cancer. Besides the possibility of screening both healthy or at-risk asymptomatic patient groups for early detection, diagnosis and treatment of cancer, this wave of studies has made it clear that cfDNA analysis holds considerable promise as a surrogate marker for multiple indications in oncology (Fig. 1), including (i) staging and prognosis, (ii) tumor localization, (iii) initial therapy stratification, (iv) monitoring response to local or systemic therapies in patients with a defined diagnosis, (v) monitoring minimal residual disease and relapse following the completion of the primary therapy, and (vi) identification of acquired drug resistance mechanisms.

Although there is still much progress to be made, the feasibility of cfDNA as a marker for cancer management is underscored by two FDA-approved applications for cfDNA assays in routine clinical practice, namely the cobas EGFR Mutation Test v2, an assay designed to help clinicians identify lung cancer patients that are eligible for erlotinib or osimertinib treatment [8,9], and Epi proColon, a colorectal cancer (CRC) screening test based on the methylation status of the SEPT9 promotor [10].

The potential clinical utility of cfDNA has been covered extensively in previous publications [7,11–19]. This review will explore the most recent advancements concerning each of the aforementioned indications, and provide an up-to-date view on the clinical scope of cfDNA analysis in cancer. In addition, it will address the importance of method standardization. It will also highlight the current gaps in our knowledge concerning the biology of cfDNA, a factor which may represent a substantial hindrance to the rapid translation of basic research to routine clinical practice.

\* Corresponding author.

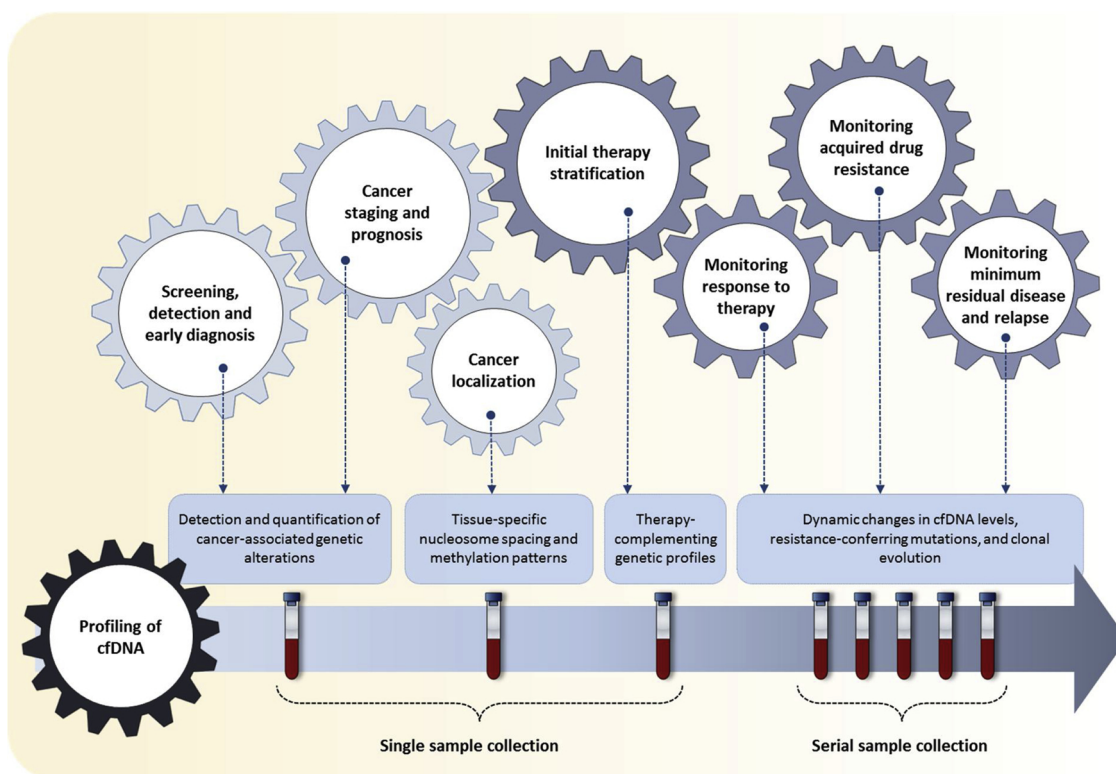
E-mail address: [holdenrieder@dhm.mhn.de](mailto:holdenrieder@dhm.mhn.de) (S. Holdenrieder).

<https://doi.org/10.1016/j.bdq.2019.100087>

Received 21 November 2018; Received in revised form 26 February 2019; Accepted 11 March 2019

Available online 18 March 2019

2214-7535/ © 2019 The Authors. Published by Elsevier GmbH. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).



**Fig. 1.** Potential applications of cell-free DNA in cancer management. Profiling of cancer-associated genetic alterations in cell-free DNA (cfDNA) may enable the large-scale screening of healthy or at-risk population groups for the early detection of multiple cancers. Furthermore, accumulating data demonstrates the potential clinical value of cfDNA analysis for the care of cancer patients at various stages of the disease. The level of cancer-associated mutations detected in cfDNA generally corresponds with tumor burden, and quantification thereof can serve as an indicator of disease stage and outcome. Tissue-specific nucleosome-spacing patterns and methylation signatures are encoded in cfDNA, and this information is useful for diagnosis, but can also be harnessed to pinpoint the location of tumors, especially those of unknown primary. Novel targeted therapies are only effective when specific pathways are altered in cancer cells. Detection of these mutations in cfDNA can, therefore, guide the selection of matched therapies. Compared to tumor biopsies, cfDNA has been shown to provide a better representation of the complete genetic landscape of a tumor. In addition, cfDNA offers the additional benefit of serial sampling, which allows the longitudinal assessment of dynamic changes in the concentration of cfDNA, the identification of acquired resistance-conferring mutations, and the tracking of clonal evolution. This makes it possible to monitor response of the cancer to therapy, detect the emergence of acquired resistance, as well as monitor and predict minimal residual disease and recurrence following surgery or therapy.

## 2. Origin, physical characteristics, and fluctuation of cfDNA

Despite the ubiquity of cfDNA in human body fluids, its molecular origin remains poorly understood. Excluding exogenous sources of cfDNA, several possible sources and cognate mechanisms have been proposed. First, early studies suggested that cfDNA enters circulation following the lysis of cells on the interface between a tumor and circulation. However, this was disproved after it was shown that the concentration of cfDNA in the blood of cancer patients is greater than could be accounted for by the mass of cells present [20,21]. Second, it was proposed that cfDNA may originate from the destruction of tumor micrometastases and circulating cancer cells. However, this was proven to be false since specific mutations in the cfDNA from CRC patients could not be correlated with mutations in the cells of the Ficoll layer where micrometastatic cells should be present [22,23]. There remain three more possible sources that may account for the occurrence of cfDNA, namely: apoptosis, necrosis, and active cellular secretion.

Sizing of cfDNA often generates a “ladder” pattern representing apoptotic fragmentation. The majority of DNA produced by apoptosis has a modal size of ~166 bp, which corresponds to 147 bp of DNA wrapped around a nucleosome plus the stretch of DNA on Histone H1 that links two nucleosome cores. However, depending on nuclease action apoptosis can also produce longer DNA fragments, representing di-, tri-, or poly-nucleosomes [24–29]. Using different experimental approaches, recent studies have demonstrated smaller fragment sizes (as short as 90 bp) for tumor-derived cfDNA compared to wild-type cfDNA

[30–34]. The cause of this shortening is not fully understood, but possible explanations include (i) differences in nucleosome wrapping or the mode of nuclease action between different tissue types, (ii) differential cfDNA size recovery of different extraction methods, or (iii) biases introduced during commonly used single-stranded DNA library preparation methods (reviewed in [7]). In contrast to apoptosis, cfDNA fragments larger than 10,000 bp are often observed in cancer patients, indicating an origin from necrosis [29,35–38]. An argument against necrosis as the primary pathway for cfDNA release comes from observations that cfDNA levels decrease by approximately 90% following radiation therapy. If necrosis ensued, a surge and not a decline in cfDNA concentration would be expected [21,39]. However, it can also be argued that the reduction of cfDNA levels may be due to radiation-induced inhibition of cfDNA release pathways in healthy cells [40]. It may be of interest to note that other forms of cell death (e.g., pyroptosis, autophagy, phagocytosis, mitotic catastrophe, and NETosis) can also serve as sources for cfDNA. For example, NETosis results in the decondensation of chromatin, cell lysis, disruption of nuclear membranes, and finally the liberation of neutrophil extracellular DNA traps (NETs), which is composed of extracellular DNA fibers and defense-related substances and invading microorganisms [41]. However, the characteristics of the cfDNA derived from these mechanisms have not yet been elucidated [42].

Mitochondrial DNA (mtDNA) can also be released into blood circulation during the above-mentioned cellular clearance or repair processes [43–46]. While the structure of cell-free mtDNA (cf-mtDNA) is

still not well characterized, it can be present either in naked form or associated with internal and external mitochondrial membrane fragments [47]. Unlike the typical three-mode size signature of autosomal cfDNA produced by apoptosis, cf-mtDNA has been shown to be more fragmented; typically ranging between 40 and 300 bp [33,45,48–50]. This size range can be ascribed to the absence of nucleosome-associated histone proteins (i.e., the absence of higher order packaging), which render “naked” cf-mtDNA exposed to enzymatic cleavage.

In contrast to cellular destruction, early studies indicated that a significant fraction of cfDNA is derived from active cellular secretions [21,51–55]. Recent *in vitro* cell culture studies have demonstrated the presence of cfDNA in culture medium at levels which do not correlate with the processes of apoptosis, necrosis, or DNA replication. Moreover, these cfDNA fragments are in the range of 1000–3000 bp, which is a size not typically associated with apoptosis or necrosis [56–58]. This provides further evidence for the active release of cfDNA. While the exact mechanisms involved in the active release of cfDNA remain unclear, it is possible that cfDNA is released as a consequence of genomic instability [59]. In keeping with this, a recent paper reported for the first time the presence of extrachromosomal circular DNA in human blood [60]. This species of DNA molecules is typically extruded from the nucleus as double minutes, which are secondary nuclear structures that form as a result of DNA amplification induced by chromosomal instability [61,62]. This finding has been corroborated by another research group that demonstrated the presence of a heterogeneous population of extrachromosomal circular DNA, ranging between 30 and 20,000 bp, in human blood [63]. Another form of active or regulated release includes DNA fragments associated with extracellular vesicles, such as exosomes. These vesicles range in size between 30 and 100 nm and carry cfDNA fragments that range between 150 and 6000 bp [64–66], however, the exact ratio of cfDNA bound to the exterior surface vs those localized in the interior are yet to be determined. Nevertheless, the commonly held assumption that apoptosis is the main origin and most relevant fraction of cfDNA in human blood may be restrictive and should be reconsidered.

There is undoubtedly a great dearth of knowledge surrounding the origin and molecular properties of cfDNA. Although a large fraction of cfDNA has been shown to originate from apoptosis, it is becoming clear that cfDNA is released into circulation by multiple mechanisms. Moreover, each of these mechanisms are modulated by a wide range of biological and environmental factors (many of which are inextricably linked by a complex interplay of cellular and physiological interactions) that are virtually unique to each individual. Variables may include age, gender, ethnicity, body-mass-index, organ health, smoking, physical activity, diet, glucose levels, oxidative stress, medication status, infections, menstruation, and pregnancy [42,67,68]. Besides the mechanism of release, the characteristics of cfDNA are greatly influenced by the rate of its clearance. Studies have estimated the half-life of cfDNA in circulation between 16 min and 2.5 h [69–71], but this requires further confirmation in various settings (e.g., healthy vs diseased; before surgery vs after surgery; at rest vs after exercise). Although the mechanisms by which cfDNA is cleared from blood remains poorly understood, it may be achieved by DNase I activity [72,73], renal excretion into the urine [74–76], and uptake by the liver and spleen followed by macrophagic degradation [77,78]. Clearance by these mechanisms may be further influenced by the association of cfDNA with protein complexes, extracellular vesicles, and the binding of individual cfDNA fragments to several serum proteins (e.g., Albumin, transferrin, fibrin, fibrinogen, prothrombin, globulins, C-reactive protein, HDL, Ago2, and SAA) (reviewed in [67]). Moreover, cfDNA can be recognized by various cell-surface DNA-binding proteins and be transported into cells for possible degradation to mononucleotides or for transportation into the nucleus. Interestingly, the binding of cfDNA to cell-surface receptors is dependent on pH and temperature, and can be inhibited by various substances [79]. Therefore, the rate of cfDNA uptake by different cells may also affect the rate of its clearance.

