

[0003a] In one aspect, the present invention refers to a method of analyzing a biological sample from a female subject pregnant with a fetus, the biological sample including cell-free DNA molecules from the female subject and the fetus, the method comprising: analyzing the cell-free DNA molecules from the biological sample, wherein analyzing a cell-free DNA molecule includes: determining a location of the cell-free DNA molecule in a genome, wherein the genome corresponds to the fetus or the female subject; and determining whether the cell-free DNA molecule is methylated at one or more sites; for each of the one or more sites: determining a respective number of cell-free DNA molecules that are methylated at the site; calculating a measured methylation level of cell-free DNA molecules in the biological sample based on the respective numbers of cell-free DNA molecules methylated at the one or more sites; obtaining one or more calibration data points, wherein each calibration data point specifies a gestational age corresponding to a calibration methylation level, and wherein the one or more calibration data points are determined from a plurality of calibration samples obtained during pregnancy with known gestational ages and including cell-free DNA molecules; comparing the measured methylation level to a calibration methylation level of at least one calibration data point; and estimating a gestational age of the fetus based on the comparing.

[0003b] In another aspect, the present invention refers to a method of analyzing a biological sample from a female subject pregnant with a fetus, the biological sample including cell-free DNA molecules from the female subject and the fetus, the method comprising: analyzing the cell-free DNA molecules from the biological sample, wherein analyzing each cell-free DNA molecule of the cell-free DNA molecules includes: determining a genomic position in a reference genome corresponding to at least one end of the cell-free DNA molecule; measuring a sample vector for the biological sample, wherein each value of the sample vector corresponds to a frequency that the cell-free DNA molecules end on a respective site of a set of sites, wherein the set of sites is determined based on numbers of cell-free DNA molecules that end on various sites in one or more first calibration samples that include maternal-derived DNA molecules and fetal-derived DNA molecules; obtaining a plurality of representative vectors, wherein each representative vector corresponds to a gestational age and includes representative frequencies of cell-free DNA molecules that end on the set of sites, and wherein the plurality of representative vectors are determined from one or more second calibration samples obtained during pregnancy with known gestational ages; comparing the sample vector to the plurality of representative vectors; and estimating the gestational age of the fetus based on the comparing.

CLAIMS:

1. A method of analyzing a biological sample from a female subject pregnant with a fetus, the biological sample including cell-free DNA molecules from the female subject and the fetus, the method comprising:

analyzing the cell-free DNA molecules from the biological sample, wherein analyzing a cell-free DNA molecule includes:

determining a location of the cell-free DNA molecule in a genome, wherein the genome corresponds to the fetus or the female subject; and

determining whether the cell-free DNA molecule is methylated at one or more sites;

for each of the one or more sites:

determining a respective number of cell-free DNA molecules that are methylated at the site;

calculating a measured methylation level of cell-free DNA molecules in the biological sample based on the respective numbers of cell-free DNA molecules methylated at the one or more sites;

obtaining one or more calibration data points, wherein each calibration data point specifies a gestational age corresponding to a calibration methylation level, and wherein the one or more calibration data points are determined from a plurality of calibration samples obtained during pregnancy with known gestational ages and including cell-free DNA molecules;

comparing the measured methylation level to a calibration methylation level of at least one calibration data point; and

estimating a gestational age of the fetus based on the comparing.

2. The method of claim 1, wherein the one or more calibration data points are determined by:

for each calibration sample of the plurality of calibration samples:

determining a calibration methylation level based on numbers of cell-free DNA molecules methylated at the one or more sites; and

determining the one or more calibration data points based on the known gestational ages and the calibration methylation levels of the plurality of calibration samples.

3. The method of claim 1, further comprising:
generating an alarm message when the estimated gestational age does not match a gestational age determined using another technique.

4. The method of claim 3, wherein the other technique includes an ultrasound-based method or a last menstrual period-based method.

5. The method of claim 1, further comprising:
identifying a plurality of fetal DNA molecules from the biological sample, each fetal DNA molecule having a fetal-specific allele,
wherein analyzing the cell-free DNA molecules from the biological sample includes analyzing only the plurality of fetal DNA molecules.

6. The method of claim 1, wherein calculating the measured methylation level of cell-free DNA molecules in the biological sample is based on a fraction of fetal DNA molecules in the biological sample.

7. The method of claim 1, further comprising:
determining a measured statistical value based on counts of cell-free DNA molecules corresponding to various sizes in the biological sample, wherein each calibration data point of the one or more calibration data points specifies the gestational age corresponding to (1) the calibration methylation level and (2) a calibration statistical value determined using the plurality of calibration samples,
wherein comparing the measured methylation level to the calibration methylation level of at least one calibration data point comprises comparing the measured methylation level and the measured statistical value to a calibration methylation level and a calibration statistical value of at least one calibration data point.

