

# What fraction of cellular DNA turnover becomes cfDNA?

Ron Sender, Elad Noor, Ron Milo , Yuval Dor 

Weizmann Institute of Science, Rehovot 7610001, Israel • Department of Developmental Biology and Cancer Research, Institute for Medical Research Israel Canada, the Hebrew University-Hadassah Medical School; Jerusalem, Israel

 [https://en.wikipedia.org/wiki/Open\\_access](https://en.wikipedia.org/wiki/Open_access)

 <https://creativecommons.org/licenses/by/4.0/>

## Reviewed Preprint

Published from the original preprint after peer review and assessment by eLife.

[About eLife's process](#)

## Reviewed preprint posted

August 24, 2023 (this version)

## Posted to bioRxiv

June 6, 2023

## Sent for peer review

May 26, 2023

## Abstract

Cell-free DNA (cfDNA) tests use small amounts of DNA in the bloodstream as biomarkers. While it is thought that cfDNA is largely released by dying cells, the proportion of dying cells' DNA that reaches the bloodstream is unknown. Here we integrate estimates of cellular turnover rates to calculate the expected amount of cfDNA. By comparing this to the actual amount of cell type-specific cfDNA, we estimate the proportion of DNA reaching plasma as cfDNA. We demonstrate that <10% of the DNA from dying cells is detectable in plasma, and the ratios of measured to expected cfDNA levels vary a thousand-fold among cell types, often reaching well below 0.1%. The analysis suggests that local clearance, presumably via phagocytosis, takes up most of the dying cells' DNA. Insights into the underlying mechanism may help to understand the physiological significance of cfDNA and improve the sensitivity of liquid biopsies.

### eLife assessment

This **important** study makes a bold step towards understanding what fraction of DNA that is liberated from different tissues in a healthy human is found in circulation as cell-free DNA. Unfortunately, the evidence for the conclusions is presently **incomplete**, but with additional controls, this could become a major achievement for reference in understanding changes in cell-free DNA in disease states.

## Introduction

The human body is in a constant state of cellular turnover, with an estimated  $0.3 \times 10^{12}$  cells being replaced daily, two-thirds of which are erythrocytes (Sender and Milo, 2021 ). Apoptosis is the primary mechanism of cell death in tissues with rapid turnover, in which cells are disposed of orderly and engulfed by phagocytic cells to recycle their resources.

Apoptosis involves the fragmentation of DNA through two distinct mechanisms. The first mechanism occurs within the apoptotic cell, where endonucleases break down chromatin into nucleosomal units. The second mechanism is carried out by phagocytic cells, which engulf and degrade the DNA of apoptotic cells in order to prevent the release of potentially immunogenic intracellular materials (Nagata, 2005 .

It has been known since the late 1940s that cell-free DNA (cfDNA) fragments can be found in the circulation of healthy and diseased individuals (Bendich et al., 1965 [↗](#); Mandel and Metais, 1948 [↗](#)). The fragments are typically nucleosome-size (165 base pairs), likely representing molecules that were not completely degraded during the process of cell death. With the advent of next-generation sequencing (NGS), cfDNA has become a clinically-useful biomarker for various applications. These include non-invasive prenatal testing (via detection of fetal chromosomal abnormalities via sampling of maternal cfDNA), cancer monitoring (via detection of oncogenic mutations in plasma, termed circulating tumor DNA [ctDNA]), and monitoring of allogeneic organ transplants (via detection of donor-derived SNPs in cfDNA) (Heitzer et al., 2019 [↗](#)). Perhaps the greatest promise of cfDNA is in cancer diagnostics – a blood test that can allow early detection at an actionable stage, real-time assessment of treatment response, detection of recurrence, and identification of specific genetic mutations to inform treatment decisions (Jamshidi et al., 2022 [↗](#); Wan et al., 2017 [↗](#)).

Regardless of somatic mutations, the mere presence of cfDNA from a given tissue is of great value as it often correlates with tissue-specific injury (Gala-Lopez et al., 2018 [↗](#); Heitzer et al., 2020 [↗](#); Lehmann-Werman et al., 2018 [↗](#), 2016 [↗](#); Zemmour et al., 2018 [↗](#)). Multiple layers of epigenetic information allow the inference of the tissue origins of cfDNA. For example, the size, fragmentation patterns, and exact end position of cfDNA fragments, nucleosome positions reflected in the relative abundance of promoter sequences, and histone modification patterns all allow tracing cfDNA molecules to their tissue origin (Lo et al., 2021 [↗](#); Moss et al., 2018 [↗](#); Oberhofer et al., 2022 [↗](#); Zhou et al., 2022 [↗](#)). One particularly sensitive approach is using DNA methylation patterns, a stable determinant of cell identity preserved on cfDNA. Deconvolution of cfDNA methylomes using a reference atlas of human cell type-specific methylomes has revealed various tissues' relative and absolute contribution to cfDNA in health and disease. Under baseline conditions, over 90% of cfDNA originates in blood cells (neutrophils, megakaryocytes, monocytes, lymphocytes, and erythroblasts), with vascular endothelial cells and hepatocytes being the only solid tissue source (Loyfer et al., 2023 [↗](#)). In both homeostatic and pathologic conditions, the exact mechanism by which cfDNA is released is not fully understood but is thought to involve cell death. Whether cfDNA can be released from cells that remain alive after the event is controversial (Stroun et al., 2001 [↗](#)). Two striking examples of such a scenario are megakaryocytes and erythroblasts, whose physiological function is to release anuclear cells, namely platelets and erythrocytes (Moss et al., 2022 [↗](#)).

