[0004] The degree of jagged ends, which may be the quantity or the length of jagged ends, in a sample may reflect the level of a condition in an individual. For example, the degree of jagged ends may be related to a disease, a disorder, a pregnancy-related condition. The jagged ends may be used to determine the fractional concentration of clinically-relevant DNA in a sample. The age of an individual may be related to the degree of jagged ends. Jagged ends from specific tissues may be analyzed, and the degree of jagged ends may determine a level of cancer.

[0005] The degree of jagged ends may be measured in various ways. For example, the jagged ends may be repaired using methylated or unmethylated nucleotides, and the resulting change in the level of methylation can indicate the presence and/or length of a jagged end. In some cases, methylated cytosines can be used in end repair to measure the exact length of a jagged end. As another example, the degree of jagged ends may also be determined by aligning portions of the fragments to a reference genome or a complementary strand or measuring other signals from nucleotides added through end repair.

[0005a] In one aspect, the present disclosure refers to a method of analyzing a biological sample obtained from an individual, the method comprising:

analyzing a plurality of cell-free nucleic acid molecules of the biological sample, wherein each cell-free nucleic acid molecule of the plurality of cell-free nucleic acid molecules is double-stranded with a first strand having a first portion and a second strand, wherein the first portion of the first strand of at least some of the plurality of cell-free nucleic acid molecules has no complementary portion from the second strand, is not hybridized to the second strand, and is at a first end of the first strand, and wherein the analyzing comprises for each cell-free nucleic acid molecule of the plurality of cell-free nucleic acid molecules:

measuring a property that is proportional to a length of the first strand that overhangs the second strand, wherein the property is measured by analyzing the first strand, the second strand, or the first strand and the second strand;

determining a jagged end value using the measured properties of the plurality of cell-free nucleic acid molecules, wherein the jagged end value provides a collective measure that a strand overhangs another strand in the plurality of cell-free nucleic acid molecules;

comparing the jagged end value to a reference value, wherein the reference value is determined using at least one reference sample selected from the group consisting of one or more reference samples of subjects that have a condition, and one or more reference samples of subject that do not have the condition; and

determining ~~a~~ the level of a condition of the individual based on the comparison, wherein the condition is a cancer or systemic lupus erythematosus (SLE)~~an auto-immune disease~~.

[0005b] In another aspect, the present disclosure refers to a method of determining a fraction of clinically-relevant DNA in a biological sample obtained from an individual, the method comprising:

analyzing a plurality of cell-free nucleic acid molecules of the biological sample, wherein each cell-free nucleic acid molecule of the plurality of cell-free nucleic acid molecules is double-stranded with a first strand having a first portion and a second strand, wherein the first portion of the first strand of at least some of the plurality of cell-free nucleic acid molecules has no complementary portion from the second strand, is not hybridized to the second strand, and is at a first end of the first strand, and wherein the analyzing comprises for each cell-free nucleic acid molecule of the plurality of cell-free nucleic acid molecules:

measuring a property that is proportional to a length of the first strand that overhangs the second strand, wherein the property is measured by analyzing the first strand, the second strand, or the first strand and the second strand;

determining a jagged end value using the measured properties of the plurality of cell-free nucleic acid molecules, wherein the jagged end value provides a collective measure that a strand overhangs another strand in the plurality of cell-free nucleic acid molecules;

comparing the jagged end value to a reference value; and

determining the fraction of clinically-relevant DNA in the biological sample based on the comparison.

[0005c] In another aspect, the present disclosure refers to a method of determining an age of an individual using a biological sample obtained from the individual, the method comprising:

analyzing a plurality of cell-free nucleic acid molecules of the biological sample, each nucleic acid molecule of the plurality of cell-free nucleic acid molecules being double-

stranded with a first strand having a first portion and a second strand, wherein the first portion of the first strand of at least some of the plurality of cell-free nucleic acid molecules has no complementary portion from the second strand, is not hybridized to the second strand, and is at a first end of the first strand, and wherein the analyzing comprises for each cell-free nucleic acid molecule of the plurality of cell-free nucleic acid molecules:

measuring a property that is proportional to a length of the first strand that overhangs the second strand, wherein the property is measured by analyzing the first strand, the second strand, or the first strand and the second strand;

determining a jagged end value using the measured properties of the plurality of cell-free nucleic acid molecules, wherein the jagged end value provides a collective measure that a strand overhangs another strand in the plurality of cell-free nucleic acid molecules;

comparing the jagged end value to a reference value, wherein the reference value is obtained using nucleic acid molecules from one or more reference subjects having known ages; and

determining the age of the individual based on the comparison.

