[0004] In further examples, the estimated gestational age may be compared with a gestational age determined using another technique for the biological sample, such as the gestational age determined based on medical history or based on ultrasonic diagnosis. An alarm message can be generated when the estimated gestational age does not match the gestational age determined using other techniques. A corrective action may be taken if the mismatch is caused by administrative errors. Further diagnosis and treatment may be conducted if the mismatch may be caused by pathological reasons, such as pregnancy-associated disorders, such as preeclampsia, preterm labor, intrauterine growth restriction (IUGR), or fetal chromosomal aneuploidies.

[0004a] In one aspect, provided herein is 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; obtaining one or more 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 one or more representative vectors are determined from one or more calibration samples with known gestational ages; comparing the sample vector to the one or more representative vectors; and estimating the gestational age of the fetus based on the comparing.

[0004b] In another aspect, provided herein is a computer product comprising a computer readable medium storing a plurality of instructions for controlling a computer system to perform an operation of the methods as disclosed herein.

[0004c] In another aspect, provided herein is a system comprising: the computer product as disclosed herein; and one or more processors for executing instructions stored on the computer readable medium.

[0004d] In another aspect, provided herein is a system comprising means for performing the methods as disclosed herein.

[0004e] In another aspect, provided herein is a system configured to perform the methods as disclosed herein.

[0004f] In another aspect, provided herein is a system comprising modules that respectively perform the methods as disclosed herein.

[0005] Other examples are directed to systems and computer readable media associated with methods described herein.

[0006] A better understanding of the nature and advantages of examples of the present disclosure may be gained with reference to the following detailed description and the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

[0007] FIG. 1 a plot of a size distribution of circulating cell-free DNA in maternal plasma, according to certain embodiments of the present invention.

[0008] FIG. 2A shows a plot of size distributions of DNA fragments in two maternal plasma samples (1st trimester pregnancies) with different fractional fetal DNA concentrations, according to certain embodiments of the present invention. .

[0009] FIG. 2B shows a plot of size distributions of DNA fragments in two maternal plasma samples (2nd trimester pregnancies) with different fractional fetal DNA concentrations, according to certain embodiments of the present invention.

[0010] FIG. 3 is a plot showing a proportion of DNA fragments that have 150 or less base pairs for samples having various fetal DNA percentage in maternal plasma, according to certain embodiments of the present invention.

[0011] FIG. 4A shows a bar chart of percentage of methylated CpG sites in Autosomes among samples collected during pregnancy, according to certain embodiments of the present invention.

WHAT IS CLAIMED IS:

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 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;

obtaining one or more 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 one or more representative vectors are determined from one or more calibration samples with known gestational ages;

comparing the sample vector to the one or more representative vectors; and

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

2. 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.

3. The method of claim 1, wherein analyzing the cell-free DNA molecules from the biological sample includes analyzing only fetal DNA molecules in the biological sample.

4. The method of claim 3, further comprising:

enriching the fetal DNA molecules based on sizes and methylation levels of the fetal DNA molecules.

5. The method of claim 1, wherein:

the one or more representative vectors are determined using principal component analysis, 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 one or more representative vectors are configured to provide largest eigenvalues of a covariance matrix.

6. The method of claim 1, wherein comparing the sample vector to the one or more representative vectors includes:

determining a distance between the sample vector and each representative vector of the one or more representative vectors; or

performing a two or more dimensional principal component analysis on the sample vector and the one or more representative vectors.

7. The method of claim 1, 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.

8. The method of claim 1, wherein the set of sites is determined based on numbers of cell-free DNA molecules that end on various sites in one or more samples.

9. The method of claim 8, wherein the set of sites comprises top 0.5% most prevalent cell-free DNA molecule ending positions in the one or more samples.

10. The method of claim 8, wherein the one or more samples are the one or more calibration samples.

11. The method of any one of claims 1 to 10, wherein the one or more representative vectors comprise a plurality of reference vectors.

12. 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.

13. A system comprising:

the computer product of claim 12; and

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

14. A system comprising means for performing any of the methods above.

15. A system configured to perform any of the above methods.

16. A system comprising modules that respectively perform the steps of any of the above methods.