Liquid biopsy for cancer screening, patient stratification and monitoring

Graham Brock, Elena Castellanos-Rizaldos, Lan Hu, Christine Coticchia, Johan Skog

Exosome Diagnostics, Cambridge, MA 02139, USA

Correspondence to: Johan Skog. Exosome Diagnostics, 840 Memorial Drive, Cambridge, MA 02139, USA. Email: Johan@exosomedx.com.

Abstract: Molecular characterization of a patient's tumor to guide treatment decisions is increasingly being applied in clinical care and can have a significant impact on disease outcome. These molecular analyses, including mutation characterization, are typically performed on tissue acquired through a biopsy at diagnosis. However, tumors are highly heterogeneous and sampling in its entirety is challenging. Furthermore, tumors evolve over time and can alter their molecular genotype, making clinical decisions based on historical biopsy data suboptimal. Personalized medicine for cancer patients aims to tailor the best treatment options for the individual at diagnosis and during treatment. To fully enable personalized medicine it is desirable to have an easily accessible, minimally invasive way to determine and follow the molecular makeup of a patient's tumor longitudinally. One such approach is through a liquid biopsy, where the genetic makeup of the tumor can be assessed through a biofluid sample. Liquid biopsies have the potential to help clinicians screen for disease, stratify patients to the best treatment and monitor treatment response and resistance mechanisms in the tumor. A liquid biopsy can be used for molecular characterization of the tumor and its non-invasive nature allows repeat sampling to monitor genetic changes over time without the need for a tissue biopsy. This review will summarize three approaches in the liquid biopsy field: circulating tumor cells (CTCs), cell free DNA (cfDNA) and exosomes. We also outline some of the analytical challenges encountered using liquid biopsy techniques to detect rare mutations in a background of wild-type sequences.

Keywords: Liquid biopsy; exosome; circulating tumor cell (CTC); cell free DNA (cfDNA); nucleic acids

Submitted May 11, 2015. Accepted for publication Jun 11, 2015.
doi: 10.3978/j.issn.2218-676X.2015.06.05
View this article at: http://dx.doi.org/10.3978/j.issn.2218-676X.2015.06.05

Introduction

The science of noninvasive disease monitoring has advanced greatly since circulating cell free DNA (cfDNA) was first reported in body fluids by Mandel and Metais (1). Since then, the evolution of sensitive cfDNA detection technologies has enabled the development of liquid biopsies with many clinical applications. For example, in oncology, the use of liquid biopsy allows for patient stratification (companion diagnostics), screening, monitoring treatment response and detection of minimal residual disease after surgery/recurrence.

Liquid biopsies have grown in importance because, the genetic profile of tumors can affect how well they respond to a certain treatment. However, this characterization is currently achieved through a biopsy despite the inherent problems in procurement of tissue samples and the limitations of tumor analyses. For example, the invasive nature of a biopsy poses a risk to patients and can have a significant cost. Tumor sampling from some cancer types also remains difficult resulting in inadequate amount of tissue for genetic testing. In the case of advanced or metastatic non-small cell lung cancers (NSCLC) as many as 31% of cases do not have accessible tissue (2). Even when tissue can be collected, preservation methods such as formalin fixation can cause C > T transitions through deamination of cytosine, potentially leading to false positive results for genetic tests (3). Finally, due to tumor heterogeneity, biopsies often suffer from sample bias (4).

More concerning with respect to guiding treatment decisions; biopsies will only inform of the genotype at that
time-point. However, it is known that tumors are very dynamic and can change their dominant mutation pattern or acquire new mutations, especially after the selective pressure of drug treatment. This could be particularly unfavorable when stratifying patients to a specific targeted therapy based on historical mutation profiles of past tumor biopsies. In another example, approximately 50% of NSCLC patients become resistant to tyrosine kinase inhibitor therapy through an epidermal growth factor receptor (EGFR) T790M mutation (5,6), significantly only <5% of NSCLC patient have this mutation detectable in the primary biopsy (7). Another study showed that 38% of colorectal cancers with wild-type Kirsten rat sarcoma viral oncogene homolog (KRAS) developed mutations in this gene after anti-EGFR therapy as rapidly as 6 months after treatment (8).