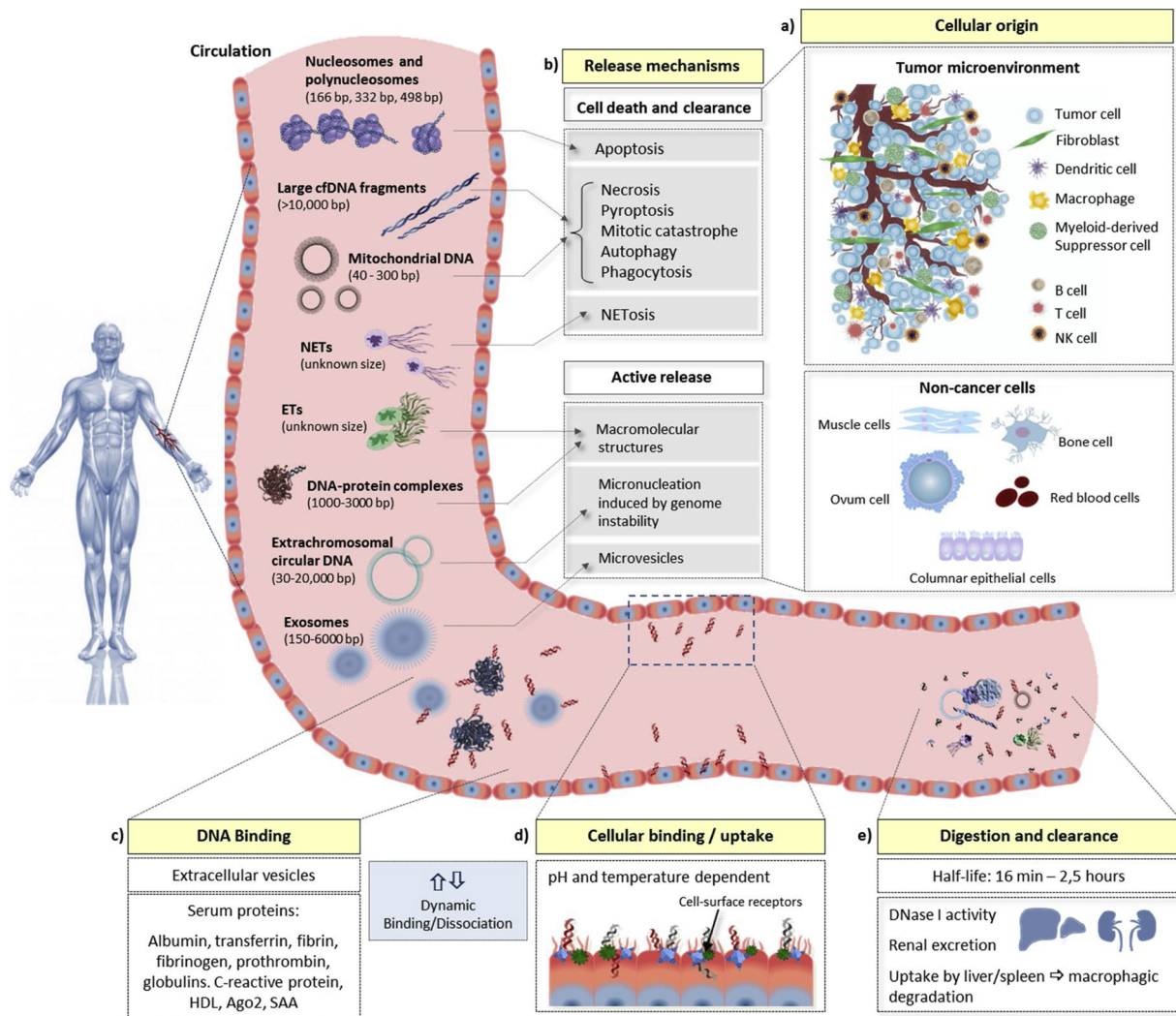
Furthermore, in cancer cfDNA does not originate only from tumor cells. It also originates from cells of the tumor microenvironment, as well as other non-cancer cells (e.g., endothelial and immune cells) from various parts of the body [67]. It seems to be the case that all cells are capable of, and are likely, continuously releasing cell-specific DNA into the extracellular environment (it has yet to be found absent in *in vitro* studies). An important point in this regard is that the concentration of cfDNA from tumor microenvironment cells and other “healthy” cells, the concentration of tumor-derived DNA, and the abundance of genetic alterations in tumors varies significantly between individuals (reviewed in [67]). For diagnosis it may, therefore, be sufficient to look only at apoptosis-derived cfDNA originating from cancer cells. However, to better estimate tumor dynamics, mutation load, progression or assess the efficacy of treatment, the best approach may be to determine the proportion of aberrant vs wild-type DNA, including all forms of cfDNA.

A related issue of possible concern is the phenomenon of genetic mosaicism, a term used to describe the presence of two or more cell populations with different genotypes within one individual [80,81]. It is generally assumed that all of the somatic cells in a higher organism contain an exact replica of the entire genetic code, and that it is subject to change only by virtue of random mutations due to replication errors and inevitable damage to the genome (reviewed in [82]). However, there is accumulating evidence that the genome is continuously formatted by both intentional and incidental rearrangements, including duplications, deletions and insertions in both the germline and somatic cells, in both healthy and diseased states. This is possible because of compartmentalization, which creates a unique environmental niche for individual organs, tissues and cells, allowing adaptation/diversification according to localized conditions. Genetic diversification is achieved by mechanisms such as non-allelic homologous recombination (NAHR), non-homologous end-joining (NHEJ), and Fork Stalling and Template Switching (FoSTeS), which have been associated with transposon mobilization and insertion, drug-induced gene duplication, retroviral mutagenesis, and the action of mini-satellites and small RNA molecules (for a concise review, refer to [83]). The implications of genetic mosaicism for downstream cfDNA analyses are further discussed in Section 5.1.1.

For these reasons, the aggregate cfDNA profile present in a single blood sample comprises a muddled blend of both “wild-type” and genetically and epigenetically altered DNA fragments released by various cells from different tissues and organs by different mechanisms under different environmental pressures. Not only does this large population of background DNA make it very difficult to detect cancer-associated alterations and make comparisons between individuals, it also significantly complicates the elucidation of the biological properties and functions of cfDNA *in vivo*. The magnitude of this issue is more elegantly argued in recent review articles in which the heterogeneity of blood samples is illustrated by highlighting numerous putative sources and causes that result in the presence of cfDNA in the extracellular environment [42,67]. Although it is not obvious how these issues can be solved, it is clear that basic research into the fundamental biology of cfDNA is needed to interpret the associations between the characteristics of cfDNA and the clinical manifestations of cancer, and that these associations should be given consideration during clinical validation experiments. The possible ways in which preanalytical processing of DNA can be affected by its diverse characteristics are briefly discussed in Section 4. A summary of the putative sources, corresponding characteristics, and dynamics of cfDNA in circulation is outlined in Fig. 2.

### 3. Roles of cfDNA in the pathology of cancer

Early studies have demonstrated that cfDNA fragments are able to enter neighboring or distal cells [84–88], and accumulating evidence suggests that cfDNA is capable of altering the biology of recipient cells. This phenomenon has been implicated in the oncogenic transformation of normal cells and the development of metastases [85,89–93]. While it is currently not understood exactly how cfDNA elicits these effects,



**Fig. 2.** Characteristics of cell-free DNA in the human body. (a) In cancer patients, cell-free DNA (cfDNA) originates from multiple sources, including cancer cells, cells from the tumor microenvironment, and non-cancer cells from other regions of the body (e.g., hematopoietic stem cells, muscle cells, and epithelial cells). (b) DNA can be liberated from these cells via different mechanisms, most prominently apoptosis, necrosis, and active secretion, although other forms of cell death and clearance may contribute. The physical characteristics of the cfDNA produced by these different mechanisms vary considerably. Apoptosis causes the systematic cleavage of chromosomal DNA into multiples of 160–180 bp stretches, resulting in the extracellular presence of mono- (~166 bp) and poly-nucleosomes (332 bp, 498 bp). Necrosis results in nuclear chromatin clumping and non-specific digestion, producing cfDNA fragments that are typically larger than 10,000 bp. Mitochondrial DNA can also be released by these mechanisms and typically range between 40 and 300 bp. CfDNA derived from active cellular secretions have been shown to range between 1000 and 3000 bp, while cfDNA originating from extrachromosomal circular DNA ranges between 30 and 20,000 bp. Exosomes, which also enter circulation via regulated release, carry DNA ranging between 150 and 6000 bp. In addition, recent evidence has demonstrated the presence of DNA fragments smaller than 166 bp in blood, which might represent degraded products of any of the aforementioned cfDNA types. Extracellular levels of cfDNA is, therefore, highly dependent on the rate of its release from cells. However, once present in circulation, cfDNA levels are further influenced by its (c) dynamic association and disassociation with extracellular vesicles and several serum proteins, (d) rate of binding, dissociation and internalization by cells, which is dependent on pH, temperature, and can be inhibited by certain substances (such as heparin), and (e) the rate of digestion or clearance, including the activity of DNase I, renal excretion into urine, and uptake by the liver and spleen.

there are some interesting lines of evidence that can be considered. First, it is possible that cfDNA promotes metastasis in recipient cells by inducing the overexpression of several pro-metastatic genes through the TLR9/MYD88 independent pathway [94–96]. Second, we have recently shown that the majority of cfDNA released by cultured human bone osteosarcoma (143B) cells consist of repetitive DNA and is enriched in specific transposons (e.g., LINE-1) that are currently active in the human genome. Considering its inherent mobility, the lateral transfer and aberrant insertion of transposons into a host genome can disrupt coding regions, splice signals, and activate oncogenes. For example, it has been demonstrated that the insertion of a single hot L1 transposon into a tumor suppressor gene can initiate tumor formation [97]. Last, it is also possible that these effects are facilitated by the cellular uptake of

exosomes. The production of exosomes has been shown to be markedly elevated in cancer [98], which can affect recipient cells in a variety of ways, e.g., development of acquired resistance to chemotherapeutic agents [99]. In contrast to the above observations, one *in vitro* study suggests that the lateral transfer of cfDNA derived from healthy cells can potentially halt the proliferation of cancer cells [100]. However, this requires further experimental verification.

The lateral transfer of cfDNA is also involved in the augmented resistance of cancer cells against radiation- and chemotherapy [101–103]. During radiation therapy oxidized DNA is generated. When oxidized cfDNA is then released and assimilated by neighboring cells it induces DNA breaks, stimulates the biogenesis of reactive oxygen species, and activates DNA damage response pathways. In support of this,

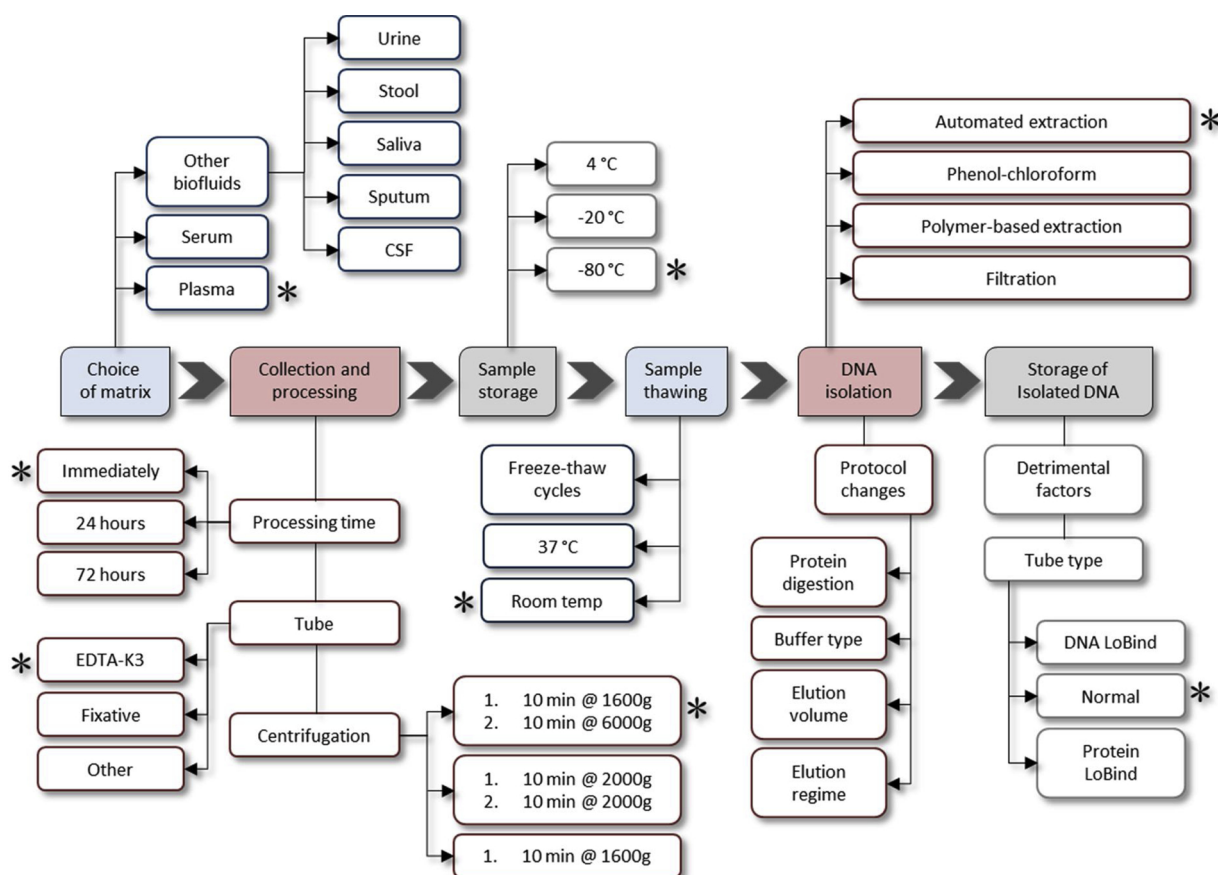


Fig. 3. A robust preanalytical workflow for tumor-derived cell-free DNA analysis. This workflow was carefully formulated by selecting the most optimal preanalytical steps (indicated by the asterisks) from a wide range of alternative steps that are reported in the literature.

