

METHODS FOR DETERMINING RESPONSIVENESS TO CANCER THERAPY

The present application claims the benefit of priority of U.S. provisional application no. 63/621,749 filed January 17, 2024, which is incorporated by reference herein in its entirety.

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH

[0001] This invention was made with government support under grants CA006973, CA062924, CA121113, CA233259, CA237624 and CA271896 awarded by the National Institutes of Health. The government has certain rights in the invention.

FIELD

[0002] The disclosure is related to methods that rapidly determine disease progression in patients receiving immune checkpoint inhibitor therapy. Methods that distinguish between different cancer subtypes noninvasively are also provided.

BACKGROUND

[0003] Disease monitoring for patients receiving immune checkpoint inhibitors (ICIs) has relied on conventional computed tomography (CT) imaging as the standard of care to guide clinical decisions (1, 2). Despite the advantages of CT-imaging, this approach faces several challenges such as the logistics of image timing taken every 6-8 weeks that may fail to detect the early signs of disease progression, potential misclassification of atypical tumor responses such as pseudo-progression (3) or hyperprogression (4), and the clinically challenging therapeutic decisions for patients with stable or potentially progressive disease categories. There is an unmet clinical need for new approaches that rapidly and reliably detect the early signs of disease progression in this setting.

[0004] Analyses of cell-free DNA (cfDNA) through non-invasive liquid biopsies provide a new opportunity to detect and serially monitor therapeutic responses in a timely manner. To date, most studies that evaluate liquid biopsies in the neoadjuvant or adjuvant setting of patients treated with immunotherapy do so by using targeted next-generation deep sequencing assays to detect molecular responses (5–12). However, these current technologies face several challenges including (i) the initial complexity of target panel design across different cancer types, (ii) high volume of plasma needed, (iii) high-cost associated with deep sequencing and capture panels to

detect tumor-specific mutations, (iv) the lack of any alterations in a subset of patients that could be evaluated by targeted panels, (v) potentially low to no abundance of detectable mutations in the blood for the small number of potentially identifiable alterations, (vi) the necessity of analyzing matched tumor tissue to identify tumor-specific changes, and (vii) analysis of white blood cells for the removal of variants derived from clonal hematopoiesis (5–7, 13, 14). These potential translational limitations indicate the strong need for a less complex and economically viable approach for tumor monitoring that is not challenged by clonal hematopoiesis and do not necessitate invasive tumor biopsies.

SUMMARY

[0005] Embodiments are directed to methods for determining responsiveness of a subject to a cancer therapeutic.

[0006] We have now demonstrated in human cancer patients the use of DNA fragmentomics to assess treatment, including to monitor response to immunotherapy in cancer patients.

[0007] In certain aspects, a method of determining responsiveness to a cancer therapy in a subject, comprising:

- a) obtaining a biological sample from the subject at one or more times after administration of the cancer therapy;
- b) isolating cell free DNA (cfDNA) from the subject's biological samples;
- c) conducting low coverage whole genome sequencing of the subject's isolated cfDNA to generate genomic libraries of sequenced cfDNA fragments;
- d) determining a cfDNA fragmentome profile of the subject's sequenced DNA fragments comprising evaluating one or more cfDNA fragment characteristics;
- e) evaluating the cfDNA fragmentome profiles of a reference panel from other subjects without cancer to the subject's cfDNA fragmentome profile prior to administration of a cancer therapy (baseline), at one or more time points during therapy, after therapy or combinations thereof;

thereby, determining responsiveness to the cancer therapy in the subject. The reference panel suitably is from one or more subjects without cancer that are other or different than the subject the biological sample is obtained at one or more times after administration of the cancer therapy.

[0008] In certain embodiments, the one or more cfDNA fragment characteristics comprises calculating ratios of small to large cfDNA fragments, sequences of one or more cfDNAs, median cfDNA fragment sizes, fragment size distribution, mutant allele frequencies, fragment length, fragment size distribution, fragment end motifs, preferred end coordinates, breakpoint motifs, methylation frequencies or combinations thereof.

[0009] In certain embodiments, a small cfDNA fragment comprises about 80 base pairs (bp) to about 150 bp. In certain embodiments, a large cfDNA fragment comprises about 151 bp to about 300 bp.

[0010] In certain embodiments, the small to large cfDNA ratios are GC corrected.

[0011] In certain embodiments, the mutant allele frequencies in the subject's cfDNA fragmentomes are determined utilizing a hierarchical regression model algorithm.

[0012] In certain embodiments, the cfDNA fragmentome profile comprises the sequence coverage of small cfDNA fragments in windows across the genome. In certain embodiments, the cfDNA fragmentome profile comprises the sequence coverage of large cfDNA fragments in windows across the genome. In certain embodiments, the cfDNA fragmentome profile comprises the sequence coverage of small and large cfDNA fragments in windows across the genome. In certain embodiments, the cfDNA fragmentome profiles in individuals with cancer are altered across the genome prior to treatment and alterations are decreased in subjects responding to the cancer therapy.

[0013] In certain embodiments, the cancer therapy comprises immune checkpoint inhibition, a surgical therapy, chemotherapy, radiation therapy, cryotherapy, hyperthermia treatment, phototherapy, radioablation therapy, hormonal therapy, immunotherapy, small molecule therapy, receptor kinase inhibitor therapy, anti-angiogenic therapy, cytokine therapy, monoclonal antibodies, siRNA, miRNA, antisense oligonucleotides, ribozymes or gene therapy.

[0014] In certain embodiments, the cancer therapy comprises an immune checkpoint inhibitor.

[0015] In certain embodiments, the cancer is a solid tumor. In certain embodiments, the cancer is metastatic.

[0016] In certain embodiments, the method further comprises conducting blood cell counts. In certain embodiments, the biological sample is a plasma sample.

[0017] In another aspect, a method of determining tumor burden in a subject during therapy, is provided and comprises:

a) training a machine learning algorithm with cell free DNA (cfDNA) fragmentome profiles of subjects treated with a cancer therapy, wherein cfDNA tumor fraction levels are independently measured for each sample;

b) evaluating cfDNA fragmentome profiles of subjects undergoing cancer therapy utilizing the trained machine learning algorithm to generate a score relative to a reference or baseline score, wherein the subject's score is decreased as compared to the reference score if the tumor burden is decreased.

[0018] In certain embodiments, the subject's cfDNA fragmentome profiles are assayed at different time points comprising prior to administration of a cancer therapy (baseline), at one or more time points during therapy, after therapy or combinations thereof.

[0019] In certain embodiments, the subject's cfDNA fragmentome profile comprises calculating ratios of small to large cfDNA fragments, sequences of one or more cfDNAs, median cfDNA fragment sizes, fragment size distribution, mutant allele frequencies, fragment length, fragment size distribution, fragment end motifs, preferred end coordinates, breakpoint motifs, methylation frequencies or combinations thereof.

[0020] In certain embodiments, a small cfDNA fragment comprises about 80 base pairs (bp) to about 150 bp.

[0021] In certain embodiments, a large cfDNA fragment comprises about 151 bp to about 300 bp. In certain embodiments, the small to large cfDNA ratios are GC corrected.

[0022] In certain embodiments, the mutant allele frequencies in the subject's cfDNA fragmentomes are determined utilizing a hierarchical regression model algorithm.

[0023] In certain embodiments, the cfDNA fragmentome profile comprises the sequence coverage of small cfDNA fragments in windows across the genome.

[0024] In certain embodiments, the cfDNA fragmentome profile comprises the sequence coverage of large cfDNA fragments in windows across the genome.

[0025] In certain embodiments, the cfDNA fragmentome profile comprises the sequence coverage of small and large cfDNA fragments in windows across the genome.

[0026] In certain embodiments, the step of comparing can include comparing the cfDNA fragmentome profile to a reference cfDNA fragmentome profile over the whole genome. In certain

embodiments, the step of comparing can include comparing the cfDNA fragmentome profile to a reference cfDNA fragmentome profile over a subgenomic interval.

[0027] In certain embodiments, the reference cfDNA fragmentome profile can be a cfDNA fragmentome profile of a healthy mammal. In certain embodiments, the reference cfDNA fragmentome profile can be generated by determining a cfDNA fragmentome profile in a sample obtained from a healthy mammal. In certain embodiments, the reference DNA fragmentome pattern can be a reference nucleosome cfDNA fragmentome profile.

[0028] In certain embodiments, the cfDNA fragmentome profile can include a median fragment size, where a median fragment size of the cfDNA fragmentome profile is shorter than a median fragment size of the reference cfDNA fragmentome profile. In certain embodiments, the cfDNA fragmentome profile can include a fragment size distribution, where a fragment size distribution of the cfDNA fragmentome profile differs by at least 10 nucleotides as compared to a fragment size distribution of the reference cfDNA fragmentome profile.

[0029] In certain embodiments, the cfDNA fragmentome profiles in individuals with cancer are altered across the genome prior to treatment and alterations are decreased in subjects responding to the cancer therapy. In certain embodiments, the cancer therapy comprises immune checkpoint inhibition, a surgical therapy, chemotherapy, radiation therapy, cryotherapy, hyperthermia treatment, phototherapy, radioablation therapy, hormonal therapy, immunotherapy, small molecule therapy, receptor kinase inhibitor therapy, anti-angiogenic therapy, cytokine therapy, monoclonal antibodies, siRNA, miRNA, antisense oligonucleotides, ribozymes or gene therapy. In certain embodiments, the cancer therapy comprises an immune checkpoint inhibitor.

[0030] In certain embodiments, the cancer is a solid tumor. In certain embodiments, the cancer is metastatic. In certain embodiments, the method further comprises conducting blood cell counts.

[0031] In certain embodiments a DELFI (DNA evaluation of fragments for early interception) score is generated, wherein the principle component analysis is incorporated into a machine learning predictive model to generate a score for each subject (DELFI score(s)). In addition to the principal component features, all 39 z-scores were evaluated in a logistic regression model with a LASSO penalty. The optimized LASSO penalty in our analysis was obtained by resampling using the caret R package. References herein to “DELFI score” are values determined by this above specified procedure. See, for example, Mathios, D., Johansen, J.S., Cristiano, S. et

al. Detection and characterization of lung cancer using cell-free DNA fragmentomes. Nat Commun 12, 5060 (2021). doi.org/10.1038/s41467-021-24994-w, incorporated herein by reference in its entirety.

[0032] In certain embodiments, the DELFI scores for non-cancer individuals are less than about 0.3. In certain embodiments, the DELFI scores for stage I cancer are between about 0.3 to less than 0.5. In certain embodiments, the DELFI scores for stage II cancer are between about 0.5 to less than 0.8. In certain embodiments, the DELFI scores for stage III cancer are between about 0.8 to less than 0.99. In certain embodiments, the DELFI scores for stage IV cancer are about 0.99 or greater. In certain embodiments, the DELFI score for stage I cancer is about 0.35. In certain embodiments, the DELFI score for stage II cancer is about 0.75. In certain embodiments, the DELFI score for stage III cancer is about 0.9. In certain embodiments, the DELFI score for stage IV cancer is about 0.99.

[0033] In certain embodiments, the cancer subjects display widespread genome-wide variation.

[0034] In certain embodiments, a method of determining responsiveness to a cancer therapy in a subject, during therapy, comprises obtaining a biological sample from the subject at various times after administration of the cancer therapy; isolating cell free DNA (cfDNA) from the subject's biological samples; conducting blood cell counts prior to administration of therapy, during administration of therapy and after therapy; conducting low coverage whole genome sequencing of the subject's isolated cfDNA to generate genomic libraries of sequenced cfDNA fragments; determining a cfDNA fragmentome profile of the subject's sequenced DNA fragments comprising evaluating one or more cfDNA fragment characteristics; evaluating the cfDNA fragmentome profiles of a reference panel from subjects without cancer to the subject's cfDNA fragmentome profile prior to administration of a cancer therapy (baseline), at one or more time points during therapy, after therapy or combinations thereof; analyzing blood cell counts and comparing blood cell counts to the subject's cfDNA fragmentation profile, wherein the subject's score is decreased as compared to the reference score if the tumor burden is decreased. In certain embodiments, analysis of differences between baseline and longitudinal timepoints of fragmentation profiles and blood counts to evaluate molecular responses. In certain embodiments, results obtained from blood cell counts and fragmentation profiles clinically classify patients that have stable disease and to determine whether the therapy is of benefit to the subject. In certain

embodiments, the method further comprise RECIST classification over course of cancer therapy to further identify tumor responsiveness.

[0035] In certain embodiments, a method of diagnosing and differentiating between breast cancer subtypes, comprises obtaining a biological sample from the subject at various times after administration of the cancer therapy; isolating cell free DNA (cfDNA) from the subject's biological samples; conducting low coverage whole genome sequencing of the subject's isolated cfDNA to generate genomic libraries of sequenced cfDNA fragments; determining a cfDNA fragmentome profile of the subject's sequenced DNA fragments comprising evaluating one or more cfDNA fragment characteristics; evaluating the cfDNA fragmentome profiles of a reference panel from subjects without cancer; assay low coverage whole genome sequences and transcription binding sites, thereby diagnosing and differentiating between breast cancer subtypes. In certain embodiments, changes in transcription factor binding site coverages is diagnostic of a breast cancer subtype. In certain embodiments, assaying of low coverage whole genome sequences and transcription binding sites is conducted at baseline and later timepoints in response to therapy. In certain embodiments, the transcription binding sites comprise: ESR1 (estrogen receptor 1; Ensembl:ENSG00000091831 MIM:133430; AllianceGenome:HGNC:3467), AR (androgen receptor; Ensembl:ENSG00000169083 MIM:313700; AllianceGenome:HGNC:644), GATA3 (Ensembl:ENSG00000107485 MIM:131320; AllianceGenome:HGNC:4172), FOXA1 Forkhead Box A1; HGNC: 5021 NCBI Gene: 3169 Ensembl: ENSG00000129514 OMIM®: 602294 UniProtKB/Swiss-Prot: P55317), TLE3 (HGNC: 11839 NCBI Gene: 7090 Ensembl: ENSG00000140332 OMIM®: 600190 UniProtKB/Swiss-Prot: Q04726) or combinations thereof.

[0036] Definitions

[0037] Unless otherwise defined, all terms (including technical and scientific terms) used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. It will be further understood that terms, such as those defined in commonly used dictionaries, should be interpreted as having a meaning that is consistent with their meaning in the context of the relevant art and will not be interpreted in an idealized or overly formal sense unless expressly so defined herein.

[0038] As used herein, the singular forms “a”, “an” and “the” are intended to include the plural forms as well, unless the context clearly indicates otherwise. Furthermore, to the extent that the terms “including”, “includes”, “having”, “has”, “with”, or variants thereof are used in either

the detailed description and/or the claims, such terms are intended to be inclusive in a manner similar to the term “comprising.”

[0039] The term “about” or “approximately” means within an acceptable error range for the particular value as determined by one of ordinary skill in the art, which will depend in part on how the value is measured or determined, i.e., the limitations of the measurement system. For example, “about” can mean within 1 or more than 1 standard deviation, per the practice in the art. Alternatively, “about” can mean a range of up to 20%, up to 10%, up to 5%, or up to 1% of a given value or range. Alternatively, particularly with respect to biological systems or processes, the term can mean within an order of magnitude within 5-fold, and also within 2-fold, of a value. Where particular values are described in the application and claims, unless otherwise stated the term “about” meaning within an acceptable error range for the particular value should be assumed.

[0040] The terms “aligned”, “alignment”, “mapped” or “aligning”, “mapping” refer to one or more sequences that are identified as a match in terms of the order of their nucleic acid molecules to a known sequence from a reference genome. Such alignment can be done manually or by a computer algorithm, examples including the Efficient Local Alignment of Nucleotide Data (ELAND) computer program distributed as part of the Illumina Genomics Analysts pipeline. The matching of a sequence read in aligning can be a 100% sequence match or less than 100% (non-perfect match).

[0041] The term “cancer” as used herein is meant, a disease, condition, trait, genotype or phenotype characterized by unregulated cell growth or replication as is known in the art; including liver cancer (including hepatocellular carcinoma (HCC)), lung cancer (including non-small cell lung carcinoma), gastric cancer, colorectal cancer, as well as, for example, leukemias, e.g., acute myelogenous leukemia (AML), chronic myelogenous leukemia (CML), acute lymphocytic leukemia (ALL), and chronic lymphocytic leukemia, AIDS related cancers such as Kaposi's sarcoma; breast cancers; bone cancers such as Osteosarcoma, Chondrosarcomas, Ewing's sarcoma, Fibrosarcomas, Giant cell tumors, Adamantinomas, and Chordomas; Brain cancers such as Meningiomas, Glioblastomas, Lower- Grade Astrocytomas, Oligodendrocytomas, Pituitary Tumors, Schwannomas, and Metastatic brain cancers; cancers of the head and neck including various lymphomas such as mantle cell lymphoma, non-Hodgkins lymphoma, adenoma, squamous cell carcinoma, laryngeal carcinoma, gallbladder and bile duct cancers, cancers of the retina such as retinoblastoma, cancers of the esophagus, gastric cancers, multiple myeloma, ovarian cancer,

uterine cancer, thyroid cancer, testicular cancer, endometrial cancer, melanoma, bladder cancer, prostate cancer, pancreatic cancer, sarcomas, Wilms' tumor, cervical cancer, head and neck cancer, skin cancers, nasopharyngeal carcinoma, liposarcoma, epithelial carcinoma, renal cell carcinoma, gallbladder adeno carcinoma, parotid adenocarcinoma, endometrial sarcoma, multidrug resistant cancers; and proliferative diseases and conditions, such as neovascularization associated with tumor angiogenesis.