8. The method of claim 7, wherein the measured statistical value includes a proportion of cell-free DNA molecules with a size greater than or less than a threshold number of base pairs in the biological sample.

9. The method of claim 1, wherein the measured methylation level is determined by:

$$MD = \frac{M}{M + U},$$

wherein M is a count of cell-free DNA molecules methylated at one or more CpG sites and U is a count of cell-free DNA molecules unmethylated at the one or more CpG sites.

10. The method of claim 1, wherein the plurality of calibration samples includes maternal plasma samples from pregnant female subjects at various pregnancy stages and wherein the biological sample is a plasma sample.

11. The method of claim 1, wherein the one or more sites are a plurality of sites, and wherein the measured methylation level is determined based on the respective numbers of cell-free DNA molecules methylated at the plurality of sites.

12. A method of analyzing a biological sample from a female subject pregnant with a fetus, the biological sample including cell-free DNA molecules from the female subject and the fetus, the method comprising:

analyzing the cell-free DNA molecules from the biological sample, wherein analyzing each cell-free DNA molecule of the cell-free DNA molecules includes:

determining a genomic position in a reference genome corresponding to at least one end of the cell-free DNA molecule;

measuring a sample vector for the biological sample, wherein each value of the sample vector corresponds to a frequency that the cell-free DNA molecules end on a respective site of a set of sites, wherein the set of sites is determined based on numbers of cell-free DNA molecules that end on various sites in one or more first calibration samples that include maternal-derived DNA molecules and fetal-derived DNA molecules;

obtaining a plurality of representative vectors, wherein each representative vector corresponds to a gestational age and includes representative frequencies of cell-free DNA molecules that end on the set of sites, and wherein the plurality of representative vectors are determined from one or more second calibration samples obtained during pregnancy with known gestational ages;

comparing the sample vector to the plurality of representative vectors; and
estimating the gestational age of the fetus based on the comparing.

13. The method of claim 12, further comprising generating an alarm message when the estimated gestational age does not match a gestational age determined using another technique.

14. The method of claim 12, wherein analyzing the cell-free DNA molecules from the biological sample includes analyzing only fetal DNA molecules in the biological sample.

15. The method of claim 14, further comprising:
enriching the fetal DNA molecules based on sizes and methylation levels of the fetal DNA molecules.

16. The method of claim 12, wherein:
the plurality of representative vectors is determined using principal component analysis; and
the plurality of representative vectors are configured to provide largest eigenvalues of a covariance matrix.

17. The method of claim 12, wherein:
the plurality of representative vectors is determined using linear discriminant analysis, logistic regression, machine learning algorithms, k-means clustering, K-nearest neighbors, support vector machine, artificial neural network, or random decision forests; and
the plurality of representative vectors are configured to provide largest eigenvalues of a covariance matrix.

18. The method of claim 12, wherein comparing the sample vector to the plurality of representative vectors includes:
determining a distance between the sample vector and each representative vector of the plurality of representative vectors; or
performing a two or more dimensional principal component analysis on the sample vector and the plurality of representative vectors.

19. The method of claim 12, wherein analyzing the cell-free DNA molecules from the biological sample includes performing a massively parallel paired-end sequencing on the cell-free DNA molecules in the biological sample.

20. The method of claim 12, wherein the set of sites comprises top 0.5% most prevalent cell-free DNA molecule ending positions in the one or more first calibration samples.

21. A computer product comprising a computer readable medium storing a plurality of instructions for controlling a computer system to:

analyze a cell-free DNA molecules from a biological sample from a female subject pregnant with a fetus, the biological sample including cell-free DNA molecules from the female subject and the fetus, wherein analyzing a cell-free DNA molecule includes:

determining a location of the cell-free DNA molecule in a genome,
wherein the genome corresponds to the fetus or the female subject; and

determining whether the cell-free DNA molecule is methylated at one or more sites;

determine, for each of the one or more sites, a respective number of cell-free DNA molecules that are methylated at the site;

calculate a measured methylation level of cell-free DNA molecules in the biological sample based on the respective numbers of cell-free DNA molecules methylated at the one or more sites;

obtain one or more calibration data points, wherein each calibration data point specifies a gestational age corresponding to a calibration methylation level, and wherein the one or more calibration data points are determined from a plurality of calibration samples obtained during pregnancy with known gestational ages and including cell-free DNA molecules;

compare the measured methylation level to a calibration methylation level of at least one calibration data point; and

estimate a gestational age of the fetus based on the comparing.

22. 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 above.

23. A system comprising:
the computer product of claim 22; and
one or more processors for executing instructions stored on the computer readable medium.

24. A computer system including processors that are programmed to implement any of the methods of any of claims above.