



The Opportunities and Challenges of Molecular Tagging Next-Generation Sequencing in Liquid Biopsy

Giuseppa De Luca¹ · Mariella Dono¹

Accepted: 20 June 2021 / Published online: 5 July 2021
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2021

Abstract

Liquid biopsy (LB) is a promising tool that is rapidly evolving as a standard of care in early and advanced stages of cancer settings. Next-generation sequencing (NGS) methods have become essential in molecular diagnostics and clinical laboratories dealing with LB analytes, i.e., cell-free DNA and RNA. The sensitivity and high-throughput capacity of NGS enable us to overcome technical issues that are mainly attributable to low-abundance (below 1% mutated allelic frequency) tumour genetic material circulating within biological fluids. In this context, the introduction of unique molecular identifiers (UMIs), also known as molecular barcodes, applied to various NGS platforms greatly improved the characterization of rare genetic alterations, as they resulted in a drastic reduction in background noise while maintaining high levels of positive predictive value and sensitivity. Different UMI strategies have been developed, such as single (e.g., safe-sequencing system, Safe-SeqS) or double (duplex-sequencing system, Duplex-Seq) strand-based labelling, and, currently, considerable results corroborate their potential implementation in a routine laboratory. Recently, the US Food and Drug Administration approved the clinical use of two comprehensive UMI-based NGS assays (FoundationOne Liquid CDx and Guardant360 CDx) in cfDNA mutational assessment. However, to definitively translate LB into clinical practice, UMI-based NGS protocols should meet certain feasibility requirements in terms of cost-effectiveness, wet laboratory performance and easy access to web-source and bioinformatic tools for downstream molecular data.

Key Points

Liquid biopsy is proposed as a minimally invasive tool with potential prognostic, predictive and diagnostic purposes.

Unique molecular identifier strategies applied to NGS methodologies are a valuable molecular tool for improving one of the main challenges of liquid biopsy, i.e., detection of rare alleles in routine practice.

Further progress in UMI-based NGS panel performance is urgently needed to make their use feasible and reproducible in multiple cancer clinical settings.

1 Liquid Biopsy as a Clinical Tool

Liquid biopsy (LB) refers in general to circulating biomarkers, including cell-free DNA (cfDNA), cell-free RNA (cfRNA), circulating tumour cells (CTCs), extracellular vesicles (EVs) and tumour-educated platelets (TEPs). Although LB does not yet represent a complete substitute for tumour tissue, the hope of the clinical scientific community is that, in the near future, it can become decisive in cancer management. The LB approach offers significant advantages over others primarily because of its less invasive nature, lower cost, real-time insights into tumour status, and, in some cases, the ability to overcome the issue of tumour heterogeneity [1].

The analysis of LB analytes, mainly circulating tumour DNA and CTCs, has opened new avenues in different clinical settings, from early to metastatic disease, and for multiple purposes, such as the detection of gene drivers for patient eligibility/resistance to a treatment and/or to predict prognosis in different tumour types, including lung, breast, colorectal and prostate cancers [2].

For example, advanced lung cancer is the best example for utilization of the LB tool in two specific phases of the disease course: first-line therapeutic administration with

✉ Mariella Dono
maria.dono@hsanmartino.it

Giuseppa De Luca
giuseppa.deluca@hsanmartino.it

¹ Molecular Diagnostic Unit, IRCCS Ospedale Policlinico San Martino, 16132 Genova, Italy

detection of druggable *EGFR* targets for cases with limited and/or unavailable tumour tissue and at progression with detection of *EGFR* Thr790Met resistant mutation [3, 4].

Recently, the application of ctDNA analysis has received more attention in advanced breast cancer for the detection of *PIK3CA* mutations and patient selection for alpelisib therapy [5], as well as in colorectal cancer, where serial ctDNA samples can be a reliable device to precisely guide the timing of anti-*EGFR* therapy [6–8].

In addition to mutational analysis, some epigenetic changes, such as DNA methylation profiles, can be explored in LB, as shown in the elegant paper of Luo, who identified a methylation score able to distinguish healthy controls from colorectal cancer patients [9].

Although less abundant and more unstable than cfDNA, cfRNA is also attracting increasing interest for diagnostic, treatment-prediction, or even prognostic purposes. Indeed, cfRNA can be used to identify complex genomic rearrangements such as fusion gene transcripts in NSCLC (*ALK* and *ROS1*) or alternative splice variants (exon 14-skipping *MET* gene) as well as information about the relative expression levels of key genes [10].

It was recently found that the levels of PD-L1, measured from cfRNA in patients with NSCLC prior to immune checkpoint inhibitor treatment, are significantly associated with the response to immunotherapy, and that the changes in PD-L1 RNA over time can signal the onset of progressive disease before CT scan documentation [11]. In other studies, the expression level of PD-L1 in the plasma of cancer patients had a high degree of concordance with the corresponding tissues [12, 13].

In addition to ctDNA/RNA, circulating tumour cells (CTCs) are another main but rare component of LB, and are released into circulation from primary and/or multiple metastases in cancer patients. However, recently, thanks to advances in isolation technologies, some of the biological and molecular aspects of CTCs have been elucidated. In breast cancer, the clinical value of CTC analysis has assumed a very important prognostic role in the metastatic setting, and similar findings have been obtained in prostate cancer [14–16]. Although CTC mutational profiling has not yet reached the same clinical utility, the idea that CTCs may potentially be complementary to cfDNA is intriguing. Indeed, the transcriptional program of CTCs has been revealed at the single-cell level by advanced RNA sequencing (RNA-seq), and some prognostic signatures and resistance mechanisms have been identified in various cancer types [17–21]. Nonetheless, more studies are needed to elucidate the clinical utility of both ctDNA and CTCs to launch them as surrogate endpoints for clinical studies in which new drugs are being tested.

More recently, the search for both tumour mutational burden through ctDNA analysis and PD-L1 assessment of

CTCs for immune checkpoint inhibitor treatment as first-line therapy in the NSCLC setting is a clear example of how combinatory studies of the two LB components would aid in the treatment of cancer patients [22].

Here, we discuss the technical aspects and benefits of unique molecular identifier (UMI) strategies in next-generation sequencing (NGS) methodologies when applied to LB. Notably, we will focus our discussion on cfDNA (and cfRNA), with a prospective view on CTCs.

EVs and TEPs will not be further discussed since no definitive significant data were obtained from these analytes, and little is known about their clinical application. Comprehensive studies are available elsewhere [1, 23–25].

2 The Next-Generation Sequencing Approach to Solving LB Technical Requirements

NGS has revolutionized molecular biology in the LB field, thus representing an important alternative to the current real-time and digital PCR approaches [26–31]. NGS assays can be used to investigate the mutational status of a restricted set of genes (targeted panels) or to sequence the entire genome (WGS) or the entire exome (WES). At present, NGS-targeted panels have found a large consensus among clinical laboratories for the analysis of liquid biopsy, since they display greater sensitivity, a faster turnaround time, and lower costs than other approaches.

Several advantages have accelerated the spread of NGS applications into routine practice: (1) the ability to interrogate multiple genes simultaneously and to identify a broader range of genomic alterations (base substitutions, insertions and deletions, copy number alterations, gene rearrangement, gene expression and epigenomic variants) without sensitivity/specificity being affected, and (2) the ability to use small quantities of starting genetic material for library generation in multiplexed reactions. This issue is of paramount importance when applied to ctDNA analysis since it represents only a minor fraction (< 0.1–10%) of the total circulating cfDNA, usually derived from normal cells [30].

