

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
17 September 2009 (17.09.2009)

PCT

(10) International Publication Number  
**WO 2009/114543 A2**

(51) International Patent Classification:

C12Q 1/68 (2006.01) C07H 21/04 (2006.01)  
C12N 15/11 (2006.01)

(21) International Application Number:

PCT/US2009/036683

(22) International Filing Date:

10 March 2009 (10.03.2009)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

61/035,711 11 March 2008 (11.03.2008) US

(71) Applicant (for all designated States except US): SE-  
QUENOM, INC. [US/US]; 3595 John Hopkins Court,  
San Diego, CA 92121 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): OETH, Paul, An-  
drew [US/US]; 4140-162 Via Candidiz, San Diego, CA  
92130 (US). MAHBOUBI, Payam [US/US]; 2568 Alba-  
tross St, #5J, San Diego, CA 92101 (US). LEE, Min,  
Seob [KR/US]; 7824 Vista Lazanja, San Diego, CA  
92127 (US).

(74) Agents: GRANT, Bruce, D. et al.; Grant Anderson Llp,  
C/o Portfoliioip, P.o. Box 52050, Minneapolis, MN 55402  
(US).

(81) Designated States (unless otherwise indicated, for every

kind of national protection available): AE, AG, AL, AM,  
AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ,  
CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ,  
EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN,  
HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR,  
KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME,  
MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO,  
NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG,  
SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA,  
UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every

kind of regional protection available): ARIPO (BW, GH,  
GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM,  
ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ,  
TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE,  
ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV,  
MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR),  
OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML,  
MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished  
upon receipt of that report (Rule 48.2(g))

(54) Title: NUCLEIC ACID-BASED TESTS FOR PRENATAL GENDER DETERMINATION

(57) Abstract: Provided herein are compositions, processes and kits for noninvasive, early determination of fetal sex from, and/or amount of fetal nucleic acid in, an extracellular nucleic acid sample from a pregnant female. Such compositions, processes and kits are useful for detection of low genomic copy numbers of male fetal nucleic acid in a high copy number background of female nucleic acid, thereby determining the sex of a fetus and/or amount of fetal nucleic acid in a sample.



WO 2009/114543 A2

## NUCLEIC ACID-BASED TESTS FOR PRENATAL GENDER DETERMINATION

Related Patent Application

5 This patent application claims the benefit of United States Patent Application No. 61/035,711 filed on March 11, 2008, entitled NUCLEIC ACID-BASED TESTS FOR PRENATAL GENDER DETERMINATION, naming Paul Oeth and Payam Mahboubi as inventors, and designated by attorney docket no. SEQ-6018-PV. The entire content of the foregoing patent application is incorporated herein by reference, including all text, tables and drawings.

10

Field of the Invention

The invention relates in part to compositions, methods and kits for determining from an extracellular nucleic acid sample obtained from a pregnant female (i) the sex of a fetus, and/or (ii) the amount of fetal nucleic acid.

15

Background

Early determination of fetal sex is medically relevant when there is a family history of sex dependent or X-linked disorders. The presence of cell free fetal (cff) DNA in maternal plasma allows for fetal genotyping at an earlier time point than either amniocentesis or chorionic villus sampling with lower risk to the fetus (Lo et al., Lancet 350, 485-487 (1997)). However, the total amount and relative proportion of cff DNA in maternal plasma is lower in earlier stages of pregnancy. Because of low relative amounts of cff DNA during the first trimester, assays for fetal sex determination must have extraordinary sensitivity and specificity, and must be able to detect Y-chromosomal sequences at very low copy numbers.

25

Summary

30 Provided herein are compositions and processes for noninvasive, early determination of fetal sex. More specifically, the compositions and processes are useful for the detection of low genomic copy numbers of male DNA in a high copy number background of female DNA, thereby determining the sex of a fetus. The assay comprises a multiplexed PCR reaction to amplify a portion of the albumin (ALB) gene (which serves as a positive amplification control) and 8 Y-chromosomal targets to confirm the presence of male Y-chromosomal sequences. The multiple

35

Y-chromosomal regions provide increased sensitivity while maintaining high specificity. The assay was designed to target gene-based Y-chromosome sequences (termed “ampliconic”) primarily with testis-specific expression (see Table 1). With the exception of SRY and RBMY, Y-chromosomal sequences with evolutionary similarity to X-chromosomal sequences were avoided.

5 Any PCR-based detection method may be used to determine the fetal sex using the methods and compositions provided herein. In some embodiments, PCR amplification is followed by a primer extension reaction (for example, Sequenom’s TypePLEX™ assay) and detection of the extension products using mass spectrometry (for example, Sequenom’s MassARRAY® System).

10 Methods described herein may be performed in conjunction with other noninvasive prenatal tests, for example, but not limited to, detecting the presence of fetal nucleic in a sample, determining the relative amount of fetal nucleic acid in a sample, testing for chromosomal abnormalities or determining fetal blood type or RhD compatibility, where each of the assays may be performed alone or in combination with methods of the present invention. Examples of RhD and fetal

15 identifier assays are provided in U.S. Patent Application Number 12/027,954, which was filed February 7, 2008, and is hereby incorporated by reference.

Provided herein are compositions and methods for determining fetal sex. In some embodiments, the compositions and methods of the invention may be used to determine the presence or

20 absence of Y-chromosome nucleic acid in a sample from a pregnant female. In related embodiments, compositions and methods described herein may be used to determine the presence or absence of any of the genes in Table 1.

Also provided herein are compositions and methods that can be used to analyze a nucleic acid

25 sample for the presence or absence of one or more Y-chromosome genes from Table 1, comprising the steps of amplifying the one or more Table 1 genes (or amplicons therein) with one or more primer pairs provided in Table 3; determining the presence or absence of the amplification products from the amplification reaction, thereby determining the sex of the fetus where the presence of one or more Y-chromosome amplification products indicates the presence

30 of a male fetus and conversely the absence of one or more Y-chromosome amplification products indicates the presence of a female fetus. In related embodiments, the sample is blood from a pregnant female. In certain embodiments, the genes are analyzed in a multiplexed amplification reaction. In related embodiments, two or more multiplexed assays are performed in parallel. In certain embodiments, the sample is blood, plasma or serum from a pregnant female. In related

35 embodiments, the sample contains fetal nucleic acid and maternal nucleic acid. In related

embodiments, the primer pairs in Table 3 comprise a tag sequence to improve multiplexing. In certain embodiments, the presence or absence of amplification products is determined by mass spectrometry. In some embodiments, the presence or absence of amplification products is determined by detection of hybridization of the amplification products to a gene chip. In certain  
5 embodiments, the presence or absence of amplification products is determined by real time-PCR (alternatively called RT-PCR or Q-PCR).

In some embodiments, each primer of the amplification primer pair may comprise the entire sequence shown or only the non-underlined sequence, where the underlined portion of the primer  
10 is a tag sequence for improved multiplexing and the non-underlined portion is a sequence-specific primer sequence. The tag sequence may be any tag sequence known in the art that improves multiplexing. In certain embodiments, the invention in part includes primers having nucleotide sequences substantially identical to a nucleotide sequence of primers provided herein, for example, about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%,  
15 93%, 94%, 95%, 96%, 97%, 98%, 99% or 99% or more identical, or having only 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 mismatched nucleotides (e.g., mismatches are determined when the nucleotide sequences are aligned) and further where the primers still specifically hybridize to a given Y-chromosome region (e.g., gene). For example, one or more bases of a primer sequence may be changed or substituted, for example with an inosine, but the primer still maintains the same  
20 specificity and plexing ability.

