

CLAIMS

What is claimed is:

5 1. A method of amplifying target loci in a nucleic acid sample, the method comprising:

 (a) contacting the nucleic acid sample with a library of test primers that simultaneously hybridize to at least 1,000 different target loci to produce a reaction mixture; and

 (b) subjecting the reaction mixture to primer extension reaction conditions to produce
10 amplified products comprising target amplicons.

 2. The method of claim 1, wherein at least 5,000 different target loci are amplified.

 3. The method of claim 1, wherein at least 10,000 different target loci are amplified.

15 4. The method of claim 1, wherein at least 20,000 different target loci are amplified.

 5. The method of claim 1, wherein at least 30,000 different target loci are amplified.

20 6. The method of claim 1, wherein at least 90% of the amplified products are target amplicons.

 7. The method of claim 1, wherein at least 95% of the amplified products are target amplicons.

25 8. The method of claim 1, wherein at least 99% of the amplified products are target amplicons.

 9. The method of claim 1, wherein at least 90% of the targeted loci are amplified.

30 10. The method of claim 1, wherein at least 95% of the targeted loci are amplified.

 11. The method of claim 1, wherein at least 99% of the targeted loci are amplified.

12. The method of claim 1, wherein less than 20% of the amplified products are test primer dimers.

5

13. The method of claim 1, wherein less than 10% of the amplified products are test primer dimers.

10

14. The method of claim 1, wherein less than 1% of the amplified products are test primer dimers.

15. The method of claim 1, wherein less than 0.1% of the amplified products are test primer dimers.

15

16. The method of claim 1, wherein the test primers are selected from a library of candidate primers based on one or more parameters.

20

17. The method of claim 16, wherein the test primers are selected from a library of candidate primers based at least in part on the ability of the candidate primers to form primer dimers.

18. The method of claim 17, wherein the selection comprises

25

(i) calculating on a computer an undesirability score for most or all of the possible combinations of two candidate primers from the library, wherein each undesirability score is based at least in part on the likelihood of dimer formation between the two candidate primers;

(ii) removing the candidate primer with the highest undesirability score from the library of candidate primers;

(iii) if the candidate primer removed in step (ii) is a member of a primer pair, then removing the other member of the primer pair from the library of candidate primers; and

30

(iv) optionally repeating steps (ii) and (iii)..

19. The method of claim 17, wherein the selection comprises

(i) calculating on a computer an undesirability score for most or all of the possible combinations of two candidate primers from the library, wherein each undesirability score is based at least in part on the likelihood of dimer formation between the two candidate primers;

(ii) removing from the library of candidate primers the candidate primer that is part of the greatest number of combinations of two candidate primers with an undesirability score above a first minimum threshold;

(iii) if the candidate primer removed in step (ii) is a member of a primer pair, then removing the other member of the primer pair from the library of candidate primers; and

(iv) optionally repeating steps (ii) and (iii).

20. The method of claim 19, comprising further reducing the number of candidate primers remaining in the library by decreasing the first minimum threshold used in step (ii) to a lower second minimum threshold and repeating steps (ii) and (iii) until the undesirability scores for the candidate primer combinations remaining in the library are all equal to or below the second minimum threshold, or until the number of candidate primers remaining in the library is reduced to a desired number.

21. The method of claim 19, comprising increasing the first minimum threshold used in step (ii) to a higher second minimum threshold and repeating steps (ii) and (iii) until the undesirability scores for the candidate primer combinations remaining in the library are all equal to or below the second minimum threshold, or until the number of candidate primers remaining in the library is reduced to a desired number.

22. The method of claim 18 or 19, wherein a candidate primer is selected out of a group of two or more candidate primers with equal undesirability scores for removal from the library based on one or more parameters.

23. The method of claim 1, wherein the concentration of each test primer is less than 100 nM.

24. The method of claim 1, wherein the concentration of each test primer is less than 10 nM.

25. The method of claim 1, wherein the concentration of each test primer is less than 2 nM.

26. The method of claim 1, wherein the library of test primers comprises at least 1,000 test primer pairs that include a forward test primer and a reverse test primer where each pair of test primers hybridize to a target locus.

27. The method of claim 1, wherein the library of test primers comprises at least 1,000 individual test primers that each hybridize to a different target locus.