The accuracy of cfDNA results is not sufficient, especially when background mutation noise from ageing individuals can be interpreted erroneously as a “real” cancer-derived variation. This phenomenon is mainly derived from the possible presence of acquired mutations by clonal populations of myeloid/haematopoietic cells, also known as clonal haematopoiesis of indeterminate potential (CHIP). These mutations occur in genes usually altered in haematological disorders and even in other genes, such as *KRAS*, *NRAS*, *BRAF* and *PIK3CA* [32–34]. Since recent studies suggest that rates of CHIP mutations are high under small mutant

allele fractions [35, 36], this finding unavoidably would lower the specificity of potential tumour-derived alterations. Although this event has particular relevance when ultradeep sequencing is applied in the early stage of disease, the ability to distinguish between tumour-derived mutations and those of haematological origin is relevant for a correct patient diagnosis and eligibility for targeted therapies as well as the identification of a mutation as a result of tumour heterogeneity and/or resistance subclones [37, 38]. To this end, sequencing of matched peripheral blood mononuclear cells would be preferable.

3 UMI-Based NGS Strategies Improve the Discrimination Between Real Low-Frequency Genetic Alterations and False Positives

In the past decade, next-generation sequencing platforms have moved significantly from a research-only level to molecular diagnostics, drastically impacting the routine genotyping of tumours. In particular, the most common NGS platforms currently in use worldwide are the Illumina and Ion Torrent (ThermoFisher Scientific) systems, which are different in terms of technologies and subsequent run time, sequencing output, read length and error rates. The Illumina platform offers hybridization capture-based sequencing by synthesis, whereas Ion Torrent uses targeted amplicon-based NGS and non-optical semiconductor technology. Each platform provides high sequencing accuracy and different error rates because of the specific technologies employed, and these are estimated to be approximately 0.25–0.8% in Illumina and 1.1–2% in Ion Torrent sequencing for indel variations. In general, standard NGS with both platforms has been successfully applied to tumour tissues and is suitable for LB analysis. However, strategies have been adopted to correct and to lower the possible errors occurring at every step of the NGS workflow, which renders low-level mutation detection (<1% variant allele frequency) difficult, especially in the presence of a limited amount of genetic material.

In addition, LB samples require high (deep and/or ultradeep) sequencing to ensure an adequate number of NGS “reads” to validate a variant call; thus, the need to filter out polymerase, sequencing and miscalling errors, and the avoidance of “false calls” is urgent [39, 40]. The above considerations should alert precision medicine laboratories when choosing the most suitable NGS approach for LB genotyping applications.

One revolutionary strategy introduced and upgraded more recently has been the introduction of unique molecular identifiers (UMIs) or unique identifiers (UIDs), also known as

“molecular barcodes” or “random barcodes”, i.e., random oligomers ligated to DNA/RNA molecules, which has drastically lowered the false-positive rate [41–44]. Indeed, UMIs “tag” the original DNA molecules as distinct identities, referred to as “tag families”, i.e., common groups sharing the same UMI. After sequencing, a bioinformatics pipeline will recognize amplified reads and collapse together only those with the same UMI. If an alteration is present in all the reads within a family compared to the genomic reference, that mutation will be considered a true event.

Two different template tagging approaches can be realized depending on the incorporation of UMI into single or double DNA strands. The first considers each strand molecule as an independent family, and forward and reverse DNA sides will be tagged by two different barcodes (Fig. 1). Consequently, as shown in Fig. 1a, each tagged strand gives rise to independent families, each separately aligned on the genomic reference. Although single-stranded consensus techniques in general reduce artefacts, certain mistakes, known as “jackpot errors”, persist [45]. These errors usually occur during the first round of amplification and thus can be propagated to all subsequent produced reads.

The second alternative UMI ligation technology provides DNA double-strand tagging that produces an original DNA side back together in the original fragment after paired-end sequencing (Fig. 1b). In this case, we have correspondence between the two strands, thus implying that only if a variant is present in both chains will be further considered.

It is obvious that this last approach provides high confidence of variant calling and in some instances may reach higher sensitivity with detection of 0.01% mutants compared to the single-strand UMI method.

4 UMI-Based Technology Evolution and its Application in LB

Since its discovery, UMI technology has gone through several developmental stages, with a wide range of applications in quantitative sequencing as well as in variant detection.

In 2004, Laird and colleagues developed a model of “hairpin-bisulfite PCR” with the aim of analysing patterns of cytosine methylation by using a hairpin linker targeted and ligated to restriction-enzyme-cleaved genomic DNA before bisulfite conversion and PCR amplification, which was subsequently modified with the addition of molecular sequence barcodes [46–48].

In 2010, the first paper reported the use of molecular tagging technology applied to NGS with the intent of labelling sequence reads derived from a common progenitor molecule [49–55]. The timeline in Fig. 2 summarizes the main steps in the development of UMI-based technologies in the last decade. Among these pivotal studies, Kinde et al., published in

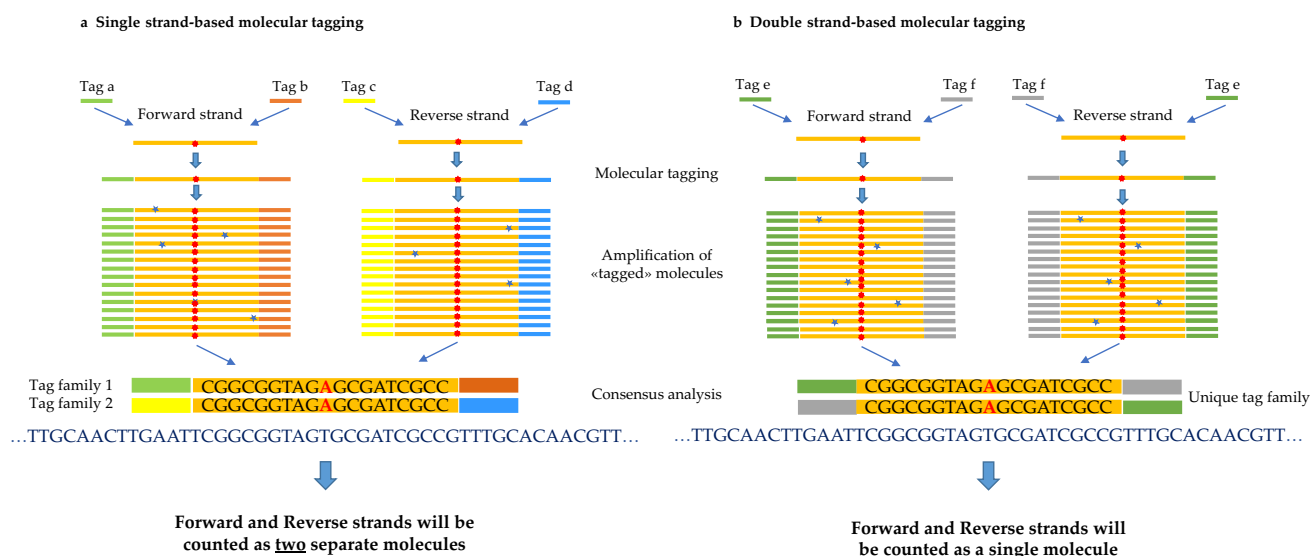


Fig. 1 An overview of unique molecular identifiers (UMIs) single- and double-strand strategies applied to a DNA molecule. The whole process is simplified by the representation of library preparation using UMIs, starting from a DNA molecule. The first step is UMI ligation to the ends of each template. The tagged fragments are amplified, and multiple reads are produced. Generally, during this step, various errors are generated, which are mainly due to nucleotide misincorporation by the polymerase enzyme. Other errors will occur during the sequencing and variant calling phases. UMIs enable a drastic reduction of the effects of these events and thus of the over-

all background noise. **A** In the single-strand-based UMI approach, the forward strand is tagged by a pair of UMIs (**a**, **b**) different from those of the reverse strand (**c**, **d**). This will generate two independent tag read families with no possibility of returning to the original DNA template molecule. **B** When a double-strand-based approach is considered, both strands are tagged with the same couple of UMIs (**e**, **f**). Since the 5' and 3' UMIs of the forward strand will correspond to the 5' and 3' UMIs of the reverse strand, both templates will be collapsed to a unique tag family and individually counted