It has recently been demonstrated that cell-free chromatin derived from dying cancer cells are laterally transferred to both bystander cells and cells of distant organs, followed by genomic incorporation and the induction of significant DNA damage and inflammation [104]. This can impair genome stability and undermine the homeostatic capacity of the tumor microenvironment, which are potent stimuli for oncogenic transformation. Although there is currently no clear answer on how cfDNA is integrated into the human genome, research suggests that it may occur through non-homologous end-joining (NHEJ) [90,105,106].

Although the role of cfDNA in human biology remains largely unexplored, the above-mentioned studies suggest that further inquiry into the functions of cfDNA may provide a new framework for a deeper understanding of some of the molecular mechanisms that underlie the pathology of cancer.

#### 4. Considerations for optimizing pre-analytical procedures

As discussed in Section 2 and illustrated in Fig. 2, cfDNA does not originate solely from tumor cells. It also comes from tumor microenvironment cells and other non-cancer cells from various parts of the body. CfDNA can then be released from these cells via different mechanisms, resulting in the presence of a cfDNA population with diverse physical properties in circulation. It logically follows that these cfDNA molecules are differently affected by various preanalytical steps, especially when different sample processing protocols [107], storage conditions [108], and extraction methods are used. For example, cfDNA obtained with non-hybridization-based extraction methods, which do not filter and capture DNA fragments of a specific size, have revealed very different size profiles [109–111]. Moreover, in our lab we have compared two commonly used automated cfDNA extraction methods (unpublished results). Using the same samples in each case, one method

recovered one cfDNA population (166 bp), while the other method recovered two cfDNA populations (166 bp, and 30–100 bp, respectively). Also using plasma samples, a manual extraction kit has been shown to recover one cfDNA population of ~166 bp, and one population in the range of 1000–3000 bp, while automated extraction methods were typically poor at isolating these larger DNA fragments (see application note: MagMAX Cell-Free DNA Isolation Kit). This is probably because automated methods have been tailored to extract DNA with a modal size of 166 bp [111,112], based on the assumption that any DNA fragments larger than this is a byproduct of germline DNA contamination. While studies suggest that selection of short fragments may enrich for tumor-derived cfDNA and improve error correction during next generation sequencing, it is important not to undermine the potential value of other cfDNA fragments. As mentioned in Section 2, *in vitro* studies suggest that these larger fragments, or at least a fraction of them, may originate from an active release mechanism [57–59]. In addition, a recent study has demonstrated that these larger fragments harbor tumor-associated mutations that were also detected in the short cfDNA fragments [113]. It is thus very important to discriminate between these different cfDNA species in a sample before analysis. Consolidating data from analyses of cfDNA obtained by different extraction methods with corresponding histopathological data will likely provide new insights into the biological characteristics of cfDNA.

Aside from recovering different sizes of cfDNA, different extraction methods also yield different concentrations of cfDNA [109,110,114]. The concentration of cfDNA molecules, particularly those of tumor origin, is often very low, especially in certain cancer types and low burden disease. Therefore, any loss of sampled material will reduce the sensitivity of downstream molecular analyses. In this regard, it may be useful to select a cfDNA extraction method that delivers the highest yield, or to perform size-selection following extraction. Furthermore,

cfDNA yield can be maximized by selecting optimal preanalytical steps. An in-depth discussion of all the preanalytical variables that may affect downstream cfDNA analyses has been undertaken in other publications [115–117]. Here, we will give a brief summary of a robust preanalytical workflow that may be considered (Fig. 3):

The primary matrix used for cfDNA analysis is blood plasma. However, it should be noted that cfDNA isolated from other biofluids which are in closer proximity to the tumor may provide a better representation of disease status in particular regions of interest (see Section 5.1.1). Blood should be collected in EDTA-containing tubes (e.g., EDTA-K3) to prevent coagulation; heparin and citrate could inhibit PCR and should be avoided. To prevent lysis of peripheral blood cells and release of germline DNA, blood samples must be processed immediately after venipuncture by two rounds of refrigerated centrifugation; first at a slow speed to remove cells (e.g., 10 min at 1600 x g), followed by a high-speed centrifugation of the carefully collected supernatant (e.g., 10 min at 6000 x g). Processing can be delayed to a maximum of 4–6 h. If this is not possible, tubes containing fixative agents that conserve membrane stability are recommended. These tubes will prevent cell lysis for several days, including during shipping. However, agitation of samples during a delay in processing should be minimized. Following processing, the samples can be aliquoted and stored at -80 °C. Before cfDNA extraction plasma samples can be thawed at room temperature, however since there are no recommendations on plasma thawing temperature in the literature it is not yet clear if this is optimal. Freeze-thaw cycles should be avoided. For maximum yield of tumor-derived cfDNA, a method tailored for the selective extraction of small cfDNA fragments (i.e. < 166 bp) is recommended, or size selection for short DNA fragments can be performed. To avoid loss of isolated cfDNA, samples should be stored in tubes with a low DNA-adsorption quality.

It is important to note that there is currently no consensus among researchers regarding these preanalytical steps [118–120]. This is a major source of conflicting data in the literature. Therefore, the importance of establishing a robust preanalytical workflow for cfDNA analysis, in conjunction with universal equivalence of procedures, cannot be overstated.

## 5. Utility of cfDNA in clinical oncology

### 5.1. Early diagnosis of cancer, disease staging and prognosis

#### 5.1.1. Diagnosis

The arrival of cancer immunotherapy has revolutionized the treatment of patients with a variety of tumors, significantly prolonging the survival of many patients; especially those with advanced cancers (see Section 5.2). However, the success of these therapies can be improved greatly by the implementation of clinical modalities that enable large asymptomatic population screening, early diagnosis, and localized treatment of a broad spectrum of both common cancers and deadly tumor types [121,122]. However, despite its promise and years of research on the topic, early stage diagnosis of cancer remains an extremely challenging task. The small size of early-stage solid tumors makes it difficult to distinguish it from normal anatomic and biochemical variation in a non-invasive manner. This uncertainty inevitably results in the construction of detection algorithms that result either in false-negative results, or introduces the risk of high false-positive rates, over-diagnosis, and, consequently, overtreatment [123,124]. Recent research suggests that technological advances in the analysis of cfDNA may overcome these challenges, and possibly expedite the early, sensitive, and accurate diagnosis of cancer [125]. Numerous studies have demonstrated the possibility that cancer-associated mutations can be detected in cfDNA in (i) early-stage disease [126–131], (ii) before the presence of symptoms [132–135], and (iii) up to 2 years before cancer diagnosis [136,137]. However, while these studies demonstrate the potential of cfDNA as a marker for the early detection and diagnosis of cancer, there are significant challenges that need to be overcome before it can be applied in a clinical setting:

**5.1.1.1. Technological and analytical limitations.** The average mutant allele fraction (MAF) is a metric that denotes the ratio between the amounts of mutant alleles versus wild-type alleles in a sample, and generally correlates with tumor burden. The sensitivity of the technologies used to detect tumor-derived cfDNA is thus aptly expressed by the range of the MAF that they are capable of detecting. The MAF detection limits of traditional quantitative PCR (qPCR) methods range between 10%–20%. However, many PCR-based variations have been developed to increase sensitivity, such as allele-specific amplification (AS-PCR) [138], allele-specific non-extendable primer blocker PCR (AS-NEPB-PCR) [139], peptide nucleic acid-locked nucleic acid (PNA-LNA) PCR clamp [140], and co-amplification at lower denaturation temperature (COLD-PCR) [141]. COLD-PCR, for example, has been shown to be able to detect a MAF of 0.1% [141]. In contrast, digital PCR (dPCR) methods, which include microfluidic-based droplet digital PCR (ddPCR) and BEAMing (beads, emulsions, amplification and magnetics) [77,142–148], are able to quantify cfDNA with extreme sensitivity (0.001%–0.05% MAF); in some cases only one mutant cfDNA fragment can be detected in 5 mL of sample [149]. This is optimal for inexpensive absolute quantification of cfDNA and for the detection of cancer hot-spot mutations. However, as these assays are based on differences in the binding affinities of mutant and wild-type alleles and generally require primers or probes that target specific mutations or loci, multiplexing capacity is limited.

Genome-wide sequencing methods (e.g., Plasma-Seq [150], PARE [151], FAST-SeqS [152], and mFAST-SeqS [153]) requires a relatively large amount of cfDNA for detection (5–10% MAF), and can be used for detecting cancer-specific copy-number alterations. Targeted sequencing approaches using hybrid capture (e.g., exome sequencing [154], CAPP-Seq [130,155], and digital sequencing [156]) or PCR amplicons (e.g., TAm-Seq [157], Enhanced TAm-Seq [158], NG-TAS [159], and Safe-SeqS [160]) are more sensitive, and can be used to query a larger number of loci. Whole exome (~50 Mb) sequencing has a detection limit of approximately 5% MAF [154], while smaller off-the-shelf multiplex gene sequencing panels are at least 5 times more sensitive (1% MAF) [161–163]. Recent studies have shown that the MAF detection limit for targeted approaches can be reduced to 0.1% by minimizing the background error rates of sequencing, or by increasing the signal-to-noise ratio of samples [155,158,160]. This can be achieved by molecular barcoding, which involves the allocation of a unique identifier (UID) to each DNA template followed by the amplification of each uniquely tagged template to create UID families, or by performing independent assays on multiple aliquots of the total amount of cfDNA that was recovered from plasma [157,164]. Using cancer- or patient-specific multiplexed panels in conjunction with targeted sequencing methods, MAFs as low as 0.01%–0.5% can be detected [130,157,158]. For example, the CAPP-Seq assay demonstrated a MAF detection limit of 0.02% and correctly identified mutations in 100% of stage II or greater NSCLC patients, but only detected mutations in 50% of stage I patients [130]. By combining barcoding with an integrated digital error suppression computational tool for the *in silico* elimination of highly stereotypical sequencing errors, the sensitivity of the CAPP-Seq method was increased approximately 15-fold (~0.004% MAF) [155].

A realistic goal for the early diagnosis of cancer is the detection of a tumor with a diameter of approximately 5 mm (0.07 cm<sup>3</sup>), as this represents an asymptomatic stage, localized, less likely to progress, and curable. However, it seems unlikely that this goal can be met with the current cfDNA analysis strategies described above. Based on recent calculations, a tumor with a 5 mm diameter corresponds to a ratio of tumor-derived cfDNA to normal cfDNA of less than 1–100,000 copies (MAF of 0.001%) [165]. Based on observations that 1 mL of plasma from healthy subjects contains approximately 3000 whole-genome equivalents [34,166], the total amount of whole-genome equivalents in 3 L of plasma, which represents the whole blood circulation, should contain about 9,000,000 copies. Thus, in the entire cfDNA population, only one cancer genome will originate from a 1 mm diameter tumor,

making the chances of extracting one tumor-derived cfDNA fragment from a 10 mL blood sample very low. Indeed, this data suggests that current methods are only able to reliably detect tumors with a diameter greater than 1 cm ( $0.5 \text{ cm}^3$ ) [165]. Although this is also considered an early stage tumor, it usually corresponds with patients that already show clinical signs and symptoms of cancer. Moreover, tumors of this size can be identified through imaging [124]. It should be noted that calculations for estimating the concentration of tumor-derived cfDNA, such as those above, are often based only on extrapolation from tumor size. However, the amount of cfDNA released by tumors is not only dependent on size, but also on turnover activity, proliferation rate, vascularization, and perfusion. Therefore, different tumor types of the same size can release different amounts of cfDNA. Moreover, the final concentration of cfDNA in circulation is influenced by the rate of its degradation and clearance (see Section 2).

