



# Signatures Beyond Oncogenic Mutations in Cell-Free DNA Sequencing for Non-Invasive, Early Detection of Cancer

Subhajoti De\*

Rutgers Cancer Institute of New Jersey, New Brunswick, NJ, United States

Early detection of cancer saves lives, but an effective detection strategy in public health settings requires a delicate balance - periodic screening should neither miss rapidly progressing disease nor fail to detect rare tumors at unusual locations; on the other hand, even a modest false positive rate carries risks of over-diagnosis and over-treatment of relatively indolent non-malignant disease. Genomic profiling of cell-free DNA from liquid biopsy using massively parallel sequencing is emerging as an attractive, non-invasive screening platform for sensitive detection of multiple types of cancer in a single assay. Genomic data from cell-free DNA can not only identify oncogenic mutation status, but also additional molecular signatures related to potential tissue of origin, the extent of clonal growth, and malignant disease states. Utilization of the full potential of the molecular signatures from cfDNA sequencing data can guide clinical management strategies for targeted follow-ups using imaging or molecular marker-based diagnostic platforms and treatment options.

**Keywords:** cancer, early detection, non-invasive, cell-free DNA, sequencing

## OPEN ACCESS

### Edited by:

Anupama Yadav,  
Dana-Farber Cancer Institute,  
United States

### Reviewed by:

Ramakrishna Sompallae,  
The University of Iowa, United States  
Anindya Dey,  
University of Oklahoma Health  
Sciences Center, United States

### \*Correspondence:

Subhajoti De  
subhajoti.de@rutgers.edu

### Specialty section:

This article was submitted to  
Human and Medical Genomics,  
a section of the journal  
Frontiers in Genetics

**Received:** 17 August 2021

**Accepted:** 13 September 2021

**Published:** 14 October 2021

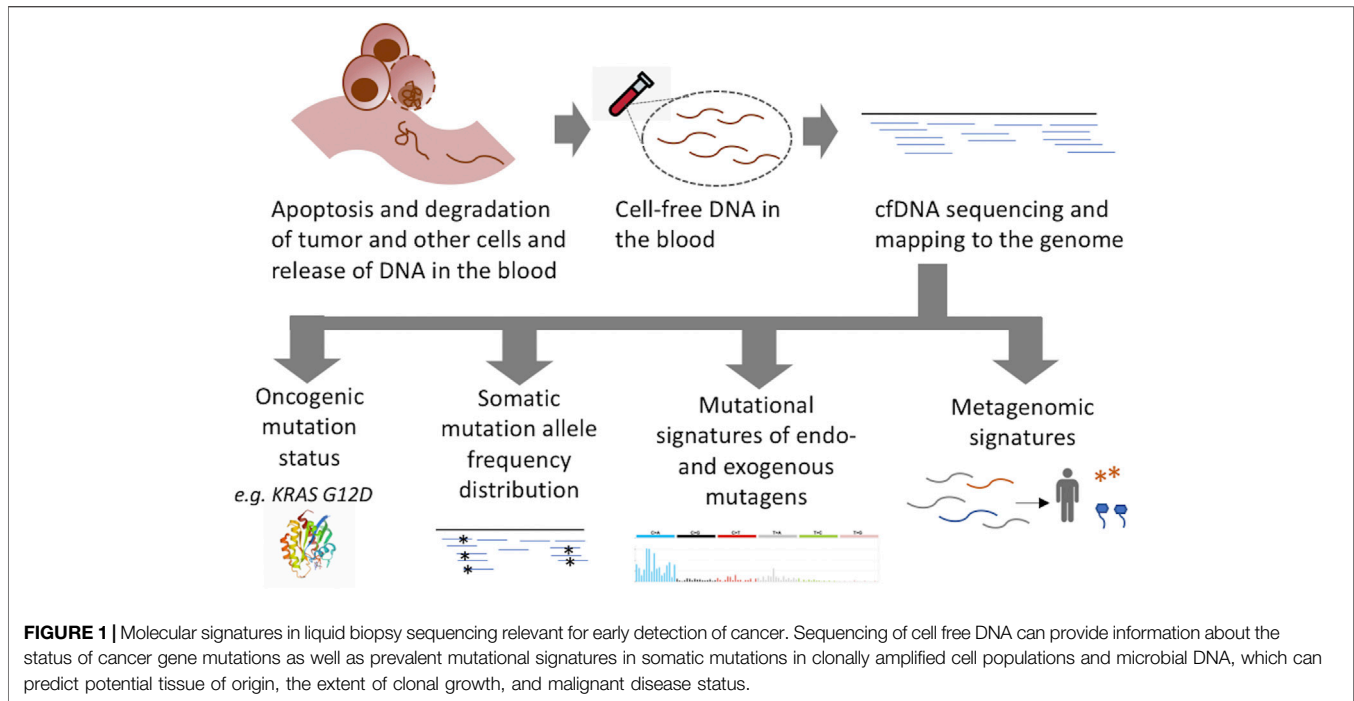
### Citation:

De S (2021) Signatures Beyond  
Oncogenic Mutations in Cell-Free DNA  
Sequencing for Non-Invasive, Early  
Detection of Cancer.  
Front. Genet. 12:759832.  
doi: 10.3389/fgene.2021.759832

## INTRODUCTION

Early cancer diagnosis significantly reduces cancer mortality and improves the chances of a favorable response to treatment (Campbell et al., 2016). Currently, the routine early screening for early detection is cancer type-specific and typically involves low dose computed tomography (CT), X-ray, or testing for cancer-specific molecular markers (e.g. tumor-specific antigen); some of these platforms have moderate accuracy and limited predictive power. For instance, In the National Lung Screening Trial (NLST) involving 53,454 persons at high risk for lung cancer at 33 U.S. medical centers, screening with low-dose computed tomography (CT) had a 96.4% false-positive rate. Unfortunately, a high false-positive rate translates to increased risks of over-diagnosis and over-treatment of relatively indolent non-malignant disease and reduces public confidence. Furthermore, a majority of these screening platforms are cancer-type specific, such that a negative screening result for a given cancer type does not eliminate the risk of undetected early malignancy elsewhere in the body. Therefore, an accurate, ideally non-invasive, pan-cancer screening strategy has been a long-standing interest in the clinical cancer research community, and can fundamentally benefit clinical management strategies for early detection of cancer.

Tumors release circulating tumor cells, cell-free DNA, and other cellular contents in blood and tissue-relevant body fluids (e.g. saliva, urine, and sputum for oral, bladder, and lung cancer



respectively), and recent advances in genomics and multi-omics approaches have enabled sensitive detection of tumor-derived cell-free DNA with high specificity and sensitivity. As a major connective tissue, blood carries cell-free DNA and other molecules throughout the body, such that tumor-derived cell-free DNA for a majority of common cancer types can be detected in blood from a liquid biopsy. Indeed, next-generation sequencing-based molecular profiling of cell-free DNA (cfDNA) in blood has emerged as a promising, non-invasive method for advanced, high precision assessment of multi-cancer risk in a single assay (Crowley et al., 2013; Wan et al., 2017). Notably, in some cases, cfDNA profiling has detected cancer recurrence several weeks, months, or even years in advance relative to the current standard of care (Abbosh et al., 2017; Chen et al., 2020), which has far-reaching implications for early detection and clinical management.