Several reports have indicated there are difficulties in detecting tumor derived mutations in plasma, while others have been able to efficiently isolate circulating tumor derived nucleic acid in both metastatic and non-metastatic disease (9-11). This discrepancy is likely due to the methodologies used for detection of the mutation, as the allelic fraction of tumor derived circulating DNA varies from less than 0.01% (or undetectable) to over 90% (12,13). In addition, the amount of recoverable DNA varies significantly (over 3 logs) between patients with an average of about 17 ng of DNA per mL of plasma from advanced-stage cancers (14), corresponding to roughly 5,000 haploid genome equivalents.

Recent technological developments and the downstream analytics being applied to liquid biopsies are now capable of reproducibly detecting mutations at very low allelic frequencies. Advances have also been made in droplet digital PCR (ddPCR) (15), next-generation sequencing (NGS) (16), beads, emulsion, amplification and magnetics (BEAMing) (13), amplification of refractory mutation system (ARMS) (17), co-amplification at lower denaturation temperature-PCR (COLD-PCR) and its derivatives (18,19) and PointMan™ DNA enrichment technology (20), to name but a few.

Ultimately the choice of platforms and required detection limit will depend on the clinical sample being analyzed, as the most sensitive methods are reported to detect allelic frequencies of as little as 0.01%, providing a theoretical lower limit to detect one mutated copy in a background of 10,000 wild-type alleles (13). Thus, this level of sensitivity requires samples/patients where at least 10,000 target alleles enter the downstream analytical assay.

Although technically challenging, an inherent advantage of liquid biopsies over other traditional tissue-based methodologies is the enablement of longitudinal monitoring which could help clinical oncologists gain a broader molecular understanding of the disease. This review will focus on the application of genetic profiling of tumor associated RNA and DNA derived from biofluids.

**Approaches to liquid biopsy analysis**

**Circulating tumor cells (CTCs)**

CTCs are cells shed into the vasculature from a primary tumor and may constitute seeds for subsequent growth of additional tumors (metastasis) in distant organs. They have been detected in various metastatic carcinomas for example breast, prostate, lung, and colorectal cancer (21,22) but are extremely rare in healthy subjects and patients with nonmalignant diseases (23). Clinical evidence indicates that patients with metastatic lesions are more likely to have CTCs amenable to isolation but their frequency is low, often ~1-10 CTCs per mL of whole blood (24). As 1 mL of blood contains ~7×10^6 white blood cells and ~5×10^6 red blood cells (25), technologies capable of reproducibly isolating a single CTC from the background of all other blood components are fundamental. While such levels of sensitivity are challenging, there are several novel developments in this area. These include positive selection, negative selection, physical properties or even enrichment-free assays to efficiently isolate these rare CTCs (26,27).

Typically, CTCs are defined as cells with an intact viable nucleus, cytokeratin positive, epithelial cell adhesion molecule (EpCAM) positive and with the absence of CD45. Unfortunately EpCAM and other markers are not always expressed on CTCs and are down-regulated by processes such as epithelial to mesenchymal transition (28). In addition, non-tumor epithelial cells are known to circulate in the blood of patients with prostatitis (29) or patients undergoing surgery (30). From a technical standpoint, the heterogeneity of CTCs is a major challenge and this has led to alternative strategies of CTC enrichment, such as the CTC-iChip (31), which do not rely on tumor antigen expression.

Sequencing the genetic material from CTCs has demonstrated that, even when the isolated cell(s) fit the phenotypic criteria of being a CTC, the majority are not cancer cells. One study developed a protocol to recover the CTC enriched samples from the cartridge of the Veridex platform and found that from 37 NSCLC patients, the
Cell free DNA (cfDNA)

There is currently an intensive research effort to understand the utility of cfDNA in various clinical fields such as cancer research (38,39), non-invasive prenatal testing (40) and transplant rejection diagnostics (41). Initial studies in cancer patients reported that cfDNA concentration in serum was significantly increased in comparison to healthy individuals (42), and it was suggested that this correlated with malignancy (43).