From a practical perspective, the amount of cfDNA (typically ~1000 genome equivalents per ml of plasma) is a major barrier to a sensitive diagnosis of diseases – particularly cancer – at an early stage. Beyond maximization of the volume of blood drawn and the number of markers tested in parallel, understanding and eventually manipulating the local release and systemic clearance of cfDNA hold great potential for improving the sensitivity of tests. For example, recent studies have suggested pharmacologic approaches for blocking the removal of cfDNA from the systemic circulation, leading to a transient elevation in cfDNA concentration (Tabrizi et al., 2023 [↗](#)). The efficiency and determinants of local cfDNA release to circulation have not been examined.

In this study, we use recent estimates of cellular turnover rates (Sender and Milo, 2021 [↗](#)) to calculate the expected amount of DNA resulting from cell death from each cell type at a given time. By comparing this to the amount of cell type-specific cfDNA present in the plasma, taking into account estimates of systemic cfDNA clearance rate, we estimate the fraction of DNA that reaches the plasma as cfDNA.

## Cellular turnover

Normal cellular turnover data for all cell types except hepatocytes and megakaryocytes were obtained from estimates provided by Sender & Milo, 2021 [\[1\]](#). We estimated the cellular turnover of megakaryocytes in two ways. First, using the number of megakaryocytes in the bone marrow (Harrison, 1962 [\[2\]](#); Noetzli et al., 2019 [\[3\]](#)) maturation time of around five days (Machlus and Italiano, 2013). Second, using the production of platelets (Harker and Finch, 1969 [\[4\]](#)) and an average number of platelets produced per megakaryocyte (Harker and Finch, 1969 [\[4\]](#); Kaufman et al., 1965 [\[5\]](#); Trowbridge et al., 1984 [\[6\]](#)). Cellular turnover of hepatocytes was calculated based on (Heinke et al., 2022 [\[7\]](#)) by combining estimates for the number of cells and the death rate for the different ploidy groups (See Dataset 1).

## Tissue-specific cfDNA concentration

The concentrations of plasma cfDNA derived from specific cell types were obtained from two studies that used deconvolution of the plasma methylome using a human cell type methylation atlas (Loyfer et al., 2023 [\[8\]](#); Moss et al., 2018 [\[9\]](#)).

## Estimation of the potential DNA flux

We estimated the potential cfDNA plasma levels if all the DNA from the dying cells had reached the bloodstream. Our estimate utilized the calculated cellular turnover rate and data regarding the ploidy of the cells, the volume of blood plasma, and the half-life of cfDNA molecules in the blood. For each cell type, we defined the cellular turnover in units of cells per day  $d$  and the ploidy (average number of sets of chromosomes)  $p_c$ . We used blood plasma volume  $V = 3L$  (ICRP, 2002 [\[10\]](#)), the mean lifespan of cfDNA molecules in the blood  $\tau = \frac{\text{half-life}}{\ln(2)} = 0.7h$  (Diehl et al., 2008 [\[11\]](#); Lo et al., 1999 [\[12\]](#); To et al., 2003 [\[13\]](#); Yao et al., 2016 [\[14\]](#)) and haploid genome mass  $m_h = 3.2 \cdot 10^{-12} g$  (Piovesan et al., 2019 [\[15\]](#))

The expected level of cfDNA levels was calculated according to the formula:

$$X_c = \frac{d \cdot p_c \cdot \tau}{V_{\text{plasma}}} \quad (1)$$

Where  $X_c$  is given in units of Genome equivalents/ml (Units of g/ml could be obtained by multiplication by  $m_h$ ).

Intuitively, the potential cfDNA level is obtained by calculating the amount of DNA in dying cells at a given moment (defined by the mean lifespan of cfDNA in plasma) when considering the total volume of the plasma.

Ultimately, we compared the measured cfDNA levels to the potential DNA flux estimates to determine the fraction of DNA that reaches the blood.

## Uncertainties estimate

Standard error was collected or calculated for each value used. In several cases, such as the half-life of plasma cfDNA, the value's uncertainty was big and best described as a multiplication factor of error (uncertainty of a variable with lognormal distribution). To enable simple error propagation, we transformed all errors to be expressed in terms of multiplication error, approximating linear (normal) error by a lognormal error using the formula  $f_{\text{error}} = \sqrt{\frac{\mu + \sigma}{\mu - \sigma}}$  where  $\mu$ ,  $\sigma$  are the mean and standard error in the linear (normal) model. Thus, we model the

uncertainty around the value  $\mu$  as a random variable  $\mu_x$  with a lognormal distribution with a shape parameter of  $s = \ln(f_{\text{error}})$ . By definition, the shape parameter describes the standard error of the exponential transformed random variable, defined as  $e^{\frac{x}{s}}$  (which is distributed normally).

Error propagation of multiplication of two values with multiplication error was done analytically using the formula:  $f_{x \cdot y} = e^{\sqrt{\ln(f_x)^2 + \ln(f_y)^2}}$ . The formula is based on the fact that the multiplication of two lognormal variables also distributes lognormally with a shape parameter that is equal to the root of the sum of the squares of the original shape factors.