Provided herein are compositions and methods to analyze a nucleic acid sample for the presence or absence of one or more Y-chromosome genes from Table 1, comprising the steps of amplifying the one or more Y-chromosome genes with the primer pairs provided in Table 3;  
25 hybridizing (e.g., annealing) one or more extend primers to the amplification products of the first step (e.g., examples of extend primers are provided in Table 3); performing a primer extension reaction; and analyzing the primer extension products to determine the sex of a fetus. In certain embodiments, the presence or absence of primer extension products is determined by mass spectrometry. In certain embodiments, the presence or absence of primer extension products is  
30 determined by any method known in the art.

The invention is not limited by the detection method; therefore, the amplification products and/or primer extension products may be detected by any detection method, which includes but is not limited to, the use of hybridization probes and quantitative real time polymerase chain reaction  
35 (QRT-PCR), digital PCR, electrophoresis, pyrosequencing, primer extension, microarrays, chips

and sequencing. In certain embodiments, detection of amplification products and/or extension products can be carried out using the single tube methods described in US Patent Application 11/950,395, which was filed December 4, 2007, and is hereby incorporated by reference. In certain embodiments, whole genome sequencing methods may be utilized to detect nucleic acid  
5 from the Y-chromosome. Examples of whole genome sequencing methods include, but are not limited to, nanopore-based sequencing methods, sequencing by synthesis and sequencing by ligation.

In some embodiments, a primer extension reaction includes incorporation of a chain terminating  
10 nucleotide. In related embodiments, the chain terminating nucleotide is a dideoxynucleotide, dideoxybromouridine or acyclonucleotide. In certain embodiments, the extension reaction comprises incorporation of a deoxynucleotide, a dideoxynucleotide or a combination thereof. In certain embodiments, the extension reaction comprises incorporation of a labeled nucleotide. In related embodiments, the extension reaction comprises using a mixture of labeled and unlabeled  
15 nucleotides. In related embodiments, the labeled nucleotide is labeled with a molecule selected from the group consisting of radioactive molecule, fluorescent molecule, mass label, antibody, antibody fragment, hapten, carbohydrate, biotin, derivative of biotin, phosphorescent moiety, luminescent moiety, electrochemiluminescent moiety, chromatic moiety, and moiety having a detectable electron spin resonance, electrical capacitance, dielectric constant and electrical  
20 conductivity. In related embodiments, the labeled nucleotide is labeled with a fluorescent molecule.

In some embodiments the sample is blood from a pregnant female. In certain embodiments, the biological sample is from any animal, including but not limited to, human, non-human, mammal,  
25 reptile, cattle, cat, dog, goat, swine, pig, monkey, ape, gorilla, bull, cow, bear, horse, sheep, poultry, mouse, rat, lion, tiger, fish, dolphin, whale, and shark, or any animal or organism that may be subjected to prenatal gender determination.

Methods described herein are useful across a range of gestational ages (see Example 1). Thus  
30 in some embodiments, a sample is obtained from a human pregnant female when the fetus is at a gestational age selected from the group consisting of: 0-4, 4-8, 8-12, 12-16, 16-20, 20-24, 24-28, 28-32, 32-36, 36-40, 40-44, 44-48, 48-52, and more than 52 weeks. In related embodiments, the sample is obtained through non-invasive means. In certain embodiments, the nucleic acid is obtained from plasma from blood. In certain embodiments, the nucleic acid is obtained from

serum from blood. In some embodiments, the biological sample contains cellular elements or cellular remnants in maternal blood.

In certain embodiments, the fetal nucleic acid may be extracted from maternal body fluids, sometimes whole blood, and often plasma or serum, using e. g. DNA extraction methods such as, but not limited to, gelatin extraction method; silica, glass bead, or diatom extraction method; guanidinium thiocyanate acid- phenol based extraction methods; guanidinium thiocyanate acid based extraction methods; guanidine-hydrochloride based extraction methods; methods using centrifugation through cesium chloride or similar gradients; phenol-chloroform based extraction methods; and/or other available DNA extraction methods, as are known in the art for use in extraction of intracellular DNA, including commercially available DNA extraction methods, e. g. by using or adapting or modifying the methods of Boom et al. (1990, J. Clin. Microbiol. 28: 495-503); Cheung et al. (1994, J. Clin. Microbiol. 32: 2593-2597); Boom et al. (1991, J. Clin. Microbiol. 29: 1804-1811); Chomczynski and Sacchi (1987, Analytical Biochem. 162: 156-159); Chomczynski, (1993, Biotech. 15: 532-537); Chomczynski and Mackey (1995, Biotechniques 19: 942-945); Chomczynski and Mackey (1995, Anal. Biochem. 225: 163-164); Chirgwin et al. (1979, Biochem. 18: 5294-5299); Fournie et al. (1986 Anal. Biochem. 158: 250-256); and W097/35589.

While the invention is not limited by how the sample is obtained, methods and compositions described herein are particularly useful for assaying samples obtained by non-invasive means, which may contain lower amounts of fetal nucleic acid to be assayed. In related embodiments, the sample is processed to selectively enrich fetal nucleic acid. In certain embodiments, a sample is enriched or relatively enriched for fetal nucleic acid. Methods for enriching a sample for a particular species of nucleic acid are described, for example, in International Patent Application Number PCT/US07/69991, filed May 30, 2007, International Patent Application Number PCT/US2007/071232, filed June 15, 2007, International Patent Application Number PCT/US2008/074689 and International Patent Application Number PCT/US2008/074692 and International Patent Application Number PCT/EP05/012707, which are all hereby incorporated by reference.

Multiplexed reactions can be utilized in methods described herein to improve throughput and reduce cost. Thus, provided herein are optimized methods for performing a primer mass extension assay, including an optimized PCR amplification reaction that produces amplified targets for subsequent multiplexed primer mass extension genotyping analysis using mass spectrometry. Also provided herein are optimized methods for performing multiplexed

amplification reactions and multiplexed primer mass extension reactions in a single reaction vessel (e.g., single well or pit of plate or chip, or single reaction tube) to further increase the throughput and reduce the cost per genotype for primer mass extension reactions. Certain nucleic acid target-region amplification and primer mass extension genotyping reactions have  
5 been optimized herein to permit moderate to high level multiplexing reactions with greater efficiency and accuracy, while at the same time not adversely affecting the mass spectrometry analysis of mass extension products.

In some embodiments, the amplification primers provided in Table 3 comprises a 5' tag  
10 (underlined) and a gene-specific sequence. The tag can be used to assist in the amplification of the nucleic acids. The primer tags may serve to stabilize the primer during amplification or they may serve as universal primer sites. More specifically, once the target gene nucleic acids of Table 1 have been PCR amplified using the primers, primers to the tags are used to further amplify the sequences in certain embodiments. In some embodiments, both amplification steps  
15 are performed simultaneously. As will be appreciated by those skilled in the art, primers without the 5' tag can be used in the method of the invention to amplify the target gene nucleic acids. Alternatively, the primer sequences can comprise different tag sequences than the tags indicated in the Table. Tag sequences useful for multiplex amplification reactions are known in the art.

