Communications Medicine

https://doi.org/10.1038/s43856-026-01630-6

Article in Press

Impact of preanalytical factors on liquid biopsy in the canine cancer model

Received: 18 March 2025

Accepted: 21 April 2026

Cite this article as: Megquier, K., Husted, C., Rhoades, J. *et al.* Impact of preanalytical factors on liquid biopsy in the canine cancer model. *Commun Med* (2026). https://doi.org/10.1038/s43856-026-01630-6

Kate Megquier, Christopher Husted, Justin Rhoades, Michelle E. White, Diane P. Genereux, Frances L. Chen, Kan Xiong, Euijin Kwon, Ross Swofford, Corrie Painter, Viktor A. Adalsteinsson, Cheryl A. London, Heather L. Gardner & Elinor K. Karlsson

Impact of preanalytical factors on liquid biopsy in the canine cancer model

Kate Megquier[1], Christopher Husted[1,2,3], Justin Rhoades[1], Michelle E. White[1], Diane P. Genereux[1], Frances L. Chen[1,2], Kan Xiong[1], Euijin Kwon[2,3], Ross Swofford[2], Corrie Painter[1], Viktor A. Adalsteinsson[1], Cheryl A. London[4], Heather L. Gardner[4]*, Elinor K. Karlsson[1,2,5]*

[1] Broad Institute of MIT and Harvard, Cambridge, MA, USA
[2] Genomics and Computational Biology, UMass Chan Medical School, Worcester, MA, USA
[3] Morningside Graduate School of Biomedical Sciences, UMass Chan Medical School, Worcester, MA, USA
[4] Tufts Cummings School of Veterinary Medicine, North Grafton, MA, USA
[5] Program in Molecular Medicine, UMass Chan Medical School, Worcester, MA, USA

*Authors contributed equally

Correspondence:
Elinor K. Karlsson, PhD
UMass Chan Medical School
368 Plantation Street, AS4.1065
Worcester MA 01605
Elinor.Karlsson@umassmed.edu

Kate Megquier, DVM, PhD
Broad Institute of MIT and Harvard
75 Ames Street, 4009B
Cambridge, MA 02142
kmegq@broadinstitute.org

Keywords: liquid biopsy, cell-free DNA, cancer, comparative, canine, optimization

Abbreviations:
cfDNA: cell-free DNA
ng: nanograms
mL: milliliters
µL: microliters
bp: basepairs
Mb: megabases
PCR: polymerase chain reaction
ULPWGS: ultra low-pass whole-genome sequencing
WGS: whole-genome sequencing

ABSTRACT

Background: While liquid biopsy has potential to transform cancer diagnostics through minimally-invasive detection and monitoring of tumors, the impact of preanalytical factors such as the timing and anatomical location of blood draw is not well understood.

Methods: To address this gap, we leveraged pet dogs with spontaneous cancer as a model system for liquid biopsy of plasma cell-free DNA (cfDNA), as their compressed disease timeline facilitates rapid diagnostic benchmarking. We examined key cfDNA metrics, including DNA concentration in plasma, as well as tumor fraction and fragment size ratio derived from ultra low pass whole genome sequencing.

Results: We show that liquid biopsy metrics in dogs are consistent with reported human metrics. The tumor content of samples is slightly higher when blood is obtained from a central vein closer to the tumor. Metrics also differ between lymphoma and non-hematopoietic cancers, supporting cancer-type-specific interpretation. Disease status tracks with liquid biopsy findings over the course of treatment over both short (hours to days) and long (weeks to months) time frames, and trends of increased tumor fraction and other metrics are observed prior to clinical relapse in dogs with lymphoma and osteosarcoma.

Conclusions: Together, these data support the utility of pet dogs with cancer as a relevant system for advancing liquid biopsy platforms.

PLAIN LANGUAGE SUMMARY

Liquid biopsy allows the use of body fluids such as blood to diagnose and monitor patients with cancer. Dogs develop several cancers that are similar to those in humans, but their faster disease progression enables rapid data collection. We used data from pet dogs with cancer to ask whether blood draw techniques and patient characteristics impact the reliability of liquid biopsy. As in humans, findings from liquid biopsy in dogs match results of conventional biopsy. Also, blood drawn from the vein in a dog's neck (jugular vein) contains more DNA from tumors than blood drawn from leg veins. In dogs that had lymphoma or osteosarcoma, liquid biopsy detects changes in cancer burden over time that match clinical disease. Our findings show that the information from liquid biopsy in dogs with cancer mirrors that found in human patients, making them a good model system to help improve approaches across both species.

INTRODUCTION

Liquid biopsy is a powerful, minimally-invasive method for sampling tumor-derived materials including cells and nucleic acids from patient biofluids. The approach can be applied to any biofluid, including blood, saliva, urine, cerebrospinal fluid, and lymph, and characterization can involve multiple techniques including measurement of concentration, assessment of fragment sizes, next-generation sequencing, PCR, and other approaches (reviewed in[1–3]). Liquid biopsy has the potential to enable precision cancer treatment and monitoring by matching patients to targeted therapies[4–7], identifying molecular relapse prior to clinical relapse[8–12], and detecting emerging mutations that confer therapeutic resistance[13]. Here we focus on characterization of cell-free DNA (cfDNA) isolated from blood plasma, which contains DNA from both normal and tumor cells. Tumor content can be assessed by several methodologies, including ultra-low-pass whole genome sequencing (ULP-WGS)[14], an approach which is well-established in human oncology research. We use the term “liquid biopsy” to refer to this approach.

While methodologies have advanced substantially, it remains unclear how a variety of preanalytical variables impact plasma cfDNA yield and diagnostic results. These include patient-specific factors, cancer type, and sample collection protocols, among others[15]. The performance of liquid biopsy is reliant on collecting a sufficient amount of tumor DNA, which may make up only a small fraction of total cfDNA from a given blood sample. Increasing the yield of tumor DNA can therefore substantially impact test sensitivity and specificity[15,16]. Recent data has shown that cfDNA levels vary by collection tube type[17], cancer type[18], by time of day[19,20], and timing of physical exercise[21–24]. However, there are currently no standardized recommendations or guidelines regarding sample collection to mitigate variability and optimize reliability, accuracy, and cost efficiency in the clinical setting[25,26].

Discerning the influence of preanalytical factors on liquid biopsies is challenging in human studies as clinical protocols typically dictate that blood be drawn from the median cubital vein[27]. Moreover, it is often difficult to intensively sample human patients over short timelines and as such, samples may be collected weeks, months, or even years apart[28]. While mouse models have been used to investigate liquid biopsy application, several constraints exist, including very small sample volumes, limited options for blood draw site, and challenges following individual mice over the course of diagnosis, remission and eventual therapeutic resistance[29–31]. Importantly, mouse cancer models also do not capture the environmental and lifestyle diversity of human cohorts[29]. Mouse models also tend to be more genetically homogeneous[29].

Use of pet dogs with cancer as a relevant comparative and preclinical model has recently gained traction as a way to more effectively bridge translation of findings to human patients[32,33]. Cancers in pet dogs develop spontaneously and possess similar histologies, clinical characteristics, and genomic alterations as human cancers[32], and it is estimated that over 1.5 million dogs are diagnosed with cancer each year[34,35]. With their shorter lifespans and disease course, longitudinal clinical studies in dogs can be completed in 1-2 years, instead of the 5-10+ years necessary for similar human studies[32,36,37]. Dogs with cancer can be treated with the same modalities used in human medicine, including chemotherapy, stereotactic/conformal radiation therapy, immunotherapy, and small molecule inhibitors[32]. However, veterinary clinical protocols are more amenable to modification and evaluation of novel approaches in the research setting[19]. A growing number of studies have demonstrated feasibility of liquid biopsy of plasma cell-free DNA in dogs using a variety of analytic techniques, the majority focusing on real-time and digital PCR[38–45], with fewer examining fragment size[46,47], nucleosomes[48,49], targeted sequencing of SNVs[47,50], or whole genome sequencing/ultra low pass whole genome sequencing (ULPWGS) for evaluation of broad copy number changes and tumor fraction[47,51].

Here, we leverage ULPWGS of plasma cfDNA to examine liquid biopsy parameters in dogs to demonstrate the potential utility of the dog cancer model for optimizing liquid biopsy protocols and to identify preanalytical factors that may impact liquid biopsy for further study. Throughout the

study, we focus on three key cell-free DNA metrics: (1) the concentration of cfDNA in plasma; (2) tumor fraction; and (3) fragment size ratio. Broadly, our study is divided into three main sections. (1) comparing cfDNA performance and metrics between dogs and humans, including assessment of the concordance of somatic mutations between tumor and matched cfDNA samples, in which we show that liquid biopsy metrics in dogs are consistent with reported human metrics. (2) Examining the effects of preanalytical variables on cfDNA metrics, finding that the tumor content of samples was slightly higher when blood was obtained from a central vein closer to the tumor and that metrics differed between lymphoma and non-hematopoietic cancers, supporting cancer-type-specific interpretation. (3) Short and longer-term longitudinal monitoring in canine patients with lymphoma or osteosarcoma in which we show that disease status tracks with liquid biopsy findings over the course of treatment over both short (hours to days) and long (weeks to months) time frames, and trends of increased tumor fraction and other metrics are observed prior to clinical relapse in dogs with lymphoma and osteosarcoma (Figure 1).

METHODS

Ethical approval

Ethical approval for sample collection was obtained from the Institutional Animal Care and Use Committees at Cummings School of Veterinary Medicine at Tufts University (TCSVM), The Ohio State University School of Veterinary Medicine, and the Broad Institute of MIT and Harvard. All samples were collected with owner consent.

Blood sampling, plasma fractionation, and storage

All blood samples were collected by veterinary medical professionals using standard protocols. We were not able to control for all possible variables in the phlebotomy procedure itself. However, the majority (N=378, 88%) of samples were collected by the Clinical Research Shared Resource at the Tufts University Foster Hospital for Small Animals, and as such many of the same technicians performed the blood draws and standard techniques (including using only 20 or 22 gauge needles) were used. Further study will be needed to assess the effect of needle gauge, restraint, and other variables in phlebotomy technique.

The Ohio State University Biobank. Blood samples were collected in EDTA tubes, and plasma was fractionated within 30 minutes and stored in aliquots of approximately 0.5 mL at -80°C. *TCSVM Comparative Pathology and Genomics Resource.* Blood was collected into a Streck cell-free DNA blood collection tube. Plasma was fractionated following the manufacturer's protocol, and stored in approximately 1 mL aliquots at -80°C. *Broad Institute Genomics Platform*. Blood collected in Streck or EDTA tubes was centrifuged at 1900 g for 10 minutes and plasma was transferred to a second tube before further centrifugation at 15000 g for 10 minutes. Plasma was stored at -80°C until cfDNA extraction. Collection site is annotated in Supplementary Data 1.

DNA extraction and QC

Cell-free DNA. For all samples, plasma was separated from blood cells (plasma volume ranging from approximately 0.5 mL to 5 mL), and cfDNA isolated from the plasma. *Blood Biopsy group.* Frozen plasma from TCSVM was thawed and centrifuged at 15000 g for 10 minutes. If a sample did not meet the preferred starting input (2.1mL), 1x PBS was added to bring the volume up. cfDNA was extracted using the QIAsymphony DSP Circulating DNA Kit (QIAGEN) according to the manufacturer's instructions. The extracted cfDNA was quantified using the PicoGreen (Life Technologies) assay and then frozen at -20°C until ready for further processing. *Broad Genomics Platform.* If a sample did not meet the preferred starting input (6.3mL), 1x PBS was added to bring the volume up. cfDNA was extracted using the QIAsymphony DSP Circulating DNA Kit according to the manufacturer's instructions and quantified with the PicoGreen assay. *TCSVM Comparative Pathology and Genomics Resource.* Frozen plasma (up to 4mL) was extracted using the Apostle MiniMax cfDNA extraction kit (Apostle, LLC) on a Biomek i7 Hybrid liquid handling system

(Beckman Coulter) per manufacturer instructions. cfDNA was quantified using PicoGreen or Qubit high sensitivity 1x dsDNA assay (Invitrogen), per manufacturer instructions. cfDNA was stored at -80C until library construction. The amount of input plasma was quantified in order to calculate cfDNA concentration per mL plasma. In a subset of samples (N=80/317), plasma amounts were estimated based on the number of approximately 0.5 or 1mL aliquots.

