The Journal of Molecular Diagnostics, Vol. 26, No. 5, May 2024

ELSEVIER

the Journal of Molecular Diagnostics
jmdjournal.org

# Assay Validation of Cell-Free DNA Shallow Whole-Genome Sequencing to Determine Tumor Fraction in Advanced Cancers



Micah Rickles-Young,* Gabriel Tinoco,[†‡] Junko Tsuji,* Sam Pollock,* Marcy Haynam,[‡§] Heather Lefebvre,[‡§] Kristyn Glover,[‡§] Dwight H. Owen,[†‡] Katharine A. Collier,[†‡] Gavin Ha,[¶] Viktor A. Adalsteinsson,* Carrie Cibulskis,* Niall J. Lennon,* and Daniel G. Stover[†‡§]

*From the Genomics Platform,* Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, Massachusetts; the Division of Medical Oncology,[†] The Ohio State University College of Medicine, Columbus, Ohio; the Ohio State University Comprehensive Cancer Center,[‡] Columbus, Ohio; the Stefanie Spielman Comprehensive Breast Center,[§] Columbus, Ohio; and the Public Health Sciences Division,[¶] Fred Hutchinson Cancer Research Center, Seattle, Washington*

Accepted for publication January 18, 2024.

Address correspondence to Daniel G. Stover, M.D., Ohio State University Comprehensive Cancer Center, James Cancer Hospital, Columbus, OH 43210; or Niall J. Lennon, Ph.D., The Broad Institute of Harvard and MIT, 415 Main St., Cambridge, MA 02142. E-mail: daniel.stover@osumc.edu or nlennon@broadinstitute.org.

Blood-based liquid biopsy is increasingly used in clinical care of patients with cancer, and fraction of tumor-derived DNA in circulation (tumor fraction; TFx) has demonstrated clinical validity across multiple cancer types. To determine TFx, shallow whole-genome sequencing of cell-free DNA (cfDNA) can be performed from a single blood sample, using an established computational pipeline (ichorCNA), without prior knowledge of tumor mutations, in a highly cost-effective manner. We describe assay validation of this approach to facilitate broad clinical application, including evaluation of assay sensitivity, precision, repeatability, reproducibility, pre-analytic factors, and DNA quality/quantity. Sensitivity to detect TFx of 3% (lower limit of detection) was 97.2% to 100% at 1× and 0.1× mean sequencing depth, respectively. Precision was demonstrated on distinct sequencing instruments (HiSeqX and NovaSeq) with no observable differences. The assay achieved prespecified 95% agreement of TFx across replicates of the same specimen (repeatability) and duplicate samples in different batches (reproducibility). Comparison of samples collected in EDTA and Streck tubes from single venipuncture in 23 patients demonstrated that EDTA or Streck tubes were comparable if processed within 8 hours. On the basis of a range of DNA inputs (1 to 50 ng), 20 ng cfDNA is the preferred input, with 5 ng minimum acceptable. Overall, this shallow whole-genome sequencing of cfDNA and ichorCNA approach offers sensitive, precise, and reproducible quantitation of TFx, facilitating assay application in clinical cancer care. *(J Mol Diagn 2024, 26: 413—422; https://doi.org/10.1016/j.jmoldx.2024.01.014)*

Next-generation sequencing technologies are capable of detecting small amounts of cell-free DNA (cfDNA) shed by normal and malignant cells into the circulation, making the blood an accessible repository for real-time genomic profiles of tumors. There are growing data that the fraction of tumor-derived DNA among total cfDNA (tumor fraction; TFx) is prognostic and may be useful in cancer detection, and changes in TFx may serve as an early identifier for patients responding, or failing to respond, to therapy.[1—9]

Early efforts to calculate cfDNA TFx focused on a personalized assay approach, in which mutations are identified through tumor sequencing, then mutation-specific primers synthesized to probe cfDNA from plasma.[2,10—12] This approach requires both tumor biopsy and germline sequencing, complex analyses, and time-consuming probe generation. Personalized assays and panel sequencing approaches are expensive (Figure 1A). Analysis of whole-exome sequencing (WES) of cfDNA provides a reliable approach to determine TFx via algorithms, such as TITAN[13] or ABSOLUTE,[14] and may serve as a gold standard, but it is expensive and depends on adequate tumor DNA. As an

Supported by National Cancer Institute UH2CA239105 (G.H., V.A.A, N.J.L., and D.G.S.); and the Gerstner Family Foundation (V.A.A.).

M.R.-Y. and G.T. contributed equally to this work.

N.J.L. and D.G.S. equally directed this work.


https://doi.org/10.1016/j.jmoldx.2024.01.014

alternative, ultra—low-pass (ULP) or shallow whole-genome sequencing (WGS) assay is a low-cost approach to determine TFx.

ULP-WGS allows rapid, precise quantitation of tumor DNA content in plasma as well as somatic copy number alterations (SCNAs) from a fraction (approximately 5%) of the DNA derived from a single blood sample, without prior knowledge of tumor mutations, all in a highly cost-effective manner. In ULP-WGS, small fragment DNA is extracted from plasma, with WGS performed at shallow coverage (typical median depth, 0.1× to 1×), then quantification of tumor content in cfDNA is performed through a computational pipeline, ichorCNA, which uses a hidden Markov probabilistic two-component mixture model to derive TFx and SCNAs from ULP-WGS data (Figure 1B). We published our application of ULP-WGS on 1439 blood samples from 520 patients with metastatic prostate or breast cancers, demonstrating scalability, precision, and reproducibility.[15] In subsequent publications, we and others have evaluated the clinical significance of TFx, demonstrating significant association with prognosis in multiple cancers.[2,3] Work by our group and others in metastatic breast cancer demonstrated that the TFx is strongly associated with metastatic survival,[2] and more recent work using a commercial assay suggested that TFx is prognostic across common tumor types.[4] Turner and colleagues showed that the change in tumor DNA content in plasma from day 1 to 15 of fulvestrant plus cyclin-dependent kinase 4/6 inhibitor palbociclib, defined by *PIK3CA* mutant copies/mL, was highly associated with progression-free survival.[16] These efforts demonstrate that cfDNA TFx and TFx dynamics are promising as prognostic and predictive genomic biomarkers for metastatic breast and other cancers.