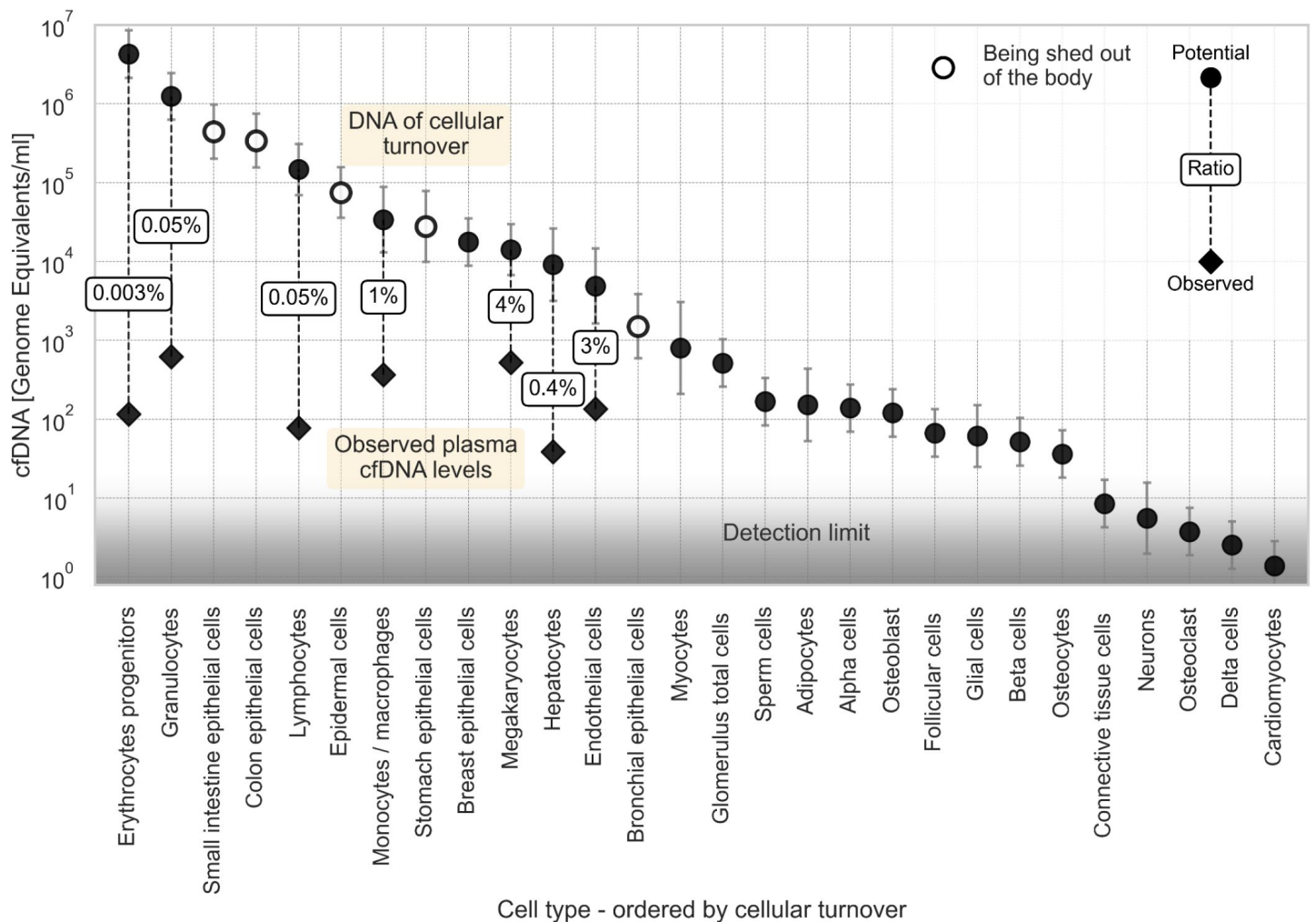
Error propagation for the summation of variables with non-linear uncertainty was calculated using bootstrapping by drawing 1000 samples from the distribution describing the uncertainties of the values.

## Results

Based on atlases of human cell type-specific methylation signatures, Moss et al. and Loyfer et al. analyzed the main cell types contributing to plasma cfDNA. They found the primary sources of plasma cfDNA to be blood cells: granulocytes, megakaryocytes, macrophages, and/or monocytes (the signature could not differentiate between the last two), lymphocytes, and erythrocyte progenitors. Other cells that had detectable contributions are endothelial cells and hepatocytes. Qualitatively, these cells represent most of the leading cell types in cellular turnover, as shown in Sender & Milo 2021 [\[1\]](#) (Sender and Milo, 2021 [\[2\]](#)). Epithelial cells of the gastrointestinal tract, the lung, and the skin are other cell types that significantly contribute to cellular turnover. Dying cells in these tissues are shed into the gut lumen, the air spaces, or out of the skin (note that while DNA from gut and lung epithelial cells can be found in stool and bronchoalveolar lavage, the fate of DNA from skin cells is not known). This arrangement may explain why DNA from these cell types is not represented in plasma cfDNA in healthy conditions. Therefore, it appears that cells with high cfDNA plasma levels are those with relatively high turnover that is not being shed out of the body.

We used the cellular turnover estimates of these cell types to calculate the potential amount of DNA discarded from each cell type. We derived the potential levels of cfDNA in the plasma (see Methods). By comparing this data to the measured levels of cfDNA in plasma (Loyfer et al., 2023 [\[3\]](#); Moss et al., 2018 [\[4\]](#)), we could calculate the fraction of potential DNA presented as cfDNA in the plasma, as illustrated in **Figure 1** [\[5\]](#). The results indicate that less than 4% of the DNA of dying cells reaches the plasma. The ratios of measured to expected cfDNA levels vary a thousand-fold, ranging from 1:30 (megakaryocytes and endothelial cells) to  $1:3 \times 10^4$  (erythrocyte progenitors).

In general, around 1000 genome equivalents of cfDNA are found in the plasma in healthy individuals. The limit for detection of a cell type of specific origin depends on the assay in use. General essays using deconvolution have a sensitivity of around 1%, e.g., ~10 genome equivalents (Loyfer et al., 2023 [\[3\]](#)). Targeted essays using markers for a specific cell type at a deep coverage can improve the sensitivity to around 0.1%, e.g., ~1, genome equivalent. The gradient in **Figure 1** [\[5\]](#) depicts this range of sensitivities. The low ratios of measured to potential cfDNA described for the mentioned cell types indicate that cells with lower cellular turnover, such as skeletal myocytes, adipocytes, and pancreatic beta cells, are not being detected in the plasma of healthy individuals because their plasma levels are lower than the sensitivities of existing essays. Notably, a comparison of potential cfDNA plasma levels of breast epithelial cells in healthy women to the limit of detection reveals the breast as an outlier. This might suggest that dying breast epithelial cells' local DNA utilization mechanism is extremely efficient.