The studies discussed above suggest that the maximum sensitivity for cfDNA analyses is limited for localized cancers. However, a recent study has shown that the sensitivity for the detection of pancreatic cancers can be increased by querying a combination of four protein biomarkers and mutated KRAS cfDNA in parallel [167]. Subsequently, the authors expanded this approach by assessing a panel of 61 amplicons (16 genes) and 39 proteins (using PCR- and immune-assays, respectively) to detect eight different solid tumors (ovarian, colorectal, pancreas, liver, stomach, esophageal, lung, and breast cancer) before the emergence of distant metastases [126]. For five of the cancers, the sensitivities ranged between 69% and 98%, while the specificity was greater than 99%. While these are still not optimum values for early diagnosis, and were not confirmed in a screening cohort, this study provides a proof-of-concept and has laid the experimental groundwork for the further development of multi-analyte blood tests to screen for various cancers at once. Besides proteins, it is likely that other circulating biomarkers, such as extracellular vesicles [168], mitochondrial DNA [43,50], miRNAs [169], mRNA transcripts [170], and metabolites [171] could be combined in a similar fashion to increase the sensitivity of tests for cancer detection and localization.

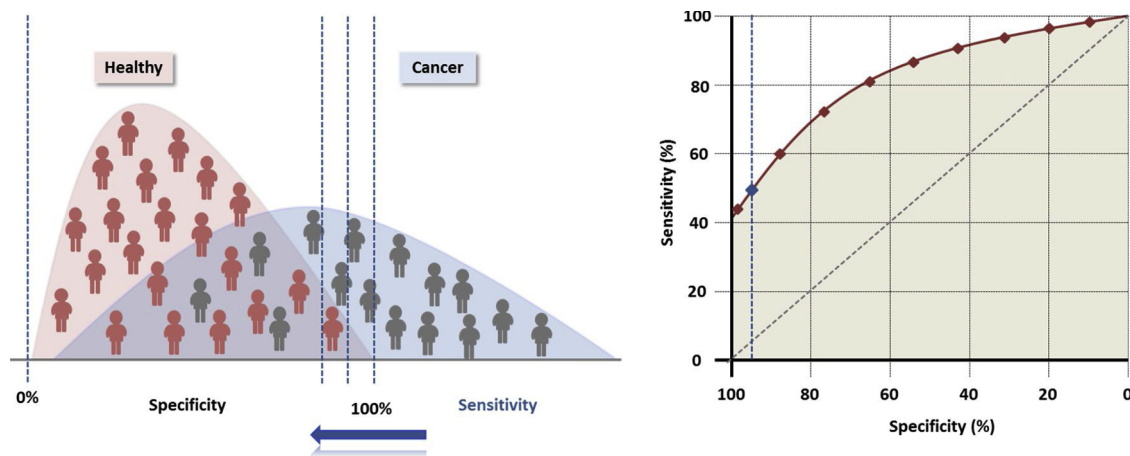
Another approach that could potentially overcome the technical limitations of mutation-based cfDNA detection methods, is to query genomic features of cfDNA that are altered on a larger scale in cancer, such as differential methylation patterns [172–174] and post-translational histone modifications [175,176]. For example, methylation signatures differ among tissues and cell types [177,178], and recent studies have shown that this information is encoded in cfDNA. Harnessing this data could increase the probability of detecting cancer, aid in pinpointing the location of a tumor, and may be useful for identifying cancers of unknown primary origin. Several studies have detected organ-specific and tumor-associated DNA methylation signatures in plasma by focusing on signatures of one tissue [135,179–181]. In one study, a method for deconvoluting genome-wide bisulfite sequencing data was developed to produce ‘tissue methylation maps’ of cfDNA for pregnant women, cancer patients, transplantation patients, and healthy subjects, revealing percentage contributions by different tissues [135]. The clinical potential of this method has been substantiated by other studies [182–184], and has provided some insight regarding the composition, degradation and variation of cfDNA in urine [185]. However, some major drawbacks of bisulfite sequencing is the high level (84–96%) of degradation of input DNA during the bisulfite conversion process [186], high costs, and the recovery of limited information due to the low abundance of CpGs in the genome. This can be overcome by using alternative and bisulfite-free sequencing methods, such as hydroxymethylcytosine sequencing [187] or methylated DNA immunoprecipitation-sequencing (MeDIP-seq) [188]. An important paper has recently demonstrated that an optimized protocol for MeDIP-seq can be used to detect tumor-derived cfDNA in early-stage pancreatic cancer, as well as classify several tumor types [172]. Although these approaches are promising, both the accuracy of the method for predicting the tissue of origin and its specificity and sensitivity as a

diagnostic test is currently too low to be used in most clinical settings. This can be improved by applying the method to a large cohort of subjects, or by the further development of models for predicting differentially methylated regions by using machine learning and both biological and simulated training data [189,190]. Similar to methylation patterns, nucleosome occupancy patterns also differ between tissues, have also been found to be encoded in cfDNA, and can also potentially be used to classify tumors of unknown primary or cancers that require invasive biopsies for definitive diagnosis [191,192]. However, all of the above-mentioned approaches need large-scale confirmation before it will become useful in clinical practice.

**5.1.1.2. Sampling noise.** Another major factor that limits detection sensitivity is sampling noise due to limited blood availability. There are ways in which this can be improved. First, larger volumes of plasma can be collected using plasmapheresis, or the yield of cfDNA can be increased by implanted devices that selectively capture cfDNA. Second, preanalytical steps relating to the collection and processing of blood samples can be optimized (see Section 4). Moreover, since recent studies have demonstrated a shorter size for tumor-derived cfDNA than normal cfDNA, experimental or *in silico* selection of shorter fragments should enhance the sensitivity of downstream analyses [30–34]. Indeed, a recent landmark study of 200 cancer patients has shown that selection of these shorter fragments improved the detection of tumor-derived cfDNA by more than two-fold in approximately 95% of cases, and more than four-fold in approximately 10% of cases [30]. Third, extraction of cfDNA from other biofluids that are in closer proximity to the tumor of interest may provide a much higher ratio of tumor-derived cfDNA to normal DNA in comparison with plasma [193], as shown in studies of CRC (stool) [194], oral cancer (saliva) [195], lung cancer (sputum) [137,196], bladder cancer (urine) [142], ovarian and endometrial cancers (cervical smears [131] and uterine lavage [197]), brain tumors (cerebrospinal fluid) [198–201], and esophageal cancers (esophageal brushings) [202].

**5.1.1.3. Issues concerning diagnostic specificity and sensitivity.** The specificity and sensitivity of cfDNA tests for the screening of asymptomatic individuals is crucial. Sensitivity indicates the proportion of correctly identified positive results in a patient group, while specificity describes the proportion of correctly identified negative results. The diagnostic performance of cfDNA is best shown by the complete profile of sensitivity and specificity using ROC curves (Fig. 4). This figure gives the sensitivity and specificity at all possible cut-off points, and is particularly useful for comparing the performance of a wide range of biomarkers with each other. Meaningful measures include: (i) the area under the curve (AUC); which should ideally be close to 1.0 and indicates no discriminative potential of the marker if it is close to 0.5; (ii) the sensitivity at a defined specificity; and (iii) an optimized sensitivity-specificity combination that is reflected in the figure by the point closest to the upper-left corner. Careful consideration should be given to the selection of groups that will be compared using this approach. Best curves will result if patients with advanced cancer disease are compared with young healthy individuals. However, it is clinically more relevant to distinguish between equally aged individuals (with suspicious symptoms) that suffer from an early cancer or a benign disease. In this case the curves will be less optimistic [203].

It should be pointed out that positive predictive values (PPV) and negative predictive values (NPV) are more informative than sensitivity and specificity in terms of screening. While PPV indicates the probability of disease if the value is positive, NPV gives the probability of being disease-free if the value is negative. Since the frequency of cancer cases is often very low, PPV may be low even if the sensitivity and specificity is higher than 90%. For example, if a population of 100,000 individuals is screened for a cancer with a prevalence of 1 in 4000 (such as ovarian or pancreatic cancer), approximately 25 people will be affected, while the remaining will be unaffected. Even if the screening test



**Fig. 4.** The value distributions of most cancer biomarkers show an overlap of cancer patients and healthy individuals. The diagnostic performance of a biomarker is best illustrated by ROC curves. To establish this graph, the portion of correctly identified negative controls (specificity) and correctly identified positive cancer patients (sensitivity) are identified for all possible cut-off points (decreasing stepwise from 100% specificity). The area under the curve (AUC) and the sensitivity at a fixed specificity (e.g. 95%) are most informative for the comparison of diagnostic markers. As control groups, healthy individuals and patients with differential diagnostically relevant benign diseases are considered.

correctly identifies all affected individuals (thus 100% sensitivity), 99% specificity will yield close to 1000 false-positive results and a PPV of only 2.4%. Even if the specificity can be increased to 99.9%, the test will still yield 100 false-positive results and a PPV of 20% [165]. This can lead to significant overdiagnosis, expensive tests and undue overtreatment. Although these aspects seem obvious, some of the points are often ignored in cfDNA studies, i.e. either inappropriate controls are chosen or numbers of patients, particularly in monitoring studies, are too small to allow general and robust conclusions. Strategies for improving this includes the screening of high-risk groups (e.g., screening smokers for lung cancer or elderly people for CRC), and combining methods (e.g., using biomarkers for prescreening to stratify individuals at higher cancer risk for subsequent colonoscopy or low-dose CT scan).

**5.1.1.4. Ambiguity surrounding cancer-associated mutations.** Although the sensitivity of cfDNA tests can be improved through technical advances [155], other biological factors may become restrictive. As discussed above, confident cancer-detection using cfDNA necessitates a high PPV of detected mutations for cancer. However, several recent studies have reported the presence of cancer-associated genomic alterations in both tissue biopsies and cfDNA from healthy individuals without cancer [136,155,204–211]. Some key insights in this regard came from the analysis of whole exome-sequencing data of the DNA isolated from the peripheral-blood cells of 12,380 individuals. Clonal hematopoiesis was observed in 10% of patients over the age of 65 years, but only in 1% of patients under the age of 50 years. Moreover, it was shown that a portion of mutated genes in patients with myeloid cancers are also mutated in healthy individuals that never developed cancer [211]. Therefore, if non-tumorigenic clones that bear mutations that are typically associated with cancer expand enough to contribute a substantial amount of DNA to the total cfDNA pool of healthy individuals it could induce biological noise, which poses a significant challenge to the development of cfDNA screening tests. To gain a better understanding of the biological and clinical implications of this finding, the clinical outcomes of healthy patients that present with cancer-associated mutant cfDNA should be assessed. In line with these observations, genetic mosaicism presents another obstacle to the development of highly specific cancer tests (Section 2). Our understanding of this phenomenon will be greatly improved by the completion of large-scale single-cell sequencing projects [212,213]. Apart from the abovementioned, other factors can also affect the outcome of population screening programs. For example, screening programs are often not practical for the detection of rapid-growing

tumors, since patients with negative test results in the first round may test positive with disseminated disease in the next screening round. Contrarily, early detection, overdiagnosis and overtreatment of slow-growing tumors that may remain idle for decades can result in undue harm to patients [165].

In summary, theoretical and empirical findings suggest a limited potential of current cfDNA analysis strategies for early-stage cancer-diagnostics. While the sensitivity of these methods can be moderately enhanced by incremental technological improvements, it is becoming clear that major improvements will likely be driven by biologic based discoveries. Although currently unsuitable for cancer screening, these methods demonstrate sufficient sensitivity for later-disease staging, assessing tumor burden, and prognosis. Furthermore, it holds considerable promise for guiding therapy selection, evaluating response to therapy, monitoring minimum residual disease, and predicting recurrence (Fig. 1). These potential applications of cfDNA are explored in the following sections.

#### 5.1.2. Tumor staging and prognosis

In the past three decades, numerous studies have indicated that cfDNA levels are generally higher in cancer patients compared to healthy subjects (reviewed in [4]). Moreover, the concentration of cfDNA in circulation has been demonstrated to correlate with tumor size [214–218], disease stage [129,130], and metastatic burden [157]. In a landmark study of 640 patients with different cancer types at varying disease stages, it was shown that cfDNA levels are approximately 100 times higher in patients with stage IV disease compared to patients with stage I disease, providing a rough proportion for estimating tumor size from cfDNA concentration [129]. In stage I patients, less than 10 mutant cfDNA copies were present per 5 mL of plasma, while patients with advanced cancers had more than 100 mutant cfDNA copies per 5 mL of plasma. By correlating a higher frequency of ctDNA mutations or variant allele frequency with a greater tumor volume (determined by CT volumetric analysis) subsequent studies have developed a more sensitive metric for estimating tumor size [214,215]. In patients with high-grade serous ovarian cancer, mutant alleles in plasma increased by 0.08% and by 6 mutant copies per mL of plasma for every cubic centimeter of the tumor [215]. In NSCLC, a variant allele frequency of 0.1% was shown to correspond with a tumor size of approximately 10 cm<sup>3</sup> (27 mm in diameter), which was also the minimum size that permitted adequate sensitivity of ctDNA tests [214]. As discussed in Section 5.1.1 (a), this is considerably larger than an early stage asymptomatic tumor, making this technique unsuitable for



early stage detection. However, it may still be useful for disease staging, estimating tumor burden, and monitoring response of the tumor to treatment (see Section 5.2.2)

Besides these striking differences, tumor-derived cfDNA levels were found to vary considerably between patients with the same cancer type and disease stage [129]. Apart from differences in the metastatic spread or tumor burden, this variability can be partially explained by a wide range of biologic and physiologic factors that modulate the rate of cfDNA release and clearance, which are liable to significant intra- and inter-individual variation (see Section 4). Additional explanations may include (i) poor tumor vascularization, which could, on one hand, restrict DNA release, and on the other hand enhance cfDNA release as a consequence of hypoxia and cell death; (ii) histological differences, which could affect the dynamics and type of pathway for cell death and DNA release, (iii) perfusion, (iv) turnover activity, and (v) proliferation rate. In addition, the release of cfDNA from a tumor into circulation can be limited by biological compartments. For example, the concentration of mutant cfDNA molecules in the CSF of patients with primary brain tumors has been shown to be considerably higher than in plasma. While there is no direct experimental evidence, the authors have proposed that the movement of cfDNA into circulation is restricted by the blood-brain barrier [129,200].