The ULP-WGS approach for cfDNA interrogation offers potential for a cost-effective genomic biomarker from simple blood draws without the need for tumor biopsy, a truly

**A**

| Technology | % Genome Sequenced | SNV Detection | Tumor Content | Computational Analysis | Cost |
|---|---|---|---|---|---|
| ULP-WGS | 100% | NA | TFx to ~3% | Easy | |
| Genotyping/ ddPCR | <0.001% | ~1/1000 | VAF | Easy | |
| Targeted Panel | <0.1% (100s genes) | ~1/1000 | VAF | Moderate | |
| WES | 1% | ~1/50 | TFx to ~0.1% | Moderate | |
| WGS | 100% | ~1/20 (~60X) | TFx to ~1% | Complex | |
| Personal/ Bespoke Assay | Pre: 1%-100% Assay:<0.001% | ~1/100,000 | TFx to <<0.1% | Complex | |

**B**

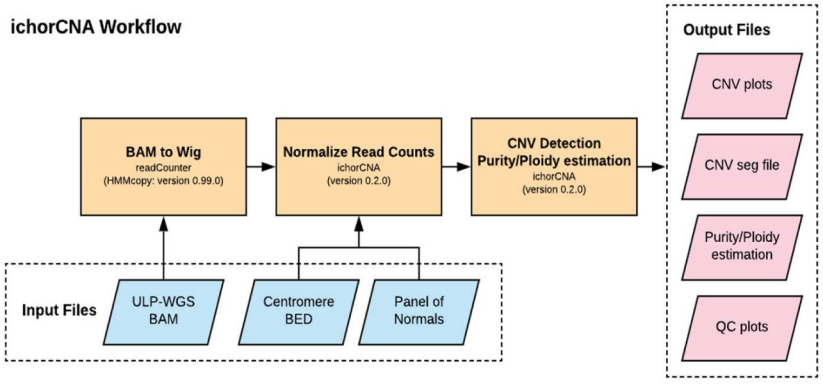



**Figure 1** Cell-free DNA sequencing and computational approaches. **A:** Overview of example approaches for cell-free DNA sequencing in cancer, which reflect varying coverage of the genome, distinct sensitivities to detect specific mutations/single-nucleotide variants (SNVs), ranging tumor content detection limits, and wide variation in computational complexity and cost. **B:** Schematic overview of the ichorCNA pipeline for tumor fraction (TFx) determination from shallow whole-genome sequencing (WGS). CNV, copy number variation; ddPCR, digital droplet PCR; NA, not applicable; QC, quality control; ULP, ultra—low pass; VAF, variant allele fraction; WES, whole-exome sequencing.

patient-centered diagnostic. In addition, radiologic imaging remains the standard for assessing tumor response, yet the cost and complexity remain a significant burden. Cost-effective liquid biopsy approaches that could augment or reduce the number of scans have the potential to add significant value to the health care system.

Herein, we report the clinical validation of ULP-WGS to determine TFx, with a goal to establish the performance characteristics of the assay. The intended use of the ULP-WGS TFx assay and ichorCNA pipeline is to determine the TFx in blood samples from individuals with metastatic cancer to provide information to researchers for use in clinical practice or the analysis of clinical trials. This validation exclusively focuses on establishing validity for the wet laboratory and computational processes, including analytical sensitivity and specificity, precision, and limitations of the assay.

## Materials and Methods

### Sample Identification

Patients with metastatic biopsy-proven breast cancer enrolled on clinical data and biospecimen banking protocols were identified for analyses, as described previously.[2] The study was approved by the institutional review boards of the Dana-Farber Cancer Institute (Boston, MA) and The Ohio State University (Columbus, OH) and was conducted in accordance with the Declaration of Helsinki. All patients provided written consent. Samples included in analyses were as follows: 3 patients with eight dilutions each for quantitative dilutions (Supplemental Table S1), 10 patients with six replicates each for repeatability/reproducibility (Supplemental Table S2), 24 patients with three replicates each for tube type/time to processing (Supplemental Table S3), and 7 patients with four cfDNA input amounts for input titration (Supplemental Table S4); samples for quantitative dilutions and cfDNA input experiments were from the repeatability/reproducibility cohort, resulting in a total of 34 unique patients. A total of 52 healthy donor plasma samples were included, 20 for updated panel of normal, 24 for reproducibility and repeatability experiments, and 8 for dilution experiments.

### Blood Sample Collection, Processing, and Sequencing

Venous blood samples were processed to component parts within 4 hours of collection (unless otherwise specified), through standard density gradient centrifugation. Plasma samples were subjected to an additional high-speed spin at 19,000 $\times$ *g* for 10 minutes, and plasma was frozen in 1- to 2-mL aliquots at $-80^{\circ}$C until further processing. Frozen aliquots of plasma were thawed at room temperature and subjected to high-speed spin if not previously performed. As described previously,[15] cell-free DNA was extracted from 4 to 6 mL of plasma using the Qiagen (Hilden, Germany) Circulating DNA kit on the QIAsymphony liquid handling system, and DNA quantification was performed. Up to 50 ng of cfDNA input (input as specified), or approximately 1000 to 10,000 haploid genome equivalents, was used for ULP-WGS. Constructed sequencing libraries were pooled (2 μL of each $\times$ 96 per pool) and sequenced using 150 bp paired-end runs over 1 $\times$ lane on a HiSeqX (Illumina, San Diego, CA) or NovaSeq (Illumina) instrument to average genome-wide fold coverage of 0.1$\times$ to 1$\times$; there was no difference in library preparation for HiSeqX versus NovaSeq sequencing.

### Computational Determination of Tumor Fraction via ichorCNA

Segment copy number and TFx were derived with ichorCNA (Figure 1B).[15] ichorCNA quantifies TFx in cfDNA without prior knowledge of somatic single-nucleotide variants or SCNAs in patients' tumors from ULP-WGS by simultaneously predicting segments of SCNA and estimating TFx while accounting for subclonality and tumor ploidy. Briefly, after assessment of read coverage and read count normalization for GC content and mappability using HMMcopy, a hidden Markov model is used for copy number prediction and TFx estimation from the tumor-normal cfDNA admixture. For the analysis pipeline null model, an additional panel of 20 independent healthy donors were sequenced using the NovaSeq platform to generate a panel of normal noise model using recommended ichorCNA parameters. The NovaSeq panel of normal demonstrated similar noise reduction performance to the panel of normal existing within the ichorCNA package, suggesting no significant difference based on sequencing platform. These additional healthy donor data were not needed for successful routine ichorCNA use. Quality metric GC Map Correction MAD (mean absolute deviation), a metric within the ichorCNA package, represents the median absolute deviation in coverage between adjacent copy number bins after the coverage corrected for mappability and GC content.

### Statistical Analysis and Data Visualization

All statistical analyses and data visualizations were performed in R version 4.0.3 (R Project for Statistical Computing, *https://www.r-project.org*), including sensitivity and specificity. For both reproducibility and repeatability cohorts, mean-normalized SD (MNSD) was defined as the ratio of the SD of the tumor fraction estimate between replicates to the mean. For tube type analyses, mean absolute percentage error was calculated using the Metrics package in R (*https://github.com/mfrasco/Metrics*), as the mean of the absolute value of the difference between nonzero Streck and EDTA TFx estimates divided by the EDTA TFx estimate for samples within each time point.

### Data Availability

All sequencing data supporting the conclusions of this article are available through the database of genotypes and phenotypes (*https://www.ncbi.nlm.nih.gov/gap*, cacession numbers: phs003500.v1.p1, last accessed March 1, 2024).

