Investigating cell-free DNA in liquid biopsy

Cell-free DNA (cfDNA) is currently trending as a biomarker from liquid biopsy in several clinical applications, including oncology, organ and transplant medicine, and non-invasive prenatal testing (NIPT). Due to the small amount of cfDNA found in circulation, there is a need to use efficient, highly sensitive technologies, such as next-generation sequencing (NGS), to detect these biomarkers. This white paper discusses the challenges and opportunities in measuring cfDNA from liquid samples.

cytiva.com

Cell-free DNA (cfDNA) is currently trending as a biomarker for liquid biopsy in several clinical applications, including oncology, organ and transplant medicine, and non-invasive prenatal testing (NIPT). cfDNA comprises various forms of unencapsulated DNA freely circulating the bloodstream, including circulating tumor DNA (ctDNA) and cell-free fetal DNA (cffDNA). Due to the small amount of cfDNA found in circulation, there is need to use efficient, highly sensitive technologies, such as NGS, to detect these biomarkers. However, the NGS workflow—isolation, library preparation, and sequencing—can present its own sensitivity challenges in clinical application. For example, with cfDNA extraction, the release of genomic DNA from lysed or apoptotic cells contaminates the limited amount of cfDNA in a sample, thereby diluting the concentration of the cfDNA used in an assay. This white paper discusses some of these challenges and opportunities to measure cfDNA and ctDNA from blood.

Nucleic acids in cancer

Nucleic acid isolation and purification is a fundamental requirement in biological research. High-quality DNA is essential for enabling scientists across a plethora of fields to conduct life science and medical research. Automation and technological advances in DNA isolation and purification have lowered the cost and time needed for DNA sequencing and diagnostics. This is driving extensive changes through those specialties where the utilization of nucleic acids has gone far beyond just the storage of genetic information and protein synthesis.

The ability to identify tumor genotype variations between patients, called interpatient heterogeneity, has driven recent therapeutic advances in oncology. The process can help predict the clinical response and guide both conventional and novel treatments. It can also inform clinical trial enrolment. Researchers and clinicians are now able to identify intratumoral heterogeneity: subpopulations of cancer cells with distinct genomes in different regions of tumor. These subpopulations can arise during tumor growth due to microenvironmental pressures such as nutrient availability, a reduced oxygen supply (hypoxia), or radio-, chemo-, or immune-therapy treatment (1).

Sequencing technologies provide the ability to characterize intratumor heterogeneity at diagnosis, monitor subpopulation dynamics during treatment, and identify the emergence of resistant cells during disease progression. However, interpatient and intratumor heterogeneity can pose challenges for the design and enrolment of patients onto clinical trials that use genomic selection criteria. These criteria can include the presence or absence of a specific mutation, for example, *EGFR* amplification (2).

Tissue biopsy

Currently, cancer treatment is based on the accurate tissue diagnosis of samples collected by either a needle biopsy or surgical excision. After collection, light microscopy is used to reveal the sample histopathology. If they are known, these samples may also be immunostained for important biomarkers before finally performing further molecular tests. Once completed, the results of these investigations enable a cancer diagnosis and the initiation of treatment.

When there are cases where solid tumors recur after treatment at the primary site, or metastasize to distant sites, it is rare to take additional tissue biopsies to guide further ongoing treatment. Rather, the systemic treatment of patients with relapsed or metastatic disease continues, generally based on the biomarkers identified in the original tissue biopsy. However, because of intratumor heterogeneity and selective pressure throughout tumor treatment, these biomarkers might no longer represent the current disease, making further therapy ineffective. This situation has been observed in many cancers such as melanoma patients with *BRAF* mutations (3) and adult high-grade glioma patients with *EGFR* mutation (4).

Genome-scale technologies provide an unbiased characterization of clonal heterogeneity within tumors far beyond a specific genetic locus or a set of loci. Next-generation sequencing (NGS) technology has enabled the systematic detection of single nucleotide mutations as well as the identification of rare subpopulations only present in a small fraction of the overall tumor mass.