Whole blood. (Broad Blood Biopsy group only) gDNA was extracted from approximately 200 µL whole blood (N=5 samples, 2 replicates each) using the QIAsymphony DSP DNA mini kit. The extracted gDNA was sheared to 150 bp using a Covaris ultrasonicator (gDNA only, cfDNA was not sheared), quantified using the Qubit dsDNA HS kit, then frozen at -20^{o}C until ready for further processing.

ULPWGS library construction

Library preparation was performed using a commercially available kit (KAPA HyperPrep Kit with Library Amplification) and IDT's duplex UMI adapters. Unique 8-base dual index sequences embedded within the p5 and p7 primers (IDT) were added during PCR. Enzymatic clean-ups were performed using AMPure XP beads (Beckman Coulter). Input DNA was normalized to be within the range of 25-52.5 ng in 50 uL of TE buffer (10mM Tris HCl 1mM EDTA, pH 8.0) according to PicoGreen quantification (*Broad Genomics platform)*, or 1-10ng molarity (*TCSVM Comparative Pathology and Genomics Resource)*. Library quantification was performed using the Invitrogen Quant-It broad range dsDNA quantification assay kit (Thermo Scientific) with a 1:200 PicoGreen dilution (*Broad Genomics platform)*, PicoGreen (Life Technologies) (*Broad Blood Biopsy group)*, or TapeStation (*TCSVM Comparative Pathology and Genomics Resource).* Libraries were sequenced on Illumina HiSeq 2500 rapid run (100 bp paired-end reads, *Broad Blood Biopsy group and Broad Genomics Platform*), HiSeq X (151 bp paired-end reads, *Broad Genomics Platform*), NovaSeq 6000 (151 bp paired-end reads, *Broad Genomics Platform*), or NextSeq 2000 (151 bp paired-end reads, *TCSVM Comparative Pathology and Genomics Resource*). ULPWGS sequencing metrics are presented in Supplementary Data 3n.

Alignment and preprocessing of ULPWGS data

Broad Genomics Platform and *Blood biopsy group*. ULPWGS sequencing data was preprocessed following the GATK best practices[52]. Sequencing data was aligned to the CanFam3.1 reference genome[53] using bwa[54], and duplicate reads were flagged using Picard Tools MarkDuplicates. *TCSVM Comparative Pathology and Genomics Resource.* ULPWGS sequencing data were demultiplexed using BCL convert (v2.4.0) and aligned to the CanFam4 reference genome[55] using BWA Aligner (v1.1.4) on the Illumina BaseSpace sequencing hub.

Fragment sizes

Fragment size counts were calculated from bam files using the Deeptools2[56] bamPEFragmentSize command. Fragment size ratio was calculated as the ratio of the count of fragments between 100-150 base pairs (bp) (short) to the count of fragments between 151-220 bp (long) for each sample. Fragment size ratios are provided in Supplementary Data 1.

Evaluation of tumor fraction

The tumor fraction of cfDNA ULPWGS data was assessed using the ichorCNA tool[12], modified to run on data aligned to the canine genome, and to allow the manual definition of sample sex due to the lack of a Y chromosome in the CanFam3.1 genome build. Briefly, ichorCNA uses read depth information from ULPWGS data and either a paired normal sample or a panel of normals to identify copy number changes in large (1Mb) bins across the genome (corrected for GC content and mappability differences between bins). Drawing on the assumption that variation in read depth across the genome (after subtraction of inherited copy number changes from a paired normal or panel of normals) reflects somatic copy number changes in tumor DNA, the tool uses

a hidden Markov model to simultaneously infer the proportion of tumor DNA present in the sample and whether each copy number change is clonal or subclonal[12].

HMMcopy[57] was used to produce read depth counts in 1 Mb intervals across the genome. Reference files for GC and mappability correction were generated using the HMMcopy utils package[58]. Putative centromere locations were assigned as regions with 80% or more repetitive sequence, based on a genome scan for four centromeric repeats in 5 kb windows across the CanFam3.1[59] or CanFam4[60] genomes. The pseudoautosomal region was filtered out to prevent spurious copy number gains after chrX normalization in male dogs. As the first tumor fraction solution provided by ichorCNA is not always the best based on heuristics developed by the Broad Blood Biopsy team, we developed a set of rules to select the best solution in an unbiased way, prioritizing lower fraction SCNAs subclonal, lower copy number levels, and concordance of SCNAs among samples from the same individual.

Panel of normals creation

Small panel. We sequenced cfDNA from blood plasma and genomic DNA (gDNA) from PBMCs from five healthy dogs. We generated a panel of normals (PoN) using both cfDNA and gDNA samples and found that the gDNA PoN showed less noise (lower mean absolute deviation) than the cfDNA PoN. For this reason, we used the gDNA PoN for subsequent analysis with ichorCNA. We realigned these bam files to CanFam4, and this small panel was used as the panel of normals for files aligned to CanFam4. *Large panel*. We also created a panel of normals using previously-published normal whole-genome sequencing data from 92 dogs[61,62] (Supplementary Data 4), along with the five PBMC samples included in the small panel. This was used as the panel of normals for samples aligned to CanFam3.1.

Canine whole exome capture design

The capture panel for whole exome sequencing (WES) was designed by the Vertebrate Genomics group at the Broad Institute, together with the Broad Genomics Platform, who provided input on the regions targeted in the Broad Human Somatic Exome Capture. The Bioinformatics team at Twist Bioscience then designed the baits to capture the targeted regions. This panel targets 20,257 genes derived from the Ensembl canine gene annotation Release 99, human cancer-specific targets (translated from hg19 to CanFam3.1 using liftOver[63]), and selected gene promoter sites and UTRs. The canine exome design targets approximately 36.6 Mb with 40.5 Mb of probe territory[64] and is available as a product from Twist Bioscience.

Whole exome hybrid capture and sequencing.

Pre-existing cfDNA ULPWGS libraries were quantified and pooled prior to selection. The library pool along with the Canine Exome Capture v2 probe set were used as input into the xGEn Hyb and Wash kit (IDT). Post capture, libraries were sequenced on the Illumina HiSeq X platform using a 2x151 bp run, to a target depth of 150x coverage, achieving a mean target coverage of 131x. WES sequencing metrics are presented in Supplementary Data 5n.

Alignment and preprocessing of whole-exome sequencing data

Broad Genomics Platform. WES data was preprocessed following the GATK best practices. Sequencing data was aligned to the CanFam3.1 reference genome[53] using bwa[54] version 0.7.15-r1140, and duplicate reads were flagged using Picard Tools MarkDuplicates. A VCF of known canid germline variation (34,191,821 SNPs and 11,943,064 INDELs from 674 dogs and other canids[61]) was used as the germline reference for BQSR. Preprocessing was performed using the "processing-for-variant-discovery" pipeline, with GATK version 4.1.8.1 and genomes-in-the-cloud version 2.3.1-1512499786.

Simple somatic mutation calling

Simple somatic variant calling was performed on whole genome sequencing (WGS) data from 24 osteosarcoma and paired normal samples, as well as WES data from paired cfDNA samples from six of the same dogs. WGS data was subset to the targeted exome regions using the Samtools[65,66] view command. Data was preprocessed in the Terra Cloud Platform using the "Processing-for-variant-discovery" showcase workflow. To ensure the robustness of the somatic mutation calls, we used three different callers: Mutect2 v4.1.8.1[67], Strelka2 v2.9.10[68], and Octopus v0.7.4[69]. Each tumor or cfDNA sample was run with its matched normal sample, and mutations called by two or more of the tools were retained for downstream analysis. A VCF of germline variants in 674 dogs was used as a population variant resource. Mutect2 was run with the additional filter "minimum median read length" of 10 and the additional arguments: "20 downsample stride", "6 max reads per alignment start", and "6 max suspicious reads per alignment start". Tumor-cfDNA pair WLOS_0015 was excluded from subsequent concordance analysis because we detected only a single mutation in the tumor. This sample had a higher proportion of mutational calls flagged as potential cross-sample contamination by Mutect2 (70%; the other five tumor samples had a range of 26% - 40%), which may have contributed to the low number of variants passing filtration.

Consensus of somatic mutations analysis

Raw mutation calls from each tool were filtered to include only mutations which passed all filters using the Bcftools[66] view function. Consensus mutation calls from the three tools were created using Bcftools isec. SnpEff[70] was used to annotate the calls and SnpSift[71] was used to extract annotations into tab-delimited format. The intersection of the consensus mutation calls for each tumor - cfDNA pair was obtained using the Bcftools isec command.

Disease stage

Dogs in the longitudinal studies of lymphoma and osteosarcoma had consistent staging including thoracic radiographs and complete blood counts in addition to physical examination as part of clinical trials at the Cummings School of Veterinary Medicine at Tufts University. While a numbered staging system is not used clinically for canine osteosarcoma, dogs with osteosarcoma were all enrolled with the primary tumor present and no evidence of metastatic disease on physical examination or thoracic radiographs. Dogs with lymphoma were assigned stage based on lymph node measurements on physical examination, thoracic radiographs ± abdominal ultrasound, and complete blood count.

Tumor size

Information on tumor size was available from a limited number of dogs. In dogs with lymphoma, tumor size was calculated as the sum of the longest dimensions in centimeters of the five pairs of lymph nodes commonly measured on physical examination (mandibular, prescapular, axillary, inguinal, and popliteal). In dogs with other cancer types, tumor size was the sum of the longest dimensions of masses measured on physical examination or thoracic radiographs. These tumor measurements are not expected to capture the total disease burden, but are expected to be correlated with total disease burden.

Sample exclusion

No *a priori* rules for sample exclusion were established. Seven samples were excluded post hoc for technical reasons, including: inadequate sequencing data (N=2), abnormal ichorCNA plots (N=2 replicates), ambiguous labeling (N=1), suspected library preparation error (N=1), and high contamination levels flagged by the Mutect2 pipeline (N=1). In addition, one individual was excluded from the blood draw site analysis due to having no evidence of disease at the time of blood draw (N=2 samples).

Statistics and reproducibility

Statistical tests were performed using the R programming language (versions 4.2.1 and 4.5.2)[72]. For all tests, correction for multiple hypothesis testing was performed using the Benjamini-

Hochberg method[73], and significance was confirmed using permutation-based and robust (outliers trimmed) methods to ensure that significance was not an artifact of non-normality or driven by outliers. *Correlation tests*. The Spearman correlation test was used for correlations between two variables and we report bootstrapped confidence intervals. *Paired tests*. To compare differences in means between independent or paired groups of samples, we used the non-parametric Wilcoxon signed-rank test for cfDNA metrics, which tended not to be normally distributed. For demographic variables, we used paired Welch's t-tests[74] or Fisher's Exact Test for categorical data. *Multivariable ANOVA and Linear Mixed Models*. To assess the influence of multiple variables on cfDNA metrics, we performed multivariable ANOVA analyses on the first sample from each dog. For analyses including multiple samples from the same dog, we used linear mixed models incorporating the dog ID as a random effect. The effect size for dog ID in these models was calculated as the Intraclass Correlation Coefficient[75]. The effect of cfDNA metrics on progression free survival were calculated as univariable and multivariable Cox proportional hazards models[76]. Results of all statistical tests are provided in the supplementary tables. Large language models ChatGPT 4o[77] and Gemini 3 Pro[78] were used to edit statistical and figure code. All code was reviewed by the authors.