## Results

### Sensitivity of ULP-WGS versus WES Gold Standard in Determining Tumor Fraction

The prespecified goal was to achieve 95% sensitivity to detect TFx of 3%, with TFx of 4% acceptable. True positive was TFx >3% via ULP-WGS when TFx via WES was also >3%. True negative was TFx <3% via ULP-WGS when TFx via WES was also <3%. False positive was TFx >3% via ULP-WGS when TFx via WES was <3%. False negative was TFx <3% via ULP-WGS when TFx via WES was >3%. In this quantitative TFx cohort, using $N = 3$ patient cfDNA samples with previously established TFx estimates and $N = 2$ healthy donor blood samples from research blood components for serial dilution, the authors analyzed eight dilution levels per patient, performed in triplicate (Supplemental Table S1). The gold standard TFx was determined through circulating tumor DNA (ctDNA) whole-exome sequencing. The MNSD increased significantly below 3%, thus establishing 3% TFx as the lower limit of detection (Figure 2A), in line with published work based on *in silico* experimentation.[15] The authors determined that sensitivity ranged from 97.2% to 100% at both 1× and 0.1× mean depth down-sampling (Figure 2B). For lower limit of detection, as detailed in this paragraph, the authors had excellent sensitivity/specificity at 3% and 4% TFx.

Although ichorCNA was optimized for resolution of TFx ≤10%, typical ctDNA levels for most advanced cancers, certain cancers may have higher levels, particularly triple-negative breast cancer (cancer source type for most samples), as previously reported.[2] At TFx >10%, performance was similar for two samples (alias EAQYU and HKB8G), but the authors noted differences between WES gold standard and ULP-WGS in one sample (E9E2Y) (Supplemental Figure S1). The authors further investigated this phenomenon of potential underestimation of TFx and identified that copy-neutral loss of heterozygosity (CNLOH) could contribute. Because ULP-WGS lacks the resolution of WES to measure allelic copy number data, segments of the genome that exhibit CNLOH are indistinguishable from truly copy-neutral segments based solely on coverage, resulting in lower TFx estimates (Figure 2C).

### Precision of ULP-WGS on Distinct Sequencing Instruments

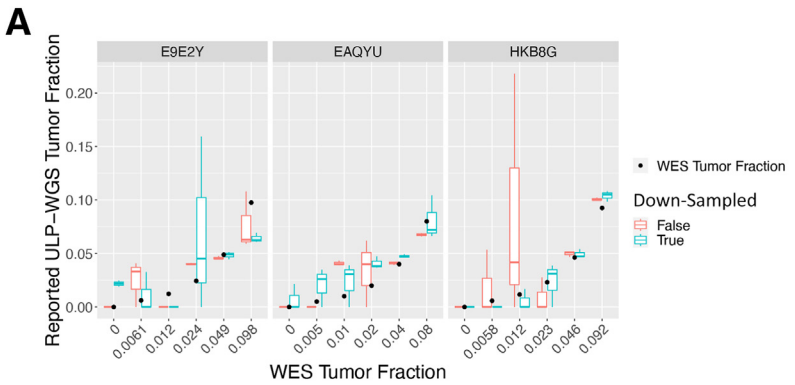
As sequencing instruments are continuously updated, the authors evaluated HiSeqX and NovaSeq instrumentation and rigorously evaluated both technologies (Figure 3, A and B). The major difference between the HiSeqX technology and NovaSeq technology is the sequencing chemistry. HiSeqX uses four-color chemistry, whereas NovaSeq uses two-color chemistry. NovaSeq offers higher yield of sequencing depending on the flow cell used. The authors assessed MAD score of the copy ratio differences between adjacent data points as an indicator of data quality and coverage; historically, in their experience, an MAD score <0.20 designates samples with sufficient data quality.[15] The authors demonstrate that MAD was highly concordant between sequencing instruments (Figure 3A). Furthermore, the authors demonstrated, by unsupervised hierarchical clustering, that instrument-specific biases were not observed (Figure 3B). In sum, these additional analyses provided confidence that, at least among these two instruments, no notable differences were observed.

### Repeatability and Reproducibility of Tumor Fraction via ULP-WGS

The authors evaluated agreement in TFx across replicates of the same specimen (repeatability) and duplicate samples in different batches (reproducibility) (Supplemental Table S2). To assess repeatability of TFx via ULP-WGS, 30 clinical specimens and 1 healthy donor sample were prepared and sequenced in triplicate, in the same batch (Figure 3C). Repeatability was assessed as the ability to consistently generate concordant TFx estimates across replicates of the same specimen. To assess reproducibility of TFx via ULP-WGS under varying conditions, 29 clinical specimens and 7 healthy donor samples were prepared in three separate batches (Figure 3D). Two batches were prepared on the same day by different technologists. One batch was prepared by the same technologist, but on a different day. For all samples >10% TFx, MNSD per replicate set was <0.19, whereas for samples >5% TFx, MNSD per replicate set was <0.35. There were three samples with TFx 3% to 5%, and MSND ranged from 0.33 to 0.54.

### Sequencing Quality Metrics for ULP-WGS Tumor Fraction Assessment

To ensure capacity across a range of outputs, the authors sought to establish reportable range, reference intervals, and quality metrics from control samples. Quality metrics were assessed across replicates within and across batches, including the following: MAD score, mean coverage, library complexity, percentage coverage/pass filter gigabase, and percentage chimera. The goal of the sequencing depth was 0.1× to 1×, and mean depth for the repeatability cohort ranged from 0.32 to 0.92, with all samples achieving depth well above the 0.1× goal (Supplemental Figure S2). Library size ranged from $3.5 \times 10^7$ to $2.1 \times 10^8$, with a at least 99.6% of reads aligned (Supplemental Figure S3). The fraction of chimeric read pairs indicates lower-quality




B

| | Quantitative Dilution Series | |
|---|---|---|
| | Sensitivity (With Error Bounds) | |
| Coverage | 3% Tumor Fraction | 4% Tumor Fraction |
| Target 1x Mean Depth | 36/36 - **100%** (90% ≤ p ≤ 100%) | 35/36 - **97.2%** (85% ≤ p ≤ 99%) |
| 0.1x Mean Depth Down-Sampling | 36/36 - **100%** (90% ≤ p ≤ 100%) | 36/36 - **100%** (90% ≤ p ≤ 100%) |

