Adapting Clinical Chemistry Plasma as a Source for Liquid Biopsies

Spencer C. Ding[1], Jingru Yu[1], Tiepeng Liao[1], Lauren S. Ahmann[2], Yvette Y. Yao[3], Chandler Ho[4], Linlin Wang[5], Benjamin A. Pinsky[1,6], Wei Gu[1,*]

[1] Department of Pathology, School of Medicine, Stanford University, Stanford, USA

[2] Department of Molecular and Medical Genetics, Oregon Health & Science University, Portland, Oregon, USA

[3] School of Medicine, University of Calgary, Calgary, Canada

[4] Clinical Laboratories, Stanford Health Care, Stanford, USA

[5] Department of Laboratory Medicine, School of Medicine, University of California San Francisco, San Francisco, USA

[6] Department of Medicine, Division of Infectious Diseases and Geographic Medicine, School of Medicine, Stanford University, Stanford, USA

* To whom correspondence may be addressed.

Wei Gu

Assistant Professor, Department of Pathology, Stanford University

3373 Hillview Ave, room 220

Palo Alto, CA 94304-1204

weigu@stanford.edu

Keywords: Clinical chemistry plasma, cell-free DNA, metagenomics, CNAs, methylomes, fragmentomics

Running head: Clinical Chemistry Plasma for Liquid Biopsy

ABSTRACT

BACKGROUND: Circulating cell-free DNA (cfDNA) has become a valuable analyte for molecular testing, but requires specialized collection tubes or immediate processing. We investigated the feasibility of using residual plasma from heparin separators, which are routinely used in clinical chemistry, as an accessible and underutilized source for cfDNA biobanking and testing.

METHODS: We analyzed matched plasma samples from healthy volunteers in two experiments: an immediate-processing tube comparison across EDTA, Streck, and heparin separators (n = 5) and a clinical-handling simulation that paired EDTA and heparin separator tubes and delayed processing at room temperature versus 4°C (n = 6). We also analyzed matched EDTA and heparin separator plasma samples from viral PCR-positive patients (Hospital Cohort; n =38). Whole-genome sequencing and genome-wide enriched methylation sequencing were performed to evaluate concordance across multiple benchmarks, including metagenomics, chromosomal copy number, methylome, and fragmentomics.

RESULTS: Under immediate processing, heparin separator plasma showed high concordance with EDTA and Streck plasma for methylation patterns (Spearman's ρ=0.65-0.70) and fragmentation features (n = 5). In the clinical-handling simulation, cfDNA integrity in heparin separators was comparable to that in EDTA at 4°C (n=6). In the Hospital Cohort, heparin separators showed a strong concordance with matched EDTA tubes for viral detection (n=38, Spearman's ρ=0.95), copy number alteration profiling (n=6, Spearman's ρ=0.72-0.96), and methylation patterns (n=12, Spearman's ρ=0.50-0.83).

CONCLUSION: Hospital residual plasma from routine clinical chemistry tests that are processed within a short pre-centrifugation window and refrigerated can provide a vast, untapped resource for cfDNA biobanking and potential testing.

INTRODUCTION

Plasma cell-free DNA (cfDNA) is widely used in clinical testing(1), including non-invasive prenatal testing (NIPT)(2–5), cancer testing(6–8), infectious disease testing(9,10), and organ transplantation monitoring(11,12). In cancer testing, cfDNA is actively studied for the detection of residual cancer disease(13) and early detection(14,15).

Reliable cfDNA analysis typically depends on specialized preservative tubes (e.g., Streck) or rapid processing of EDTA tubes(16). However, broader implementation is limited by the requirement for specific tubes or on-call personnel. In contrast, heparin separators are now ubiquitous in daily blood draws at many hospitals for routine chemistry testing, such as the basic metabolic panel. At Stanford Health Care alone, over 900,000 such tubes are processed annually. After testing, there is typically a per-tube surplus of 0.5-2mL of plasma with cfDNA, which is discarded.

Despite its availability and appeal as a residual sample source, heparinized plasma has historically been avoided for molecular testing because (i) heparin is a well-established inhibitor of PCR, discouraging its use in PCR-based assays(17–19), (ii) Genomic DNA (gDNA) leakage and (iii) cfDNA degradation. Prior studies comparing heparinized and EDTA-plasma samples have reported mixed results and often do not distinguish between gDNA contamination and cfDNA degradation in PCR measurements. Barra et al. and Gerber et al. reported elevated cfDNA levels in heparin tubes compared to specialized tubes such as PAXgene, Roche cfDNA, and Streck, even at time point 0 (20,21). In contrast, Lo et al. and Hebels et al. reported comparable cfDNA yield and integrity between EDTA and heparin plasma when samples were processed promptly(22) or stored at -80°C within 8 hours(23). Using ddPCR, van Ginkel et al. reported a lower cfDNA concentration in heparin samples compared with EDTA and serum

samples(24). Notably, most existing studies relied on PCR-based assays for overall DNA quantification rather than quantification of cfDNA sequences detectable by NGS. Tracking sequenceable cfDNA directly provides measurements required for most cfDNA testing.

In the current work, we focus on heparin separator tubes, which differ from plain heparin tubes in that they contain a gel barrier that physically separates plasma from cellular components after centrifugation. Common chemistry tests are typically performed on-site at hospitals and affiliated clinics. These tubes are processed through high-throughput, automated clinical workflows designed to support rapid turnaround in a regulated environment. Moreover, typical NGS workflows include multiple purification steps that reduce heparin and gDNA contamination before the final PCR reaction, and adapter-ligation workflows further reduce gDNA carryover. Together, these features motivate an evaluation of whether ubiquitous heparin separators are a fit for modern sequencing-based cfDNA analysis.

To assess the feasibility of using heparin separators for cfDNA testing, we conducted whole-genome sequencing (WGS) and the FLEXseq methylation assay(25) on paired samples collected from the same venipuncture in both heparin separators and gold standard EDTA (and/or Streck) tubes. Our evaluation spanned multiple benchmarks relevant to a broad assortment of cfDNA tests, including metagenomics(9,26), methylation profiling(25,27), copy number alterations (CNAs)(28), epigenetic tissue-of-origin mapping(27), and cfDNA fragmentomics(29).

RESULTS

Heparin separators preserve cfDNA integrity under ideal conditions

To evaluate the potential of repurposing leftover plasma from heparin separators, we conducted a controlled Tube Comparison Study (n=5) in the Healthy Cohort using matched samples collected from the same blood draw in heparin separators and in gold-standard EDTA and Streck tubes (Fig. 1a, Table S1).

The heparin separator samples showed coverage profiles comparable to the matched EDTA samples (Supplementary Fig.1, Spearman's $\rho = 1.00$). Fragment size distributions were highly similar across all tube types, with a modal peak at ~166 bp, consistent with known cfDNA profiles (Fig. 1b)(4,31). End motif analysis of 256 4-mers showed comparable rankings (Fig. 1c), suggesting that heparin separators do not perturb fragmentation patterns when processed immediately.

To assess the preservation of epigenomic features, we calculated methylation levels at individual CpG sites using the FLEXseq methylation assay's output. CpG sites with ≥30× coverage showed correlations between heparin separator and EDTA (Spearman's ρ=0.70, Fig. 1d, left), and between heparin separator and Streck (Spearman's ρ=0.69, Fig. 1d, right). Tissue deconvolution revealed highly concordant cell type proportions across all collection tube types (Fig. 1e, P = 0.86, PERMANOVA), with consistent results across all five cases (Heparin separator versus EDTA: Spearman's ρ= 0.65-0.70; heparin separator versus Streck: Spearman's ρ= 0.65-0.69; Supplemental Fig. 2). These findings show that heparin separators preserve the biological status of methylation under ideal processing conditions.