20 In certain embodiments, the amplification primers allow for sequence specific amplification. For example, the PCR primers are designed to discriminate against amplification of similar genes or paralogs that are on other chromosomes by taking advantage of sequence differences between the target nucleic acids of Table 1 and any paralogs from other chromosomes.

25 In particular embodiments, a sequence tag is attached to a plurality of primary and secondary primer pairs provided in Table 3. The sequence tag can be attached to either one or both of the primary and secondary primers from each pair. Typically, the sequence tag is attached to the primary and secondary primer of each pair. The sequence tags used herein can range from 5 up to 20, from 5 up to 30, from 5 up to 40, or from 5 up to 50 nucleotides in length, with a sequence  
30 tag of 10-mer length being particularly useful in the methods provided herein. The sequence tag need not be the same sequence for each primer pair in a multiplexed amplification reaction, nor the same sequence for a primary and secondary primer within a particular amplification pair. In a particular embodiment, the sequence tag is the same for each primer in the multiplexed amplification reaction. For example, in certain embodiments, the sequence tag is a 10-mer, such  
35 as -ACGTTGGATG-, and is attached to the 5' end of each primary and secondary primer. In

particular embodiments of the methods provided herein, only a single primer pair is used to amplify each particular nucleic acid target-region (e.g., a "universal primer").

5 Methods and compositions described herein may be combined with other prenatal tests, including the use of fetal identifiers to detect the presence or absence of fetal nucleic acid in a maternal sample. Thus In some embodiments, fetal identifier compositions and methods are also provided for analyzing a plurality of polymorphisms in a nucleic acid sample of fetal origin; and analyzing a plurality of polymorphisms in a nucleic acid sample of maternal origin, whereby the presence of at least one polymorphism in the nucleic acid sample of fetal origin, which is not present in the  
10 nucleic acid sample of maternal origin, confirms the presence of fetal nucleic acid in the fetal nucleic acid sample. In related embodiments, the presence of at least one polymorphism in the nucleic acid sample of fetal origin, which is not present in the nucleic acid sample of maternal origin, is a paternally-inherited allele. In certain embodiments, the same polymorphisms are analyzed in fetal nucleic acid and maternal nucleic acid. In certain embodiments, the  
15 polymorphism is heterozygous. The plurality of polymorphisms may include 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000 or more polymorphisms. In related embodiments, the polymorphism is a single nucleotide polymorphism (SNP), insertion/deletion, short tandem repeats (STRs), RFLPs or any other alternate form of a gene, genomic DNA or non-coding region  
20 of DNA that occupies the same position on a chromosome. The polymorphism may be naturally-occurring or synthetic. Synthetic polymorphisms may include alternative forms introduced on a synthetic oligonucleotide that serve as a competitor or control. In some embodiments, the sequence variation falls in a restriction site, whereby one allele is susceptible to digestion by a restriction enzyme and the one or more other alleles are not. In certain embodiments, the  
25 sequence variation is a methylation site.

In certain embodiments, a method for detecting the presence or absence of fetal nucleic acid in a sample comprises obtaining or possessing a nucleic acid sample known to be of maternal origin and suspected of comprising fetal nucleic acid; analyzing the nucleic acid sample to determine  
30 the maternal genotype of at one or more nucleotide polymorphisms; and analyzing the nucleic acid sample to determine the fetal genotype of one or more nucleotide polymorphisms, where a fetal genotype possessing a paternally-inherited allele indicates the presence of fetal nucleic acid. In related embodiments, the maternal genotypes are determined from DNA that is substantially free of fetal nucleic acid. For example, in the case when the sample is blood, the maternal  
35 genotypes may be determined from the portion of the blood that comprises nucleated maternal



cells (e.g., white blood cells). In some embodiments, the DNA that is substantially free of fetal nucleic acid is from peripheral blood mononuclear cells. In certain embodiments, the amount of fetal DNA is determined by comparing the relative amount of paternally-inherited alleles to maternally-inherited alleles in fetal nucleic acid.

5

The invention in part also provides compositions and methods to determine the relative amount of target nucleic acid in a sample (e.g., fetal nucleic acid in a pregnant female sample). In some embodiments, compositions and methods described herein may be used to quantify the relative amount of the alleles at a heterozygous polymorphic site, where the heterozygous polymorphic site has been identified by determining the sequence of alleles at a polymorphic site from template DNA obtained from a maternal sample, where the relative amount is expressed as a ratio, and the ratio indicates the relative amount of fetal nucleic acid present in the maternal sample. In certain embodiments, the polymorphic site is an insertion/deletion, STR or RFLP.

10

15 In certain embodiments, the heterozygous polymorphic site is part of a restriction site, and a restriction enzyme is introduced that can discriminate between the alleles of the polymorphic site. The maternal allele is digested and the nucleic acid comprising the paternal allele is relatively enriched in some embodiments.

20

In certain embodiments, the total copy number of nucleic acid molecules for the human serum albumin (ALB) gene is determined. Methods for determining the total copy number of nucleic acid present in a sample comprise detecting albumin-specific extension products and comparing the relative amount of the extension products to competitors introduced to the sample. In related embodiments, Provided herein are compositions and methods to determine the relative amount of fetal DNA in a sample (e.g., plasma of a pregnant woman carrying a male fetus), which comprises annealing one or more albumin gene sequences to the fetal DNA; and analyzing the primer extension products to determine the relative amount of ALB extension products. The assay is useful to measure how much nucleic acid (e.g., total copy number) is present in a sample or loaded into a particular reaction. The assay may serve as an internal control and a guide to the likelihood of success for a particular PCR reaction.

25

30

In certain embodiments, the methods and compositions of the invention are used to detect Y-chromosome nucleic acid at low copy numbers. In some embodiments, the fetal nucleic acid copy number is about 5 to about 1000 copies, and sometimes about 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39,

35

40, 45, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 500, 600, 700, 800, 900 or 1000 or more.

5 In certain embodiments, methods described herein include steps to reduce the introduction of non-fetal Y-chromosome. For example, samples sometimes are collected and/or processed by a female.

Embodiments of the invention are described further in the following brief description of the drawings, detailed description and claims.

10

#### Brief Description of the Drawings

FIG. 1A-1I provide the location design of the fetal sex amplification and extend primers within the target nucleic acids.

15

FIG. 2 shows results from the maternal plasma sample analysis. Samples were assayed in either August 2007 (20070829) or November 2007 (20071130) as indicated by the light and dark gray bars. The top panel shows the number of Y-target calls for each sample. Bottom panel shows the ALB call for each sample to indicate successful PCR. Samples within brackets at the bottom indicate fetal gender phenotype information for those where it was known.

20

FIG. 3 shows the results from the analysis of non-pregnant female plasma DNA. Top panel shows the number of Y-target calls for each replicate non-pregnant female plasma DNA or NTC. Bottom panel shows the presence or absence of ALB call for each sample to indicate successful PCR. (NTC = non- template control).