The correlation between tumor-derived cfDNA levels and tumor stage/size indicates the potential prognostic value of cfDNA. This is exemplified by several studies that have demonstrated a relationship between cfDNA levels and the outcome and survival of cancer patients [219–225]. Moreover, some studies have shown that cfDNA often provides a better indication of outcome than other tumor markers [215,226–228]. This relationship is especially pronounced in patients that have received treatment. Higher levels of cfDNA following treatment often correlate with reduced survival rates and therapy resistance. In contrast, low cfDNA levels following therapy generally correspond with a positive response to treatment, which is often detected earlier than traditional detection methods. These studies are discussed in Section 5.2.2.

## 5.2. Guiding therapy selection

### 5.2.1. Initial therapy stratification

There is a strong clinical need for assessing the mutational status of cancer patients that are in line for receiving new targeted antibody or tyrosine kinase inhibitor (TKI) therapies, as these therapies are only effective when specific pathways are altered in cancer cells. Thus, tumor tissues from patients with lung cancer, melanoma, and colorectal cancer are regularly examined to stratify them for matched molecular therapies [17,229]:

In non-small cell lung carcinoma (NSCLC), the TKIs gefitinib and erlotinib directed against the intracellular part of the epidermal growth factor receptor (EGFR) showed only benefit in patients with an activating mutation (L858R or exon 19 deletion) in the EGFR gene, identifying EGFR tissue mutation analysis as stratification tool for TKI treatment of NSCLC patients [230–232]. However, although better than chemotherapy, TKI response rates in preselected patients were only around 70% in first-line and 50% in second-line treatment [233,234]. Recent whole genome analyses indicate that this is due considerable genetic heterogeneity, including spatial (either within a tumor or between primary tumor, lymph node and distant metastases) and temporal variability [235]. Furthermore, resistance to TKI by other mutations such as EGFR (T790M) that prevents erlotinib binding, or downstream mutations such as KRAS, PIK3CA, ALK and BRAF may also occur [17,234]. Identification of these mutations enables the use of alternative targeted drugs, such as crizotinib in case of the presence of an ALK-EML fusion gene [229,236]. Similarly to lung cancer, patients suffering from malignant melanoma will only benefit from inhibitor of the serine-threonine protein kinase BRAF therapy (vemurafenib or dabrafenib) in the presence of an activating BRAF mutation (V600E)

[237,238]. As resistance to vemurafenib will develop by activation of the MAP kinase pathway, MEK-inhibitors (trametinib) show some efficacy in these cases [239]. Finally, patients with colorectal cancer are unlikely to benefit from anti-EGFR antibody therapies (cetuximab or panitumumab) if mutations of the KRAS gene are present [240,241].

While pre-therapeutic assessment of the mutational status in tumor tissue is now a well-established auxiliary component in routine patient management, some inherent disadvantages of tumor biopsies have come to light in recent years. Progressive improvements of the techniques used in the analysis of biopsied material have illuminated the bias of single tumor snap-shots. Sequence analyses of biopsies taken from different portions of a primary tumor and its metastases in one patient have demonstrated markedly different mutational profiles, and revealed extensive intra- and intertumoral evolution [235]. Therefore, when tumor heterogeneity is unaccounted for it could prevent accurate classification, introducing bias into the selection and efficacy of personalized treatments. While it is now clear that a single biopsy is likely to underestimate the mutational landscape of a tumor, it is usually impractical to perform several biopsies on one patient, especially in cases when it is extremely invasive (e.g., glioblastoma and lymphoma). Procedural complications have been found to occur in up to 16% of needle biopsies [242], and often fail to obtain sufficient material for high quality genomic profiling [243]. In addition, invasive tissue biopsies are not supposed to be performed on patients with recurrent disease, metastatic disease, or multi-morbidity, and usually do not provide meaningful information in these cases. Recent studies, however, have shown that these limitations of tumor biopsies could potentially be overcome by mutational profiling of tumor-derived cfDNA. Since cfDNA is released from all cells in the human body it should allow the characterization of the complete genomic architecture of a patient's cancer. Indeed, in numerous studies spanning a wide range of cancers, profiling of tumor-derived cfDNA was able to identify mutations that were not detected by genotyping of the corresponding tissue [157,163,220,244–250].

However, before cfDNA can be utilized for this purpose with confidence, it should not only demonstrate a high concordance with results obtained from tissue biopsies, but also exhibit superior specificity and sensitivity. Using various methods, both high sensitivity/specificity and high concordance was demonstrated in several studies for various cancers and various mutations [70,77,138,221,250–260] (Table 1), and demonstrates higher sensitivity/specificity than tumor biopsies in most cases. In other studies, however, either a high concordance but low sensitivity or specificity [261–265], or both low concordance and sensitivity or specificity was obtained [260,266–271]. While it is clear that the extent to which cfDNA can be used for therapy stratification requires further investigation, these tests are gradually making their way into the clinic. For example, in 2016, the FDA approved the cobas EGFR Mutation Test v2 for routine clinical use. This test can be used to screen for 42 sensitizing mutations in the EGFR gene, which helps clinicians to identify NSCLC patients that would benefit from EGFR TKIs as initial therapy [8,9].

### 5.2.2. Monitoring response of cancer to therapy

The short half-life of cfDNA, together with the non-invasiveness of venipuncture relative to tissue biopsies and imaging makes cfDNA an ideal marker for monitoring the response of cancer to therapy, or to monitor disease burden after surgery. Numerous studies that have assessed patients during the course of treatment have demonstrated a correlation between lower cfDNA levels and a positive response to treatment in various cancers [273–276], including colorectal [70,277,278], ovarian [157,218], breast [227,279], NSCLC [130,280–283], melanoma [222,284–289], and brain tumors [290]. Similarly, higher levels of cfDNA in cancer patients generally correspond with poor response to treatment, therapy resistance, higher risk of relapse, and reduced survival rates [215,227,289–295]. In many of these studies, it was shown that cfDNA was able to monitor response to

**Table 1**  
Therapy resistance mutations detected in tumor-derived cell-free DNA vs. tissue.

Cancer type	Resistance Genes	Number of patients	Detection in cfDNA	Concordance	Reference
Colorectal cancer	KRAS	206	Sensitivity: 87.2% Specificity: 99.2%	95%	[272]
Colorectal cancer	RAS	329	Sensitivity: 92.9% Specificity: 97.7%	94.8%	[256]
Colorectal cancer	KRAS, BRAF, NRAS	100	Not reported	97%	[250]
Colorectal cancer	RAS (KRAS and NRAS)	98	Sensitivity: 90.4% Specificity: 93.5%	91.8%	[257]
Metastatic colorectal cancer	KRAS	95	Sensitivity: 92% Specificity: 98%	96%	[138]
Metastatic colorectal cancer	RAS	115	Sensitivity: 96.4% Specificity: 90%	93%	[258]
Metastatic colorectal cancer	BRAF	95	Sensitivity & Specificity: 100%	100%	[138]
Non-small-cell lung cancer	EGFR	18	Sensitivity: 81.8% Specificity: 85.7%	83.3%	[252]
Non-small-cell lung cancer	EGFR (T790 M)	61	Sensitivity: 100% Specificity: 98.3%	98.4%	[262]
Pancreatobiliary carcinomas	KRAS	26	Sensitivity: 92.3% Specificity: 100%	90.3%	[253]
Melanoma Gastric cancer	BRAF V600K	732	Sensitivity: 81% Specificity: 99%	Not reported	[221, 260]
Metastatic castration resistantprostate cancer (mCRPC)	Various	70	Not reported	Not reported	
		34		88.9%	[259]

therapy more sensitively than traditional or alternative detection methods [220,226–228,296,297]. For example, in a recent study on pancreatic cancer patients receiving gemcitabine-based chemotherapy, rapid and complete reduction of KRAS cfDNA (often within the first week) were mainly observed in patients with response to therapy, while non-responders presented either with continuously high KRAS cfDNA levels, moderate decreases, or re-increases after temporary decreases. For detection of progressive disease, the sensitivity (83%) and specificity (100%) obtained in this study outperformed the kinetics of the established tumor markers CA 19-9 and CEA. Moreover, this observation emphasizes the importance of defining clinically meaningful time-points for serial cfDNA assessment, in order not to overlook the dynamic changes during the first phase of treatment [298].

5.2.3. Identification and monitoring of acquired drug resistance

As discussed in Section 5.2.1, drugs that exploit genetic vulnerabilities in tumors have shown great potential for improving the treatment of cancer patients. However, the effectiveness of these drugs is often diminished by acquired resistance [299,300]. This resistance may emerge either as a result of *de novo* mutations or the expansion of a sub-clonal population of cells with pre-existing resistance [301]. Therefore, understanding the underlying mechanisms that are involved in the development of acquired resistance is crucial for preventing/controlling it, but this remains poorly understood and there is currently no effective way to detect these resistant clones or to monitor clonal populations over time. As numerous studies have shown that tumor-derived cfDNA better reflects the complete genetic landscape of the tumor compared to tissue biopsies (especially in metastatic disease) (Section 5.2.1), and also offers the additional benefit of longitudinal sampling, analysis of cfDNA represents a promising modality for sequential monitoring of the molecular response of cancer during targeted therapy.

First, cfDNA can be used to monitor the development of resistance by screening for known resistance-conferring mutations [154,274,297,300,302–310]. In a recent study, the fraction of BRAF V600E mutations in cfDNA was shown to correlate with the response of non-melanoma cancers to BRAF inhibitor (BRAFi) combination (EGFRi/MEKi) therapy. On average the presence of these mutations was detected five weeks prior to radiological evidence. In addition, longitudinal monitoring of BRAF revealed a reduction in allele fraction after 4 and 12 weeks after therapy, and correlated with progression-free survival (PFS) and overall survival (OS). Two patients with no detectable mutations throughout therapy showed prolonged survival [274]. Similar results have been obtained in earlier larger-cohort studies [221,251].

Second, serial sampling and characterization of cfDNA can identify novel resistance mutations/mechanisms [129,154,246,247,274,311–313]. Sequence analysis of tumor-derived cfDNA from metastatic CRC patients treated with FOLFOX and dasatinib, with or without cetuximab, provided interesting insights. Before treatment, 37 out of 42 patients harbored RAS/BRAF mutations. Following treatment, RAS/BRAF mutations were detected in 41 out of 42 patients. Interestingly, longitudinal profiling of tumor-derived cfDNA from 21 patients that had RAS/BRAF mutations at baseline, revealed that 11 of these patients developed additional point mutations following treatment [247]. Using a whole genome sequencing (plasma-Seq) approach, another CRC study of ten patients has shown that acquired resistance to anti-EGFR therapy was modulated by gains of KRAS (in 4 patients), MET (in 2 patients), and ERB1 (1 patient), while overrepresentation of the EGFR gene was an indicator of good anti-EGFR efficacy. In this study, no novel acquired mutations in KRAS, BRAF, PIK3CA, and EGFR were found using ultra-sensitive deep sequencing [314]. This demonstrates the utility of combining different strategies. In a study of metastatic castrate-resistant prostate cancer (CRPC), several patients treated with abiraterone acetate plus prednisone or enzalutamide, novel agents that target the androgen receptor (AR) pathway, presented with amplified

AR sequences in cfDNA and elevated levels of prostate specific antigen (PSA). This suggested that the status of AR copy number in circulation can be used as a predictor of response to therapy in CRPC patients. However, it was found that only 50% of the patients with AR amplification at baseline showed PSA response, with one patient surviving progression-free for more than 20 months [305]. An interesting finding that relates to this is the observation that copy number variations on cfDNA specific for the primary tumor were detected in the blood of breast cancer patients up to 12 years after diagnosis, despite no other evidence of disease, indicating dormancy of breast cancer cells [315]. Therefore, while it may be a useful indicator in some cases, copy number alone cannot be used to differentiate between treatment responders and non-responders. A possible auxiliary approach may include absolute quantification of cell-free mRNA levels of the AR-V7 splice variant. In a recent study, high levels of AR-V7 mRNA in blood were detectable in 15 out of 85 metastatic CRPC patients, none of which achieved a PSA response. Furthermore, high levels of cell-free AR-V7 mRNA levels were associated with (i) shorter PSA-PFS, (ii) shorter clinical PFS, and (iii) shorter OS [316].