[0006] 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

[0007] FIG. 1 shows a method of using jagged end values to analyze a biological sample according to embodiments of the present invention.

[0008] FIG. 2 shows one example for assessing the degree of 5' overhangs according to embodiments of the present invention.

[0009] FIG. 3. Illustration of the calculation of methylation levels along a DNA molecule after mapping to the human reference genome according to embodiments of the present invention.

[0010] FIG. 4 shows a method of analyzing a biological sample obtained from an individual to calculate a jagged end value using methylation levels according to embodiments of the present invention.

WHAT IS CLAIMED IS:

1. A method of analyzing a biological sample obtained from an individual, the method comprising:

analyzing a plurality of cell-free nucleic acid molecules of the biological sample, wherein each cell-free nucleic acid molecule of the plurality of cell-free nucleic acid molecules is double-stranded with a first strand having a first portion and a second strand, wherein the first portion of the first strand of at least some of the plurality of cell-free nucleic acid molecules has no complementary portion from the second strand, is not hybridized to the second strand, and is at a first end of the first strand, and wherein the analyzing comprises for each cell-free nucleic acid molecule of the plurality of cell-free nucleic acid molecules:

measuring a property that is proportional to a length of the first strand that overhangs the second strand, wherein the property is measured by analyzing the first strand, the second strand, or the first strand and the second strand;

determining a jagged end value using the measured properties of the plurality of cell-free nucleic acid molecules, wherein the jagged end value provides a collective measure that a strand overhangs another strand in the plurality of cell-free nucleic acid molecules;

comparing the jagged end value to a reference value, wherein the reference value is determined using at least one reference sample selected from the group consisting of one or more reference samples of subjects that have a condition, and one or more reference samples of subject that do not have the condition; and

determining a level of ~~a~~ the condition of the individual based on the comparison, wherein the condition is a cancer or systemic lupus erythematosus (SLE)~~an auto-immune disease~~.

2. The method of claim 1, wherein the first end is a 5' end.

3. The method of claim 1 or 2, further comprising:

measuring sizes of cell-free nucleic acid molecules, wherein the plurality of cell-free nucleic acid molecules has sizes within a specified range.

4. The method of claim 3, wherein:

the plurality of cell-free nucleic acid molecules is a first plurality of cell-free nucleic acid molecules, and

the specified range is a first specified range,

the method further comprising:

measuring the property of a strand of each cell-free nucleic acid molecule of a second plurality of cell-free nucleic acid molecules, wherein the second plurality of cell-free nucleic acid molecules has sizes with a second specified range,

wherein determining the jagged end value comprises calculating a ratio using the measured properties of the first plurality of cell-free nucleic acid molecules and the measured properties of the second plurality of cell-free nucleic acid molecules.

5. The method of claim 3, wherein the specified range is 140 to 160 bp.

~~5~~6. The method of claim 1 or 2, wherein the property is a methylation status at one or more sites at end portions of the first strands, the second strands, or the first strands and the second strands of each of the plurality of cell-free nucleic acid molecules, and wherein the jagged end value includes a methylation level over the plurality of cell-free nucleic acid molecules at one or more sites of end portions of the first strands, the second strands, or the first strands and the second strands.

~~6~~7. The method of claim ~~5~~6, wherein a higher methylation level is correlated with a longer length of the first strand that overhangs the second strand.

~~7~~8. The method of claim 1 or 2, further comprising:

analyzing the plurality of cell-free nucleic acid molecules to produce reads,

aligning the reads to a reference genome,

wherein:

the cell-free nucleic acid molecules each have reads within a certain distance range relative to a transcription start site.

~~8~~9. The method of any one of claims 1 to ~~7~~8, wherein the measured property is length.

10. The method of any one of claims 1 to 8, wherein the measured property is a methylation status at one or more sites.

~~9~~11. The method of claim 1, wherein the reference value is determined using one or more reference samples of subjects that have the condition.