**Figure 1**

### cfDNA as a fraction of homeostatic cell turnover.

Estimates for the DNA flux from homeostatic cell turnover were made based on Sender & Milo, 2021 [\[1\]](#) and converted to units of potential cfDNA plasma concentration (see Methods). Cell types are ordered by their estimated cellular turnover. Empty markers represent cell types being shed out of the body upon turnover. Observed levels of cfDNA (Loyfer et al., 2023 [\[2\]](#); Moss et al., 2018 [\[3\]](#)) are shown for cell types found in the plasma, with labels depicting the ratio between potential and measured cfDNA concentrations. Circles, potential representation in cfDNA; diamonds, observed concentration of cfDNA. The assay's detection limit is presented by a gradient (from around ten genome equivalents for deconvolution assay to around one genome equivalent for targeted assays). Error bars represent the standard error of the mean of the uncertainty, approximated by a lognormal distribution (see Methods). Uncertainty regarding the measured cfDNA is smaller than the marker size.

In this study, we report a surprising, dramatic discrepancy between the measured levels of cfDNA in the plasma and the potential DNA flux from dying cells. One hypothetical explanation for that discrepancy is the limited sensitivity of typical cfDNA assays to short DNA fragments, which may contribute a significant fraction of the overall cfDNA mass. Regular cfDNA analysis shows a size distribution concentrated around a length of 165 base pairs (bp). The sizes in ctDNA vary more, but most are longer than 100 bp (Alcaide et al., 2020 [DOI](#); Udomruk et al., 2021 [DOI](#)). A recent study suggested a significant fraction of single-strand ultrashort fragments (length of 25-60 bp) (Cheng et al., 2022 [DOI](#)). However, the total amount of DNA contained in these fragments is less than that of the longer regular cfDNA fragments (Cheng et al., 2022 [DOI](#)), arguing against ultrashort fragments as an explanation for the “missing” cfDNA material.

An alternative hypothetical explanation is that most DNA of dying enters the bloodstream but is rapidly degraded or taken up. Accounting for the tissue-specific DNA concentration found in the blood, we can estimate the half-life of cfDNA in the bloodstream in that case, based on the cellular turnover rate. This calculation suggests that the half-life of cfDNA in the bloodstream should be only a few seconds to a few minutes. However, previous research using various methods (mostly the decay of fetal cfDNA in maternal plasma after birth) has shown that the half-life of DNA in the bloodstream ranges from 15-120 minutes, orders of magnitude higher than this estimate suggest (Diehl et al., 2008 [DOI](#); Lo et al., 1999 [DOI](#); To et al., 2003 [DOI](#); Yao et al., 2016 [DOI](#)). In addition, a systemic clearance mechanism cannot explain the differential representation of cfDNA from different cell types relative to their turnover rate.

Therefore, the low fraction of DNA measured as cfDNA suggests that less than a few percent of the DNA from dying cells reaches the bloodstream. An unknown mechanism utilizes the rest with a tissue-specific efficiency. A potential explanation is that tissue-resident phagocytes are degrading the DNA after apoptosis, similar to what has been shown for extruded erythroblast nuclei (Yoshida et al., 2005 [DOI](#)). Studies in mice have revealed that lysosomal DNases in macrophages play a role in cleaving chromosomal DNA during apoptosis (McIlroy et al., 2000 [DOI](#)), as well as the degradation of erythroid nuclei in erythroblastic islands (Yoshida et al., 2005 [DOI](#)). The first DNA fragmentation occurs within the apoptotic cells, resulting in nucleosomal unit multiples of about 180 bp (Matassov et al., 2004 [DOI](#)). Therefore, up to 200-bp cfDNA fragments in plasma may indicate that phagocytes have not further degraded these fragments.

A comparison between the different types of cells shows a trend in which less DNA flux from cells with higher turnover gets to the bloodstream. In particular, a tiny fraction ( $1 \text{ in } 3 \times 10^4$ ) of DNA from erythroid progenitors arrives at the plasma, indicating an extreme efficiency of the DNA recovery mechanism. Erythroid progenitors are arranged in erythroblastic islands. Up to a few tens of erythroid progenitors surround a single macrophage that collects the nuclei extruded during the erythrocyte maturation process (pyrenocytes) (Chasis and Mohandas, 2008 [DOI](#)). The amount of DNA discarded through the maturation of over 200 billion erythrocytes per day (Sender and Milo, 2021 [DOI](#)) exceeds all other sources of homeostatic discarded DNA. Our findings indicate that the organization of dedicated erythroblastic islands functions highly efficiently regarding DNA utilization. The overall trend of higher turnover resulting in a lower cfDNA to DNA flux ratio may indicate similar design principles, in which the utilization of DNA is better in tissues with higher turnover. However, our analysis is limited to only several cell types (due to cfDNA test and deconvolution sensitivities), and extrapolation to cells with lower cell turnover is problematic.

A thorough explanation for the gap between the estimated DNA flux from dying cells and the measured cfDNA data requires more research. Since macrophages play a prominent role in the phagocytosis of dying cells, we hypothesize that the local uptake of cfDNA by activated macrophages is responsible for the uptake of most DNA from dying cells. An interesting implication of this possibility is that cfDNA levels are expected to be highly sensitive to



perturbations in the local clearance mechanism. In other words, elevated levels of cfDNA from a given cell type may represent a disruption of local macrophages rather than an actual increase in the rate of cell death.