Code availability

The code needed to run ichorCNA on the canine genome, as well as custom scripts used for figure generation and statistical analyses are available on GitHub at [79].

RESULTS

Cohort overview

We evaluated cfDNA samples from a diverse set of 145 dogs—129 with cancer and 16 presumed-healthy controls (Supplementary Table 1). Overall, the dogs in our cohort tended to be larger (29.8 (mean) ±10.9 kg (standard deviation)), older (8±3 years), and more likely to be purebred (65.5%) and spayed or neutered (93%) than has been reported in the overall U.S. dog population[80,81] (Supplementary Data 1). Dogs with cancer were not statistically different from the healthy dogs in our cohort in terms of sex (p_{fisher}=0.6), spay/neuter status (p_{fisher}=0.06; only 6.25% of dogs were intact), whether they had ancestry from one or multiple breeds (p_{fisher}=0.18), weight ($p_{Wilcoxon}$=0.556), or body condition score ($p_{Wilcoxon}$=0.704) (Supplementary Figures S1A-E, Supplementary Data 2), although given the small number of controls we lack power to detect small to moderate differences between the two groups. Dogs with cancer were older than healthy dogs (8.5±2.7) years old for dogs with cancer and 4.4±3.2 years old for healthy dogs, $p_{Wilcoxon}=3.73x10^{-5}$) (Supplementary Figure S1F).

Sample collection

Plasma cell-free DNA was isolated from venous blood collected in Streck or EDTA blood collection tubes. The number of samples collected from each dog varied from 1 to 11 depending on the study cohort (Figure 2A), totaling 428 individual cfDNA samples. We performed ultra-low-pass whole-genome sequencing (ULPWGS) of cfDNA at a target sequencing depth of 0.1x, achieving an overall coverage of 0.32±0.34 (range: 0.0085 - 1.56x) (Supplementary Data 3).

Cell-free DNA metrics

We calculated three metrics from cfDNA samples and the resulting ultra-low-pass sequencing data: plasma cfDNA concentration (nanograms cfDNA per mL of plasma), tumor fraction, and genome-wide fragment size ratio. Tumor fraction is the fraction of the cfDNA derived from the tumor, as estimated by the ichorCNA tool[12]. Fragment size ratio is the ratio of small (100 - 150 bp) to large (151 - 220 bp) DNA fragments genome-wide. A higher fragment size ratio indicates a greater proportion of short fragments, potentially reflecting a higher tumor fraction or increased apoptosis. Analyses reported below include only the first sample from each dog in our cohort

unless otherwise noted, *e.g.,* multiple samples were part of the experimental design (blood draw site, time of day, longitudinal analyses) (Supplementary Data 1).

Liquid biopsy metrics in dogs match human studies

The canine cfDNA was nucleosomal in modal length, corresponding to approximately the length of DNA wound around a single nucleosome (147 bp) plus linker DNA (166-167 bp)[82,83]. The size distribution also displays a periodicity of approximately 10 bp below nucleosomal sizes, consistent with DNase cleavage at the twists of the DNA double helix and closely matching reported size distributions in human[83] and canine[84] cfDNA (Supplementary Figure 2). The modal fragment lengths for healthy dogs and dogs with lymphoma were both 164 bp, while dogs with osteosarcoma had slightly smaller modal fragment size (144 bp). Consistent with reports from human liquid biopsies[85], lymphoma samples appeared to have a relative enrichment of long dinucleosomal DNA (~353 bp).

We compared canine cell-free DNA metrics to those from published human studies (Supplementary Data 6)[12,86]. While we were not able to match the cancer types represented or control for disease stage, treatment protocols, and other factors, we nevertheless found no evidence of significant differences in plasma cfDNA concentration, tumor fraction, and fragment size ratio between dogs and humans. Examining cancer type- species combinations with at least ten samples each (including first samples per individual only), species had no significant effect on tumor fraction (ANOVA ges=0; p=0.73), and a small effect on cfDNA concentrations (ANOVA ges=0.02; $p=6.6x10^{-5}$) and fragment size ratio (ges=0.12; $p=9.53x10^{-9}$) (Supplementary Figure 3A-D, Supplementary Data 7). In contrast, cancer type had a significant effect on all three metrics: cfDNA concentration (ges=0.23; $p=1.63x10^{-36}$), tumor fraction (ges=0.2; $p=2.91x10^{-26}$) and fragment size ratio (ges=0.2; $p=1.72x10^{-9}$).

While the average tumor fraction was lower in humans than in dogs (0.14±0.183 in humans vs 0.299±0.289 in dogs; $p_{t\text{-}test}=2.8x10^{-4}$), this significant difference was eliminated when we removed the dog lymphoma samples (average tumor fraction in dog non-lymphoma cancers = 0.178±0.191 in dogs; p=0.085)(Supplementary Figure 3E). While the effect of species on fragment size remains when limiting the analysis to canine EDTA samples only, some of this difference may be attributable to differences in sample handling and storage (including time to plasma separation), DNA extraction protocols, and library construction techniques[86]. Further study will be needed to determine whether there are differences between the two species when controlling for all clinical and preanalytical variables.

Technical variation in the amount, concentration and fragment sizes of the tumor DNA detected in the cfDNA samples was extremely low, with all three measures significantly and strongly correlated in 47 pairs of technical replicates (Supplementary Figure S4F-H, Supplementary Table 2). Consistent with benchmarking in human samples[87], tumor fraction estimates were also strongly correlated (Spearman's rho=0.69, $p=2.2x10^{-5}$) between libraries made using higher and lower DNA inputs from 32 cfDNA samples from dogs with osteosarcoma. Fragment size ratios were less strongly correlated across input replicates (Spearman's rho=0.56, $p=1x10^{-3}$), suggesting that this metric may be more affected by input DNA amount (Supplementary Figure 5, Supplementary Table 3). As expected, the tumor fraction was positively correlated with the plasma cfDNA concentration in liquid biopsies from dogs with cancer (Spearman's rho=0.556; $p=1.07x10^{-4}$; N=108) but not in healthy dogs (Spearman's rho=0.335; p=0.204; N=16)(Supplementary Figure 6, Supplementary Table 4). cfDNA concentration was negatively correlated with fragment size ratio, especially in dogs without cancer (Spearman's rho=−0.65, p=0.006; N=16). Tumor fraction was negatively correlated with fragment size ratio in lymphomas

(Spearman's rho -0.38; p=0.012; N=43) and positively correlated with fragment size ratio in other cancers (Spearman's rho 0.681; $p=2.53x10^{-12}$; N=86).

Tumor-cfDNA concordance

cfDNA samples yielded more somatic mutations than matched tumor tissue

Liquid biopsies in dogs capture somatic changes observed in tumor tissue collected from the same dog at the same time, consistent with findings in humans[88]. For 24 dogs with osteosarcoma[62], we identified large somatic copy number changes in the tumor (60x WGS) and the cfDNA (ULPWGS) compared to normal muscle tissue (30x WGS) using ichorCNA (Supplementary Data 8, 9). Copy number changes were significantly correlated between tumor and matched cfDNA samples (Spearman's rho = 0.63; $p<2.2x10^{-16}$, N=24). The correlation between sample pairs with cfDNA tumor fractions ≥ 0.1 was even stronger (rho=0.72; $p<2.2x10^{-16}$; N=14) and nearly identical to that observed in humans (0.763 in 22 sample pairs of metastatic breast cancer[12]) (Figure 2B, Supplementary Figure 7, Supplementary Table 5).

To investigate whether simple somatic mutations (SNVs and indels) are captured by liquid biopsies, we developed a custom whole exome sequencing (WES) panel[64]. Our design targets 37.6 Mb and includes 40.5 Mb of probes covering 20,257 genes as well as non-coding regulatory regions implicated in human cancers. The probes are available as a product through Twist Bioscience as a resource for other comparative canine genomics studies. We performed WES on six of the 24 osteosarcoma samples. We called somatic mutations using Mutect2[67], Strelka2[68], and Octopus[69] and kept mutations identified by two or more callers for further analysis (Supplementary Data 10). One sample was excluded from further analysis because there was evidence of cross-sample contamination.

We assessed the concordance of clonal and subclonal simple somatic mutations between tumor and paired cfDNA samples (Figure 2C, Supplementary Table 6). Clonal mutations were defined as having an allele fraction greater than or equal to 0.9 times the tumor fraction as estimated by ichorCNA. Concordance between tumor and cfDNA was high for four sample pairs, with an average of 88% of clonal mutations found in the tumor also seen in the cfDNA (86%, 92%, 100%, and 100%) and an average of 86% of clonal cfDNA mutations also observed in the tumor (66.7%, 88.9%, 100%, and 100%). Concordance of subclonal mutations was lower, averaging 79% for tumor mutations found in cfDNA, and 39% of cfDNA mutations found in the tumor. Recurrent mutations in common human (and putative canine) driver genes were found in the cfDNA samples both as clonal and subclonal events. For example, SNVs in *TP53* were identified in the cfDNA of five dogs, and were also identified in the matched tumor sample in four of those dogs. Three of these mutations were clonal and two subclonal. *SETD2* mutations were identified in the cfDNA of five dogs, with four identified in both the tumor and cfDNA. Two of these mutations were clonal in both the tumor and cfDNA, while three mutations were subclonal.

One sample (WLOS_0017) had much lower concordance, with just 33% of clonal and 8.3% subclonal tumor mutations seen in the cfDNA, and 18.2% of clonal and 1.4% of subclonal cfDNA mutations seen in the tumor. The cause of the low concordance in this sample is unknown. Tumor fractions (tumor: 0.6, cfDNA: 0.28), read depth, and percentage duplicate reads were within the range observed in the other samples, and we confirmed that the tumor and cfDNA were from the same individual using Bam-Matcher[89] (Supplementary Table 6, Supplementary Figure 8). It is possible that this dog had a higher burden of metastatic disease that differed genetically from the primary, or that there was a second primary tumor, which is not uncommon in dogs[90].

In all five cfDNA/tumor pairs, we detected a larger number of somatic mutations in cfDNA WES than in tumor (74 vs. 34 average variants, paired one-tailed t-test $p=1.92x10^{-4}$) (Figure 2C). This finding is consistent with previous work in humans that cfDNA captures a more complete

representation of tumor mutations than single-tumor biopsies[91]. However, more work is needed to confirm whether the observed differences in mutational burden are affected by differences in sequencing approach or other technical artifacts. Consistent with this, the cfDNA also tended to have lower allelic fractions (paired t-test p=0.013) and higher MATH (Mutant-Allele Tumor Heterogeneity)[92] scores (paired t-test p=0.05) compared to the tumor (Supplementary Figure 9F,G, Supplementary Tables 7, 8). Some genes had recurrent subclonal mutations seen only in the cfDNA, including *MCM6* (five mutations in four dogs) and *MLLT3* (four mutations in the 5' untranslated region in three dogs). More work is needed to discern whether these mutations are biologically more common in metastatic tumors, or whether they are present in the primary tumors and missed by sampling or other technical artifacts.

Assessment of preanalytical variables

Liquid biopsy metrics were affected by disease status and cancer type

Dogs with cancer had significantly elevated values on all three cfDNA metrics compared to healthy controls (Figure 3 A-C), consistent with human studies[12,86]. About half of the variation in liquid biopsy metrics was explained by differences among individual dogs (Supplementary Data 11), which could reflect variation in clinical features, other environmental factors, or inherited genetics. None of the patient characteristics, including breed, sex (with spay/neuter status), weight, body condition score, or age, explained the effect, consistent with results from human studies that found a weak or mixed effect of age, sex, and weight[93].