25

#### Detailed Description

Early determination of fetal sex is medically relevant, for example, when there is a family history of sex dependent or X-linked disorders such as congenital adrenal hyperplasia (CAH), Duchennes muscular dystrophy, hemophilia or Fragile X. The presence of fetal nucleic acid in maternal plasma allows for fetal genotyping at an earlier time point than either amniocentesis or chorionic villus sampling and with lower secondary risk to the fetus. However, cell-free fetal nucleic acid constitutes only a minority of the total DNA in maternal plasma (typically less than 6%) (Lo YM et. al., Am J Hum Genet. 62, 768-75. 1998). Additionally, the total amount and

35

relative proportion of fetal nucleic acid in maternal plasma is lowest in the first trimester when fetal sex determination is most relevant, especially in cases of CAH. Therefore, assays for fetal sex determination must be highly sensitive and specific.

5 In certain embodiments, assays comprises a multiplexed PCR to amplify a portion of the albumin (ALB) gene (which serves as a positive amplification control) and 8 Y-chromosomal targets to confirm the presence of male Y-chromosomal sequences. The Y-chromosomal markers used in the assay design were chosen from targets in the male specific regions of the Y-chromosome, and cover sites in both the long and short arm of the Y-chromosome. Markers include the genes and transcriptional units indicated in Table 1. An emphasis was placed on multi-copy loci because the nature of these targets is predicted to allow higher rates of detection in situations with low fetal copy numbers typically seen in maternal plasma at early stages of pregnancy. The multiple Y-chromosomal regions provide increased sensitivity while maintaining high specificity in such embodiments. The assays were designed to target gene-based Y-  
 10 chromosome sequences (termed “ampliconic”) primarily with testis-specific expression (Table 1). With the exception of SRY and RBMY, Y-chromosomal sequences with evolutionary similarity to X-chromosomal sequences were avoided.

TABLE 1

Class	Gene/Txn Symbol	Gene/Txn Name	No. copies	Tissue expression	X-linked homologue	Autosomal homologue
ampliconic	HSFY	Heat shock transcription factor Y	2	testis	---	---
	RBMY	RNA-binding motif Y	6	testis	RBMX	---
	PRY	PTP-BL related Y	2	testis	---	---
	BPY2	Basic protein Y 2	3	testis	---	---
	XKRY	XK related Y	2	testis	---	---
	CDY	Chromodomain	4	testis	---	CDYL
	TTY	Non-coding transcription units		---	---	---
X-degenerate	CYORF14	Chromosome Y open reading frame 14		---	---	---
	SRY	Sex determining region Y	1	testis	SOX3	---

20 In some embodiments, PCR amplification is followed by a primer extension reaction and detection of the extension products using mass spectrometry. Other detection methods may be used to detect the amplification products or extension products.

In certain assay embodiments, initial testing on male and female genomic DNA showed the ability of the assay system to discriminate the presence or absence of Y-chromosomal DNA.

5 Secondary studies were performed using DNA mixtures with as low as 20 genomic copies of either male or female DNA mixed with a 50-fold excess of female genomic DNA. Out of 88 total DNA mixtures, 40 female:female DNA mixtures were used to evaluate spurious amplification of individual Y-chromosomal targets. In 48 male:female DNA mixtures, ALB was amplified in all cases and all 8 Y-chromosomal target regions were detected consistently. These results demonstrate high sensitivity and specificity for accurately detecting Y-chromosomal DNA. Initial testing of the assay using maternal plasma and non-pregnant female plasma samples has been performed and shows a high degree of reproducibility between replicate sample aliquots of maternal plasma and high specificity as determined by the lack of Y-chromosomal target detection in plasma DNA isolated from non-pregnant female donors. Compared to other cell-free fetal nucleic acid tests for SRY, the use of methods and compositions described herein increases specificity of the assay while maintaining high sensitivity with utility for early determination of fetal sex in a clinical setting.

As used herein, "sample" refers to a composition containing a material to be detected or analyzed. Samples include "biological samples", which refer to any material obtained from a living source, for example, an animal such as a human or other mammal. The sample may be obtained through invasive (e.g., amniocentesis) or non-invasive (e.g., blood draw) means. In some embodiments, the sample is obtained non-invasively. The biological sample can be in any form that potentially includes fetal nucleic acid, including a solid material such as a tissue, cells, a cell pellet, a cell extract, or a biological fluid such as urine, whole blood, plasma, serum, interstitial fluid, vaginal swab, pap smear, peritoneal fluid, sweat, saliva, follicular fluid, breast milk, non-milk breast secretions, cerebral spinal fluid, seminal fluid, lung sputum, amniotic fluid, a mouth wash containing buccal cells, synovial fluid, or any other fluid sample produced by the subject. If desired, solid materials can be mixed with a fluid or purified or amplified or otherwise treated. Samples examined using the methods described herein can be treated in one or more purification steps in order to increase the purity of the desired cells or nucleic acid in the sample. Samples also can be examined using the methods described herein without any purification steps to increase the purity or relative concentration of desired cells or nucleic acid. As used herein, the term "blood" encompasses whole blood or any fractions of blood, such as serum and plasma as conventionally defined. In certain embodiments, nucleic acid is isolated from a sample that is cell free, acellular or extracellular (e.g., blood plasma, blood serum, urine).

The terms "nucleic acid" and "nucleic acid molecule" may be used interchangeably throughout the disclosure. The terms refer to a deoxyribonucleotide (DNA), ribonucleotide polymer (RNA), RNA/DNA hybrids and polyamide nucleic acids (PNAs) in either single- or double-stranded form, and unless otherwise limited, would encompass known analogs of natural nucleotides that can function in a similar manner as naturally occurring nucleotides. In the case of RNA or mRNA, a DNA copy (cDNA) of the RNA transcript of interest often is synthesized. This synthesis often is achieved by reverse transcription, which can be carried out as a separate step in certain embodiments. In a homogeneous reverse transcription-polymerase chain reaction (RT-PCR) embodiment, a modification of the polymerase chain reaction for amplifying RNA.

As used herein, the term "multiplexing" refers to simultaneous amplification and/or primer mass extension reaction of more than one oligonucleotide or primer (*e.g.*, in a single reaction container); or the simultaneous analysis of more than one oligonucleotide, in a single mass spectrometric or other mass measurement, *i.e.*, a single mass spectrum or other method of reading sequence. Examples of different multiplex schemes (MP1, MP2, MP3 and MP4) are provided in Table 3.

As used herein, the phrase "target nucleic acid" refers to one or more nucleic acids, such as genomic DNA, from which one or more regions or loci are to be amplified. As used herein, the phrase "nucleic acid-target region" refers to the region-specific areas or loci of a target nucleic acid (*e.g.*, UTR, exon or intron) that are amplified for subsequent fetal sex determination. For example, Figures 1A-1I provide target nucleic acid regions from the genes listed in Table 1.

As used herein, the term "polymorphism" refers to the coexistence of more than one form or allele of a nucleic acid, such as a chromosome, or portion thereof. For example, a portion or locus of a gene at which there are at least two different alleles, *i.e.*, two different nucleotide sequences, is referred to as a polymorphic loci, site or region of a gene. A polymorphic loci can be a single nucleotide (*e.g.*, SNP) or can be several nucleotides in length (*e.g.*, insertions or deletions). Accordingly, polymorphism includes substitutions, insertions, duplications and deletions of nucleotides. A polymorphism can also refer to a particular nucleotide(s) or nucleotide sequence occurring at a particular polymorphic site.