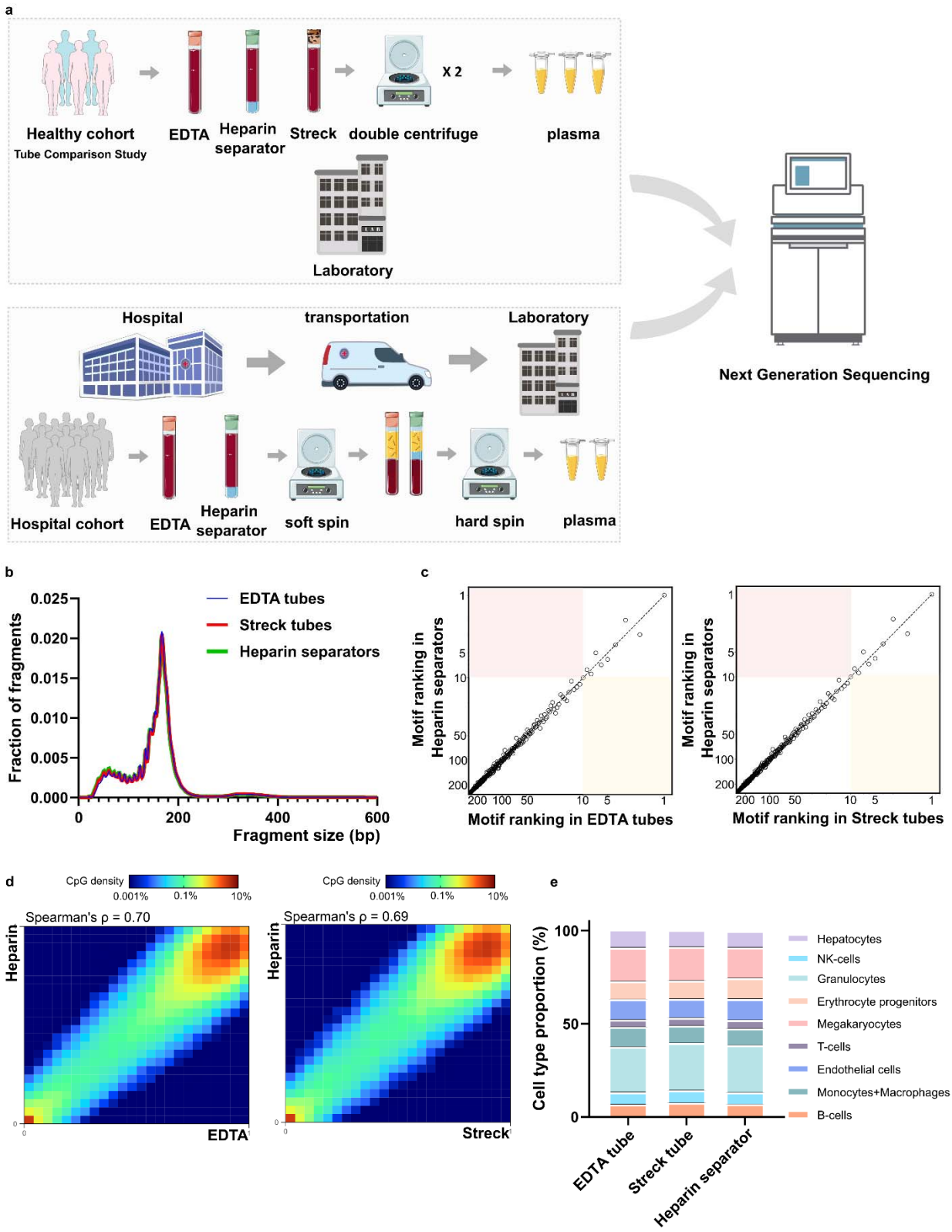


Figure 1. Overall schematic and Healthy Cohort's comparative performance

a. Schematic illustration of the experimental design across two cohorts. The Healthy Cohort was a controlled experiment where plasma samples from healthy volunteer donors were collected freshly in EDTA, Streck, and heparin separators. The Hospital Cohort plasma samples were retrospectively collected in EDTA and heparin separators as part of patient care and leftover from routine clinical testing. The Hospital Cohort samples underwent an initial soft spin during the routine clinical workflow and then a hard spin after refrigerated storage. Processed samples were analyzed using whole-genome sequencing (WGS) and/or FLEXseq methylation testing.

b. Healthy Cohort fragment size distribution of cfDNA collected across the three tubes.

c. Healthy Cohort comparison of end motif rankings in heparin separator samples versus EDTA or Streck tubes.

d. Healthy Cohort heatmap and correlation of methylation beta values in paired samples from a healthy donor (P140). Heparin separators (y-axis) are compared to EDTA and Streck tubes (x-axis).

e. Healthy Cohort estimation of cell-type proportions using methylation deconvolution for each collection tube.

To better reflect real-world handling procedures, we performed a controlled delay study in 6 healthy volunteers using paired EDTA and heparin separator tubes from the same blood draw (Fig. 2a, Table S2). When compared with the corresponding control samples that were processed at time 0, minimal changes in fragment size distributions were observed across EDTA samples at all incubation times, at either room temperature or 4 °C (Fig. 2b). Heparin separator samples stored at 4°C at all incubation times also show similar profiles (Fig. 2b). While in the heparin separator samples stored at room temperature showed a time-dependent shift towards shorter

fragment size, which indicates increasing degradation (Fig. 2b). As extreme control, incubation at 37 °C for 24 hours, as used in Barra et al., produced drastic fragmentation changes, reflecting a lot of degradation (Fig. 2b). Together, these results indicate that temperature before the first spin is a key determinant of cfDNA integrity in heparin separator tubes.

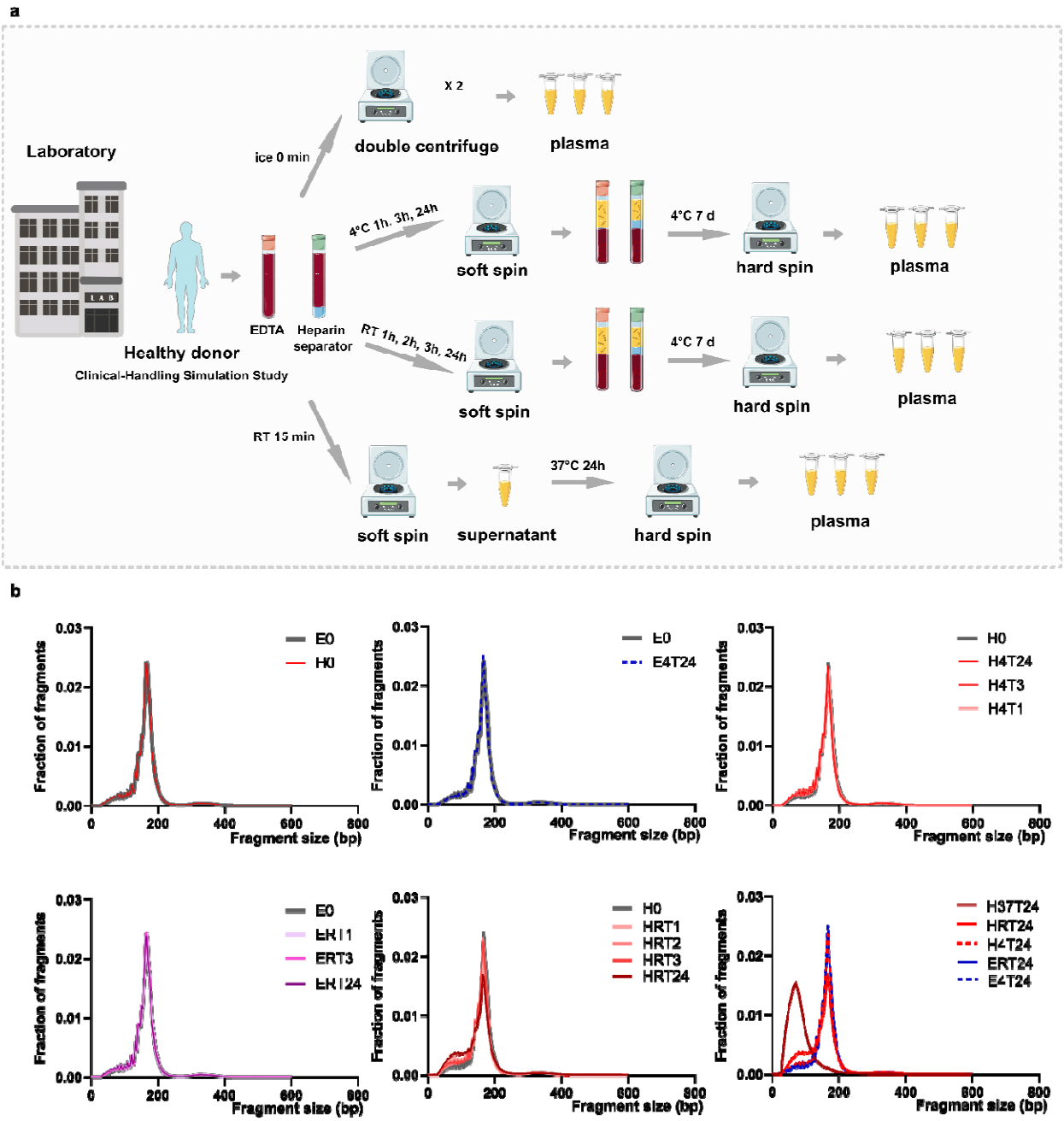


Figure 2. Degradation evaluation in the Clinical-Handling Simulation Study

a. Schematic of the Clinical-Handling Simulation Study. Whole blood from a healthy donor was collected in EDTA tubes and heparin separator tubes and processed either immediately (ice, 0 min; double-centrifuged plasma) or after defined delays under different temperature conditions. Plasma was generated by a soft spin followed by a hard spin, including short delays at 4°C or room temperature, extended storage at 4°C for 7 days, and a stress condition at 37°C for 24 h. b. cfDNA fragment size distributions across all conditions, shown as fraction of fragments versus fragment size. Condition codes denote tube type and handling condition, where E indicates EDTA, H indicates heparin separator, 4 indicates 4°C, RT indicates room temperature, 37 indicates 37°C, and T indicates the incubation time in hours, with 0 indicating immediate processing.

Hospital Cohort: Residual plasma from Heparin separators after clinical usage

To evaluate the feasibility of heparin separator tubes for cfDNA analysis, we identified 38 matched pairs of viral PCR–positive plasma samples collected in both heparin separator and EDTA tubes from the same blood draw. These cases were retrieved by querying the electronic medical record systems at Stanford (n=32) and UCSF (n=6) for samples that tested positive by the Clinical Virology Laboratories. All samples underwent cfDNA extraction, followed by whole-genome sequencing (WGS) to evaluate viral load, copy number alterations (CNAs), and fragmentation profiles, as well as methylation profiling via FLEXseq.

Leftover plasma from the heparin separators enables viral load detection

We performed metagenomic sequencing to compare viral load measurements and assess concordance between plasma collected in EDTA and heparin separators (Fig. 3). WGS achieved a median coverage of 4.33× (IQR: 2.89-4.88×). Viral read counts normalized to the number of

reads aligned to the human genome were highly correlated between plasma collection tube types (Spearman's ρ=0.95, $P < 0.001$), with no significant difference observed (Wilcoxon signed-rank test, $P = 0.27$). This high level of concordance was consistently observed across institutions and among distinct viral types.