Controlling for inter-dog variability as a random effect in a linear mixed model across all samples, disease status (presence or absence of disease plus response criteria) was strongly associated with cfDNA concentration ($p=2x10^{-21}$), tumor fraction ($p=4x10^{-25}$), and fragment size ratio (p=0.005) (Figure 3D, Supplementary Data 11). To evaluate factors that did not vary within samples from the same dog and limit confounding caused by dog specific factors, we assessed the first sample from all dogs via ANOVA, finding that cancer type was significantly associated with cfDNA concentration (p=0.045) and tumor fraction (p=0.003) depending on the model (Figure 3D, Supplementary Data 11). Consistent with prior studies[42,45,94], dogs with lymphoma had higher cfDNA concentrations and higher tumor fractions than either healthy controls or dogs with other types of cancer (Figure 3B-D). This could be due to the hematopoietic nature of lymphoma, or differences in stage between the cancer types. The majority of dogs with lymphoma were enrolled as part of a clinical trial with Stage III disease or greater (meaning that more systemic disease was present). In contrast, the majority of dogs with osteosarcoma (the non-lymphoma cancer type with the largest number of samples) had a primary tumor without detectable metastasis at enrollment (localized disease). Sample collection from dogs with other cancer types was not standardized to a specific treatment time point.

The effect size of dog breed was large, but did not reach significance in the statistical model overall (Figure 3D). Because few breeds were represented by more than one dog in our data set (Supplementary Figure 10A), we could not distinguish a potential effect of these breeds from other characteristics that differ between individuals. Comparing the two most highly represented breeds (Labrador retriever and golden retriever) and all other dogs revealed a difference in fragment size ratios between Labrador retrievers and other breeds (p=) on univariable analysis, but this was not significant on multivariable ANOVA (Supplementary Figure 10, Supplementary Data 12). Sequencing group had a small, nominally significant effect on fragment size ratio in one model (p=0.032), but this was not significant after multiple testing correction (Figure 3D, Supplementary Data 11).

Tumor size was associated with cfDNA metrics

In a subset of dogs with lymphoma (n=33 dogs, 131 samples), we were able to control for tumor size as estimated by peripheral lymph node measurements. Controlling for inter-dog variability across all longitudinal samples, disease status and measured tumor size were associated with

DNA concentration (p_{size}=5x10^{-6}, p_{status}=0.009), tumor fraction (p_{size}=2x10^{-6}, p_{status}=0.001), and fragment size ratio (p_{size}=0.019, p_{status}=0.004)(Supplementary Figure 11, Supplementary Data 13). As disease stage (the clinical score noting extent of disease spread in the body) only varied between dogs, we assessed its effect across the first sample from all dogs in this cohort. Tumor size was significantly associated with cfDNA concentration (p=0.039), however, none of the covariates were significant after multiple testing correction. On univariable analysis, DNA concentration (p=0.013) and fragment size ratio (p=0.015) varied with disease stage, although larger cohorts will be needed to further investigate this trend (Supplementary Figure 12, Supplementary Table 9).

We were also able to estimate tumor size from dogs with other cancer types at the first sample time point based on physical examination or radiology measurements (n=30 dogs). These included dogs with osteosarcoma (n=13 dogs), soft tissue sarcoma (n=3) anal gland adenocarcinoma (n=2), hemangiosarcoma (n=2), and others. On univariable analysis across one sample per dog, tumor size was significantly associated with DNA concentration (R=0.46, p=0.004) and tumor fraction (R=0.49, p=0.002) in lymphomas and tumor fraction (R=0.37, p=0.047) in other cancer types (Supplementary Figure 13, Supplementary Table 10). After adjusting for the effects of sequencing group, tumor fraction was no longer significantly associated with tumor size in other cancer types, however fragment size ratio was significantly associated in opposite directions in lymphomas (R=-0.33, p=0.045) and other cancer types (R=0.49, p=0.006) (Supplementary Figure 13, Supplementary Table 10).

The effect of white blood cell count on cfDNA metrics differed by cancer type

White blood cells are a primary source of cfDNA, even in healthy individuals, so it is possible that differences in metrics for lymphomas as compared to other cancer types are driven by the elevated (neoplastic) white blood cell counts characteristic of this cancer type[95,96]. Complete blood count (CBC) data was available from 66 dogs with cancer (34 with lymphoma, 22 with osteosarcoma, 10 other cancers), in some cases from multiple time points, for a total of 267 samples (Supplementary Data 14). We found that white blood cell count was positively correlated with DNA concentration (R=0.76, p=1.58x10^{-7}), and tumor fraction (R=0.65, p=2.86x10^{-5}), and negatively correlated with fragment size ratio (R=-0.47, p=0.0053) in lymphomas. There was no significant correlation between cfDNA metrics and white blood cell count in other cancer types overall (Figure 4, Supplementary Table 11). These findings suggest that in dogs with lymphoma, DNA released from the circulating neoplastic white blood cells either *in vivo* or during the sample collection process may increase both the cfDNA concentration and the fraction of cfDNA that is tumor derived, and that the tumor DNA from these cells is less fragmented than might be expected if it had traveled from a non-blood-cell source. This hypothesis is supported by the significantly lower fragment size ratios of Stage V lymphomas, in which circulating neoplastic lymphocytes are present, versus other lymphomas of lower stages.

Blood collection tube type affected sample quality

Whether a sample was collected in a conventional EDTA tube or a specialized blood collection tube has a small effect on the amount and quality of cell-free DNA collected. Streck tubes are designed for liquid biopsy applications and contain specialized preservatives to prevent cell lysis[17]. For five dogs, each with a different type of cancer, we collected into both EDTA and Streck tubes at the same visit. There was no effect of tube type on the detected tumor fraction ($mean_{EDTA}$= 0.22, $mean_{Streck}$=0.23, paired $p_{Wilcoxon}$=0.313) (Supplementary Figure 14, Supplementary Data 15). While not significant, fragment length ratios were modestly higher in the Streck tube samples (smaller fragments) ($mean_{EDTA}$= 0.80, $mean_{Streck}$=0.88, paired $p_{Wilcoxon}$=0.063), potentially consistent with decreased white cell lysis or inhibition of DNase activity in the Streck samples[26]. We were not able to assess the effect of tube type on cfDNA concentration because we didn't have total plasma quantities for these five paired samples, so we expanded this comparison to consider all samples where the blood collection tube was known

(N_{EDTA}=39, N_{Streck}=93). After excluding the dogs with lymphoma, which tended to have higher DNA concentrations and tumor fractions, and for which samples were predominantly collected into Streck tubes (N=34) rather than EDTA tubes (N=4), we found that Streck tubes yielded significantly lower concentrations ($p_{Wilcoxon}=4.66x10^{-6}$), potentially indicative of less white cell lysis. Most samples collected in EDTA tubes were processed within 30 minutes, however, in some cases EDTA samples from our pilot were fractionated up to 24 hours post collection. The observation that dogs with lymphoma have higher cfDNA concentrations and tumor fractions was not attributable to tube type.

A small subset of samples (N=12, 2.4%) had modal fragment lengths <100bp, most commonly at 19bp (N=8). While we saw no difference in modal fragment lengths between Streck and EDTA tube samples taken from the same dogs (N=6), in the cohort as a whole 13% (N=5) of EDTA samples had a modal fragment length <100bp, while only 1% of samples (N=7) collected in Streck tubes had a modal fragment length <100bp, suggesting that the fragmentation may have occurred after collection, and that it was more pronounced in EDTA samples.

Blood draws from a central vein yielded more tumor-derived cfDNA

Given the short half-life of cfDNA in the circulation, we asked whether the characteristics of the central or peripheral vasculature or proximity of the venipuncture site to the tumor might affect the amount of tumor-derived cfDNA sampled (and therefore the diagnostic quality of the liquid biopsy). This is difficult to assess in human patients, where a single sample is generally drawn at any given time point, and the venipuncture site is standardized (typically drawn from the median cubital or cephalic veins[28]). Despite this being a potentially simple procedure to optimize, only a few case reports in the human literature exist for adjusting sampling site to increase tumor-derived cfDNA yield[97], and none of these reports assessed tumor fraction.

To assess whether the vein used for liquid biopsy sampling affects liquid biopsy metrics, we collected blood samples from three different veins (jugular, cephalic, and saphenous veins) in the same dog at a single time point in a cohort of 10 dogs with various cancer types, and collected blood samples from two different veins (jugular and either cephalic or saphenous) at a single time point in a cohort of nine dogs (N=19 dogs, 48 samples). All blood draws were performed using the same phlebotomy procedures, via percutaneous venipuncture with a sterile intravenous needle (20 or 22 gauge). The distances between the blood draw site and the tumor were estimated from all known tumor sites present in each dog, including any metastases, and all dogs underwent staging tests in addition to physical examination to determine the extent of the disease in the body. We mapped the approximate tumor and blood draw site locations onto a standardized canine anatomy diagram and measured the distance between the tumor and blood draw sites. We considered both the “simple” straight-line distance and the “circulation” distance, which traces the path of the blood flow through the circulatory system from the tumor to the blood draw site (Supplementary Data 16).

We performed linear mixed models using dog ID as a random effect (controlling for factors that differed between dogs such as dog size or conformation), to assess whether vein or distance were associated with cfDNA metrics. On multivariable analysis, blood draw site (vein) had a significant impact on tumor fraction (p=0.0034, ges=0.239) and fragment size ratio (p=0.0033, ges=0.237) (Figure 5, Supplementary Table 12). While only fragment size ratio remained significant in the robust test after trimming outliers, filtering out dogs where all samples had tumor fractions below the limit of detection for ichorCNA in humans (0.03)[12] yielded slightly lower p-values for these metrics (tumor fraction, p=0.0016, ges=0.346, fragment size ratio, p=0.0033, ges=0.237) with cfDNA concentration and fragment size ratio being significant or borderline significant in the robust test (Supplementary Table 12). These findings are suggestive and will need to be validated in larger studies. Because the jugular vein was always closer to the tumor than the saphenous vein ($p<2.22x10^{-16}$) or the cephalic vein ($p=1.9x10^{-8}$) (Supplementary Figure

15,16), we cannot rule out that there may also be an effect of circulation distance on cfDNA metrics.

Time of day had no significant effect on cfDNA metrics

Studies in humans have not conclusively determined whether the time of day of sampling has an effect on cfDNA yields[19,20,98,99]. However, circadian rhythm and sample timing has been shown to be important in blood cell[19,100,101] and circulating tumor cell counts[102]. We examined the effect of the time of blood sampling in a cohort of 10 dogs with various cancer types. Among three samples taken at intervals of 2 to 3.5 hours (Supplementary Data 17), we found no significant effect of time of day on either cfDNA concentration (p_{LMM}=0.766), tumor fraction (p_{LMM}=0.371), or fragment size ratio (p_{LMM}=0.187) (Supplementary Figure S=17A-C, Supplementary Table 13). We also saw no difference between morning and afternoon blood draws on any metric (Supplementary Figure 17D-E). It is possible, however, that longer sampling intervals (*e.g.,* overnight) would show a more significant effect.

