

regions on chromosome 14 [8]. Jensen et al, on the other hand, only focused their statistical analysis on the deleted region on chromosome 22 [9]. Thus, from the data presented by Peters et al and Jensen et al, it is not clear if the approach would be robust enough for a genomewide survey of microdeletions or microduplications, or indeed for the noninvasive determination of a fetal karyotype.

[0005] Lo et al reported that fetal single nucleotide polymorphisms (SNPs) can be genotyped in a genomewide scale using maternal plasma DNA sequencing [10]. In particular, these investigators have demonstrated that SNP alleles and mutations for single gene disorders that are inherited by a fetus from its mother can be elucidated by a process called relative haplotype dosage analysis [10]. Fan et al confirmed the robustness of relative haplotype dosage analysis and used this approach to detect a ~2.85 Mb deletion inherited by a fetus from its mother [11]. There are two concerns for using this method for the clinical implementation of noninvasive prenatal karyotyping. First, this method requires maternal haplotyping to be performed which would require additional analytical steps [12,13] or pedigree analysis. Second, it is unclear if this method could be used to detect de novo subchromosomal deletion or duplication.

BRIEF SUMMARY

[0006] Disclosed herein are methods, systems, and apparatus for detecting microamplifications or microdeletions in the genome of a fetus. In some embodiments, the method comprises receiving sequence tags for each of a plurality of DNA fragments in a biological sample; determining genomic positions for the sequence tags; determining whether the density of DNA in each of a plurality of genomic regions is aberrantly high or low; identifying as a microamplification a set of consecutive genomic regions having aberrantly high density; and identifying as a microdeletion a set of consecutive genomic regions having aberrantly low density. The biological sample may be a blood sample obtained noninvasively from a female subject pregnant with the fetus.

[0006a] In one aspect, the present invention provides a method of identifying microamplifications or microdeletions in a genome of a fetus by analyzing a biological sample obtained from a female subject pregnant with the fetus, the biological sample

including cell-free DNA from the fetus and from the female subject, the method comprising:

receiving one or more sequence tags for each of a plurality of DNA fragments in the biological sample;

determining genomic positions for the sequence tags;

for each of a plurality of genomic regions:

determining, with a computer system, a respective amount of DNA fragments within the genomic region from sequence tags having genomic positions within the genomic region;

normalizing the respective amount to obtain a respective density; and

comparing the respective density to a reference density to identify whether the respective density is statistically different from the reference density;

determining whether any of the genomic regions identified to have a respective density statistically different from the reference density is consecutive with

another genomic region identified to have a respective density statistically different from the reference density;

when at least N first genomic regions identified to have respective densities statistically higher than the reference density are consecutive, identifying the first consecutive genomic regions as a microamplification, N being an integer equal to or greater than two;

when at least N second genomic regions identified to have respective densities statistically lower than the reference densities are consecutive, identifying the second consecutive genomic regions as a microdeletion.

[0006b] In another aspect, the present invention provides a computer product comprising a non-transitory computer readable medium storing a plurality of instructions that when executed control a computer system to identify microamplifications or microdeletions in a genome of a fetus by analyzing a biological sample obtained from a female subject pregnant with the fetus, the biological sample including cell-free DNA from the fetus and from the female subject, the method comprising:

receiving one or more sequence tags for each of a plurality of DNA fragments in a biological sample;
determining genomic positions for the sequence tags;
for each of a plurality of genomic regions:
5 determining a respective amount of DNA fragments within the genomic region from sequence tags having genomic positions within the genomic region;
normalizing the respective amount to obtain a respective density; and
comparing the respective density to a reference density to identify whether the respective density is statistically different from the reference density;
10 determining whether any of the genomic regions identified to have a respective density statistically different from the reference density is consecutive with another genomic region identified to have a respective density statistically different from the reference density;
when at least N first genomic regions identified to have respective
15 densities statistically higher than the reference density are consecutive, identifying the first consecutive genomic regions as a microamplification, N being an integer equal to or greater than two;
when at least N second genomic regions identified to have respective densities statistically lower than the reference densities are consecutive, identifying the second
20 consecutive genomic regions as a microdeletion.

BRIEF DESCRIPTION OF THE DRAWINGS

[0007] FIG 1 is a flowchart showing a method of identifying microamplifications or microdeletions in a genome of a fetus.