cfDNA profiling for tracking cancer progression in clinical genomic settings typically focuses on the detection of known oncogenic mutations using high depth sequencing of targeted cancer gene panel, and multiple FDA-approved tests are now available in the clinic (Ignatiadis et al., 2021). When the mutational makeup of the tumor is already known, it is cost-effective and ideal for tracking the progression of disease during treatment and potential relapse. However, when the genomic makeup of the tumor is unknown, as in the case of early detection, potential challenges need to be taken into consideration. First, oncogenic mutations are not unique to tumors; sensitive assays have routinely detected many cancer gene mutations in non-malignant tissues, especially in older individuals (Aghili et al., 2014; Alborelli et al., 2019; Liggett et al., 2019). Second, sequential accumulation of oncogenic mutations is a continuous process during pre-malignant and malignant transformation, such that

oncogenic mutation status may not always identify the tipping point in cancer progression and/or the initiation of a malignant disease (De and Ganesan, 2017; Srivastava et al., 2018). Third, many oncogenic mutations are not cancer type-specific (Lawrence et al., 2013), and hence oncogenic mutation status alone may not uniquely identify the underlying cancer type. However, Genomic data from cell-free DNA also carries molecular signatures related to the cell of origin, malignant disease states, and the extent of neoplastic growth (Figure 1), which can complement information regarding oncogenic mutations for targeted follow-ups.

## Molecular Signatures in Cell-free DNA

Tumor and non-tumor tissues release cellular material including DNA fragments in the blood. cfDNA carries several attributes that are informative about the source tissue contexts, such that systematic analysis of those properties can provide further information about the disease potential and tissue of origin. First, a tumor genome has  $10^3$ – $10^5$  somatic mutations, only a small minority of which are classic oncogenic mutations, while the others are passenger mutations – which are typically classified as the variants of unknown significance (VUS). Since cancer is a disease marked by clonal growth, the somatic mutations of malignant origin are likely to be present in clonally amplified cell populations (Podlaha et al., 2012); therefore unlike the baseline somatic mutations from other, non-tumor tissues, the cancer-associated mutations are likely to be present at a relatively higher allele frequency in cfDNA. Somatic mosaicism (Moss et al., 2018) i.e., presence of clonal growth in healthy somatic tissues, low tumor DNA fraction in cfDNA, as well as the low depth of coverage of cfDNA sequencing could potentially complicate the inferences. Nonetheless, the VUS vastly

outnumber the oncogenic mutations, such that a joint analysis of allele frequency distributions of the VUS and oncogenic driver mutations from cfDNA sequencing data can provide increased statistical power to predict the presence of clonally amplified cell populations carrying the observed set of oncogenic mutations in the patient's blood.

Second, the somatic mutations typically carry mutational signatures of tissue-dependent carcinogenic exposure and DNA repair defects. For instance, DNA derived from lung tissues e.g. in the liquid biopsy fraction typically carries mutational signatures of smoking and other pollutants (Abbosh et al., 2017). Since a liquid biopsy sample carries an admixture of DNA fragments derived from potential tumor tissues as well as hematopoietic and other non-malignant tissues (Hu et al., 2020), which are present even in healthy individuals, an analysis of the mutational signatures in cfDNA can help identify the presence of nucleic acid contents from unexpected tissue types (Gerstung et al., 2020). In addition, cfDNA fragments in the blood typically are 120–220 bp, or multiples thereof, with a maximum peak at 167 bp – which corresponds to the length of DNA wrapped around a single nucleosome in the source tissue, and therefore the genome-wide landscape of cfDNA fragments provides information about epigenome of the cell of origin.

Third, although cancer progression is a gradual process, often progressing through localized dysplasia and hyperplasia stages for years before neoplastic transformation, not all clonal growths progress to malignant disease (Podlaha et al., 2012; Gerstung et al., 2020; Singh et al., 2020). The neoplastic transformation typically involves overcoming replicative senescence and genomic instability, accompanied by invasive clonal growth. Genomic instability leads to distinct mutational signatures in somatic mutations, which are uncommon in normal tissues. For instance, the mutational signature SBS3, SBS8, and SBS40, which are associated with replication stress and genomic instability (Maley et al., 2017; Alexandrov et al., 2020; Gerstung et al., 2020) – rarely occur in normal or benign tissues, but are widely present in cancer genomes. Collectively, the presence of oncogenic mutations, as well as the somatic mutations with distinct mutational signatures and inferred chromatin makeups in cfDNA can help identify potentially malignant clonal growth in a narrow set of probable tissue types with significant contributions in the cfDNA.

