1998;339:1734-8; Saito H et al., *Lancet* 2000;356:1170; Chiu RW et al., *Lancet* 2002;360:998-1000).

[0004] However, the detection of certain paternally-inherited mutant alleles can be difficult, e.g., gene deletion, inversion, mutations in repetitive elements, and homologous genes, even with excessive depths of sequencing. Further, it can be difficult to detect maternally-inherited mutations, particularly if no genetic information is available from the father.

SUMMARY

[0004a] In one aspect, the present invention provides a method of determining a portion of a fetal genome of a fetus inherited from a pregnant mother using a biological sample obtained from the pregnant mother, the pregnant mother having a maternal genome with a first maternal haplotype and a second maternal haplotype in a chromosomal region, where the biological sample comprises a mixture of maternal and fetal DNA fragments, the method comprising:

based on an analysis of DNA in one or more other samples from the pregnant mother:

determining the first maternal haplotype to have first alleles at a plurality of loci in the chromosomal region, the maternal genome being heterozygous at the plurality of loci, and

determining the second maternal haplotype to have second alleles at the plurality of loci in the chromosomal region, the second alleles being different than the first alleles;

selecting a set of the plurality of loci, wherein selecting the set of the plurality of loci does not use any measurements of a paternal allele;

analyzing a plurality of cell-free DNA fragments from the biological sample obtained from the pregnant mother, wherein analyzing each DNA fragment of the plurality of cell-free DNA fragments includes:

identifying a location of the DNA fragment in a reference genome, wherein at least a portion of the plurality of cell-free DNA fragments is located at the set of the plurality of loci; and

determining an allele of the DNA fragment;

identifying a first group of DNA fragments in the biological sample as having one of the first alleles at one of the set of the plurality of loci based on the identified locations and the determined alleles for the first group of DNA fragments;

identifying a second group of DNA fragments in the biological sample as having one of the second alleles at one of the set of the plurality of loci based on the identified locations and the determined alleles for the second group of DNA fragments;

calculating, by a computer system, a first value of the first group of DNA fragments, the first value defining a property of a size or a number of the DNA fragments of the first group;

calculating, by the computer system, a second value of the second group of DNA fragments, the second value defining a property of a size or a number of the DNA fragments of the second group;

computing a separation value between the first value and the second value;

determining that the fetus inherited the first maternal haplotype when the separation value is greater than a first threshold; and

determining that the fetus inherited the second maternal haplotype when the separation value is less than a second threshold, wherein the first and second values or the first and second thresholds are determined based on likelihoods of the plurality of loci being homozygous in the fetus and/or a father of the fetus, and wherein the first threshold is greater than the second threshold.

[0004b] In another aspect, the present invention provides a method for detecting a mutation in a fetal genome of a fetus inherited from a pregnant mother using a biological sample obtained from the pregnant mother, the pregnant mother having a maternal genome with a first maternal haplotype and a second maternal haplotype in a chromosomal region, wherein the biological sample contains a mixture of maternal and fetal DNA fragments, the method comprising:

sequencing DNA molecules from a cellular maternal sample, wherein the DNA molecules overlap with the chromosomal region and are at least 1 kb long to obtain long sequence reads of at least 1 kb from both chromosomal copies in the chromosomal region;

constructing the first maternal haplotype using a first set of long sequence reads that share alleles at a plurality of loci in the chromosomal region, the first maternal haplotype having first alleles at the plurality of loci;

constructing the second maternal haplotype using a second set of long sequence reads that share alleles at the plurality of loci in the chromosomal region, the second maternal haplotype having second alleles at the plurality of loci;

identifying the mutation at a first location in the first maternal haplotype in the chromosomal region;

selecting a set of the plurality of loci based on the first location of the mutation, wherein the set of the plurality of loci is selected to be within a specified distance of the first location of the mutation, and wherein the specified distance is the at least 1 kb used in sequencing the DNA molecules;