Third, serial profiling of cfDNA can identify resistant sub-clones before the onset of clinical progression and enable earlier intervention [161,214,244,245,250,279,297,302,310,317,318]. Some good examples of this come from studies that assessed mutations that confer resistance to anti-EGFR therapy in CRC. For example, in 6 of 10 CRC patients with resistance to cetuximab and panitumumab, new KRAS mutations were detected up to 4 months before an increase of CEA was detected, and nine months prior to radiological relapse diagnosis. While the tumor cells showed a resistance to EGFR inhibitors, they remained sensitive to a combination of EGFR and MEK inhibitors, which enabled early and individual therapy adjustment [300]. Similarly, KRAS mutations were found in 9 out of 24 patients whose tumors were initially KRAS wild-type and who were treated with panitumumab monotherapy. While these mutations generally occurred 5–6 months after the start of therapy, mathematical modeling indicated that the mutations were present in expanded subclones already before the commencement of the panitumumab treatment [299]. Similar results have been obtained by a different research group [310]. Building on these studies, a recent study combined genomic profiling of serially collected cfDNA and corresponding tissue biopsies with mathematical modelling of cancer evolution. First, it was demonstrated that aberrations of the RAS pathway were present in pretreatment cfDNA, but was not detected in the corresponding tissue biopsy. Moreover, resistance to cetuximab was shown to often be polyclonal in nature and is reflected by tissue and plasma. Lastly, this approach enabled the authors to predict the expected time to treatment failure in individual patients [318].

Lastly, in some cases where a specific treatment is interrupted, a decline or complete disappearance of acquired resistance mutations is observed. When these cancer patients are rechallenged with the same therapy, they often show a markedly improved clinical response and tumor regression, which demonstrates that resistance to certain therapies can be reversed by interruption of treatment [319]. In a study of CRC patients treated with the EGFR-specific antibodies cetuximab or panitumumab, patients that developed primary or acquired resistance to therapy presented with cfDNA alterations in several genes (KRAS, NRAS, MET, ERBB2, FLT3, EGFR and MAP2K1). Interestingly, after stopping the treatment cell-free KRAS mutant clones declined. Analysis of CRC cells showed that populations with decayed mutant KRAS clones regained drug sensitivity [250]. In a recent CRC study, these cell-free mutant KRAS clones have also been shown to disappear after treatment discontinuation. In addition, they remained undetectable for a period of 12 months and were still not detected in fourth-line treatment. Remarkably, rechallenging therapy was only successful if mutant KRAS clones were not present in cfDNA [318]. These studies show that cfDNA can potentially be used to monitor the dynamic adaptive changes of the genome during intermittent drug schedules, identify patients that may be eligible for rechallenging therapy, as well as monitor the efficacy of

rechallenging therapies.

In summary, these studies show that analysis of cfDNA samples collected before and after treatment can provide an expanded view of the genetic response of a patients' tumor, including the dynamic changes in the mutational landscape as well as the heterogeneity that develops due to the selective pressure of therapy. In addition, it is clear that the potential of cfDNA analyses is enhanced when combined with results obtained from tissue biopsies and integrated with mathematical modelling of tumor evolution. When these studies are validated in larger patient cohorts, cfDNA can provide knowledge that will allow prediction of individualized response to therapy, the early adoption of alternative therapies, as well as enable the development of novel therapeutic approaches.

#### 5.2.4. Monitoring minimal residual disease (MRD) and relapse

A vital question in cancer management is whether further therapy should be administered to cancer patients following tumor resection. The most commonly used method for differentiating between disease-free patients and those with MRD after surgery is still based on the tumor-node-metastasis staging system, but this only gives a rough estimation. Since undetected and untreated MRD leads to recurrence, most patients with high-risk clinical and pathological criteria are indiscriminately subjected to pre-emptive adjuvant therapy, irrespective of the fact that some patients may have already been cured by the primary surgery and/or radiotherapy. This can lead to overtreatment, inflicting adverse effects on the wellbeing of cancer survivors. Liquid profiling of tumor-derived cfDNA has great potential in this regard. The highly sensitive methods used for cfDNA analysis described thus far also have promise for the detection of MRD and prediction of recurrence. In important earlier work, it was shown that tumor-derived cfDNA levels decrease by 99% in 24 h after complete surgical tumor resection in CRC. Conversely, in cases of incomplete resections tumor-derived cfDNA levels either reduced only slightly or increased significantly. Moreover, persistently high mutation values after surgery indicated residual disease. CfDNA showed more pronounced dynamics and had a higher predictive value for tumor recurrence than the conventional tumor marker CEA. If patients were monitored after successful surgery, measurable cfDNA levels after 1–2 months accurately identified patients with later tumor recurrence [70]. Similarly, promising results were obtained with metastatic breast cancer patients in whom the tumor-related mutations PIK3CA and TP53 were found in plasma cfDNA in 97% of cases, while CTCs and CA 15–3 were detected only in 87% and 78% of cases, respectively. Tumor-derived cfDNA levels correlated better with tumor burden and indicated tumor recurrence more accurately (89%) than CTCs (37%) or CA 15–3 (50%). Thereby cfDNA provided the earliest measure of treatment response in 53% of the progressive patients with an average lead time of 5 months to recurrence detection [227]. Numerous subsequent studies have reported similar findings in lung cancer [320–322], NSCLC [284,287,296,323], CRC [277,324], hematological malignancies [317,325–327], Hodgkin lymphoma [245], breast cancer [146,293], melanoma [289], bladder cancer [292], and ovarian cancer [215].

In a study of 55 early breast cancer patients receiving neoadjuvant chemotherapy, detection of tumor-derived cfDNA in plasma collected at a single time-point after surgery was generally predictive of MRD and metastatic relapse, but was insufficient in some cases. The accuracy of MRD detection and relapse was improved by mutational tracking of tumor-derived cfDNA in serially collected blood samples [313]. Similar results were reported in a prospective study of 230 patients with resected stage II CRC. In this study it was demonstrated that the absence of tumor-derived cfDNA after surgery was predictive of recurrence-free survival at 3 years, while the presence of tumor-derived cfDNA in serially collected samples correlated with 100% recurrence [328]. While these studies have demonstrated the capacity of cfDNA to detect MRD and identify patients that are at high-risk for recurrence, the assays were not able to detect tumor-derived cfDNA in more than half of the

patients that eventually experienced recurrence. Moreover, these assays are patient and mutation-specific. In order to overcome these limitations, a recent study utilized CAPP-seq, an NGS-based method that tracks multiple mutations per patient at very low limits of detection ( $\sim 0.002\%$ ) and does not require personalization [130,155], for the identification of MRD in patients with localized stage I-III lung cancer. Remarkably, tumor-derived cfDNA was detected in the first post-treatment blood samples of 94% of evaluable patients that experienced recurrence. Importantly, this detection preceded radiographic progression in 72% of patients by approximately 5 months. Moreover, in more than half of these patients, mutational profiles that are associated with favorable responses to TKI- or immunotherapy were identified [329]. Taken together, the above studies demonstrate the potential of cfDNA as a highly sensitive indicator of MRD and micrometastases. Therefore, perusal of cfDNA could enable the identification of patients that are likely to benefit from adjuvant therapy, and could spare cured cancer patients from the detrimental effects of undue treatment. In addition, the detection of early relapse via cfDNA could potentially guide the use of personalized targeted therapies.

#### 5.2.5. Combining tissue biopsies and cfDNA-based liquid profiling

The term “liquid biopsy” is currently used to describe the analysis of cfDNA or circulating tumor cells (CTCs) in connection with the identification of tumor characteristics [330,331]. In our opinion, this term undermines the diagnostic importance of tissue biopsies. The latter can provide the pathologist with important information on the complexity of the tumor, including necrosis, signs of hypoxia, vascularization and vascular status of the tumor bed, stroma reaction of the normal tissue, and immune cell infiltration. Moreover, a tumor biopsy can distinguish important phenotypical features of tumor cells and their heterogeneity. On the other hand, a liquid biopsy can only provide information on the molecular characteristics of the tumor itself. Besides cfDNA and CTCs, various markers at the protein-level have been used for detecting tumor characteristics in blood for many years. Indeed, the differentiation between tumor and normal tissue, or the characterization of the prognostically relevant profile of a tumor, can be done at all levels of diagnostically available biomolecules. Therefore, the multi-parametric assessment of tumor characteristics in biofluids may be better expressed by the term “liquid profiling”. This also makes it clear that the analysis is performed on blood or other body fluids and not on biopsied tissue. Therefore, based on this clear distinction, best patient management encompasses the combination of pre-therapeutic tissue biopsies and serial blood collections for liquid profiling (Fig. 5).

## 6. Concluding remarks

Cell-free DNA demonstrates immense potential as a versatile biomarker in oncology, and marks a new point of departure in the application of molecular methods for the development of comprehensive clinical tests based on non-invasive personal and precision medicine. As a next step, the clinical validity and utility of cfDNA needs to be investigated in cohorts with an appropriate number of individuals. Whereas clinical validity refers to the usefulness of cfDNA for diagnostic purposes in comparison with established methods, clinical utility refers to improving the outcome of patients – which is most challenging considering the myriad of determinants that influence the course of disease. For differential diagnosis of cancer from non-malignant diseases in patients with suspicious symptoms or radiological findings, the choice of appropriate control groups is essential to achieve a realistic view of the diagnostic performance of a marker. Depending on the depth of subgroup analyses regarding stage and histology, 50 to several hundreds of cancer patients as well as a similar number of age- and sex-matched controls with benign findings need to be included. Sensitivities at clinically relevant specificity rates as well as ROC-curves should be compared with established diagnostic tools to test the superior or additive value of cfDNA markers.

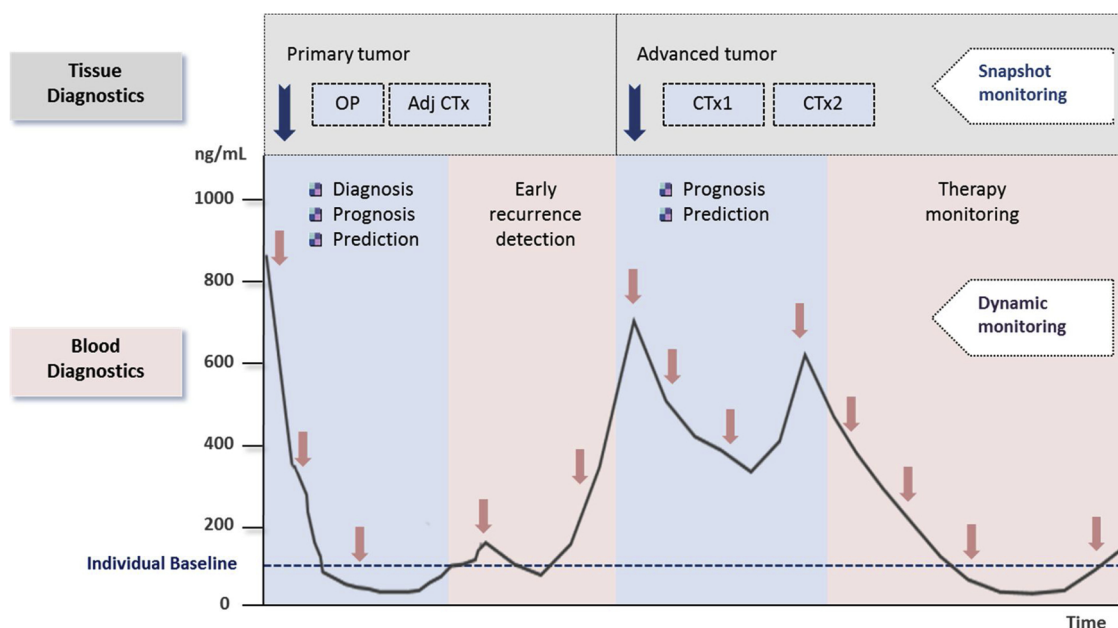
Early diagnosis or screening is a completely different setting as it assumes cancer detection in an asymptomatic population. Since the prevalence of cancer even in high risk groups (such as smokers) is often low and amounts up to only 1 in 100 or 1000 individuals, the size of the population that needs to be investigated is generally large, i.e., more than 10,000 individuals. Importantly, PPVs and NPVs are much more relevant in this context than sensitivity and specificity. Depending on cancer prevalence even markers with high sensitivities and specificities of 90% can have very poor PPVs and create high absolute numbers of false-positive findings [203] (see Section 5.1.1. c). Nevertheless, such markers may be valuable as a prescreening tool in high risk patients before other more expensive or invasive screening methods are applied. The combination of cfDNA markers with other biomarker classes, such as tumor-associated proteins has recently shown considerable improvement of diagnostic accuracy even if the cohort tested was not a screening population [126].