2011, introduced the use of a random tag sequence in a PCR primer [53]. This multiplexed PCR-based NGS strategy used UMI to group redundant read results into PCR families and is currently known as the Safe-Sequencing System ("Safe-SeqS"). The single-stranded DNA fragments are labelled with a primer containing a 14-bp degenerate sequence, and any mutation found will be considered "true" only if it is present in multiple duplicates with a reduction of ~20-fold of error recurrence [53]. Safe-SeqS was successfully applied in clinical cohorts and was shown to be able to detect > 75% of patients with multiple advanced cancer types and was highly specific in the detection of minimal residual disease after surgery in colorectal cancer patients [56]. Together with the exogenous UID used to label DNA molecules during library preparation, the use of endogenous UMI derived from the sequencing of the DNA fragment end was the clear strength of this strategy. Similarly, by using patient-specific multiplex PCR-NGS formulated together with the Natera[®] companion, the Abbosh group in 2017 identified both driver genes and passenger variants (TRACERx study). This study aimed to evaluate the presence of ctDNA in plasma samples of NSCLC patients after surgery (in a minimal residual disease context) and was essentially based on preceding knowledge of the genotype signature of each patient through deep exome sequencing [57].

In 2012, Schmitt and colleagues introduced a duplex sequencing (Duplex-Seq) UMI approach, further improving error correction by using the sequence complementarity of double-stranded DNA to filter out recognizable false variants that do not appear in either strand of the original fragment [58].

In contrast with the multiplexed PCR-based NGS test described above, two other NGS approaches, i.e., cancer personalized profiling by deep sequencing (CAPP-Seq) and targeted error-correction sequencing (TEC-Seq), are hybrid capture-based methods [59, 60]. Subsequently, improvement by integrated digital error suppression (IDES) applied to CAPP-Seq technology further lowered the limit of detection from 0.02% to ~ 0.001% and allowed early-cancer-stage ctDNA analysis [61]. Similarly, a modified CAPP-seq technology involving dual-index sample barcodes and error-correcting duplexes was used for cancer screening applications, and the detection of very low amounts of ctDNA was demonstrated to be strongly prognostic in early-stage lung cancers prior to treatment [62].

New dual-matched indexed adapters with UMI were employed by MacConaill in 2018 to correct sample index misassignment and the "jumping PCR" phenomenon, which can result in false-positive calls, particularly with the Illumina platform. The use of dual-indexed UMI adapters leads

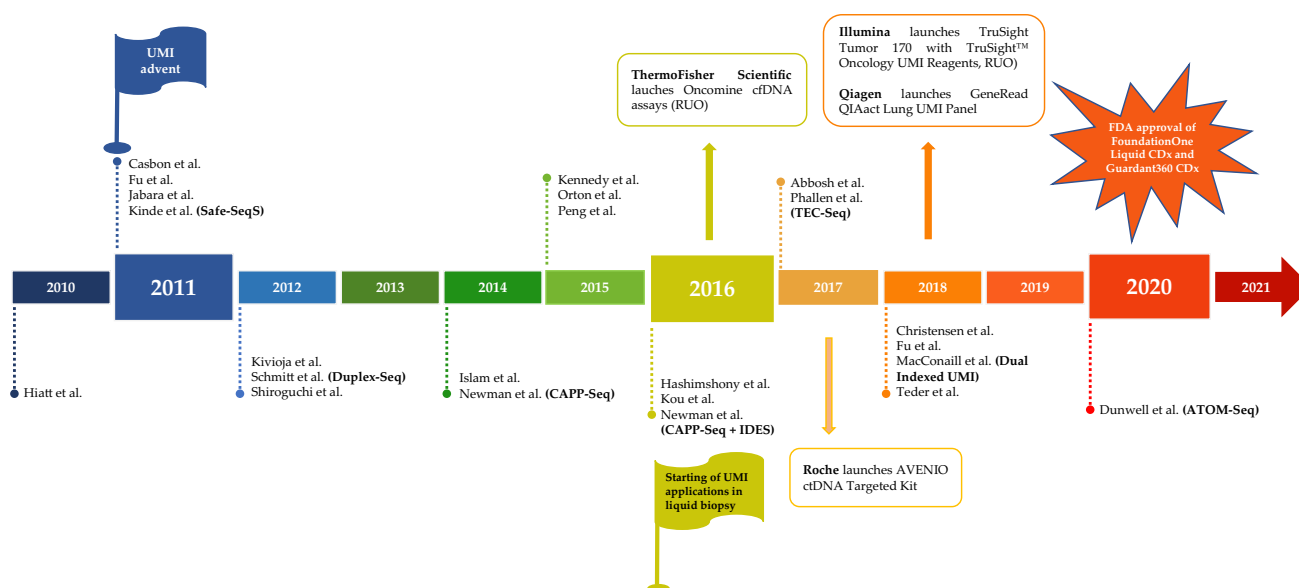


Fig. 2 A timeline of major achievements in unique molecular identifier (UMI) strategies. *Safe-SeqS* safe-sequencing system, *Duplex-seq* duplex sequencing, *CAPP-Seq* + *IDES* Cancer Personalized Profiling by Deep Sequencing + Integrated Digital Error Suppression,

TEC-Seq targeted error-correction sequencing, *FDA* US Food and Drug Administration, *ATOM-Seq* Adaptor Template Oligo Mediated Sequencing, *RUO* Research Use Only

to a drastic drop in false positives (~30-fold) and a considerable increase in the predictive positive value (PPV) from 26 to 92% when low-frequency variants (0.5–1%) are considered [63].

In 2020, Dunwell and colleagues developed a novel technology, called adaptor template oligo-mediated sequencing (ATOM-Seq), for the analysis of both cfDNA and cfRNA. Briefly, the idea was to use the original acid nucleic molecule as a primer and to capture it by extending it in the 3' direction. The extended DNA/RNA molecule was combined with a synthetic oligo, named an adaptor template oligo (ATO), bound to a single-strand random sequence at the 5' end of the ATO, which functions as a true UMI [64]. The PCR procedures and primer extension are conceptually similar to those previously designed by other groups and here have been optimized to fit LB sample analysis [65, 66].

5 Commercial UMI-Based NGS Testing for cfDNA/cfRNA Analysis

In 2020, the US Food and Drug Administration (FDA) approved two hybrid capture-based UMI-based NGS assays, i.e., FoundationOne Liquid CDx and Guardant360 CDx [67, 68]. FoundationOne Liquid CDx includes > 300 genes implicated in cancer growth and genomic biomarkers for microsatellite instability. A > 99% analytical sensitivity was reported for SNVs, indels, and gene rearrangements at a > 0.5% allelic frequency, and a 99% positive predictive

value was reported for all alterations. Clinical validation of the FoundationOne Liquid CDx test was conducted in > 30 cancer types and included > 7500 tests and > 30,000 unique variants across > 300 genes, and the results demonstrated a 95% limit of detection of 0.40% variant allele fraction for restricted substitutions and insertions/deletions, 0.37% variant allele fraction for select rearrangements, and 21.7% and 30.4% tumour fraction for copy number amplifications and copy number losses, respectively [69].