Tissue biopsy limitations

Tissue biopsy can provide information about the cancer tissue architecture and permit further molecular and histological tests. However, this approach has several limitations. Tissue biopsy is highly invasive and carries a substantial cost, both in the time needed to collect the sample and the fiscal cost in obtaining and analyzing the material. The surgical resection of primary tumors or metastatic lesions provide large volumes of tissue for assessment, and these samples are routinely formalin fixed and paraffin embedded (FFPE) to preserve histology. This can present a challenge for immunostaining as well as making the extraction of high-quality DNA challenging and expensive.

The characterization of metastatic lesions through a core needle biopsy can identify clonal evolution. However, sampling bias is a concern in these biopsies as they only enable the analysis of a limited region of a tumor. Consequently, this collection approach does not capture the heterogeneity of the whole tumor. Furthermore, tissue and needle biopsy are also inherently weak at detecting subpopulations. These subpopulations could be rare and are highly unlikely to be identified by either biopsy approach. Moreover, multiple biopsies are not feasible for many patients, for example, the

elderly or those with multiple medical conditions. As a result, an alternative sampling approach is needed to take advantage of NGS ability to sequence multiple genes simultaneously and detect rare variants in a pool of many genes.

Liquid biopsy

Recently, there have been substantial breakthroughs in biopsy collection and clinical sequencing studies using liquid biopsy (Fig 1). These are a considerably less invasive sampling method compared to other tissue collection modalities. Liquid biopsy can be sampled from a diverse range of fluids that include blood, urine, cerebrospinal fluid, saliva, stool, and lavage fluids. Due to their comparatively reduced invasiveness, liquid biopsy raises the possibility of an alternative approach for cancer diagnosis and patient care. As sample availability is substantially higher, liquid biopsy might also enable the earlier detection of disease progression. These samples might reveal changes even before clinicians could observe these by conventional imaging approaches such as X-ray and magnetic resonance imaging (MRI), or by blood protein marker changes in the patient (5, 6).

Liquid biopsy offers high specificity and efficiency for monitoring tumor changes or metastatic disease progression. They can reveal precise DNA mutations that are directly associated to specific neoplasms (7, 8). The liquid biopsy sampling method also enables easy monitoring of the patients treatment response(s) to both conventional and novel therapies (9). Due to the appreciably lower fiscal cost and time required to perform them, the approach supports the screening of at-risk population groups by simple routine blood sampling. This substantially increases the possibility for the early detection of many cancers and, as a result, could increase overall patient survival through timely initiation of treatment and surgery (10, 11). Similarly, liquid biopsy tests can be repeated as often as is necessary to monitor a patients' progress during therapy. The potential impact of this methodology in oncology is only just starting to be realized.

Fig 1. Liquid biopsy is changing the face of oncology.

The challenges presented by cell-free DNA

Circulating cell-free DNA (cfDNA) is the degraded DNA fragments released to the blood plasma (Fig 2), first described by Mandel and Métais in 1948. There are various forms of cfDNA freely circulating the bloodstream, including circulating tumor DNA (ctDNA) and cell-free fetal DNA (cffDNA). Elevated levels of cfDNA are observed in cancer, especially in advanced stages of the disease, and the detection and molecular characterization of this has enabled researchers to gain new insights into the mechanism of cancer. However, the detection of ctDNA when surrounded by normal, non-cancerous cells and non-neoplastic cfDNA presents many challenges.
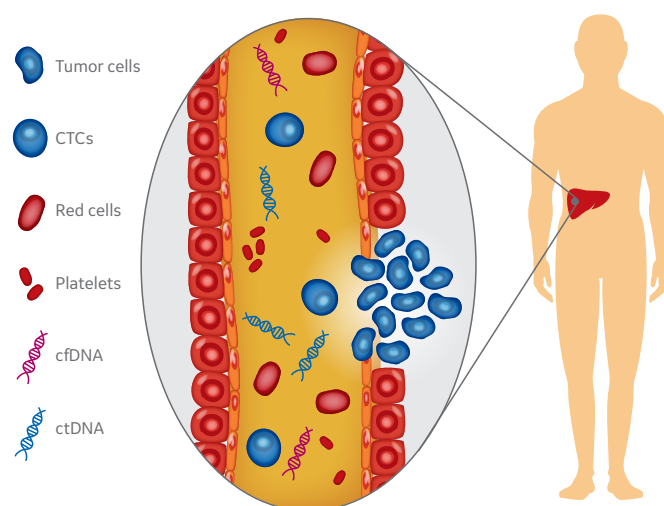