[0042] The term “cell free nucleic acid,” “cell free DNA,” or “cfDNA” refers to nucleic acid fragments that circulate in an individual's body (e.g., bloodstream) and originate from one or more healthy cells and/or from one or more cancer cells. Additionally, cfDNA may come from other sources such as viruses, fetuses, etc.

[0043] The term “cfDNA sequence coverage” refers to the average number of cfDNA molecules overlapping a specific position.

[0044] The term “circulating tumor DNA” or “ctDNA” refers to nucleic acid fragments that originate from tumor cells or other types of cancer cells, which may be released into an individual's bloodstream as result of biological processes such as apoptosis or necrosis of dying cells or actively released by viable tumor cells.

[0045] As used herein, the terms “comprising,” “comprise” or “comprised,” and variations thereof, in reference to defined or described elements of an item, composition, apparatus, method, process, system, etc. are meant to be inclusive or open ended, permitting additional elements, thereby indicating that the defined or described item, composition, apparatus, method, process, system, etc. includes those specified elements--or, as appropriate, equivalents thereof--and that other elements can be included and still fall within the scope/definition of the defined item, composition, apparatus, method, process, system, etc.

[0046] “Diagnostic” or “diagnosed” means identifying the presence or nature of a pathologic condition. Diagnostic methods differ in their sensitivity and specificity. The “sensitivity” of a diagnostic assay is the percentage of diseased individuals who test positive (percent of “true positives”). Diseased individuals not detected by the assay are “false negatives.” Subjects who are not diseased and who test negative in the assay, are termed “true negatives.” The “specificity” of a diagnostic assay is 1 minus the false positive rate, where the “false positive” rate is defined as the proportion of those without the disease who test positive. While a particular

diagnostic method may not provide a definitive diagnosis of a condition, it suffices if the method provides a positive indication that aids in diagnosis.

[0047] An “effective amount” as used herein, means an amount which provides a therapeutic or prophylactic benefit.

[0048] As used herein, the terms “fragmentation profile,” “fragmentome profile,” “position dependent differences in fragmentation patterns,” and “differences in fragment size and coverage in a position dependent manner across the genome” are equivalent and can be used interchangeably. In some embodiments, determining a cfDNA fragmentation profile in a mammal can be used for identifying a mammal as having cancer. For example, cfDNA fragments obtained from a mammal (e.g., from a sample obtained from a mammal) can be subjected to low coverage whole- genome sequencing, and the sequenced fragments can be mapped to the genome (e.g., in non- overlapping windows) and assessed to determine a cfDNA fragmentation profile. As described herein, a cfDNA fragmentation profile of a mammal having cancer is more heterogeneous (e.g., in fragment lengths) than a cfDNA fragmentation profile of a healthy mammal (e.g., a mammal not having cancer). As such, this disclosure also provides methods and materials for assessing, monitoring, and/or treating mammals (e.g., humans) having, or suspected of having, cancer. In some embodiments, this document provides methods and materials for identifying a mammal as having cancer. For example, a sample (e.g., a blood sample) obtained from a mammal can be assessed to determine the presence and, optionally, the tissue of origin of the cancer in the mammal based, at least in part, on the cfDNA fragmentation profile of the mammal. In some embodiments, methods and materials for monitoring a mammal as having cancer are provided. For example, a sample (e.g., a blood sample) obtained from a mammal can be assessed to determine the presence of the cancer in the mammal based, at least in part, on the cfDNA fragmentation profile of the mammal. In some embodiments, methods and materials for identifying a mammal as having cancer and administering one or more cancer treatments to the mammal to treat the mammal are provided. For example, a sample (e.g., a blood sample) obtained from a mammal can be assessed to determine if the mammal has cancer based, at least in part, on the cfDNA fragmentation profile of the mammal, and one or more cancer treatments can be administered to the mammal.

[0049] The term “genomic nucleic acid,” or “genomic DNA,” refers to nucleic acid including chromosomal DNA that originates from one or more healthy (e.g., non-tumor) cells or

tumor cells. In various embodiments, genomic DNA can be extracted from a cell derived from a blood cell lineage, such as a white blood cell (WBC).

[0050] “Optional” or “optionally” means that the subsequently described event or circumstance can or cannot occur, and that the description includes instances where the event or circumstance occurs and instances where it does not.

[0051] As used in this specification and the appended claims, the term “or” is generally employed in its sense including “and/or” unless the content clearly dictates otherwise.

[0052] “Parenteral” administration of an immunogenic composition includes, e.g., subcutaneous (s.c.), intravenous (i.v.), intramuscular (i.m.), or intrasternal injection, or infusion techniques.

[0053] The terms “patient” or “individual” or “subject” are used interchangeably herein, and refers to a mammalian subject to be treated, with human patients being preferred. In some embodiments, the methods of the invention find use in experimental animals, in veterinary application, and in the development of animal models for disease, including, but not limited to, rodents including mice, rats, and hamsters, and primates.

[0054] The term “reference genome” as used herein may refer to a digital or previously identified nucleic acid sequence database, assembled as a representative example of a species or subject. Reference genomes may be assembled from the nucleic acid sequences from multiple subjects, sample or organisms and does not necessarily represent the nucleic acid makeup of a single person. Reference genomes may be used to for mapping of sequencing reads from a sample to chromosomal positions. For example, a reference genome used for human subjects as well as many other organisms is found at the National Center for Biotechnology Information at ncbi.nlm.nih.gov.

[0055] The term “read segment” or “read” refers to any nucleotide sequences including sequence reads obtained from an individual and/or nucleotide sequences derived from the initial sequence read from a sample obtained from an individual.

[0056] The terms “sample,” “patient sample,” “biological sample,” and the like, encompass a variety of sample types obtained from a patient, individual, or subject and can be used in a diagnostic, prognostic and/or monitoring assay. The patient sample may be obtained from a healthy subject, a diseased patient, or a patient with lung cancer. In certain embodiments, a sample that is “provided” can be obtained by the person (or machine) conducting the assay, or it

can have been obtained by another, and transferred to the person (or machine) carrying out the assay. Moreover, a sample obtained from a patient can be divided and only a portion may be used for diagnosis. Further, the sample, or a portion thereof, can be stored under conditions to maintain sample for later analysis. The definition specifically encompasses blood and other liquid samples of biological origin (including, but not limited to, peripheral blood, serum, plasma, cord blood, amniotic fluid, cerebrospinal fluid, urine, saliva, stool and synovial fluid), solid tissue samples such as a biopsy specimen or tissue cultures or cells derived therefrom and the progeny thereof. In certain embodiment, a sample comprises cerebrospinal fluid. In a specific embodiment, a sample comprises a blood sample. In another embodiment, a sample comprises a plasma sample. In yet another embodiment, a serum sample is used. The definition of “sample” also includes samples that have been manipulated in any way after their procurement, such as by centrifugation, filtration, precipitation, dialysis, chromatography, treatment with reagents, washed, or enriched for certain cell populations. The terms further encompass a clinical sample, and also include cells in culture, cell supernatants, tissue samples, organs, and the like. Samples may also comprise fresh-frozen and/or formalin-fixed, paraffin-embedded tissue blocks, such as blocks prepared from clinical or pathological biopsies, prepared for pathological analysis or study by immunohistochemistry.

[0057] The term “sequence reads” refers to nucleotide sequences read from a sample obtained from an individual. Sequence reads can be obtained through various methods known in the art.

[0058] As defined herein, a “therapeutically effective” amount of a compound or agent (i.e., an effective dosage) means an amount sufficient to produce a therapeutically (e.g., clinically) desirable result. The compositions can be administered from one or more times per day to one or more times per week, including once every other day. The skilled artisan will appreciate that certain factors can influence the dosage and timing required to effectively treat a subject, including but not limited to the severity of the disease or disorder, previous treatments, the general health and/or age of the subject, and other diseases present. Moreover, treatment of a subject with a therapeutically effective amount of the compounds of the invention can include a single treatment or a series of treatments.

[0059] As used herein, the terms “treat,” “treating,” “treatment,” and the like refer to reducing or ameliorating a disorder and/or symptoms associated therewith. It will be appreciated

that, although not precluded, treating a disorder or condition does not require that the disorder, condition or symptoms associated therewith be completely eliminated.

[0060] Genes: All genes, gene names, and gene products disclosed herein are intended to correspond to homologs from any species for which the compositions and methods disclosed herein are applicable. It is understood that when a gene or gene product from a particular species is disclosed, this disclosure is intended to be exemplary only, and is not to be interpreted as a limitation unless the context in which it appears clearly indicates. Thus, for example, for the genes or gene products disclosed herein, are intended to encompass homologous and/or orthologous genes and gene products from other species.

[0061] Ranges: throughout this disclosure, various aspects of the invention can be presented in a range format. It should be understood that the description in range format is merely for convenience and brevity and should not be construed as an inflexible limitation on the scope of the invention. Accordingly, the description of a range should be considered to have specifically disclosed all the possible subranges as well as individual numerical values within that range. For example, description of a range such as from 1 to 6 should be considered to have specifically disclosed subranges such as from 1 to 3, from 1 to 4, from 1 to 5, from 2 to 4, from 2 to 6, from 3 to 6 etc., as well as individual numbers within that range, for example, 1, 2, 2.7, 3, 4, 5, 5.3, and 6. This applies regardless of the breadth of the range.

[0062] Any compositions or methods provided herein can be combined with one or more of any of the other compositions and methods provided herein.

BRIEF DESCRIPTION OF THE FIGURES

[0063] **FIGS. 1A-1B** show a schematic of patients and an embodiment of the liquid biopsy approach. **FIG. 1A:** Patients with liquid biopsy draws at baseline (BL) and after treatment with immunotherapy are evaluated for cell-free DNA fragmentation patterns. Individuals responding to immunotherapy demonstrate decreased tumor derived DNA molecules in the blood compared to patients that do not respond. The changes in DELFI-TF are monitored per patient and categorized by molecular observations. These decisions may provide the option the adjustment of therapeutic modality for patients identified as not responding. **FIG. 1B:** Cohort characteristics for all 50 individuals. Wk, week; ICI, immune checkpoint inhibitor; OS, overall survival; PFS, progression-free survival; DL, dose level.

[0064] **FIGS. 2A-2D** demonstrate that the genome-wide fragmentation profiles reflect responses during immune checkpoint inhibition. **FIG. 2A:** cfDNA fragmentation profiles of all 50 patients across different cancer types at the baseline blood collection reveal genome-wide alterations (blue) evident by low correlation to profiles of non-cancer individuals. Patients are stratified by RECISTv1.1 clinical evaluation post-immunotherapy week 10 indicating fragmentation profiles of non-responders (n=20) and responders (n=7). **FIG. 2B:** Heatmap displaying genome-wide size and coverage fragment feature complexity for patients at baseline as well as for clinical responders and non-responders as determined by RECISTv1.1. Patients with higher DELFI-TF values revealed higher variability in fragmentation features but were not affected by clinical characteristics. **FIG. 2C:** Distribution of DELFI-TF values across serial blood collection timepoints and different cancer types revealed lower values after ICI therapy (Welch two sample test: Baseline (BL) vs after entinostat, p-value = 0.74; BL vs after ICIs, p-value = 0.0054). **(d)** Analysis of overall survival using the median cut-off threshold at baseline DELFI-Tumor Fraction (log-rank, p = 0.0093). BL, baseline; BOR, best overall response; ECOG, Eastern Cooperative Oncology Group Performance Status Scale.

[0065] **FIGS. 3A-3F** show the response to entinostat and ICI therapy using s RECIST and DELFI-TF. **FIGS. 3A 3B:** DELFI-TF values of individuals who were categorized by RECIST as responders and non-responders. Molecular responders (those with a $\geq 50\%$ reduction in DELFI-TF between baseline and week 10) in (A) are distinguished using a diagonal line. **FIG. 3C:** Time to best response for DELFI-TF was 2.3 months compared to 9.1 months for CT-scan. **FIG. 3D:** Week 10 RECIST detected 7 partial responders and 20 patients experiencing stable or progressive disease (median 6.4 vs. 6.7 months; HR = 1.7; 95% CI 0.54-5.3; p = 0.35). **FIG. 3E:** Categorizing patients using DELFI-TF molecular response at week 10 revealed 12 molecular responders and 15 non-responders. Molecular responses detected by DELFI-TF had improved progression-free survival 25.8 months vs non-responders 4.5 months (log-rank p = 0.0014; HR = 5.5; 95% CI 1.7-17; p = 0.004). **FIG. 3F:** At week 10 after immunotherapy 16 patients were considered to have stable disease of which 6 of 10 were detected to have a molecular response (Log-rank p = 0.012).

[0066] **FIGS. 4A-4D** demonstrate the fragmentome chromatin features during immune checkpoint inhibition and for noninvasive subtyping patients with breast cancer. **FIG. 4A:** Breast cancer subset analysis for chromosome 22 displaying A/B compartments evaluating responses to immunotherapy. A TCGA based tissue breast cancer signal (Track 1), extracted breast plasma

component (Track 2) and fragmentation profile for breast cancer non-responder (Track 3) demonstrate a high odds concordance. Concordance between the baseline plasma fragmentation profile (Track 4) to plasma in responders (Track 5) demonstrates a shift to reflect non-cancer (Track 6) as well as lymphoblastoid HiC compartments (track 7), also indicated by change in log odds for all chromosomes (right panel). Breast cancer patients clinically classified as non-responders (track 3) had plasma A/B compartments indicative of increased log odds concordance with the tissue TCGA BRCA (track 1) chromatin. **FIG.4B:** Schema for determining **(FIG. 4C)** top transcription factor binding sites that differentiate HR+ from TNBC with multiple ESR1 hits along with other transcription factors detected for subtyping. Statistically significant decreases in log2FC indicate lower mean coverage among TNBC patients (Bonferroni adjusted $p \leq 0.03$ labeled). **FIG.4D:** Overall performance of subtyping breast cancer subtypes HR+ vs TNBC using the transcription factor binding sites of ESR1, AR, GATA3, FOXA1, and TLE3.

[0067] **FIGS. 5A-5H** demonstrate combining counts of blood cells with DELFI-TF to predict clinical response. **FIG. 5A:** RECISTv1.1 and **(FIG. 5B)** molecular categorization of non-responders compared to responders for neutrophil to lymphocyte ratios demonstrate an increase after immunotherapy (Wilcoxon, RECISTv1.1 $p=0.032$, Molecular $p=0.009$). Further RECISTv1.1 and molecular categorization of non-responders and responders for monocytes **(FIGS. 5C, 5D)** and basophils **(FIGS. 5E, 5F)** reveal lower counts after immunotherapy (monocyte: Wilcoxon, RECISTv1.1 p -value = 0.0085, molecular p -value = 0.0071; basophil: Wilcoxon, RECISTv1.1 p -value = 0.02, molecular p -value = 0.034). **FIGS. 5G, 5H:** Combining NLR plus peripheral blood counts of monocytes and basophils after ICIs with DELFI-TF, DELFI-multi increased the ability to detect durable clinical benefit.

[0068] **FIG. 6** shows graphically reduction in DELFI-TF scores after ICI therapy for individuals evaluated across all time points. DELFI-TF scores for 27 individuals at baseline (week 0), after entinostat treatment (week 2) and after ICI therapy (week 10).

[0069] **FIG. 7** shows longitudinal measures of DELFI-TF scores and overall survival for patients evaluated across all time points. DELFI-TF values at baseline collection (week 0), after entinostat treatment (week 2), and after ICI therapy (week 10). Tumor type and RECIST categorization is indicated above each case, with the overall survival indicated at right.

[0070] **FIG. 8** shows comparison of response using DELFI-TF and compared to RECIST week 10 categorization. Tumor lesion changes by RECIST between baseline and week 10 compared to DELFI-TF revealed a significant correlation (Pearson's correlation, $R=0.7$, $p<0.001$).

[0071] **FIG. 9** shows genome-wide fragmentation profiles are markedly different between molecular responders and non-responders during entinostat and ICI therapy. DELFI-TF at week 10 categorized patients who did ($n=12$) or did not have a molecular response ($n = 15$) during therapy. Molecular responders demonstrated highly uniform genome-wide fragmentation, while non-responders displayed aberrant profiles. Genome-wide alterations are indicated by the correlation to a non-cancer median, with fragmentation profiles that are less correlated shaded in dark blue compared to those with a high correlation colored in light grey.