analyzing a plurality of cell-free DNA fragments from the biological sample obtained from the pregnant mother, wherein analyzing each DNA fragment of the plurality of cell-free DNA fragments includes:

identifying a location of the DNA fragment in a reference genome, wherein at least a portion of the plurality of cell-free DNA fragments is located at the set of the plurality of loci; and

determining an allele of the DNA fragment;

determining paternal alleles inherited by the fetus from a father at the set of the plurality of loci, wherein the paternal alleles correspond to the first alleles or the second alleles, and wherein the set of the plurality of loci is further selected based on locations that the paternal alleles are determined;

identifying a first group of DNA fragments in the biological sample as having one of the first alleles at one of the set of the plurality of loci based on the identified locations and the determined alleles for the first group of DNA fragments;

identifying a second group of DNA fragments in the biological sample as having one of the second alleles at one of the set of the plurality of loci based on the identified locations and the determined alleles for the second group of DNA fragments;

calculating, by a computer system, a first value of the first group of DNA fragments, the first value defining a property of a size or a number of the DNA fragments of the first group;

calculating, by the computer system, a second value of the second group of DNA fragments, the second value defining a property of a size or a number of the DNA fragments of the second group;

computing a separation value between the first value and the second value; and determining whether the fetus inherited the mutation on the first maternal haplotype based on a comparison of the separation value to a cutoff value and based on whether the paternal alleles correspond to the first alleles or the second alleles, wherein the fetus is determined to have inherited the mutation on the first maternal haplotype when the comparison of the separation value to the cutoff value indicates the first maternal haplotype is overrepresented.

[0004c] In yet another aspect, the present invention provides a method for detecting a mutation in a fetal genome of a fetus inherited from a father using a biological sample obtained from a pregnant mother of the fetus, the father having a paternal genome with a first paternal haplotype and a second paternal haplotype in a chromosomal region, wherein the biological sample contains a mixture of maternal and fetal DNA fragments, the method comprising:

sequencing DNA molecules from a cellular paternal sample, wherein the DNA molecules overlap with the chromosomal region and are at least 1 kb long, to obtain long sequence reads from both chromosomal copies in the chromosomal region, the long sequence reads being at least 1 kb long;

constructing the first paternal haplotype using a first set of long sequence reads that share alleles at a plurality of loci in the chromosomal region, the first paternal haplotype having first alleles at the plurality of loci;

constructing the second paternal haplotype using a second set of long sequence reads that share alleles at the plurality of loci in the chromosomal region, the second paternal haplotype having second alleles at the plurality of loci;

identifying the mutation at a first location in the first paternal haplotype in the chromosomal region;

selecting a set of the plurality of loci based on the first location of the mutation and based on a maternal genome of the pregnant mother being homozygous at the set of the plurality of loci, wherein the maternal genome is homozygous for the first alleles at a first subset of the set of the plurality of loci, wherein the maternal genome is homozygous for the second alleles at a second subset of the set of the plurality of loci, wherein the set of the plurality of loci is selected to be within a specified distance of the first location of the mutation, and wherein the specified distance is the at least 1 kb used in sequencing the DNA molecules;

analyzing a plurality of cell-free DNA fragments from the biological sample obtained from the pregnant mother, wherein analyzing each DNA fragment of the plurality of cell-free DNA fragments includes:

identifying a location of the DNA fragment in a reference genome, wherein at least a portion of the plurality of cell-free DNA fragments is located at the set of the plurality of loci; and

determining an allele of the DNA fragment;

identifying a first group of DNA fragments in the biological sample as having one of the first alleles at one of the first subset of the set of the plurality of loci based on the identified locations and the determined alleles for the first group of DNA fragments;

identifying a second group of DNA fragments in the biological sample as having one of the second alleles at one of the second subset of the set of the plurality of loci based on the identified locations and the determined alleles for the second group of DNA fragments;

calculating, by a computer system, a first amount of the first group of DNA fragments;

calculating, by the computer system, a second amount of the second group of DNA fragments;

computing a separation value between the first amount and the second amount; and

determining whether the fetus inherited the mutation on the first paternal haplotype based on a comparison of the separation value to a cutoff value, wherein the fetus is determined to have inherited the mutation on the first paternal haplotype when the comparison of the separation value to the cutoff value indicates the first paternal haplotype is overrepresented.

