have to be identified before cancer can be detected. Thus, only screening for well-known point mutations is possible for screening, or the specific point mutations for a patient must be identified through previous invasive biopsied for a previously detector tumor.

SUMMARY

[0006] Examples can analyze cell-free DNA for disease detection (e.g., cancer) of a particular organ. Some examples may use an organ-associated sample that is from a particular organ or passes through the particular organ, as may occur, for example, in urine, saliva, blood, and stool samples. The analysis may be performed in various ways.

[0006a] In one aspect, the present disclosure refers to a method of analyzing a biological sample of an organism, the biological sample including a mixture of cell-free DNA molecules from a plurality of tissues types, including a first tissue type, the method comprising:

identifying N genomic sites, N being an integer greater than or equal to 10;

for each of M tissue types:

obtaining N tissue-specific methylation levels at the N genomic sites, N being greater than or equal to M, wherein the tissue-specific methylation levels form a matrix A of dimensions N by M, wherein one of the M tissue types corresponds to a first diseased tissue type corresponding to a first disease of a first organ;

analyzing, by a computer system, a plurality of cell-free DNA molecules from the biological sample, the plurality of cell-free DNA molecules being at least 1,000 cell-free DNA molecules, wherein analyzing a cell-free DNA molecule includes:

identifying a location of the cell-free DNA molecule in a reference genome corresponding to the organism;

identifying a set of the plurality of cell-free DNA molecules that are each located at any one of the N genomic sites,

measuring N mixture methylation levels at the N genomic sites using the set of the plurality of cell-free DNA molecules;

determining, by the computer system, a first fractional contribution of the first diseased tissue type in the mixture using the N mixture methylation levels and the N tissue-specific methylation levels of each of the M tissue types; and

using the first fractional contribution of the first diseased tissue type to determine a level of the first disease for the first organ in the organism.

[0006b] In one aspect, the present disclosure refers to a method of analyzing a urine sample of an organism, the urine sample including DNA originating from normal cells and potentially from cells associated with cancer, wherein at least some of the DNA is cell-free in the urine sample, the method comprising:

for each size of a plurality of sizes:

measuring an amount of DNA fragments from the urine sample corresponding to the size;

calculating a first value of a first parameter based on the amounts of DNA fragments at multiple sizes, the first parameter providing a statistical measure of a size profile of DNA fragments in the urine sample, wherein the first parameter increases with increasing size of the DNA fragments;

comparing the first value to a reference value; and

determining a classification of a level of bladder cancer based on the comparing.

[0006c] In one aspect, the present disclosure refers to a method of analyzing a urine sample from a renal pelvis of a kidney of an organism, the urine sample comprising DNA, wherein at least some of the DNA is cell-free, the method comprising:

for each size of a plurality of sizes:

measuring an amount of DNA fragments from the urine sample corresponding to the size;

calculating a first value of a first parameter based on the amounts of DNA fragments at multiple sizes, the first parameter providing a statistical measure of a size profile of DNA fragments in the urine sample;

comparing the first value to a reference value; and

determining a classification of a level of inflammation in the kidney based on the comparing.

[0006d] In another aspect, the present disclosure refers to a method of identifying cancer in a first organ of an organism by analyzing a first sample and a blood sample of the organism, the first sample and the blood sample both including DNA originating from normal cells and

potentially from cells associated with cancer, wherein the first sample is from a first organ or passes through the first organ as the first sample exits the organism and is different from the blood sample, and wherein at least some of the DNA is cell-free in both the first sample and the blood sample, the method comprising:

analyzing a plurality of DNA molecules from the first sample and the blood sample, wherein analyzing a DNA molecule includes:

identifying a location of the DNA molecule in a genome of the organism, and optionally determining whether the DNA molecule is methylated at one or more sites;

for each chromosomal region of a plurality of chromosomal regions of the organism:

determining a classification of whether the chromosomal region exhibits an aberration of at least one of a copy number aberration or a methylation aberration for each of the first sample and the blood sample by:

identifying a respective group of DNA molecules from the respective sample as being from the chromosomal region based on the identified locations, the respective group including at least one DNA molecule located at each of a plurality of loci of the chromosomal region;

calculating, with a computer system, a respective value of the respective group of DNA molecules, the respective value defining a property of the DNA molecules of the respective group, the property being at least one of a copy number or a methylation level;

comparing the respective value to a reference value;

determining a first level of cancer based on whether a first amount of chromosomal regions classified as exhibiting an aberration for the first sample is above a first threshold;

determining a second level of cancer based on whether a second amount of chromosomal regions classified as exhibiting an aberration for the blood sample is above a second threshold; and

determining that the organism has cancer of the first organ when the first level of cancer indicates that the organism has cancer and the second level of cancer indicates that the organism does not have cancer.

[0006e] In another aspect, the present disclosure refers to a method of analyzing a urine sample of an organism, the urine sample including DNA originating from normal cells and potentially from cells associated with cancer, wherein at least some of the DNA is cell-free in the urine sample, the method comprising:

analyzing a plurality of DNA molecules from the urine sample, wherein analyzing a DNA molecule includes:

identifying a location of the DNA molecule in a genome of the organism;

for each chromosomal region of a plurality of chromosomal regions of the organism:

determining a classification of whether the chromosomal region exhibits an aberration of at least one of a copy number aberration or a methylation aberration by:

identifying a group of DNA molecules from the urine sample as being from the chromosomal region based on the identified locations, the group including at least one DNA molecule located at each of a plurality of loci of the chromosomal region;

calculating, with a computer system, a value of the group of DNA molecules, the value defining a property of the DNA molecules of the group, the property being at least one of a copy number or a methylation level;

comparing the value to a reference value;

determining a first level of cancer based on whether a first amount of chromosomal regions classified as exhibiting a copy number aberration for the urine sample is above a first threshold;

determining a second level of cancer based on whether a second amount of chromosomal regions classified as exhibiting the methylation aberration for the urine sample is above a second threshold;

determining a third level of cancer based on whether a fractional contribution of a tumor tissue is above a third threshold;

determining that the organism has cancer when at least one of the first level of cancer, the second level of cancer, or the third level of cancer indicate that the organism has cancer.