28. The method of claim 1, wherein the GC content of the test primers is between 30% and 80%, inclusive.

29. 10. The method of claim 1, wherein the range of GC content of the test primers is less than 20%.

30. The method of claim 1, wherein the melting temperature of the test primers is between 40 °C and 80 °C, inclusive.

31. The method of claim 1, wherein the range of melting temperatures of the test primers is less than 5 °C.

32. The method of claim 1, wherein the length of the test primers is between 17 and 35 nucleotides, inclusive

33. The method of claim 1, wherein the test primers comprise a tag that is not target specific.

34. The method of claim 33, wherein the tag forms an internal loop structure.

35. The method of claim 35, wherein the tag is between two DNA binding regions.

36. The method of claim 34, wherein the test primers comprise a 5' region that is specific for a target locus, an internal region that is not specific for the target locus and forms loop structure, and a 3' region that is specific for the target locus.

5 37. The method of claim 36, wherein the length of the 3' region is at least 7 nucleotides.

38. The method of claim 37, wherein the length of the 3' region is between 7 and 20 nucleotides, inclusive.

10 39. The method of claim 33, wherein the test primers comprise a 5' region that is not specific for a target locus followed by a region that is specific for a target locus, an internal region that is not specific for the target locus and forms a loop structure, and a 3' region that is specific for the target locus.

15 40. The method of claim 1, wherein range of the length of the target amplicons is less than 15 nucleotides.

20 41. The method of claim 1, wherein the primer extension reaction conditions are polymerase chain reaction conditions (PCR).

42. The method of claim 41, wherein the length of the annealing step is greater than 10 minutes.

25 43. The method of claim 1, further comprising determining the presence or absence of at least one target amplicon.

30 44. The method of claim 1, further comprising determining the sequence of at least one target amplicon.

45. The method of claim 1, wherein the target loci are present on the same chromosome of interest.

46. The method of claim 1, wherein at least some of the target loci are present on different chromosomes nucleic acids of interest.

47. The method of claim 1, wherein the nucleic acid sample comprises fragmented or digested nucleic acids.

48. The method of claim 1, wherein nucleic acid sample comprises genomic DNA, cDNA, or mRNA.

49. The method of claim 1, wherein nucleic acid sample comprises DNA from a single cell.

50. The method of claim 1, wherein nucleic acid sample comprises or is derived from blood, plasma, saliva, sperm, cell culture supernatant, mucus secretion, dental plaque, gastrointestinal tract tissue, stool, urine, or forensic sample.

51. The method of claim 1, wherein the target loci are segments of human nucleic acids.

52. The method of claim 1, wherein the target loci comprise single nucleotide polymorphisms.

53. A method of selecting test primers from a library of candidate primers, the method comprising:

(a) calculating on a computer an undesirability score for most or all of the possible combinations of two candidate primers from the library, wherein each undesirability score is based at least in part on the likelihood of dimer formation between the two candidate primers;

(b) removing the candidate primer with the highest undesirability score from the library of candidate primers;

(c) if the candidate primer removed in step (b) is a member of a primer pair, then removing the other member of the primer pair from the library of candidate primers; and

(d) optionally repeating steps (b) and (c), thereby selecting a library of test primers.

54. A method of selecting test primers from a library of candidate primers, the method comprising:

(a) calculating on a computer an undesirability score for most or all of the possible combinations of two candidate primers from the library, wherein each undesirability score is based at least in part on the likelihood of dimer formation between the two candidate primers;

(b) removing from the library of candidate primers the candidate primer that is part of the greatest number of combinations of two candidate primers with an undesirability score above a first minimum threshold;

(c) if the candidate primer removed in step (b) is a member of a primer pair, then removing the other member of the primer pair from the library of candidate primers; and

(d) optionally repeating steps (b) and (c), thereby selecting a library of test primers.

55. The method of claim 54, comprising further reducing the number of candidate primers remaining in the library by decreasing the first minimum threshold used in step (b) to a lower second minimum threshold and repeating steps (b) and (c) until the undesirability scores for the candidate primer combinations remaining in the library are all equal to or below the second minimum threshold, or until the number of candidate primers remaining in the library is reduced to a desired number.