[0072] **FIG. 10** shows the breast cancer subtype fragmentation profiles at baseline and then stratified by molecular response at week 10. Overall survival for patients assessed by RECIST or DELFI-TF at week 10. Evaluation of response to therapy between baseline and after entinostat and ICI at week 10 using RECIST (a) or DELFI-TF molecular response (b). A significant separation in OS was observed for molecular but not clinical responders.

[0073] **FIGS. 11A-11B** shows low DELFI-TF scores identified in patients with durable stable disease. Assessment of DELFI-TF scores and length of stable disease (> 6 months or <6 months) in patients determined to have stable disease at week 10 or as best overall response. (**FIG. 11A**) At week 10, patients with durable clinical response had lower DELFI-TF scores compared to those who did not have durable stable disease ($n = 16$; Welch two sample t-test, $p\text{-value} = 0.01$). (**FIG. 11B**) Similarly, among patients who were classified as having stable disease as a best overall response, patients with durable clinical response had lower DELFI-TF scores compared to those who had non-durable stable disease ($n = 11$; Welch two sample t-test, $p\text{-value} < 0.001$).

[0074] **FIGS. 12A-12E** show analysis of cfDNA fragmentomes in separate validation cohort of lung cancer patients. **FIG 12A**: Individuals with no previous treatment examined at baseline (BL) ($n=21$) demonstrated genome-wide alterations in cfDNA fragmentation. Patients analyzed after immune checkpoint inhibition (ICI) at 6 ($n=15$) or 11 ($n=11$) weeks had fragmentation profiles that were closer to individuals without cancer. Blue and grey colors indicate correlation to median fragmentation profiles of a control group of individuals without cancer ($n = 145$). **FIG. 12B**: DELFI-TF scores evaluated for patients at baseline or at different time points after initiation of ICI and/or chemotherapy. **FIG. 12C**: DELFI-TF scores at baseline and week 11

for individuals who had a RECIST clinical CR/PR revealed a significant reduction 11 weeks after initiation of therapy (p-value = 0.016) while (**FIG. 12D**) patients with RECIST clinical SD/PD did not show a significant reduction. (**FIG. 12E**) DELFI-TF correctly categorized patients with RECIST clinical PR/CR compared to those who were SD/PD. CR, complete response; PR, partial response; SD, stable disease; PD, progressive disease.

[0075] **FIG. 13** shows graphically DELFI-TF scores in longitudinal samples in external validation cohort.

[0076] **FIGS. 14A-D** show cfDNA fragmentation during entinostat and ICI therapy for **patients** with breast cancer subtypes. **FIG. 14A**: RECIST classification of response at week 10 for Hormone positive (HR+) and triple-negative breast cancer (TNBC) breast cancer patients. **FIG. 14B**: Molecular categorization at week 10 of the same HR+ and TNBC patients. Comparison of patients with HR+ or TNBC did not reveal a difference in progression free survival. **FIG. 14C**: Genome-wide fragmentation profiles of breast cancer patients with either HR+ (n=14) or TNBC (n=14) at baseline demonstrated genome-wide aberrations that were not correlated to noncancer profiles. **FIG. 14D**: At week 10, patients HR+ or TNBC who had a molecular response had a higher correlation to that of a healthy profile compared to profiles in those without a molecular response.

[0077] **FIG. 15** shows comparison of cfDNA fragmentation at transcription factor binding sites for breast cancer subtypes during entinostat and ICI therapy. cfDNA coverage at five transcription factor binding sites for patients with HR+ and TNBC breast cancers at baseline or at week 10 for molecular non-responders or responders after therapy.

[0078] **FIGS. 16A-B** show absolute blood cell counts and DELFI-TF scores at baseline and after entinostat and ICI therapy. Absolute blood cell counts for basophils, eosinophils, monocytes, neutrophils, and lymphocytes with each data point representing an individual. Blood cell types do not have a noticeable difference in absolute counts between (**FIG. 16A**) baseline and (**FIG. 16B**) after immunotherapy.

DETAILED DESCRIPTION

[0079] A cell-free DNA fragmentome approach (DELFI-TF) is provided for evaluating circulating tumor fraction during therapy. DELFI-TF was used to assess 50 patients with advanced solid tumors, including 28 with breast cancer, enrolled in a multicenter phase 1 clinical trial combining entinostat and nivolumab with or without ipilimumab (NCT02453620). DELFI-TF molecular responders and non-responders at week 10 experienced a median progression-free survival (PFS) that was 25.8 and 4.5 months, respectively (HR=5.5, 95% CI=1.7-17, p=0.004) and improved overall survival (OS) (HR=3.4, 95% CI=1-12, p=0.05). In contrast, imaging analyses using RECIST at 10 weeks was not associated with PFS or OS differences for patients who had a complete or partial response compared to stable or progressive disease. Among the subgroup of radiographic non-responders with stable disease, DELFI-TF distinguished the subset of molecular responders from non-responders with a PFS of 25.8 and 5.2 months, respectively (p=0.012). Cell-free DNA fragmentation profiles of patients with advanced cancer reflected cancer-specific chromatin features and could be used to non-invasively distinguish patients with hormone receptor positive or triple negative breast cancer subtypes. DELFITF could be combined with counts of immune cells for improved detection of molecular response. The approach was validated in an independent cohort of lung cancer patients (n=21) treated with immune checkpoint inhibitors and chemotherapy. These analyses demonstrated use of cfDNA fragmentomics for monitoring therapeutic response for cancer patients, including response to immunotherapy in cancer patients.

[0080] DNA Evaluation of Fragments for early Interception (DELFI)

[0081] DNA Evaluation of Fragments for early Interception (DELFI) (Cristiano S, Leal A, Phallen J, *et al.* Genome-wide cell-free DNA fragmentation in patients with cancer. *Nature* 2019;570:385-9 incorporated herein in its entirety), was used to evaluate genome-wide fragmentation patterns of cfDNA of 236 patients with breast, colorectal, lung, ovarian, pancreatic, gastric, or bile duct cancers as well as 245 healthy individuals. These analyses revealed that cfDNA profiles of healthy individuals reflected nucleosomal fragmentation patterns of white blood cells, while patients with cancer had altered fragmentation profiles. DELFI had sensitivities of detection ranging from 57% to >99% among the seven cancer types at 98% specificity and identified the tissue of origin of the cancers to a limited number of sites in 75% of embodiments. Assessing cfDNA (e.g., using DELFI) provide a screening approach for early detection of cancer, which can increase the chance for successful treatment of a patient having cancer. Assessing cfDNA (e.g.,

using DELFI) can also provide an approach for monitoring cancer, which can increase the chance for successful treatment and improved outcome of a patient having cancer. In addition, a cfDNA fragmentation profile can be obtained from limited amounts of cfDNA and using inexpensive reagents and/or instruments.

[0082] *cfDNA Fragmentome Profiles*

[0083] A cfDNA fragmentome profile can include one or more cfDNA fragmentation patterns. A cfDNA fragmentation pattern can include any appropriate cfDNA fragmentation pattern. Examples of cfDNA fragmentation patterns include, without limitation, median fragment size, fragment size distribution, ratio of small cfDNA fragments to large cfDNA fragments, and the coverage of cfDNA fragments. In some embodiments, a cfDNA fragmentation pattern includes two or more (e.g., two, three, or four) of median fragment size, fragment size distribution, ratio of small cfDNA fragments to large cfDNA fragments, and the coverage of cfDNA fragments. In some embodiments, cfDNA fragmentome profile can be a genome-wide cfDNA profile (e.g., a genome-wide cfDNA profile in windows across the genome). In some embodiments, cfDNA fragmentome profile can be a targeted region profile. A targeted region can be any appropriate portion of the genome (e.g., a chromosomal region). Examples of chromosomal regions for which a cfDNA fragmentome profile can be determined as described herein include, without limitation, a portion of a chromosome (e.g., a portion of 2q, 4p, 5p, 6q, 7p, 8q, 9q, 10q, 11q, 12q, and/or 14q) and a chromosomal arm (e.g., a chromosomal arm of 8q,13q, 11q, and/or 3p). In some embodiments, a cfDNA fragmentome profile can include two or more targeted region profiles.

[0084] In some embodiments, a cfDNA fragmentome profile can be used to identify changes (e.g., alterations) in cfDNA fragment lengths. An alteration can be a genome-wide alteration or an alteration in one or more targeted regions/loci. A target region can be any region containing one or more cancer-specific alterations. In some embodiments, a cfDNA fragmentome profile can be used to identify (e.g., simultaneously identify) from about 10 alterations to about 500 alterations (e.g., from about 25 to about 500, from about 50 to about 500, from about 100 to about 500, from about 200 to about 500, from about 300 to about 500, from about 10 to about 400, from about 10 to about 300, from about 10 to about 200, from about 10 to about 100, from about 10 to about 50, from about 20 to about 400, from about 30 to about 300, from about 40 to about 200, from about 50 to about 100, from about 20 to about 100, from about 25 to about 75, from about 50 to about 250, or from about 100 to about 200, alterations).

[0085] A cfDNA fragmentome profile can be obtained using any appropriate method. In some embodiments, cfDNA from a mammal (e.g., a mammal having, or suspected of having, cancer) can be processed into sequencing libraries which can be subjected to whole genome sequencing (e.g., low-coverage whole genome sequencing), mapped to the genome, and analyzed to determine cfDNA fragment lengths. Mapped sequences can be analyzed in non-overlapping windows covering the genome. Windows can be any appropriate size. For example, windows can be from thousands to millions of bases in length. As one non-limiting example, a window can be about 5 megabases (Mb) long. Any appropriate number of windows can be mapped. For example, tens to thousands of windows can be mapped in the genome. For example, hundreds to thousands of windows can be mapped in the genome. A cfDNA fragmentome profile can be determined within each window.

[0086] In some embodiments, methods and materials described herein also can include machine learning. For example, machine learning can be used for identifying mutation frequencies, altered fragmentome profile (e.g., using coverage of cfDNA fragments, fragment size of cfDNA fragments, coverage of chromosomes, and mtDNA).

[0087] In some embodiments, determining a cfDNA fragmentome profile in a mammal can be used for a method of determining responsiveness to a cancer therapy in a subject. For example, cfDNA fragments obtained from a mammal (e.g., from a sample obtained from a mammal) can be subjected to low coverage whole-genome sequencing, and the sequenced fragments can be mapped to the genome and assessed to determine a cfDNA fragmentome profile. As described herein, a cfDNA fragmentome profile of a mammal having cancer is more heterogeneous (e.g., in fragment lengths) than a cfDNA fragmentome profile of a healthy mammal (e.g., a mammal not having cancer). As such, also provided are methods and materials for assessing, monitoring, and/or treating mammals (e.g., humans) having, or suspected of having, cancer. In some embodiments, methods and materials are provided for identifying a mammal as having cancer. For example, a sample (e.g., a blood sample) obtained from a mammal can be assessed to determine the presence and, optionally, the tissue of origin of the cancer in the mammal based, at least in part, on the cfDNA fragmentome profile of the mammal. In some embodiments, methods and materials are provided for monitoring a mammal as having cancer. For example, a sample (e.g., a plasma sample) obtained from a mammal can be assessed to determine the presence of the cancer in the mammal based, at least in part, on the cfDNA fragmentome profile of the

mammal. In some embodiments, methods and materials are provided for identifying a mammal as having cancer and administering one or more cancer treatments to the mammal to treat the mammal. For example, a sample (e.g., a blood sample) obtained from a mammal can be assessed to determine if the mammal has cancer based, at least in part, on the cfDNA fragmentome profile of the mammal, and one or more cancer treatments can be administered to the mammal.

[0088] In some embodiments, a cfDNA fragmentome profile can be used to detect tumor-derived DNA. For example, a cfDNA fragmentome profile can be used to detect tumor-derived DNA by comparing a cfDNA fragmentome profile of a mammal having, or suspected of having, cancer to a reference cfDNA fragmentome profile (e.g., a cfDNA fragmentome profile of a healthy mammal and/or a nucleosomal DNA fragmentome profile of healthy cells from the mammal having, or suspected of having, cancer). In some embodiments, a reference cfDNA fragmentome profile is a previously generated profile from a healthy mammal. For example, methods provided herein can be used to determine a reference cfDNA fragmentome profile in a healthy mammal, and that reference cfDNA fragmentome profile can be stored (e.g., in a computer or other electronic storage medium) for future comparison to a test cfDNA fragmentome profile in mammal having, or suspected of having, cancer. In some embodiments, a reference cfDNA fragmentome profile (e.g., a stored cfDNA fragmentome profile) of a healthy mammal is determined over the whole genome. In some embodiments, a reference cfDNA fragmentome profile (e.g., a stored cfDNA fragmentome profile) of a healthy mammal is determined over a subgenomic interval.

[0089] In some embodiments, a cfDNA fragmentome profile can be used to identify a mammal (e.g., a human) as having cancer (e.g., a liver cancer, a colorectal cancer, a lung cancer, a breast cancer, a gastric cancer, a pancreatic cancer, a bile duct cancer, and/or an ovarian cancer).

[0090] In some embodiments, the methods herein differentiate between breast cancer subtypes. Breast cancer is a genetically and clinically heterogeneous disease with multiple subtypes. The classification of these subtypes has evolved over the years. The most common and widely accepted classification of breast cancer is from an immunohistochemical perspective, based on the expression of the following hormone receptors: estrogen (ER), progesterone (PR) and human epidermal growth factor (HER2). Accordingly, the following four subtypes of breast cancer are widely recognized: luminal A, luminal B, HER2-positive, and triple-negative. With the recent advances in cancer research, and an increased molecular understanding of breast cancer, the current clinical model for classification of breast cancer may benefit from the addition of several

molecular markers such as miRNAs (let-7, miR-155, miR-150, miR-153) and mutations (p53, BRCA 1 and 2 genes).

[0091] The table below shows the characteristics of subtypes of breast cancer

	Luminal A	Luminal B	HER2	TNBC
Frequency (%)	50	15	20	15
ER	Yes	Yes	Some cases	No
PR	Yes	Some cases	Some cases	No
HER	No	No	Yes	No
miRNAs	<i>Let-7f, Let-7c, miR-10, miR-29a, miR-181a, miR-223 and miR-652</i>	<i>miR155, miR-93, miR-18a, miR-135b, miR-718, miR-4516, miR-210, and miR-125b-5p</i>	<i>miR-150 and miR-142-3p</i>	<i>miR-153, miR-10b, miR-26a, and miR146a</i>
Ki67	Some cases	Some cases	High	High
Mutations	No	BRCA2	p53	p53 and BRCA1
Prognosis	Good	Middle	Middle/Bad	Bad
Therapy	Hormonal	Hormonal/ Chemo	Hormonal/Chemo/ Herceptin	Chemo/ Experimental

See, Orrantia-Borunda E, *et al.* Subtypes of Breast Cancer. In: Mayrovitz HN, editor. Breast Cancer [Internet]. Brisbane (AU): Exon Publications; 2022 Aug 6. Chapter 3. Available from: ncbi.nlm.nih.gov/books/NBK583808/ doi: 10.36255/exon-publications-breast-cancer-subtypes.

[0092] **Luminal A subtype:** Luminal A tumors are characterized by the presence of ER and/or PR and the absence of HER2 and have a low expression of cell proliferation marker Ki-67 (less than 20%). Clinically they are low grade, slow growing, and have the best prognosis with less incidence of relapse and higher survival rate. These carcinomas present a high response rate to hormone therapy (tamoxifen or aromatase inhibitors), and a more limited benefit to chemotherapy. For this reason, according to the European Society for Medical Oncology (ESMO) and National Comprehensive Cancer Network from USA (NCCN) Guidelines, the use of genetic platforms is recommended in this group to establish which patients would benefit from adjuvant chemotherapy treatment based on the risk of relapse and survival rate. Relapse is more frequent at the bone level, with a lower rate of visceral and central nervous system (CNS) relapses. Likewise, they have a longer survival in case of relapse.

[0093] **Luminal B subtype:** Luminal B tumors are of higher grade and worse prognosis compared to Luminal A. They are ER positive and can be PR negative and have a high expression of Ki67 (greater than 20%). They are generally of intermediate/high histologic grade. These tumors may benefit from hormonal therapy along with chemotherapy. The elevated Ki67 makes them grow faster than luminal A and worse prognosis. It constitutes 10–20% of luminal tumors. It has a moderately low expression of estrogen receptors, and increased expression of proliferation and cell cycle genes. It represents the group of luminal tumors with the worst prognosis. They benefit from hormone therapy and in a higher percentage from chemotherapy compared to the previous group. Although bone recurrence is frequent, they have a higher rate of visceral recurrence, and survival from diagnosis to relapse is lower.

[0094] **HER2 Subtype:** The HER2-positive group constitutes 10–15% of breast cancers and is characterized by high HER2 expression with absence of ER and PR. They grow faster than the luminal ones and the prognosis has improved after the introduction of HER2-targeted therapies. The HER2-positive subtype is more aggressive and fast-growing. Within this, two subgroups can be distinguished: luminal HER2 (E+, PR+, HER2+ and Ki-67:15–30%) and HER2-enriched (HER2+, E-, PR-, Ki-67>30%). They have a worse prognosis compared to luminal tumors, and require specific drugs directed against the HER2/neu protein, including trastuzumab, trastuzumab combined with emtansine (T-DM1), pertuzumab, and tyrosine kinase inhibitors such as lapatinib and neratinib, among others, in addition to surgery and treatment with precise chemotherapy. They have a high response rate to chemotherapy schemes. Bone localization is the most common site

for disseminated disease, and visceral relapses are also more frequent in this subgroup compared to the previous group.