The Guardant360 CDx (Guardant Health, Inc.) NGS panel includes 73 cancer-related genes with an expected sequencing depth of ~ 15,000×, which enables variant calling and filtering of SNVs, indels, CNVs, and fusions [70, 71]. Validation of the Guardant360 CDx assay in advanced tumour patients revealed 85.9% clinical sensitivity and 97% SNV detection specificity [71]. Guardant360 CDx was specifically approved for liquid biopsy application for the detection of actionable EGFR mutations in NSCLC patients for osimertinib eligibility. The clinical utility of the Guardant360 CDx assay has been demonstrated in a large cohort of 21,807 consecutive cancer patients (25,578 total samples) affected by over 50 solid tumour types [72]. More recently, Leighl et al. demonstrated through the NILE (Non-invasive versus Invasive Lung Evaluation) study, enrolling 307 metastatic lung cancer patients, the non-inferiority of the “blood first” approach in the metastatic NSCLC setting when using the Guardant360 CDx assay [73]. In this study, the authors also showed that the detection of guideline-recommended biomarkers was at a rate similar to tissue

genotyping ($P < 0.0001$) and had a better turnaround time (9 vs. 15 days; $P < 0.0001$) and biomarker discovery rate (80%) [73].

In addition to the two FDA-approved tests, other UMI-based commercial NGS options are available from ThermoFisher Scientific and Illumina, both requiring modest minimum circulating nucleic acid input (range 20–30 ng). The Safe-Seq-based NGS panel launched by Illumina and the TruSight Tumor 170 coupled with TruSight™ Oncology UMI Reagents detects a wide range of variations in different cancer types and was demonstrated to produce highly confident variant calling while maintaining a low level of error rates [74]. ThermoFisher proposes various cancer-specific (for lung, colon and breast cancers) or PanCancer solutions for ctDNA-only genotyping or broader cfDNA/cfRNA analysis, including CNV applications. As we and others recently reported, the OncoPrint Lung cfDNA Assay based on a single-strand molecular tagging approach produced optimal results [75–77]. Conclusive results in favour of UMI-based OncoPrint Lung cfDNA assays were obtained in a comparative study in the context of EGFR Thr790Met resistance detection, demonstrating that mutations with low allelic frequencies were more reliably revealed by the use of the UMI-based panel [78]. Finally, additional commercial solutions include the QIAact Lung UMI Panel from Qiagen Company and the AVENIO ctDNA Targeted Kit from Roche. The former is lung-cancer oriented, while the latter may be used for pan-cancer research applications, even if it was especially optimized for lung cancer and colorectal cancer [79, 80].

6 New Frontiers for Applications of Molecular Tagging Technology in LB: CTC Sequencing by Skipping Whole-Genome Amplification

CTCs are live cells with great potential to best describe cancer heterogeneity and its evolving biology over time [81]. Although the CTC count can have a relevant clinical impact [14, 82–87], the enumeration of CTCs alone may not be enough to fulfil precision therapy requirements since molecular profiling provides more useful information for clinicians. Indeed, CTCs often reflect tumour resistance to previous treatment and can be derived from heterogeneous subclones that affect cancer biology evolution [88].

The lack of standardized procedures for their isolation as well as for their definitive molecular characterization are still barriers to definitive CTC implementation in clinical practice [89–91]. The pre-PCR step of whole-genome amplification (WGA) employed in standard protocols for CTC molecular characterization to produce enough starting

genetic material is known to introduce much bias (e.g., allelic dropout, chimaeric molecules, low uniformity and polymerase errors), sometimes leading to the incorrect interpretation of mutational data [92–96].

We recently proposed a modified NGS workflow based on molecular tagging technology to obtain reliable molecular data of CTC mutational status while avoiding WGA and its drawbacks. In detail, we took advantage of a targeted UMI-based NGS panel already commercialized for ctDNA evaluation in breast cancer patients and prepared NGS libraries starting from DNA derived directly from cell lysis without any preamplification step. Our findings showed that this approach, if confirmed, can be an alternative tool that is able to discriminate true mutations from false positives and wild-type DNA molecules [97].

These preliminary findings were encouraging, and further studies and perhaps more efficient protocols are needed to corroborate the idea of using UMI-based NGS for this purpose.

7 Challenges to Overcome for UMI-Based NGS Implementation in Clinical Laboratories

Both clinical and technical aspects remain to be elucidated when NGS platforms are present *in loco*, and the intention is to optimally perform a UMI-based NGS procedure instead of sending the samples to centralized laboratories that in general will produce a final analysis report.

Therefore, what requirements should there be for a UMI-based NGS panel to be fully implemented in a routine clinical laboratory with minimally experienced and skilled personnel in NGS procedures?

1. *In-house validation system* Substantial variability among the different NGS assays in terms of specificity and sensitivity has been reported, and highlights the necessity and paramount importance of an extensive validation of test performance before labs are ready to offer these tests in clinical practice, including external quality assessment.
2. *Cost effectiveness* NGS panels have the advantage of interrogating multiple genomic regions of interest through a single test. However, focusing on hotspot genes of real interest (“disease-oriented” targeted UMI-based NGS panels) has the double advantage of limiting costs and increasing assay sensitivity. For example, a UMI-based panel for NSCLC should include genes such as *EGFR*, *BRAF*, *KRAS*, *ALK*, *MET*, *RET*, *ROS1* and *NTRK*, which are currently the guideline biomarkers for

patients affected, and a few other targets such as *PIK3CA* and *ERBB2*.

3. *Faster turnaround times* By quickly reviewing wet-laboratory protocols offered by, for example, Illumina and ThermoFisher, long procedures and still too many manual steps are scheduled into different working days. The library preparation phase should be fastened by decreasing the overall steps and increasing the automation levels. This optimization may have the further advantage of drastically reducing operator-dependent errors.
4. *Robust bioinformatic pipeline* Currently, each UMI approach has a dedicated data-analysis procedure. Usually, from an NGS run, it is possible to obtain raw (e.g., FASTQ) and processed sequencing outputs (e.g., BAM and VCF) corresponding to the samples tested. Some NGS platforms, for example, the Ion Torrent system (ThermoFisher), provide an already predefined pipeline with optimized variant call parameters that facilitate the post-analytic phase. This approach also has the advantage of being manageable by personnel without a machine-learning education, but parallel processing of raw data with an alternative in-house-validated pipeline is not allowed.

Generally, a reliable bioinformatic pipeline should ensure the correct filtering of sequencing errors and only include the most confident variant calls. When UMI methodologies are involved, it is necessary to introduce additional suitable tools to perform UMI quality control and trimming and accurate reads to merge to unique tag families. This last point is very relevant, as UMI sequences are more likely to accumulate sequencing errors, which could result in ambiguity about whether sequences with similar UMIs result from sequencing errors or are unique molecular events. In this direction, web source tools, i.e., pRESTO [98], MAGERI [99] and AmpUMI [100], have been developed to perform bias correction to reduce noise and to mitigate sequencing errors, with AmpUMI, which additionally helps to establish a suitable UMI length for NGS platforms [100].

8 Conclusions

The advent of UMI technologies in the NGS approach has greatly improved the accuracy and confidence of variant calls because of precise discrimination of true mutations from the background. During the last decade, powerful single- or double-strand-based UMI strategies were created, and NGS panels were developed, with recent FDA approval

of FoundationOne Liquid CDx and Guardant360 CDx assays for cfDNA assessment.

Major efforts in wet laboratory and bioinformatics procedures are required before full implementation of UMI-based NGS methodologies in clinical laboratories. Indeed, despite the encouraging results from the clinical validation of even FDA-approved assays, it is still necessary to establish their real feasibility in a routine laboratory. A UMI-based NGS panel must in fact be sustainable for laboratory turnaround times (library preparation and NGS runs), use of economic and human resources, and easily accessible bioinformatics tools.