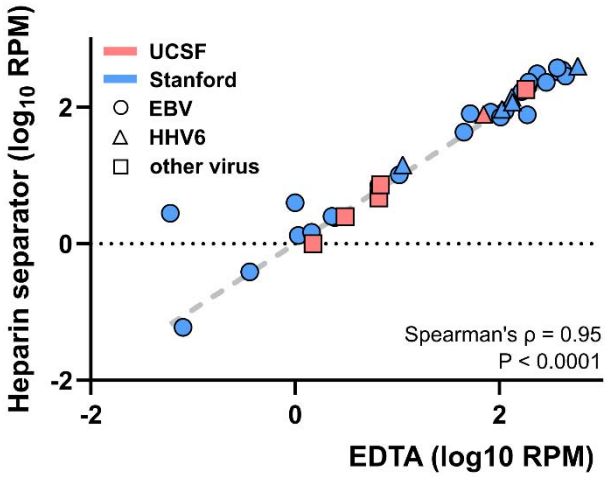


Figure 3. Viral load correlation of the Hospital Cohort

Correlation of plasma viral DNA reads detected by whole genome sequencing in paired cases collected in EDTA and heparin separators. Samples were collected at two institutions (Stanford - blue, UCSF - red, EBV - circle, HHV6 - triangle, other virus - square). RPM: reads per million human-aligned reads.

Leftover plasma from the heparin separator enables copy number profiling

CNAs are commonly observed in cancer and can be detected in cfDNA as a surrogate marker for tumor burden. To evaluate whether heparin separators preserve CNA signals, we compared the genome-wide copy number profiles of samples from heparin separators and EDTA tubes. CNA profiles were generated using a 500 kbp bin size, with a median of 97.15 million reads per sample (range: 0.73–360.73 million). Representative genome-wide CNA plots from a matched

pair (Patient 34) revealed similar patterns of chromosomal gains and losses (Fig. 4a). Genome-wide CNA plots for the remaining CNV-positive cases are shown in Supplementary file 1. To summarize concordance across all paired cases, we quantified agreement using correlation of the genome-wide log□ copy ratio and visualized the distribution stratified by CNV status (Fig. 4b). Similar results can be found in a smaller bin size of 50 kb (data uploaded to Zenodo) and a second CNV calling method, CNVkit (Supplementary Fig. 3) (32). CNV-positive samples (i.e., deduced tumor fraction ≥ 5%) exhibited consistently high concordance between tube types (n = 6, Spearman's ρ=0.72-0.96; Supplementary file 1). Tumor fraction estimates based on copy number also correlated strongly between matched samples (Spearman's ρ= 0.77; Fig. 4c). These findings demonstrate that plasma from heparin separators is suitable for CNA detection and tumor fraction estimation.

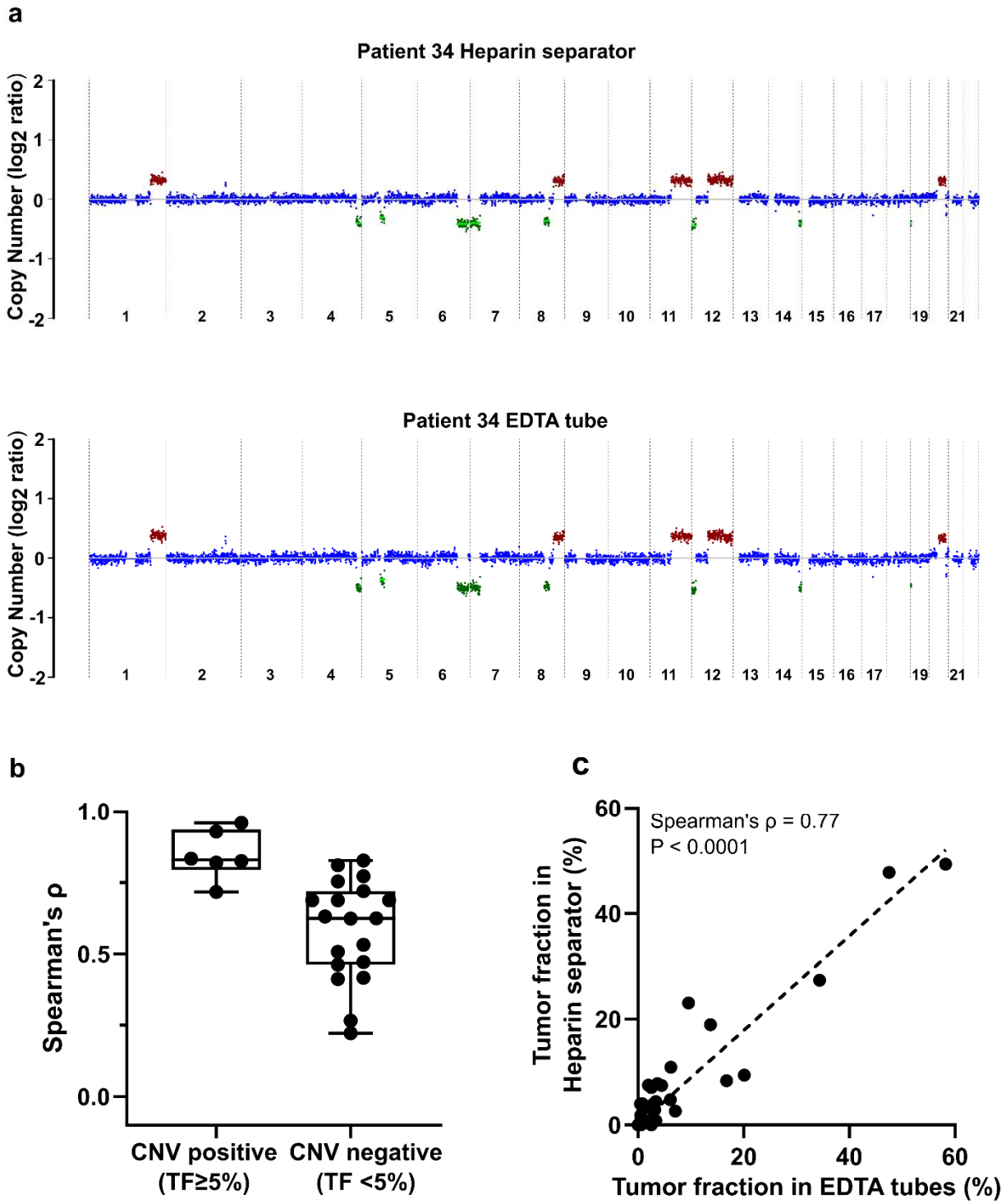


Figure 4. Copy number analysis of the Hospital Cohort

a. An example of genome-wide copy number profiles from Patient 34. Case P171 was collected in a heparin separator; Case P187 was collected in EDTA tube.

b. Boxplot of correlation of log□ copy ratio between tube types in all cases.

c. Comparison of tumor fraction in paired cases (n=32 pairs) inferred from copy number analysis.

Correlated Methylation Profiling from Heparin Separator Plasma

To assess the reliability of methylation profiling across blood collection tube types, we analyzed the available residual matched plasma samples in the Hospital Cohort (n=12). Using FLEXseq(25), a high-resolution methylation assay, we observed strong genome-wide concordance of CpG methylation levels between heparin separator and EDTA tubes (Patient 31; Spearman's ρ= 0.63; Fig. 5a; other cases in Supplementary Fig. 4). To evaluate the biological relevance of the cfDNA methylation profiles, we performed cell type deconvolution on the paired plasma(27). The results showed comparable tissue-of-origin proportions between plasma collected in heparin separator and EDTA tubes (Fig. 5b). These findings demonstrate that cfDNA methylation signatures are well preserved in heparin separator plasma and are reliable for epigenomic applications such as cell-of-origin estimation.

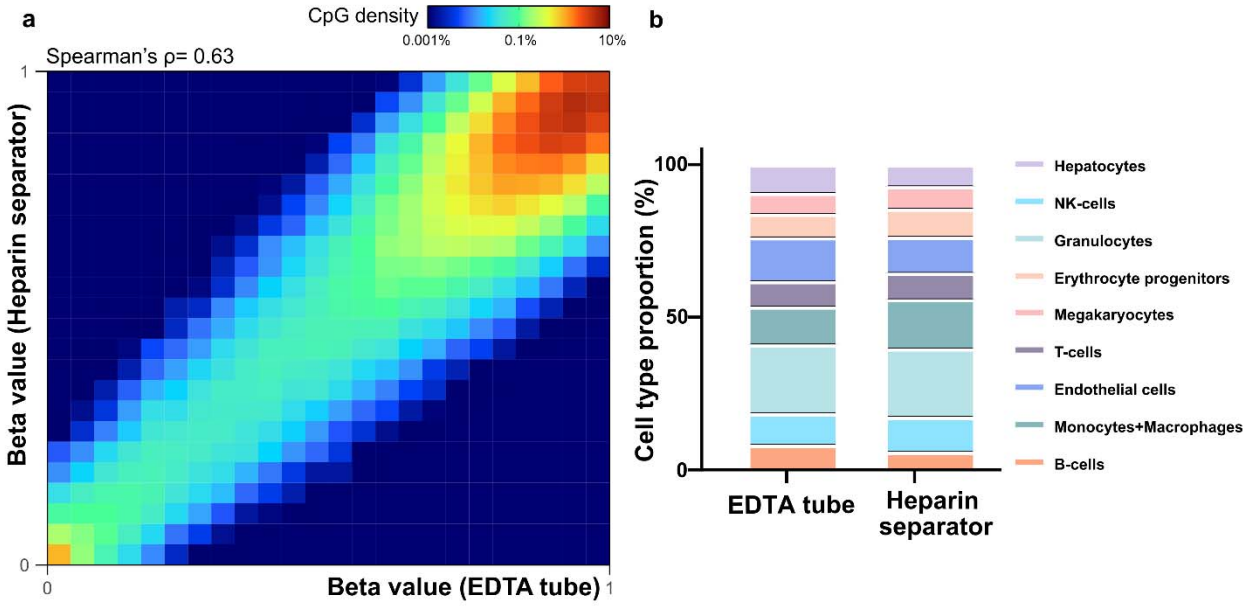


Figure 5. Methylation correlation and cell type deconvolution analysis in the Hospital Cohort

a. Heatmap showing methylation beta values in paired samples (Patient 31, P185 versus P168).

