required to capture a set of regions of interest, for example, through the use of hybridization- or amplicon-based enrichment systems [Sparks AB et al., (2012), *Am J Obstet Gynecol*, 206: 319 e311-319; Liao GJ et al. (2011), *Clin Chem*, 57: 92-101].

[0005] In addition to SNP dependent approaches, SNP-independent methods are also being explored. For example, fragment sizes of maternal plasma DNA can be used for estimating fetal DNA fraction [Yu SC et al. (2014), *Proc Natl Acad Sci U S A*, 111: 8583-8588; Kim SK et al. (2015), Prenatal diagnosis: n/a-n/a] because the fetal-derived DNA are generally shorter than maternal-derived DNA [Lo YMD et al. (2010), *Sci Transl Med*, 2: 61ra91]. However, some other conditions will influence the accuracy of the size-based fetal DNA fraction estimation, for example, systemic lupus erythematosus [Chan RW et al. (2014), *Proc Natl Acad Sci U S A*, 111: E5302-5311]. As an alternative, fetal-specific epigenetic changes, such as methylated RASSF1A and unmethylated SERPINB5 sequences, were demonstrated to be fetal markers for fetal DNA fraction prediction irrespective of genotype information [Chan KC et al. (2006), *Clin Chem*, 52: 2211-2218; Chim SS et al. (2005), *Proc Natl Acad Sci U S A,* 102: 14753-14758]. However, the analytical procedures for quantifying these epigenetic markers involve either bisulfite conversion or digestion with methylation-sensitive restriction enzymes, and may thus potentially affect the precision of these methods.

[0006] Accordingly, it is desirable for new techniques to provide fetal DNA fraction information from maternal plasma.

SUMMARY

[0006a] In one aspect, the present disclosure refers to a method of measuring a fetal DNA fraction in a biological sample of a female pregnant with a fetus, the biological sample including maternal DNA molecules and fetal DNA molecules, the method comprising performing, by a computer system:

identifying a plurality of sites based on sequence information that indicates the female is homozygous for a first allele at each site of the plurality of sites;

obtaining a plurality of reads from DNA molecules of the biological sample;

identifying locations of the plurality of reads in a reference genome;

determining a first value of a first group of DNA molecules based on the plurality of reads, wherein:

each DNA molecule of the first group of DNA molecules includes a read located at a site of the plurality of sites and exhibits a second allele different from the first allele at the site, and

the first value defines a property of the first group of DNA molecules;

determining a second value of a second group of DNA molecules, wherein:

each DNA molecule of the second group of DNA molecules includes a read located at a site of the plurality of sites and exhibits the first allele at the site, and

the second value defines the property of the second group of DNA molecules;

determining a parameter value of a parameter from the first value and the second value;

comparing the parameter value to a calibration point that is determined using at least one other sample having a known fetal DNA fraction and a calibration value corresponding to separate measurement of the parameter in the at least one other sample; and

computing the fetal DNA fraction based on the comparison.

[0006b] In another aspect, the present disclosure refers to a method of measuring a fetal DNA fraction in a biological sample of a female pregnant with a fetus, the biological sample including maternal DNA molecules and fetal DNA molecules, the method comprising performing, by a computer system:

identifying a plurality of sites based on sequence information that indicates the female is homozygous for a first allele at each site of the plurality of sites;

obtaining a plurality of reads from DNA molecules of the biological sample;

identifying locations of the plurality of reads in a reference genome;

determining a first amount of reads, wherein:

each read of the first amount of reads is located at a site of the plurality of sites, and

each read exhibits a second allele different from the first allele at the site;

determining a second amount of reads at the plurality of sites, wherein:

each read of the second amount of reads is located at a site of the plurality of sites, and

each read exhibits the first allele at the site;

determining a non-maternal allele fraction from the first amount and the second amount;

obtaining a calibration point that is determined using another sample having a known fetal DNA fraction and a measured non-maternal allele fraction; and

computing the fetal DNA fraction using the calibration point and the non-maternal allele fraction.

[0006c] In another aspect, the present disclosure refers to a method of measuring a fetal DNA fraction in a biological sample of a female pregnant with a fetus, the biological sample including maternal DNA molecules and fetal DNA molecules, the method comprising performing, by a computer system:

identifying a plurality of sites based on sequence information that indicates the female is homozygous for a first allele at each site of the plurality of sites;

obtaining a plurality of reads from DNA molecules of the biological sample;

identifying locations of the plurality of reads in a reference genome;

determining sizes of DNA molecules of the biological sample;

determining a first size value of a first group of DNA molecules, wherein:

each DNA molecule of the first group of DNA molecules includes a read located at a site of the plurality of sites and exhibits a second allele different from the first allele at the site, and

the first size value corresponds to a statistical value of a first size distribution of the first group of DNA molecules;

determining a second size value of the second group of DNA molecules, wherein:

each DNA molecule of the second group of DNA molecules includes a read located at a site of the plurality of sites and exhibits the first allele at the site, and

the second size value corresponds to a statistical value of a second size distribution of the second group of DNA molecules;

determining a parameter value from the first size value and the second size value;

comparing the parameter value to a calibration point that is determined using at least one other sample having a known fetal DNA fraction and a calibration value corresponding to separate measurement of the parameter in the at least one other sample; and

computing the fetal DNA fraction based on the comparison.