As used herein, the term "genotyping" refers to the process of determining the particular nucleotide or nucleotides (*e.g.*, sequence variation) either present or absent at a particular

polymorphic loci or genomic location. As used herein, the term "genotype" refers to the identity of the alleles or non-homologous variants present in an individual or sample. The term "genotyping a sample" or "genotyping an individual" refers to determining a specific allele or specific nucleotide(s) in a sample or carried by an individual at particular region(s).

5

As used herein, the term "allele" refers to a variant at a particular locus, namely a nucleic acid sequence variant (e.g., sequence variations such as single nucleotide polymorphisms, copy number variations, short tandem repeats (STRs), insertion/deletion polymorphisms and methylation sites), at the particular locus of interest. A locus is a site on a chromosome or, for example, a transcript thereof. When a subject has two identical alleles of a polymorphic region within a gene, the subject is said to be homozygous for the allele. When a subject has two different alleles of a polymorphic region within a gene, the subject is said to be heterozygous for the allele.

#### 15 Amplification

In some embodiments, it may be desirable to amplify a target sequence using any of several nucleic acid amplification procedures (described in greater detail herein). Nucleic acid amplification may be particularly beneficial when target sequences exist at low copy number, or the target sequences are non-host sequences and represent a small portion of the total nucleic acid in the sample (e.g., fetal nucleic acid in a maternal nucleic acid background). In some embodiments, amplification of target sequences may aid in detection of one or more Y-chromosome nucleotide sequences, for example.

25 Nucleic acid amplification often involves enzymatic synthesis of nucleic acid amplicons (copies), which contain a sequence complementary to a nucleotide sequence species being amplified. An amplification product (amplicon) of a particular nucleotide sequence species (e.g., target sequence) is referred to herein as an "amplified nucleic acid species." Amplifying target sequences and detecting the amplicon synthesized, can improve the sensitivity of an assay, since  
30 fewer target sequences are needed at the beginning of the assay, and can improve detection of target sequences.

The terms "amplify", "amplification", "amplification reaction", or "amplifying" refers to any in vitro process for multiplying the copies of a target sequence of nucleic acid. Amplification sometimes  
35 refers to an "exponential" increase in target nucleic acid. However, "amplifying" as used herein

can also refer to linear increases in the numbers of a select target sequence of nucleic acid, but is different than a one-time, single primer extension step. In some embodiments a limited amplification reaction, also known as pre-amplification, can be performed. Pre-amplification is a method in which a limited amount of amplification occurs due to a small number of cycles, for example 10 cycles, being performed. Pre-amplification can allow some amplification, but stops amplification prior to the exponential phase, and typically produces about 500 copies of the desired nucleotide sequence(s). Use of pre-amplification may also limit inaccuracies associated with depleted reactants in standard polymerase chain reaction (PCR) reactions, and also may reduce amplification biases due to nucleotide sequence or species abundance of the target. In some embodiments, a one-time primer extension may be used may be performed as a prelude to linear or exponential amplification. In some embodiments, amplification of the target nucleic acid may not be required, due to the use of ultra sensitive detections methods (e.g., single nucleotide sequencing, sequencing by synthesis and the like).

Based on the 5' and 3' primers that are chosen, amplification often serves to restrict and define a target-region or locus of the genome subject to analysis. Amplification can be by any process known, and in particular embodiments, includes the use of PCR. The phrase "simultaneous amplification" as used herein refers to amplification of two or more nucleic acid target-regions at the same time. Simultaneous amplification often is performed within the same amplification mixture. As used herein, the phrase "simultaneous amplification" refers to the multiplexed amplification of two or more loci or nucleic acid target-regions in a single reaction mixture. Simultaneous amplification therefore encompasses 5 or more, 6 or more, 7 or more, 8 or more, 9 or more, 10 or more, 11 or more, 12 or more, 13 or more, 14 or more, 15 or more, 16 or more, 17 or more, 18 or more, 19 or more, 20 or more, 30 or more, 40 or more, 50 or more, 60 or more, 70 or more, 80 or more, 100 or more, 200 or more, 500 or more, 1000 or more, 2000 or more amplification reactions. Amplification of each particular target-region occurs in parallel at the same time. Although it is contemplated herein that the simultaneous amplifications can occur in separate reaction mixtures, for the methods provided herein the simultaneous amplification reactions typically occur in the same single reaction. Likewise multiplexed primer mass extension refers to the simultaneous extension of 2 or more extend primers in a single reaction mixture. Accordingly, multiplexed primer mass extension therefore encompasses 5 or more, 6 or more, 7 or more, 8 or more, 9 or more, 10 or more, 11 or more, 12 or more, 13 or more, 14 or more, 15 or more, 16 or more, 17 or more, 18 or more, 19 or more, 20 or more, 30 or more, 40 or more, 50 or more, 60 or more, 70 or more, 80 or more, 100 or more, 200 or more, 500 or more, 1000 or more, 2000 or more primer mass extension reactions. Multiplexed amplification and primer mass

extension reactions also encompass 21, 22, 23, 24, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 60, 70, 80, 100, 1000 or more reactions.

Where amplification may be desired, any suitable amplification technique can be utilized. Non-limiting examples of methods for amplification of polynucleotides include, polymerase chain reaction (PCR); ligation amplification (or ligase chain reaction (LCR)); amplification methods based on the use of Q-beta replicase or template-dependent polymerase (see US Patent Publication Number US20050287592); helicase-dependant isothermal amplification (Vincent et al., "Helicase-dependent isothermal DNA amplification". EMBO reports 5 (8): 795–800 (2004)); strand displacement amplification (SDA); thermophilic SDA nucleic acid sequence based amplification (3SR or NASBA) and transcription-associated amplification (TAA). Non-limiting examples of PCR amplification methods include standard PCR, AFLP-PCR, Allele-specific PCR, Alu-PCR, Asymmetric PCR, Colony PCR, Hot start PCR, Inverse PCR (IPCR), In situ PCR (ISH), Intersequence-specific PCR (ISSR-PCR), Long PCR, Multiplex PCR, Nested PCR, Quantitative PCR, Reverse Transcriptase PCR (RT-PCR), Real Time PCR, Single cell PCR, Solid phase PCR, combinations thereof, and the like. Reagents and hardware for conducting PCR are commercially available.

In some embodiments, amplification of Y-chromosome loci described herein may be accomplished by any suitable method available to one of skill in the art or selected from the listing above (e.g., ligase chain reaction (LCR), transcription-mediated amplification, and self-sustained sequence replication or nucleic acid sequence-based amplification (NASBA)). More recently developed branched-DNA technology may also be used to amplify the signal of the Y-chromosome loci described herein. For a review of branched-DNA (bDNA) signal amplification for direct quantification of nucleic acid sequences in clinical samples, see Nolte, Adv. Clin. Chem. 33:201-235, 1998.

Amplification also can be accomplished using digital PCR, in certain embodiments (e.g., Kalinina and colleagues (Kalinina et al., "Nanoliter scale PCR with TaqMan detection." Nucleic Acids Research. 25; 1999-2004, (1997); Vogelstein and Kinzler (Digital PCR. Proc Natl Acad Sci U S A. 96; 9236-41, (1999); PCT Patent Publication No. WO05023091A2 (incorporated herein in its entirety); US Patent Publication No. 20070202525 (incorporated herein in its entirety)). Digital PCR takes advantage of nucleic acid (DNA, cDNA or RNA) amplification on a single molecule level, and offers a highly sensitive method for quantifying low copy number nucleic acid. Systems for digital amplification and analysis of nucleic acids are available (e.g., Fluidigm® Corporation).