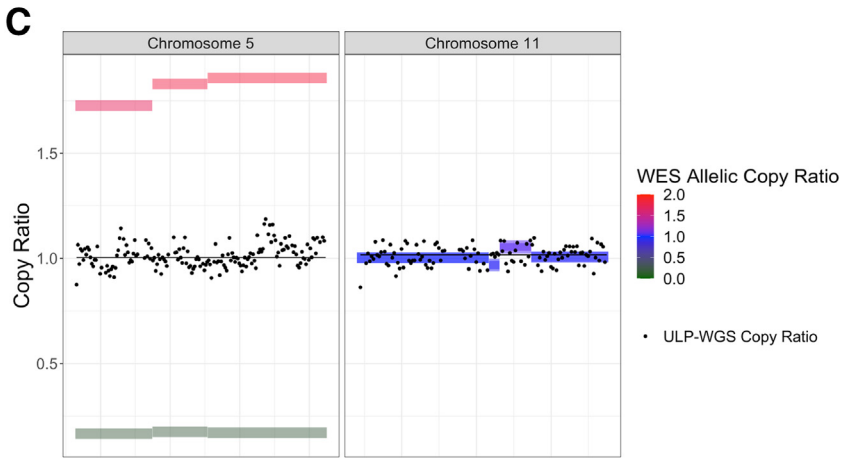



**Figure 2** Sensitivity of tumor fraction determination. **A:** Using three patient cell-free DNA samples with previously established tumor fraction (TFx) estimates and two healthy donor blood samples from research blood components for serial dilution, eight experimental dilution levels per patient were analyzed, performed in triplicate (**red boxed areas**). In addition, *in silico* down-sampling was performed to mimic experimental dilution (**blue boxed areas**). The gold standard TFx was determined through circulating tumor DNA whole-exome sequencing (WES; black dots). **B:** Sensitivity ranged from 97.2% to 100% at both 1× and 0.1× mean depth down-sampling. Error bounds indicated in parentheses. **C:** Copy-neutral loss of heterozygosity may impact TFx estimates, as seen in chromosome 5 (**left panel**) versus chromosome 11 (**right panel**) within a single sample C-519473, part of the input titration cohort. TFx by ultra–low-pass whole-genome sequencing (ULP-WGS)/ichorCNA was 62%, whereas TFx via WES/ABSOLUTE was 82%. Allelic copy ratio via ABSOLUTE from whole-exome sequencing indicated in bars (colored by copy ratio) versus copy ratio via ichorCNA from ultra–low-pass whole-genome sequencing in black dots. Chromosome 5 reflects balanced copy number events inferred as copy neutral by ichorCNA, resulting in an artificially lower TFx estimate.

libraries but may also represent structural rearrangement events present in tumor samples. It was observed that higher TFx is correlated with higher rates of chimeric reads, but across all samples in the repeatability set, the fraction of chimeric reads was <1.65% of total read pairs (Supplemental Figure S4). The MAD score ranged from 0.024 to 0.050 (Supplemental Figure S5). Coverage per pass-filter gigabase sequenced is a measure of sequencing efficiency for this process and ranged from 0.16 to 0.19 in the repeatability TFx cohort, which is within the range of expected values (Supplemental Figure S6).

## Impact of Pre-Analytic Collection Tube Type and Time to Processing

To ensure that the authors can effectively analyze samples collected in either EDTA or Streck BCT tubes, the authors directly compared tube types. The authors obtained 10 mL blood in one EDTA and one Streck tube (from a single venipuncture) from 23 total patients (Supplemental Table S3). The 2018 American Society of Clinical Oncology/College of American Pathologists "Joint Review on ctDNA Analysis in Patients with Cancer" noted, "EDTA tubes need to be processed...within 6 hours."[17,p.1633] With <4 hours being already established as an acceptable time window, the authors tested longer times between collection and processing: 3 unique patients for <4 hours from collection to processing, 8 unique patients for 6 to 8 hours from collection to processing, and 12 unique patients for 12 to 24 hours from collection to processing. The authors observed no significant difference between EDTA and Streck in quality metrics (Figure 4, A–C). TFx was similar across samples, although there was a single sample with high TFx and variance in the <4 hours' group (Figure 4D), and the shortest interval (<4 hours) had a Streck versus EDTA mean absolute percentage error of 4.7% (Figure 4E). These

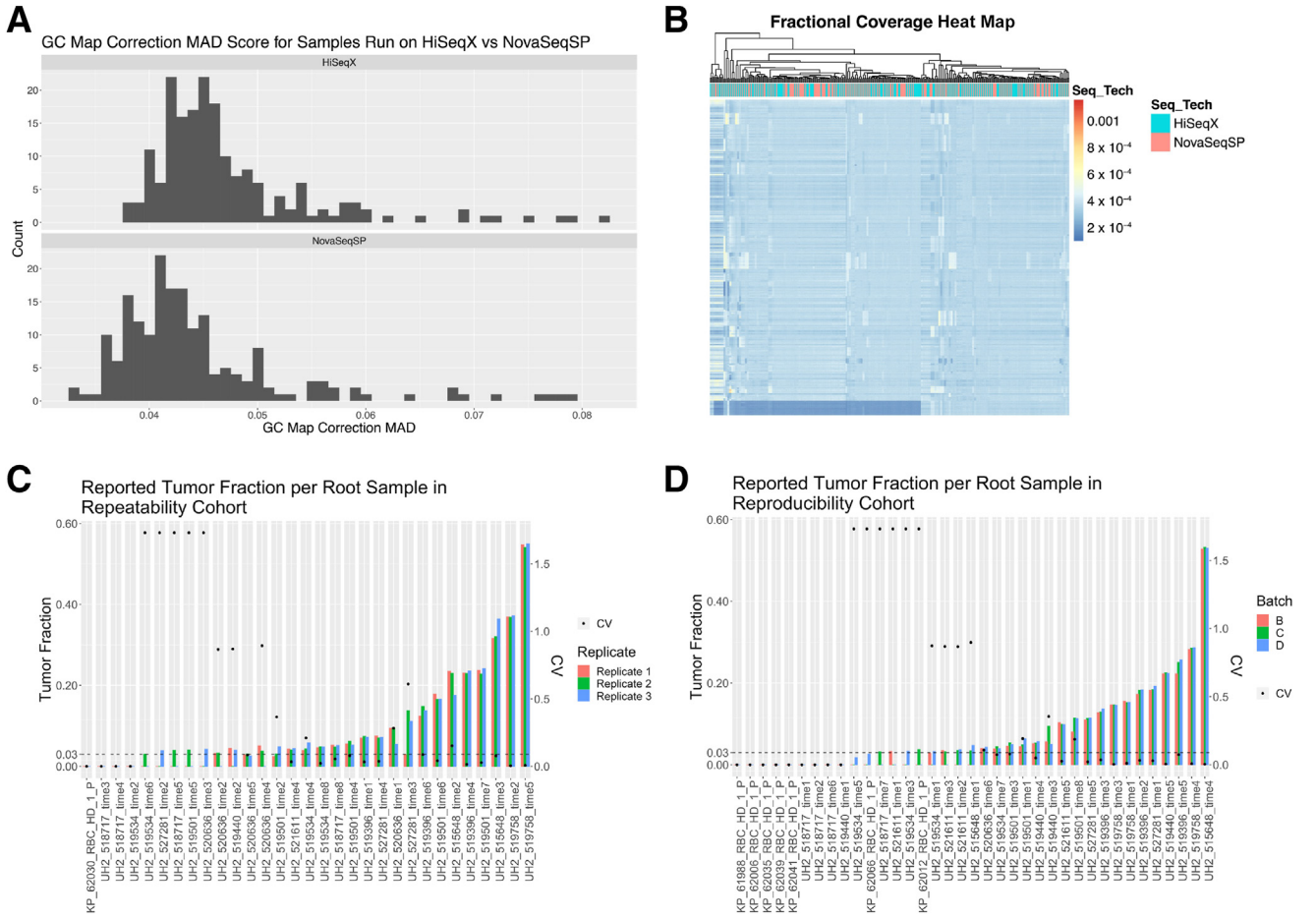



**Figure 3** Precision, repeatability, and reproducibility of tumor fraction via shallow whole-genome sequencing. **A:** GC-corrected mean absolute deviation (MAD) score of samples run on Illumina HiSeqX (**top panel**) versus NovaSeq (**bottom panel**) had similar distribution across sequencers. **B:** Unsupervised hierarchical clustering (average linkage, Euclidean distance) of samples run on either HiSeqX or NovaSeq demonstrates no clustering by machine, with color bar above heat map indicating sequencing instrument (HiSeqX, blue; NovaSeq, pink). **C** and **D:** Repeatability and reproducibility was assessed on 60 clinical specimens and 8 healthy donor samples, with repeatability assessed across replicates of the same specimen and reproducibility assessed via two batches prepared on the same day by different technologists and one batch was prepared by the same technologist, but on a different day. Tumor fraction (left axis) of sample dilution series run in triplicate. CV indicated by black dots (right axis).