Most cfDNA in plasma is reportedly fragmented, around 150-180 bp in length (44) with a higher prevalence of tumor associated mutations in the shorter fragments (9). In fact, when analyzing the mutation abundance with massively parallel sequencing a significant correlation was found between mutations and fragments less than 150 bp (44). Notably, the size of the majority of cfDNA fragments overlaps well with the size of histone DNA (45).

The entry of cfDNA into the bloodstream is thought to originate from a cell following apoptosis or necrosis. Late stage cancer patients also have an increased level of cfDNA in plasma, however, most of this DNA is wild-type and believed to be from non-malignant cells and tumor stroma (9). It has also been suggested that the mutant fraction of cfDNA is derived from necrotic neoplastic cells phagocytized by macrophages, which then release digested DNA, a phenomena not seen in macrophages that engulf apoptotic cells (14). The extensive background of wild-type DNA limits the ability of downstream analytical platforms to detect tumor-derived mutation, presenting technical challenges for the use of cfDNA in liquid biopsies. While cell-free tumor DNA analyses are capable of examining the genetic or epigenetic changes that originate in tumor DNA (such as mutations, translocations, amplifications, indels and methylation abnormalities), they cannot analyze the tumor RNA transcriptome or proteome.

However, an advantage of cfDNA is that it can be analyzed from bio-banked biofluids, such as frozen plasma. In addition, a direct comparison of mutation detection on cfDNA vs. CTCs showed a higher abundance of the mutation on the cfDNA from the same patient (39). Finally, recent large studies comparing the effectiveness of cfDNA analysis to tissue biopsy in NSCLC showed the clinical value of the liquid biopsy approach (46). This positive result led to an approval to use cfDNA analysis for EGFR mutation analysis for IRESSA® in Europe (in patients where a tumor sample was not evaluable), making it the first EGFR tyrosine kinase inhibitor for which cfDNA testing is included in the label.

Although promising, challenges remain when using cfDNA to characterize the mutation status of a tumor. In addition to the low copy number of mutant alleles, the median half-life of cfDNA in circulation ranges from 15 minutes to a few hours (47). Also, the total concentration of cfDNA in the blood of cancer patients varies considerably (48) with tumor specific mutations ranging from undetectable (less than 1 copy per 5 mL of blood) to patients with over hundred thousand copies of the mutation per ml of plasma (39). Thus, the challenge of how to maximize the yield of the cfDNA and pair this with a platform sensitive enough to detect rare variants in the background of wild-type DNA remains. Optimally, the ability to detect mutations in plasma should not be limited to a subpopulation of patients with very high mutant copy numbers in circulation. While many analytical platforms report the mutation load with an allelic frequency compared to the wild-type DNA, platforms relying solely on the allelic frequency without recording the number of mutations have limitations. The allelic frequency is affected by the amount of wild-type DNA not related to the tumor. Therefore, it is important to consider the processes that affect the amount of wild-type DNA in circulation. For example, exercise increases cfDNA levels 10-fold (49) and other pre-analytical variables such as blood collection and extraction protocols affect the amount and size range of cfDNA fragments in a sample (50). Delays in blood processing, blood storage
temperature, agitation of the sample and shipment can all cause wild-type cfDNA release from lysed nucleated blood cells and effect the allelic frequency (51). For the same reason, plasma is often preferred over serum because of the potential for cell lysis during blood coagulation (52).

Exosomes

The exosome field has grown exponentially the last few years impacting various areas of research. Studies demonstrating that exosomes are actively released vesicles (carrying RNA, DNA and protein) and can function as inter-cellular messengers, have contributed to their elevated recognition in the scientific community (53-64). A recent review outlining the biological properties of exosomes and other extracellular vesicles (EVs) highlights these developments (65). However, with respect to nomenclature, the exosome field still lags behind as the definition and characterization of EV types are not yet firmly established (66). The majority of exosomes range in size from 30-200 nanometer in diameter and are isolated from all biofluids, including serum (60), plasma, saliva, urine and cerebrospinal fluid (67).