In some embodiments, where RNA nucleic acid species may be used for detection of certain nucleotide sequences (e.g., fetal nucleotide sequences), a DNA copy (cDNA) of the RNA transcripts of interest can be synthesized prior to the amplification step. The cDNA copy can be synthesized by reverse transcription, which may be carried out as a separate step, or in a homogeneous reverse transcription-polymerase chain reaction (RT-PCR), a modification of the polymerase chain reaction for amplifying RNA. Methods suitable for PCR amplification of ribonucleic acids are described by Romero and Rotbart in *Diagnostic Molecular Biology: Principles and Applications* pp. 401-406; Persing et al., eds., Mayo Foundation, Rochester, Minn., 1993; Egger et al., *J. Clin. Microbiol.* 33:1442-1447, 1995; and U.S. Pat. No. 5,075,212.

Use of a primer extension reaction also can be applied in methods of the invention. A primer extension reaction operates, for example, by discriminating nucleic acid sequences, SNP alleles for example, at a single nucleotide mismatch (e.g., a mismatch between paralogous sequences, or SNP alleles). The terms "paralogous sequence" or "paralogous sequences" refer to sequences that have a common evolutionary origin but which may be duplicated over time in the genome of interest. Paralogous sequences may conserve gene structure (e.g., number and relative position of introns and exons and sometimes transcript length), as well as sequence. Therefore, the methods described herein can be used to detect sequence mismatches in SNP-alleles or in evolutionarily conserved regions that differ by one or more point mutations, insertions or deletions (both will hereinafter be referred to as "mismatch site" or "sequence mismatch").

The mismatch may be detected by the incorporation of one or more deoxynucleotides and/or dideoxynucleotides to a primer extension primer or oligonucleotide, which hybridizes to a region adjacent to the SNP site (e.g., mismatch site). The extension oligonucleotide generally is extended with a polymerase. In some embodiments, a detectable tag or detectable label is incorporated into the extension oligonucleotide or into the nucleotides added on to the extension oligonucleotide (e.g., biotin or streptavidin). The extended oligonucleotide can be detected by any known suitable detection process (e.g., mass spectrometry; sequencing processes). In some embodiments, the mismatch site is extended only by one or two complementary deoxynucleotides or dideoxynucleotides that are tagged by a specific label or generate a primer extension product with a specific mass, and the mismatch can be discriminated and quantified.

For embodiments using primer extension to amplify a target sequence, the primer extension is not limited to a single round of extension, and is therefore distinguished from "one-time primer

extension” described above. Non-limiting examples of primer extension or oligonucleotide extension methods suitable for use with embodiments described herein are described in U.S. Patent Nos. 4,656,127; 4,851,331; 5,679,524; 5,834,189; 5,876,934; 5,908,755; 5,912,118; 5,976,802; 5,981,186; 6,004,744; 6,013,431; 6,017,702; 6,046,005; 6,087,095; 6,210,891; and 5 WO 01/20039, for example.

A generalized description of an amplification process is presented herein. Primers and target nucleic acid are contacted, and complementary sequences anneal to one another, for example. Primers can anneal to a nucleic acid, at or near (e.g., adjacent to, abutting, and the like) a target sequence of interest. A reaction mixture, containing all components necessary for full enzymatic functionality, is added to the primer – target nucleic acid hybrid, and amplification can occur under suitable conditions. Components of an amplification reaction may include, but are not limited to, e.g., primers (e.g., individual primers, primer pairs, primer sets and the like) a polynucleotide template (e.g., nucleic acid containing a target sequence), polymerase, nucleotides, dNTPs and 15 the like. In some embodiments, non-naturally occurring nucleotides or nucleotide analogs, such as analogs containing a detectable label (e.g., fluorescent or colorimetric label) may be used, for example. Polymerases can be selected by a person of ordinary skill and include polymerases for thermocycle amplification (e.g., Taq DNA Polymerase; Q-Bio™ Taq DNA Polymerase (recombinant truncated form of Taq DNA Polymerase lacking 5’-3’exo activity); SurePrime™ 20 Polymerase (chemically modified Taq DNA polymerase for "hot start" PCR); Arrow™ Taq DNA Polymerase (high sensitivity and long template amplification)) and polymerases for thermostable amplification (e.g., RNA polymerase for transcription-mediated amplification (TMA) described at World Wide Web URL “gen-probe.com/pdfs/tma\_whiteppr.pdf”). Other enzyme components can be added, such as reverse transcriptase for transcription mediated amplification (TMA) reactions, 25 for example.

The terms “near” or “adjacent to” when referring to a nucleotide target sequence refers to a distance or region between the end of the primer and the nucleotide or nucleotides of interest. As used herein adjacent is in the range of about 5 nucleotides to about 500 nucleotides (e.g., about 30 5 nucleotides away from nucleotide of interest, about 10, about 20, about 30, about 40, about 50, about 60, about 70, about 80, about 90, about 100, about 150, about 200, about 250, about 300, about 350, about 400, about 450 or about 500 nucleotides from a nucleotide of interest).

Each amplified nucleic acid species independently can be about 10 to about 1000 base pairs in length in some embodiments. In certain embodiments, an amplified nucleic acid species is about 35

20 to about 250 base pairs in length, sometimes is about 50 to about 150 base pairs in length and sometimes is about 100 base pairs in length. Thus, in some embodiments, the length of each of the amplified nucleic acid species products independently is about 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 5 114, 116, 118, 120, 125, 130, 135, 140, 145, 150, 175, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950 or 1000 base pairs (bp) in length.

An amplification product may include naturally occurring nucleotides, non-naturally occurring nucleotides, nucleotide analogs and the like and combinations of the foregoing. An amplification 10 product often has a nucleotide sequence that is identical to or substantially identical to a target sequence or complement thereof. A "substantially identical" nucleotide sequence in an amplification product will generally have a high degree of sequence identity to the nucleotide sequence species being amplified or complement thereof (e.g., about 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 15 96%, 97%, 98%, 99% or greater than 99% sequence identity), and variations sometimes are a result of infidelity of the polymerase used for extension and/or amplification, or additional nucleotide sequence(s) added to the primers used for amplification.

PCR conditions can be dependent upon primer sequences, target abundance, and the desired 20 amount of amplification, and therefore, one of skill in the art may choose from a number of PCR protocols available (see, e.g., U.S. Pat. Nos. 4,683,195 and 4,683,202; and PCR Protocols: A Guide to Methods and Applications, Innis et al., eds, 1990. PCR often is carried out as an automated process with a thermostable enzyme. In this process, the temperature of the reaction mixture is cycled through a denaturing region, a primer-annealing region, and an extension 25 reaction region automatically. Machines specifically adapted for this purpose are commercially available. A non-limiting example of a PCR protocol that may be suitable for embodiments described herein is, treating the sample at 95°C for 5 minutes; repeating forty-five cycles of 95°C for 1 minute, 59°C for 1 minute, 10 seconds, and 72°C for 1 minute 30 seconds; and then treating the sample at 72°C for 5 minutes. Multiple cycles frequently are performed using a commercially 30 available thermal cycler. Suitable isothermal amplification processes known and selected by the person of ordinary skill in the art also may be applied, in certain embodiments.