b. Estimated cell-type proportions in the paired samples (P185 versus P168)

Comparable cfDNA yield deduced from the sequencing result

We quantified cfDNA concentrations (normalized to plasma volume) in the 25 matched pairs from the Tube Comparison Study in the Healthy Cohort and the Hospital Cohort with sufficient leftover DNA. Qubit measurements revealed variability across samples (0.14–43.19 ng/mL; Fig. 6a & b) and poor correlation between EDTA and heparin separator data. However, Qubit measurements are related to both addressable (intact short fragments) and non-addressable (genomic length or broken) DNA, whereas NGS output is related to only addressable DNA. To enable precise measurements of the addressable DNA, we spiked in lambda DNA at a constant quantity across all samples at the beginning of the library preparation and normalized human-aligned read counts to this internal control. This strategy produces a measure (Human/Lambda ratio) proportional to the addressable DNA content of the sample. This normalized ratio yielded strong concordance between EDTA and heparin separator samples (R^2 = 0.998; Fig. 6c), indicating that library preparation effectively mitigated preanalytical variability.

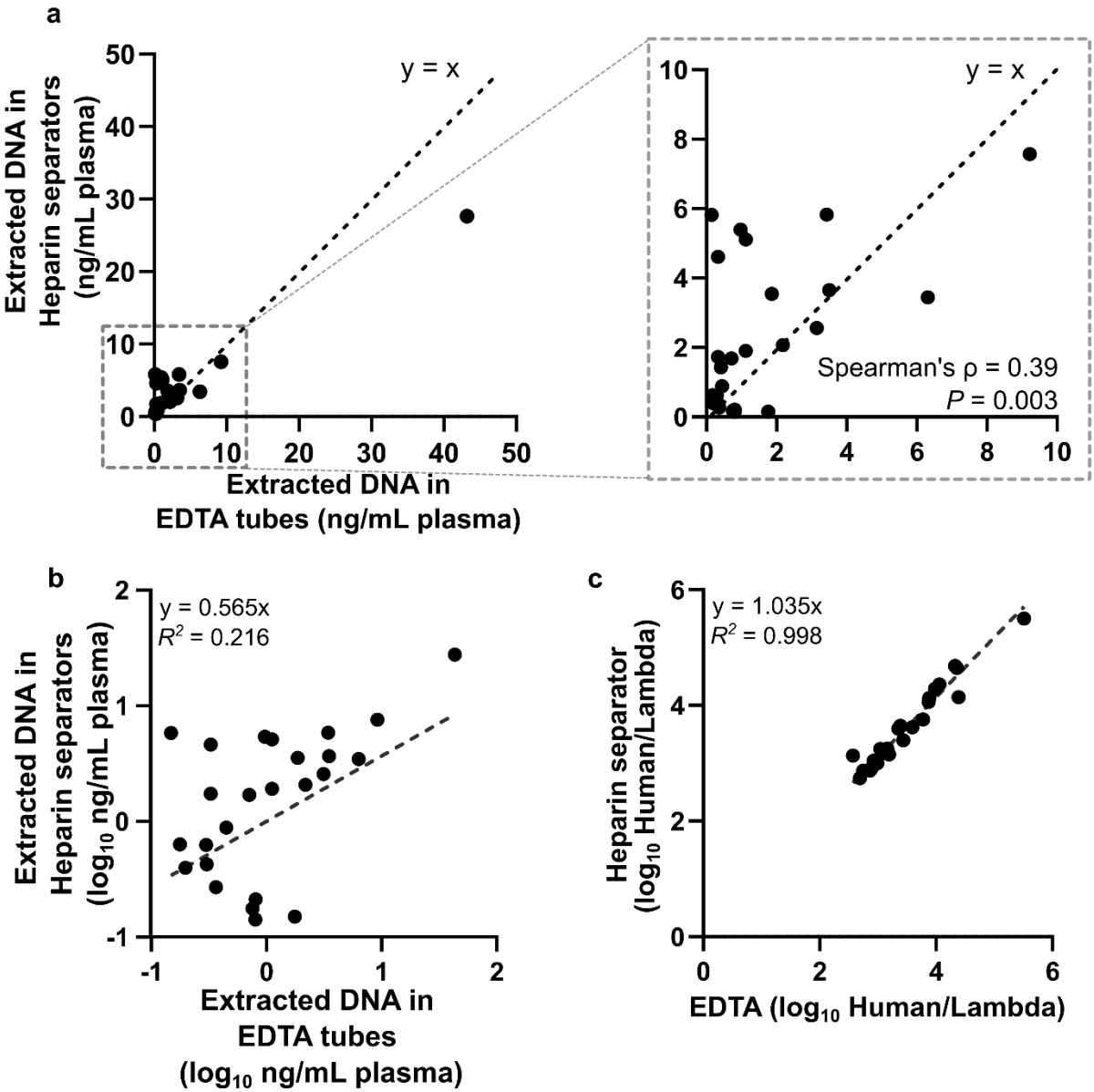


Figure 6. Correlation of DNA quantity

a. Scatter plot of the normalized extracted DNA concentration based on Qubit reading. Data is from 20 individuals of the Hospital Cohort and 5 individuals from the Healthy Cohort using heparin separators and EDTA tubes. Normalized DNA concentration (ng/mL plasma) in heparin separator samples (y-axis) vs. EDTA tube samples (x-axis), quantified by dsDNA Qubit and normalized to the input volume of plasma.

b. Scatter plot of the log-scale normalized concentration (Qubit readout) in the heparin separators (y-axis) vs. the EDTA tube samples (x-axis).

c. Scatter plot of the ratio of human reads to a constant lambda phage DNA spike-in detected in the heparin separator sample (y-axis) vs. the EDTA tube samples (x-axis). This spike-in measurement is proportional to the addressable DNA quantity of the sample.

Leftover plasma from the heparin separator alters fragmentation profiles

Fragmentation features of cfDNA, such as size distribution and end motifs, have been linked to nuclease activity and are known to reflect host physiology and pathology(33–35). To evaluate whether heparin separators preserve these fragmentation features, we examined two well-characterized fragmentomic features: size distribution and end-motif. Both EDTA and heparin separator plasma samples exhibited expected peak sizes near 70 bp and 170 bp (Supplementary Fig. 5). However, heparin separator plasma showed a modest increase in short fragments and a slightly diminished ~166 bp peak, indicating a shorter size distribution. Analysis of 4-mer motifs at the 5′ ends of cfDNA fragments revealed significant differences in four of the six most frequent motifs (i.e., CCCA, CCAG, CCTG, and CAAA; $P < 0.05$, paired Student's t-test), suggesting disrupted cleavage preference. These altered fragmentation profiles in the Hospital Cohort suggest that plasma samples collected in heparin separators may have influenced cfDNA fragmentation signatures.

DISCUSSION

Our study explores the untapped potential of repurposing residual plasma from heparin separators as a practical and underutilized resource for circulating cfDNA-based research and diagnostic applications. Using matched samples from the Healthy Cohort and a real-world Hospital Cohort, we demonstrate that the leftover plasma is suitable for metagenomic viral detection, copy number profiling, and genome-wide methylation analysis, with performance comparable to gold-standard collection methods.

Use of heparinized plasma has been historically limited by concerns over PCR inhibition, genomic DNA leakage, and cfDNA degradation(17–21). Our study helps clarify these concerns in the context of sequencing-based assays. First, despite the well-recognized PCR-inhibitory effect of heparin, a strong concordance was observed in viral load, CNV profiles, and methylation patterns. This suggests that residual heparin did not pose a practical limitation for these NGS-based assays, which, unlike PCR assays, undergo multiple rounds of DNA purification in the workflow. Second, NGS internal controls gave precise measurements of sequencable DNA and indicated comparable amounts of sequenceable cfDNA in matched EDTA and heparin separator samples. This suggests that the confounding DNA (e.g., gDNA carryover) had minimal effect. Third, degradation risk was directly reflected by fragmentomic analyses. Prior work showed that room-temperature incubation of heparinized blood can alter cfDNA fragmentation patterns, potentially via enhanced nuclease activity(36). However, controlled experiments mimicking clinical scenarios demonstrated that the storage temperature before the first spin is a key determinant of fragmentation integrity. Moreover, our controlled handling experiments indicate that minimizing the pre-centrifugation interval and maintaining refrigeration prior to the first spin are key to preserving cfDNA integrity in heparin separator

tubes. In practice, achieving a short pre-centrifugation window is generally more feasible for inpatient collections near the core laboratory than for outpatient draws at remote clinics, which may require additional cold-chain controls.