Exosomes and other EVs are particularly interesting as cancer biomarkers since they are stable carriers of genetic material and proteins from their cell of origin. They are also thought to be part of the disease process, for example, tumor exosomes have been shown to stimulate tumor cells growth, suppress the immune response and induce angiogenesis (60,68) and even be part of the metastatic process (63,69). Exosome release is also an active process and tumor cells can shed tens of thousands of vesicles per day resulting in hundreds of billions of vesicles per mL of plasma (55). The two mechanisms by which exosomes are released, either involve the formation of multivesicular bodies (MVB) and direct budding at the plasma membrane, or a process more akin to a retrovirus particle leaving the cell (Figure 1) (70).

In the early decades of exosome research, it was thought that they contained only protein and lipids. However, it has since been shown that exosomes are highly stable packages of RNA from the cell of origin (61). The finding that exosomes contain RNA with tumor specific mutations, can be isolated from biofluid samples and stored for many years in the freezer has opened up new opportunities in the field of diagnostics (60,71). Recent publications have also examined the DNA associated with exosomes and shown its utility for detection of gene amplifications as well as mutations (55,64,72).

Due to the size of an exosome, on average just over 100 nanometers, the entire transcriptome cannot be packaged inside every vesicle. By way of comparison, retrovirus particles with a similar size can package only around 10 kb (73), so it is likely that a single vesicle of that size carries only a limited number of transcripts. However, exosomes are extremely abundant \( \times 10^{11} \) per mL of plasma and when isolating the vesicle fraction, most of the transcriptome can be detected (74). Exosomal RNA can be used for mutation detection (55,60,71,72) as well as global profiling of most types of RNA (74), and the profile alone (without mutation characterization) can be utilized for diagnostics (58,75,76).
The precipitous release of exosomes by cancer cells seems to correspond to activation of the mitogen-activated protein kinases (MAPK) pathway frequently upregulated in tumor cells (77). Tumor derived mutations can be detected in exosomes from cerebrospinal fluid (67), serum (60), plasma (64) as well as in urine (71). However, as exosomes are released by all cells, they are particularly useful to profile not only mutations in cancer but also RNA profiles in inflammatory (78), metabolic (79), cardiovascular (80), neurodegenerative (81) and other disease processes.

Exosomes also carry surface markers from the cell of origin, which can be used for enrichment strategies, similar to CTCs (75). For example, characterization and analysis of exosome surface proteins hold great promise for the ability to identify, separate, sort and enrich exosomes originating from diverse cell sources. While the development of methods that allow for the routine analysis of exosome surface proteins has been a challenge, a number of recent advances have demonstrated potential. Immunoaffinity-bead based capture methods, microfluidic chip methods and antibody-based exosome arrays using both label and label-free detection platforms have all successfully exploited specific exosome surface proteins. This has enabled the capture, enrichment and characterization of unique populations of exosomes in the blood of healthy donors and of patients with pancreatic cancer (82), ovarian cancer (83), lung cancer (84,85). Surface protein-based exosome isolation methods combined with exosomal RNA extraction and qPCR detection assays have proven to be rapid and sensitive enough to monitor therapeutic response and resistance using exosomes from the blood of patients with glioblastoma (86,87).

In addition, the rapid advancement of a novel method of nanoscale fluorescence activated cell sorting call nanoFACS has further advanced methods of exosome isolation and sorting and allowed for the study of discrete, free, individual exosomes from body fluids (88). This technique and variants thereof hold great promise for future diagnostic applications where isolation and examination of individual exosomes is paramount. Finally, in addition to proteins, analysis of exosome protein-to-lipid ratios can be used to further isolate and characterize subpopulations of exosomes in body fluids (89).