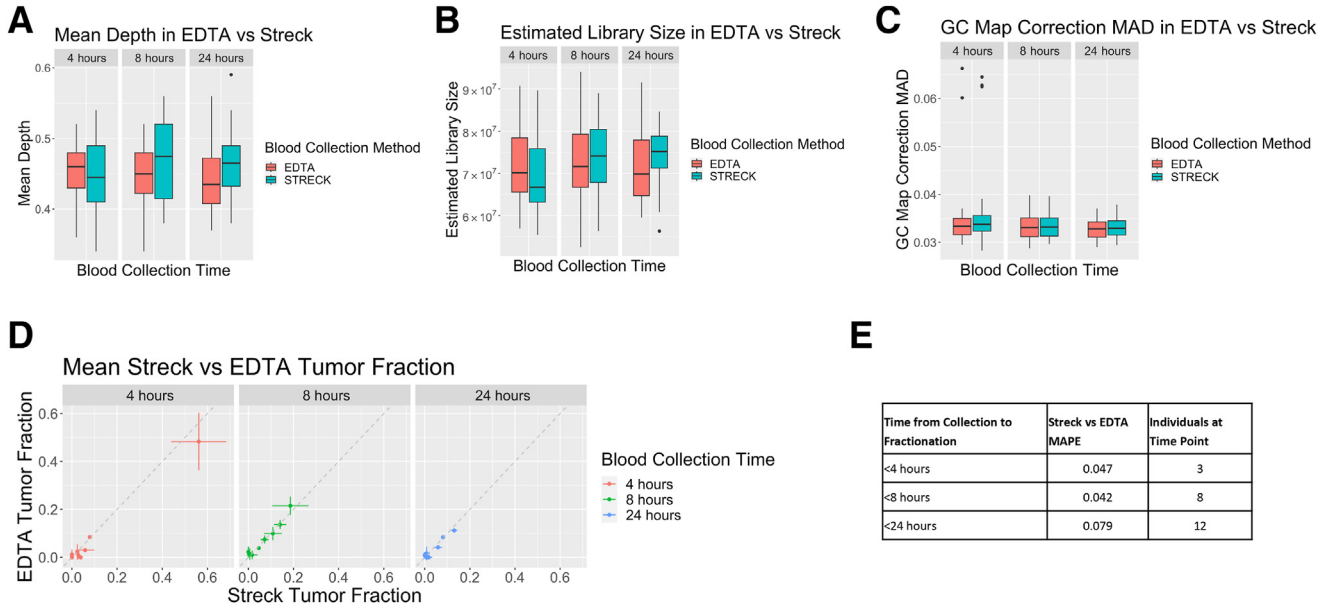
data demonstrate that EDTA or Streck tubes, if processed within 8 hours, achieve similar results and would be acceptable. Longer duration for Streck tubes per package insert is also acceptable.

### Impact of DNA Quality and Quantity on ULP-WGS Tumor Fraction Determination

The research ULP-WGS TFx assay before this study used total small fragment DNA extracted from 4 mL of plasma (from approximately 10 mL of whole blood) and 5 ng DNA input for library preparation (minimum input, 3 ng). To assess the impact of DNA input and whether the authors could get optimal results with lower input, the authors evaluated a range of DNA inputs to align with US Food and Drug Administration standard 10-fold variation (Supplemental Table S4). Using samples with high DNA content (>500 ng total cfDNA), the authors systematically tested inputs of 1/5/20/50 ng (in triplicate). The observed TFx was consistent even at 1-ng input (Figure 5A), and all inputs of 20/50 ng achieved library size $>4 \times 10^7$ (Figure 5B), depth (mean, > 0.25) (Figure 5C), MAD <0.20 (100% of samples) (Figure 5D), and coverage per pass filter gigabase (>0.15) (Figure 5E), whereas 1-ng inputs failed to achieve each metric in at least one sample. On the basis of these results, the preferred input is 20 ng cfDNA, with 5 ng as a minimum acceptable amount (quality metrics for input titration cohort provided in Supplemental Table S5).

## Discussion

Cancer liquid biopsy is a rapidly evolving diagnostic approach for the detection of circulating analytes reflective of cancer activity through minimally invasive blood analysis.[4,18] With the widespread adoption and regulatory approval of multiple multigene assays, ctDNA analysis has been increasingly incorporated into clinical practice. The




| Time from Collection to Fractionation | Streck vs EDTA MAPE | Individuals at Time Point |
|---|---|---|
| <4 hours | 0.047 | 3 |
| <8 hours | 0.042 | 8 |
| <24 hours | 0.079 | 12 |

**Figure 4** Pre-analytic factors and impact of tumor fraction determination from cell-free DNA. **A**–**C:** Comparison of sequencing metrics from samples collected in two different tube types (EDTA versus Streck) from a single venipuncture in 23 patients. Samples were processed after varying times of incubation at room temperature: <4 hours (4 hours in figure), 4 to 8 hours (8 hours), or 12 to 24 hours (24 hours). **A**–**C:** There were no significant differences in library size (**A**), mean depth of sequencing (**B**), or mean absolute deviation (MAD; **C**). **D:** Correlation coefficients of tumor fraction in EDTA (*y* axis) versus Streck (*x* axis) of paired samples. **E:** Mean absolute percentage error (MAPE) of EDTA versus Streck by time point. Number of individuals per time point indicated.

clinical application of cfDNA analysis for patient care is limited by the amount of ctDNA; TFx varies widely between individuals, cancer types, tumor burden, and disease activity; and all biological and clinical factors that influence TFx have yet to be fully identified.[9,19–24] Given the cost and complexity of many clinical assays, validated low-cost, high-throughput assays to assess TFx are needed. ULP-WGS allows rapid, precise quantitation of TFx of the DNA derived from a single blood sample, without prior knowledge of tumor mutations, all in a highly cost-effective manner. ULP-WGS offers multiple potential clinical applications: i) TFx as a stand-alone prognostic assay,[2,3,15] ii) TFx dynamics to potentially assess treatment response, iii) low-cost, low-input quality assessment of ctDNA for downstream assays, such as targeted panel sequencing,[15] and iv) orthogonal algorithms to assess ctDNA features, such as homologous recombination deficiency,[25] nucleosome positioning,[26] or cancer type–specific SCNAs.[2,27] This analysis rigorously assessed ctDNA TFx via ULP-WGS assay sensitivity, specificity, repeatability, precision, reportable range, reference intervals, and quality metrics as well as pre-analytic factors, including tube type and detection thresholds with respect to DNA quantity and quality, validating its performance.