Limitations of using residual heparin separators from Chemistry tests include limited residual plasma volume (0.5–2 mL), which may not meet the input requirements of applications such as single nucleotide variants (SNV) or minimal residual disease detection. Similarly, the resultant limited sequencing depth is not optimized for high-resolution CNV detection but rather for larger-scale alterations. Moreover, variability in transport and storage conditions prior to processing, as observed in the Hospital Cohort, can affect cfDNA integrity. The observed alteration in fragmentation profiles highlights the need for cold-chain transportation protocols to minimize variability and maintain cfDNA quality. Therefore, we anticipate that this approach is optimal when specimens are from within the hospital and constrained to rapid turnaround times rather than from remote clinics. Future studies should also validate this approach in larger, clinically diverse populations and assess its integration into routine preanalytical workflows.

In summary, residual plasma from heparin separators can serve as a viable source for cfDNA analysis, with stable results for copy number variation, viral load, and methylation profiling despite fragmentomic sensitivity to preanalytical variability. We anticipate these specimens may be suitable for hospital biobanking, rapid and on-site testing, and scenarios where obtaining immediate informed consent in an emergent case is difficult. Repurposing residual plasma from routine blood tests could greatly simplify and expand access to cfDNA across a diverse range of patients.

METHODS

Study design and subject selection

The Healthy Cohort consisted of two complementary experiments: (i) a Tube Comparison Study (n=5; 28–40 years old) performed under immediate processing to assess tube-dependent effects, and (ii) a Clinical-Handling Simulation Study (n=6; 29-45 years old) that mimics the real-world delayed processing, evaluating the effect of pre-centrifugation delays at room temperature versus 4°C. All healthy adult volunteers provided written informed consent as approved under Stanford IRB (71230). In the Tube Comparison Study, whole blood from each donor was collected via a single venipuncture into EDTA, Streck, and heparin separators. In the Clinical-Handling Simulation Study, paired blood samples were collected into EDTA and heparin separators.

The Hospital Cohort consisted of clinical samples from two institutions, Stanford and UCSF. At Stanford, 32 matched pairs of remnant samples were obtained (collected between 2022 and 2025) under a no-patient-contact protocol (IRB 58461). Patients positive for Epstein-Barr virus (EBV), human herpesvirus 6 (HHV6), BK virus, adenovirus, or Torque Teno Virus were identified based on routine clinical viral PCR testing performed by the Stanford or UCSF Clinical Virology Laboratories. These viral tests were typically conducted in the context of post-transplant surveillance or infectious disease diagnostics. Residual EDTA plasma samples were processed by the clinical laboratory within four hours of collection. Matched heparin separator plasma specimens drawn at the same time were retrieved from the Stanford Clinical Chemistry Laboratory. Only matched plasma pairs with volumes exceeding 0.5 mL were included in the study. Additional sample pairs from UCSF (n=6) were obtained (collected between 2017 and 2020) under a similar no-patient-contact IRB protocol (CA-0161925).

Sample collection and processing

Healthy Cohort: In the Tube Comparison Study, samples were processed immediately after blood collection with two centrifugations (1,600 × g for 10 minutes and 16,000 × g for 10 minutes). In the Clinical-Handling Simulation Study (Fig. 2a), samples were incubated at either room temperature or 4°C for 0, 1, 2, 3, or 24 hours, similar to the pre-analytical journey of a blood draw prior to the first centrifugation (1,600 × g, 10 minutes). Post-spin tubes were stored at 4°C for one week to mimic the residual plasma handling process in the hospital, followed by a second centrifugation at 16,000 × g for 10 minutes. An additional subset of heparin separator samples (n=2) was incubated at 37°C for 24 hours as a degradation control.

Hospital Cohort: Upon arrival, plasma samples in EDTA tubes and heparin separators were immediately centrifuged at 1300 g for 10 minutes using a Cobas 8100 automated workflow. Following completion of routine clinical testing, the residual plasma was stored and transported at approximately 4°C for up to 8 days before further processing. EDTA plasma was separated into a different container, and heparin separator plasma was kept in the same original collection tube, but physically isolated above the gel separator. These procedures were conducted by the hospital staff as part of routine clinical care in accordance with a CLIA-compliant standard operating procedure. The plasma supernatant was identified by a research team, deidentified, transferred to our research lab at refrigeration, and subjected to a second centrifugation at 16,000 × g for 10 minutes. Plasma aliquots from both EDTA tubes and heparin separators were stored at –80°C until downstream analysis. UCSF cases were processed as previously described(26).

DNA extraction

Cell-free DNA (cfDNA) was extracted from 0.5 to 1 mL of plasma using the Maxwell RSC 48 instrument (Promega) with the Maxwell RSC ccfDNA Plasma Kit (AS1480), following the manufacturer's instructions. Samples were processed in parallel in batches. cfDNA

concentrations were quantified using the Varioskan spectrophotometer (Thermo Fisher Scientific). UCSF cases were extracted as previously described(26).

Whole genome sequencing, DNA library preparation

A median of 18.65 ng of extracted cfDNA (range: 3.00- 51.70 ng) and 2.5 ng of fragmented Lambda DNA (NEB #N3013) as internal control was used for whole-genome sequencing (WGS) library preparation with the NEBNext Ultra II DNA Library Prep Kit (New England Biolabs, #E7645L), following the manufacturer's protocol. PCR primers were NEBNext Multiplex Oligos for Illumina (NEB #E6444L). Adapter-ligated DNA was amplified for 20 cycles using the Q5 master mix, following the manufacturer's instructions. UCSF cases were processed as previously described(26).

FLEXseq DNA library preparation

FLEXseq DNA libraries were prepared as previously described. Briefly, extracted cfDNA molecules were adapter-ligated, dephosphorylated, digested with MspI (NEB, #R0106T), adapter-ligated a second time with custom adapters, subjected to NEBNext Enzymatic Methyl-seq Conversion Module (New England Biolabs, #E7125L), and PCR amplified.

Sequencing and alignment

Paired-end sequencing was performed on the NovaSeq X platform (Illumina) at the University of Chicago Genomics Core Facility using the 25B Reagent Kit (300 cycles). For WGS data, sequencing adapters were trimmed using Cutadapt v3.5 with the following parameters: -a AGATCGGAAGAGC -A AGATCGGAAGAGC --minimum-length 25. Trimmed reads were aligned to the hg38 human reference genome, as well as EBV, HHV6, and other viral genomes using BWA v0.7.17. PCR duplicates were removed using Samtools.

FLEXseq data were processed following the protocol described by Yu et al.(25). Briefly, reads were trimmed to remove adapters and filtered based on quality scores. High-quality reads were aligned to the hg38 genome using Bismark v0.23.0. Unconverted and duplicated reads were removed using Bismark. All SNP-containing regions were masked for de-identification using Bedtools.

Downstream analyses

***Viral load analysis*:** The ratio of reads aligned to viral genomes (e.g., EBV, HHV6) relative to every one million reads aligned to the human genome (hg38) was calculated and log-transformed. Spearman correlation analysis was performed between matched tube types, with the linear regression constrained to pass through the origin (0, 0).

***CNA analysis*:** Copy number alterations were analyzed using ichorCNA on autosomal chromosomes(30). HMMcopy's readCounter was used to generate 500 kb bin-level Wig files from aligned BAM files, which included only reads with a mapping quality score >20. Reference files for GC content (gcWig), mappability score (mapWig), and centromere positions were obtained from the ichorCNA GitHub repository corresponding to the selected bin size. Panel-of-normal (PoN) Wig files were generated separately using the cases in the Healthy Cohort in the EDTA and heparin separator groups, respectively. Tumor purity was estimated based on ichorCNA outputs.

***Methylation correlation analysis*:** Methylation calls were extracted using the bismark_methylation_extractor. For each CpG site, the methylation level was calculated as the ratio of methylated reads to total coverage depth. CpG sites with coverage ≥30× were retained for correlation analysis between matched samples.

***Tissue-of-origin analysis*:** Tissue-of-origin deconvolution was performed using cell type-specific unmethylated markers as previously described by Loyfer et al.(27). The top 250 unmethylated CpG markers for each cell type were extracted from the reference dataset. Markers located on sex chromosomes or overlapping with common single nucleotide polymorphisms (SNPs) were excluded. Deconvolution was conducted using the uxm function based on a panel of nine reference cell types: B cells, T cells, natural killer (NK) cells, granulocytes, monocytes/macrophages, endothelial cells, erythroid progenitors, megakaryocytes, and hepatocytes.

***cfDNA fragmentomics*:** Fragment size distributions were generated for each sample by calculating the frequency across fragment lengths between 50 and 600 bp. End motif frequencies were also calculated for each sample. The Shapiro-Wilk test was performed separately for each motif in different tubes (EDTA, Heparin separator, and Streck) to assess whether the frequency distributions followed a normal distribution. For the Healthy Cohort, differences among the three collection tubes were evaluated using the Friedman test. For the Hospital Cohort, statistical significance between tube types was assessed using the Paired Student's t-test. To control the false discovery rate (FDR), multiple testing corrections were applied using the Benjamini-Hochberg procedure across all tested motifs. Significance was determined based on FDR-adjusted *P*-values.

DATA AVAILABILITY

Raw WGS and genome-wide enriched methylation sequencing data are available for the Healthy Cohort (n = 11) with informed consent. FASTQ files were uploaded to the Sequence Read Archive (SRA), with the Bioproject ID of PRJNA1260066. Viral reads detailed statistic files (n=32), methylation calls (n=17), copy number plots (n=28), and fragment length files (n=34) are available on Zenodo (https://doi.org/10.5281/zenodo.15367093).