[0095] TNBC subtype: Triple-negative breast cancer is ER-negative, PR-negative, and HER2-negative. They constitute about 20% of all breast cancers. It is most common among women under 40 years of age, and in African-American women. The TNBC subtype is further classified into several additional subgroups including basal-like (BL1 and BL2), claudin-low, mesenchymal (MES), luminal androgen receptor (LAR), and immunomodulatory (IM), the first two being the most frequent with 50–70% and 20–30% of cases. Moreover, each of these has unique clinical outcomes, phenotypes, and pharmacological sensitivities. TNBC presents an aggressive behavior and 80% of breast cancer tumors (tumor suppressor gene BRCA1 and BRCA2) belong to this group. The risk of developing TNBC varies with genetics, race, age, overweight and obesity, breastfeeding patterns, and parity. TNBC is characterized by its aggressiveness, early relapse, and a greater tendency to present in advanced stages. It presents a high proliferation rate, alteration in DNA repair genes and increased genomic instability. Histologically, it is a poorly differentiated, highly proliferative, heterogeneous neoplasm, including subsets of variable prognosis. Immunohistochemically, they are subdivided into basal and non-basal TNBC; the former characterized by expression of cytokeratins (CK)5/6 and human epidermal growth factor receptor type 1 (EGFR1), while the non-basal do not express CK5/6 cytokeratins.

[0096] A cfDNA fragmentome profile can include a cfDNA fragment size pattern. cfDNA fragments can be any appropriate size. For example, cfDNA fragment can be from about 50 base pairs (bp) to about 400 bp in length.

[0097] A cfDNA fragmentome profile can include a cfDNA fragment size distribution. As described herein, a mammal having cancer can have a cfDNA size distribution that is more variable than a cfDNA fragment size distribution in a healthy mammal. In some embodiments, a size distribution can be within a targeted region. A healthy mammal (e.g., a mammal not having cancer) can have a targeted region cfDNA fragment size distribution of about 1 or less than about 1. In some embodiments, a mammal having cancer can have a targeted region cfDNA fragment size distribution that is longer (e.g., 10, 15, 20, 25, 30, 35, 40, 45, 50 or more bp longer, or any number of base pairs between these numbers) than a targeted region cfDNA fragment size distribution in a healthy mammal. In some embodiments, a mammal having cancer can have a targeted region cfDNA fragment size distribution that is shorter (e.g., 10, 15, 20, 25, 30, 35, 40, 45, 50 or more bp

shorter, or any number of base pairs between these numbers) than a targeted region cfDNA fragment size distribution in a healthy mammal. In some embodiments, a size distribution can be a genome-wide size distribution. A healthy mammal (e.g., a mammal not having cancer) can have very similar distributions of short and long cfDNA fragments genome wide. In some embodiments, a mammal having cancer can have, genome-wide, one or more alterations (e.g., increases and decreases) in cfDNA fragment sizes. The one or more alterations can be any appropriate chromosomal region of the genome. For example, an alteration can be in a portion of a chromosome. Examples of portions of chromosomes that can contain one or more alterations in cfDNA fragment sizes include, without limitation, portions of 2q, 4p, 5p, 6q, 7p, 8q, 9q, 10q, 11q, 12q, and 14q. For example, an alteration can be across a chromosome arm (e.g., an entire chromosome arm).

[0098] A cfDNA fragmentome profile can include a ratio of small cfDNA fragments to large cfDNA fragments and a correlation of fragment ratios to reference fragment ratios. As used herein, with respect to ratios of small cfDNA fragments to large cfDNA fragments, a small cfDNA fragment can be from about 100 bp in length to about 150 bp in length. As used herein, with respect to ratios of small cfDNA fragments to large cfDNA fragments, a large cfDNA fragment can be from about 151 bp in length to 220 bp in length. A mammal having cancer can have a correlation of fragment ratios (e.g., a correlation of cfDNA fragment ratios to reference DNA fragment ratios such as DNA fragment ratios from one or more healthy mammals) that is lower (e.g., 2-fold lower, 3-fold lower, 4-fold lower, 5-fold lower, 6-fold lower, 7-fold lower, 8-fold lower, 9-fold lower, 10-fold lower, or more) than in a healthy mammal. A healthy mammal (e.g., a mammal not having cancer) can have a correlation of fragment ratios (e.g., a correlation of cfDNA fragment ratios to reference DNA fragment ratios such as DNA fragment ratios from one or more healthy mammals) of about 1 (e.g., about 0.96). In some embodiments, a mammal having cancer can have a correlation of fragment ratios (e.g., a correlation of cfDNA fragment ratios to reference DNA fragment ratios such as DNA fragment ratios from one or more healthy mammals) that is, on average, lower than a correlation of fragment ratios (e.g., a correlation of cfDNA fragment ratios to reference DNA fragment ratios such as DNA fragment ratios from one or more healthy mammals) in a healthy mammal.

[0099] A cfDNA fragmentome profile can include coverage of all fragments. Coverage of all fragments can include windows (e.g., non-overlapping windows) of coverage. In some

embodiments, coverage of all fragments can include windows of small fragments (e.g., fragments from about 100 bp to about 150 bp in length). In some embodiments, coverage of all fragments can include windows of large fragments (e.g., fragments from about 151 bp to about 220 bp in length).

[00100] In certain embodiments, a cfDNA fragmentome profile can be used to identify the molecular origins of cfDNA in patients and identify genomic and chromatin features associated with fragmentation changes.

[00101] In some embodiments, a cfDNA fragmentome profile can be used to identify the tissue of origin of a cancer (e.g., a liver cancer, a colorectal cancer, a lung cancer, a breast cancer, a gastric cancer, a pancreatic cancer, a bile duct cancer, or an ovarian cancer). For example, a cfDNA fragmentome profile can be used to identify a localized cancer. When a cfDNA fragmentome profile includes a targeted region profile, one or more alterations described herein can be used to identify the tissue of origin of a cancer. In some embodiments, one or more alterations in chromosomal regions can be used to identify the tissue of origin of a cancer.

[00102] A cfDNA fragmentome profile can be obtained using any appropriate method. In some embodiments, cfDNA from a mammal (e.g., a mammal having, or suspected of having, cancer) can be processed into sequencing libraries which can be subjected to whole genome sequencing (e.g., low-coverage whole genome sequencing), mapped to the genome, and analyzed to determine cfDNA fragment lengths. Mapped sequences can be analyzed in non-overlapping windows covering the genome. Windows can be any appropriate size. For example, windows can be from thousands to millions of bases in length. As one non-limiting example, a window can be about 5 megabases (Mb) long. Any appropriate number of windows can be mapped. For example, tens to thousands of windows can be mapped in the genome. For example, hundreds to thousands of windows can be mapped in the genome. A cfDNA fragmentome profile can be determined within each window.

[00103] In some embodiments, methods and materials described herein also can include machine learning. For example, machine learning can be used for identifying an altered fragmentome profile (e.g., using coverage of cfDNA fragments, fragment size of cfDNA fragments, coverage of chromosomes, and mtDNA). Various machine learning algorithms can be used to analyze the fragmentation profiles. For example, to distinguish healthy from cancer patients using fragmentation profiles, a stochastic gradient boosting model is used (gbm; see, e.g.,

Friedman et al., 2001 Ann Stat 29:1189-1232; and Friedman et al., 2002 Comput Stat Data An 38:367-378). GC-corrected total and short fragment coverage for all 504 bins can be centered and scaled for each sample to have mean 0 and unit standard deviation. Additional features included Z-scores for each of the 39 autosomal arms and mitochondrial representation (log10-transformed proportion of reads mapped to the mitochondria). To estimate the prediction error of this approach, 10-fold cross-validation can be used as described elsewhere (see, e.g., Efron et al., 1997 J Am Stat Assoc 92, 548-560). Feature selection, performed only on the training data in each cross-validation run, removed bins that were highly correlated (correlation > 0.9) or had near zero variance. Stochastic gradient boosted machine learning can be implemented using the R package gbm package. To average over the prediction error from the randomization of patients to folds, the 10-fold cross validation procedure can be repeated.

[00104] In some embodiments, a machine learning model is a neural network (NN). In certain embodiments, the neural network is a convolutional neural network, a recurrent neural network, or a deep learning neural network. In some embodiments, the machine learning model is a random forest, logistic regression, or an unsupervised clustering model. In other embodiments of neural network models, the model can also be a deep neural network (DNN) with multiple locally and fully connected hidden layers, or a high-order neural network (HONN). For DNN, a Restricted Boltzmann Machine (RBM) can be used to pre-train the neural nodes of input and connecting layers. For HONN, a mean-covariance RBM can be used to pre-train the neural nodes of input and connecting layers.

[00105] In certain embodiments, the machine learning model is a Random forest regression mode. This model was initially trained to estimate mutant allele frequency of tumor-specific RAS/BRAF variants measured by digital droplet PCR, using 692 longitudinal plasma samples collected from the CAIRO5 phase III clinical trial. The features of this model included the DELFI score (Cristiano, A. et al., Genome-wide cell-free DNA fragmentation in patients with cancer. Nature 570, 385–389 (2019) incorporated herein by reference in its entirety), DELFI divergence, and mixture model components (J. Carey, et al., Modeling cell-free DNA fragment size densities for non-invasive detection of cancer. J. Clin. Orthod. 39, 3058–3058 (2021)). The DELFI divergence was computed as one-minus the Pearson correlation of each sample to the median fragmentation profile of a reference set comprised of individuals without cancer (16). See also

Iris van't Erve et al., Cancer treatment monitoring using cell-free DNA fragmenomes, Nature Communications volume 15, Article number: 8801 (2024).

[00106] In certain embodiments, a computer system for obtaining access to database files and executing one or more software programs may include a server, a data storage device, a network, and a user interface device. The server may also be a hypervisor-based system executing one or more guest partitions hosting operating systems with modules having server configuration information. In a further embodiment, the system may include a storage controller, or a storage server configured to manage data communications between the data storage device and the server or other components in communication with the network. In an alternative embodiment, the storage controller may be coupled to the network.

[00107] In certain embodiments, the user interface device is referred to broadly and is intended to encompass a suitable processor-based device such as a desktop computer, a laptop computer, a personal digital assistant (PDA) or tablet computer, a smartphone or other mobile communication device having access to the network. In a further embodiment, the user interface device may access the Internet or other wide area or local area network to access a web application or web service hosted by the server and may provide a user interface for enabling a user to enter or receive information. The network may facilitate communications of data between the server and the user interface device. The network may include any type of communications network including, but not limited to, a direct PC-to-PC connection, a local area network (LAN), a wide area network (WAN), a modem-to-modem connection, the Internet, a combination of the above, or any other communications network now known or later developed within the networking arts which permits two or more computers to communicate.

[00108] In certain embodiments, a computer system comprises a central processing unit ("CPU") coupled to the system bus. The CPU may be a general purpose CPU or microprocessor, graphics processing unit ("GPU"), and/or microcontroller. The CPU may execute the various logical instructions according to the present embodiments. The computer system may also include random access memory (RAM), which may be synchronous RAM (SRAM), dynamic RAM (DRAM), synchronous dynamic RAM (SDRAM), or the like. The computer system may utilize RAM to store the various data structures used by a software application. The computer system may also include read only memory (ROM) which may be PROM, EPROM, EEPROM, optical storage, or the like. The ROM may store configuration information for booting the computer

system. The RAM and the ROM hold user and system data, and both the RAM and the ROM may be randomly accessed.

[00109] The computer system may also include an I/O adapter, a communications adapter, a user interface adapter, and a display adapter. The I/O adapter and/or the user interface adapter may, in certain embodiments, enable a user to interact with the computer system. In a further embodiment, the display adapter may display a graphical user interface (GUI) associated with a software or web-based application on a display device, such as a monitor or touch screen.

[00110] The I/O adapter may couple one or more storage devices, such as one or more of a hard drive, a solid state storage device, a flash drive, a compact disc (CD) drive, a floppy disk drive, and a tape drive, to the computer system. The data storage may be a separate server coupled to the computer system through a network connection to the I/O adapter. The communications adapter may be adapted to couple the computer system to the network, which may be one or more of a LAN, WAN, and/or the Internet. The user interface adapter couples user input devices, such as a keyboard, a pointing device, and/or a touch screen to the computer system. The display adapter may be driven by the CPU to control the display on the display device.

[00111] The computer system is provided as an example of one type of computing device that may be adapted to perform the functions of the server and/or the user interface device. For example, any suitable processor-based device may be utilized including, without limitation, personal data assistants (PDAs), tablet computers, smartphones, computer game consoles, and multi-processor servers to implement various embodiments and/or steps the cancer detection models disclosed herein. Moreover, various embodiments of the cancer detection methods of the present disclosure may be implemented on application specific integrated circuits (ASIC), very large scale integrated (VLSI) circuits, or other circuitry. In fact, persons of ordinary skill in the art may utilize any number of suitable structures capable of executing logical operations according to the described embodiments. For example, the computer system may be virtualized for access by multiple users and/or applications.

[00112] Various methods, steps, calculations of parameters disclosed herein if implemented in firmware and/or software, the various functions described above may be stored as one or more instructions or code on a computer-readable medium. Examples include non-transitory computer-readable media encoded with a data structure and computer-readable media encoded with a computer program. Computer-readable media includes physical computer storage media. A

storage medium may be any available medium that can be accessed by a computer. By way of example, and not limitation, such computer-readable media can comprise RAM, ROM, EEPROM, CD-ROM or other optical disk storage, magnetic disk storage or other magnetic storage devices, or any other medium, e.g., cloud based, that can be used to store desired program code in the form of instructions or data structures and that can be accessed by a computer. Disk and disc include compact discs (CD), laser discs, optical discs, digital versatile discs (DVD), floppy disks and blu-ray discs. Generally, disks reproduce data magnetically, and discs reproduce data optically. Combinations of the above should also be included within the scope of computer-readable media.

[00113] In addition to storage on computer-readable medium, instructions and/or data may be provided as signals on transmission media included in a communication apparatus. For example, a communication apparatus may include a transceiver having signals indicative of instructions and data. The instructions and data are configured to cause one or more processors to implement the functions embodied herein.

[00114] Response Evaluation Criteria in Solid Tumors (RECIST)

[00115] Assessment of the change in tumor burden is an important feature of the clinical evaluation of cancer therapeutics: both tumor shrinkage and disease progression are useful endpoints in clinical trials. Response evaluation criteria in solid tumors (RECIST) was introduced by a small international working group in February 2000 to facilitate, improve and standardize the evaluation and the reporting of objective tumor outcomes in early clinical trials investigating new anti-cancer agents. In comparison to earlier response assessment systems, the new criteria gave much more detailed recommendations on how to assess tumor lesions, how to report responses, and also took into account recent developments in medical imaging techniques. RECIST uses a unidimensional measure (the longest diameter) to quantify measurable tumor lesions as opposed to the bidimensional product (longest diameter multiplied by its perpendicular), which was commonly employed by earlier iterations of response criteria (P. Therasse, *et al.*, “RECIST revisited: A review of validation studies on tumour assessment”, *European Journal of Cancer*, Volume 42, Issue 8, 2006, pp. 1031-1039, ISSN 0959-8049, doi.org/10.1016/j.ejca.2006.01.026).

[00116] Response criteria are standardized and can be used at different time points to classify response into the categories of complete response, partial response, stable disease, or disease progression. At the trial level, categorical responses for all patients are summated into image-based trial endpoints. These outcome measures, including objective response rate (ORR)

and progression-free survival (PFS), are characteristics that can be derived from imaging and can be used as surrogates for overall survival (OS).

[00117] Response evaluation criteria in solid tumors (RECIST) forms the basis for PFS determination and defines progressive disease as at least a 20% increase in the sum of diameters of up to 5 target lesions (2 lesions/organ), taking as reference the smallest sum on study and an absolute lesion increase of at least 5 mm or the appearance of new lesions. A complete response is the disappearance of all target lesions, and a partial response (PR) is defined as at least a 30% decrease in the sum of the target lesions. Stable disease is defined as fitting the criteria neither for progressive disease nor a PR (Villaruz LC, Socinski MA. The clinical viewpoint: definitions, limitations of RECIST, practical considerations of measurement. *Clin Cancer Res.* 2013 May 15;19(10):2629-36. doi: 10.1158/1078-0432.CCR-12-2935. PMID: 23669423; PMCID: PMC4844002).