Declarations

Funding Not applicable.

Conflict of interest The authors declare no conflicts of interest.

Ethics approval Not applicable.

References

1. Bardelli A, Pantel K. Liquid biopsies, What we do not know (yet). *Cancer Cell*. 2017;31(2):172–9.
2. Ignatiadis M, Sledge GW, Jeffrey SS. Liquid biopsy enters the clinic—implementation issues and future challenges. *Nat Rev Clin Oncol*. 2021. <https://doi.org/10.1038/s41571-020-00457-x> (Epub ahead of print. PMID: 33473219.).
3. Oxnard GR, Thress KS, Alden RS, Lawrance R, Paweletz CP, Cantarini M, et al. Association between plasma genotyping and outcomes of treatment with osimertinib (AZD9291) in advanced non-small-cell lung cancer. *J Clin Oncol*. 2016;34:3375–82.
4. US Food & Drug Administration. Premarket approval P150044—Cobas EGFR MUTATION TEST V2.FDA. 2016. <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpma/pma.cfm?id=P150044>.
5. André F, Ciruelos E, Juric D, Loibl S, Campone M, Mayer IA, et al. Alpelisib for PIK3CA-mutated, hormone receptor-positive advanced breast cancer. *N Engl J Med*. 2019;380:1929–40.
6. Santini D, Vincenzi B, Addeo R, Garufi C, Masi G, Scartozzi M, et al. Cetuximab rechallenge in metastatic colorectal cancer patients: how to come away from acquired resistance? *Ann Oncol*. 2012;23(9):2313–8. <https://doi.org/10.1093/annonc/mdr623>.
7. Liu X, George GC, Tsimberidou AM, Naing A, Wheler JJ, Kopetz S, et al. Retreatment with anti-EGFR based therapies in metastatic colorectal cancer: impact of intervening time interval and prior anti-EGFR response. *BMC Cancer*. 2015;15:713.
8. Parseghian CM, Loree JM, Morris VK, Liu X, Clifton KK, Napolitano S, et al. Anti-EGFR-resistant clones decay exponentially after progression: implications for anti-EGFR re-challenge. *Ann Oncol*. 2019;30(2):243–9.
9. Luo H, Zhao Q, Wei W, Zheng L, Yi S, Li G, et al. Circulating tumor DNA methylation profiles enable early diagnosis,

- prognosis prediction, and screening for colorectal cancer. *Sci Transl Med*. 2020;12(524):533. <https://doi.org/10.1126/scitranslmed.aax7533> (Erratum in: *Sci Transl Med*. 2020;12(540): PMID: 31894106).
10. Berger LA, Janning M, Velthaus JL, Ben-Batalla I, Schatz S, Falk M, et al. Identification of a high-level MET amplification in CTCs and cTNA of an ALK-positive NSCLC patient developing evasive resistance to crizotinib. *J Thorac Oncol*. 2018;13(12):e243–6. <https://doi.org/10.1016/j.jtho.2018.08.2025> (Epub 2018 Sep 8 PMID: 30205165).
 11. Raez L, Danenberg K, Hunis B, Castrellon A, Jaimes Y, Velez M, et al. P203b–039 Cell-Free (cf) DNA and cRNA levels in plasma of lung cancer patients indicate disease status and predict progression. *Topic Biomark J Thorac Oncol*. 2017;12(1):959. <https://doi.org/10.1016/j.jtho.2016.11.1320>.
 12. Ishiba T, Hoffmann AC, Usher J, Elshimali Y, Sturdevant T, Dang M, et al. Frequencies and expression levels of programmed death ligand 1 (PD-L1) in circulating tumor RNA (ctRNA) in various cancer types. *Biochem Biophys Res Commun*. 2018;500(3):621–5. <https://doi.org/10.1016/j.bbrc.2018.04.120> (Epub 2018 Apr 30 PMID: 29679564).
 13. Escors D, Gato-Cañás M, Zuazo M, Arasanz H, García-Granda MJ, Vera R, et al. The intracellular signalosome of PD-L1 in cancer cells. *Signal Transduct Target Ther*. 2018;3:26. <https://doi.org/10.1038/s41392-018-0022-9>. PMID: 30275987; PMCID: PMC6160488.
 14. Bidard FC, Peeters DJ, Fehm T, Nolé F, Gisbert-Criado R, Mavroudis D, et al. Clinical validity of circulating tumour cells in patients with metastatic breast cancer: a pooled analysis of individual patient data. *Lancet Oncol*. 2014;15(4):406–14. [https://doi.org/10.1016/S1470-2045\(14\)70069-5](https://doi.org/10.1016/S1470-2045(14)70069-5) (Epub 2014 Mar 11 PMID: 24636208).
 15. Lorente D, Olmos D, Mateo J, Bianchini D, Seed G, Fleisher M, et al. Decline in circulating tumor cell count and treatment outcome in advanced prostate cancer. *Eur Urol*. 2016;70(6):985–92. <https://doi.org/10.1016/j.eururo.2016.05.023> (Epub 2016 Jun 9. PMID: 27289566; PMCID: PMC5568108).
 16. Cristofanilli M, Pierga JY, Reuben J, Rademaker A, Davis AA, Peeters DJ, et al. The clinical use of circulating tumor cells (CTCs) enumeration for staging of metastatic breast cancer (MBC): International expert consensus paper. *Crit Rev Oncol Hematol*. 2019;134:39–45.
 17. González-Silva L, Quevedo L, Varela I. Tumor functional heterogeneity unraveled by scRNA-seq technologies: (Trends in Cancer). *Trends Cancer*. 2021;7(3):265. <https://doi.org/10.1016/j.trecan.2021.02.001>.
 18. Tellez-Gabriel M, Heymann MF, Heymann D. Circulating tumor cells as a tool for assessing tumor heterogeneity. *Theranostics*. 2019;9(16):4580–94. <https://doi.org/10.7150/thno.34337>. PMID: 31367241; PMCID: PMC6643448.
 19. Miyamoto DT, Zheng Y, Wittner BS, Lee RJ, Zhu H, Broderick KT, et al. RNA-Seq of single prostate CTCs implicates non-canonical Wnt signaling in antiandrogen resistance. *Science*. 2015;349(6254):1351–6. <https://doi.org/10.1126/science.aab0917>. PMID: 26383955; PMCID: PMC4872391.
 20. Lang JE, Ring A, Porras T, Kaur P, Forte VA, Mineyev N, et al. RNA-Seq of circulating tumor cells in Stage II–III breast cancer. *Ann Surg Oncol*. 2018;25(8):2261–70. <https://doi.org/10.1245/s10434-018-6540-4> (Epub 2018 Jun 4. PMID: 29868978; PMCID: PMC7065419).
 21. Franses JW, Philipp J, Missios P, Bhan I, Liu A, Yashaswini C, et al. Pancreatic circulating tumor cell profiling identifies LIN28B as a metastasis driver and drug target. *Nat Commun*. 2020;11(1):3303. <https://doi.org/10.1038/s41467-020-17150-3>. PMID: 32620742; PMCID: PMC7335061.
 22. Hofman P, Heeke S, Alix-Panabières C, Pantel K. Liquid biopsy in the era of immuno-oncology: is it ready for prime-time use for cancer patients? *Ann Oncol*. 2019;30(9):1448–59. <https://doi.org/10.1093/annonc/mdz196> (PMID: 31228184).
 23. Yu W, Hurley J, Roberts D, Chakraborty SK, Enderle D, Noerholm M, Breakefield XO, Skog JK. Exosome-based liquid biopsies in cancer: opportunities and challenges. *Ann Oncol*. 2021;32(4):466–77. <https://doi.org/10.1016/j.annonc.2021.01.074> (Epub 2021 Feb 4 PMID: 33548389).
 24. De Rubis G, Rajeev Krishnan S, Bebawy M. Liquid biopsies in cancer diagnosis, monitoring, and prognosis. *Trends Pharmacol Sci*. 2019;40(3):172–86. <https://doi.org/10.1016/j.tips.2019.01.006> (Epub 2019 Feb 5 PMID: 30736982).
 25. Best MG, Sol N, Int Veld SGJG, Vancura A, Muller M, Niemeijer AN, et al. Swarm intelligence-enhanced detection of non-small-cell lung cancer using tumor-educated platelets. *Cancer Cell*. 2017;32(2):238–2529. <https://doi.org/10.1016/j.ccell.2017.07.004>. PMID: 28810146; PMCID: PMC6381325.
 26. Kimura H, Suminoe M, Kasahara K, Sone T, Araya T, Tamori S, et al. Evaluation of epidermal growth factor receptor mutation status in serum DNA as a predictor of response to gefitinib (IRESSA). *Br J Cancer*. 2007;97(6):778–84. <https://doi.org/10.1038/sj.bjc.6603949>.
 27. Miyazawa H, Tanaka T, Nagai Y, Matsuoka M, Huqun SA, Udagawa K, et al. Peptide nucleic acid-locked nucleic acid polymerase chain reaction clamp-based detection test for gefitinib-refractory T790M epidermal growth factor receptor mutation. *Cancer Sci*. 2008;99(3):595–600. <https://doi.org/10.1111/j.1349-7006.2007.00706.x>.
 28. Lang AH, Drexel H, Geller-Rhomberg S, Stark N, Winder T, Geiger K, et al. Optimized allele-specific real-time PCR assays for the detection of common mutations in KRAS and BRAF. *J Mol Diagn*. 2011;13(1):23–8.
 29. Vogelstein B, Kinzler KW. PCR digital. *Proc Natl Acad Sci USA*. 1999;96:9236–41.
 30. Diehl F, Schmidt K, Choti MA, Romans K, Goodman S, Li M, et al. Circulating mutant DNA to assess tumor dynamics. *Nat Med*. 2008;14(9):985–90.
 31. Hindson BJ, Ness KD, Masquelier DA, Belgrader P, Heredia NJ, Makarewicz AJ, et al. High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. *Anal Chem*. 2011;83:8604–10.
 32. Steensma DP, Ebert BL. Clonal hematopoiesis as a model for premalignant changes during aging. *Exp Hematol*. 2020;83:48–56. <https://doi.org/10.1016/j.exphem.2019.12.001>.
 33. Acuna-Hidalgo R, Sengul H, Steehouwer M, van de Vorst M, Vermeulen SH, Kiemeny LALM, Veltman JA, Gilissen C, Hoischen A. Ultra-sensitive sequencing identifies high prevalence of clonal hematopoiesis-associated mutations throughout adult life. *Am J Hum Genet*. 2017;101(1):50–64. <https://doi.org/10.1016/j.ajhg.2017.05.013>.
 34. Hu Y, Ulrich BC, Supplee J, Kuang Y, Lizotte PH, Feeney NB, Guibert NM, Awad MM, Wong KK, Jänne PA, Pawletz CP, Oxnard GR. False-positive plasma genotyping due to clonal hematopoiesis. *Clin Cancer Res*. 2018;24(18):4437–43. <https://doi.org/10.1158/1078-0432.CCR-18-0143> (Epub 2018 Mar 22. PMID: 29567812).
 35. Swanton C, Venn O, Aravanis A, Hubbell E, Maddala T, Beaussang JF, et al. Prevalence of clonal hematopoiesis of indeterminate potential (CHIP) measured by an ultra-sensitive sequencing