Comparing the DNA flux involved in the homeostatic cellular turnover of specific cell types to the sensitivity of cfDNA essays reveals some current limitations in the field. Cell types such as adipocytes, cardiomyocytes, and pancreatic beta cells are not represented in the cfDNA of healthy individuals. Our analysis suggests that current essays need to be more sensitive to identify the minute amount of those. Moreover, the quantitative analysis can predict potential cell types with a non-neglectable contribution to the plasma cfDNA and allow their focused study. The current analysis call for a focus on breast epithelial cells and myocytes, as their potential cfDNA levels are relatively higher than the detection limit. Previous research regarding their contribution to cfDNA has used highly sensitive essays but found no contribution in healthy individuals (Loyfer et al., 2023 [↗](#); Moss et al., 2020 [↗](#), 2018 [↗](#)). This might indicate a highly effective mechanism for the utilization of DNA from dying cells, for example, local degradation of myonuclei within the syncytium of skeletal muscle fibers.

Quantitative characterization of the abundance of macrophages concerning the cellular death rate in different tissues could improve our understanding of the DNA clearance mechanism and the role of phagocytes.

Illuminating the discrepancy between the dying cells' DNA flux and the measured cfDNA levels may open the door for research with clinical potential. The sensitivity of the assays and the amount of available DNA limit the utility of liquid biopsy, particularly in early disease detection. Better characterization of the mechanism which limits available plasma cfDNA could lead to potential interventions that increase the fraction of DNA flux arriving at the plasma and thus improve the sensitivity of liquid biopsies based on cfDNA.

## Acknowledgements

We thank Yinon Bar-On, Lior Greenspon, and Yuval Rosenberg for valuable feedback on this manuscript. Funding: This research was generously supported by the Mary and Tom Beck Canadian Center for Alternative Energy Research, the Schwartz-Reisman Collaborative Science Program, the Ullmann Family Foundation, and the Yotam Project (RM). This research was supported by grants from the Helmsley Charitable Trust, JDRF, NIDDK, and Grail (YD). Prof. Yuval Dor has filed patents on cfDNA analysis. Prof. Ron Milo is the Head of the Mary and Tom Beck Canadian Center for Alternative Energy Research and the Charles and Louise Gartner Professorial Chair incumbent.

## Data and code availability

All study data are included in the article and Dataset S1. All code is available in Jupyter notebooks at <https://gitlab.com/milo-lab-public/cfdna-and-cellular-turnover> [↗](#)

## References

- Alcaide M, Cheung M, Hillman J, Rassekh SR, Deyell RJ, Batist G, Karsan A, Wyatt AW, Johnson N, Scott DW, Morin RD (2020) **Evaluating the quantity, quality and size distribution of cell-free DNA by multiplex droplet digital PCR** *Sci Rep* **10**
- Bendich A, Wilczok T, Borenfreund E. (1965) **CIRCULATING DNA AS A POSSIBLE FACTOR IN ONCOGENESIS** *Science* **148**:374–376
- Chasis JA, Mohandas N. (2008) **Erythroblastic islands: niches for erythropoiesis** *Blood* **112**:470–478
- Cheng J, Morselli M, Huang W-L, Heo YJ, Pinheiro-Ferreira T, Li F, Wei F, Chia D, Kim Y, He H-J, Cole KD, Su W-C, Pellegrini M, Wong DTW (2022) **Plasma contains ultrashort single-stranded DNA in addition to nucleosomal cell-free DNA** *iScience* **25**
- Diehl F, Schmidt K, Choti MA, Romans K, Goodman S, Li M, Thornton K, Agrawal N, Sokoll L, Szabo SA, Kinzler KW, Vogelstein B, Jr Diaz LA (2008) **Circulating mutant DNA to assess tumor dynamics** *Nat Med* **14**:985–990
- Gala-Lopez BL, Neiman D, Kin T, O’Gorman D, Pepper AR, Malcolm AJ, Pianzin S, Senior PA, Campbell P, Glaser B, Dor Y, Shemer R, Shapiro AMJ (2018) **Beta Cell Death by Cell-free DNA and Outcome After Clinical Islet Transplantation** *Transplantation* **102**:978–985
- Harker LA, Finch CA (1969) **Thrombokinetis in man** *J Clin Invest* **48**:963–974
- Harrison WJ (1962) **The total cellularity of the bone marrow in man** *J Clin Pathol* **15**:254–259
- Heinke P, Rost F, Rode J, Trus P, Simonova I, Lázár E, Feddema J, Welsch T, Alkass K, Salehpour M, Zimmermann A, Seehofer D, Possnert G, Damm G, Druid H, Bruschi L, Bergmann O. (2022) **Diploid hepatocytes drive physiological liver renewal in adult humans** *Cell Syst* **13**:499–507
- Heitzer E, Auinger L, Speicher MR (2020) **Cell-Free DNA and Apoptosis: How Dead Cells Inform About the Living** *Trends Mol Med* **26**:519–528
- Heitzer E, Haque IS, Roberts CES, Speicher MR (2019) **Current and future perspectives of liquid biopsies in genomics-driven oncology** *Nat Rev Genet* **20**:71–88
- ICRP (2002) **Basic Anatomical and Physiological Data for Use in Radiological Protection Reference Values (No. 89). ICRP**
- Jamshidi A, Liu MC, Klein EA, Venn O, Hubbell E, Beausang JF, Gross S, Melton C, Fields AP, Liu Q, Zhang N, Fung ET, Kurtzman KN, Amini H, Betts C, Civello D, Freese P, Calef R, Davydov K, Fayzullina S, Hou C, Jiang R, Jung B, Tang S, Demas V, Newman J, Sakarya O, Scott E, Shenoy A, Shojaei S, Steffen KK, Nicula V, Chien TC, Bagaria S, Hunkapiller N, Desai M, Dong Z, Richards DA, Yeatman TJ, Cohn AL, Thiel DD, Berry DA, Tummala MK, McIntyre K, Sekeres MA, Bryce A, Aravanis AM, Seiden MV, Swanton C. (2022) **Evaluation of cell-free DNA approaches for multi-cancer early detection** *Cancer Cell* **40**:1537–1549
- Kaufman RM, Airo R, Pollack S, Crosby WH (1965) **Circulating megakaryocytes and platelet release in the lung** *Blood* **26**:720–731