[00118] Methods of Treatment

[00119] When monitoring a mammal having, or suspected of having, cancer as described herein, the monitoring can be before, during, and/or after the course of a cancer treatment. Methods of monitoring provided herein can be used to determine the efficacy of one or more cancer treatments and/or to select a mammal for increased monitoring. In some embodiments, the monitoring can include identifying a cfDNA fragmentome profile as described herein. For example, a cfDNA fragmentome profile can be obtained before administering one or more cancer treatments to a mammal having, or suspected of having, cancer, one or more cancer treatments can be administered to the mammal, and one or more samples can be analyzed during the course of the cancer treatment. In some embodiments, a cfDNA fragmentome profile can change during the course of cancer treatment (e.g., any of the cancer treatments described herein). For example, a cfDNA fragmentome profile indicative that the mammal has cancer can change to a cfDNA fragmentome profile indicative that the mammal does not have cancer. Such a cfDNA fragmentome profile change can indicate that the cancer treatment is working. Conversely, a cfDNA fragmentome profile can remain static (e.g., the same or approximately the same) during the course of cancer treatment (e.g., any of the cancer treatments described herein). Such a static cfDNA fragmentome profile can indicate that the cancer treatment is not working.

[00120] In some embodiments, the monitoring can include conventional techniques capable of monitoring one or more cancer treatments (e.g., the efficacy of one or more cancer treatments).

In some embodiments, a mammal selected for increased monitoring can be administered a diagnostic test (e.g., any of the diagnostic tests disclosed herein) at an increased frequency compared to a mammal that has not been selected for increased monitoring. For example, a mammal selected for increased monitoring can be administered a diagnostic test at a frequency of twice daily, daily, bi-weekly, weekly, bi-monthly, monthly, quarterly, semi-annually, annually, or any at frequency therein. In some embodiments, a mammal selected for increased monitoring can be administered a one or more additional diagnostic tests compared to a mammal that has not been selected for increased monitoring. For example, a mammal selected for increased monitoring can be administered two diagnostic tests, whereas a mammal that has not been selected for increased monitoring is administered only a single diagnostic test (or no diagnostic tests). In some embodiments, a mammal that has been selected for increased monitoring can also be selected for further diagnostic testing. Once the presence of a tumor or a cancer (e.g., a cancer cell) has been identified (e.g., by any of the variety of methods disclosed herein), it may be beneficial for the mammal to undergo both increased monitoring (e.g., to assess the progression of the tumor or cancer in the mammal and/or to assess the development of one or more cancer biomarkers such as mutations), and further diagnostic testing (e.g., to determine the size and/or exact location (e.g., tissue of origin) of the tumor or the cancer). In some embodiments, one or more cancer treatments can be administered to the mammal that is selected for increased monitoring after a cancer biomarker is detected and/or determining a cfDNA fragmentome profile of the subject's sequenced DNA fragments by evaluating one or more cfDNA fragment characteristics; evaluating the cfDNA fragmentome profiles of a reference panel from subjects without cancer to the subject's cfDNA fragmentome profile prior to administration of a cancer therapy (baseline), at one or more time points during therapy, after therapy or combinations thereof; to determine responsiveness to the cancer therapy in the subject. Any of the cancer treatments disclosed herein or known in the art can be administered. For example, a mammal that has been selected for increased monitoring can be further monitored, and a cancer treatment can be administered if the presence of the cancer cell is maintained throughout the increased monitoring period. Additionally, or alternatively, a mammal that has been selected for increased monitoring can be administered a cancer treatment, and further monitored as the cancer treatment progresses. In some embodiments, the subject's cfDNA fragmentomic profile will provide cause to administer a different cancer treatment (e.g., a

resistance mutation may arise in a cancer cell during the cancer treatment, which cancer cell harboring the resistance mutation is resistant to the original cancer treatment).

[00121] When a mammal is identified as having cancer as described herein (e.g., based, at least in part, on the cfDNA fragmentome profile of the mammal), the identifying can be before and/or during the course of a cancer treatment. Methods of identifying a mammal as having cancer provided herein can be used as a first diagnosis to identify the mammal (e.g., as having cancer before any course of treatment) and/or to select the mammal for further diagnostic testing. In some embodiments, once a mammal has been determined to have cancer, the mammal may be administered further tests and/or selected for further diagnostic testing. In some embodiments, methods provided herein can be used to select a mammal for further diagnostic testing at a time period prior to the time period when conventional techniques are capable of diagnosing the mammal with an early-stage cancer. For example, methods provided herein for selecting a mammal for further diagnostic testing can be used when a mammal has not been diagnosed with cancer by conventional methods and/or when a mammal is not known to harbor a cancer. In some embodiments, a mammal selected for further diagnostic testing can be administered a diagnostic test (e.g., any of the diagnostic tests disclosed herein) at an increased frequency compared to a mammal that has not been selected for further diagnostic testing. For example, a mammal selected for further diagnostic testing can be administered a diagnostic test at a frequency of twice daily, daily, bi-weekly, weekly, bi-monthly, monthly, quarterly, semi-annually, annually, or any at frequency therein. In some embodiments, a mammal selected for further diagnostic testing can be administered a one or more additional diagnostic tests compared to a mammal that has not been selected for further diagnostic testing. For example, a mammal selected for further diagnostic testing can be administered two diagnostic tests, whereas a mammal that has not been selected for further diagnostic testing is administered only a single diagnostic test (or no diagnostic tests). In some embodiments, the diagnostic testing method can determine the presence of the same type of cancer (e.g., having the same tissue or origin) as the cancer that was originally detected (e.g., based, at least in part, on the cfDNA fragmentome profile of the mammal). Additionally, or alternatively, the diagnostic testing method can determine the presence of a different type of cancer as the cancer that was original detected. In some embodiments, the diagnostic testing method is a scan. In some embodiments, the scan is a computed tomography (CT), a CT angiography (CTA), an esophagram

(a Barium swallow), a Barium enema, a magnetic resonance imaging (MRI), a PET scan, an ultrasound (e.g., an endobronchial ultrasound, an endoscopic ultrasound), an X-ray, a DEXA scan.

[00122] In some embodiments, the diagnostic testing method is a physical examination, such as an anoscopy, a bronchoscopy (e.g., an autofluorescence bronchoscopy, a white-light bronchoscopy, a navigational bronchoscopy), a colonoscopy, a digital breast tomosynthesis, an endoscopic retrograde cholangiopancreatography (ERCP), an esophagogastroduodenoscopy, a mammography, a Pap smear, a pelvic exam, a positron emission tomography and computed tomography (PET-CT) scan. In some embodiments, a mammal that has been selected for further diagnostic testing can also be selected for increased monitoring. Once the presence of a tumor or a cancer (e.g., a cancer cell) has been identified (e.g., by any of the variety of methods disclosed herein), it may be beneficial for the mammal to undergo both increased monitoring (e.g., to assess the progression of the tumor or cancer in the mammal and/or to assess the development of one or more cancer biomarkers such as mutations), and further diagnostic testing (e.g., to determine the size and/or exact location of the tumor or the cancer). In some embodiments, a cancer treatment is administered to the mammal that is selected for further diagnostic testing after a cancer biomarker is detected and/or after the cfDNA fragmentome profile of the mammal has not improved or deteriorated. Any of the cancer treatments disclosed herein or known in the art can be administered. For example, a mammal that has been selected for further diagnostic testing can be administered a further diagnostic test, and a cancer treatment can be administered if the presence of the tumor or the cancer is confirmed. Additionally, or alternatively, a mammal that has been selected for further diagnostic testing can be administered a cancer treatment and can be further monitored as the cancer treatment progresses. In some embodiments, after a mammal that has been selected for further diagnostic testing has been administered a cancer treatment, the additional testing will reveal one or more cancer biomarkers. In some embodiments, such one or more cancer biomarkers (e.g., mutations) will provide cause to administer a different cancer treatment (e.g., a resistance mutation may arise in a cancer cell during the cancer treatment, which cancer cell harboring the resistance mutation is resistant to the original cancer treatment).

[00123] When treating a mammal having cancer as described herein, the mammal can be administered one or more cancer treatments. A cancer treatment can be any appropriate cancer treatment. One or more cancer treatments described herein can be administered to a mammal at any appropriate frequency (e.g., once or multiple times over a period of time ranging from days to

weeks). Examples of cancer treatments include, without limitation immune checkpoint inhibitors, adjuvant chemotherapy, neoadjuvant chemotherapy, histone deacetylase (HDAC) inhibitors such as 5-azacytidine and entinostat, radiation therapy, hormone therapy, cytotoxic therapy, immunotherapy, adoptive T cell therapy (e.g., chimeric antigen receptors and/or T cells having wild-type or modified T cell receptors), targeted therapy such as administration of kinase inhibitors (e.g., kinase inhibitors that target a particular genetic lesion, such as a translocation or mutation), (e.g. a kinase inhibitor, an antibody, a bispecific antibody), signal transduction inhibitors, bispecific antibodies or antibody fragments (e.g., BiTEs), monoclonal antibodies, immune checkpoint inhibitors, surgery (e.g., surgical resection), or any combination of the above. In some embodiments, a cancer treatment can reduce the severity of the cancer, reduce a symptom of the cancer, and/or to reduce the number of cancer cells present within the mammal.

[00124] In some embodiments, a cancer treatment can include an immune checkpoint inhibitor. Non-limiting examples of immune checkpoint inhibitors include nivolumab (Opdivo), pembrolizumab (Keytruda), atezolizumab (tecentriq), avelumab (bavencio), durvalumab (imfinzi), ipilimumab (yervoy).

EXAMPLES

[00125] EXAMPLE 1: MONITORING RESPONSES TO IMMUNE CHECKPOINT INHIBITION IN ADVANCED SOLID TUMORS USING GENOME-WIDE cfDNA FRAGMENTOMES

[00126] In this study, the utility of DELFI-TF was validated for measuring response and prognosis in patients from a phase I clinical trial evaluating entinostat (a histone deacetylase [HDAC] inhibitor) (17) and nivolumab with or without ipilimumab (immune checkpoint inhibitors [ICIs]) (18) in advanced solid tumors (NCT02453620) (19). Response assessment based on CT-imaging criteria was compared to DELFI-TF, underlying molecular fragmentome characteristics during therapy were identified, and these analyses were integrated with blood cell counts for ICI monitoring strategies.

[00127] Results

[00128] Study population

[00129] We evaluated the application of cfDNA fragmentome analyses for the purpose of molecularly monitoring treatment response in patients receiving entinostat and nivolumab with or without ipilimumab in advanced solid tumors from a phase I clinical trial (NCT02453620) (FIG. 1) (18). Plasma was obtained from 50 participants who had a blood draw at initiation of the clinical

trial (week 0), after entinostat run-in (week 2), and after receiving combined entinostat and immune checkpoint blockade (week 10) (**FIG. 1**). The 50 patients (47 female, 3 male, median age 59 years) with breast (n = 28), salivary gland (n = 7), gastrointestinal (n = 6), gynaecological (n = 4), sarcoma (n = 3), or lung cancer (n = 2) received a median of four prior lines of therapy (**FIG. 1b**, **Table 1 above**). Of these patients, only 54% (27/50) (25 female, 2 male, median age 59 years) had both a baseline and week 10 time point sample available for longitudinal liquid biopsy analyses because a subset of the advanced solid tumor patients did not reach week 10. The patients with both samples available included patients with breast (n = 16), salivary gland (n = 3), gastrointestinal (n = 3), gynaecological (n = 1), sarcoma (n = 2), and lung cancer (n = 2) who received a median of 3 prior lines of therapy. The subjects evaluable at both time points were representative of all subjects with regard to age distribution and prior lines of therapy. We isolated 4-5 ml of plasma from each of the 50 patients and examined cfDNA using low coverage whole-genome sequencing (~2x coverage). The analyses in this cohort were subsequently validated in a separate cohort (n=21) of lung cancer patients with advanced disease treated with ICI and/or chemotherapy as standard of care (**Table 2 above**).

[00130] Liquid biopsy-based monitoring of disease status

[00131] We analyzed approximately 35 million paired reads per sample and evaluated cfDNA fragmentation profiles (cfDNA fragmentomes) (19–21), comprising the compendium of genome-wide cfDNA fragment characteristics for the samples analyzed. The underlying fragmentomes were measured at baseline as well as after combined HDAC and ICI and were separated by response category as determined by conventional imaging response evaluation criteria in solid tumors (RECIST v1.1) (**FIGS. 2A, 2B**). Comparison of fragmentation profiles of a reference panel of individuals without cancer (n=54) to patient profiles at baseline or after therapy revealed a decreased correlation to patients at baseline or who were RECIST non-responders (n=20) (baseline mean $r = 0.62$, clinical non-responder mean $r = 0.79$, Pearson correlation coefficient) (**FIG. 2A**). Fragmentation features appeared to be broadly altered across the genome in the vast majority of patients at baseline and were much more consistent in individuals responding to therapy regardless of patient demographic or clinical characteristics (mean $r = 0.94$, Pearson correlation coefficient) (**FIG. 2B**). The high correlation between individuals without cancer and treatment responders suggested that genome-wide fragmentomes were able to molecularly identify patients determined to be responders by RECIST criteria whereas

fragmentation profiles among RECIST non-responders revealed aberrant profiles indicative of disease burden (**FIG. 2A**).

[00132] *DELFI-TF and clinical outcome*

[00133] The DELFI-TF machine learning model utilizing genome-wide fragmentome features was applied to this current phase I immunotherapy cohort of patients and DELFI-TF scores were calculated at baseline (week 0, n=50 patients), after run-in entinostat (week 2, n=48 patients), and after immune checkpoint inhibition (week 10, n=27 patients).

[00134] Median DELFI-TF values were similar at baseline (DELFI-TF = 0.11) and at the 2-week entinostat time point (DELFI-TF = 0.16) (p-value = 0.48, Welch two sample t-test), but were markedly reduced at the week 10 time point (DELFI-TF=0.0065) (Welch two sample t-test, p-value = 0.038) (**FIG. 2C**). Evaluation of overall survival in these patients demonstrated a survival benefit for patients with DELFI-TF scores below the observed median value at the baseline timepoint (log-rank, p = 0.0093) (**FIG. 2D**). Similar results were obtained when evaluating the subset of patients who had blood samples across all three time points (n=27) (**FIG. 6, FIG. 7**). These results suggested that DELFI-TF scores prior to initiation of therapy may have important prognostic value for cancer patients, similar to previous analyses using targeted panels for evaluation of tumor-derived cfDNA mutations (10, 22).

[00135] We then examined the change in DELFI-TF scores for each individual between the baseline sample and the 10-week sample obtained after 8 weeks of treatment with entinostat and immune checkpoint blockade. We hypothesized that a drop of 50% or greater of DELFI-TF scores would indicate a molecular response representative of decreased tumor burden, similar to criteria used for mutant allele fraction assessment (10, 23, 24), while individuals without this drop would be considered to be non-responders. Using this approach, we observed that imaging and DELFI-TF changes were in agreement in categorizing the majority of patients: all but one of the individuals with RECIST determined complete or partial clinical response (CR/PR) had molecular responses while for those patients with progressive disease (PD) all but one did not have a molecular response (**FIG. 3A, FIG. 3B, FIG. 8**). The exception among CR/PR patients was for an individual with a sarcoma that had very low DELFI-TF values that were reduced by only -37%, while the exception among PD patients occurred in an individual whose tumor sum of longest diameters decreased after therapy but was categorized as PD due to development of a new lesion.

Individuals categorized as having SD included both those with and without molecular responses (FIG. 3B).

[00136] We next examined progression-free survival (PFS) for patients who were categorized for response to therapy using either changes in RECIST or DELFI-TF between baseline and the week 10 time point. RECIST assessments categorized 7 patients as having PR and 20 patients as having SD/PD. Patients measured with RECIST and categorized as PR had very similar outcomes to patients with SD/PD (median PFS 6.4 vs. 6.7 months; hazard ratio, HR = 1.7, 95% CI=0.54-5.3, p=0.36) (FIG. 3D). In contrast, DELFI-TF assessments identified 12 molecular responders and 15 molecular non-responders who experienced a median PFS of 25.8 and 4.5 months, respectively (HR = 5.5, 95% CI=1.7-17, p=0.004) (FIG. 3E, FIG. 9), demonstrating an improved categorization by DELFI-TF compared to imaging analyses. DELFI-TF assessments were similarly more predictive of overall survival than changes in RECIST between baseline and week 10 time points (FIG. 10).

[00137] We hypothesized that the inability of imaging to accurately predict progression free survival at 10 weeks after treatment initiation was due to the incorrect categorization of a subset of individuals as having stable disease, and evaluated whether molecular response assessments could improve outcome prediction in these individuals. Among the subgroup of individuals with radiographic stable disease, the median PFS for those with a DELFI-TF molecular response (n = 6) was 25.8 months compared to 5.2 months for those without a molecular response (n = 10) (log-rank, p = 0.012) (FIG. 3F). Additionally, patients with SD at week 10 who had a durable clinical response (≥ 6 months) had lower DELFI-TF values compared to individuals without a durable clinical response (< 6 months) (Welch two sample t-test, p-value = 0.01) (FIG. 11).