Fourth, it is increasingly becoming evident that the tumor microenvironment does not only include tumor, immune, and stromal cells, but also diverse microbial entities including bacteria, fungi, and viruses (Poore et al., 2020). While earlier works identified tumor-associated microbiome predominantly in selected tumor types, recent studies suggest that tumor-associated microbiome might be present in most cancer types (Sr et al., 2020). Furthermore, microbial DNA from the tumor microenvironment could be detected alongside DNA from host tissues during unbiased total cfDNA sequencing from a liquid biopsy. Since different cancer tissue types have distinct microbial makeups (Sr et al., 2020), annotation of microbial DNA could provide additional clues about the potential sources of

genetic material in blood. Indeed, it seems possible to correctly predict tumor types from microbial signatures in blood alone (Sr et al., 2020). Therefore, total DNA sequencing from blood can provide orthogonal metagenomic signatures characteristics of different cancer types, adding yet another line of evidence for the detection of the disease.

While the above attributes could be derived from the standard cfDNA sequencing, additional insights could be gained from profiling microRNA and DNA methylation states, and other biomarkers, which typically require dedicated assays. Taken together, in addition to the detection of oncogenic mutations the emerging molecular signatures in cfDNA sequencing (Figure 1) can provide multi-faceted insights into tumor etiology and disease status, and guide imaging or molecular marker-based targeted follow-ups and treatment options.

## Advances in cfDNA Sequencing and Genomic Analysis

A number of genomic approaches and computational methods have been developed for sensitive detection of somatic mutations and other molecular signatures from cfDNA sequencing. Recent developments have increased the yield and quality of cfDNA through a rational design (De, 2011), and it has been possible to isolate tumor-derived DNA from diverse types of body fluids including sputum, saliva, and urine to inform about cancer progression (Ignatiadis et al., 2021; Jj et al., 2020). In parallel, using specialized genomic and computational approaches it is now possible to detect somatic mutations at an ultra-low allele frequency (Mw et al., 2016; Sr et al., 2020). Moreover, integrative analysis of driver and passenger somatic mutations from cfDNA has been able to infer clonal and subclonal mutations, and also clonal architecture in the primary tumor for some patients (Abbosh et al., 2017)–which opens up the possibility of tracking clonal dynamics in vivo in real-time. Elegant computational analyses of cfDNA fragment patterns in patient samples have provided quantitative estimates of the likely tissue of origin of cfDNA (Schwarzenbach et al., 2011). More recently, Poore et al. used an innovative approach to identify cancer types based on microbial signatures in blood (Cc et al., 2020), which has significance for cancer diagnostics. Collectively, these and other advances provide a rich toolbox to track cancer progression non-invasively from blood, even at an early stage of the disease and recurrence.

## DISCUSSION

Genomic profiling of cfDNA for diagnostic purposes, especially at an early stage of cancer is not trivial. Tumor DNA fraction in cfDNA in blood might be very low during the early non-invasive stage of the disease and can vary considerably, depending on the location of the tumor in the tissue contexts. It also remains challenging to assign the cell of origin to individual cfDNA molecules, potentially complicating etiological inferences. Large cohorts are required to assess the cancer type-specific accuracy of

cfDNA profiling in the early detection setting. In any case, the benefits appear to outweigh the potential limitations. The technologies are maturing rapidly, and non-invasive cancer progression using cfDNA profiling has already moved from bench to bedside. A number of FDA-approved providers routinely offer services, such as the Guardant360 and FoundationOne Liquid CDx platforms to track cancer progression using liquid biopsy sequencing. This year, GRAIL Inc. has launched the Galleri platform under a CLIA-waiver and can be used to screen people of age 50 years or older at elevated risk of cancer.

While the current focus remains on the detection of somatic mutations in liquid biopsy, the signatures discussed above could be extracted from the clinical genomics workflow for cfDNA sequencing without additional sample preparation and processing and can provide additional insights that could further guide clinical decision-making. While liquid biopsy-based tracking of cancer progression is already integrated into the clinical management for cancer patients, leveraging it for non-invasive screening for early detection of pan-cancer risk at a low cost in a standardized setting will have important implications for clinical management of high-risk patient population and public health at large.