Longitudinal monitoring

Longitudinal cfDNA metrics correlated with disease status

Detection of molecular relapse is an important clinical application of liquid biopsy and is under active development[103,104]. It is estimated that a tumor must reach a size of approximately 1.5 cm in diameter (approximately 1.77 mL volume[105]) before it is detectable by clinical imaging, thus, using liquid biopsy to identify molecular relapse before the tumor is visible on imaging may improve patient outcomes. Such studies take years in human patient cohorts, but can be performed much more quickly in canine patients, who have an accelerated clinical course and shorter lifespans. In order to evaluate the comparative utility of the dog model to help optimize longitudinal patient monitoring, we assessed cfDNA dynamics and dog clinical status over short (hours to days) and long (weeks to months) timescales. We collected multiple cfDNA samples over the course of treatment from dogs with lymphoma or osteosarcoma enrolled in clinical trials at the Tufts Foster Hospital for Small Animals, noting, at each time point, disease status (whether gross disease was present and whether there was a response to therapy) based on patient clinical status and RECIST criteria (Response Evaluation Criteria in Solid Tumors)[106], which standardize the designation of response based on changes in tumor size. Samples collected longitudinally at appointments where treatment was administered were always taken pre-treatment. We did not collect paired tumor biopsies, although in some cases samples were collected prior to surgical biopsy or lymph node aspiration for clinical use.

cfDNA metrics change within hours of L-asparaginase treatment

Little is known from either mouse models[107] or human patients[108] about the dynamics of cfDNA over short time frames after treatment. Time scales for blood collection are often on weekly to monthly or even quarterly intervals whereas the half-life of cfDNA is on the order of minutes to hours[109]. Identifying the post-treatment interval in which liquid biopsy samples have the highest tumor content, or in which cfDNA metrics are most predictive of response could help to maximize tumor-derived cfDNA yield and provide an early assessment of treatment efficacy. To examine cfDNA dynamics over short (6 hours to 1 week) time periods, we recruited four dogs with lymphoma (three with multicentric diffuse large B-cell lymphoma (DLBCL) and one with epitheliotropic lymphoma). For each dog, we took blood samples 6-12 hours apart during the 24 hours prior to treatment with one dose of L-asparaginase, then again during the 24 hours following treatment, and, finally, during the 24 hour period one week later. We assessed the relationship between cfDNA metrics from the three multicentric samples and the sum of the five lymph node sizes measured for RECIST[106] staging each day (in the three dogs with DLBCL), as well as the clinical assessment of the attending veterinarian (Supplementary Data 18).

Relative changes in all three metrics qualitatively followed the clinical responses of the dogs (Figure 6, Supplementary Figure 18). Fragment size ratios increased by 12 hours post-treatment in all dogs (Figure 6D), consistent with an increase in apoptosis and tumor DNA release into the circulation immediately post-treatment. For three of the dogs, fragment size ratios then decline, while in the dog who had a good clinical response overall, fragment size ratio remained elevated at 48 hours post-treatment (Figure 6D). These findings suggest that very early post-treatment changes and kinetics may be more informative than a single sample taken one week post-treatment, although a larger cohort will be needed to further assess whether short-term cfDNA kinetics post-treatment are predictive of response to therapy.

Changes in cfDNA metrics are detected at clinical relapse in lymphoma and osteosarcoma

The clinical utility of using liquid biopsy metrics, including tumor fraction and cfDNA concentration, for monitoring molecular response to assess the effectiveness of therapy and as predictors for outcomes such as progression-free survival in human oncology is a promising and active area of investigation[110,111]. We therefore asked how well liquid biopsy values compare with clinical metrics across a full course of cancer treatment, which can be completed in much shorter time frames than human studies — approximately one year for canine lymphoma and osteosarcoma patients.

Longitudinal data was collected from a total of 34 dogs with DLBCL and 21 dogs with appendicular osteosarcoma undergoing treatment as part of clinical trials at the Cummings School of Veterinary Medicine at Tufts University (Supplementary Data 1). Samples from 3 - 5 clinical visits were collected from the dogs with lymphoma, including screening/day 0, day 7 or 14, day 21 or 28, the visit prior to clinical relapse (defined retrospectively), and clinical relapse. A subset of six dogs had not yet relapsed, so the sample from the latest visit was analyzed. For dogs with osteosarcoma, samples from 3 - 8 time points were collected every two weeks for the first month (including screening visit and amputation at week 2), then monthly thereafter until relapse. Again, for a subset of 6 dogs that had yet to relapse, samples up to the latest time point collected were included.

We assessed pre- and post-amputation time points from the 21 dogs with osteosarcoma to benchmark the ability of tumor fraction to identify a clinically well-defined decrease in tumor burden. The median change in tumor fraction was 0.135 +/-0.154, with 91% of dogs showing a decrease in tumor fraction (Supplementary Data 19). We noted similar trends in fragment size ratios in 71.4% of dogs. Because all dogs underwent amputation to remove the primary tumor, we were unable to compare these trends between dogs that had the intervention and those that did not.

cfDNA metrics were significantly associated with disease status across all time points in both cancer types (Supplementary Figures S19-S23). We used paired linear mixed models with dog ID as a random effect to compare initial presentation (“disease present”) to intermediate time points with no clinical evidence of disease (“no evidence of disease”) and progressive disease (Figure 7, Supplementary Figure 24, Supplementary Table 13). Dogs were considered to have “no evidence of disease” if they achieved a complete response and did not yet have progressive disease based on lymph node measurements and complete blood counts (lymphoma) or if they were post-amputation and no metastases had been detected clinically on physical examination or thoracic radiographs (osteosarcoma). Between the “disease present” and “no evidence of disease” time points, where cfDNA metrics would be expected to decrease, cfDNA concentration significantly decreased in lymphomas ($p=6x10^{-8}$) (Figure 7A), tumor fraction significantly decreased in lymphoma ($p=2.1x10^{-7}$) and osteosarcoma ($p=2.4x10^{-5}$). Between the “no evidence of disease” and “progressive disease” time points, where cfDNA metrics would be expected to increase, we found that cfDNA concentration ($p=1.9x10^{-5}$) and tumor fraction ($p=2.7x10^{-5}$) significantly increased in lymphoma (Figure 7B). These changes in cfDNA metrics suggest that MRD was increasing with disease burden and molecular evidence of relapse was detectable at the clinical relapse time point. Changes in fragment size ratio were not as predictive, with nominal

changes seen in lymphomas and osteosarcomas, but only the decrease in fragment size ratio between “disease present” and “no evidence of disease” remained significant on robust testing ($p=0.0049$, $p_{robust}=0.0071$, Supplementary Table 13)

Changes in cfDNA metrics are detected prior to clinical relapse in lymphoma and osteosarcoma

In some cases, altered cfDNA metrics were identified at time points where dogs had no clinical evidence of disease, suggesting that minimal residual disease (MRD) was being detected by liquid biopsy at these time points (Supplementary Figures 19-23). To investigate whether evidence of molecular relapse prior to clinical relapse could be identified, we assessed trends in cfDNA metrics from screening to relapse. We identified 17 dogs (9 with lymphoma and 8 with osteosarcoma) with four or more longitudinal samples, including one time point with no evidence of disease and the clinical relapse time point (with at least one intervening time point) (Supplementary Data 20). We found that there were patterns of increasing tumor fraction with increases continuing at each time point from initial increase until clinical relapse. This pattern of increasing tumor fraction was seen in 8/9 (89%) of dogs with lymphoma and 5/8 (63%) of dogs with osteosarcoma. These trends started an average of 41 days prior to clinical relapse in lymphoma and 30 days prior in the osteosarcomas, although it should be noted that the sampling interval was dictated by the frequency of clinical trial visits. We identified similar trends in cfDNA concentration in lymphomas, which increased in 2 dogs an average of 30 days earlier than clinical relapse, and in fragment size ratios, which increased in 2 dogs with osteosarcoma an average of 30 days before clinical relapse, suggesting that these metrics may also be useful for patient monitoring.

While we cannot confirm the presence of disease at all time points where cfDNA metrics suggest relapse, we were able to examine a subset of samples where disease presence was not ascertained clinically at the time, but was recognized to have been present retrospectively. In four dogs with osteosarcoma, pulmonary nodules were either (1) present in thoracic radiographs but only identified retrospectively, or (2) evidence of metastases was equivocal and not classified as metastases until progression was seen at a later clinical visit. In all cases, the tumor fraction at the time point where disease progression was later identified was higher than at the previous time point (linear mixed model, $p=0.053$, Supplementary Figure 25, Supplementary Table 15), suggesting that liquid biopsy could provide additional clinically relevant information prior to clinical relapse.

Using Cox proportional hazards models, we investigated whether cfDNA metrics were predictive of outcome. We implemented both univariable and multivariable models, correcting for age, sex and spay-neuter status, weight, body condition score, breed, disease status, and tumor size). In lymphomas, cfDNA concentration at Day 0 was associated with progression-free survival (PFS) in the univariable model ($p=0.036$) (Supplementary Figure 26A, Supplementary Table 16). At Day 7, when all dogs had achieved either a complete or partial response, cfDNA concentration ($p_{univariable}=0.00465$, $p_{multivariable}=0.0042$) and tumor fraction ($p_{univariable}=5.6\times10^{-4}$, $p_{multivariable}=0.0197$) were significantly associated with PFS in both the univariable and multivariable models. Baseline and post-amputation time points were not significantly associated with outcome in osteosarcoma. Our results lend weight to previous canine studies that tumor-derived cfDNA levels can be predictive of outcome[38,39,44], identify residual disease after curative-intent surgery[47,51], and predict relapse when patients were monitored longitudinally with serial liquid biopsies[42,45].

DISCUSSION

In this study, we demonstrated that spontaneous canine cancer is a powerful comparative model for optimizing liquid biopsy techniques. To validate the canine model, we confirmed that several key metrics for dog cancers—including tumor-cfDNA concordance, cfDNA plasma concentration, tumor fraction, and fragment size ratios—closely match previous reports from human patients.

Leveraging the higher incidence of cancer types such as lymphoma and osteosarcoma in dogs, their rapid clinical course, and flexibility in veterinary medicine with respect to phlebotomy and treatment protocols (when compared to human medicine), we identified factors affecting cfDNA metrics that have not been reported previously or that are not yet commonly considered in liquid biopsy optimization. These include differences between lymphoma and nonhematopoietic cancers and the impact of central venous sampling and circulatory distance on tumor fraction.

While numerous studies have demonstrated the feasibility of liquid biopsy in pet dogs with cancer[38–47,50,51,94,112–114], ours is the first to specifically evaluate canine liquid biopsy performance in the context of comparable human studies. The modal fragment lengths of canine cfDNA samples were nucleosomal, as reported in human cfDNA. We found that despite differences in cancer type, and our inability to control for disease stage and other factors, the ranges of cfDNA metrics in dogs with and without cancer were similar to metrics in humans with and without cancer, that concordance of somatic mutations between tumor and cfDNA samples was very similar to that reported in humans[12], and that more mutations are identified in cfDNA than in tumor tissue, as has been reported in humans[91]. Taken together, our findings indicate that liquid biopsy performance is broadly similar across dogs and humans, verifying that spontaneous cancers in dogs are an excellent model system for exploring novel methodologies to advance liquid biopsy performance and accuracy.

We identified key differences between cfDNA metrics between lymphomas and non-hematopoietic cancers that point to cancer type as an important consideration in clinical interpretation of liquid biopsy findings. Overall, lymphomas tended to have higher cfDNA concentrations and tumor fractions but lower fragment size ratios than non-hematopoietic cancer types. Total white blood cell count was associated with higher cfDNA concentration and tumor fraction and lower fragment size ratios in lymphomas, but not other cancer types. These findings suggest that while cfDNA from normal white blood cells dilutes cfDNA with normal germline DNA, in individuals with lymphoma, circulating neoplastic lymphocytes instead contribute tumor DNA. This is further supported by our finding that white cell count is correlated with disease stage and response criteria. It has been reported that tumor cfDNA tends to be more highly fragmented than normal cfDNA[115]. Our findings suggest cfDNA contributed by white blood cells directly in the bloodstream is less fragmented than DNA from extravascular sources.