[00138] We examined the time CT-scans took to reach best overall response compared to the timing of DELFI-TF assessments and found that CT-scans detected clinical responses an average of 6.8 months later than molecular responses through fragmentome analyses (Welch two sample t-test, p = 0.048) (FIG. 3C). In particular, 5 of the 27 patients (19% of patients (95% CI: 6.3%-38.1%)) required additional time after week 10 to reach best overall RECIST response, comprising two individuals who were classified as SD at week ten but were later determined to have PR, and three individuals who were ultimately determined to have SD. These analyses highlight the potential of the DELFI-TF approach for assessment of therapeutic response of

patients treated with immune checkpoint inhibition, including for individuals initially thought to have radiographic stable disease.

[00139] Cell-free DNA fragmentomes in patients with breast cancer

[00140] The majority of patients evaluated in this phase I trial had breast cancer (n = 28) and evaluation of this subset of patients for progression-free survival using RECIST and DELFI-TF criteria reveal similar improvement of molecular responses compared to radiographic assessments at the 10 week time point (RECIST log rank p =0.28; HR=2.3, 95% CI=0.47-11; p=0.3; DELFI-TF log rank p=0.014, HR=5; 95% CI=1.2-20; p=0.026) (**FIGS. 14A, 14B**). The DELFI-TF assessment was reflected in changes in genome-wide fragmentation profiles during therapy in patients with either hormone receptor positive (HR+) or triple negative breast cancers (TNBC) (**FIGS. 14C, 14d**). These well-defined groups of breast cancer patients provided an opportunity to explore additional biological characteristics of cfDNA fragmentation.

[00141] We have previously shown that cfDNA fragmentation in patients with cancer reflects the chromatin states from cancer cells as well as from white blood cells (21). We hypothesized that in the current trial cfDNA from patients with breast cancer at baseline or who were not responsive to therapy would have a fragmentome profile that reflected the chromatin states of breast cancer tissue, while those individuals who were responsive or without cancer would have profiles similar to those of normal white blood cells. We found that fragmentation profiles of patients who did not have a molecular response at week 10 closely resembled open and closed genome compartments obtained through genome-wide methylation analyses of breast cancers, while individuals who were responsive at this time point or who did not have cancer had profiles similar to genome compartments obtained from high throughput sequencing chromosome conformation capture (Hi-C) of non-cancer lymphoblast cells (**FIG. 4A**). Interestingly, patients who did not respond to therapy had a stronger breast cancer chromatin signal in cfDNA than patients at baseline. These analyses suggest that cfDNA fragmentomes from individuals with breast cancer represent a mixture of cfDNA profiles of chromatin compartments of cells from peripheral blood as well as those from breast cancer, and that fragmentome analyses capture changes in chromatin representation during therapy.

[00142] Because of the distinct treatments for different breast cancer subtypes (25), we evaluated the possibility of identifying different breast cancer histologies using genome-wide cfDNA fragmentation profiles. We assessed regions of the genome from chromatin

immunoprecipitation experiments available from the reMap database (n=5620 experiments) and compared cfDNA coverage differences between HR+ and TNBC at transcription factor binding sites (**FIG. 4B**). Regions with the largest changes in cfDNA coverage between these subtypes revealed decreased cfDNA coverage at binding sites for ESR1, AR, GATA3, FOXA1, and TLE3 in HR+ compared to TNBC patients samples (**FIG. 4C**). All of these genes are known to be highly expressed and involved in luminal breast cancer and reflects the ability to identify subtype specific transcriptome differences in cfDNA (26–31). cfDNA fragmentation coverage of binding sites of individual transcription factors had high performance for distinguishing between HR+ and TNBC patients (AUCs for ESR1 = 0.92, AR = 0.92, GATA3 = 0.91, FOXA1 = 0.90 and TLE3=0.91) (**FIG. 4D**). Interestingly, cfDNA coverage differences at these binding sites were altered for molecular responders compared to non-responders at week 10 after initiation of immune checkpoint blockade therapy (p= 0.014, Wilcoxon test) (**FIG. 15**). These analyses provide initial evidence that low-coverage cfDNA fragmentation can be used to distinguish breast cancer subtypes noninvasively.

[00143] DELFI-TF and tumor immune cell responses

[00144] We also explored enhancing DELFI-TF using blood cell counts as studies have indicated that neutrophil to lymphocyte ratios (NLRs) as well as changes in levels of other immune cells in patients with advanced cancers may be prognostic, with high NLR (32, 33) as well as high monocyte or basophil levels linked to poor outcomes (34). Accordingly, we expanded our analyses to consider complete blood count studies for the majority of the patients (n=27) analyzed at baseline (week 0), after entinostat treatment (week 2), and after immune checkpoint therapy (week 10) (**FIGS. 16A, 16B**). We found that patients who had clinical RECIST responses or DELFI-TF molecular responses to immunotherapy demonstrated lower NLRs than those without responses after immune checkpoint inhibition (RECISTv1.1 p-value=0.032, molecular p-value=0.009) (**FIGS. 5A, 5B**). Similarly, patients who had clinical or molecular responses after immunotherapy had lower median monocyte and basophil counts compared to patients without such responses (monocyte p=0.0071; basophil p=0.034) (**FIGS. 5C, 5D, 5E, 5F**). Development of a DELFI-multi model that included the DELFI-TF model scores, absolute counts of monocytes and basophils, as well as NLR levels demonstrated an enhanced ability to identify patients with improved progression-free or overall survival (log rank, p-value < 0.0001) (**FIGS. 5G, 5H**). These data

highlight the potential benefit of combining DELFI-TF and blood cell counts to improve early detection of molecular responses.

[00145] Validation of DELFI-TF in an external patient population treated with ICI

[00146] To validate the performance of the cfDNA DELFI-TF fragmentome approach we evaluated an additional group of patients (n=21) with late-stage lung cancer treated with ICI and/or chemotherapy (**Table 2 above**). The individuals had no prior lines of therapy and had blood collections available at baseline (n=21), week 6 (n=15), and week 11 (n=11). As expected, baseline genome-wide cfDNA fragmentation profiles demonstrated aberrant patterns across the genome (**FIG. 12A**). When evaluating longitudinal changes in between baseline and subsequent time points, it became evident that DELFI-TF values at baseline were significantly higher to those from the time points after initiation of therapy, including at week 6 and 11 (**FIG. 12B**). Utilizing the same threshold of a 50% reduction in the DELFI-TF score for molecular responders, we found that the approach accurately identified all of the individuals at 11 weeks after initiation of therapy who were categorized as clinical responders (CR/PR) using RECIST 1.1 (p=0.016, Wilcoxon) (**FIG. 12C**), while individuals who were deemed to be clinical non-responders (SD/PD) did not have a similar reduction of DELFI-TF values (p=0.17, Wilcoxon) (**FIG. 12D**). The change in DELFI-TF scores between clinical responders and non-responders after initiation of therapy in this population (p<0.036, Wilcoxon) (**FIG. 12E**) was similar to that observed in the initial cohort (**FIG. 3B**) suggesting the approach was generalizable across patients with cancer treated with ICI.

Table 1

	Baseline n = 50	Week 2 n = 48	Week 10 n = 27
Age, n (%)			
> 50	35 (70%)	33 (69%)	18 (67%)
≤ 50	15 (30%)	15 (31%)	9 (33%)
Sex, n (%)			
Male	3 (6%)	3 (6%)	2 (7%)
Female	47 (94%)	45 (94%)	25 (93%)
Best Overall RECIST, n (%)			
CR	1 (2%)	1 (2%)	1 (4%)
PR	7 (14%)	6 (13%)	7 (26%)
SD	18 (34%)	18 (34%)	12 (44%)
PD	15 (30%)	13 (27%)	7 (26%)
NI	-	2 (4%)	-
Discontinued	9 (18%)	8 (17%)	-
Histology, n (%)			
Adenoid cyst	4 (8%)	4 (8%)	2 (7%)
Apocrine	1 (2%)	1 (2%)	1 (4%)
Bile duct	3 (6%)	3 (6%)	2 (7%)
Bowel	1 (2%)	1 (2%)	1 (4%)
Breast	28 (56%)	26 (54%)	16 (59%)
Cervix	1 (2%)	1 (2%)	-
Colorectal	1 (2%)	1 (2%)	-
Endometrial	1 (2%)	1 (2%)	1 (4%)
Gastric	1 (2%)	1 (2%)	-
Liposarcoma	1 (2%)	1 (2%)	1 (4%)
Lung	2 (4%)	2 (4%)	2 (7%)
Ovarian	1 (2%)	1 (2%)	-
Pancreatic	2 (4%)	2 (4%)	-
Phecoma	1 (2%)	1 (2%)	-
Sarcoma	1 (2%)	1 (2%)	1 (4%)
Uterine	1 (2%)	1 (2%)	-

Table 2. Immunotherapy validation cohort patient characteristics

Clinical features	
Median age (range)	73 (54-89)
Sex	
Male	11 (52%)
Female	10 (48%)
Median BMI (range)	24.5 (17-37)
Stage	
Stage III	7 (33%)
Stage IV	14 (67%)
Histology	
NSCLC	
Adenocarcinoma	11 (52%)
Squamous cell carcinoma	5 (24%)
Neuroendocrine	1 (5%)
Non-small cell carcinoma (C34_)	1 (5%)
SCLC	
Carcinoma	
Malignant lymphoma, large B-cell	2 (9%)
	1 (5%)
Longitudinal samples	
Timepoint 0 (Baseline)	21 (37.5%)
Timepoint 1 (week 6)	21 (37.5%)
Timepoint 2 (week 11)	14 (25%)
PD-L1 expression	
Strong positive	5 (24%)
Weak positive	7 (33%)
Negative	4 (19%)
NA	5 (24%)
EGFR mutation status	
Wild type	11 (52%)
NA	10 (48%)
Post-baseline Treatment type	
Immunotherapy	6 (28.5%)
Immunotherapy + chemotherapy	9 (43%)
Chemotherapy	6 (28.5%)
RECIST Overall response	
Complete response	1 (5%)
Partial response	2 (9%)
Stable disease	7 (33%)
Progressive disease	9 (43%)
Divergent event	2 (9%)

[00147] Discussion

[00148] In this study, we validate the utility of genome-wide cfDNA fragmentomes through the DELFI-TF approach to monitor therapeutic response in patients from a phase I clinical trial

evaluating entinostat and nivolumab with or without ipilimumab in advanced solid tumors (NCT02453620) (18). We show the ability of DELFI-TF to rapidly detect molecular responses that are more predictive of PFS than conventional CT-imaging. Our analyses revealed that cfDNA fragmentomes reflected chromatin states of cancer during therapy and could be used to distinguish breast cancer subtypes. The integration of DELFI-TF with blood cell analyses resulted in a combined model that may improve detection of response. The DELFI-TF approach was further validated in a separate group of patients treated with ICI. To our knowledge, this is the first fragmentation-based approach that provides an estimate of tumor fraction to detect primary response or resistance to immunotherapy.

[00149] Some of the strengths of our study include the evaluation of the existing DELFI-TF approach in a prospective clinical trial setting in which patient samples were collected and preserved systematically for the purpose of the analyses described in this study. The alignment in timing at baseline and week 10 of both liquid biopsies and CT-scans for clinical RECIST provided an opportunity for a robust comparison to the existing standard approach not usually available in other studies. Additionally, because the fragmentome based test only requires plasma samples, our approach obviates the need for matched tumor tissue or white blood cells to identify tumor-specific changes or remove non-tumor derived mutations which are common confounders of other approaches (13, 14, 19). The use of genome-wide features in DELFI-TF increases the applicability of the approach compared to targeted panels that are limited by the number of assessable genes in specific tumor types for which they were designed. As an example, the use of fragmentome analyses to detect and distinguish different breast cancer subtypes noninvasively suggests that the approach may be useful for therapeutic stratification in a tumor type that have historically been challenging for liquid biopsy analyses (6, 35).

[00150] Some of the limitations of the study include the modest sample size of 50 individuals in this phase I trial, as well as that not all individuals had matching plasma available before and after therapy. This concern is mitigated by the independent validation of this approach in an additional group of patients with lung cancer also treated with ICI. Additionally, our study only had three liquid biopsy timepoints available for analysis, and additional longitudinal timepoints may be helpful for determining optimal timing to monitor response. Validation of the DELFI-TF approach in larger clinical trials will be needed prior to clinical use. Despite these challenges, we were able to evaluate the early changes in DELFI-TF fragmentation profiles of

patients both at baseline as well as after initiation of immune checkpoint blockade and compare this to overall and progression-free survival. Although analysis of the effect of entinostat after 2-week run-in did not lead to a detectable difference in DELFI-TF scores, this is consistent with a previous study showing no detectable variations in DNA accessibility with this therapy (36).

[00151] Overall, we present a machine-learning fragmentome-based approach for monitoring patients during therapy that is inexpensive, does not rely on few mutations, and is not limited by the need for invasive tumor biopsies or parallel analyses of white blood cells. With the increasing number of new indications for immune checkpoint blockade, including the recent FDA approval of pembrolizumab for TNBC patients (37), there is an emerging need for effective monitoring approaches in these settings. These analyses provide an initial proof-of-concept framework that may have broad applicability in the setting of monitoring disease in patients receiving immunotherapy.

[00152] Example 2: Materials and Methods

[00153] Patient and sample collection characteristics

[00154] Blood samples were collected from patients with advanced solid tumors participating within the multisite, open-label, phase I clinical trial (NCT02453620, ETCTN-9844)(18). Eligibility criteria and clinical trial design are previously described (18). Briefly, individuals with advanced solid unresectable or metastatic tumors without curative or palliative treatment options remaining were assessed for eligibility to receive ICI. The trial design follows a 3+3 dose-escalation with four dose levels (DL): DL1/2, entinostat 3-5 mg weekly plus nivolumab 3 mg/kg every 2 weeks and DL3/4 administered at entinostat 3-5mg weekly plus nivolumab 3 mg/kg every 2 weeks plus ipilimumab 1 mg/kg every 6 weeks(18). Image-assessment for measurable or evaluable/non-measurable disease per RECIST version 1.1 was implemented. Blood samples were collected at baseline (week 0), post run-in with entinostat (week 2), and after 8 weeks of combination therapy dependent on the assigned dose level (week 10). Peripheral blood was collected in EDTA tubes and plasma was freshly isolated and subsequently stored at -70°C or below for downstream analyses. Blood samples for the validation cohort were collected in Streck tubes from patients with advanced lung cancer by Indivumed (Hamburg, Germany) and plasma was isolated in a similar fashion.

[00155] Genomic library construction

[00156] cfDNA was isolated using the Qiagen QIAamp Circulating Nucleic Acids Kit (Qiagen GmbH) from approximately 4 to 5 mls of plasma. Extracted cfDNA was eluted into low bind tubes at a volume of 52ul's of Qiagen buffer AVE and quantified using the Bioanalyzer 2100 (Agilent Technologies). Genomic libraries for next-generation sequencing were constructed using an input of 15 to 125 ng of cfDNA as previously described(20). The NEBNext DNA Library Prep Kit for Illumina (New England Biolabs) was utilized for library generation containing four essential modifications to the manufacturer's guidelines: (i) the library purification steps use the on-bead AMPure XP (Beckman Coulter) approach to minimize sample loss during elution and tube transfer steps; (ii) NEBNext End Repair, A-tailing and adaptor ligation enzyme and buffer volumes were adjusted as appropriate to accommodate on-bead AMPure XP purification; (iii) Illumina dual index adaptors were used in the ligation reaction; and (iv) cfDNA libraries were amplified with Phusion Hot Start Polymerase. Whole-genome samples underwent a 4 cycle PCR amplification following DNA ligation. In total nine genomic library batches were completed for next-generation sequencing preparation. Each batch was designed to contain the available longitudinal timepoints per patient. The final library batches included a technical replicate of nucleosomal DNA derived from nuclease-digested human peripheral blood mononuclear cells (PBMCs) to evaluate reproducibility across batches.

[00157] Whole-genome sequencing and alignment

[00158] Genomic libraries from each patient across longitudinal treatment timepoints were sequenced at 100bp paired-end (200 cycles) for a target of 1-2x coverage per genome on the Hiseq 2500 or Novaseq 6000 platform. Adaptor sequences were trimmed with fastp, and reads were aligned against the human reference genome (hg19) with Bowtie2. Duplicate reads were removed with Sambamba(20). Paired reads were filtered to only retain reads with a mapq score of at least 30 or greater. Additional filtering was conducted of read pairs overlapping to remove reads aligned to the Duke Excluded Regions blacklist (<https://genome.ucsc.edu/cgi-bin/hgTrackUi?db=hg19&g=wgEncodeMapability>). The hg19 reference genome was tiled into non-overlapping 5 MB bins; bins containing fragments with an average GC content <0.3 or mappability <0.9 were excluded.

[00159] Genome-wide fragmentome analyses and computation of DELFI-TF

[00160] Sequenced cfDNA fragments were evaluated for short (100-150 bp) to long (151-220 bp) ratios across 473 5MB bins that were GC-corrected using the nonparametric fragment

level model described in our previous study(20). To compute DELFI-TF we leveraged a Random forest regression model that was initially trained to estimate mutant allele frequency of tumor-specific variants measured by digital droplet PCR in 692 longitudinal plasma samples collected from the CAIRO5 phase III clinical trial (15). The features of this model are previously described (15).