- assay: exploratory analysis of the Circulating Cancer Genome Atlas (CCGA) study. *J Clin Oncol*. 2018;36:12003.
36. Young AL, Challen GA, Birmann BM, Druley TE. Clonal haematopoiesis harbouring AML-associated mutations is ubiquitous in healthy adults. *Nat Commun*. 2016;22(7):12484.
 37. Chan HT, Chin YM, Nakamura Y, Low S-K. Clonal hematopoiesis in liquid biopsy: from biological noise to valuable clinical implications. *Cancers*. 2020;12(8):2277. <https://doi.org/10.3390/cancers12082277>.
 38. Chan HT, Nagayama S, Chin YM, Otaki M, Hayashi R, Kiyotani K, et al. Clinical significance of clonal hematopoiesis in the interpretation of blood liquid biopsy. *Mol Oncol*. 2020;14(8):1719–30. <https://doi.org/10.1002/1878-0261.12727>.
 39. Mehrotra M, Singh RR, Chen W, Huang RSP, Almohammedsalim AA, Barkoh BA, et al. Study of preanalytic and analytic variables for clinical next-generation sequencing of circulating cell-free nucleic acid. *J Mol Diagn*. 2017;19(4):514–24.
 40. Demuth C, Winther-Larsen A, Madsen AT, Meldgaard P, Sorensen BS. A method for treatment monitoring using circulating tumour DNA in cancer patients without targetable mutations. *Oncotarget*. 2018;9(57):31066–76.
 41. Islam S, Zeisel A, Joost S, La Manno G, Zajac P, Kasper M, et al. Quantitative single-cell RNA-seq with unique molecular identifiers. *Nat Methods*. 2014;11:163–6.
 42. Best K, Oakes T, Heather JM, Shawe-Taylor J, Chain B. Computational analysis of stochastic heterogeneity in PCR amplification efficiency revealed by single molecule barcoding. *Sci Rep*. 2015;5:14629.
 43. Kou R, Lam H, Duan H, Ye L, Jongkam N, Chen W, et al. Benefits and challenges with applying unique molecular identifiers in next generation sequencing to detect low frequency mutations. *PLoS ONE*. 2016;11(1):e0146638.
 44. Hashimshony T, Senderovich N, Avital G, Klochendler A, de Leeuw Y, Anavy L, et al. CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq. *Genome Biol*. 2016;17:77. <https://doi.org/10.1186/s13059-016-0938-8>.
 45. Salk JJ, Schmitt MW, Loeb LA. Enhancing the accuracy of next-generation sequencing for detecting rare and subclonal mutations. *Nat Rev Genet*. 2018;19(5):269–85. <https://doi.org/10.1038/nrg.2017.117> (Epub 2018 Mar 26. PMID: 29576615; PMCID: PMC6485430).
 46. Laird CD, Pleasant ND, Clark AD, Sneed JL, Hassan KM, Manley NC, et al. Hairpin-bisulfite PCR: assessing epigenetic methylation patterns on complementary strands of individual DNA molecules. *Proc Natl Acad Sci USA*. 2004;101(1):204–9.
 47. Miner BE, Stöger RJ, Burden AF, Laird CD, Hansen RS. Molecular barcodes detect redundancy and contamination in hairpin-bisulfite PCR. *Nucleic Acids Res*. 2004;32(17):e135.
 48. McCloskey ML, Stöger R, Hansen RS, Laird CD. Encoding PCR products with batch-stamps and barcodes. *Biochem Genet*. 2007;45(11–12):761–7.
 49. Hiatt JB, Patwardhan RP, Turner EH, Lee C, Shendure J. Parallel, tag-directed assembly of locally derived short sequence reads. *Nat Methods*. 2010;7(2):119–22. <https://doi.org/10.1038/nmeth.1416>.
 50. Casbon JA, Osborne RJ, Brenner S, Lichtenstein CP. A method for counting PCR template molecules with application to next-generation sequencing. *Nucleic Acids Res*. 2011;39(12): e81. <https://doi.org/10.1093/nar/gkr217>.
 51. Fu GK, Hu J, Wang PH, Fodor SP. Counting individual DNA molecules by the stochastic attachment of diverse labels. *Proc Natl Acad Sci USA*. 2011;108(22):9026–31. <https://doi.org/10.1073/pnas.1017621108>.
 52. Jabara CB, Jones CD, Roach J, Anderson JA, Swanstrom R. Accurate sampling and deep sequencing of the HIV-1 protease gene using a Primer ID. *Proc Natl Acad Sci USA*. 2011;108:20166–71. <https://doi.org/10.1073/pnas.1110064108>.
 53. Kinde I, Wu J, Papadopoulos N, Kinzler KW, Vogelstein B. Detection and quantification of rare mutations with massively parallel sequencing. *Proc Natl Acad Sci USA*. 2011;108:9530–5.
 54. Kivioja T, Vähärautio A, Karlsson K, Bonke M, Enge M, Linnarsson S, et al. Counting absolute numbers of molecules using unique molecular identifiers. *Nat Methods*. 2011;9(1):72–4. <https://doi.org/10.1038/nmeth.1778>.
 55. Shiroguchi K, Jia TZ, Sims PA, Xie XS. Digital RNA sequencing minimizes sequence-dependent bias and amplification noise with optimized single-molecule barcodes. *Proc Natl Acad Sci USA*. 2012;109(4):1347–52. <https://doi.org/10.1073/pnas.1118018109>.
 56. Bettgowda C, Sausen M, Leary RJ, Kinde I, Wang Y, Agrawal N, et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. *Sci Transl Med*. 2014;6(224):224ra24. <https://doi.org/10.1126/scitranslmed.3007094>.
 57. Abbosh C, Birkbak NJ, Wilson GA, Jamal-Hanjani M, Constantin T, Salari R, et al. Phylogenetic ctDNA analysis depicts early-stage lung cancer evolution. *Nature*. 2017;545(7655):446–51.
 58. Schmitt MW, Kennedy SR, Salk JJ, Fox EJ, Hiatt JB, Loeb LA. Detection of ultra-rare mutations by next-generation sequencing. *Proc Natl Acad Sci USA*. 2012;109(36):14508–13. <https://doi.org/10.1073/pnas.1208715109>.
 59. Newman AM, Bratman SV, To J, Wynne JF, Eclov NC, Modlin LA, et al. An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. *Nat Med*. 2014;20(5):548–54. <https://doi.org/10.1038/nm.3519>.
 60. Phallen J, Sausen M, Adleff V, Leal A, Hruban C, White J, et al. Direct detection of early-stage cancers using circulating tumor DNA. *Sci Transl Med*. 2017;9(403):2415. <https://doi.org/10.1126/scitranslmed.aan2415>.
 61. Chaudhuri AA, Chabon JJ, Lovejoy AF, Newman AM, Stehr H, Azad TD, et al. Early detection of molecular residual disease in localized lung cancer by circulating tumor DNA profiling. *Cancer Discov*. 2017;7(12):1394–403. <https://doi.org/10.1158/2159-8290.CD-17-0716> (Epub 2017 Sep 24. PMID: 28899864; PMCID: PMC5895851).
 62. Chabon JJ, Hamilton EG, Kurtz DM, Esfahani MS, Moding EJ, Stehr H, et al. Integrating genomic features for non-invasive early lung cancer detection. *Nature*. 2020;580(7802):245–51. <https://doi.org/10.1038/s41586-020-2140-0> (Epub 2020 Mar 25 PMID: 32269342).
 63. MacConaill LE, Burns RT, Nag A, Coleman HA, Slevin MK, Giorda K, et al. Unique, dual-indexed sequencing adapters with UMIs effectively eliminate index cross-talk and significantly improve sensitivity of massively parallel sequencing. *BMC Genomics*. 2018;19:30. <https://doi.org/10.1186/s12864-017-4428-5>.
 64. Dunwell TL, Dailey SC, Yu J, Becker PW, Scaife S, Richman SD, et al. Adaptor template oligo-mediated sequencing (ATOM-Seq): a versatile and ultra-sensitive UMI-based NGS library preparation technology, for use with cfDNA and cfRNA. *bioRxiv*. 2020. <https://doi.org/10.1101/2020.07.10.152405>.
 65. Zheng Z, Liebers M, Zhelyazkova B, Cao Y, Panditi D, Lynch KD, et al. Anchored multiplex PCR for targeted next-generation sequencing. *Nat Med*. 2014;20(12):1479–84. <https://doi.org/10.1038/nm.3729>.