Fig 2. Circulating cell-free DNA (cfDNA) is the degraded DNA fragments released to the blood plasma.

The approaches to measure this circulating DNA can require the selective enrichment of circulating tumor cells (CTCs). This can also entail the removal of other nucleated cells, while maintaining the viability of the CTCs. Cell-free DNA refers to all non-encapsulated DNA in the blood stream (or other liquid biopsy material) and is thought to be released during apoptosis (programmed cell death), necrosis (non-programmed cell death), or by active secretion (12). Circulating tumor DNA is differentiated in that it originates from a tumor cell instead of a non-neoplastic cell.

A second, critical difference between cell-free DNA from non-neoplastic cells and ctDNA is the size of the DNA fragments. The enzymatic cleavage of DNA during apoptosis in non-neoplastic cells produces cfDNA fragments which are, on average 166 base pairs in length. This size is equivalent to DNA wrapped around a single nucleosome. In contrast, ctDNA fragments are approximately 90-150 base pairs, distinctly shorter in comparison to non-neoplastic cfDNA fragments (13).
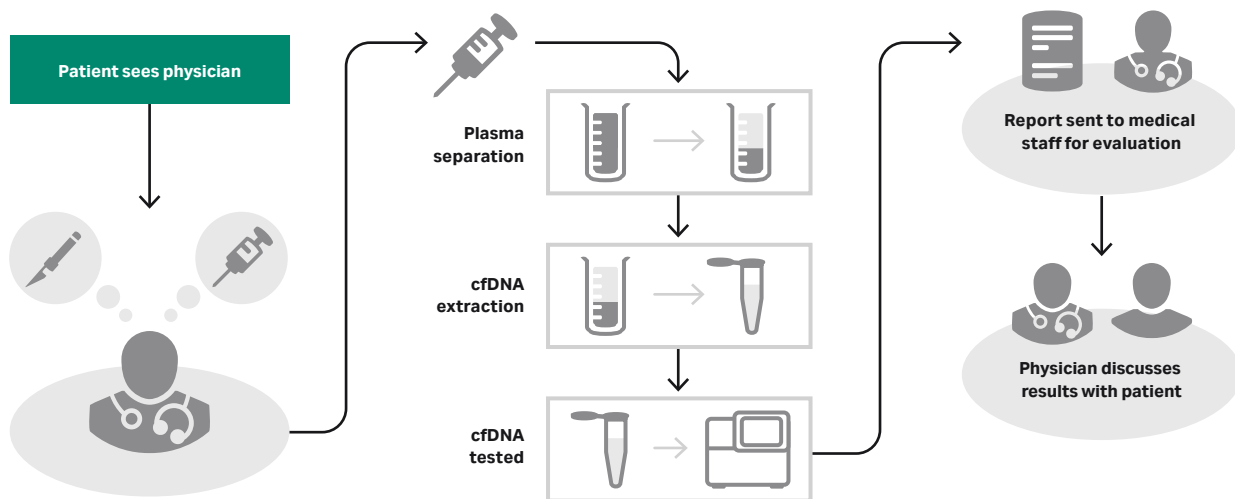
Fig 3. An overview of the cell-free DNA process from initial patient referral, through sample testing and reporting, to diagnosis and treatment initiation.