~~10~~12. The method of claim 1, wherein the reference value is determined using one or more reference samples of subjects that do not have the condition.

~~11~~13. The method of claim 1, -wherein a machine learning model is used to perform the comparing of the jagged end value to the reference value and the determining of the level of the condition of the individual.

14. The method of claim 1, wherein the condition is the cancer.

15. The method of claim 14, wherein the cancer is hepatocellular carcinoma or colorectal cancer.

16. The method of claim 1, wherein the condition is SLE.

~~12~~17. A method of determining a fraction of clinically-relevant DNA in a biological sample obtained from an individual, the method comprising:

analyzing a plurality of cell-free nucleic acid molecules of the biological sample, wherein each cell-free nucleic acid molecule of the plurality of cell-free nucleic acid molecules is double-stranded with a first strand having a first portion and a second strand, wherein the first portion of the first strand of at least some of the plurality of cell-free nucleic acid molecules has no complementary portion from the second strand, is not hybridized to the second strand, and is at a first end of the first strand, and wherein the analyzing comprises for each cell-free nucleic acid molecule of the plurality of cell-free nucleic acid molecules:

measuring a property that is proportional to a length of the first strand that overhangs the second strand, wherein the property is measured by analyzing the first strand, the second strand, or the first strand and the second strand;

determining a jagged end value using the measured properties of the plurality of cell-free nucleic acid molecules, wherein the jagged end value provides a collective measure that a strand overhangs another strand in the plurality of cell-free nucleic acid molecules;

comparing the jagged end value to a reference value; and

determining the fraction of clinically-relevant DNA in the biological sample based on the comparison.

~~13~~18. The method of claim ~~12~~17, further comprising:

treating the plurality of cell-free nucleic acid molecules by a library preparation protocol before measuring the property,

wherein:

the reference value is obtained using nucleic acid molecules from one or more reference subjects having a known fraction of clinically-relevant DNA, and

the nucleic acid molecules from the one or more reference subjects are treated by the library preparation protocol.

~~14~~19. The method of claim ~~12~~17, wherein the clinically-relevant DNA comprises fetal DNA.

20. The method of claim 17, wherein the clinically-relevant DNA comprises, tumor DNA.

21. The method of claim 17, wherein the clinically-relevant DNA comprises~~,~~ ~~or~~ transplant DNA.

~~15~~22. The method of claim ~~12~~17, wherein the reference value is determined from one or more calibration samples having a known fraction of clinically-relevant DNA and whose jagged end value has been measured.

~~16~~23. The method of claim ~~13~~18, wherein the reference value is determined from a calibration curve that is fit to calibration data points of a plurality of calibration samples, each of the calibration data points including a measured jagged end value and a measured fraction of clinically-relevant DNA of one of the plurality of calibration samples.

~~17~~24. A method of determining an age of an individual using a biological sample obtained from the individual, the method comprising:

analyzing a plurality of cell-free nucleic acid molecules of the biological sample, each nucleic acid molecule of the plurality of cell-free nucleic acid molecules being double-stranded with a first strand having a first portion and a second strand, wherein the first portion of the first strand of at least some of the plurality of cell-free nucleic acid molecules has no complementary portion from the second strand, is not hybridized to the second strand, and is at a first end of the first strand, and wherein the analyzing comprises for each cell-free nucleic acid molecule of the plurality of cell-free nucleic acid molecules:

measuring a property that is proportional to a length of the first strand that overhangs the second strand, wherein the property is measured by analyzing the first strand, the second strand, or the first strand and the second strand;

determining a jagged end value using the measured properties of the plurality of cell-free nucleic acid molecules, wherein the jagged end value provides a collective measure that a strand overhangs another strand in the plurality of cell-free nucleic acid molecules;

comparing the jagged end value to a reference value, wherein the reference value is obtained using nucleic acid molecules from one or more reference subjects having known ages; and

determining the age of the individual based on the comparison.

~~18. The method of claim 17, wherein the reference value is obtained using nucleic acid molecules from one or more reference subjects having known ages.~~

~~19~~25. The method of claim ~~17~~24, wherein the reference value is determined from a calibration curve that is fit to calibration data points of a plurality of calibration samples, each of the calibration data points including a measured jagged end value and a known age of a subject from which one of the plurality of calibration samples was obtained.