Lehmann-Werman R, Magenheimer J, Moss J, Neiman D, Abraham O, Piyanzin S, Zemmour H, Fox I, Dor T, Grompe M, Landesberg G, Loza B-L, Shaked A, Olthoff K, Glaser B, Shemer R, Dor Y. (2018) **Monitoring liver damage using hepatocyte-specific methylation markers in cell-free circulating DNA** *JCI Insight* **3** <https://doi.org/10.1172/jci.insight.120687>

Lehmann-Werman R, Neiman D, Zemmour H, Moss J, Magenheimer J, Vaknin-Dembinsky A, Rubertsson S, Nellgård B, Blennow K, Zetterberg H, Spalding K, Haller MJ, Wasserfall CH, Schatz DA, Greenbaum CJ, Dorrell C, Grompe M, Zick A, Hubert A, Maoz M, Fendrich V, Bartsch DK, Golan T, Ben Sasson SA, Zamir G, Razin A, Cedar H, Shapiro AMJ, Glaser B, Shemer R, Dor Y. (2016) **Identification of tissue-specific cell death using methylation patterns of circulating DNA** *Proc Natl Acad Sci U S A* **113**:E1826–34

Loyfer N, Magenheimer J, Peretz A, Cann G, Bredno J, Klochendler A, Fox-Fisher I, Shabi-Porat S, Hecht M, Pelet T, Moss J, Drawshy Z, Amini H, Moradi P, Nagaraju S, Bauman D, Shveiky D, Porat S, Dior U, Rivkin G, Or O, Hirshoren N, Carmon E, Pikarsky A, Khalaileh A, Zamir G, Grinbaum R, Abu Gazala M, Mizrahi I, Shussman N, Korach A, Wald O, Izhar U, Erez E, Yutkin V, Samet Y, Rotnemer Golinkin D, Spalding KL, Druid H, Arner P, Shapiro AMJ, Grompe M, Aravanis A, Venn O, Jamshidi A, Shemer R, Dor Y, Glaser B, Kaplan T. (2023) **A DNA methylation atlas of normal human cell types** *Nature* **613**:355–364

Lo YMD, Han DSC, Jiang P, Chiu RWK (2021) **Epigenetics, fragmentomics, and topology of cell-free DNA in liquid biopsies** *Science* **372** <https://doi.org/10.1126/science.aaw3616>

Lo YM, Zhang J, Leung TN, Lau TK, Chang AM, Hjelm NM (1999) **Rapid clearance of fetal DNA from maternal plasma** *Am J Hum Genet* **64**:218–224

Machlus KR, Jr Italiano JE (2013) **The incredible journey: From megakaryocyte development to platelet formation** *J Cell Biol* **201**:785–796

Mandel P, Metais P. (1948) **[Nuclear Acids In Human Blood Plasma]** *C R Seances Soc Biol Fil* **142**:241–243

Matassov D, Kagan T, Leblanc J, Sikorska M, Zakeri Z. (2004) **Measurement of apoptosis by DNA fragmentation** *Methods Mol Biol* **282**:1–17

McIlroy D, Tanaka M, Sakahira H, Fukuyama H, Suzuki M, Yamamura K, Ohsawa Y, Uchiyama Y, Nagata S. (2000) **An auxiliary mode of apoptotic DNA fragmentation provided by phagocytes** *Genes Dev* **14**:549–558

Moss J, Ben-Ami R, Shai E, Kalish Y, Klochender A, Cann G, Glaser B, Arad A, Shemer R, Dor Y. (2022) **Megakaryocyte and erythroblast DNA in plasma and platelets** *bioRxiv* <https://doi.org/10.1101/2022.10.03.510502>

Moss J, Magenheimer J, Neiman D, Zemmour H, Loyfer N, Korach A, Samet Y, Maoz M, Druid H, Arner P, Fu K-Y, Kiss E, Spalding KL, Landesberg G, Zick A, Grinshpun A, Shapiro AMJ, Grompe M, Wittenberg AD, Glaser B, Shemer R, Kaplan T, Dor Y. (2018) **Comprehensive human cell-type methylation atlas reveals origins of circulating cell-free DNA in health and disease** *Nat Commun* **9**

Moss J, Zick A, Grinshpun A, Carmon E, Maoz M, Ochana BL, Abraham O, Arieli O, Germansky L, Meir K, Glaser B, Shemer R, Uziely B, Dor Y. (2020) **Circulating breast-derived DNA allows universal detection and monitoring of localized breast cancer** *Ann Oncol* **31**:395–403

Nagata S. (2005) **DNA degradation in development and programmed cell death** *Annu Rev Immunol* **23**:853–875

Noetzli LJ, French SL, Machlus KR (2019) **New Insights Into the Differentiation of Megakaryocytes From Hematopoietic Progenitors** *Arterioscler Thromb Vasc Biol* **39**:1288–1300

Oberhofer A, Bronkhorst AJ, Uhlig C, Ungerer V, Holdenrieder S. (2022) **Tracing the Origin of Cell-Free DNA Molecules through Tissue-Specific Epigenetic Signatures** *Diagnostics (Basel)* **12** <https://doi.org/10.3390/diagnostics12081834>

Piovesan A, Pelleri MC, Antonaros F, Strippoli P, Caracausi M, Vitale L. (2019) **On the length, weight and GC content of the human genome** *BMC Res Notes* **12**

Sender R, Milo R. (2021) **The distribution of cellular turnover in the human body** *Nat Med* **27**:45–48