[0006d] In another aspect, the present disclosure refers to a method for measuring a fetal DNA fraction in a biological sample of a female pregnant with a fetus, the biological sample including maternal DNA molecules and fetal DNA molecules, the method comprising:

receiving a first data set including a first plurality of reads from DNA molecules of a first sample of the female pregnant with the fetus;

identifying locations of the first plurality of reads in a reference genome;

identifying a first set of loci based on the first data set and the identified locations, wherein no locus of the first set of loci exhibits more than one allele;

determining a first amount of loci in the first set of loci;

receiving a second data set including a second plurality of reads from DNA molecules of the biological sample;

identifying locations of the second plurality of reads in the reference genome;

identifying a second set of loci based on the second data set and the identified locations, wherein:

each one of the second set of loci is one of the first set of loci, and

each locus of the second set of loci exhibits an allele different from the allele exhibited in the first set of loci;

determining a second amount of loci in the second set of loci in the second data set;

determining a normalized parameter of the first amount and the second amount;

comparing the parameter value to a calibration point that is determined using at least one other sample having a known fetal DNA fraction and a calibration value corresponding to separate measurement of the parameter in the at least one other sample; and

computing the fetal DNA fraction based on the comparison.

[0006e] In another aspect, the present disclosure refers to a method for measuring a fetal DNA fraction in a biological sample of a female pregnant with a fetus, the biological sample including maternal DNA molecules and fetal DNA molecules, the method comprising:

receiving a first data set from a first plurality of reads from a first plurality of DNA molecules;

identifying locations of the first plurality of reads in a reference genome;

determining sizes of the DNA molecules corresponding to the first plurality of reads;

identifying a first set of loci in the first data set, wherein DNA molecules comprising reads in each of the first set of loci have a first size distribution and have a first size value of the first size distribution exceeding a first size threshold;

determining a first amount of loci in the first set of loci;

receiving a second data set from a second plurality of reads from a second plurality of DNA molecules of the biological sample;

identifying locations of the second plurality of reads in the reference genome;

determining sizes of the DNA molecules corresponding to the second plurality of reads;

identifying a second set of loci in the second data set, wherein:

each one of the second set of loci is one of the first set of loci, and

DNA molecules comprising reads in each of the second set of loci have a second size distribution and have a second size value of the second distribution exceeding a second size threshold in an opposite direction than the first size value exceeds the first size threshold;

determining a second amount of loci in the second set of loci in the second data set;

determining a normalized parameter of the first amount and the second amount;

comparing the parameter value to a calibration point that is determined using at least one other sample having a known fetal DNA fraction and a calibration value corresponding to separate measurement of the parameter in the at least one other sample; and

computing the fetal DNA fraction based on the comparison.

[0006f] 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 as disclosed herein.

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

the computer product as disclosed herein; and

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

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

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

[0006j] 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] Examples of the present disclosure provide methods, systems, and apparatus for deducing the fetal DNA fraction in maternal plasma. The fetal DNA fraction can be determined without specifically determining paternal or fetal genotypes. A separate parameter may be determined, and a calibration curve can be used to determine the actual fetal DNA fraction. For example, a ratio can be determined of an amount of reads having an allele nominally identified as a non-maternal allele and an amount of reads having an allele nominally identified as a maternal allele. As another example, a ratio can be determined of an amount of loci exhibiting a nominal non-maternal allele and an amount of homozygous maternal loci, as determined from a separate dataset. A difference in size of reads may also be used. The loci (sites) may be limited to known heterozygous loci in a population.

[0008] Maternal genotype information may be obtaining from a maternal-only DNA sample or may be assumed from a sequencing (e.g., at shallow-depth) of a biological sample having both maternal and fetal DNA molecules. The actual or assumed maternal genotype information can be combined with sequencing of DNA molecules from the biological sample. Even though it may not be definitively known whether the mother is homozygous at a particular loci or whether the fetus is heterozygous, examples can use reads at such sites in determining the separate parameter, which is a difference from previous techniques. Any errors are shown to be consistent, and thus compensate by the calibration curve, which can be generated once using a separate technique to determine the fetal DNA fraction.

[0009] Because sequencing may be at shallow depths, a locus may have only few reads and may fail to exhibit a non-maternal allele even if a non-maternal allele is present. However, normalized parameters that characterize non-maternal alleles sequenced can be used to provide an accurate estimate of the fetal DNA fraction, even if the amount of non-maternal alleles at a locus or at all loci is not representative of the fetal DNA fraction. These normalized parameters may include the amount of sequence reads with the non-maternal allele or the amount of loci with the non-maternal allele. Methods described herein may not need high-depth sequencing or enrichment of specific regions. As a result, these methods can be integrated into widely used non-invasive prenatal testing and other diagnostics.

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

[0011] A better understanding of the nature and advantages of embodiments of the present invention

BRIEF DESCRIPTION OF THE DRAWINGS

[0012] **FIG. 1** is a schematic illustration of measuring fractional fetal DNA concentration using a maternal genotype according to embodiments of the present invention.

[0013] **FIG. 2A** shows a linear regression model of the actual fetal DNA fraction and the non-maternal allele fraction constructed from a training data set from a first data set according to embodiments of the present invention.