Both cfDNA from non-cancerous cells and ctDNA released from cancer cells can be detected in blood plasma. The concentration of ctDNA varies among cancer patients and is influenced by tumor type, tumor location, and disease stage. Independent of the tumor type, the frequency of ctDNA is very low (14). Circulating cell-free DNA is rapidly eliminated from the body, and with a short half-life of approximately 2.5 hours, ctDNA is hypothesized as a real-time biomarker in cancer research.

Crucially, ctDNA sequencing can reveal genetic information about the tumor cell(s) of origin. It was possible to selectively amplify *KRAS*-mutated alleles in the cfDNA extracted from pancreatic carcinoma patients by allele-specific polymerase chain reaction (PCR), then confirm by Sanger sequencing (15). The presence of amplifiable *KRAS*-mutated alleles was an exclusive feature of pancreatic tumor liquid biopsies, where none was observed with respect to healthy samples in these studies.

cfDNA technology requirements and available systems

Relative to the total volume of blood the quantity of cfDNA is low. Within this, the level of ctDNA is even lower. The cfDNA is also frequently very fragmented (16). This high level of fragmentation presents challenges during the DNA isolation process where the widespread loss of small size DNA fragments is common. As a result, the workflow, including cfDNA isolation and amplification, requires highly sensitive quantitative methods with the minimization of sample loss.

There are technologies in the market able to extract cfDNA from various liquid biopsy samples. However, there is a distinct lack of consensus regarding which extraction method is optimal for the efficient capture of such DNA. This is a critical component frequently associated with

the disparities reported in the literature, and includes the reported total concentration of plasma or serum DNA (17), in addition to the DNA integrity measurement for patient diagnosis or prognosis. As previously mentioned, while the ability to detect *KRAS*-mutated alleles is published, this has also been shown to vary depending on the chosen DNA isolation method(s) (18).

The most critical requirement for the evaluation of cfDNA from a diagnostic perspective is a robust, optimal workflow (Fig 3). This must incorporate an extraction process that consistently purifies all cfDNA fragments with the same efficiency, maximizing yield and minimizing the presence of PCR inhibitors and genomic DNA contamination. This will permit the next-generation sequencing (NGS) of the purified cfDNA and enable an accurate diagnosis.

Historically, silica column-based systems have been the mainstay for nucleic acid isolation, utilizing the binding affinity of DNA molecules in specific buffers to columns in collection tubes. However, an alternate approach, now one of the most effective for the enrichment of cfDNA, involves magnetic enrichment with functionalized magnetic beads (19).

Today, the most robust and routine methods employed in cfDNA purification are based on either magnetic beads or silica-based membranes. These methods provide high purity and quality extracted DNA for sensitive applications such as NGS analysis. The strengths and benefits of these methods are beginning to be realized. For example, one study has revealed that the cfDNA variations in *EGFR, PIK3CA* and *TP53* was associated with improved survival in patients with metastatic breast and lung cancer (20). In a second study, the levels of cfDNA were seen to be more closely correlated with the overall clinical response in breast cancer patients compared to either CTCs or the serum level of the circulating antigen biomarker: CA15-3. Both studies highlight the need for reliable isolation and quantitation processes (21).

As with CTCs, one of the greatest technical challenges is the identification of very small amounts of ctDNA within the whole cfDNA pool, where it accounts for less than one percent of the total circulating free DNA in the blood. Even with exceptional sample processing and collection, standard sequencing techniques, such as Sanger sequencing or pyrosequencing, can detect cfDNA only among patients with heavy tumor burden. Technological advances such as the introduction of digital PCR and NGS have enabled the detection of cfDNA derived from tumors in a far more sensitive and consistent manner. With the falling cost of conducting NGS, the high-coverage DNA sequencing of important cancer genes is now becoming accessible to the clinician using liquid biopsies.

cfDNA extraction and NGS from liquid biopsies

The process of obtaining, preparing, and sequencing cfDNA includes several stages where optimization can improve the overall workflow and minimize bottlenecks.