## REFERENCES

- Abbosh, C., Birkbak, N. J., Wilson, G. A., Jamal-Hanjani, M., Constantin, T., Salari, R., et al. (2017). Phylogenetic ctDNA Analysis Depicts Early-Stage Lung Cancer Evolution. *Nature* 545, 446–451. doi:10.1038/nature22364
- Aghili, L., Foo, J., DeGregori, J., and De, S. (2014). Patterns of Somatic Acquired Amplifications and Deletions in Apparently normal Tissues of Ovarian Cancer Patients. *Cell Rep.* 7, 1310–1319. doi:10.1016/j.celrep.2014.03.071
- Alborelli, I., Generali, D., Jermann, P., Cappelletti, M. R., Ferrero, G., Scaggiante, B., et al. (2019). Cell-free DNA Analysis in Healthy Individuals by Next-Generation Sequencing: a Proof of Concept and Technical Validation Study. *Cell Death Dis.* 10, 534–611. doi:10.1038/s41419-019-1770-3
- Alexandrov, L. B., au, fnm., Kim, J., Haradhvala, N. J., Huang, M. N., Tian Ng, A. W., et al. (2020). The Repertoire of Mutational Signatures in Human Cancer. *Nature* 578, 94–101. doi:10.1038/s41586-020-1943-3
- Campbell, J. D., Mazzilli, S. A., Reid, M. E., Dhillon, S. S., Platero, S., Beane, J., et al. (2016). The Case for a Pre-cancer Genome Atlas (PCGA). *Cancer Prev. Res.* 9, 119–124. doi:10.1158/1940-6207.capr-16-0024
- Cc, V., Rr, Y., Mr, F., Lindsey N, W., Tan, Li., Jesse J, S., et al. (2020). Direct Quantification of In Vivo Mutagenesis and Carcinogenesis Using Duplex Sequencing. *Proc. Natl. Acad. Sci. U. S. A.* 117, 33414–33425. doi:10.1073/pnas.2014145117
- Chen, X., Gole, J., Gore, A., He, Q., Lu, M., Min, J., et al. (2020). Non-invasive Early Detection of Cancer Four Years before Conventional Diagnosis Using a Blood Test. *Nat. Commun.* 11, 3475–3510. doi:10.1038/s41467-020-17316-z
- Crowley, E., Di Nicolantonio, F., Loupakis, F., and Bardelli, A. (2013). Liquid Biopsy: Monitoring Cancer-Genetics in the Blood. *Nat. Rev. Clin. Oncol.* 10, 472–484. doi:10.1038/nrclinonc.2013.110
- De, S. (2011). Somatic Mosaicism in Healthy Human Tissues. *Trends Genet.* 27, 217–223.
- De, S., and Ganesan, S. (2017). Looking beyond Drivers and Passengers in Cancer Genome Sequencing Data. *Ann. Oncol.* 28, 938–945. doi:10.1093/annonc/mdw677
- Gerstung, M., au, fnm., Jolly, C., Leshchiner, I., Dentre, S. C., Gonzalez, S., et al. (2020). The Evolutionary History of 2,658 Cancers. *Nature* 578, 122–128. doi:10.1038/s41586-019-1907-7
- Hu, X., Xu, Z., and De, S. (2020). Characteristics of Mutational Signatures of Unknown Etiology. *NAR Cancer* 2, zcaa026. doi:10.1093/narcan/zcaa026

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

## AUTHOR CONTRIBUTIONS

SD conceived the project, wrote the manuscript, and approved the submitted version.

## FUNDING

SD acknowledges funding from the NIH (R01GM129066 and R21CA248122).

## ACKNOWLEDGMENTS

SD thanks members of Rutgers Cancer Institute of New Jersey for helpful discussion.