A key goal of this validation work was to determine lower limit of detection and assess repeatability and reproducibility. By using experimental and *in silico* dilution experiments as well as replicates of the same specimen across settings, it was determined that sensitivity to detect TFx of 3% ranged from 97.2% to 100% at both 1× and 0.1× mean depth down-sampling. MNSD, a marker of variability, increased significantly below 3%, thus establishing 3% TFx as our lower limit of detection, in line with our published work based on *in silico* experimentation.[15] Thus, the ULP-WGS TFx assay reliably assesses ctDNA positive (TFx $\geq$ 3%) versus negative (TFx $<$ 3%) samples.

A comparison of TFx between ULP-WGS and WES revealed varied estimates at high TFx, likely due to the presence of CNLOH (Figure 2C). ichorCNA was optimized for detecting cancer and estimating TFx when the expected TFx is low (ie, <20%). When TFx is higher (eg, TFx $>$ 20%), small relative variability may be acceptable because those samples will clearly be ctDNA positive. We also hypothesize that subclonal CNA events could impact estimated TFx. This may impact future work on ctDNA dynamics, particularly in high TFx cancer types, like triple-negative breast cancer, and a goal is for future iterations of ichorCNA to address the potential effects of CNLOH or other similar factors.

We also investigated sequencing performance metrics in the context of pre-analytic variability, including tube type, time to processing, and DNA input. As in other studies and guidelines,[17] EDTA blood tubes and Streck tubes had comparable performance (within 5% mean absolute percentage error of TFx estimates) for time points within 8 hours to fractionation. This analytic difference is determined to be acceptable, and either tube may be used to collect blood for this assay. Furthermore, low cfDNA input was found to be required (5 ng acceptable), reinforcing that this assay can be applied using only a fraction of cfDNA from a single plasma sample, potentially reserving the remainder for downstream assays. The study demonstrated good

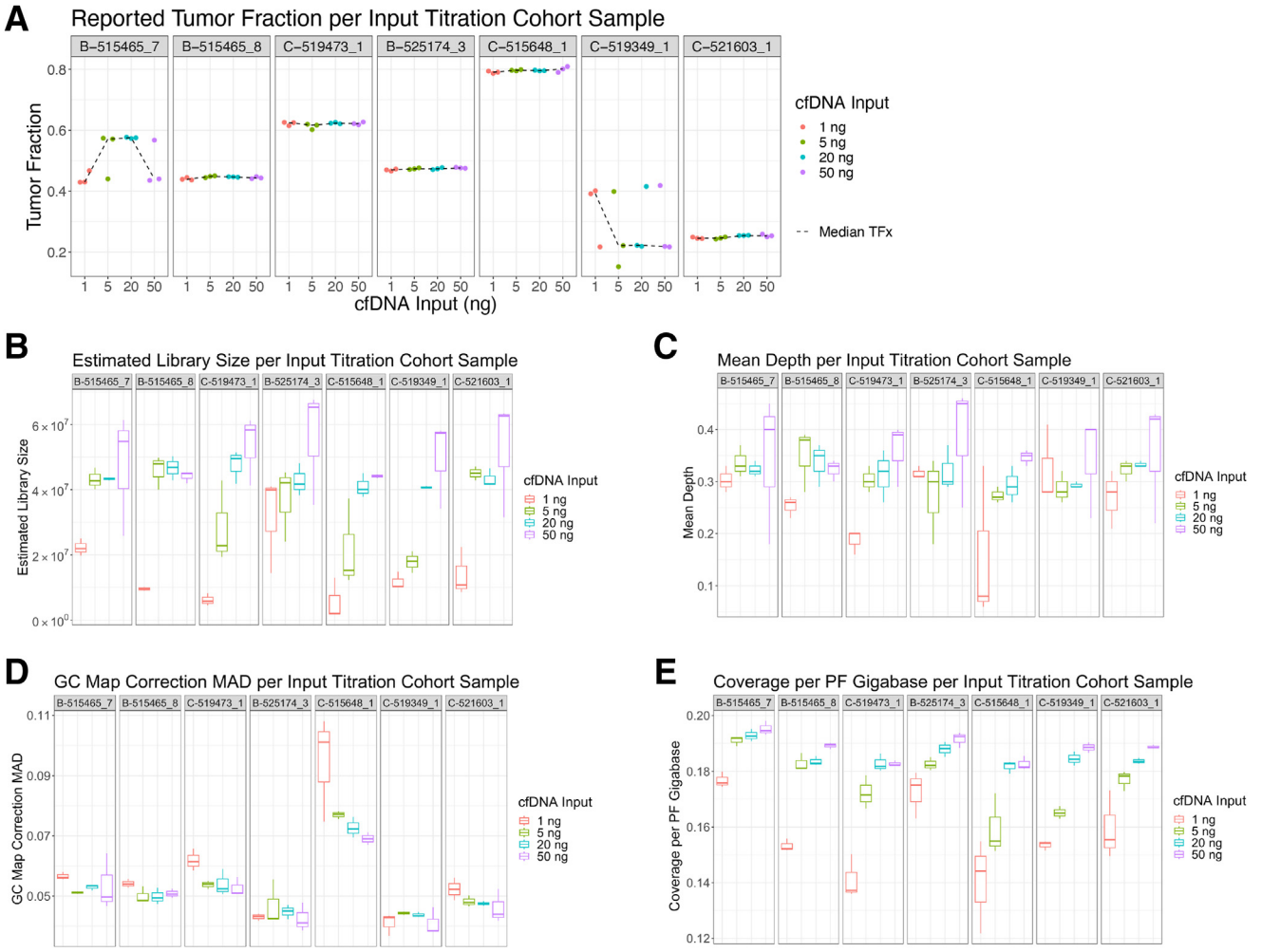



**Figure 5** Impact of DNA input for shallow whole-genome sequencing on tumor fraction determination from cell-free DNA. Four DNA input amounts (1/5/20/50 ng) were evaluated from seven distinct samples, each with adequate cell-free DNA content for all input amounts. **A:** Calculated tumor fraction was similar across time points for most samples, with greatest variation at 1 ng. **B**–**E:** Sequencing metrics of distinct DNA input amounts demonstrate greater variation for 1 ng than other input amounts across metrics, including library size (**B**), mean sequencing depth (**C**), mean absolute deviation (MAD; **D**), and coverage (**E**). cfDNA, cell-free DNA; PF, pass filter.

performance using distinct sequencing instrumentation, important given the dynamic nature of sequencing technologies over time.