The demand for ctDNA measurement and detection in clinical trials has necessitated the development of high-throughput methods for fast, reproducible, and efficient practices throughout the entire process. These protocols are becoming widely available and comparisons are now revealing that these latest approaches, in tandem with automated systems, can replace the older, more laborious platforms, especially when high-throughput cfDNA isolation is required (23).

cfDNA sample preparation and isolation

Irrespective of a manual or an automated approach, the sample preparation and DNA isolation stages are critical for the entire cfDNA workflow. To extract cfDNA from a liquid biopsy blood sample, plasma should ideally be obtained from a cfDNA blood collection tube. This is to prevent the blood sample clotting and minimize cell lysis and artefacts. However, standard blood collection tubes with EDTA, Heparin or Citrate can be used with increased chances of genomic DNA contamination from lysed cells.

The immediate processing of blood by centrifugation might not always be possible in the clinic. A centrifuge might not be available on-site, or the blood samples might be taken over time during a clinic, particularly where a screening service or study is being carried out. This delay greatly increases the risk of genomic DNA contamination within the sample. An additional centrifugation step of the plasma prior to cfDNA isolation can mitigate this risk (24). Contamination of cfDNA with genomic DNA does not significantly affect the overall pipeline *per se*, and this may not be detected during sample quality control prior to NGS. However, if the amount of ctDNA is reported as a ratio between ctDNA and cfDNA, this will be under-estimated and the results inconsistent.

It is vital to minimize genomic DNA contamination. There are several different approaches to assess this with many based on DNA capillary electrophoresis methodologies. These methods enable an estimation of DNA fragment sizes. This is important because all PCR-based methods, including qPCR, digital PCR, and NGS are sensitive to DNA fragmentation. The fragmentation of cfDNA is regulated by nucleosome positioning and this in turn is cell type-dependent (25).

In addition to genomic DNA contamination, liquid biopsy can contain many PCR inhibitors, including heparin and immunoglobulin G (26, 27). The liquid biopsy sample itself and the extraction strategy applied will determine the purity of the isolated cfDNA. A balance needs to be struck between obtaining a high yield of cfDNA and the minimization of PCR inhibitors. The NGS process will fail if the inhibitor concentration is high. However, should the isolated cfDNA sample pass NGS quality control while PCR inhibitors are present, and the sequencing run carried out, the data obtained and analyzed could incorporate biases. This will result in the inaccurate quantitation of ctDNA within the sample.

These pre-analytical steps, including blood sampling and cfDNA isolation, have a strong influence on assay sensitivity. Many studies have compared the cfDNA yield between various extraction methods, but the quantitation of recovery is challenging to estimate. Quality control assays throughout these stages are essential to ensure an optimized liquid biopsy workflow and reproducible data. Quality control helps to standardize workflows, providing information about PCR inhibition, genomic DNA contamination, and the amount of cfDNA that can be amplified in individual samples. Each is critical when measuring ctDNA during clinical trials for patient diagnosis, screening, and treatment monitoring.

Enrichment and clean up during library preparation

Magnetic bead-based purification methods are commonly used during library prep. These beads are compatible with both manual and automated cfDNA isolation platforms as well as being scalable. These properties enable the researcher or clinician the option to use higher and lower liquid biopsy sample volumes for cfDNA extraction. Importantly, magnetic bead-based methods also enable phenol-free extraction, which notably reduces the possibility of sample contamination while also maximizing sample quality.

NGS library preparation concludes with a clean-up step of the PCR-amplified cfDNA samples. This is required because there is the need to remove unincorporated primers and dNTPS from the PCR reaction(s) before NGS. These protocols are optimized for DNA clean-up, recovery, and size selection, with several commercial kits available and designed with NGS specifically in mind.

cfDNA sequencing and analysis

Circulating cell-free DNA is opening an exciting new area in diagnostics. Some challenges remain, but these are gradually being addressed with improvements in cfDNA isolation, library preparation, sequencing and bioinformatics. Low pass whole genome sequencing is sometimes used for screening purposes, but due to the nature of the sample, the sequenced genome has many gaps and depth of coverage is not sufficient for many applications.