66. Xu C, Nezami Ranjbar M, Wu Z, Di Carlo J, Wany Y. Detecting very low allele fraction variants using targeted DNA sequencing and a novel molecular barcode-aware variant caller. *BMC Genomics*. 2017;18:5. <https://doi.org/10.1186/s12864-016-3425-4>.
67. US Food and Drug Administration. FoundationOne® Liquid CDx (F1 Liquid CDx). FDA (2020). https://www.accessdata.fda.gov/cdrh_docs/pdf20/P200016A.pdf.
68. US Food and Drug Administration. Guardant360® CDx. FDA (2020). https://www.accessdata.fda.gov/cdrh_docs/pdf20/P200010A.pdf.
69. Woodhouse R, Li M, Hughes J, Delfosse D, Skoletsky J, Ma P, et al. Clinical and analytical validation of FoundationOne Liquid CDx, a novel 324-Gene cfDNA-based comprehensive genomic profiling assay for cancers of solid tumor origin. *PLoS ONE*. 2020;15(9):e0237802. <https://doi.org/10.1371/journal.pone.0237802>. PMID:32976510;PMCID:PMC7518588.
70. Lanman RB, Mortimer SA, Zill OA, Sebisano D, Lopez R, Blau S, et al. Analytical and clinical validation of a digital sequencing panel for quantitative, highly accurate evaluation of cell-free circulating tumor DNA. *PLoS ONE*. 2015;10(10):e0140712. <https://doi.org/10.1371/journal.pone.0140712>. PMID:26474073;PMCID:PMC4608804.
71. Odegaard JI, Vincent JJ, Mortimer S, Vowles JV, Ulrich BC, Banks KC, et al. Validation of a plasma-based comprehensive cancer genotyping assay utilizing orthogonal tissue- and plasma-based methodologies. *Clin Cancer Res*. 2018;24(15):3539–49. <https://doi.org/10.1158/1078-0432.CCR-17-3831> (Epub 2018 Apr 24 PMID: 29691297).
72. Zill OA, Banks KC, Fairclough SR, Mortimer SA, Vowles JV, Mokhtari R, et al. The landscape of actionable genomic alterations in cell-free circulating tumor DNA from 21,807 advanced cancer patients. *Clin Cancer Res*. 2018;24(15):3528–38. <https://doi.org/10.1158/1078-0432.CCR-17-3837> (Epub 2018 May 18 PMID: 29776953).
73. Leighl NB, Page RD, Raymond VM, Daniel DB, Divers SG, Reckamp KL, et al. Clinical utility of comprehensive cell-free DNA analysis to identify genomic biomarkers in patients with newly diagnosed metastatic non-small cell lung cancer. *Clin Cancer Res*. 2019;25(15):4691–700. <https://doi.org/10.1158/1078-0432.CCR-19-0624> (Epub 2019 Apr 15 PMID: 30988079).
74. Illumina 2018. Detection of low-frequency variants in cfDNA using TruSight Tumor 170 with the TruSight UMI Toolkit. <https://www.illumina.com/content/dam/illumina-marketing/documents/products/technotes/trusight-umi-trusight-tumor-170-app-note-1000000050427.pdf>.
75. De Luca G, Lastraioli S, Conte R, Mora M, Genova C, Rossi G, et al. Performance of the oncomine™ lung cfDNA assay for liquid biopsy by NGS of NSCLC patients in routine laboratory practice. *Appl Sci*. 2020;10:2895.
76. Dono M, De Luca G, Lastraioli S, Anselmi G, Dal Bello MG, Coco S, et al. Tag-based next generation sequencing: a feasible and reliable assay for EGFR T790M mutation detection in circulating tumor DNA of non small cell lung cancer patients. *Mol Med*. 2019;25(1):15. <https://doi.org/10.1186/s10020-019-0082-5>.
77. Papadopoulou E, Tsoulos N, Tsantikidi K, Metaxa-Mariatou V, Stamou PE, Kladi-Skandalis A, et al. Clinical feasibility of NGS liquid biopsy analysis in NSCLC patients. *PLoS ONE*. 2019;14(12):e0226853. <https://doi.org/10.1371/journal.pone.0226853>. PMID:31860648;PMCID:PMC6924668.
78. Vollbrecht C, Lehmann A, Lenze D, Hummel M. Validation and comparison of two NGS assays for the detection of EGFR T790M resistance mutation in liquid biopsies of NSCLC patients. *Oncotarget*. 2018;9:18529–39.
79. Weber S, Spiegl B, Perakis SO, Ulz CM, Abuja PM, Kashofer K, et al. Technical evaluation of commercial mutation analysis platforms and reference materials for liquid biopsy profiling. *Cancers (Basel)*. 2020;12(6):1588. <https://doi.org/10.3390/cancers12061588>. PMID:32560092;PMCID:PMC7352370.
80. Vessies DC, Linders TC, Ramkisoensing KL, Nederlof PM, Meijer GA, Monkhorst K, et al. Technical validation of Roche AVENTIO sequencing platform for liquid biopsies [abstract]. In: Proceedings of the American Association for Cancer Research Annual Meeting 2019; 2019 Mar 29–Apr 3; Atlanta, GA. Philadelphia (PA): AACR; Cancer Res 2019;79(13 Suppl):Abstract nr 1381.
81. Alix-Panabieres C, Pantel K. Circulating tumor cells: liquid biopsy of cancer. *Clin Chem*. 2013;59:110–8.
82. Huang X, Gao P, Song Y, Sun J, Chen X, Zhao J, et al. Meta-analysis of the prognostic value of circulating tumor cells detected with the cell search system in colorectal cancer. *BMC Cancer*. 2015;15:202. <https://doi.org/10.1186/s12885-015-1218-9>. PMID:25880692;PMCID:PMC4389311.
83. Tol J, Koopman M, Miller MC, Tibbe A, Cats A, Creemers GJ, et al. Circulating tumour cells early predict progression-free and overall survival in advanced colorectal cancer patients treated with chemotherapy and targeted agents. *Ann Oncol*. 2010;21(5):1006–12. <https://doi.org/10.1093/annonc/mdp463> (Epub 2009 Oct 27 PMID: 19861577).
84. Krebs MG, Sloane R, Priest L, Lancashire L, Hou JM, Greystoke A, et al. Evaluation and prognostic significance of circulating tumor cells in patients with non-small-cell lung cancer. *J Clin Oncol*. 2011;29(12):1556–63. <https://doi.org/10.1200/JCO.2010.28.7045> (Epub 2011 Mar 21 PMID: 21422424).
85. Sparano J, O'Neill A, Alpaugh K, Wolff AC, Northfelt DW, Dang CT, et al. Association of circulating tumor cells with late recurrence of estrogen receptor-positive breast cancer: a secondary analysis of a randomized clinical trial. *JAMA Oncol*. 2018;4(12):1700–6. <https://doi.org/10.1001/jamaoncol.2018.2574>. PMID:30054636;PMCID:PMC6385891.
86. Alix-Panabieres C, Pantel K. Clinical applications of circulating tumor cells and circulating tumor DNA as liquid biopsy. *Cancer Discov*. 2016;6:479–91.
87. Lianidou E, Pantel K. Liquid Biopsies. *Genes Chromosomes Cancer*. 2019;58:219–32.
88. Ma N, Jeffrey SS. Deciphering cancer clues from blood. *Science*. 2020;367(6485):1424–5. <https://doi.org/10.1126/science.abb0736> (PMID: 32217712).
89. Cristofanilli M, Budd GT, Ellis MJ, Stopeck A, Matera J, Miller MC, et al. Circulating tumor cells, disease progression, and survival in metastatic breast cancer. *N Engl J Med*. 2004;351:781–91.
90. Hayes DF, Cristofanilli M, Budd GT, Ellis MJ, Stopeck A, Miller MC, et al. Circulating tumor cells at each follow-up time point during therapy of metastatic breast cancer patients predict progression-free and overall survival. *Clin Cancer Res*. 2006;12:4218–24.
91. Kanwar N, Hu P, Bedard P, Clemons M, McCready D, Done SJ. Identification of genomic signatures in circulating tumor cells from breast cancer. *Int J Cancer*. 2015;137:332–44.
92. Lasken RS, Egholm M. Whole genome amplification: abundant supplies of DNA from precious samples or clinical specimens. *Trends Biotechnol*. 2003;21:531–5.