Exosome investigations have focused on the important physiologic and pathophysiologic functions of these vesicles in micro-metastasis, angiogenesis and immune modulation (63,90) and as a means for detection of tumor specific mutations in biofluids. Consequently, in 2012, interest in this new field increased when the National Institute of Health (NIH) dedicated the large strategic Common Fund to study these new entities of extracellular RNA. The goal of this effort is to better understand how exosomes can be utilized for biomarkers and therapeutics as well as understanding this new mechanism of intercellular communication (http://commonfund.nih.gov/Exrna/index).

**Mutation detection and RNA profiling**

Analysis of nucleic acids present in bodily fluids can provide a better understanding of the disease, as summarized in Table 1.

<table>
<thead>
<tr>
<th>Analysis capability</th>
<th>Examples</th>
<th>CTCs</th>
<th>cfDNA</th>
<th>Exosomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mutations</td>
<td>Point mutations, InDels, amplifications, deletions, translocations</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Epigenetic modifications</td>
<td>Methylation patterns</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>RNA transcription profiles</td>
<td>Levels/activity of mRNA, microRNA, long non-codingRNA, RNA splice variants</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Phenotypic studies of cells from the tumor</td>
<td>Cell morphology, protein localization, in vivo studies</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Inflammatory response, stromal and other systemic changes</td>
<td>Inflammatory RNA and protein markers</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Analysis of RNA as well as DNA and protein profiles from tumor cells</td>
<td>Separate or in combination</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Can utilize biobanked samples</td>
<td>Frozen plasma, urine and other biofluids</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

CTCs, circulating tumor cells; cfDNA, cell free DNA; InDels, insertions/deletions.
has evolved, the clinical applications of liquid biopsies have improved significantly. Examples of these analytical platforms include BEAMing (13), ARMS (17), and ddPCR (15). These platforms were developed specifically for the detection of extremely rare alleles and are used when the mutation type and position is known. Other platforms such as ice-COLD PCR and targeted resequencing using NGS platforms can detect rare allelic frequencies even when the type and location of the mutation in the gene is undefined. Targeted resequencing is becoming increasingly popular since it can easily accommodate larger panels of genes to cover the actionable mutations in cancer that have significant diagnostic, prognostic or therapeutic implications for a specific therapy. Initially, the inherent error rate of NGS platforms made it difficult to identify very rare alleles (<1%), but strategies using paired-end sequencing and background correction have enabled detection of allelic frequencies at or below 0.1% (91). Incorporation of unique identifiers to each target enables highly sensitive digital sequencing capable of quantifying the number of mutated reads as well as their allelic frequency (92,93).

RNA profiling from biofluids also poses numerous challenges. However, the discovery that exosomes contained RNA made it possible to separate the fragile RNA from the large amounts of RNases and PCR inhibitors that are present in most biofluids. As cell-free RNA in blood is immediately degraded, RNAs in serum and plasma are either protected inside vesicles like an exosome, in protein complexes with the Ago2 protein (94) or associated with HDL particles (95) as outlined in Figure 2. Most of the early studies were limited to the more abundant short (~22 nt) regulatory microRNAs. The levels of these microRNAs are tightly regulated in normal cells and dysregulation has been implicated in a number of human diseases e.g., cardiovascular (96) neurological and is strongly linked to cancer development and progression as reviewed by Croce (97). However, although robust and readily detectable, microRNAs represent only a minor fraction of the transcriptome. By contrast, if the appropriate methods are used, the nucleic acids in exosomes can be isolated and the entire transcriptome interrogated.