56. The method of claim 54, comprising increasing the first minimum threshold used in step (b) to a higher second minimum threshold and repeating steps (b) and (c) until the undesirability scores for the candidate primer combinations remaining in the library are all equal to or below the second minimum threshold, or until the number of candidate primers remaining in the library is reduced to a desired number.

57. The method of claim 53 or 54, wherein a candidate primer is selected out of a group of two or more candidate primers with equal undesirability scores for removal from the library of candidate primers based on one or more other parameters.

58. The method of claim 53 or 54, wherein the undesirability scores are based at least in part on one or more parameters selected from the group consisting of heterozygosity rate of the target locus, disease prevalence associated with a polymorphism or mutation at the target

locus, disease penetrance associated with a polymorphism or mutation at the target locus, specificity of the candidate primer for the target locus, size of the candidate primer, melting temperature of the target amplicon, GC content of the target amplicon, amplification efficiency of the target amplicon, and size of the target amplicon.

5

59. The method of claim 58, wherein the undesirability scores are based at least in part on one or more parameters selected from the group consisting of heterozygosity rate of the target locus, specificity of the candidate primer for the target locus, size of the candidate primer, , melting temperature of the target amplicon, GC content of the target amplicon, amplification efficiency of the target amplicon, and size of the target amplicon; and wherein the test primers are used to simultaneously amplify at least 1,000 target loci in a sample comprising maternal DNA from the pregnant mother of a fetus and fetal DNA to determine the presence or absence of a fetal chromosome abnormality.

15

60. The method of claim 59, comprising ligating a universal primer binding site to the DNA molecules in the sample; amplifying the ligated DNA molecules using at least 1,000 specific primers and a universal primer to produce a first set of amplified products; and amplifying the first set of amplified products using at least 1,000 pairs of specific primers to produce a second set of amplified products.

20

61. The method of claim 58, wherein the undesirability scores are based at least in part on one or more parameters selected from the group consisting of heterozygosity rate of the target locus, specificity of the candidate primer for the target locus, size of the candidate primer, melting temperature of the target amplicon, GC content of the target amplicon, amplification efficiency of the target amplicon, and size of the target amplicon; and wherein the test primers are used to simultaneously amplify at least 1,000 target loci in sample comprises DNA from an alleged father of a fetus and to simultaneously amplify the target loci in a sample comprising maternal DNA from the pregnant mother of a fetus and fetal DNA to establish whether the alleged father is the biological father of the fetus.

30

62. The method of claim 58, wherein the undesirability scores are based at least in part on one or more parameters selected from the group consisting of heterozygosity rate of the target locus, specificity of the candidate primer for the target locus; size of the candidate

primer, and size of the target amplicon; and wherein the test primers are used to simultaneously amplify at least 1,000 target loci in a forensic nucleic acid sample using an annealing time of at least 5 minutes.

5 63. The method of claim 58, comprising using the test primers to simultaneously amplify at least 1,000 target loci in a control nucleic acid sample to produce a first set of target amplicons and to simultaneously amplify the target loci in a test nucleic acid sample to produce a second set of target amplicons; and comparing the first and second sets of target amplicons to determine whether a target locus is present in one sample but absent in the other,
10 or whether a target locus is present at different levels in the control sample and the test sample.

 64. The method of claim 63, wherein the test sample is from an individual suspected of having a disease or phenotype of interest, or an increased risk for a disease or phenotype of interest; and wherein one or more of the target loci comprise a polymorphism associated with
15 an increased risk for the disease or phenotype of interest, or associated with the disease or phenotype of interest.

 65. The method of claim 58, comprising using the test primers to simultaneously amplify at least 1,000 target loci in a control sample comprising RNA to produce a first set of
20 target amplicons and to simultaneously amplify the target loci in a test sample comprising RNA to produce a second set of target amplicons; and comparing the first and second sets of target amplicons to determine the presence or absence of a difference in the RNA expression levels between the control sample and the test sample.

25 66. The method of claim 65, wherein the RNA is mRNA.

 67. The method of claim 65, wherein the test sample is from an individual suspected of having cancer or an increased risk for cancer; and wherein one or more of the target loci comprises a polymorphism or other mutation associated with an increased risk for cancer or
30 associated with cancer.

68. The method of claim 65, wherein the test sample is from an individual diagnosed with cancer; and wherein a difference in the RNA expression level between the control sample and test sample indicates a target locus comprises a polymorphism or other mutation associated with an increased or decreased risk for cancer.