More commonly, targeted approaches are used, looking at subsets of mutations which are being designed into diagnostic panels (Fig 4). This is where the real potential can be realized. Here, the sensitivity of the tests is key to performance, and being able to reliably detect mutations at low variant allele frequencies (VAR) ultimately determines the quality and utility of the assay.

The first requirement is to obtain as much of the cfDNA as possible from the sample. In most cases, the volume of sample will be limited, putting more emphasis on the reliability and performance of the isolation system to recover as much cfDNA as possible with minimal gDNA interference. Secondly, to identify variants with low frequencies, they need to be discoverable above the limits of detection; therefore process errors must be understood and controlled effectively. This can be done purely on the computational side, but such methods have their limitations, so focus has moved more to the library preparation for solutions. An example of this is the use of unique molecular identifiers (UMIs) as employed in the Safe-SeqS system (Illumina), where every single stranded DNA molecule is individually barcoded prior to final amplification. As every starting molecule has a unique identifier, we can trace the DNA copies back to their origin, which allows most amplification errors to be recognized and removed from the analysis. With such limited amounts of input DNA being typical in most liquid biopsy applications, artifacts cannot be completely eliminated, but incremental improvements are leading to valuable increases in sensitivity.

Summary and conclusions

The liquid biopsy is a revolutionary approach that is opening previously unexpected opportunities in oncology. It empowers the detection and isolation of CTC, cfDNA, and ctDNA, drawing on the strength of NGS technologies now available. Many difficulties have been overcome, but there is still no standardized approach for liquid biopsy processing and cfDNA analysis. Substantial progress has been made regarding sample preparation and isolation. However, continued progress in reducing genomic DNA contamination and PCR inhibitors is needed to further improve sensitivity.

Library preparation advances have tackled some of the bottlenecks and hurdles such as cfDNA enrichment and clean-up after PCR amplification. While there are products and approaches optimized for DNA extraction, clean-up, and NGS, there remains a clear need for exclusive cfDNA extraction kits and reagents that allow standardization of the workflow process. The current open systems enable a user to select components from multiple manufacturers and develop a cost-effective cfDNA protocol for their own purpose.

Looking ahead, larger long-term studies will be required to overcome the current conceptual and analytical limits of ct/cfDNA for clinical practice. With an increasing amount of data derived from liquid biopsy, bioinformatic analysis together with the inherent complexity of cancer, consistent, reliable, and reproducible cfDNA and ctDNA methods are needed. Incorporation of big data management with artificial intelligence should also be integrated to make sure the promise of liquid biopsy and cfDNA studies can be met in the clinic.

Fig 4. Targeted approaches are used, looking at subsets of mutations which are being designed into diagnostic panels.

References

1. Vasan N. *et al.* A view on drug resistance in cancer. *Nature.* **Nov;575** (7782):299–309, (2019).

2. https://clinicaltrials.gov/, https://clinicaltrials.gov/ct2/show/NCT03382795

3. Cheng L. *et al.* Molecular testing for BRAF mutations to inform melanoma treatment decisions: a move toward precision medicine. *Mod Pathol.* **Jan;31**(1):24–38, (2018).

4. Lee E. *et al.* Comparison of glioblastoma (GBM) molecular classification methods. *Semin Cancer Biol.* **Dec;53**:201–211, (2018).

5. Dawson, S.J. *et al* Analysis of circulating tumor DNA to monitor metastatic breast cancer. *N. Engl. J. Med.* **368**, 1199–1209, (2013).

6. Diehl, F. *et al.* Circulating mutant DNA to assess tumor dynamics. *Nat. Med.* **14**, 985–990, (2008).

7. Palmirotta, R. *et al.* Liquid biopsy of cancer: A multimodal diagnostic tool in clinical oncology. *Ther Adv Med Oncol.,* **10 Aug** 29:1758835918794630, (2018).