In some embodiments, multiplex amplification processes may be used to amplify target sequences, such that multiple amplicons are simultaneously amplified in a single, homogenous 35 reaction. As used herein "multiplex amplification" refers to a variant of PCR where simultaneous

amplification of many target sequences in one reaction vessel may be accomplished by using more than one pair of primers (e.g., more than one primer set). Multiplex amplification may be useful for analysis of deletions, mutations, and polymorphisms, or quantitative assays, in some embodiments. In certain embodiments multiplex amplification may be used for detecting paralog  
5 sequence imbalance, genotyping applications where simultaneous analysis of multiple markers is required, detection of pathogens or genetically modified organisms, or for microsatellite analyses. In some embodiments multiplex amplification may be combined with another amplification (e.g., PCR) method (e.g., nested PCR or hot start PCR, for example) to increase amplification specificity and reproducibility. In some embodiments, multiplex amplification processes may be  
10 used to amplify the Y-chromosome loci described herein.

In certain embodiments, nucleic acid amplification can generate additional nucleic acid species of different or substantially similar nucleic acid sequence. In certain embodiments described herein, contaminating or additional nucleic acid species, which may contain sequences substantially  
15 complementary to, or may be substantially identical to, the target sequence, can be useful for sequence quantification, with the proviso that the level of contaminating or additional sequences remains constant and therefore can be a reliable marker whose level can be substantially reproduced. Additional considerations that may affect sequence amplification reproducibility are;  
20 PCR conditions (number of cycles, volume of reactions, melting temperature difference between primers pairs, and the like), concentration of target nucleic acid in sample (e.g. fetal nucleic acid in maternal nucleic acid background, viral nucleic acid in host background), the number of chromosomes on which the nucleotide species of interest resides (e.g., paralogous sequences or SNP-alleles), variations in quality of prepared sample, and the like. The terms “substantially reproduced” or “substantially reproducible” as used herein refer to a result (e.g., quantifiable  
25 amount of nucleic acid) that under substantially similar conditions would occur in substantially the same way about 75% of the time or greater, about 80%, about 85%, about 90%, about 95%, or about 99% of the time or greater.

In some embodiments, amplification may be performed on a solid support. In some  
30 embodiments, primers may be associated with a solid support. In certain embodiments, target nucleic acid (e.g., template nucleic acid or target sequences) may be associated with a solid support. A nucleic acid (primer or target) in association with a solid support often is referred to as a solid phase nucleic acid.

In some embodiments, nucleic acid molecules provided for amplification are in a “microreactor”. As used herein, the term “microreactor” refers to a partitioned space in which a nucleic acid molecule can hybridize to a solid support nucleic acid molecule. Examples of microreactors include, without limitation, an emulsion globule (described hereafter) and a void in a substrate. A  
5 void in a substrate can be a pit, a pore or a well (e.g., microwell, nanowell, picowell, micropore, or nanopore) in a substrate constructed from a solid material useful for containing fluids (e.g., plastic (e.g., polypropylene, polyethylene, polystyrene) or silicon) in certain embodiments. Emulsion globules are partitioned by an immiscible phase as described in greater detail hereafter. In some  
10 embodiments, the microreactor volume is large enough to accommodate one solid support (e.g., bead) in the microreactor and small enough to exclude the presence of two or more solid supports in the microreactor.

The term “emulsion” as used herein refers to a mixture of two immiscible and unblendable substances, in which one substance (the dispersed phase) often is dispersed in the other  
15 substance (the continuous phase). The dispersed phase can be an aqueous solution (i.e., a solution comprising water) in certain embodiments. In some embodiments, the dispersed phase is composed predominantly of water (e.g., greater than 70%, greater than 75%, greater than 80%, greater than 85%, greater than 90%, greater than 95%, greater than 97%, greater than 98% and greater than 99% water (by weight)). Each discrete portion of a dispersed phase, such as an  
20 aqueous dispersed phase, is referred to herein as a “globule” or “microreactor.” A globule sometimes may be spheroidal, substantially spheroidal or semi-spheroidal in shape, in certain embodiments.

The terms “emulsion apparatus” and “emulsion component(s)” as used herein refer to apparatus  
25 and components that can be used to prepare an emulsion. Non-limiting examples of emulsion apparatus include without limitation counter-flow, cross-current, rotating drum and membrane apparatus suitable for use by a person of ordinary skill to prepare an emulsion. An emulsion component forms the continuous phase of an emulsion in certain embodiments, and includes without limitation a substance immiscible with water, such as a component comprising or  
30 consisting essentially of an oil (e.g., a heat-stable, biocompatible oil (e.g., light mineral oil)). A biocompatible emulsion stabilizer can be utilized as an emulsion component. Emulsion stabilizers include without limitation Atlox 4912, Span 80 and other biocompatible surfactants.

In some embodiments, components useful for biological reactions can be included in the  
35 dispersed phase. Globules of the emulsion can include (i) a solid support unit (e.g., one bead or

one particle); (ii) sample nucleic acid molecule; and (iii) a sufficient amount of extension agents to elongate solid phase nucleic acid and amplify the elongated solid phase nucleic acid (e.g., extension nucleotides, polymerase, primer). Inactive globules in the emulsion may include a subset of these components (e.g., solid support and extension reagents and no sample nucleic acid) and some can be empty (i.e., some globules will include no solid support, no sample nucleic acid and no extension agents).

Emulsions may be prepared using known suitable methods (e.g., Nakano et al. "Single-molecule PCR using water-in-oil emulsion;" Journal of Biotechnology 102 (2003) 117-124). Emulsification methods include without limitation adjuvant methods, counter-flow methods, cross-current methods, rotating drum methods, membrane methods, and the like. In certain embodiments, an aqueous reaction mixture containing a solid support (hereafter the "reaction mixture") is prepared and then added to a biocompatible oil. In certain embodiments, the reaction mixture may be added dropwise into a spinning mixture of biocompatible oil (e.g., light mineral oil (Sigma)) and allowed to emulsify. In some embodiments, the reaction mixture may be added dropwise into a cross-flow of biocompatible oil. The size of aqueous globules in the emulsion can be adjusted, such as by varying the flow rate and speed at which the components are added to one another, for example.

The size of emulsion globules can be selected by the person of ordinary skill in certain embodiments based on two competing factors: (i) globules are sufficiently large to encompass one solid support molecule, one sample nucleic acid molecule, and sufficient extension agents for the degree of elongation and amplification required; and (ii) globules are sufficiently small so that a population of globules can be amplified by conventional laboratory equipment (e.g., thermocycling equipment, test tubes, incubators and the like). Globules in the emulsion can have a nominal, mean or average diameter of about 5 microns to about 500 microns, about 10 microns to about 350 microns, about 50 to 250 microns, about 100 microns to about 200 microns, or about 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 200, 300, 400 or 500 microns in certain embodiments.

In certain embodiments, amplified nucleic acid species in a set are of identical length, and sometimes the amplified nucleic acid species in a set are of a different length. For example, one amplified nucleic acid species may be longer than one or more other amplified nucleic acid species in the set by about 1 to about 100 nucleotides (e.g., about 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 30, 40, 50, 60, 70, 80 or 90 nucleotides longer).