Clinically, approaches to quantify TFx have been incorporated into commercially available ctDNA assays, including Guardant360,[28] FoundationOne Liquid CDX,[4] and cf-IMPACT,[29] among others. In one study analyzing the FoundationOne Liquid CDX assay, Reichert et al[4] evaluated TFx in a real-world setting among 1725 patients with four common cancer types, including metastatic castration-resistant prostate cancer, metastatic breast cancer, advanced non–small-cell lung cancer, or metastatic colorectal cancer. At a prespecified threshold of 10%, TFx was found to have a consistent and independent association with survival across cancer types.[4] Further work from the same group/assay on a larger dataset of 23,482 samples demonstrates that, not surprisingly, the TFx correlates with the sensitivity to detect driver alterations in ctDNA.[30] Specifically, they detect driver alterations identified in tissue biopsy from the same patient in 58% to 86% of patients, yet consistently at or near 100% in cases with TFx $\geq$10%.[30] The cf-IMPACT group assessed whether cfDNA TFx estimation through low-pass, shallow whole-genome sequencing, fragment size analysis, or both could facilitate the interpretation of negative cfDNA results.[29] Of the 47 samples without alterations detected and low TFx (z-score $<$5), 29 had sufficient material to be re-analyzed using a less comprehensive but more sensitive assay, suggesting that estimation of cfDNA TFx can facilitate the interpretation of cfDNA and help guide the selection of subsequent analysis in patients with negative results.[29] TFx is now routinely reported on FoundationOne Liquid CDX and other clinical assay reports, reinforcing the broad potential utility of TFx as a clinical assay.

In this validation analysis, we rigorously assessed multiple aspects of the ULP-WGS–based TFx assay, but there

are certainly limitations. All samples used in this validation were from patients with advanced/metastatic breast cancer. Overall, the Broad Institute platform has profiled >10,000 plasma sample from a diverse set of cancer types without obvious performance concerns within any individual cancer type, but additional validation across cancer types may be warranted. We identified the potential for TFx underestimation in the setting of CNLOH and subclonal copy number events, which is an important finding for the field and hopefully addressable in future versions of ichorCNA. For the gold standard TFx, the purity was determined from WES using ABSOLUTE,[14] but we acknowledge that the performance of ABSOLUTE was designed on higher tumor content tissue samples so alternative gold standard options could be considered. This assay also focuses on determination of TFx from ULP-WGS, and specific and/or driver tumor mutations, tumor mutation burden, clonal hematopoiesis, or any of the potential orthogonal metrics, such as specific CNAs, were not assessed.

## Conclusions

TFx determination through ULP-WGS of cell-free DNA and ichorCNA offers precise, repeatable, reproducible quantitation of TFx, facilitating assay application in clinical cancer care.

## Acknowledgments

We acknowledge clinical support from Catherine Carson, Katherine Weber, Celia Garr, and Danielle Nickell.

## Author Contributions

G.H., V.A.A., C.C., N.J.L., and D.G.S. designed the study; S.P., M.H., H.L., K.G., D.H.O., K.A.C., and D.G.S. collected and coordinated samples; M.R.-Y., G.T., J.T., S.P., K.A.C., G.H., V.A.A., C.C., N.J.L., and D.G.S. collected and analyzed data; and all authors wrote and approved the manuscript. All authors guarantee this work and, as such, had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

## Disclosure Statement

G.H. has received intellectual property rights/is a patent holder for Broad Institute. V.A.A. has received intellectual property rights/is a patent holder for Broad Institute; and is on advisory boards for AGCT GmbH and Bertis, Inc. D.G.S. is on the advisory board for Novartis.

## Supplemental Data

Supplemental material for this article can be found at *http://doi.org/10.1016/j.jmoldx.2024.01.014*.

## References

1. O'Leary B, Cutts RJ, Huang X, Hrebien S, Liu Y, André F, Loibl S, Loi S, Garcia-Murillas I, Cristofanilli M, Bartlett CH, Turner NC: Circulating tumor DNA markers for early progression on fulvestrant with or without palbociclib in ER+ advanced breast cancer. J Natl Cancer Inst 2020, 113:309−317
2. Stover DG, Parsons HA, Ha G, Freeman SS, Barry WT, Guo H, Choudhury AD, Gydush G, Reed SC, Rhoades J, Rotem D, Hughes ME, Dillon DA, Partridge AH, Wagle N, Krop IE, Getz G, Golub TR, Love JC, Winer EP, Tolaney SM, Lin NU, Adalsteinsson VA: Association of cell-free DNA tumor fraction and somatic copy number alterations with survival in metastatic triple-negative breast cancer. J Clin Oncol 2018, 36:543−553
3. Choudhury AD, Werner L, Francini E, Wei XX, Ha G, Freeman SS, Rhoades J, Reed SC, Gydush G, Rotem D, Lo C, Taplin ME, Harshman LC, Zhang Z, O'Connor EP, Stover DG, Parsons HA, Getz G, Meyerson M, Love JC, Hahn WC, Adalsteinsson VA: Tumor fraction in cell-free DNA as a biomarker in prostate cancer. JCI Insight 2018, 3:e122109
4. Reichert ZR, Morgan TM, Li G, Castellanos E, Snow T, Dall'Olio FG, Madison RW, Fine AD, Oxnard GR, Graf RP, Stover DG: Prognostic value of plasma circulating tumor DNA fraction across four common cancer types: a real-world outcomes study. Ann Oncol 2023, 34:111−120
5. Dandachi N, Posch F, Graf R, Suppan C, Klocker EV, Müller HD, Lindenmann J, Terbuch A, Heitzer E, Balic M: Longitudinal tumor fraction trajectories predict risk of progression in metastatic HR+ breast cancer patients undergoing CDK4/6 treatment. Mol Oncol 2021, 15:2390−2400
6. Avanzini S, Kurtz DM, Chabon JJ, Moding EJ, Hori SS, Gambhir SS, Alizadeh AA, Diehn M, Reiter JG: A mathematical model of ctDNA shedding predicts tumor detection size. Sci Adv 2020, 6:eabc4308
7. Kurtz DM, Soo J, Co Ting Keh L, Alig S, Chabon JJ, Sworder BJ, Schultz A, Jin MC, Scherer F, Garofalo A, Macaulay CW, Hamilton EG, Chen B, Olsen M, Schroers-Martin JG, Craig AFM, Moding EJ, Esfahani MS, Liu CL, Dührsen U, Hüttmann A, Casasnovas R-O, Westin JR, Roschewski M, Wilson WH, Gaidano G, Rossi D, Diehn M, Alizadeh AA: Enhanced detection of minimal residual disease by targeted sequencing of phased variants in circulating tumor DNA. Nat Biotechnol 2021, 39:1537−1547
8. Polski A, Xu L, Prabakar RK, Kim JW, Shah R, Jubran R, Kuhn P, Cobrinik D, Hicks J, Berry JL: Cell-free DNA tumor fraction in the aqueous humor is associated with therapeutic response in retinoblastoma patients. Transl Vis Sci Technol 2020, 9:30
9. Bredno J, Lipson J, Venn O, Aravanis AM, Jamshidi A: Clinical correlates of circulating cell-free DNA tumor fraction. PLoS One 2021, 16:e0256436
10. Dawson SJ, Tsui DW, Murtaza M, Biggs H, Rueda OM, Chin SF, Dunning MJ, Gale D, Forshew T, Mahler-Araujo B, Rajan S, Humphray S, Becq J, Halsall D, Wallis M, Bentley D, Caldas C, Rosenfeld N: Analysis of circulating tumor DNA to monitor metastatic breast cancer. N Engl J Med 2013, 368:1199−1209
11. Schiavon G, Hrebien S, Garcia-Murillas I, Cutts RJ, Pearson A, Tarazona N, Fenwick K, Kozarewa I, Lopez-Knowles E, Ribas R, Nerurkar A, Osin P, Chandarlapaty S, Martin LA, Dowsett M, Smith IE, Turner NC: Analysis of ESR1 mutation in circulating tumor DNA demonstrates evolution during therapy for metastatic breast cancer. Sci Transl Med 2015, 7:313ra182
12. Garcia-Murillas I, Schiavon G, Weigelt B, Ng C, Hrebien S, Cutts RJ, Cheang M, Osin P, Nerurkar A, Kozarewa I, Garrido JA, Dowsett M, Reis-Filho JS, Smith IE, Turner NC: Mutation tracking in circulating