[00161] Analysis of cfDNA fragmentation at transcription factor binding sites

[00162] Chromatin immunoprecipitation followed by next-generation sequencing of 5,620 transcription factor experiments were accessed from the ReMap 2020 database(38). The mean coverage from low coverage (1-2x) genome-wide sequencing at each position (-3,000 to +3,000 with regard to peak center) was computed for all peaks for each sample. Relative coverage was then calculated for every sample by taking the mean coverage ± 100 bp window divided by the mean coverage in a ± 250 bp window of the surrounding 2,750 bp downstream and upstream of the binding sites. A pairwise t-test was performed comparing the relative coverage of all transcription factors between hormone receptor positive and triple negative breast cancer samples. Bonferroni multiple test correction was applied to identify statistically significant transcription factors using the limma R package(39). Volcano plots depict log2 fold change in relative coverage versus log10 adjusted p-values. For subtyping, multiple representations of a transcription factor indicate a hit from multiple distinct ChIP-seq experiments. For transcription factors associated with molecular response, transcription factors with 6,000 or more genomic binding sites were assessed.

[00163] Chromatin Structure analysis

[00164] A/B compartments for breast cancer tissue and lymphoblastoid cells were obtained from https://github.com/Jfortin1/TCGA_AB_Compartments as well as from https://github.com/Jfortin1/HiC_AB_Compartments(40). The two reference tracks were compared to identify informative 100kb bins, defined as bins where the chromatin domain differed between the two reference tracks or the magnitude difference in eigenvalues that corresponded to a z-score greater than 1.96 or less than -1.96 ($p=.05$) across all eigenvalue differences. The median fragmentation profiles for the six breast cancer samples with the highest estimated tumor fraction by ichorCNA (41) and 10 randomly selected individuals without cancer from another cohort were calculated. This information was used to extract an estimated median breast component in the plasma weighted by the ichor score of the individual plasma samples as previously performed for cfDNA analyses of patients with liver cancer (21). The overall fragmentation profiles at different

timepoints were then compared to the reference A/B compartments to identify changes in chromatin structure.

[00165] Statistical analyses

[00166] All statistical analyses were performed with R Statistical Software (version 4.2.1 Foundation for Statistical Computing, Vienna, Austria).

[00167] References

1. L. H. Schwartz, S. Litière, E. de Vries, R. Ford, S. Gwyther, S. Mandrekar, L. Shankar, J. Bogaerts, A. Chen, J. Dancey, W. Hayes, F. S. Hodi, O. S. Hoekstra, E. P. Huang, N. Lin, Y. Liu, P. Therasse, J. D. Wolchok, L. Seymour, RECIST 1.1-Update and clarification: From the RECIST committee. *Eur. J. Cancer* 62, 132–137 (2016).
2. L. Seymour, J. Bogaerts, A. Perrone, R. Ford, L. H. Schwartz, S. Mandrekar, N. U. Lin, S. Litière, J. Dancey, A. Chen, F. S. Hodi, P. Therasse, O. S. Hoekstra, L. K. Shankar, J. D. Wolchok, M. Ballinger, C. Caramella, E. G. E. de Vries, RECIST working group, iRECIST: guidelines for response criteria for use in trials testing immunotherapeutics. *Lancet Oncol.* 18, e143–e152 (2017).
3. L. F. Chai, E. Prince, V. G. Pillarisetty, S. C. Katz, Challenges in assessing solid tumor responses to immunotherapy. *Cancer Gene Ther.* 27, 528–538 (2020).
4. C. Solinas, M. Porcu, Z. Hlavata, P. De Silva, M. Puzzoni, K. Willard-Gallo, M. Scartozzi, L. Saba, Critical features and challenges associated with imaging in patients undergoing cancer immunotherapy. *Crit. Rev. Oncol. Hematol.* 120, 13–21 (2017).
5. V. Anagnostou, P. M. Forde, J. R. White, N. Niknafs, C. Hruban, J. Naidoo, K. Marrone, I. K. A. Sivakumar, D. C. Bruhm, S. Rosner, J. Phallen, A. Leal, V. Adleff, K. N. Smith, T. R. Cottrell, L. Rhymee, D. N. Palsgrove, C. L. Hann, B. Levy, J. Feliciano, C. Georgiades, F. Verde, P. Illei, Q. K. Li, E. Gabrielson, M. V. Brock, J. M. Isbell, J. L. Sauter, J. Taube, R. B. Scharpf, R. Karchin, D. M. Pardoll, J. E. Chaft, M. D. Hellmann, J. R. Brahmer, V. E. Velculescu, Dynamics of Tumor and Immune Responses during Immune Checkpoint Blockade in Non-Small Cell Lung Cancer. *Cancer Res.* 79, 1214–1225 (2019).
6. J. Phallen, M. Sausen, V. Adleff, A. Leal, C. Hruban, J. White, V. Anagnostou, J. Fiksel, S. Cristiano, E. Papp, S. Speir, T. Reinert, M.-B. W. Orntoft, B. D. Woodward, D. Murphy, S. Parpart-Li, D. Riley, M. Nesselbush, N. Sengamalay, A. Georgiadis, Q. K. Li, M. R. Madsen, F. V. Mortensen, J. Huiskens, C. Punt, N. van Grieken, R. Fijneman, G. Meijer, H. Husain, R. B. Scharpf, L. A. Diaz Jr, S. Jones, S. Angiuoli, T. Ørntoft, H. J. Nielsen, C. L. Andersen, V. E. Velculescu, Direct detection of early-stage cancers using circulating tumor DNA. *Sci. Transl. Med.* 9 (2017), doi:10.1126/scitranslmed.aan2415.
7. J. Phallen, A. Leal, B. D. Woodward, P. M. Forde, J. Naidoo, K. A. Marrone, J. R. Brahmer, J. Fiksel, J. E. Medina, S. Cristiano, D. N. Palsgrove, C. D. Gocke, D. C. Bruhm, P. Keshavarzian, V. Adleff, E. Weihe, V. Anagnostou, R. B. Scharpf, V. E. Velculescu, H. Husain, Early Noninvasive Detection of Response to Targeted Therapy in Non-Small Cell Lung Cancer. *Cancer Res.* 79, 1204–1213 (2019).
8. S. V. Bratman, S. Y. C. Yang, M. A. J. Iafolla, Z. Liu, A. R. Hansen, P. L. Bedard, S. Lheureux, A. Spreafico, A. A. Razak, S. Shchegrova, M. Louie, P. Billings, B. Zimmermann, H.

Sethi, A. Aleshin, D. Torti, K. Marsh, J. Eagles, I. Cirlan, Y. Hanna, D. L. Clouthier, S. C. Lien, P. S. Ohashi, W. Xu, L. L. Siu, T. J. Pugh, Personalized circulating tumor DNA analysis as a predictive biomarker in solid tumor patients treated with pembrolizumab. *Nat Cancer* 1, 873–881 (2020).

9. B. Y. Nabet, M. S. Esfahani, E. J. Moding, E. G. Hamilton, J. J. Chabon, H. Rizvi, C. B. Steen, A. A. Chaudhuri, C. L. Liu, A. B. Hui, D. Almanza, H. Stehr, L. Gojenola, R. F. Bonilla, M. C. Jin, Y.-J. Jeon, D. Tseng, C. Liu, T. Merghoub, J. W. Neal, H. A. Wakelee, S. K. Padda, K. J. Ramchandran, M. Das, A. J. Plodkowski, C. Yoo, E. L. Chen, R. B. Ko, A. M. Newman, M. D. Hellmann, A. A. Alizadeh, M. Diehn, Noninvasive Early Identification of Therapeutic Benefit from Immune Checkpoint Inhibition. *Cell* 183, 363-376.e13 (2020).

10. Q. Zhang, J. Luo, S. Wu, H. Si, C. Gao, W. Xu, S. E. Abdullah, B. W. Higgs, P. A. Dennis, M. S. van der Heijden, N. H. Segal, J. E. Chaft, T. Hembrough, J. C. Barrett, M. D. Hellmann, Prognostic and Predictive Impact of Circulating Tumor DNA in Patients with Advanced Cancers Treated with Immune Checkpoint Blockade. *Cancer Discov.* 10, 1842–1853 (2020).

11. W. Al Zoughbi, J. Fox, S. Beg, E. Papp, E. Hissong, K. Ohara, L. Keefer, M. Sigouros, T. Kane, D. Bockelman, D. Nichol, E. Patchell, R. Bareja, A. Karandikar, H. Alnajjar, G. Cerqueira, V. B. Guthrie, E. Verner, J. Manohar, N. Greco, D. Wilkes, S. Tagawa, M. S. Malbari, K. Holcomb, K. W. Eng, M. Shah, N. K. Altorki, A. Sboner, D. Nanus, B. Faltas, C. N. Sternberg, J. Simmons, Y. Houvras, A. M. Molina, S. Angiuoli, O. Elemento, J. M. Mosquera, Validation of a Circulating Tumor DNA-Based Next-Generation Sequencing Assay in a Cohort of Patients with Solid tumors: A Proposed Solution for Decentralized Plasma Testing. *Oncologist* 26, e1971–e1981 (2021).

12. R. Woodhouse, M. Li, J. Hughes, D. Delfosse, J. Skoletsy, P. Ma, W. Meng, N. Dewal, C. Milbury, T. Clark, A. Donahue, D. Stover, M. Kennedy, J. Dacpano-Komansky, C. Burns, C. Vietz, B. Alexander, P. Hegde, L. Dennis, 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).

13. A. Leal, N. C. T. van Grieken, D. N. Palsgrove, J. Phallen, J. E. Medina, C. Hruban, M. A. M. Broeckaert, V. Anagnostou, V. Adleff, D. C. Bruhm, J. V. Canzoniero, J. Fiksel, M. Nordsmark, F. A. R. M. Warmerdam, H. M. W. Verheul, D. J. van Spronsen, L. V. Beerepoot, M. M. Geenen, J. E. A. Portielje, E. P. M. Jansen, J. van Sandick, E. Meershoek-Klein Kranenbarg, H. W. M. van Laarhoven, D. L. van der Peet, C. J. H. van de Velde, M. Verheij, R. Fijneman, R. B. Scharpf, G. A. Meijer, A. Cats, V. E. Velculescu, White blood cell and cell-free DNA analyses for detection of residual disease in gastric cancer. *Nat. Commun.* 11, 525 (2020).

14. I. van 't Erve, J. E. Medina, A. Leal, E. Papp, J. Phallen, V. Adleff, E. J. Chiao, A. S. Arun, K. Bolhuis, J. K. Simmons, A. Karandikar, K. C. Valkenburg, M. Sausen, S. V. Angiuoli, R. B. Scharpf, C. J. A. Punt, G. A. Meijer, V. E. Velculescu, R. J. A. Fijneman, Metastatic colorectal cancer treatment response evaluation by ultra-deep sequencing of cell-free DNA and matched white blood cells. *Clin. Cancer Res.* (2022), doi:10.1158/1078-0432.CCR-22-2538.

15. van 't Erve, I., Lumbard, K., Alipanahi, B., Skidmore, Z.L., Cancer treatment monitoring using cell-free DNA fragmentomes. *Submitted* (2023).

16. B. J. Christmas, C. I. Rafie, A. C. Hopkins, B. A. Scott, H. S. Ma, K. A. Cruz, S. Woolman, T. D. Armstrong, R. M. Connolly, N. A. Azad, E. M. Jaffee, E. T. Roussos Torres, Entinostat

Converts Immune-Resistant Breast and Pancreatic Cancers into Checkpoint-Responsive Tumors by Reprogramming Tumor-Infiltrating MDSCs. *Cancer Immunol Res* **6**, 1561–1577 (2018).

17. A. J. Korman, S. C. Garrett-Thomson, N. Lonberg, The foundations of immune checkpoint blockade and the ipilimumab approval decennial. *Nat. Rev. Drug Discov.* **21**, 509–528 (2022).

18. E. T. Roussos Torres, C. Rafie, C. Wang, D. Lim, A. Brufsky, P. LoRusso, J. P. Eder, V. Chung, M. Downs, M. Geare, R. Piekarz, H. Streicher, L. Anforth, M. A. Rudek, Q. Zhu, S. Besharati, A. Cimino-Mathews, R. A. Anders, V. Stearns, E. M. Jaffee, R. M. Connolly, Phase I Study of Entinostat and Nivolumab with or without Ipilimumab in Advanced Solid Tumors (ETCTN-9844) *Clin. Cancer Res.* **27**, 5828–5837 (2021).

19. S. Cristiano, A. Leal, J. Phallen, J. Fiksel, V. Adleff, D. C. Bruhm, S. Ø. Jensen, J. E. Medina, C. Hruban, J. R. White, D. N. Palsgrove, N. Niknafs, V. Anagnostou, P. Forde, J. Naidoo, K. Marrone, J. Brahmer, B. D. Woodward, H. Husain, K. L. van Rooijen, M.-B. W. Ørntoft, A. H. Madsen, C. J. H. van de Velde, M. Verheij, A. Cats, C. J. A. Punt, G. R. Vink, N. C. T. van Grieken, M. Koopman, R. J. A. Fijneman, J. S. Johansen, H. J. Nielsen, G. A. Meijer, C. L. Andersen, R. B. Scharpf, V. E. Velculescu, Genome-wide cell-free DNA fragmentation in patients with cancer. *Nature* **570**, 385–389 (2019).

20. D. Mathios, J. S. Johansen, S. Cristiano, J. E. Medina, J. Phallen, K. R. Larsen, D. C. Bruhm, N. Niknafs, L. Ferreira, V. Adleff, J. Y. Chiao, A. Leal, M. Noe, J. R. White, A. S. Arun, C. Hruban, A. V. Annapragada, S. Ø. Jensen, M.-B. W. Ørntoft, A. H. Madsen, B. Carvalho, M. de Wit, J. Carey, N. C. Dracopoli, T. Maddala, K. C. Fang, A.-R. Hartman, P. M. Forde, V. Anagnostou, J. R. Brahmer, R. J. A. Fijneman, H. J. Nielsen, G. A. Meijer, C. L. Andersen, A. Mellemaard, S. E. Bojesen, R. B. Scharpf, V. E. Velculescu, Detection and characterization of lung cancer using cell-free DNA fragmentomes. *Nat. Commun.* **12**, 5060 (2021).

21. Z. H. Foda, A. V. Annapragada, K. Boyapati, D. C. Bruhm, N. A. Vulpescu, J. E. Medina, D. Mathios, S. Cristiano, N. Niknafs, H. T. Luu, M. G. Goggins, R. A. Anders, J. Sun, S. H. Meta, D. L. Thomas, G. D. Kirk, V. Adleff, J. Phallen, R. B. Scharpf, A. K. Kim, V. E. Velculescu, Detecting Liver Cancer Using Cell-Free DNA Fragmentomes. *Cancer Discov.* **13**, 616–631 (2023).

22. S. Li, H. Lai, J. Liu, Y. Liu, L. Jin, Y. Li, F. Liu, Y. Gong, Y. Guan, X. Yi, Q. Shi, Z. Cai, Q. Li, Y. Li, M. Zhu, J. Wang, Y. Yang, W. Wei, D. Yin, E. Song, Q. Liu, Circulating Tumor DNA Predicts the Response and Prognosis in Patients With Early Breast Cancer Receiving Neoadjuvant Chemotherapy. *JCO Precis Oncol* **4** (2020), doi:10.1200/PO.19.00292.

23. S. B. Goldberg, A. Narayan, A. J. Kole, R. H. Decker, J. Teysir, N. J. Carriero, A. Lee, R. Nemati, S. K. Nath, S. M. Mane, Y. Deng, N. Sukumar, D. Zeltermann, D. J. Boffa, K. Politi, S. N. Gettinger, L. D. Wilson, R. S. Herbst, A. A. Patel, Early Assessment of Lung Cancer Immunotherapy Response via Circulating Tumor DNA. *Clin. Cancer Res.* **24**, 1872–1880 (2018).

24. S. Weber, P. van der Leest, H. C. Donker, T. Schlange, W. Timens, M. Tamminga, S. O. Hasenleithner, R. Graf, T. Moser, B. Spiegl, M.-L. Yaspo, L. W. M. M. Terstappen, G. Sidorenkov, T. J. N. Hiltermann, M. R. MEDINA ET AL., PAGE 13 Speicher, E. Schuurin, E. Heitzer, H. J. M. Groen, Dynamic Changes of Circulating Tumor DNA Predict Clinical Outcome in Patients With Advanced Non-Small-Cell Lung Cancer Treated With Immune Checkpoint Inhibitors. *JCO Precision Oncology*, 1540–1553 (2021).