Stroun M, Lyautey J, Lederrey C, Olson-Sand A, Anker P. (2001) **About the possible origin and mechanism of circulating DNA apoptosis and active DNA release** *Clin Chim Acta* **313**:139–142

Tabrizi S, Martin-Alonso C, Xiong K, Blewett T, Sridhar S, An Z, Patel S, Rodriguez-Aponte S, Naranjo CA, Wang S-T, Shea D, Golub TR, Bhatia SN, Adalsteinsson V, Christopher Love J. (2023) **An intravenous DNA-binding priming agent protects cell-free DNA and improves the sensitivity of liquid biopsies** *bioRxiv* <https://doi.org/10.1101/2023.01.13.523947>

To EWH, Chan KCA, Leung S-F, Chan LYS, To K-F, Chan ATC, Johnson PJ, Lo YMD (2003) **Rapid clearance of plasma Epstein-Barr virus DNA after surgical treatment of nasopharyngeal carcinoma** *Clin Cancer Res* **9**:3254–3259

Trowbridge EA, Martin JF, Slater DN, Kishk YT, Warren CW, Harley PJ, Woodcock B. (1984) **The origin of platelet count and volume** *Clin Phys Physiol Meas* **5**:145–170

Udomruek S, Orrapin S, Pruksakorn D, Chaiyawat P. (2021) **Size distribution of cell-free DNA in oncology** *Crit Rev Oncol Hematol* **166**

Wan JCM, Massie C, Garcia-Corbacho J, Mouliere F, Brenton JD, Caldas C, Pacey S, Baird R, Rosenfeld N. (2017) **Liquid biopsies come of age: towards implementation of circulating tumour DNA** *Nat Rev Cancer* **17**:223–238

Yao W, Mei C, Nan X, Hui L. (2016) **Evaluation and comparison of in vitro degradation kinetics of DNA in serum, urine and saliva: A qualitative study** *Gene* **590**:142–148

Yoshida H, Kawane K, Koike M, Mori Y, Uchiyama Y, Nagata S. (2005) **Phosphatidylserine-dependent engulfment by macrophages of nuclei from erythroid precursor cells** *Nature* **437**:754–758

Zemmour H, Planer D, Magenheimer J, Moss J, Neiman D, Gilon D, Korach A, Glaser B, Shemer R, Landesberg G, Dor Y. (2018) **Non-invasive detection of human cardiomyocyte death using methylation patterns of circulating DNA** *Nat Commun* **9**

Zhou Q, Kang G, Jiang P, Qiao R, Lam WKJ, Yu SCY, M-JL Ma, Ji L, Cheng SH, Gai W, Peng W, Shang H, Chan RWY, Chan SL, Wong GLH, Hiraki LT, Volpi S, Wong VWS, Wong J, Chiu RWK, Chan KCA, Lo YMD (2022) **Epigenetic analysis of cell-free DNA by fragmentomic profiling** *Proc Natl Acad Sci U S A* **119**

## Author information

### Ron Sender

Weizmann Institute of Science, Rehovot 7610001, Israel  
ORCID iD: [0000-0002-1165-9818](https://orcid.org/0000-0002-1165-9818)

### Elad Noor

Weizmann Institute of Science, Rehovot 7610001, Israel

### Ron Milo

Weizmann Institute of Science, Rehovot 7610001, Israel  
**For correspondence:** [ron.milo@weizmann.ac.il](mailto:ron.milo@weizmann.ac.il)  
ORCID iD: [0000-0003-1641-2299](https://orcid.org/0000-0003-1641-2299)

### Yuval Dor

Department of Developmental Biology and Cancer Research, Institute for Medical Research Israel Canada, the Hebrew University-Hadassah Medical School; Jerusalem, Israel  
**For correspondence:** [yuvald@ekmd.huji.ac.il](mailto:yuvald@ekmd.huji.ac.il)  
ORCID iD: [0000-0003-2456-2289](https://orcid.org/0000-0003-2456-2289)

## Editors

Reviewing Editor

### Jonathan Cooper

Fred Hutchinson Cancer Research Center, United States of America

Senior Editor

### Jonathan Cooper

Fred Hutchinson Cancer Research Center, United States of America

## Reviewer #1 (Public Review):

Summary:

Sender et al describe a model to estimate what fraction of DNA becomes cell-free DNA in plasma. This is of great interest to the community, as the amount of DNA from a certain tissue (for example, a tumor) that becomes available for detection in the blood has important implications for disease detection.

However, the authors' methods do not consider important variables related to cell-free DNA shedding and storage, and their results may thus be inaccurate. At this stage of the paper, the methods section lacks important detail. Thus, it is difficult to fully assess the manuscript and its results.

Strengths:

The question asked by the authors has potentially important implications for disease diagnosis. Understanding how genomic DNA degrades in the human circulation can guide towards ways to enrich for DNA of interest or may lead to unexpected methods of conserving cell-free DNA. Thus, the question "how much genomic DNA becomes cfDNA" is of great interest to the scientific and medical community. Once the weaknesses of the manuscript are addressed, I believe this manuscript has the potential to be a widely used resource.

#### Weaknesses:

There are two major weaknesses in how the analysis is presented. First, the methods lack detail. Second, the analysis does not consider key variables in their model.

Issues pertaining to the methods section.

The current manuscript builds a flux model, mostly taking values and results from three previous studies:

- 1- The amount of cellular turnover by cell type, taken from Sender & Milo, 2021
- 2- The fractions of various tissues that contribute DNA to the plasma, taken from Moss et al, 2018 and Loyfer et al, 2023

My expertise lies in cell-free DNA, and so I will limit my comments to the manuscripts in (2).