5

69. The method of claim 53 or 54, further comprising, prior to step (b), removing a primer pair from the library that produces a target amplicon that overlaps with a target amplicon produced by another primer pair.

10

70. The method of claim 53 or 54, wherein the candidate primers remaining in the library are capable of simultaneously amplifying at least 1,000 different target loci.

71. The method of claim 53 or 54, further comprising:

15

(e) contacting a nucleic acid sample comprising target loci with the candidate primers remaining in the library to produce a reaction mixture; and

(f) subjecting the reaction mixture to primer extension reaction conditions to produce amplified products comprising target amplicons.

20

72. A library of primers that simultaneously amplifies at least 1,000 different target loci such that less than 30% of the amplified products are test primer dimers.

73. A library of primers that simultaneously amplifies at least 1,000 different target loci such that at least 80% of the amplified products are target amplicons.

25

74. A library of primers that simultaneously amplifies target loci such that at least 80% of the targeted loci out of at least 1,000 different target loci are amplified.

30

75. The library of any one of claims 72-74, wherein at least 5,000 different target loci are amplified.

76. The library of any one of claims 72-74, wherein at least 10,000 different target loci are amplified.

77. The library of any one of claims 72-74, wherein at least 20,000 different target loci are amplified.

78. The library of any one of claims 72-74, wherein at least 30,000 different target loci are amplified.

79. The library of any one of claims 72-74, wherein at least 90% of the amplified products are target amplicons.

80. The library of any one of claims 72-74, wherein at least 95% of the amplified products are target amplicons.

81. The library of any one of claims 72-74, wherein at least 99% of the amplified products are target amplicons.

82. The library of any one of claims 72-74, wherein at least 90% of the targeted loci are amplified.

83. The library of any one of claims 72-74, wherein at least 95% of the targeted loci are amplified.

84. The library of any one of claims 72-74, wherein at least 99% of the targeted loci are amplified.

85. The library of any one of claims 72-74, wherein less than 20% of the amplified products are primer dimers.

86. The library of any one of claims 72-74, wherein less than 10% of the amplified products are primer dimers.

87. The library of any one of claims 72-74, wherein less than 1% of the amplified products are primer dimers.

88. The library of any one of claims 72-74, wherein less than 0.1% of the amplified products are primer dimers.

89. The library of any one of claims 72-74, comprising at least 1,000 primer pairs, wherein each primer pair includes a forward primer and a reverse primer that hybridize to a target locus.

90. The library of any one of claims 72-74, comprising at least 1,000 individual primers that hybridize to different target loci.

91. A kit for amplifying target loci in a nucleic acid sample comprising (i) the library of any one of claims 72-74 and (iii) instructions for using the library to amplify the target loci.

92. A method for determining a ploidy status of a chromosome in a gestating fetus, the method comprising:

(a) contacting a nucleic acid sample with a library of primers that simultaneously hybridize to at least 1,000 different polymorphic loci to produce a reaction mixture; wherein the nucleic acid sample comprises maternal DNA from the mother of the fetus and fetal DNA from the fetus; and

(b) subjecting the reaction mixture to primer extension reaction conditions to produce amplified products;

(c) measuring the amplified products with a high throughput sequencer to produce sequencing data;

(d) calculating, on a computer, allele counts at the polymorphic loci based on the sequencing data;

(e) creating on a computer a plurality of ploidy hypotheses each pertaining to a different possible ploidy state of the chromosome;

(f) building on a computer a joint distribution model for the expected allele counts at the polymorphic loci on the chromosome for each ploidy hypothesis;

(g) determining on a computer a relative probability of each of the ploidy hypotheses using the joint distribution model and the allele counts; and

(h) calling the ploidy state of the fetus by selecting the ploidy state corresponding to the hypothesis with the greatest probability.