Figure 2 Circulating nucleic acids are coming from a wide range of cellular processes. It is important to optimize the sample processing for the particular target and understand where the RNA and DNA are coming from as well as their abundance. Whole blood as well as cell free plasma has multiple sources containing nucleic acids (shown in A and B respectively). Even components that lack a nucleus (like erythrocytes and thrombocytes) have been shown to carry RNA and can have cfDNA co-isolating in the preparations. *, based on a range of 0-50% of exosome RNA containing the tumor specific mutant allele (67) (and Exosome Diagnostics unpublished data).
for effective molecular profiling and mutation detection. Successful RNA profiling from biofluids requires that the contaminants, which could inhibit downstream analysis are removed. The effective purification of the exosomes can remove these contaminants making the exosome isolation platform scalable, where the sample volume input is linear to the RNA output and not affected by the increased amount of RNases that can co-purify (98). This feature is important, since scaling the volume appropriately will enable profiling also of low copy number RNAs.

Finally, special precautions need to be taken to prevent degradation during the RNA extraction procedure, as the RNA purified from exosomes and the microRNAs from Ago2 complexes will now be exposed to RNases. Measuring integrity using an exogenous spiked-in sequence of similar size and structure as the RNAs in the exosomes is recommended. Ideally, the ‘spike’ should itself be protected from RNases, for example using a synthetic vesicle added directly into the biofluid as opposed to the lysed sample.

**Discussion**

The most obvious hurdle for all forms of liquid biopsy remains the relative rarity of nucleic acid derived from a tumor against the background of normal material found in most patient samples. In fact, the majority of cell, cell free nucleic acids, microRNAs and exosomes in a liquid biopsy will have originated from normal cells with numbers fluctuating as a consequence of biological variations. Such challenges are addressed using the strategies highlighted in the methods described above. These methods are currently sensitive enough to detect very rare mutation events. However, it is critical that laboratories undertaking such methods must be scrupulous in their methodologies to avoid erroneous results. Although clichéd, the analogy of a needle in a haystack applies and is appropriate for each of these approaches.

The analysis of CTCs and exosome has benefited from developments in the field of enrichment prior to the analytical readout. While still at an early stage, a number of studies have demonstrated that protein-based isolation and enrichment methods will be an important tool both in enhancing nucleic acid based assays and as stand-alone diagnostics in the future.

Clearly, exosomes have a number of advantages for diagnostics. They enable high quality RNA to be extracted from fresh or frozen biofluids, thus increasing the scope of detectable mutations to include mutations, splice variants, fusions as well as expression based assays for mRNA, microRNA, IncRNA and other non-coding RNAs. They are also released from living cells as an active process, whereas cfDNA is released through the process of apoptosis and necrosis. On cfDNA, all genes are present at an equal level, whereas RNA originating from a highly expressed gene could occur in thousands of copies/cell. However, as mutations exist on both exosome RNA (living process) and cfDNA (dying process), utilizing a platform that can use both will have obvious advantages for detecting rare mutations. This is especially true in the case of patients who do not have an abundant amount of mutated nucleic acid in circulation.

Improvements to analytical sensitivity and specificity will address some of the current hurdles, for example, cancer patients who have very few mutations in their biofluids, likely due to biology rather than analytical sensitivity. In many cases, the mutated alleles can occur at less than 1 copy per mL of plasma. So, combining exosome RNA and cfDNA has the advantage of increasing the detection sensitivity for low frequency mutations.

For the patient there is an obvious and clear advantage to a liquid biopsy in comparison to conventional surgical methods. However, most of the studies to date have focused on detection of actionable mutations in biofluids, and this is arguably only a fraction of the capability of liquid biopsies in enabling personalized medicine. As DNA mutations will only inform of some aspects of the disease, looking at RNA expression in biofluids can help further understand processes within the cancer patient.

Cancer is a complex and dynamic disease that can change quickly. To fully deliver on the promise of personalized medicine, development of reliable and robust non-invasive platforms for the diagnosis, patient stratification and to monitor treatment response are paramount. The various liquid biopsy platforms described in this review have the potential to add tremendous value to the care of cancer patients.

**Acknowledgements**

**Disclosure:** The authors are employees of Exosome Diagnostics.

**References**

2. Kim ES, Hirsh V, Mok T, et al. Gefitinib versus docetaxel in...


57. Miranda KC, Bond DT, McKee M, et al. Nucleic acids within urinary exosomes/microvesicles are potential...