Paper by Loyfer et al (additional context):

Loyfer et al is a recent landmark paper that presents a computational method for deconvoluting tissues of origin based on methylation profiles of flow-sorted cell types. Thus, the manuscript provides a well-curated methylation dataset of sorted cell-types. The majority of this manuscript describes the methylation patterns and features of the reference methylomes (bulk, sorted cell types), with a smaller portion devoted to cell-free DNA tissue of origin deconvolution.

I believe the data the authors are retrieving from the Loyfer study are from the 23 healthy plasma cfDNA methylomes analyzed in the study, and not the re-analysis of the 52 COVID-19 samples from Cheng et al (MED 2021).

Paper by Moss et al (additional context):

Moss et al is another landmark paper that predates the Loyfer et al manuscript. The technology used in this study (methylation arrays) is outdated but is an incredible resource for the community. This paper evaluates cfDNA tissues of origin in health and different disease scenarios. Again, I assume the current manuscript only pulled data from healthy patients, although I cannot be sure as it is not described in the methods section.

This manuscript:

The current manuscript takes (I think) the total cfDNA concentration from males and females from the Moss et al manuscript (pooled cfDNA; 2 young male groups, 2 old male groups, 2 young female groups, 2 old female groups, Supplementary Dataset; "total\_cfDNA\_conc" tab). I believe this is the data used as total cfDNA concentration. It would be beneficial for all readers if the authors clarified this point.

The tissues of origin, in the supplemental dataset ("fraction" tab), presents the data from 8 cell types (erythrocytes, monocytes/macrophages, megakaryocytes, granulocytes, hepatocytes, endothelial cells, lymphocytes, other). The fractions in the spreadsheet do not match the Loyfer or Moss manuscripts for healthy individuals. Thus, I do not know what values the supplementary dataset represents. I also don't know what the deconvolution values are used for the flux model.

The integration of these two methods lack detail. Are the authors here using yields (ie, cfDNA concentrations) from Moss et al, and tissue fractions from Loyfer et al? If so, why? There are more samples in the Loyfer manuscript, so why are the samples from Moss et al. being used? The authors are also selectively ignoring cell-types that are present in healthy individuals (Neurons from Moss et al, 2018). Why?

Appraisal:

At this stage of the manuscript, I think additional evidence and analysis is required to confirm the results in the manuscript.

Impact:

Once the authors present additional analysis to substantiate their results, this manuscript will be highly impactful on the community. The field of liquid biopsies (non-invasive diagnostics) has the potential to revolutionize the medical field (and has already in certain areas, such as prenatal diagnostics). Yet, there is a lack of basic science questions in the field. This manuscript is an important step forward in asking more "basic science" questions that seek to answer a fundamental biological question.

## Reviewer #2 (Public Review):

Summary:

Cell-free DNA (cfDNA) are short DNA fragments released into the circulation when cells die. Plasma cfDNA level is thought to reflect the degree of cell-death or tissue injury. Indeed, plasma cfDNA is a reliable diagnostic biomarker for multiple diseases, providing insights into disease severity and outcomes. In this manuscript, Dr. Sender and colleagues address a fundamental question: What fraction of DNA released from cell death is detectable as plasma cfDNA? The authors use public data to estimate the amount of DNA produced from dying cells. They also utilize public data to estimate plasma cfDNA levels. Their calculations showed that <10% of DNA released is detectable as plasma cfDNA, the fraction of detectable cfDNA varying by tissue sources. The study demonstrates new and fundamental principles that could improve disease diagnosis and treatment via cfDNA.

Strengths:

1. The experimental approach is resource-mindful taking advantage of publicly available data to estimate the fraction of detectable cfDNA in physiological states. The authors did not assess if the fraction of detectable cfDNA changes in disease conditions. Nonetheless, their pioneering study lays the foundation and provides the methods needed for a similar assessment in disease states.
2. The findings of this study potentially explain discrepancies in measured versus expected tissue-specific cfDNA from some tissues. For example, the gastrointestinal tract is subject to high cell turnover and release of DNA. Yet, only a small fraction of that DNA ends up in plasma as gastrointestinal cfDNA.
3. The study proposes potential mechanisms that could account for the low fraction of detectable cfDNA in plasma relative to DNA released. This includes intracellular or tissue machinery that could "chew up" DNA released from dying cells, allowing only a small fraction to escape into plasma as cfDNA. Could this explain why the gastrointestinal track with an elaborate phagosome machinery contributes a small fraction of plasma cfDNA? Given the role of cfDNA as damage-associated molecular pattern in some diseases, targeting such a machinery may provide novel therapeutic opportunities.

Weaknesses: In vitro and in vivo studies are needed to validate these findings and define tissue machinery that contribute to cfDNA production. The validation studies should address

the following limitations of the study design: -

1. Align the cohorts to estimate DNA production and plasma cfDNA levels. Cellular turnover rate and plasma cfDNA levels vary with age, sex, circadian clock, and other factors (Madsen AT et al, EBioMedicine, 2019). This study estimated DNA production using data abstracted from a homogenous group of healthy control males (Sender & Milo, Nat Med 2021). On the other hand, plasma cfDNA levels were obtained from datasets of more diverse cohort of healthy males and females with a wide range of ages (Loyfer et al. Nature, 2023 and Moss et al., Nat Commun, 2018).
2. "cfDNA fragments are not created equal". Recent studies demonstrate that cfDNA composition vary with disease state. For example, cfDNA GC content, fraction of short fragments, and composition of some genomic elements increase in heart transplant rejection compared to no-rejection state (Agbor-Enoh, Circulation, 2021). The genomic location and disease state may therefore be important factors to consider in these analyses.
3. Alternative sources of DNA production should be considered. Aside from cell death, DNA can be released from cells via active secretion. This and other additional sources of DNA should be considered in future studies. The distinct characteristics of mitochondrial DNA to genomic DNA should also be considered.