CONFLICT OF INTEREST

The authors J.Y., L.S.A., Y.Y.Y., and W.G. filed a patent application related to FLEXseq (PCT/US2024/056111).

ACKNOWLEDGEMENT

This work was funded by an NIH K08 (CA230156) grant, a Burroughs-Wellcome CAMS Award to WG. We thank the members of the Stanford Clinical Molecular Genetic Pathology, Clinical Virology, and Clinical Chemistry Laboratories, and UCSF Clinical Virology and Clinical Chemistry Laboratories for saving residual specimens. We thank Debbie Chan and Dianna Ng for their critical insights at the inception of this project and Ruben Luo for facilitating Clinical Chemistry collections at Stanford. We thank the University of Chicago Genomics Facility (RRID: SCR_019196) for Novaseq sequencing.

REFERENCES

1. Loy C, Ahmann L, De Vlaminck I, Gu W. Liquid Biopsy Based on Cell-Free DNA and RNA. Annu Rev Biomed Eng. 2024;

2. Lo YM, Corbetta N, Chamberlain PF, Rai V, Sargent IL, Redman CW, et al. Presence of fetal DNA in maternal plasma and serum. Lancet. 1997;350:485–7.

3. Fan HC, Blumenfeld YJ, Chitkara U, Hudgins L, Quake SR. Noninvasive diagnosis of fetal aneuploidy by shotgun sequencing DNA from maternal blood. PNAS. 2008;105:16266–71.

4. Lo YMD, Chan KCA, Sun H, Chen EZ, Jiang P, Lun FMF, et al. Maternal Plasma DNA Sequencing Reveals the Genome-Wide Genetic and Mutational Profile of the Fetus. Science Translational Medicine. American Association for the Advancement of Science; 2010;2:61ra91-61ra91.

5. Fan HC, Gu W, Wang J, Blumenfeld YJ, El-Sayed YY, Quake SR. Non-invasive prenatal measurement of the fetal genome. Nature. Nature Publishing Group; 2012;487:320–4.

6. Chan KCA, Woo JKS, King A, Zee BCY, Lam WKJ, Chan SL, et al. Analysis of Plasma Epstein–Barr Virus DNA to Screen for Nasopharyngeal Cancer. New England Journal of Medicine. Massachusetts Medical Society; 2017;377:513–22.

7. Chung DC, Gray DM, Singh H, Issaka RB, Raymond VM, Eagle C, et al. A Cell-free DNA Blood-Based Test for Colorectal Cancer Screening. New England Journal of Medicine. Massachusetts Medical Society; 2024;390:973–83.

8. Heitzer E, Ulz P, Geigl JB. Circulating Tumor DNA as a Liquid Biopsy for Cancer. Clinical Chemistry. 2015;61:112–23.

9. Gu W, Miller S, Chiu CY. Clinical Metagenomic Next-Generation Sequencing for Pathogen Detection. Annual Review of Pathology: Mechanisms of Disease. Annual Reviews; 2019;14:319–38.

10. Blauwkamp TA, Thair S, Rosen MJ, Blair L, Lindner MS, Vilfan ID, et al. Analytical and clinical validation of a microbial cell-free DNA sequencing test for infectious disease. Nat Microbiol. Nature Publishing Group; 2019;4:663–74.

11. De Vlaminck I, Valantine HA, Snyder TM, Strehl C, Cohen G, Luikart H, et al. Circulating Cell-Free DNA Enables Noninvasive Diagnosis of Heart Transplant Rejection. Science Translational Medicine. American Association for the Advancement of Science; 2014;6:241ra77-241ra77.

12. Bloom RD, Bromberg JS, Poggio ED, Bunnapradist S, Langone AJ, Sood P, et al. Cell-Free DNA and Active Rejection in Kidney Allografts. J Am Soc Nephrol. 2017;28:2221–32.

13. Lennon AM, Buchanan AH, Kinde I, Warren A, Honushefsky A, Cohain AT, et al. Feasibility of blood testing combined with PET-CT to screen for cancer and guide intervention. Science. 2020;

14. Shen SY, Singhania R, Fehringer G, Chakravarthy A, Roehrl MHA, Chadwick D, et al. Sensitive tumour detection and classification using plasma cell-free DNA methylomes. Nature. Nature Publishing Group; 2018;563:579–83.

15. Jamshidi A, Liu MC, Klein EA, Venn O, Hubbell E, Beausang JF, et al. Evaluation of cell-free DNA approaches for multi-cancer early detection. Cancer Cell. 2022;40:1537-1549.e12.

16. Meddeb R, Pisareva E, Thierry AR. Guidelines for the Preanalytical Conditions for Analyzing Circulating Cell-Free DNA. Clinical Chemistry. 2019;65:623–33.

17. Beutler E, Gelbart T, Kuhl W. Interference of heparin with the polymerase chain reaction. Biotechniques. 1990;9:166.

18. Yokota M, Tatsumi N, Nathalang O, Yamada T, Tsuda I. Effects of heparin on polymerase chain reaction for blood white cells. J Clin Lab Anal. 1999;13:133–40.

19. Jung R, Lübcke C, Wagener C, Neumaier M. Reversal of RT-PCR Inhibition Observed in Heparinized Clinical Specimens. BioTechniques. Taylor & Francis; 1997;23:24–8.

20. Barra GB, Santa Rita TH, Jácomo RH, Nery LFA. Impact of Tube Additives on Baseline Cell-Free DNA, Blood Nuclease Activity, and Cell-Free DNA Degradation in Serum and Plasma Samples: A Comparative Study. LabMed. Multidisciplinary Digital Publishing Institute; 2025;2:4.

21. Gerber T, Taschner-Mandl S, Saloberger-Sindhöringer L, Popitsch N, Heitzer E, Witt V, et al. Assessment of Pre-Analytical Sample Handling Conditions for Comprehensive Liquid Biopsy Analysis. J Mol Diagn. 2020;22:1070–86.

22. Lam NYL, Rainer TH, Chiu RWK, Lo YMD. EDTA is a better anticoagulant than heparin or citrate for delayed blood processing for plasma DNA analysis. Clin Chem. 2004;50:256–7.

23. Hebels DGAJ, Georgiadis P, Keun HC, Athersuch TJ, Vineis P, Vermeulen R, et al. Performance in Omics Analyses of Blood Samples in Long-Term Storage: Opportunities for the Exploitation of Existing Biobanks in Environmental Health Research. Environ Health Perspect. 2013;121:480–7.

24. van Ginkel JH, van den Broek DA, van Kuik J, Linders D, de Weger R, Willems SM, et al. Preanalytical blood sample workup for cell-free DNA analysis using Droplet Digital PCR for future molecular cancer diagnostics. Cancer Medicine. 2017;6:2297–307.

25. Yu J, Ahmann LS, Yao YY, Gu W. Enriched Methylomes of Low-input and Fragmented DNA Using Fragment Ligation EXclusive Methylation Sequencing (FLEXseq). bioRxiv. 2024;2024.11.28.625942.

26. Gu W, Deng X, Lee M, Sucu YD, Arevalo S, Stryke D, et al. Rapid pathogen detection by metagenomic next-generation sequencing of infected body fluids. Nat Med. 2021;27:115–24.

27. Loyfer N, Magenheim J, Peretz A, Cann G, Bredno J, Klochendler A, et al. A DNA methylation atlas of normal human cell types. Nature. 2023;613:355–64.

28. Lenaerts L, Vandenberghe P, Brison N, Che H, Neofytou M, Verheecke M, et al. Genomewide copy number alteration screening of circulating plasma DNA: potential for the detection of incipient tumors. Ann Oncol. 2019;30:85–95.

29. Lo YMD, Han DSC, Jiang P, Chiu RWK. Epigenetics, fragmentomics, and topology of cell-free DNA in liquid biopsies. Science. 2021;372:eaaw3616.

30. 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. Nature Publishing Group; 2017;8:1324.

31. Fan HC, Blumenfeld YJ, Chitkara U, Hudgins L, Quake SR. Analysis of the Size Distributions of Fetal and Maternal Cell-Free DNA by Paired-End Sequencing. Clinical Chemistry. 2010;56:1279–86.

32. Talevich E, Shain AH, Botton T, Bastian BC. CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. PLOS Computational Biology. Public Library of Science; 2016;12:e1004873.