[0006f] In another aspect, the present disclosure refers to a method of analyzing a biological sample of an organism, the biological sample including DNA originating from normal cells and

potentially from cells associated with cancer, wherein at least some of the DNA is cell-free in the biological sample, the method comprising:

receiving a plurality of sequence reads corresponding to DNA molecules of the biological sample;

determining a genomic position for the sequence reads,

comparing the sequence reads to a reference genome to determine sequence reads that have one mismatch with the reference genome determining a parameter based on a count of the sequence reads having one mismatch with the reference genome ;

comparing the parameter to a threshold value;

determining a classification of a level of cancer using the comparison of the parameter to the threshold value.

[0006g] In another aspect, the present disclosure refers to a method of analyzing a urine sample of a human, the urine sample including DNA originating from normal cells and potentially from cells associated with cancer, wherein at least some of the DNA is cell-free in the urine sample, the method comprising:

obtaining a plurality of sequence reads corresponding to DNA molecules of the urine sample;

for each sequence read of the plurality of sequence reads:

classifying the sequence read as a human read when the sequence read aligns, by a computer system, to a human reference genome, or

classifying the sequence read as a first pathogen read when the sequence read aligns to a first pathogen reference genome that corresponds to a species or a genus of a first pathogen;

determining a parameter based on an amount of first pathogen reads;

comparing the parameter to a cutoff value, wherein the cutoff value is determined from a first set of reference samples that have bladder cancer and a second set of control samples that do not have bladder cancer; and

determining a classification of a level of bladder cancer using the comparison.

[0006h] In another aspect, the present disclosure refers to a computer product comprising a computer readable medium storing a plurality of instructions for controlling a computer system to perform an operation of any of the methods of the present disclosure.

[0006i] In another aspect, the present disclosure refers to a system comprising:

the computer product of the present disclosure; and

one or more processors for executing instructions stored on the computer readable medium.

[0006j] In another aspect, the present disclosure refers to a system comprising means for performing any of the methods as disclosed herein.

[0006k] In another aspect, the present disclosure refers to a system configured to perform any of the methods as disclosed herein.

[0006l] In another aspect, the present disclosure refers to a system comprising modules that respectively perform the steps of any of the methods as disclosed herein.

[0007] In some examples, methylation levels of cell-free DNA can be measured in a sample. Tissue-specific methylation patterns can be used to determine fractional contributions from different tissue types. One tissue type can be a diseased tissue type of a particular disease of a particular organ. The fractional contribution of the diseased tissue type can be used to determine a level (classification) of the particular disease in the sample. As an example, methylation patterns can be used determine a percentage of cancerous bladder lining cells, and the percentage can be used to determine a level of cancer.

[0008] In other examples, sizes of organ-associated cell-free DNA can be measured. In an example using a urine sample voided from the bladder, a size profile can be measured of cell-free DNA fragments that naturally occur in the urine. A statistical measure of the size profile can indicate that the cell-free DNA fragments are collectively longer than expected for subjects with healthy bladder tissue. The indication of longer fragments can be used to identify bladder cancer in the subject. In another example using a urine sample, the urine sample can be retrieved from the renal pelvis. A determination about whether the kidney is inflamed can be made based on the statistical measure indicating that the cell-free DNA fragments are collectively longer than expected for subjects with a non-inflamed kidney.

[0009] In other examples, two different samples can be analyzed to determine whether a particular organ has cancer. Cell-free DNA in a blood sample and organ-associated sample (e.g., urine, saliva, or a stool sample) can both be analyzed to identify chromosomal regions exhibiting a copy number aberration. If the blood sample does not indicate cancer and the organ-associated sample does, then the subject can be identified as having cancer in an organ associated with the

sample. For example, for a urine sample, the subject can be identified as having bladder cancer. Other cancers can also be detected, such as cancers of the urinary tract, transitional cell carcinomas that arise from the urothelium, and kidney cancers.

[0010] Other examples are directed to systems and computer readable media associated with methods described herein.

[0011] A better understanding of the nature and advantages of embodiments of the present invention may be gained with reference to the following detailed description and the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

[0012] FIG. 1 displays the methylation density in 19,418 differentially methylated regions identified from the kidney, urothelium, B-cells, T-cells, and neutrophils according to embodiments of the present invention. Each vertical line represents a single differentially methylated region color-coded based on methylation density. Differentially methylated regions (DMR) with a high methylation density are represented in red, whereas loci with a low methylation density are represented in yellow.

[0013] FIG. 2 is a heat map showing the methylation densities of 27,371 differentially methylated regions (DMR) in bladder tumor, B-cells, urothelium, kidney, neutrophils, and T-cells according to embodiments of the present invention.

[0014] FIG. 3 shows the proportional contribution of cfDNA derived from the blood cells and kidney in 31 urine samples from and hematopoietic stem cell transplantation (HSCT) (triangles) and renal transplant (round dots) patients according to embodiments of the present invention.

[0015] FIGS. 4A-4C show plots of methylation deconvolution for bladder cancer cases and cancer-free controls for fractions defined using different tissue-specific methylation patterns, including normal urothelium cells (FIG. 4A), urothelium and bladder cancer cells as a single tissue-specific methylation patterns (FIG. 4B), and urothelium and bladder cancer cells as separate tissue-specific methylation patterns (FIG. 4C) according to embodiments of the present invention.