CLAIMS:

1. A method of determining a portion of a fetal genome of a fetus inherited from a pregnant mother using a biological sample obtained from the pregnant mother, the pregnant mother having a maternal genome with a first maternal haplotype and a second maternal haplotype in a chromosomal region, where the biological sample comprises a mixture of maternal and fetal DNA fragments, the method comprising:

based on an analysis of DNA in one or more other samples from the pregnant mother:

determining the first maternal haplotype to have first alleles at a plurality of loci in the chromosomal region, the maternal genome being heterozygous at the plurality of loci, and

determining the second maternal haplotype to have second alleles at the plurality of loci in the chromosomal region, the second alleles being different than the first alleles;

selecting a set of the plurality of loci, wherein selecting the set of the plurality of loci does not use any measurements of a paternal allele;

analyzing a plurality of cell-free DNA fragments from the biological sample obtained from the pregnant mother, wherein analyzing each DNA fragment of the plurality of cell-free DNA fragments includes:

identifying a location of the DNA fragment in a reference genome, wherein at least a portion of the plurality of cell-free DNA fragments is located at the set of the plurality of loci; and

determining an allele of the DNA fragment;

identifying a first group of DNA fragments in the biological sample as having one of the first alleles at one of the set of the plurality of loci based on the identified locations and the determined alleles for the first group of DNA fragments;

identifying a second group of DNA fragments in the biological sample as having one of the second alleles at one of the set of the plurality of loci based on the identified locations and the determined alleles for the second group of DNA fragments;

calculating, by a computer system, a first value of the first group of DNA fragments, the first value defining a property of a size or a number of the DNA fragments of the first group;

calculating, by the computer system, a second value of the second group of DNA fragments, the second value defining a property of a size or a number of the DNA fragments of the second group;

computing a separation value between the first value and the second value;

determining that the fetus inherited the first maternal haplotype when the separation value is greater than a first threshold; and

determining that the fetus inherited the second maternal haplotype when the separation value is less than a second threshold, wherein the first and second values or the first and second thresholds are determined based on likelihoods of the plurality of loci being homozygous in the fetus and/or a father of the fetus, and wherein the first threshold is greater than the second threshold.

2. The method of claim 1, further comprising:

repeating the method for other chromosomal regions that form sliding windows that overlap with each other; and

identifying a recombination when a specified number of consecutive sliding windows indicate a change to the other maternal haplotype being inherited.

3. The method of claim 1, wherein the fetal genome is homozygous at some of the set of the plurality of loci and heterozygous at some of the set of the plurality of loci.

4. The method of claim 1, wherein the fetal genome is homozygous at 30% or more of the set of the plurality of loci.

5. The method of claim 1, wherein the first threshold and the second threshold are selected using a statistical distribution for defining a stochastic variation that estimates a standard deviation.

6. The method of claim 1, wherein the separation value includes a difference between the first value and the second value, and wherein one of the first and second threshold is positive and another of the first and second thresholds is negative.

7. The method of claim 1, wherein the separation value includes a ratio between the first value and the second value.

8. The method of claim 1, wherein the first value corresponds to a statistical value of a size distribution of the DNA fragments of the first group, and the second value corresponds to a statistical value of a size distribution of the DNA fragments of the second group.

9. The method of claim 8, wherein the first value is an average size of the DNA fragments of the first group, and the second value is an average size of the DNA fragments of the second group.

10. The method of claim 8, wherein the first value Q_{HapI} is a fraction of DNA fragments in the first group that are shorter than a cutoff size, and the second value Q_{HapII} is a fraction of DNA fragments in the second group that are shorter than the cutoff size.