93. A method of testing for an abnormal distribution of a chromosome in a sample comprising a mixture of maternal and fetal DNA, the method comprising:

(a) contacting the sample with a library of primers that simultaneously hybridize to at least 1,000 different target loci to produce a reaction mixture; wherein the target loci are from a plurality of different chromosomes; and wherein the plurality of different chromosomes comprise at least one first chromosome suspected of having an abnormal distribution in the sample and at least one second chromosome presumed to be normally distributed in the sample;

(b) subjecting the reaction mixture to primer extension reaction conditions to produce amplified products;

(c) sequencing the amplified products to obtain a plurality of sequence tags aligning to the target loci; wherein the sequence tags are of sufficient length to be assigned to a specific target locus;

(d) assigning on a computer the plurality of sequence tags to their corresponding target loci;

(e) determining on a computer a number of sequence tags aligning to the target loci of the first chromosome and a number of sequence tags aligning to the target loci of the second chromosome; and

(f) comparing on a computer the numbers from step (e) to determine the presence or absence of an abnormal distribution of the first chromosome.

94. A method for detecting the presence or absence of a fetal aneuploidy, the method comprising:

(a) contacting a sample comprising a mixture of maternal and fetal DNA with a library of primers that simultaneously hybridize to at least 1,000 different non-polymorphic target loci to produce a reaction mixture; wherein the target loci are from a plurality of different chromosomes;

(b) subjecting the reaction mixture to primer extension reaction conditions to produce amplified products comprising target amplicons;

(c) quantifying on a computer a relative frequency of the target amplicons from the first and second chromosomes of interest;

(d) comparing on a computer the relative frequency of the target amplicons from the first and second chromosomes of interest; and

5 (e) identifying the presence or absence of an aneuploidy based on the compared relative frequencies of the first and second chromosome of interest.

95. A method for establishing whether an alleged father is the biological father of a fetus that is gestating in a pregnant mother, the method comprising:

10 (a) simultaneously amplifying a plurality of polymorphic loci comprising at least 1,000 different polymorphic loci on genetic material from the alleged father to produce a first set of amplified products;

(b) simultaneously amplifying the corresponding plurality of polymorphic loci on a mixed sample of DNA originating from a blood sample from the pregnant mother to produce a
15 second set of amplified products; wherein the mixed sample of DNA comprises fetal DNA and maternal DNA;

(c) determining on a computer the probability that the alleged father is the biological father of the fetus using genotypic measurements based on the first and second sets of amplified products; and

20 (d) establishing whether the alleged father is the biological father of the fetus using the determined probability that the alleged father is the biological father of the fetus.

96. The method of claim 95, further comprising simultaneously amplifying the corresponding plurality of polymorphic loci on genetic material from the mother to produce a
25 third set of amplified products; wherein the probability that the alleged father is the biological father of the fetus is determined using genotypic measurements based on the first, second, and third sets of amplified products.

97. A method of measuring the amount of two or more target loci in a nucleic acid
30 sample, the method comprising:

(a) using PCR to amplify a nucleic acid sample comprising a first standard locus, a second standard locus, a first target locus, and a second target locus to form amplified

products; wherein the first standard locus and the first target locus have the same number of nucleotides but have a sequence that differs at one or more nucleotides; and wherein the second standard locus and the second target locus have the same number of nucleotides but have a sequence that differs at one or more nucleotides;

5 (b) sequencing the amplified products to determine a standard ratio that compares the relative amount of the amplified first standard locus compared to the amplified second standard locus; wherein the standard ratio indicates the difference in PCR efficiency for the amplification of the first standard locus and the second standard locus;

(c) determining a target ratio that compares the relative amount of the amplified first
10 target locus compared to the amplified second target locus; and

(d) adjusting the target ratio from step (c) based on the standard ratio from step (b) to determine the relative amount of the first target locus and the second target locus in the sample.

15 98. The method of claim 97, further comprising determining the absolute amount of the first target locus and the second target locus in the sample.

99. The method of claim 97, comprising determining the presence or absence of a
20 target locus in the sample.

100. The method of claim 97, comprising simultaneously amplifying at least 1,000
different target loci.

101. The method of claim 97, wherein the target loci are present on the same
25 chromosome of interest.

102. The method of claim 97, wherein at least some of the target loci are present on
different chromosomes of interest.

CLAIMS

1. A method of amplifying target loci in a nucleic acid sample, the method comprising:

(a) contacting a nucleic acid sample comprising target loci with a library of test primers comprising at least 1,000 different primer pairs to produce a reaction mixture in one reaction volume; wherein each primer pair includes a forward test primer and a reverse test primer that hybridize to the same target locus, and wherein the primers do not include molecular inversion probes (MIPs);

(b) subjecting the reaction mixture to polymerase chain reaction (PCR) conditions to produce amplified products comprising target amplicons; wherein the concentration of each test primer is less than 10 nM; wherein the length of the annealing step of the reaction conditions is greater than 10 minutes; wherein at least 1,000 different target loci are simultaneously amplified; and wherein (i) less than 20% of the amplified products are test primer dimers, (ii) at least 80% of the amplified products are target amplicons, and (iii) at least 80% of the target loci are amplified; and

(c) sequencing the amplified products.