Prognostic biomarker studies refer to the overall outcome of the patients, including overall survival (OS), progression-free survival (PFS) or disease-free survival (DFS). As patients nowadays often receive several courses of different chemo-, antibody- or immune therapies, a multitude of factors have to be considered for evaluation of OS. Even for PFS or DFS testing, several hundreds of patients have to be enrolled. Moreover, several clinical factors and biomarkers need to be included to get meaningful results, enable subgroup analyses and multimarker score development. Similar to diagnostic approaches, validation of findings in an independent cohort remains highly important.

The role of cfDNA for prediction and serial monitoring of therapy response as well as for MRD detection is one of the most challenging endeavors as it demands the enrollment of a high number of patients undergoing a specific therapy who receive regular and comprehensive clinical and radiological reviews. For serial study designs with multiple marker determinations in particular, the definition of time points and methods for cfDNA assessment of clinically relevant and meaningful individual marker changes, and the choice of appropriate outcome correlates, requires many pre- and pilot studies. Recent studies have shown fast and complete decrease of KRAS cfDNA amounts (already as soon as after one or a few weeks) in pancreatic cancer patients responding to chemotherapy while levels of progressive patients remained high or increased early after a temporary decline [298]. This underscores the importance of performing a careful and prudent study design, in order not to miss clinically relevant plasma cfDNA dynamics. However, a great potential is the involvement of cfDNA as surrogate response markers in therapeutic studies that enables enormous insights in cfDNA kinetics, even if the markers are not used for therapy stratification. A new application of cfDNA is the stratification of rechallenging therapy e.g., by EGFR receptor inhibitors in colorectal cancer. Here, the presence of plasma cfDNA RAS mutations was an unfavorable predictive marker indicating insufficient response to therapy [318].

Although the clinical utility of cfDNA may be difficult to demonstrate definitively, first clinical trials indicate the potential of cfDNA assessments to improve patient guidance through diagnostic, prognostic, and monitoring situations. However, in order to validate the clinical usefulness of cfDNA, it should be compared to biomarkers that are currently used in routine clinical practice in large-scale randomized clinical trials [332]. Only when it shows a clear and reproducible advantage over existing biomarkers, or offers additive diagnostic, predictive or monitoring information, can it be considered for broad implementation in medical practice.

Although considerable progress has been made in the last decade, there remain many hurdles that still need to be overcome. Current cfDNA profiling strategies are not sensitive enough for the simultaneous screening of multiple cancers. As illustrated in Fig. 6, there are several ways by which this can be improved, including the optimization of preanalytical steps, sampling from body fluids with a higher mutant allele fraction, enrichment of tumor-derived cfDNA following extraction, and performing independent assays on aliquoted replicates of



**Fig. 5.** The combination of pre-therapeutic tissue biopsies and serial blood collections for liquid profiling during and after therapy may improve the guidance of cancer patients considerably. The mutational status in tissue is currently required to stratify patients for certain targeted therapies although it allows only a spatially and temporally restricted “snapshot”. Genetic heterogeneity, undetectable dormant and resistant cell clones, and adverse patient conditions limit this approach. CfDNA based liquid profiling, however, can reveal the entire mutational landscape of the cancer, and can be applied serially due to the non-invasiveness of blood collection. Furthermore, it provides essential information on the dynamics of tumor biology that can be used at various time points during the course of the disease for (i) therapy stratification, (ii) assessing prognosis, (iii) monitoring therapy response, (iv) early detection of disease progression, (v) recurrence detection, and (vi) identification of acquired resistances. Together, this may provide accurate individual patient guidance and become the cornerstone of personalized cancer medicine in the future.

isolated cfDNA. While dPCR techniques are highly sensitive, multiplexing capacity is limited. Targeted sequencing approaches based on hybrid capture or PCR amplicons can be used to query a larger number of loci. Moreover, the sensitivity of targeted sequencing can be increased by *in silico* suppression of random DNA sequencing errors.

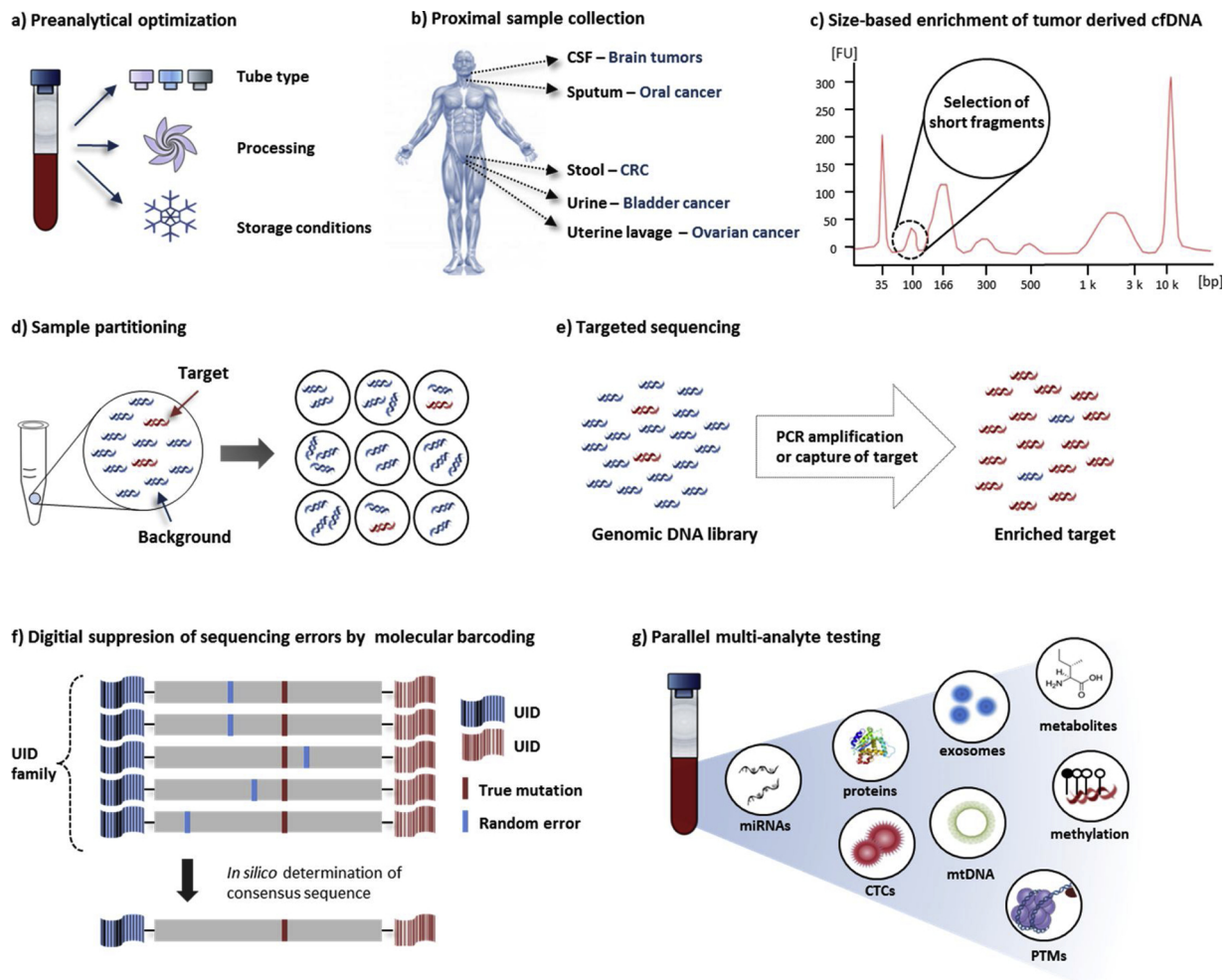
Apart from the abovementioned, recent data suggests that the sensitivity and specificity of cfDNA tests can be improved significantly by the combination of multiple biomarkers in a single parallel assessment (Fig. 6 (g)). These combinations could result either from a bottom-up approach, in which markers that are biologically complementary are assembled mainly by convenient multiplex technologies, or from a top-down approach, in which meaningful markers are extracted from a myriad of markers revealed by proteomics and genomics. While the bottom-up approach is supported by logistic regression, supporter vector machine- or neuronal network-models, the top-down approach often comprises even more complex approaches. In either case, validation in independent patient sets is paramount to confirm the findings [333,334]. Beyond the integration of many markers and biomarker classes, multiplexing will be necessary to integrate many relevant clinical questions in order to obtain a holistic view on a patient, including (i) disease characteristics, (ii) co-morbidities, (iii) unresponsiveness to specific drugs, (iv) the capacity and rate of drug metabolism, (v) unresponsiveness to toxic reactions, (vi) reactivity status of the immune system, (vii) the necessity of accompanying drugs, (viii) the interactions of diverse drugs, (ix) the development of resistances, and (x) the probability of sustained drug response and patient outcome. To address all of these questions, the future challenge will be to bring all relevant biomarker classes to a single platform. Further challenges in this regard will be to (i) facilitate quick, quality controlled, and reliable determination of marker values, (ii) integrate the resulting data in appropriate algorithms, and (iii) extract meaningful interpretations, enabling accurate decisions for patient management.

Although the above cfDNA analysis strategies are not currently suited for the early diagnosis of cancer, the methods demonstrate

considerable promise for several other applications in cancer management, including (i) later-disease staging, (ii) assessing tumor burden, (iii) prognosis, (iv) guiding therapy selection, (v) evaluating response to therapy, (vi) monitoring minimum residual disease, and (vii) predicting recurrence. Harnessing the full potential of cfDNA for these purposes requires significant refinement of the methods. While the sensitivity of these methods can be enhanced moderately by analytical standardization and incremental improvements of technology, it is becoming clear that a plateau will be reached after which major improvements will be driven mainly by an advanced understanding of the biology of cfDNA and biologic-based discoveries.

The seemingly arbitrary quantitative and qualitative fluctuation of cfDNA in the blood of an individual limits reproducible measurements, concrete interpretations of results, and interindividual comparisons. Excluding methodological reasons, this fluctuation is dependent on the rates of cfDNA release from cells and its clearance from blood. However, these mechanisms are relatively poorly understood.

An improved understanding of the rate of cfDNA release necessitates a rigorous characterization of the relative contributions of apoptosis, necrosis, other cell death mechanisms, and active secretion to the total pool of cfDNA in different settings. In addition, since each of these sources of cfDNA is highly complex in their own right, it is crucial not only to understand the underlying mechanisms, but to identify and assess all the factors that significantly affect the amount of cfDNA that is released by these processes (e.g., exercise, oxidative stress, and circadian rhythms). Similarly, an improved understanding of the rate of cfDNA clearance from circulation requires systematic *in vivo* and *in vitro* investigations of all contributing factors, including (i) the activity of DNase I, (ii) the rate of renal excretion into urine, (iii) the rate of uptake by the liver and spleen, (iv) the dynamic attachment and detachment of cfDNA to cell-surface DNA-binding proteins, and (v) the rate of cfDNA internalization by cells. Another important point that needs to be considered in this regard is that the rate of cfDNA degradation and clearance by the aforementioned mechanisms are influenced by its



**Fig. 6.** Strategies for improving the specificity and sensitivity of cell-free DNA analyses. The signal-to-noise ratio (mutant allele fraction) of samples can be increased by: (a) optimizing preanalytical steps, especially by using tube types (e.g., Streck BCT), processing procedures (immediate processing and two rounds of centrifugation), and storage conditions ( $-80^{\circ}\text{C}$ ) that minimize the release of germline DNA from peripheral blood cells; (b) sampling from body fluids that are in closer proximity to the region of interest; (c) enriching tumor-derived cell-free DNA (cfDNA) by experimental or *in silico* selection of short cfDNA fragments ( $< 166$  bp); and (d) performing independent assays on aliquoted replicates of isolated cfDNA. While dPCR methods demonstrate high sensitivity, multiplexing capacity is limited. (e) Targeted sequencing approaches based on hybrid capture or PCR amplicons can be used to query a larger number of loci. (f) In addition, the sensitivity of targeted sequencing can be improved by molecular barcoding, i.e., assigning unique identifiers (UIDs) to each DNA template during library preparation. This enables post-sequencing stratification of UID families, which allows the construction of a consensus sequence for a single molecule, thereby eliminating random errors that may have been introduced by PCR and sequencing. (g) Finally, the sensitivity and specificity of cfDNA tests could be improved significantly by the combination of multiple biomarkers in a single parallel assessment. Markers of particular interest include miRNAs, proteins, exosomes, metabolites, circulating tumor cells (CTCs), or mitochondrial DNA (mtDNA). In addition, it may be highly relevant to interrogate other features of cfDNA, such as DNA methylation changes and histone post-translational modifications (PTMs).

association with extracellular vesicles, protein complexes, and blood proteins, the levels of which can vary greatly within and between individuals at different time points. Thus, it is important to conduct fractionation experiments to determine the contribution of each cfDNA “sub-type” to the whole cfDNA population.