- Ignatiadis, M., Sledge, G. W., and Jeffrey, S. S. (2021). Liquid Biopsy Enters the Clinic - Implementation Issues and Future Challenges. *Nat. Rev. Clin. Oncol.* 18, 297–312. doi:10.1038/s41571-020-00457-x
- Jj, C., Eg, H., Dm, K., Ms, E., Ej, M., Schroers-Martin, J., et al. (2020). Integrating Genomic Features for Non-invasive Early Lung Cancer Detection. *Nature* 580, 245–251.
- Lawrence, M. S., Stojanov, P., Polak, P., Kryukov, G. V., Cibulskis, K., Sivachenko, A., et al. (2013). Mutational Heterogeneity in Cancer and the Search for New Cancer-Associated Genes. *Nature* 499, 214–218. doi:10.1038/nature12213
- Liggett, A., Sharma, A., De, S., and DeGregori, J. (2019). FERMI: A Novel Method for Sensitive Detection of Rare Mutations in Somatic Tissue. *G3 Genes Genomes Genet.* 9, 2977–2987. doi:10.1534/g3.119.400438
- Maley, C. C., Aktipis, A., Graham, T. A., Sottoriva, A., Boddy, A. M., Janiszewska, M., et al. (2017). Classifying the Evolutionary and Ecological Features of Neoplasms. *Nat. Rev. Cancer* 17, 605–619. doi:10.1038/nrc.2017.69
- Moss, J., Magenheimer, J., Neiman, D., Zemmour, H., Loyfer, N., Korach, A., et al. (2018). Comprehensive Human Cell-type Methylation Atlas Reveals Origins of Circulating Cell-free DNA in Health and Disease. *Nat. Commun.* 9, 5068–5112. doi:10.1038/s41467-018-07466-6
- Mw, S., Kircher, M., Aj, H., Rm, D., and J, S. (2016). Cell-free DNA Comprises an In Vivo Nucleosome Footprint that Informs its Tissues-Of-Origin. *Cell* 164, 57–68.
- Podlaha, O., Riester, M., De, S., and Michor, F. (2012). Evolution of the Cancer Genome. *Trends Genet.* 28, 155–163. doi:10.1016/j.tig.2012.01.003
- Poore, G. D., Kopylova, E., Zhu, Q., Carpenter, C., Fraraccio, S., Wandro, S., et al. (2020). Microbiome Analyses of Blood and Tissues Suggest Cancer Diagnostic Approach. *Nature* 579, 567–574. doi:10.1038/s41586-020-2095-1
- Schwarzenbach, H., Hoon, D. S. B., and Pantel, K. (2011). Cell-free Nucleic Acids as Biomarkers in Cancer Patients. *Nat. Rev. Cancer* 11, 426–437. doi:10.1038/nrc3066
- Singh, V. K., Rastogi, A., Hu, X., Wang, Y., and De, S. (2020). Mutational Signature SBS8 Predominantly Arises Due to Late Replication Errors in Cancer. *Commun. Biol.* 3, 421. doi:10.1038/s42003-020-01119-5
- Sr, G., Kb, E., Parpart-Li, S., Murtaza, M., Abel J, B., Mark D, P., et al. (2020). Harmonizing Cell-free DNA Collection and Processing Practices through Evidence-Based Guidance. *Clin. Cancer Res.* 26, 3104–3109.

- Srivastava, S., Ghosh, S., Kagan, J., Mazurchuk, R., Boja, E., Chuaqui, R., et al. (2018). The Making of a PreCancer Atlas: Promises, Challenges, and Opportunities. *Trends Cancer* 4, 523–536. doi:10.1016/j.trecan.2018.06.007
- Wan, J. C. M., Massie, C., Garcia-Corbacho, J., Mouliere, F., Brenton, J. D., Caldas, C., et al. (2017). Liquid Biopsies Come of Age: towards Implementation of Circulating Tumour DNA. *Nat. Rev. Cancer* 17, 223–238. doi:10.1038/nrc.2017.7

**Conflict of Interest:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

*Copyright © 2021 De. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*