93. Hughes S, Arneson N, Done S, Squire J. The use of whole genome amplification in the study of human disease. *Prog Biophys Mol Biol*. 2005;88:173–89.
94. Pinard R, deWinter A, Sarkis G, Gerstein MB, Tartaro KR, Plant RN, et al. Assessment of whole genome amplification-induced bias through high-throughput, massively parallel whole genome sequencing. *BMC Genom*. 2006;7:216.
95. Huang L, Ma F, Chapman A, Lu S, Xie XS. Single-cell whole-genome amplification and sequencing: Methodology and applications. *Annu Rev Genom Hum Genet*. 2015;16:79–102.
96. Zhu Z, Qiu S, Shao K, Hou Y. Progress and challenges of sequencing and analyzing circulating tumor cells. *Cell Biol Toxicol*. 2018;34:405–15.
97. De Luca G, Cardinali B, Del Mastro L, Lastraioli S, Carli F, Ferrarini M, et al. Optimization of a WGA-free molecular tagging-based NGS protocol for CTCs mutational profiling. *Int J Mol Sci*. 2020;21:4364.
98. Vander Heiden JA, Yaari G, Uduman M, Stern JN, O'Connor KC, Hafler DA, et al. pRESTO: a toolkit for processing high-throughput sequencing raw reads of lymphocyte receptor repertoires. *Bioinformatics*. 2014;30(13):1930–2. <https://doi.org/10.1093/bioinformatics/btu138> (Epub 2014 Mar 10. PMID: 24618469; PMCID: PMC4071206).
99. Shugay M, Zaretsky AR, Shagin DA, Shagina IA, Volchenkov IA, Shelenkov AA, et al. MAGERI: Computational pipeline for molecular-barcoded targeted resequencing. *PLoS Comput Biol*. 2017;13(5): e1005480. <https://doi.org/10.1371/journal.pcbi.1005480>. PMID:28475621;PMCID:PMC5419444.
100. Clement K, Farouni R, Bauer DE, Pinello L. AmpUMI: design and analysis of unique molecular identifiers for deep amplicon sequencing. *Bioinformatics*. 2018;34(13):i202–10. <https://doi.org/10.1093/bioinformatics/bty264>. PMID:29949956;PMCID: PMC6022702.