11. The method of claim 10, wherein the separation value includes a difference between the first value and the second value, and wherein the difference includes $\Delta Q = Q_{HapI} - Q_{HapII}$.

12. The method of claim 8, wherein the first value $F_{Hap\ I}$ and the second value $F_{Hap\ II}$ are defined for a respective haplotype as $F = \sum^{w}\text{length} / \sum^{N}\text{length}$, where $\sum^{w}\text{length}$ represents a sum of lengths of the DNA fragments of a corresponding group with a length equal to or less than a cutoff size w; and $\sum^{N}\text{length}$ represents a sum of lengths of the DNA fragments of the corresponding group with a length equal to or less than N bases, where N is greater than w.

13. The method of claim 12, wherein the separation value includes a difference between the first value and the second value, wherein the difference includes $\Delta F = F_{HapI} - F_{Hap\ II}$.

14. The method of claim 1, wherein the first value of the first group corresponds to a number of DNA fragments located at the set of the plurality of loci, and the second value of the second group corresponds to a number of DNA fragments located at the set of the plurality of loci.

15. The method of claim 1, further comprising:

determining a fractional concentration of fetal DNA in the biological sample; and

using the fractional concentration of fetal DNA to determine the first threshold and the second threshold.

16. The method of claim 1, wherein the one or more other samples are of cellular tissue from the pregnant mother, the method further comprising:

sequencing DNA molecules that overlap with the chromosomal region and that are at least 1 kb long in the one or more other samples to determine the first maternal haplotype and the second maternal haplotype.

17. The method of claim 16, wherein the sequencing includes single molecule sequencing.

18. The method of claim 16, wherein the sequencing includes linked-read sequencing of DNA molecules that are at least 1 kb long.

19. The method of claim 1, wherein selecting the set of the plurality of loci includes:

identifying a mutation at a first location in the first maternal haplotype in the chromosomal region; and

selecting the set of the plurality of loci that are within a specified distance of the first location of the mutation.

20. The method of claim 19, wherein the specified distance is 5 Mb.

21. The method of claim 19, wherein analyzing the plurality of cell-free DNA fragments from the biological sample includes:

sequencing the plurality of cell-free DNA fragments, wherein the sequencing targets a genomic window that includes the mutation.

22. The method of claim 19, wherein analyzing the plurality of cell-free DNA fragments from the biological sample includes:

using probes and/or primers that are specific to a genomic window that includes the mutation.

23. The method of claim 1, wherein selecting the set of the plurality of loci includes:

accessing a database of population statistics for a population that corresponds to the fetus and/or to the father of the fetus; and

excluding a locus having a prevalence of being heterozygous that is above a cutoff value for the population.

24. The method of claim 1, wherein the first and second thresholds are determined based on the likelihoods of the plurality of loci being homozygous in the fetus and/or the father of the fetus.

25. The method of claim 24, wherein the first and second thresholds are determined by reducing a threshold value by a proportion of the plurality of loci that have the likelihoods greater than a cutoff value.

26. The method of claim 1, wherein the first and second values are determined based on the likelihoods of the plurality of loci being homozygous in the fetus and/or the father of the fetus.

27. The method of claim 26, wherein the first and second values are determined by assigning weights to the first and second values according to the likelihoods.

28. The method of claim 1, wherein the biological sample is plasma from a blood sample, and wherein the one or more other samples includes a buffy coat from the blood sample.

29. The method of claim 1, wherein the first group includes at least one DNA fragment located at each of the set of the plurality of loci, and wherein the second group includes at least one DNA fragment located at each of the set of the plurality of loci.