Our finding that the blood draw site is associated with cfDNA metrics is important for understanding cfDNA sampling and kinetics. Because the jugular vein was always the closest site in our study, further research will be necessary to identify the underlying cause for the increase in tumor cfDNA yield (including aggregation of blood from large areas of the peripheral circulation, and potentially from unknown metastases, or lower likelihood of sampling of vascular endothelial cells[116]). In cancer types where tumor cfDNA yield is consistently low, sampling from a central vein may increase diagnostic benefit. Because the jugular site was always most proximal to the tumor, additional samples will be necessary to confirm whether there is an additional effect of circulation distance on cfDNA metrics, particularly when sampling venous drainage from the tumor. A recent study reported higher concentrations of cfDNA and more reliable identification of tumor mutations from blood samples more proximal to the tumor in three human patients[97], suggesting that selective venous sampling could lead to improved tumor-derived cfDNA yield and assay performance in humans as well.

This is the first study to evaluate cfDNA at such dense time points pre- and post-treatment. We utilized this data to identify a pattern in fragments size ratios post-treatment that may indicate increased levels of treatment-induced apoptosis. In addition, the longitudinal studies across treatment of dogs with lymphoma and osteosarcoma are the largest to date in single cancer types. While additional studies will be required to confirm these results and test the sensitivity and specificity of these metrics in the clinical setting, our findings demonstrate concordance of liquid

biopsy metrics with clinical status, and suggest their utility for identifying molecular relapse prior to clinical detection of cancer.

Key limitations of our study include the small sample size of some cohorts, including the number of healthy controls, which limits our statistical power to identify differences between healthy dogs and those with cancer, and the small number of dogs with matched cfDNA and tumor mutation calling. Lack of comprehensive staging information and clinical metadata for all dogs limited our ability to control for all possible sources of variance in every analysis. While our statistical methods ensure that our findings are robust, they cannot address structural limitations of our dataset, including collinearity between collection site, sequencing group, and extraction and sequencing protocols, as well as partial confounding between cancer type and blood collection tube type. Due to this collinearity, we cannot rule out potential biological effects involving these individual affected variables, as they cannot be separated in our analyses. Given the exploratory nature of our studies, findings will need to be confirmed in larger studies.

In conclusion, our research illuminates the potential of spontaneous cancers in pet dogs as a comparative model for liquid biopsy approaches. In addition to identifying key factors for future optimization studies, this work positions spontaneous cancers in pet dogs to be more broadly employed for studies designed to innovate, optimize and validate liquid biopsy methodologies prior to human translation.

Data Availability

Sequence data that support the findings of this study have been deposited in NCBI Sequence Read Archive (SRA) under BioProject PRJNA1151136. Fragment size counts for all samples are available on GitHub at [79]. All other data supporting the findings of this study are available within the paper and its supplementary files. The source data for Figure 2 is provided in Supplementary Data 1, 2, 8, and 10. Source data for Figure 3 is provided in Supplementary Data 1. Source data for Figure 4 is provided in Supplementary Data 1 and 14. Source data for Figure 5 is provided in Supplementary Data 1 and 16. Source data for Figure 6 is provided in Supplementary Data 1, 17, and 18. Source data for Figure 7 is provided in Supplementary Data 1.

Acknowledgements

Funding: This work was supported by the following grants from the National Institutes of Health: R01CA255319 (Karlsson, London), R37CA218570 (Karlsson, London), NIH U01CA224153 (London), and NIH U01CA224182 (London). KM was supported by F32CA247088 and K01OD036469-01A1. HLG was supported by K01OD028268-01A1. The content of this manuscript is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Support was also provided by an American Cancer Society Research Scholar Grant (Gardner, Karlsson), an MIB Agents Outsmarting Osteosarcoma grant (Gardner, London), and the Dr. Eileen L. Berman and Stanley I. Berman Foundation. The authors also acknowledge the generous support of the Gerstner Family Foundation.

The authors would like to thank Paul Maza, DVM, PhD, Senior Lecturer in Anatomy at the Cornell University College of Veterinary Medicine, for guidance on the blood draw site diagrams and circulatory distance measurements. We would also like to thank Leslie Gaffney of the Broad Institute Communications Laboratory for her work designing the overview figure.

Author contributions:

KM: conceptualization, methodology, investigation, formal analysis, visualization, data curation, writing–original draft preparation, project administration, writing–review and editing, supervision

CH: methodology, formal analysis, data curation, writing–original draft preparation, writing–review and editing

JR: software, methodology, investigation, writing–original draft preparation, writing–review and editing

MEW: methodology, investigation, formal analysis, data curation,writing–review and editing

DG: methodology, investigation, writing–review and editing

FLC: methodology, data curation, writing–review and editing

KX: methodology, investigation, formal analysis, writing–original draft preparation, writing–review and editing

EJK: formal analysis, writing–review and editing

RS: methodology

CP: conceptualization, writing–review and editing

VA: conceptualization, methodology, resources, writing–review and editing, supervision

CAL: conceptualization, resources, investigation, data curation, project administration, writing–review and editing, supervision

HLG: conceptualization, methodology, resources, investigation, data curation, project administration, writing–review and editing, supervision

EKK: conceptualization, investigation, visualization, data curation, project administration, writing–review and editing, supervision

Competing interests: EKK, CAL, and CP are members of The One Health Company advisory board. CAL is an advisor to Merck Animal Health. CP is an employee of Precede Biosciences. KM, RS, VAA, CAL, HLG, and EKK are contributors to the canine exome panel described in this work. The Broad Institute has licensed certain rights related to the canine exome panel, and KM, RS, VAA and EKK may receive licensing revenue from the Broad Institute in accordance with its licensing revenue sharing policy. VAA is a co-inventor of the MAESTRO MRD test which has been licensed to Exact Sciences and is a co-founder and advisor to Amplifyer Bio. VAA is also a co-inventor on patent applications filed by the Broad Institute on cfDNA sequencing methods utilized in this manuscript. All other authors declare no competing interest.

References

1. De Paolis, E. *et al.* Genomics and epigenomics approaches for the quantification of circulating tumor DNA in liquid biopsy: Relevance of a multimodal strategy. *Int. J. Mol. Sci.* **26**, 10982 (2025).
2. Bao, Y., Zhang, D., Guo, H. & Ma, W. Beyond blood: Advancing the frontiers of liquid biopsy in oncology and personalized medicine. *Cancer Sci.* **115**, 1060–1072 (2024).
3. Swarup, N. *et al.* Cell-free DNA: Features and attributes shaping the next frontier in liquid biopsy. *Mol. Diagn. Ther.* **29**, 277–290 (2025).
4. FDA Approves Blood Tests That Can Help Guide Cancer Treatment. *National Cancer Institute* https://www.cancer.gov/news-events/cancer-currents-blog/2020/fda-guardant-360-foundation-one-cancer-liquid-biopsy (2020).
5. Kwapisz, D. The first liquid biopsy test approved. Is it a new era of mutation testing for non-small cell lung cancer? *Ann Transl Med* **5**, 46 (2017).
6. Woodhouse, R. *et al.* Clinical and analytical validation of FoundationOne Liquid CDx, a novel 324-Gene cfDNA-based comprehensive genomic profiling assay for cancers of solid tumor origin. *PLoS One* **15**, e0237802 (2020).
7. Bauml, J. M. *et al.* Clinical validation of Guardant360 CDx as a blood-based companion diagnostic for sotorasib. *Lung Cancer* **166**, 270–278 (2022).
8. Christou, N. *et al.* Circulating Tumour Cells, Circulating Tumour DNA and Circulating Tumour miRNA in Blood Assays in the Different Steps of Colorectal Cancer Management, a Review of the Evidence in 2019. *BioMed Research International* vol. 2019 1–11 Preprint at https://doi.org/10.1155/2019/5953036 (2019).
9. Manier, S. *et al.* Whole-exome sequencing of cell-free DNA and circulating tumor cells in multiple myeloma. *Nat. Commun.* **9**, 1691 (2018).
10. Parikh, A. R. *et al.* Liquid versus tissue biopsy for detecting acquired resistance and tumor heterogeneity in gastrointestinal cancers. *Nat. Med.* **25**, 1415–1421 (2019).

11. Stover, D. G. *et al.* Association of Cell-Free DNA Tumor Fraction and Somatic Copy Number Alterations With Survival in Metastatic Triple-Negative Breast Cancer. *J. Clin. Oncol.* **36**, 543–553 (2018).
12. Adalsteinsson, V. A. *et al.* Scalable whole-exome sequencing of cell-free DNA reveals high concordance with metastatic tumors. *Nat. Commun.* **8**, 1–13 (2017).
13. Thompson, J. C. *et al.* Detection of Therapeutically Targetable Driver and Resistance Mutations in Lung Cancer Patients by Next-Generation Sequencing of Cell-Free Circulating Tumor DNA. *Clin. Cancer Res.* **22**, 5772–5782 (2016).
14. Chu, D. & Park, B. H. Liquid biopsy: unlocking the potentials of cell-free DNA. *Virchows Arch.* **471**, 147–154 (2017).
15. Lockwood, C. M. *et al.* Recommendations for Cell-Free DNA Assay Validations: A Joint Consensus Recommendation of the Association for Molecular Pathology and College of American Pathologists. *J. Mol. Diagn.* **25**, 876–897 (2023).
16. Heitzer, E., Haque, I. S., Roberts, C. E. S. & Speicher, M. R. Current and future perspectives of liquid biopsies in genomics-driven oncology. *Nat. Rev. Genet.* **20**, 71–88 (2019).
17. Kang, Q. *et al.* Comparative analysis of circulating tumor DNA stability In K3EDTA, Streck, and CellSave blood collection tubes. *Clin. Biochem.* **49**, 1354–1360 (2016).
18. Bettegowda, C. *et al.* Detection of circulating tumor DNA in early- and late-stage human malignancies. *Sci. Transl. Med.* **6**, 224ra24 (2014).
19. Tóth, K. *et al.* Circadian Rhythm of Methylated Septin 9, Cell-Free DNA Amount and Tumor Markers in Colorectal Cancer Patients. *Pathol. Oncol. Res.* **23**, 699–706 (2017).
20. Madsen, A. T., Hojbjerg, J. A., Sorensen, B. S. & Winther-Larsen, A. Day-to-day and within-day biological variation of cell-free DNA. *EBioMedicine* **49**, 284–290 (2019).
21. Atamaniuk, J. *et al.* Increased concentrations of cell-free plasma DNA after exhaustive exercise. *Clin. Chem.* **50**, 1668–1670 (2004).

22. Mavropalias, G. *et al.* Changes in plasma hydroxyproline and plasma cell-free DNA concentrations after higher- versus lower-intensity eccentric cycling. *Eur. J. Appl. Physiol.* **121**, 1087–1097 (2021).
23. Meddeb, R., Pisareva, E. & Thierry, A. R. Guidelines for the preanalytical conditions for analyzing circulating cell-free DNA. *Clin. Chem.* **65**, 623–633 (2019).
24. Ignatiadis, M., Sledge, G. W. & Jeffrey, S. S. Liquid biopsy enters the clinic - implementation issues and future challenges. *Nat. Rev. Clin. Oncol.* **18**, 297–312 (2021).
25. Cisneros-Villanueva, M. *et al.* Cell-free DNA analysis in current cancer clinical trials: a review. *Br. J. Cancer* **126**, 391–400 (2022).
26. Ungerer, V., Bronkhorst, A. J. & Holdenrieder, S. Preanalytical variables that affect the outcome of cell-free DNA measurements. *Crit. Rev. Clin. Lab. Sci.* **57**, 484–507 (2020).
27. *WHO Guidelines on Drawing Blood: Best Practices in Phlebotomy*. (World Health Organization, Geneva, 2010).
28. Jiang, C. Y. *et al.* It's not 'just a tube of blood': principles of protocol development, sample collection, staffing and budget considerations for blood-based biomarkers in immunotherapy studies. *J Immunother Cancer* **9**, (2021).
29. Cheon, D.-J. & Orsulic, S. Mouse models of cancer. *Annu. Rev. Pathol.* **6**, 95–119 (2011).
30. Parasuraman, S., Raveendran, R. & Kesavan, R. Blood sample collection in small laboratory animals. *J. Pharmacol. Pharmacother.* **1**, 87–93 (2010).
31. *Guidelines for Survival Blood Collection in Mice and Rats*. https://oacu.oir.nih.gov/system/files/media/file/2022-12/b2-Survival_Blood_Collection_Mice_Rats.pdf (2023).
32. Gardner, H. L., Fenger, J. M. & London, C. A. Dogs as a Model for Cancer. *Annu Rev Anim Biosci* **4**, 199–222 (2016).
33. LeBlanc, A. K. & Mazcko, C. N. Improving human cancer therapy through the evaluation of pet dogs. *Nat. Rev. Cancer* **20**, 727–742 (2020).