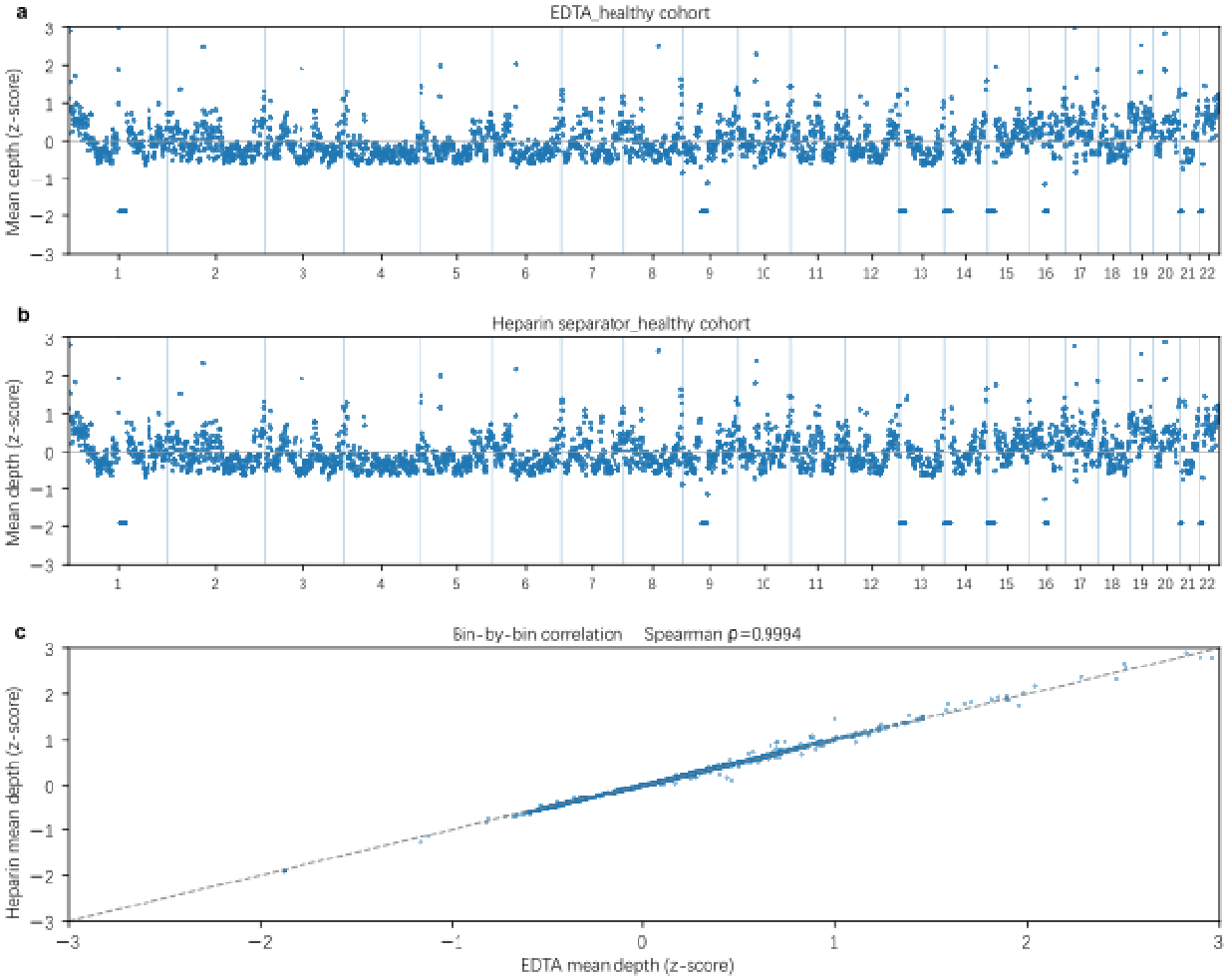
33. Jiang P, Sun K, Peng W, Cheng SH, Ni M, Yeung PC, et al. Plasma DNA End-Motif Profiling as a Fragmentomic Marker in Cancer, Pregnancy, and Transplantation. Cancer Discovery. 2020;10:664–73.

34. Ding SC, Lo YMD. Cell-Free DNA Fragmentomics in Liquid Biopsy. Diagnostics. Multidisciplinary Digital Publishing Institute; 2022;12:978.

35. Hu X, Ding SC, Jiang P. Emerging frontiers of cell-free DNA fragmentomics. Extracell Vesicles Circ Nucl Acids. 2022;3:380–92.

36. Han DSC, Ni M, Chan RWY, Chan VWH, Lui KO, Chiu RWK, et al. The Biology of Cell-free DNA Fragmentation and the Roles of DNASE1, DNASE1L3, and DFFB. The American Journal of Human Genetics. Elsevier; 2020;106:202–14.

SUPPLEMENTARY FIGURES

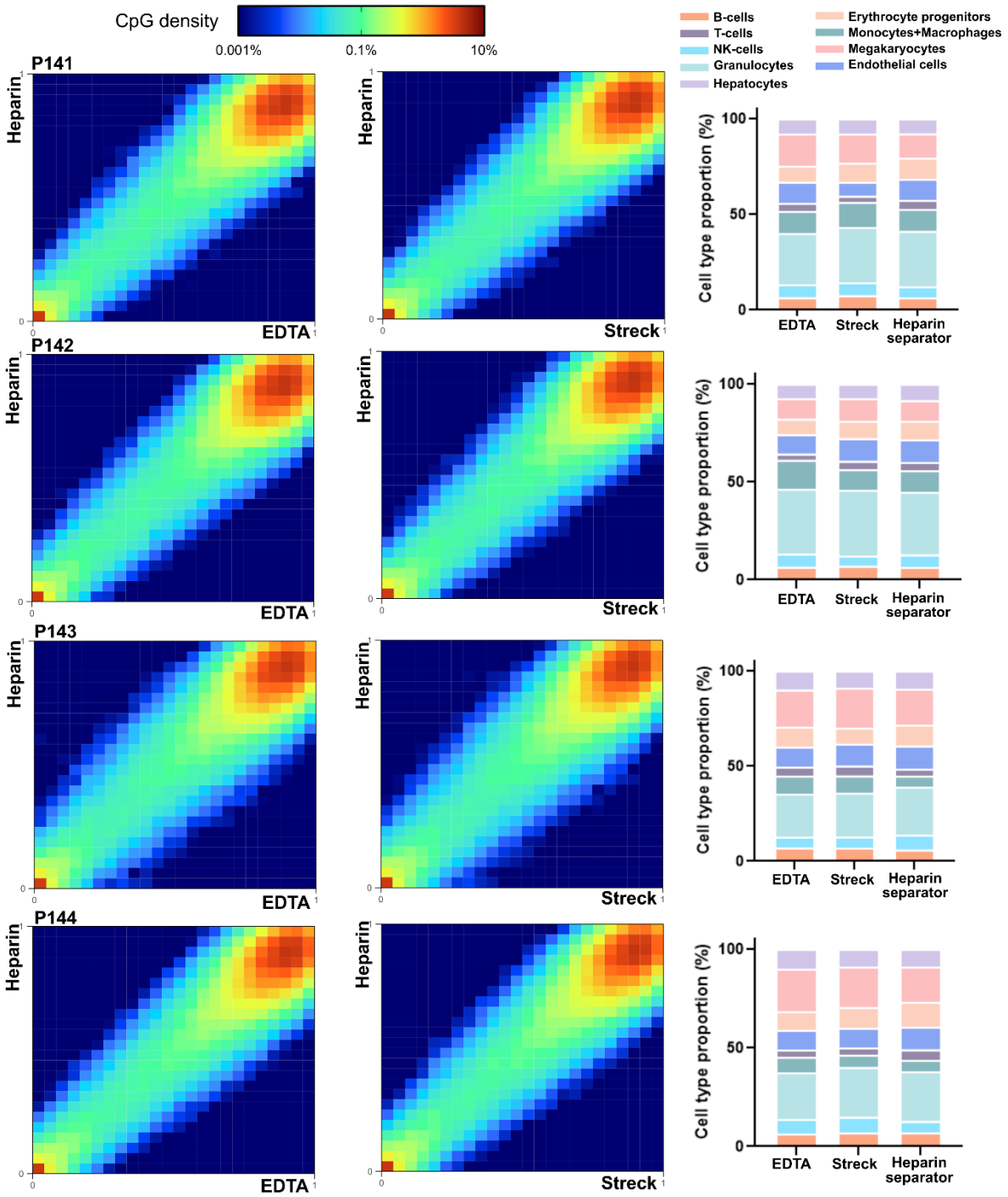


Supplementary Figure 1. Genome-wide coverage comparison between tube types

a. Genome-wide mean sequencing depth across autosomes (Chr1–22) for pooled EDTA plasma samples (n=5). Mean depth was calculated in 1 Mb bins and normalized to z-scores.

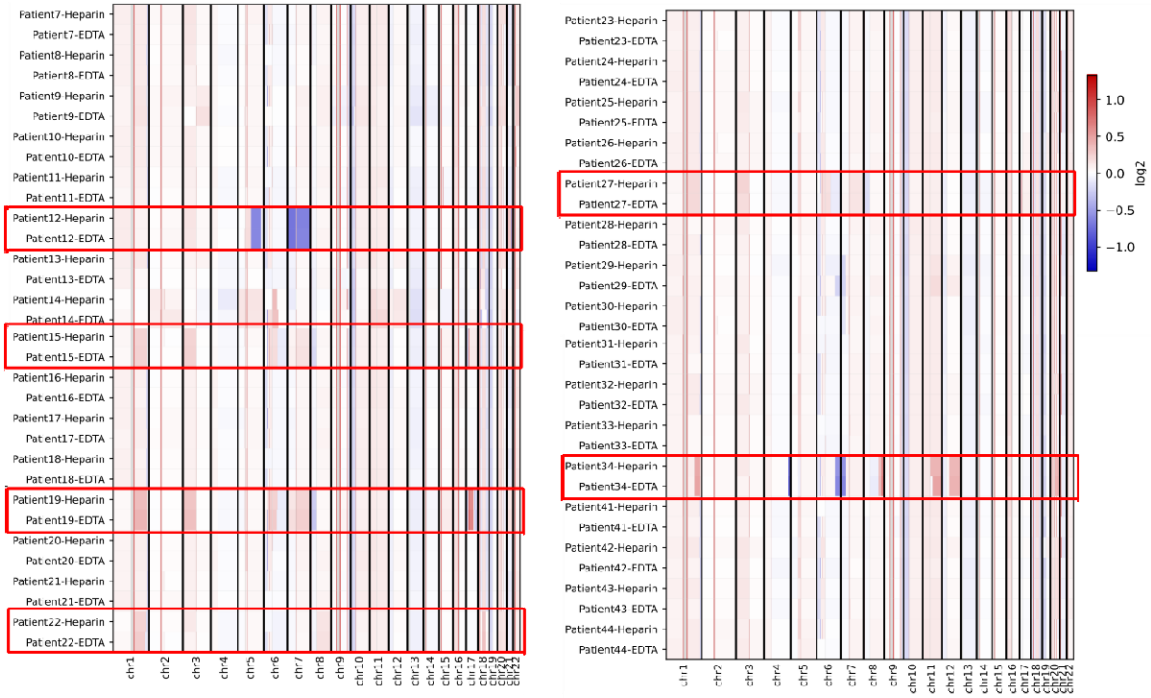
b. Genome-wide mean sequencing depth across autosomes (Chr1–22) for pooled Heparin separator plasma samples (n=5).

c. Bin-by-bin comparison of z-scored mean depth between pooled EDTA and pooled heparin separator plasma across all 1 Mb bins on Chr1–22. Each point represents one genomic bin. The dashed line indicates the identity line (y = x).