30. A method for detecting a mutation in a fetal genome of a fetus inherited from a pregnant mother using a biological sample obtained from the pregnant mother, the pregnant mother having a maternal genome with a first maternal haplotype and a second

maternal haplotype in a chromosomal region, wherein the biological sample contains a mixture of maternal and fetal DNA fragments, the method comprising:

sequencing DNA molecules from a cellular maternal sample, wherein the DNA molecules overlap with the chromosomal region and are at least 1 kb long , to obtain long sequence reads of at least 1 kb from both chromosomal copies in the chromosomal region;

constructing the first maternal haplotype using a first set of long sequence reads that share alleles at a plurality of loci in the chromosomal region, the first maternal haplotype having first alleles at the plurality of loci;

constructing the second maternal haplotype using a second set of long sequence reads that share alleles at the plurality of loci in the chromosomal region, the second maternal haplotype having second alleles at the plurality of loci;

identifying the mutation at a first location in the first maternal haplotype in the chromosomal region;

selecting a set of the plurality of loci based on the first location of the mutation, wherein the set of the plurality of loci is selected to be within a specified distance of the first location of the mutation, and wherein the specified distance is the at least 1 kb used in sequencing the DNA molecules;

analyzing a plurality of cell-free DNA fragments from the biological sample obtained from the pregnant mother, wherein analyzing each DNA fragment of the plurality of cell-free DNA fragments includes:

identifying a location of the DNA fragment in a reference genome, wherein at least a portion of the plurality of cell-free DNA fragments is located at the set of the plurality of loci; and

determining an allele of the DNA fragment;

determining paternal alleles inherited by the fetus from a father at the set of the plurality of loci, wherein the paternal alleles correspond to the first alleles or the second alleles, and wherein the set of the plurality of loci is further selected based on locations that the paternal alleles are determined;

identifying a first group of DNA fragments in the biological sample as having one of the first alleles at one of the set of the plurality of loci based on the identified locations and the determined alleles for the first group of DNA fragments;

identifying a second group of DNA fragments in the biological sample as having one of the second alleles at one of the set of the plurality of loci based on the identified locations and the determined alleles for the second group of DNA fragments;

calculating, by a computer system, a first value of the first group of DNA fragments, the first value defining a property of a size or a number of the DNA fragments of the first group;

calculating, by the computer system, a second value of the second group of DNA fragments, the second value defining a property of a size or a number of the DNA fragments of the second group;

computing a separation value between the first value and the second value; and

determining whether the fetus inherited the mutation on the first maternal haplotype based on a comparison of the separation value to a cutoff value and based on whether the paternal alleles correspond to the first alleles or the second alleles, wherein the fetus is determined to have inherited the mutation on the first maternal haplotype when the comparison of the separation value to the cutoff value indicates the first maternal haplotype is overrepresented.

31. The method of claim 30, wherein the sequencing includes linked-read sequencing of DNA molecules to reconstruct the long sequence reads from smaller linked reads, wherein the chromosomal region includes a structural variation, and wherein constructing the first maternal haplotype includes identifying reconstructed long sequence reads that each differ in length from an average length of reconstructed long sequence reads for regions before and after the structural variation by at least a specified length.

32. The method of claim 30, wherein analyzing the plurality of cell-free DNA fragments from the biological sample includes:

sequencing the plurality of cell-free DNA fragments, wherein the sequencing targets a genomic window that includes the mutation and that corresponds to the specified distance from the mutation.

33. The method of claim 30, wherein analyzing the plurality of cell-free DNA fragments from the biological sample includes:

using probes and/or primers that are specific to a genomic window that includes the mutation and that corresponds to the specified distance from the mutation.

34. The method of claim 30, wherein the first group includes at least one DNA fragment located at each of the set of the plurality of loci, and wherein the second group includes at least one DNA fragment located at each of the set of the plurality of loci.