34. Nationwide releases findings on cancer in dogs. *American Veterinary Medical Association* https://www.avma.org/news/nationwide-releases-findings-cancer-dogs.

35. Veterinary Research & Analytics Information. *Nationwide Pet Insurance* https://www.petinsurance.com/veterinarians/research/ (2019).

36. LeBlanc, A. K. *et al.* Perspectives from man's best friend: National Academy of Medicine's Workshop on Comparative Oncology. *Sci. Transl. Med.* **8**, 324ps5 (2016).

37. Paoloni, M. & Khanna, C. Translation of new cancer treatments from pet dogs to humans. *Nat. Rev. Cancer* **8**, 147–156 (2008).

38. Schaefer, D. M. W. *et al.* Quantification of plasma DNA as a prognostic indicator in canine lymphoid neoplasia. *Vet. Comp. Oncol.* **5**, 145–155 (2007).

39. Gelaleti, G. B. *et al.* Short interspersed CAN SINE elements as prognostic markers in canine mammary neoplasia. *Oncol. Rep.* **31**, 435–441 (2014).

40. Beck, J. *et al.* Genome aberrations in canine mammary carcinomas and their detection in cell-free plasma DNA. *PLoS One* **8**, e75485 (2013).

41. Lee, K.-H., Shin, T.-J., Kim, W.-H. & Cho, J.-Y. Author Correction: Methylation of LINE-1 in cell-free DNA serves as a liquid biopsy biomarker for human breast cancers and dog mammary tumors. *Sci. Rep.* **9**, 17459 (2019).

42. Tagawa, M., Shimbo, G., Inokuma, H. & Miyahara, K. Quantification of plasma cell-free DNA levels in dogs with various tumors. *Journal of Veterinary Diagnostic Investigation* vol. 31 836–843 Preprint at https://doi.org/10.1177/1040638719880245 (2019).

43. Schabort, J. J. *et al.* ANK2 Hypermethylation in Canine Mammary Tumors and Human Breast Cancer. *Int. J. Mol. Sci.* **21**, (2020).

44. Tagawa, M. *et al.* Quantitative analysis of the BRAF V595E mutation in plasma cell-free DNA from dogs with urothelial carcinoma. *PLoS One* **15**, e0232365 (2020).

45. Prouteau, A. *et al.* Circulating tumor DNA is detectable in canine histiocytic sarcoma, oral malignant melanoma, and multicentric lymphoma. *Sci. Rep.* **11**, 877 (2021).

46. Beffagna, G. *et al.* Circulating Cell-Free DNA in Dogs with Mammary Tumors: Short and Long Fragments and Integrity Index. *PLoS One* **12**, e0169454 (2017).
47. Kruglyak, K. M. *et al.* Blood-based liquid biopsy for comprehensive cancer genomic profiling using next-generation sequencing: An emerging paradigm for non-invasive cancer detection and management in dogs. *Front. Vet. Sci.* **8**, 704835 (2021).
48. Goggs, R. Effect of sample type on plasma concentrations of cell-free DNA and nucleosomes in dogs. *Vet. Rec. Open* **6**, e000357 (2019).
49. Wilson-Robles, H. *et al.* Evaluation of nucleosome concentrations in healthy dogs and dogs with cancer. *PLoS One* **15**, e0236228 (2020).
50. Lorch, G. *et al.* Identification of Recurrent Activating HER2 Mutations in Primary Canine Pulmonary Adenocarcinoma. *Clin. Cancer Res.* **25**, 5866–5877 (2019).
51. Favaro, P. F. *et al.* Feasibility and promise of circulating tumor DNA analysis in dogs with naturally-occurring sarcoma. *bioRxiv* 2020.08.20.260349 (2020) doi:10.1101/2020.08.20.260349.
52. Van der Auwera, G. A. *et al.* From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr. Protoc. Bioinformatics* **43**, 11.10.1–33 (2013).
53. Hoeppner, M. P. *et al.* An improved canine genome and a comprehensive catalogue of coding genes and non-coding transcripts. *PLoS One* **9**, e91172 (2014).
54. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
55. Wang, C. *et al.* A novel canine reference genome resolves genomic architecture and uncovers transcript complexity. *Commun Biol* **4**, 185 (2021).
56. Ramírez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–5 (2016).
57. Lai, D. & Shah, S. HMMcopy: Copy number prediction with correction for GC and mappability bias for HTS data. *R package version* **1**, (2012).

58. *Hmmcopy_utils: Tools for Extracting Read Counts and Gc and Mappability Statistics in Preparation for Running HMMCopy*. (Github).

59. *CanFam3.1.centromere.bed at Master · Chao912/Mischka*. (Github).

60. *GSD1.0_CanFam4.centromere.bed at Master · Chao912/Mischka*. (Github).

61. Morrill, K. *et al.* Ancestry-inclusive dog genomics challenges popular breed stereotypes. *Science* **in press**, (2022).

62. Gardner, H. L. *et al.* Canine osteosarcoma genome sequencing identifies recurrent mutations in DMD and the histone methyltransferase gene SETD2. *Commun Biol* **2**, 266 (2019).

63. Hinrichs, A. S. *et al.* The UCSC Genome Browser Database: update 2006. *Nucleic Acids Res.* **34**, D590–8 (2006).

64. Megquier, K. *et al. Broad UMass Chan Canine Exome Whitepaper*. https://karlssonlab.org/2024/03/27/exome/broad-umass-chan-canine-exome-whitepaper-2023/ (2023).

65. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

66. Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *Gigascience* **10**, (2021).

67. Benjamin, D. *et al.* Calling Somatic SNVs and Indels with Mutect2. *bioRxiv* 861054 (2019) doi:10.1101/861054.

68. Kim, S. *et al.* Strelka2: fast and accurate calling of germline and somatic variants. *Nat. Methods* **15**, 591–594 (2018).

69. Cooke, D. P., Wedge, D. C. & Lunter, G. A unified haplotype-based method for accurate and comprehensive variant calling. *Nat. Biotechnol.* **39**, 885–892 (2021).

70. Cingolani, P. *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly* **6**, 80–92 (2012).

71. Ruden, D. *et al.* Using Drosophila melanogaster as a Model for Genotoxic Chemical Mutational Studies with a New Program, SnpSift. *Front. Genet.* **3**, 35 (2012).
72. R Core Team. R: A language and environment for statistical computing. Preprint at (2022).
73. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)* vol. 57 289–300 Preprint at https://doi.org/10.1111/j.2517-6161.1995.tb02031.x (1995).
74. Welch, B. L. The Generalization of `Student's' Problem when Several Different Population Variances are Involved. *Biometrika* **34**, 28 (1947).
75. McGraw, K. O. & Wong, S. P. Forming inferences about some intraclass correlation coefficients. *Psychol. Methods* **1**, 30–46 (1996).
76. Cox, D. R. Regression models and life-tables. *J. R. Stat. Soc. Series B Stat. Methodol.* **34**, 187–202 (1972).
77. OpenAI. ChatGPT version 4o [Large Language Model]. *ChatGPT* https://chatgpt.com/ (2025).
78. Google. Gemini 3 Pro [Large Language Model]. https://gemini.google.com/ (2025).
79. Megquier, K. *et al. Broadinstitute/canine-Liquid-Biopsy: Canine Liquid Biopsy Release 1*. (Zenodo, 2026). doi:10.5281/ZENODO.19456313.
80. American Veterinary Medical Association & Others. AVMA pet ownership and demographics sourcebook. (Colorado State University. Libraries, 2018).
81. Montoya, M. *et al.* Life expectancy tables for dogs and cats derived from clinical data. *Front. Vet. Sci.* **10**, 1082102 (2023).
82. Fan, H. C., Blumenfeld, Y. J., Chitkara, U., Hudgins, L. & Quake, S. R. Noninvasive diagnosis of fetal aneuploidy by shotgun sequencing DNA from maternal blood. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 16266–16271 (2008).
83. Lo, Y. M. D. *et al.* Maternal plasma DNA sequencing reveals the genome-wide genetic and

mutational profile of the fetus. *Sci. Transl. Med.* **2**, 61ra91 (2010).

84. Favaro, P. F. *et al.* Feasibility of circulating tumor DNA analysis in dogs with naturally occurring malignant and benign splenic lesions. *Sci. Rep.* **12**, 6337 (2022).
85. Vogt, S. L. *et al.* Feasibility of cell-free DNA collection and clonal immunoglobulin sequencing in south African patients with HIV-associated lymphoma. *JCO Glob. Oncol.* **7**, 611–621 (2021).
86. Peneder, P. *et al.* Multimodal analysis of cell-free DNA whole-genome sequencing for pediatric cancers with low mutational burden. *Nat. Commun.* **12**, 3230 (2021).
87. Rickles-Young, M. *et al.* Assay Validation of Cell-Free DNA Shallow Whole-Genome Sequencing to Determine Tumor Fraction in Advanced Cancers. *J. Mol. Diagn.* **26**, 413–422 (2024).
88. Cescon, D. W., Bratman, S. V., Chan, S. M. & Siu, L. L. Circulating tumor DNA and liquid biopsy in oncology. *Nat Cancer* **1**, 276–290 (2020).
89. Wang, P. P. S., Parker, W. T., Branford, S. & Schreiber, A. W. BAM-matcher: a tool for rapid NGS sample matching. *Bioinformatics* **32**, 2699–2701 (2016).
90. Rebhun, R. B. & Thamm, D. H. Multiple distinct malignancies in dogs: 53 cases. *J. Am. Anim. Hosp. Assoc.* **46**, 20–30 (2010).
91. Pereira, B. *et al.* Cell-free DNA captures tumor heterogeneity and driver alterations in rapid autopsies with pre-treated metastatic cancer. *Nat. Commun.* **12**, 3199 (2021).
92. Wu, X., Song, P., Guo, L., Ying, J. & Li, W. Mutant-Allele Tumor Heterogeneity, a Favorable Biomarker to Assess Intra-Tumor Heterogeneity, in Advanced Lung Adenocarcinoma. *Front. Oncol.* **12**, 888951 (2022).
93. Hsiehchen, D., Espinoza, M., Gerber, D. E. & Beg, M. S. Clinical and biological determinants of circulating tumor DNA detection and prognostication using a next-generation sequencing panel assay. *Cancer Biol. Ther.* **22**, 455–464 (2021).
94. Kim, J. *et al.* Cell-Free DNA as a Diagnostic and Prognostic Biomarker in Dogs With

Tumors. *Front Vet Sci* **8**, 735682 (2021).