25. W. J. Gradishar, M. S. Moran, J. Abraham, R. Aft, D. Agnese, K. H. Allison, B. Anderson, H. J. Burstein, H. Chew, C. Dang, A. D. Elias, S. H. Giordano, M. P. Goetz, L. J. Goldstein, S. A. Hurvitz, S. J. Isakoff, R. C. Jankowitz, S. H. Javid, J. Krishnamurthy, M. Leitch, J. Lyons, J. Mortimer, S. A. Patel, L. J. Pierce, L. H. Rosenberger, H. S. Rugo, A. Sitapati, K. L. Smith, M. L. Smith, H. Soliman, E. M. Stringer-Reasor, M. L. Telli, J. H. Ward, K. B. Wisinski, J. S. Young, J. Burns, R. Kumar, Breast Cancer, Version 3.2022, NCCN Clinical Practice Guidelines in Oncology. *J. Natl. Compr. Canc. Netw.* **20**, 691–722 (2022).
26. L. A. Niemeier, D. J. Dabbs, S. Beriwal, J. M. Striebel, R. Bhargava, Androgen receptor in breast cancer: expression in estrogen receptor-positive tumors and in estrogen receptor-negative tumors with apocrine differentiation. *Mod. Pathol.* **23**, 205–212 (2010).
27. L. C. Collins, K. S. Cole, J. D. Marotti, R. Hu, S. J. Schnitt, R. M. Tamimi, Androgen receptor expression in breast cancer in relation to molecular phenotype: results from the Nurses' Health Study. *Mod. Pathol.* **24**, 924–931 (2011).
28. K. H. Kensler, M. M. Regan, Y. J. Heng, G. M. Baker, M. E. Pyle, S. J. Schnitt, A. Hazra, R. Kammler, B. Thürlimann, M. Colleoni, G. Viale, M. Brown, R. M. Tamimi, Prognostic and predictive value of androgen receptor expression in postmenopausal women with estrogen receptor-positive breast cancer: results from the Breast International Group Trial 1-98. *Breast Cancer Res.* **21**, 30 (2019).
29. M. Takaku, S. A. Grimm, P. A. Wade, GATA3 in Breast Cancer: Tumor Suppressor or Oncogene? *Gene Expr.* **16**, 163–168 (2015).
30. D. D. Seachrist, L. J. Anstine, R. A. Keri, FOXA1: A Pioneer of Nuclear Receptor Action in Breast Cancer. *Cancers* **13** (2021), doi:10.3390/cancers13205205.
31. L. J. Anstine, P. R. Majmudar, A. Aponte, S. Singh, R. Zhao, K. L. Weber-Bonk, F. W. Abdul-Karim, M. Valentine, D. D. Seachrist, K. E. Grennel-Nickelson, L. Cuellar-Vite, G. M. Sizemore, S. T. Sizemore, B. M. Webb, C. L. Thompson, R. A. Keri, TLE3 Sustains Luminal Breast Cancer Lineage Fidelity to Suppress Metastasis. *Cancer Res.* **83**, 997–1015 (2023).
32. M. Hwang, J. V. Canzoniero, S. Rosner, G. Zhang, J. R. White, Z. Belcaid, C. Cherry, A. Balan, G. Pereira, A. Curry, N. Niknafs, J. Zhang, K. N. Smith, L. Sivapalan, J. E. Chaft, J. E. Reuss, K. Marrone, J. C. Murray, Q. K. Li, V. Lam, B. P. Levy, C. Hann, V. E. Velculescu, J. R. Brahmer, P. M. Forde, T. Seiwert, V. Anagnostou, Peripheral blood immune cell dynamics reflect antitumor immune responses and predict clinical response to immunotherapy. *J. Immunother. Cancer* **10**, e004688 (2022).
33. M. Moschetta, M. Uccello, B. Kasenda, G. Mak, A. McClelland, S. Boussios, M. Forster, H.-T. Arkenau, Dynamics of Neutrophils-to-Lymphocyte Ratio Predict Outcomes of PD-1/PD-L1 Blockade. *Biomed Res. Int.* **2017**, 1506824 (2017).
34. S. Wen, N. Chen, Y. Hu, L. Huang, J. Peng, M. Yang, X. Shen, Y. Song, L. Xu, Elevated peripheral absolute monocyte count related to clinicopathological features and poor prognosis in solid tumors: Systematic review, meta-analysis, and meta-regression. *Cancer Med.* **10**, 1690–1714 (2021). MEDINA ET AL., PAGE 14
35. J. D. Cohen, L. Li, Y. Wang, C. Thoburn, B. Afsari, L. Danilova, C. Douville, A. A. Javed, F. Wong, A. Mattox, R. H. Hruban, C. L. Wolfgang, M. G. Goggins, M. Dal Molin, T.-L. Wang, R. Roden, A. P. Klein, J. Ptak, L. Dobbryn, J. Schaefer, N. Silliman, M. Popoli, J. T. Vogelstein, J.

D. Browne, R. E. Schoen, R. E. Brand, J. Tie, P. Gibbs, H.-L. Wong, A. S. Mansfield, J. Jen, S. M. Hanash, M. Falconi, P. J. Allen, S. Zhou, C. Bettgowda, L. A. Diaz Jr, C. Tomasetti, K. W. Kinzler, B. Vogelstein, A. M. Lennon, N. Papadopoulos, Detection and localization of surgically resectable cancers with a multi-analyte blood test. *Science* **359**, 926–930 (2018).

36. J. J. Elmer, M. D. Christensen, S. Barua, J. Lehrman, K. A. Haynes, K. Rege, The histone deacetylase inhibitor Entinostat enhances polymer-mediated transgene expression in cancer cell lines. *Biotechnol. Bioeng.* **113**, 1345–1356 (2016).

37. P. Schmid, J. Cortes, R. Dent, L. Pusztai, H. McArthur, S. Kümmel, J. Bergh, C. Denkert, Y. H. Park, R. Hui, N. Harbeck, M. Takahashi, M. Untch, P. A. Fasching, F. Cardoso, J. Andersen, D. Patt, M. Danso, M. Ferreira, M. A. Mouret-Reynier, S. A. Im, J. H. Ahn, M. Gion, S. Baron-Hay, J. F. Boileau, Y. Ding, K. Tryfonidis, G. Aktan, V. Karantza, J. O'Shaughnessy, Investigators, KEYNOTE-, Event-free Survival with Pembrolizumab in Early Triple-Negative Breast Cancer. *N. Engl. J. Med.* **386**, 556–567 (2022).

38. J. Chèneby, Z. Ménétrier, M. Mestdagh, T. Rosnet, A. Douida, W. Rhalloussi, A. Bergon, F. Lopez, B. Ballester, ReMap 2020: a database of regulatory regions from an integrative analysis of Human and Arabidopsis DNA-binding sequencing experiments. *Nucleic Acids Res.* **48**, D180–D188 (2020).

39. M. E. Ritchie, B. Phipson, D. Wu, Y. Hu, C. W. Law, W. Shi, G. K. Smyth, limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).

40. J.-P. Fortin, K. D. Hansen, Reconstructing A/B compartments as revealed by Hi-C using long-range correlations in epigenetic data. *Genome Biol.* **16**, 180 (2015).

41. V. A. Adalsteinsson, G. Ha, S. S. Freeman, A. D. Choudhury, D. G. Stover, H. A. Parsons, G. Gydush, S. C. Reed, D. Rotem, J. Rhoades, D. Loginov, D. Livitz, D. Rosebrock, I. Leshchiner, J. Kim, C. Stewart, M. Rosenberg, J. M. Francis, C.-Z. Zhang, O. Cohen, C. Oh, H. Ding, P. Polak, M. Lloyd, S. Mahmud, K. Helvie, M. S. Merrill, R. A. Santiago, E. P. O'Connor, S. H. Jeong, R. Leeson, R. M. Barry, J. F. Kramkowski, Z. Zhang, L. Polacek, J. G. Lohr, M. Schleicher, E. Lipscomb, A. Saltzman, N. M. Oliver, L. Marini, A. G. Waks, L. C. Harshman, S. M. Tolane, E. M. Van Allen, E. P. Winer, N. U. Lin, M. Nakabayashi, M.-E. Taplin, C. M. Johannessen, L. A. Garraway, T. R. Golub, J. S. Boehm, N. Wagle, G. Getz, J. C. Love, M. Meyerson, Scalable whole-exome sequencing of cell-free DNA reveals high concordance with metastatic tumors. *Nat. Commun.* **8**, 1–13 (2017).

OTHER EMBODIMENTS

[00168] From the foregoing description, it will be apparent that variations and modifications may be made to the disclosure described herein to adopt it to various usages and conditions. Such embodiments are also within the scope of the following claims.

[00169] All citations to sequences, patents and publications in this specification are herein incorporated by reference to the same extent as if each independent patent and publication was specifically and individually indicated to be incorporated by reference. By their citation of various

references in this document, Applicants do not admit any particular reference is “prior art” to their disclosure.

What is claimed:

1. A method of determining responsiveness to a cancer therapy in a subject, comprising:
 - a) obtaining a biological sample from the subject at various times after administration of the cancer therapy;
 - b) isolating cell free DNA (cfDNA) from the subject's biological samples;
 - c) conducting low coverage whole genome sequencing of the subject's isolated cfDNA to generate genomic libraries of sequenced cfDNA fragments;
 - d) determining a cfDNA fragmentome profile of the subject's sequenced DNA fragments comprising evaluating one or more cfDNA fragment characteristics;
 - e) evaluating the cfDNA fragmentome profiles of a reference panel from subjects without cancer to the subject's cfDNA fragmentome profile prior to administration of a cancer therapy (baseline), at one or more time points during therapy, after therapy or combinations thereof;thereby, determining responsiveness to the cancer therapy in the subject.
2. The method of claim 1, wherein the one or more cfDNA fragment characteristics comprises calculating ratios of small to large cfDNA fragments, sequences of one or more cfDNAs, median cfDNA fragment sizes, fragment size distribution, mutant allele frequencies, fragment length, fragment size distribution, fragment end motifs, preferred end coordinates, breakpoint motifs, methylation frequencies or combinations thereof.
3. The method of claim 2, wherein a small cfDNA fragment comprises about 80 base pairs (bp) to about 150 bp.
4. The method of claim 2 or 3, wherein a large cfDNA fragment comprises about 151 bp to about 300 bp.
5. The method of any of claims 2 to 4, wherein the small to large cfDNA ratios are GC corrected.
6. The method of any of claims 2 to 5, wherein the mutant allele frequencies in the subject's cfDNA fragmentomes are determined utilizing a hierarchical regression model algorithm.

7. The method of any of claims 2 to 6, wherein the cfDNA fragmentome profile comprises the sequence coverage of small cfDNA fragments in windows across the genome.
8. The method of any of claims 2 to 7, wherein the cfDNA fragmentome profile comprises the sequence coverage of large cfDNA fragments in windows across the genome.
9. The method of any of claims 2 to 8, wherein the cfDNA fragmentome profile comprises the sequence coverage of small and large cfDNA fragments in windows across the genome.
10. The method of any of claims 1 to 9, wherein the cfDNA fragmentome profiles in individuals with cancer are altered across the genome prior to treatment and alterations are decreased in subjects responding to the cancer therapy.
11. The method of any of claims 1 to 10, wherein the cancer therapy comprises immune checkpoint inhibition, a surgical therapy, chemotherapy, radiation therapy, cryotherapy, hyperthermia treatment, phototherapy, radioablation therapy, hormonal therapy, immunotherapy, small molecule therapy, receptor kinase inhibitor therapy, anti-angiogenic therapy, cytokine therapy, monoclonal antibodies, siRNA, miRNA, antisense oligonucleotides, ribozymes or gene therapy.
12. The method of any of claims 1 to 11, wherein the cancer therapy comprises an immune checkpoint inhibitor.
13. The method of any of claims 1 to 12, wherein the cancer is a solid tumor.
14. The method of any of claims 1 to 13, wherein the cancer is metastatic.
15. The method of any of claims 1 to 14, further comprising conducting blood cell counts.
16. A method of determining tumor burden in a subject during therapy, comprising:
 - training a machine learning algorithm with cell free DNA (cfDNA) fragmentome profiles of subjects treated with a cancer therapy, wherein cfDNA tumor fraction levels are independently measured for each sample;
 - evaluating cfDNA fragmentome profiles of subjects undergoing cancer therapy utilizing the trained machine learning algorithm to generate a score relative to a reference or baseline score;

wherein the subject's score is decreased as compared to the reference score if the tumor burden is decreased.

17. The method of claim 16, wherein the subject's cfDNA fragmentome profiles are assayed at different time points comprising prior to administration of a cancer therapy (baseline), at one or more time points during therapy, after therapy or combinations thereof.

18. The method of claim 16 or 17, wherein the subject's cfDNA fragmentome profile comprises calculating ratios of small to large cfDNA fragments, sequences of one or more cfDNAs, median cfDNA fragment sizes, fragment size distribution, mutant allele frequencies, fragment length, fragment size distribution, fragment end motifs, preferred end coordinates, breakpoint motifs, methylation frequencies or combinations thereof.

19. The method of claim 18, wherein a small cfDNA fragment comprises about 80 base pairs (bp) to about 150 bp.

20. The method of claim 18 or 19, wherein a large cfDNA fragment comprises about 151 bp to about 300 bp.

21. The method of any of claims 18 to 20, wherein the small to large cfDNA ratios are GC corrected.

22. The method of any of claims 18 to 21, wherein the mutant allele frequencies in the subject's cfDNA fragmentomes are determined utilizing a hierarchical regression model algorithm.

23. The method of any of claims 18 to 22, wherein the cfDNA fragmentome profile comprises the sequence coverage of small cfDNA fragments in windows across the genome.

24. The method of any of claims 18 to 23, wherein the cfDNA fragmentome profile comprises the sequence coverage of large cfDNA fragments in windows across the genome.

25. The method of any of claims 18 to 24, wherein the cfDNA fragmentome profile comprises the sequence coverage of small and large cfDNA fragments in windows across the genome.

26. The method of any of claims 18 to 25, wherein the cfDNA fragmentome profiles in individuals with cancer are altered across the genome prior to treatment and alterations are decreased in subjects responding to the cancer therapy.
27. The method of any of claims 16 to 26, wherein the cancer therapy comprises immune checkpoint inhibition, a surgical therapy, chemotherapy, radiation therapy, cryotherapy, hyperthermia treatment, phototherapy, radioablation therapy, hormonal therapy, immunotherapy, small molecule therapy, receptor kinase inhibitor therapy, anti-angiogenic therapy, cytokine therapy, monoclonal antibodies, siRNA, miRNA, antisense oligonucleotides, ribozymes or gene therapy.
28. The method of any of claims 16 to 27, wherein the cancer therapy comprises an immune checkpoint inhibitor.
29. The method of any of claims 16 to 28, wherein the cancer is a solid tumor.
30. The method of any of claims 16 to 29, wherein the cancer is metastatic.
31. The method of any of claims 16 to 30, further comprising conducting blood cell counts.
32. A method of determining responsiveness to a cancer therapy in a subject, during therapy, comprising:
- obtaining a biological sample from the subject at various times after administration of the cancer therapy;
 - isolating cell free DNA (cfDNA) from the subject's biological samples;
 - conducting blood cell counts prior to administration of therapy, during administration of therapy and after therapy;
 - conducting low coverage whole genome sequencing of the subject's isolated cfDNA to generate genomic libraries of sequenced cfDNA fragments;
 - determining a cfDNA fragmentome profile of the subject's sequenced DNA fragments comprising evaluating one or more cfDNA fragment characteristics;

evaluating the cfDNA fragmentome profiles of a reference panel from subjects without cancer to the subject's cfDNA fragmentome profile prior to administration of a cancer therapy (baseline), at one or more time points during therapy, after therapy or combinations thereof;

analyzing blood cell counts and comparing blood cell counts to the subject's cfDNA fragmentation profile,

wherein the subject's score is decreased as compared to the reference score if the tumor burden is decreased.

33. The method of claim 32, wherein analysis of differences between baseline and longitudinal timepoints of fragmentation profiles and blood counts to evaluate molecular responses.

34. The method of claim 33, wherein results obtained from blood cell counts and fragmentation profiles clinically classify patients that have stable disease and to determine whether the therapy is of benefit to the subject.

35. The method of any one of claims 32-34, further comprising RECIST classification over course of cancer therapy to further identify tumor responsiveness.

36. A method of diagnosing and differentiating between breast cancer subtypes, comprising: obtaining a biological sample from the subject at various times after administration of the cancer therapy;

isolating cell free DNA (cfDNA) from the subject's biological samples;

conducting low coverage whole genome sequencing of the subject's isolated cfDNA to generate genomic libraries of sequenced cfDNA fragments;

determining a cfDNA fragmentome profile of the subject's sequenced DNA fragments comprising evaluating one or more cfDNA fragment characteristics;

evaluating the cfDNA fragmentome profiles of a reference panel from subjects without cancer

assay low coverage whole genome sequences and transcription binding sites,

thereby diagnosing and differentiating between breast cancer subtypes.

37. The method of claim 36, wherein changes in transcription factor binding site coverages is diagnostic of a breast cancer subtype.

38. The method of any one of claims 35-37, wherein assaying of low coverage whole genome sequences and transcription binding sites is conducted at baseline and later timepoints in response to therapy.

39. The method of claim 38, wherein the transcription binding sites comprise : ESR1, AR, GATA3, FOXA1 or TLE3, or combinations thereof.