2. The method of claim 1, wherein at least 5,000 different target loci are amplified.

3. The method of claim 1, wherein the length of the target amplicons is less than 100 nucleotides.

4. The method of claim 1, wherein the test primers are selected from a library of candidate primers based at least in part on the ability of the candidate primers to form primer dimers.

5. The method of claim 1, wherein the sample comprises maternal DNA from the pregnant mother of a fetus and fetal DNA, and wherein the method comprises determining the presence or absence of a fetal chromosome abnormality from the sequencing data.

6. The method of claim 1, wherein the target loci are present in the human genome, or wherein the target loci comprise human single nucleotide polymorphisms.

7. The method of claim 1, wherein the nucleic acid sample comprises DNA from a tumor, transplant, or fetus.

8. The method of claim 1, wherein the sample comprises DNA from a single cell.

9. A method according to claim 1, wherein the library of test primers are selected from a library of candidate primers, by a method comprising:

(a) calculating on a computer an undesirability score for at least 90% of the possible combinations of two candidate primers from the library, wherein each undesirability score is

based at least in part on the likelihood of dimer formation between the two candidate primers;

(b) removing from the library of candidate primers the candidate primer that is part of the greatest number of combinations of two candidate primers with an undesirability score above a first minimum threshold;

(c) if the candidate primer removed in step (b) is a member of a primer pair, then removing the other member of the primer pair from the library of candidate primers; and

(d) optionally repeating steps (b) and (c), thereby selecting a library of test primers.

10. The method of claim 9, comprising either

(i) further reducing the number of candidate primers remaining in the library by decreasing the first minimum threshold used in step (b) to a lower second minimum threshold and repeating steps (b) and (c) until the undesirability scores for the candidate primer combinations remaining in the library are all equal to or below the second minimum threshold, or until the number of candidate primers remaining in the library is reduced to a desired number; or

(ii) increasing the number of candidate primers remaining in the library by increasing the first minimum threshold used in step (b) to a higher second minimum threshold and repeating steps (b) and (c) until the undesirability scores for the candidate primer combinations remaining in the library are all equal to or below the second minimum threshold, or until the number of candidate primers remaining in the library is reduced to a desired number.

11. The method of claim 9, further comprising, prior to step (b), removing a primer pair from the library that produces a target amplicon that overlaps with a target amplicon produced by another primer pair.

12. The method of claim 9, wherein the undesirability scores are based at least in part on one or more parameters selected from the group consisting of heterozygosity rate of a target locus, disease prevalence associated with a polymorphism or mutation at a target locus, disease penetrance associated with a polymorphism or mutation at a target locus, specificity of the candidate primer for a target locus, size of the candidate primer, melting temperature of a target amplicon, guanine-cytosine (GC) content of a target amplicon, amplification efficiency of a target amplicon, and size of a target amplicon.

13. The method of claim 9, comprising selecting a library of at least 1,000 different test primers and using at least 1,000 of the selected test primers to simultaneously amplify at least 1,000 different target loci.

14. The method of claim 1, wherein the melting temperatures of the primers of the primer pairs is 55 to 65°C.
15. The method of claim 1, wherein the annealing temperature comprises a temperature that is raised.
16. The method of claim 1, wherein the melting temperatures of the primers of the primer pairs is 57 to 60.5°C.
17. The method of claim 1, wherein the PCR is nested PCR.
18. The method of claim 1, wherein range of the length of the target amplicons is between 50 and 100 nucleotides.
19. The method of 1, wherein the primer pairs are designed to amplify potentially heterozygous SNP loci on chromosomes X, Y, 13, 18, and 21 and these assays are used in a single reaction to amplify cfDNA obtained from a maternal plasma sample.