95. Lui, Y. Y. N. *et al.* Predominant hematopoietic origin of cell-free DNA in plasma and serum after sex-mismatched bone marrow transplantation. *Clin. Chem.* **48**, 421–427 (2002).
96. Razavi, P. *et al.* High-intensity sequencing reveals the sources of plasma circulating cell-free DNA variants. *Nat. Med.* **25**, 1928–1937 (2019).
97. Damascelli, B., Tichà, V., Repetti, E. & Dorji, T. Beyond Standard Practice in Liquid Biopsy: Selective Venous Sampling. *J. Vasc. Interv. Radiol.* **32**, 668–671 (2021).
98. Wagner, J. T. *et al.* Diurnal stability of cell-free DNA and cell-free RNA in human plasma samples. *Sci. Rep.* **10**, 16456 (2020).
99. Korabecna, M., Horinek, A., Bila, N. & Opatrna, S. Circadian Rhythmicity and Clearance of Cell-Free DNA in Human Plasma. in *Circulating Nucleic Acids in Plasma and Serum* 195–198 (Springer Netherlands, 2011).
100. Feriel, J., Tchipeva, D. & Depasse, F. Effects of circadian variation, lifestyle and environment on hematological parameters: A narrative review. *Int. J. Lab. Hematol.* **43**, 917–926 (2021).
101. Busza, A. *et al.* 0028 Systematic Review: Time of day differences in complete blood count values. *Sleep* **46**, A12–A12 (2023).
102. Zhu, X. *et al.* In vivo flow cytometry reveals a circadian rhythm of circulating tumor cells. *Light Sci Appl* **10**, 110 (2021).
103. Spindler, K.-L. G. & Jakobsen, A. Circulating tumor DNA: Response Evaluation Criteria in Solid Tumors - can we RECIST? Focus on colorectal cancer. *Ther. Adv. Med. Oncol.* **15**, 17588359231171580 (2023).
104. Watanabe, K., Nakamura, Y. & Low, S.-K. Clinical implementation and current advancement of blood liquid biopsy in cancer. *J. Hum. Genet.* **66**, 909–926 (2021).
105. Erdi, Y. E. Limits of Tumor Detectability in Nuclear Medicine and PET. *Mol. Imaging Radionucl. Ther.* **21**, 23–28 (2012).

106. Eisenhauer, E. A. *et al.* New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur. J. Cancer* **45**, 228–247 (2009).

107. Terasawa, H. *et al.* Circulating tumor DNA dynamics analysis in a xenograft mouse model with esophageal squamous cell carcinoma. *World J. Gastroenterol.* **27**, 7134–7143 (2021).

108. Walls, G. M. *et al.* Early circulating tumour DNA kinetics measured by ultra-deep next-generation sequencing during radical radiotherapy for non-small cell lung cancer: a feasibility study. *Radiat. Oncol.* **15**, 132 (2020).

109. Diehl, F. *et al.* Circulating mutant DNA to assess tumor dynamics. *Nat. Med.* **14**, 985–990 (2008).

110. Stadler, J.-C. *et al.* Current and Future Clinical Applications of ctDNA in Immuno-Oncology. *Cancer Res.* **82**, 349–358 (2022).

111. Thompson, J. C., Scholes, D. G., Carpenter, E. L. & Aggarwal, C. Molecular response assessment using circulating tumor DNA (ctDNA) in advanced solid tumors. *Br. J. Cancer* **129**, 1893–1902 (2023).

112. Flory, A. *et al.* Clinical validation of a next-generation sequencing-based multi-cancer early detection 'liquid biopsy' blood test in over 1,000 dogs using an independent testing set: The CANcer Detection in Dogs (CANDiD) study. *PLoS One* **17**, e0266623 (2022).

113. Guil-Luna, S. *et al.* Analysis of cell-free DNA concentration, fragmentation patterns and TP53 gene expression in mammary tumor-bearing dogs: A pilot study. *Front Vet Sci* **10**, 1157878 (2023).

114. McCleary-Wheeler, A. L. *et al.* Next-generation sequencing-based liquid biopsy may be used for detection of residual disease and cancer recurrence monitoring in dogs. *Am. J. Vet. Res.* 1–8 (2023).

115. Mouliere, F. *et al.* Enhanced detection of circulating tumor DNA by fragment size analysis. *Sci. Transl. Med.* **10**, (2018).

116. Sharpe, A. N. *et al.* Isolation of endothelial colony-forming cells from blood samples

collected from the jugular and cephalic veins of healthy adult horses. *Am. J. Vet. Res.* **77**, 1157–1165 (2016).

FIGURE LEGENDS

Figure 1. Overview. Illustration of the study aims and methods. **(A)** The cell-free DNA metrics considered were cfDNA concentration in plasma, tumor fraction, and fragment size ratio. **(B)** Concordance between tumor and liquid biopsy. **(C)** Assessing the impact of variables on cfDNA metrics, including patient factors, cancer types, collection tube type, venipuncture site, and time of day. **(D)** Longitudinal patient monitoring over short and long time frames.

Figure 2. Overview of samples and tumor-cfDNA concordance. (A) Number of dogs broken down by number of samples and data types collected. **(B)** Spearman correlation (two-sided) of $\log_2$ read depth ratios between tumor WGS and cfDNA ULPWGS data in 14 dogs with cfDNA tumor fraction >10%. **(C)** Proportion of somatic mutations identified in tumor, cfDNA, and shared between both in samples from five dogs.

Figure 3. Patient factors affecting cfDNA metrics. Comparing cfDNA metrics with pairwise two-tailed t-tests between healthy dogs, all cancers, lymphoma, and non-lymphoma cancers in **(A)** cfDNA concentration, **(B)** tumor fraction, and **(C)** fragment size ratios. **(D)** Linear mixed model results showing the effect of various patient factors on cfDNA metrics in all samples, controlling for dog ID as a random effect. **(E)** ANOVA results showing the effect of various patient factors on cfDNA metrics in the first sample from each dog.

Figure 4. Correlation of white blood cell count to cfDNA metrics. White blood cell count was significantly positively associated with **(A)** cfDNA concentration and **(B)** tumor fraction, and significantly negatively associated with **(C)** fragment size ratio in lymphoma but not in osteosarcomas. In all panels, a two-sided Spearman correlation was performed, and shaded regions indicate the 95% confidence interval around the regression line.

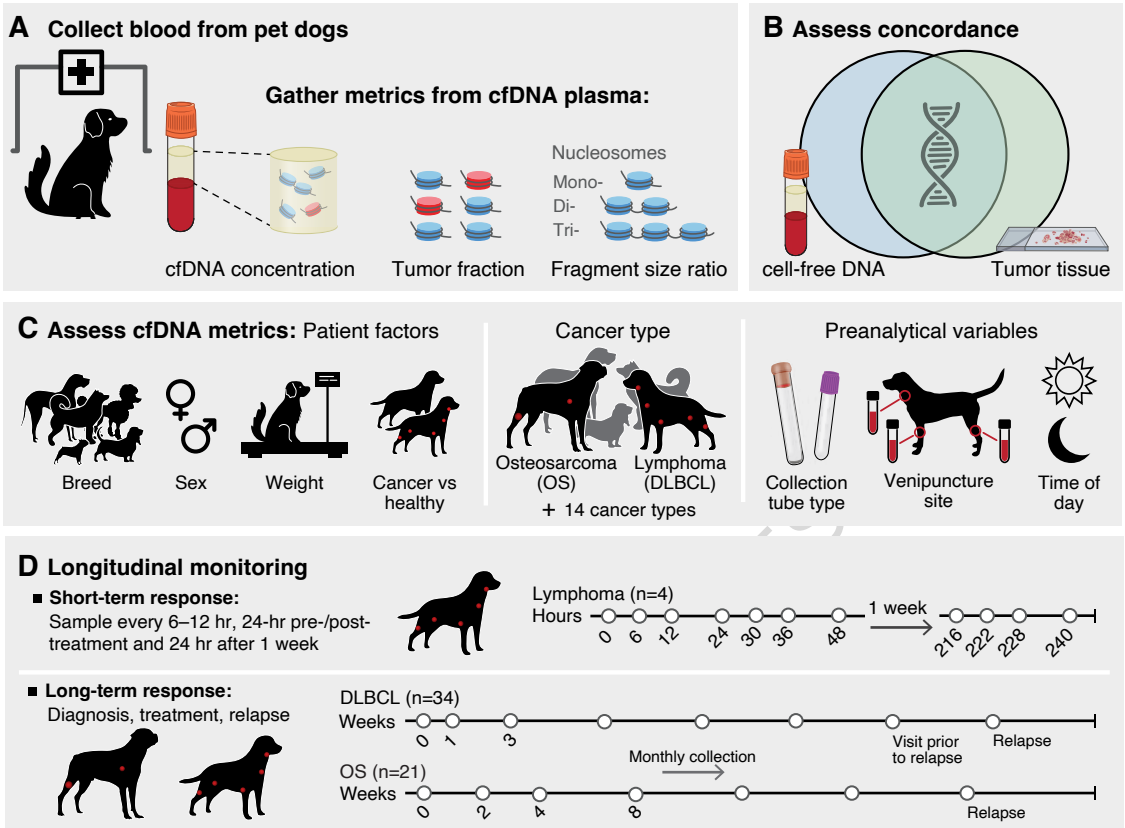
Figure 5. Effect of blood draw site on cfDNA metrics. Boxplots by blood draw site showing **(A)** cfDNA concentration, **(B)** tumor fraction, **(C)** fragment size ratios. Colors represent samples from different venipuncture sites—green: jugular vein, orange: cephalic vein, blue: saphenous vein. **(D)** Linear mixed model analysis (two-sided) showing association between blood draw site and the three cfDNA metrics controlling for dog ID as a random effect.

Figure 6. Acute response monitoring. Short-term cfDNA kinetics in four dogs with lymphoma with samples taken over 24 hours pre- and post-treatment with L-asparaginase (treatment time point indicated by pink bar) and over 24 hours one week later (dotted lines indicate break between days). **(A)** Lymph node size on days 1, 3, and 9 in three dogs with DLBCL. **(B)** Relative change in cfDNA concentration over treatment. **(C)** Relative change in tumor fraction over treatment. **(D)** Relative change in fragment size ratio over treatment. Points are colored by dog ID, line color indicates clinical response (green: poor response, blue: partial response, pink: good response).

Figure 7. Longitudinal monitoring. Plots compare initial pre-treatment time points ("disease present"), remission or post-amputation time points ("no evidence of disease"), and "progressive disease" time points once clinical disease progression was noted. **(A)** cfDNA concentration was associated with disease status in dogs with lymphoma. **(B)** Tumor fraction was significantly associated with disease status in lymphomas and osteosarcomas. **(C)** Fragment size ratio was significantly associated with disease status in osteosarcomas but not lymphomas. Linear mixed model (two-sided) with dog ID included as a random effect.

Editorial summary:
Megquier et al. evaluate pet dogs with cancer as a model for optimizing liquid biopsy via low coverage sequencing of plasma cell-free DNA. They confirm the potential of the dog model, finding that metrics are comparable to those for human cancers and reflect clinical status, then leverage the model to examine the effect of preanalytical variables.

Peer review information: *Communications Medicine* thanks the anonymous reviewers for their contribution to the peer review of this work.


A Collect blood from pet dogs
Gather metrics from cfDNA plasma:
Nucleosomes
Mono-
Di-
Tri-
cfDNA concentration
Tumor fraction
Fragment size ratio
B Assess concordance
cell-free DNA
Tumor tissue
C Assess cfDNA metrics: Patient factors
Breed
Sex
Weight
Cancer vs healthy
Cancer type
Osteosarcoma (OS)
Lymphoma (DLBCL)
+ 14 cancer types
Preanalytical variables
Collection tube type
Venipuncture site
Time of day
D Longitudinal monitoring
Short-term response:
Sample every 6–12 hr, 24-hr pre-/post-treatment and 24 hr after 1 week
Lymphoma (n=4)
Hours
0 6 12 24 30 36 48
1 week
216 222 228 240
Long-term response:
Diagnosis, treatment, relapse
DLBCL (n=34)
Weeks
0 1 3
Monthly collection
Visit prior to relapse
Relapse
OS (n=21)
Weeks
0 2 4 8
Relapse