tumor DNA predicts relapse in early breast cancer. Sci Transl Med 2015, 7:302ra133
13. Ha G, Roth A, Khattra J, Ho J, Yap D, Prentice LM, Melnyk N, McPherson A, Bashashati A, Laks E, Biele J, Ding J, Le A, Rosner J, Shumansky K, Marra MA, Gilks CB, Huntsman DG, McAlpine JN, Aparicio S, Shah SP: TITAN: inference of copy number architectures in clonal cell populations from tumor whole-genome sequence data. Genome Res 2014, 24:1881−1893
14. Carter SL, Cibulskis K, Helman E, McKenna A, Shen H, Zack T, Laird PW, Onofrio RC, Winckler W, Weir BA, Beroukhim R, Pellman D, Levine DA, Lander ES, Meyerson M, Getz G: Absolute quantification of somatic DNA alterations in human cancer. Nat Biotechnol 2012, 30:413−421
15. Adalsteinsson VA, Ha G, Freeman SS, Choudhury AD, Stover DG, Parsons HA, et al: Scalable whole-exome sequencing of cell-free DNA reveals high concordance with metastatic tumors. Nat Commun 2017, 8:1324
16. O'Leary B, Hrebien S, Morden JP, Beaney M, Fribbens C, Huang X, Liu Y, Bartlett CH, Koehler M, Cristofanilli M, Garcia-Murillas I, Bliss JM, Turner NC: Early circulating tumor DNA dynamics and clonal selection with palbociclib and fulvestrant for breast cancer. Nat Commun 2018, 9:896
17. Merker JD, Oxnard GR, Compton C, Diehn M, Hurley P, Lazar AJ, Lindeman N, Lockwood CM, Rai AJ, Schilsky RL, Tsimberidou AM, Vasalos P, Billman BL, Oliver TK, Bruinooge SS, Hayes DF, Turner NC: Circulating tumor DNA analysis in patients with cancer: American Society of Clinical Oncology and College of American Pathologists Joint Review. J Clin Oncol 2018, 36:1631−1641
18. Ignatiadis M, Sledge GW, Jeffrey SS: Liquid biopsy enters the clinic - implementation issues and future challenges. Nat Rev Clin Oncol 2021, 18:297−312
19. Lui YY, Chik KW, Chiu RW, Ho CY, Lam CW, Lo YM: Predominant hematopoietic origin of cell-free DNA in plasma and serum after sex-mismatched bone marrow transplantation. Clin Chem 2002, 48:421−427
20. Bettegowda C, Sausen M, Leary RJ, Kinde I, Wang Y, Agrawal N, et al: Detection of circulating tumor DNA in early- and late-stage human malignancies. Sci Transl Med 2014, 6:224ra224
21. Abbosh C, Birkbak NJ, Wilson GA, Jamal-Hanjani M, Constantin T, Salari R, et al: Phylogenetic ctDNA analysis depicts early-stage lung cancer evolution. Nature 2017, 545:446−451
22. Tug S, Helmig S, Deichmann ER, Schmeier-Jürchott A, Wagner E, Zimmermann T, Radsak M, Giacca M, Simon P: Exercise-induced increases in cell free DNA in human plasma originate predominantly from cells of the haematopoietic lineage. Exerc Immunol Rev 2015, 21:164−173
23. Risberg B, Tsui DWY, Biggs H, Ruiz-Valdepenas Martin de Almagro A, Dawson SJ, Hodgkin C, Jones L, Parkinson C, Piskorz A, Marass F, Chandrananda D, Moore E, Morris J, Plagnol V, Rosenfeld N, Caldas C, Brenton JD, Gale D: Effects of collection and processing procedures on plasma circulating cell-free DNA from cancer patients. J Mol Diagn 2018, 20:883−892
24. Zhou X, Cheng Z, Dong M, Liu Q, Yang W, Liu M, Tian J, Cheng W: Tumor fractions deciphered from circulating cell-free DNA methylation for cancer early diagnosis. Nat Commun 2022, 13:7694
25. Eeckhoutte A, Houy A, Manié E, Reverdy M, Bièche I, Marangoni E, Goundiam O, Vincent-Salomon A, Stoppa-Lyonnet D, Bidard FC, Stern MH, Popova T: ShallowHRD: detection of homologous recombination deficiency from shallow whole genome sequencing. Bioinformatics 2020, 36:3888−3889
26. Doebley AL, Ko M, Liao H, Cruikshank AE, Santos K, Kikawa C, Hiatt JB, Patton RD, De Sarkar N, Collier KA, Hoge ACH, Chen K, Zimmer A, Weber ZT, Adil M, Reichel JB, Polak P, Adalsteinsson VA, Nelson PS, MacPherson D, Parsons HA, Stover DG, Ha G: A framework for clinical cancer subtyping from nucleosome profiling of cell-free DNA. Nat Commun 2022, 13:7475
27. Van Roy N, Van Der Linden M, Menten B, Dheedene A, Vandeputte C, Van Dorpe J, Laureys G, Renard M, Sante T, Lammens T, De Wilde B, Speleman F, De Preter K: Shallow whole genome sequencing on circulating cell-free DNA allows reliable noninvasive copy-number profiling in neuroblastoma patients. Clin Cancer Res 2017, 23:6305−6314
28. Vu P, Khagi Y, Riviere P, Goodman A, Kurzrock R: Total number of alterations in liquid biopsies is an independent predictor of survival in patients with advanced cancers. JCO Precis Oncol 2020, 4:PO.19.00204
29. Tsui DWY, Cheng ML, Shady M, Yang JL, Stephens D, Won H, et al: Tumor fraction-guided cell-free DNA profiling in metastatic solid tumor patients. Genome Med 2021, 13:96
30. Husain H, Pavlick DC, Fendler BJ, Madison RW, Decker B, Gjoerup O, Parachoniak CA, McLaughlin-Drubin M, Erlich RL, Schrock AB, Frampton GM, Das Thakur M, Oxnard GR, Tukachinsky H: Tumor fraction correlates with detection of actionable variants across > 23,000 circulating tumor DNA samples. JCO Precis Oncol 2022, 6:e2200261