Two other biological factors can potentially obscure the results of cfDNA studies. First, the phenomenon of genetic mosaicism and the presence of cancer-associated genomic alterations in both tissue biopsies and cfDNA from healthy individuals that never develop cancer can result in false-positives and false-negatives. Our understanding of these phenomena will be improved by the completion of large-scale single-cell sequencing projects. Second, it is likely that factors such as metabolic activity and tumor vascularity result in an unequal contribution of different tumor subclones to the total cfDNA pool.

Although it is not currently obvious how these issues can be addressed, it is clear that further inquiry into the fundamental biology of cfDNA, and a firm grasp of its baseline values, is needed to interpret the

associations between changes in the characteristics of cfDNA and the clinical manifestations of cancer.

#### Conflict of interest

The authors declare to have no conflict of interest.

#### References

- [1] P. Mandel, *Les acides nucleiques du plasma sanguin chez 1 homme*, CR Seances Soc. Biol. Fil. 142 (1948) 241–243.
- [2] E.M. Tan, P.H. Schur, R.I. Carr, H.G. Kunkel, Deoxybonucleic acid (DNA) and antibodies to DNA in the serum of patients with systemic lupus erythematosus, J. Clin. Invest. 45 (1966) 1732–1740.
- [3] D. Koffler, V. Agnello, R. Winchester, H.G. Kunkel, The occurrence of single-stranded DNA in the serum of patients with systemic lupus erythematosus and other diseases, J. Clin. Invest. 52 (1973) 198–204.
- [4] M. Fleischhacker, B. Schmidt, Circulating nucleic acids (CNAs) and cancer—a survey, Biochim. Biophys. Acta (BBA)-Rev. Cancer 1775 (2007) 181–232.
- [5] M. Stroun, P. Anker, P. Maurice, J. Lyautey, C. Lederrey, M. Beljanski, Neoplastic



















- [309] S. Chandarlapaty, D. Chen, W. He, P. Sung, A. Samoila, D. You, T. Bhatt, P. Patel, M. Voi, M. Gnant, G. Hortobagyi, J. Baselga, M.E. Moynahan, Prevalence of ESR1 mutations in cell-free DNA and outcomes in metastatic breast Cancer: a secondary analysis of the BOLERO-2 clinical trial, *JAMA Oncol.* (2016), <https://doi.org/10.1001/jamaoncol.2016.1279>.
- [310] M.P. Morelli, M.J. Overman, A. Dasari, S.M.A. Kazmi, T. Mazard, E. Vilar, V.K. Morris, M.S. Lee, D. Herron, C. Eng, J. Morris, B.K. Kee, F. Janku, F.L. Deaton, C. Garrett, D. Maru, F. Diehl, P. Angenendt, S. Kopetz, Characterizing the patterns of clonal selection in circulating tumor DNA from patients with colorectal cancer refractory to anti-EGFR treatment, *Ann. Oncol.* (2015), <https://doi.org/10.1093/annonc/mdv005>.
- [311] L. Goyal, S.K. Saha, L.Y. Liu, G. Siravegna, I. Leshchiner, L.G. Ahronian, J.K. Lennerz, P. Vu, V. Deshpande, A. Kambadakone, B. Mussolin, S. Reyes, L. Henderson, J.E. Sun, E.E. Van Seventer, J.M. Gurski, S. Baltschukat, B. Schacher-Engstler, L. Barys, C. Stamm, P. Furet, D.P. Ryan, J.R. Stone, A. John lafrate, G. Getz, D.G. Porta, R. Tiedt, A. Bardelli, D. Juric, R.B. Corcoran, N. Bardeesy, A.X. Zhu, Polyclonal secondary FGFR2 mutations drive acquired resistance to FGFR inhibition in patients with FGFR2 fusion-positive cholangiocarcinoma, *Cancer Discov.* (2017), <https://doi.org/10.1158/2159-8290.CD-16-1000>.
- [312] M. Russo, S. Misale, G. Wei, G. Siravegna, G. Crisafulli, L. Lazzari, G. Corti, G. Rospo, L. Novara, B. Mussolin, A. Bartolini, N. Cam, R. Patel, S. Yan, R. Shoemaker, R. Wild, F. di Nicolantonio, A. Sartore-Bianchi, G. Li, S. Siena, A. Bardelli, Acquired resistance to the TRK inhibitor entrectinib in colorectal cancer, *Cancer Discov.* (2016), <https://doi.org/10.1158/2159-8290.CD-15-0940>.
- [313] I. Garcia-Murillas, G. Schiavon, B. Weigelt, C. Ng, S. Hrebien, R.J. Cutts, M. Cheang, P. Osin, A. Nerurkar, I. Kozarewa, J.A. Garrido, M. Dowsett, J.S. Reis-Filho, I.E. Smith, N.C. Turner, Mutation tracking in circulating tumor DNA predicts relapse in early breast cancer, *Sci. Transl. Med.* (2015), <https://doi.org/10.1126/scitranslmed.aab0021>.
- [314] S. Mohan, E. Heitzer, P. Ulz, I. Lafer, S. Lax, M. Auer, M. Pichler, A. Gerger, F. Eisner, G. Hoefler, T. Bauernhofer, J.B. Geigl, M.R. Speicher, Changes in colorectal carcinoma genomes under Anti-EGFR therapy identified by whole-genome plasma DNA sequencing, *PLoS Genet.* (2014), <https://doi.org/10.1371/journal.pgen.1004271>.
- [315] J.A. Shaw, K. Page, K. Blighe, N. Hava, D. Guttery, B. Ward, J. Brown, C. Ruangpratheep, J. Stebbing, R. Payne, C. Palmieri, S. Cleator, R.A. Walker, R.C. Coombes, Genomic analysis of circulating cell-free DNA infers breast cancer dormancy, *Genome Res.* (2012), <https://doi.org/10.1101/gr.123497.111>.
- [316] A.K. Seitz, S. Thoene, A. Bietenbeck, R. Nawroth, R. Tauber, M. Thalgott, S. Schmid, R. Secci, M. Retz, J.E. Gschwend, AR-V7 in peripheral whole blood of patients with castration-resistant prostate cancer: association with treatment-specific outcome under abiraterone and enzalutamide, *Eur. Urol.* 72 (2017) 828–834.
- [317] S. Mithraprabhu, S. Sirdesai, M. Chen, T. Khong, A. Spencer, Circulating tumour dna analysis for tumour genome characterisation and monitoring disease burden in extramedullary multiple myeloma, *Int. J. Mol. Sci.* (2018), <https://doi.org/10.3390/ijms19071858>.
- [318] K.H. Khan, D. Cunningham, B. Werner, G. Vlachogiannis, I. Spiteri, T. Heide, J.F. Mateos, A. Vatsiou, A. Lampis, M. Darvish Damavandi, H. Lote, I.S. Huntingford, S. Hedayat, I. Chau, N. Tunariu, G. Mentrasti, F. Trevisani, S. Rao, G. Anandappa, D. Watkins, N. Starling, J. Thomas, C. Peckitt, N. Khan, M. Ruge, R. Begum, B. Hezelova, A. Bryant, T. Jones, P. Proszek, M. Fassan, J.C. Hahne, M. Hubank, C. Braconi, A. Sottoriva, N. Valeri, Longitudinal liquid biopsy and mathematical modeling of clonal evolution forecast time to treatment failure in the PROSPECT-C phase II colorectal Cancer Clinical trial, *Cancer Discov.* (2018), <https://doi.org/10.1158/2159-8290.CD-17-0891>.
- [319] A.C. Seghers, S. Wilgenhof, C. Lebbé, B. Neyns, Successful rechallenge in two patients with BRAF-V600-mutant melanoma who experienced previous progression during treatment with a selective BRAF inhibitor, *Melanoma Res.* 22 (2012) 466–472.
- [320] I. Comino-Mendez, N. Turner, Predicting relapse with circulating tumor DNA analysis in lung cancer, *Cancer Discov.* (2017), <https://doi.org/10.1158/2159-8290.CD-17-1086>.
- [321] A.D. Nygaard, P.C. Holdgaard, K.L.G. Spindler, N. Pallisgaard, A. Jakobsen, The correlation between cell-free DNA and tumour burden was estimated by PET/CT in patients with advanced NSCLC, *Br. J. Cancer* 110 (2014) 363.
- [322] R. Catarino, A. Coelho, A. Araújo, M. Gomes, A. Nogueira, C. Lopes, R. Medeiros, Circulating DNA: diagnostic tool and predictive marker for overall survival of NSCLC patients, *PLoS One* 7 (2012) e38559.
- [323] H. Liang, J. Huang, B. Wang, Z. Liu, J. He, W. Liang, The role of liquid biopsy in predicting post-operative recurrence of non-small cell lung cancer, *J. Thorac. Dis.* (2018), <https://doi.org/10.21037/jtd.2018.04.08>.
- [324] D.H. Murray, E.L. Symonds, G.P. Young, S. Byrne, P. Rabbitt, A. Roy, K. Cornthwaite, C.S. Karapetis, S.K. Pedersen, Relationship between post-surgery detection of methylated circulating tumor DNA with risk of residual disease and recurrence-free survival, *J. Cancer Res. Clin. Oncol.* (2018), <https://doi.org/10.1007/s00432-018-2701-x>.
- [325] T.J. Pugh, Circulating tumour DNA for detecting minimal residual disease in multiple myeloma, *Semin. Hematol.* (2018), <https://doi.org/10.1053/j.seminhematol.2018.03.002>.
- [326] S. Nakamura, K. Yokoyama, N. Yusa, M. Ogawa, T. Takei, A. Kobayashi, M. Ito, E. Shimizu, R. Kasajima, Y. Wada, R. Yamaguchi, S. Imoto, T. Nagamura-Inoue, S. Miyano, A. Tojo, Circulating tumor DNA dynamically predicts response and/or relapse in patients with hematological malignancies, *Int. J. Hematol.* (2018), <https://doi.org/10.1007/s12185-018-2487-2>.
- [327] E.H. Rustad, E. Coward, E.R. Skytøen, K. Misund, T. Holien, T. Standal, M. Børset, V. Beisvag, O. Myklebost, L.A. Meza-Zepeda, H.Y. Dai, A. Sundan, A. Waage, Monitoring multiple myeloma by quantification of recurrent mutations in serum, *Haematologica.* (2017), <https://doi.org/10.3324/haematol.2016.160564>.
- [328] J. Tie, Y. Wang, C. Tomasetti, L. Li, S. Springer, I. Kinde, N. Silliman, M. Tacey, H.-L. Wong, M. Christie, Circulating tumor DNA analysis detects minimal residual disease and predicts recurrence in patients with stage II colon cancer, *Sci. Transl. Med.* 8 (2016) 346ra92-346ra92.
- [329] A.A. Chaudhuri, J.J. Chabon, A.F. Lovejoy, A.M. Newman, H. Stehr, T.D. Azad, M.S. Khodadoust, M.S. Esfahani, C.L. Liu, L. Zhou, F. Scherer, D.M. Kurtz, C. Say, J.N. Carter, D.J. Merriott, J.C. Dudley, M.S. Binkley, L. Modlin, S.K. Padda, M.F. Gensheimer, R.B. West, J.B. Shrager, J.W. Neal, H.A. Wakelee, B.W. Loo, A.A. Alizadeh, M. Diehn, Early detection of molecular residual disease in localized lung cancer by circulating tumor DNA profiling, *Cancer Discov.* (2017), <https://doi.org/10.1158/2159-8290.CD-17-0716>.
- [330] E.S. Lianidou, D. Mavroudis, G. Sotiropoulou, S. Agelaki, K. Pantel, What's new on circulating tumor cells? *FAO Nutr. Meet. Rep. Ser.* (2010).
- [331] L.A. Diaz, A. Bardelli, Liquid biopsies: genotyping circulating tumor DNA, *J. Clin. Oncol.* (2014), <https://doi.org/10.1200/JCO.2012.45.2011>.
- [332] M.B. Lustberg, D.G. Stover, J.J. Chalmers, Implementing liquid biopsies in clinical trials: state of affairs, opportunities, and challenges, *Cancer J. (United States)*. (2018), <https://doi.org/10.1097/PPO.0000000000000309>.
- [333] G. Sölétormos, M.J. Duffy, D.F. Hayes, C.M. Sturgeon, V. Barak, P.M. Bossuyt, E.P. Diamandis, M. Gion, P. Hyltoft-Petersen, R.M. Lamerz, D.L. Nielsen, P. Sibley, B. Tholander, M.K. Tuxen, J.M.G. Bonfrer, Design of tumor biomarker-monitoring trials: a proposal by the European Group on Tumor Markers, *Clin. Chem.* (2013), <https://doi.org/10.1373/clinchem.2011.180778>.
- [334] J.E. Dancey, K.K. Dobbin, S. Groshen, J.M. Jessup, A.H. Hruszkewycz, M. Koehler, R. Parchment, M.J. Ratain, L.K. Shankar, W.M. Stadler, L.D. True, A. Gravell, M.R. Grever, Guidelines for the development and incorporation of biomarker studies in early clinical trials of novel agents, *Clin. Cancer Res.* (2010), <https://doi.org/10.1158/1078-0432.CCR-09-2167>.