35. The method of claim 30, wherein the specified distance is 5 Mb.

36. A method for detecting a mutation in a fetal genome of a fetus inherited from a father using a biological sample obtained from a pregnant mother of the fetus, the father having a paternal genome with a first paternal haplotype and a second paternal haplotype in a chromosomal region, wherein the biological sample contains a mixture of maternal and fetal DNA fragments, the method comprising:

sequencing DNA molecules from a cellular paternal sample, wherein the DNA molecules overlap with the chromosomal region and are at least 1 kb long , to obtain long sequence reads from both chromosomal copies in the chromosomal region, the long sequence reads being at least 1 kb long;

constructing the first paternal haplotype using a first set of long sequence reads that share alleles at a plurality of loci in the chromosomal region, the first paternal haplotype having first alleles at the plurality of loci;

constructing the second paternal haplotype using a second set of long sequence reads that share alleles at the plurality of loci in the chromosomal region, the second paternal haplotype having second alleles at the plurality of loci;

identifying the mutation at a first location in the first paternal haplotype in the chromosomal region;

selecting a set of the plurality of loci based on the first location of the mutation and based on a maternal genome of the pregnant mother being homozygous at the set of the plurality of loci, wherein the maternal genome is homozygous for the first alleles at a first subset of the set of the plurality of loci, wherein the maternal genome is homozygous for the second alleles at a second subset of the set of the plurality of loci, wherein the set of the plurality of loci

is selected to be within a specified distance of the first location of the mutation, and wherein the specified distance is the at least 1 kb used in sequencing the DNA molecules;

analyzing a plurality of cell-free DNA fragments from the biological sample obtained from the pregnant mother, wherein analyzing each DNA fragment of the plurality of cell-free DNA fragments includes:

identifying a location of the DNA fragment in a reference genome, wherein at least a portion of the plurality of cell-free DNA fragments is located at the set of the plurality of loci; and

determining an allele of the DNA fragment;

identifying a first group of DNA fragments in the biological sample as having one of the first alleles at one of the first subset of the set of the plurality of loci based on the identified locations and the determined alleles for the first group of DNA fragments;

identifying a second group of DNA fragments in the biological sample as having one of the second alleles at one of the second subset of the set of the plurality ofloci based on the identified locations and the determined alleles for the second group of DNA fragments;

calculating, by a computer system, a first amount of the first group of DNA fragments;

calculating, by the computer system, a second amount of the second group of DNA fragments;

computing a separation value between the first amount and the second amount; and

determining whether the fetus inherited the mutation on the first paternal haplotype based on a comparison of the separation value to a cutoff value, wherein the fetus is determined to have inherited the mutation on the first paternal haplotype when the comparison of the separation value to the cutoff value indicates the first paternal haplotype is overrepresented.

37. The method of claim 36, wherein the sequencing includes linked-read sequencing of DNA molecules to reconstruct the long sequence reads from smaller linked reads, wherein the chromosomal region includes a structural variation, and wherein constructing the first paternal haplotype includes identifying reconstructed long sequence reads that each differ in length from an average length of reconstructed long sequence reads for regions before and after the structural variation by at least a specified length.

38. The method of claim 36, wherein analyzing the plurality of cell-free DNA fragments from the biological sample includes:

sequencing the plurality of cell-free DNA fragments, wherein the sequencing targets a genomic window that includes the mutation.

39. The method of claim 36, wherein analyzing the plurality of cell-free DNA fragments from the biological sample includes:

using probes and/or primers that are specific to a genomic window that includes the mutation.

40. The method of claim 36, wherein the first group includes at least one DNA fragment located at each of the first subset of the set of the plurality of loci, and wherein the second group includes at least one DNA fragment located at each of the second subset of the set of the plurality of loci.

41. The method of claim 36, wherein the specified distance is 5 Mb.

42. A computer product comprising a computer readable medium storing a plurality of instructions for controlling a computer system to perform the method of any one of claims 1-41.

43. A system comprising:

the computer product of claim 42; and

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

44. A system comprising means for performing the method of any one of claims 1-41.

45. A system configured to perform the method of any one of claims 1-41.

46. A system comprising modules that respectively perform the method of any one of claims 1-41.