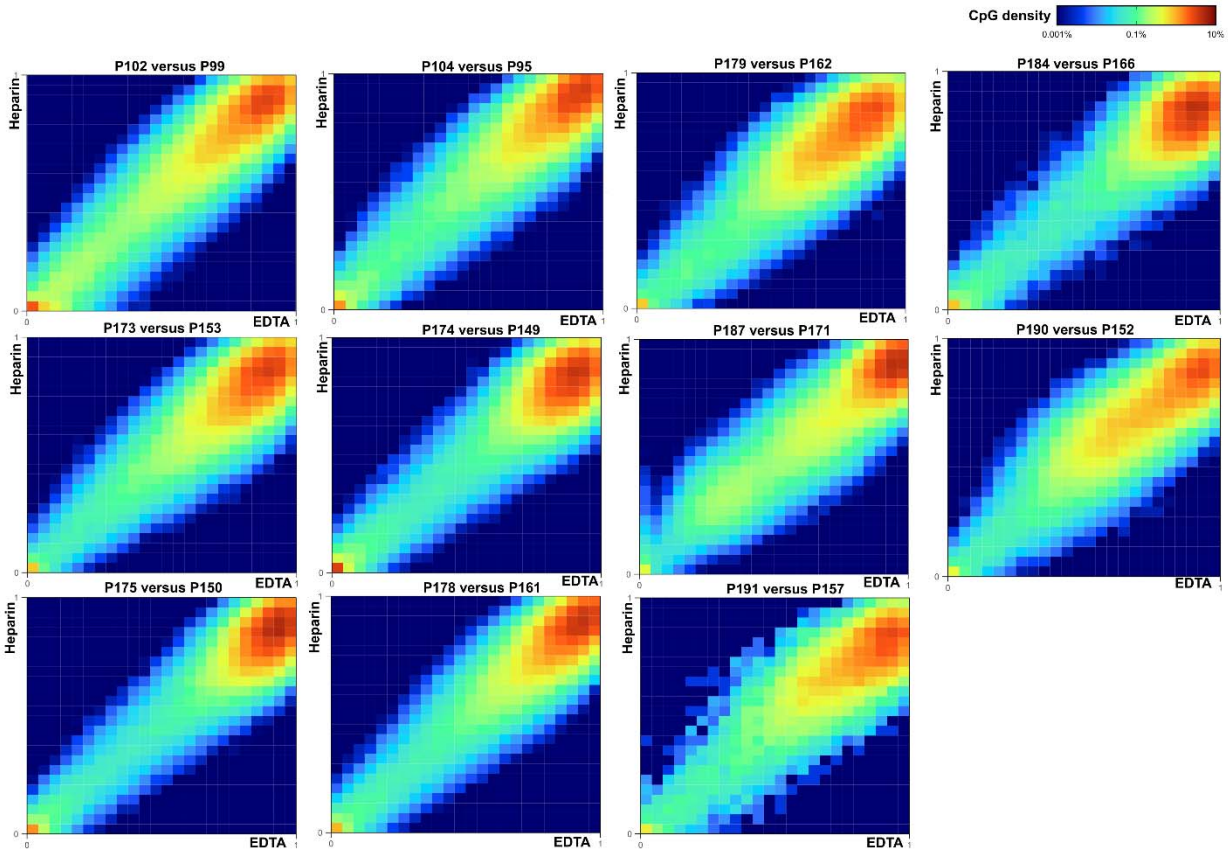
Supplementary Figure 2. Healthy Cohort methylation correlation and cell type deconvolution

Heatmap and correlation of methylation beta values with cell type deconvolution results in paired samples from five healthy donors.



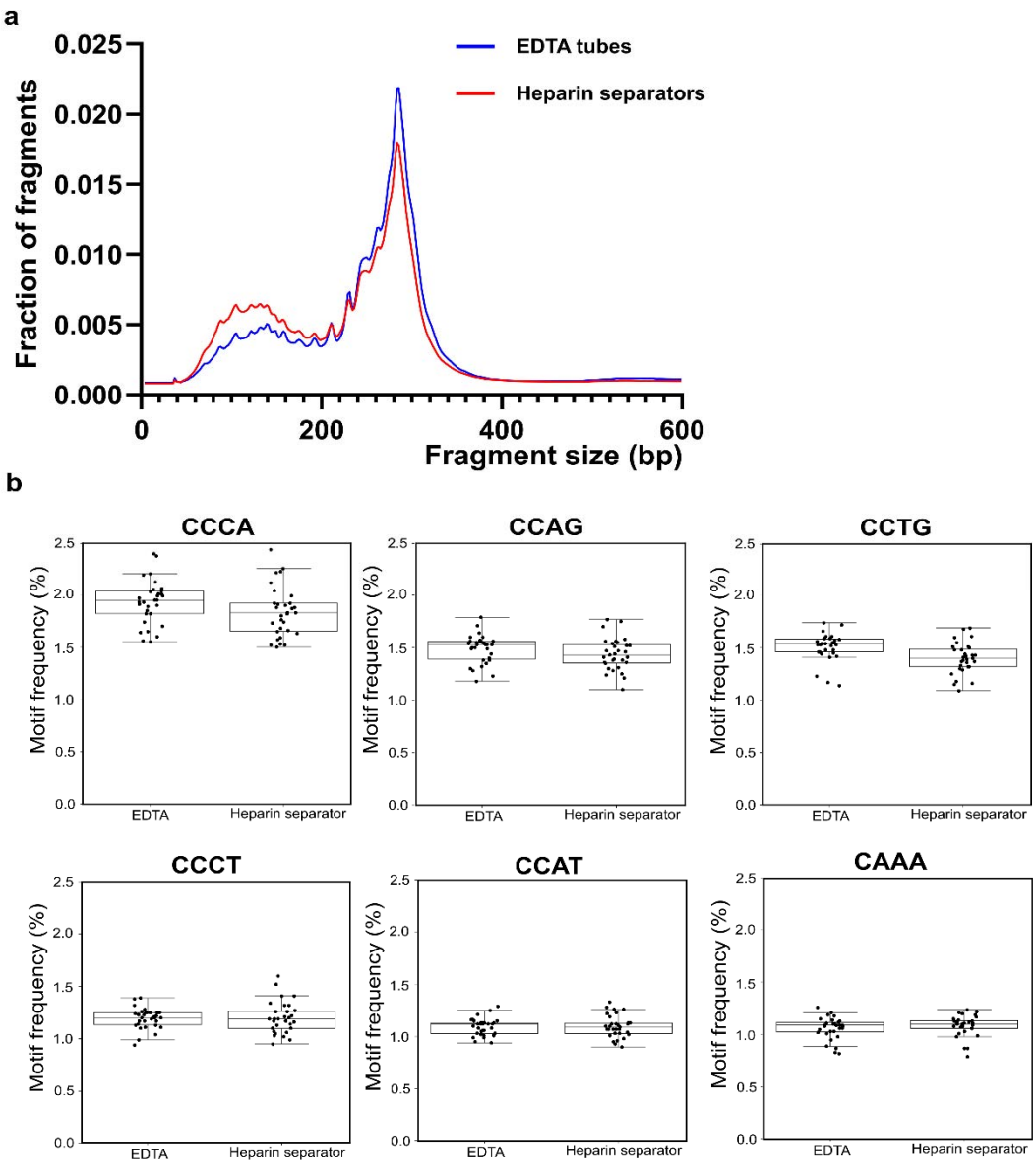
Supplementary Figure 3. CNVkit heatmap across tube types

Genome-wide copy number profiles reanalyzed using CNVkit as a secondary CNV caller. Heatmaps show log2 copy ratio across autosomes for matched EDTA and heparin separator plasma samples. Red boxes highlight CNV positive cases.



Supplementary Figure 4. Methylation correlation across Hospital Cohort cases

Heatmaps of methylation beta values in paired samples collected from 11 virus-positive patients in the Hospital Cohort. One additional case is shown in Figure 5.



Supplementary Figure 5. Fragmentomic patterns of the Hospital Cohort

a. Average cfDNA fragment size distribution collected in EDTA tubes and heparin separators.

b. Average end motif frequency comparison between EDTA tubes and heparin separators, highlighting the top six motifs: CCCA, CCAG, CCTG, CCCT, CCAT, and CAAA.