8. Fiala, C. *et al.* Utility of circulating tumor DNA in cancer diagnostics with emphasis on early detection. *BMC Med.* **16**, 166, (2018).

9. Pantel K. *et al.* Liquid biopsy and minimal residual disease - latest advances and implications for cure. *Nat Rev Clin Oncol.* **Jul;16**(7):409–424, (2019).

10. Anfossi S. *et al.* Clinical utility of circulating non-coding RNAs - an update. *Nat Rev Clin Oncol.* **Sep;15**(9):541–563, (2018).

11. Ledergor G. *et al.* Single cell dissection of plasma cell heterogeneity in symptomatic and asymptomatic myeloma. *Nat Med.* **Dec;24**(12):1867–1876, (2018).

12. Stroun M. *et al.* About the possible origin and mechanism of circulating DNA apoptosis and active DNA release. *Clin Chim Acta.* **Nov;313**(1-2):139–42, (2001).

13. Cristiano S. *et al.* Genome-wide cell-free DNA fragmentation in patients with cancer. *Nature.* **Jun;570**(7761):385–389, (2019).

14. Bettegowda, C. *et al.* Detection of circulating tumor DNA in early- and late-stage human malignancies *Sci. Transl. Med.*, **Feb 19;6**(224):224, (2014).

15. Earl J, *et al.* Circulating tumor cells (Ctc) and kras mutant circulating free Dna (cfdna) detection in peripheral blood as biomarkers in patients diagnosed with exocrine pancreatic cancer. *BMC Cancer.* **Oct 24**;15:797, (2015).

16. Jahr S, *et al.* DNA fragments in the blood plasma of cancer patients: quantitations and evidence for their origin from apoptotic and necrotic cells. *Cancer Res* **61**:1659–65, (2001).

17. Boddy JL, *et al.* Prospective study of quantitation of plasma DNA levels in the diagnosis of malignant versus benign prostate disease. *Clin Cancer Res* **11**:1394–9, (2005).

18. Wang M, *et al.* Preferential isolation of fragmented DNA enhances the detection of circulation mutated Kras DNA. *Clin Chem* **50**:211–3, (2004).

19. Ali N. *et al.* Current nucleic acid extraction methods and their implications to point-of-care diagnostics. *BioMed Res.* **2017**:9306564, (2017).

20. Savli H, *et al.* TP53, EGFR and PIK3CA gene variations observed as prominent biomarkers in breast and lung cancer by plasma cell-free DNA genomic testing. *J Biotechnol.* **Jul 20**;300:87–93, (2019).

21. Tang Z *et al.* Diagnostic Value of Serum Concentration and Integrity of Circulating Cell-Free DNA in Breast Cancer: A Comparative Study with CEA and CA15-3. *Lab Med.* **Oct 11;49**(4):323–328, (2018).

22. Song N, *et al,* Real-time bidirectional pyrophosphorolysis-activated polymerization for quantitative detection of somatic mutations. *PLoS One.* **Apr 25**;9(4), (2014).

23. van Dessel LF, *et al.* High-throughput isolation of circulating tumor DNA: a comparison of automated platforms. *Mol Oncol.* **Feb;13**(2):392–402, (2019).

24. Trigg RM, *et al.* Factors that influence quality and yield of circulating-free DNA: A systematic review of the methodology literature. *Heliyon* **25;4**(7):e00699, (2018).

25. Snyder MW *et al.* Cell-free DNA comprises an in vivo nucleosome footprint that informs its tissues-of-Origin. *Cell.* **Jan 14;**164(1–2):57-68, (2016).

26. Schrader C, *et al.* PCR inhibitors - occurrence, properties and removal. *J. Appl. Microbiol.*, **113**:1014–1026, (2012).

27. Sidstedt M, *et al.* Inhibition mechanisms of hemoglobin, immunoglobulin G, and whole blood in digital and real-time PCR. *Anal. Bioanal. Chem.* **Apr;410**(10):2569–2583, (2018).

cytiva.com/DNA-RNA-isolation