In some embodiments, a ratio can be determined for the amount of one amplified nucleic acid species in a set to the amount of another amplified nucleic acid species in the set (hereafter a "set ratio"). In some embodiments, the amount of one amplified nucleic acid species in a set is about equal to the amount of another amplified nucleic acid species in the set (i.e., amounts of amplified nucleic acid species in a set are about 1:1), which generally is the case when the number of chromosomes or the amount of DNA representative of nucleic acid species in a sample bearing each nucleotide sequence species amplified is about equal. The term "amount" as used herein with respect to amplified nucleic acid species refers to any suitable measurement, including, but not limited to, copy number, weight (e.g., grams) and concentration (e.g., grams per unit volume (e.g., milliliter); molar units). In some embodiments, the ratio of fetal nucleic acid to maternal nucleic acid (or conversely maternal nucleic acid to fetal nucleic acid) can be used in conjunction with measurements of the ratios of mismatch sequences for determination of chromosomal abnormalities possibly associated with sex chromosomes. That is, the percentage of fetal nucleic acid detected in a maternal nucleic acid background or the ratio of fetal to maternal nucleic acid in a sample, can be used in conjunction with the determination of the ratio Y-chromosome specific loci (e.g., the loci listed in Table 1) to detect chromosomal aneuploidies.

In certain embodiments, the amount of one amplified nucleic acid species in a set can differ from the amount of another amplified nucleic acid species in a set, even when the number of chromosomes in a sample bearing each nucleotide sequence species amplified is about equal. In some embodiments, amounts of amplified nucleic acid species within a set may vary up to a threshold level at which an outcome (e.g., (i) sex determination, (ii) presence or absence of Y-chromosome nucleic acid or (iii) amount of fetal nucleic acid) can be called with a confidence level of about 95% (e.g., about 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or greater than 99%). In certain embodiments, the amounts of the amplified nucleic acid species in a set vary by about 50% or less (e.g., about 45, 40, 35, 30, 25, 20, 15, 10, 5, 4, 3, 2 or 1%, or less than 1%). Thus, in certain embodiments amounts of amplified nucleic acid species in a set may vary from about 1:1 to about 1:1.5. Without being limited by theory, certain factors can lead to the observation that the amount of one amplified nucleic acid species in a set can differ from the amount of another amplified nucleic acid species in a set, even when the number of chromosomes in a sample bearing each nucleotide sequence species amplified is about equal. Such factors may include different amplification efficiency rates and/or amplification from a chromosome not intended in the assay design.

35

Each amplified nucleic acid species in a set generally is amplified under conditions that amplify that species at a substantially reproducible level. The term "substantially reproducible level" as used herein refers to consistency of amplification levels for a particular amplified nucleic acid species per unit template nucleic acid (e.g., per unit template nucleic acid that contains the particular nucleotide sequence species amplified). A substantially reproducible level varies by about 1% or less in certain embodiments, after factoring the amount of template nucleic acid giving rise to a particular amplification nucleic acid species (e.g., normalized for the amount of template nucleic acid). In some embodiments, a substantially reproducible level varies by 5%, 4%, 3%, 2%, 1.5%, 1%, 0.5%, 0.1%, 0.05%, 0.01%, 0.005% or 0.001% after factoring the amount of template nucleic acid giving rise to a particular amplification nucleic acid species.

In some embodiments amplification nucleic acid species (e.g., amplified target sequences) of primer sets may be generated in one reaction vessel. In some embodiments amplification of mismatch sequences may be performed in a single reaction vessel. In certain embodiments, mismatch sequences (on the same or different chromosomes) may be amplified by a single primer pair or set. In some embodiments target sequences may be amplified by a single primer pair or set. In some embodiments target sequences in a set may be amplified with two or more primer pairs.

## 20 Primers

Primers useful for amplification, detection, quantification and sequencing of Y-chromosome loci described herein, for example, are provided. In some embodiments the primers may be complementary to, and hybridize or anneal specifically to or near (e.g., adjacent to) sequences that flank a target region therein. In some embodiments primers are used in sets, where a set contains at least a pair. In some embodiments a set of primers may include a third or a fourth nucleic acid (e.g., two pairs of primers or nested sets of primers, for example). A plurality of primer pairs may constitute a primer set in certain embodiments (e.g., about 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95 or 100 pairs). In some embodiments a plurality of primer sets, each set comprising pair(s) of primers, may be used.

The term "primer" as used herein refers to a nucleic acid that comprises a nucleotide sequence capable of hybridizing or annealing to a target nucleic acid, at or near (e.g., adjacent to) a specific region of interest. As used herein, the term "PCR primer(s)" refers to oligonucleotides that can be used in a polymerase chain reaction (PCR) to amplify a nucleotide sequence originated from a Y-



chromosome loci described herein, for example. In certain embodiments, at least one of the PCR primers for amplification of a nucleotide sequence encoding a Y-chromosome loci described herein can be sequence-specific for the locus (e.g., sequence specific for one of the loci described in Table 1). In some embodiments, primers may be modified (e.g., addition of a universal primer sequence) to improve multiplexing.

Primers can allow for specific determination of a target nucleic acid nucleotide sequence or detection of the target nucleic acid sequence (e.g., presence or absence of a sequence or copy number of a sequence), or feature thereof, for example. Primers may also be used to detect amplification products or extension products, in certain embodiments. A primer may be naturally occurring or synthetic. The term "specific", "specifically" or "specificity", as used herein with respect to nucleic acids, refers to the binding or hybridization of one molecule to another molecule, such as a primer for a target polynucleotide sequence. That is, "specific", "specifically" or "specificity" refers to the recognition, contact, and formation of a stable complex between two molecules, as compared to substantially less recognition, contact, or complex formation of either of those two molecules with other molecules. As used herein, the term "anneal" refers to the formation of a stable complex between two molecules. The terms "primer", "oligo", or "oligonucleotide" may be used interchangeably throughout the document, when referring to primers.

A primer nucleic acid can be designed and synthesized using suitable processes, and may be of any length suitable for hybridizing to a nucleotide sequence of interest (e.g., where the nucleic acid is in liquid phase or bound to a solid support) and performing analysis processes described herein. Primers may be designed based upon a target nucleotide sequence. A primer in some embodiments may be about 10 to about 100 nucleotides, about 10 to about 70 nucleotides, about 10 to about 50 nucleotides, about 15 to about 30 nucleotides, or about 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95 or 100 nucleotides in length. A primer may be composed of naturally occurring and/or non-naturally occurring nucleotides (e.g., labeled nucleotides), or a mixture thereof. Primers suitable for use with embodiments described herein may be synthesized and labeled using known techniques. Oligonucleotides (e.g., primers) may be chemically synthesized according to the solid phase phosphoramidite triester method first described by Beaucage and Caruthers, Tetrahedron Letts., 22:1859-1862, 1981, using an automated synthesizer, as described in Needham-VanDevanter et al., Nucleic Acids Res. 12:6159-6168, 1984. Purification of oligonucleotides can be effected by native acrylamide gel electrophoresis or by anion-exchange high-performance liquid

chromatography (HPLC), for example, as described in Pearson and Regnier, J. Chrom., 255:137-149, 1983.

5 All or a portion of a primer nucleic acid sequence (naturally occurring or synthetic) may be substantially complementary to a target nucleic acid sequence, in some embodiments. As referred to herein, "substantially complementary" with respect to sequences refers to nucleotide sequences that will hybridize with each other. The stringency of the hybridization conditions can be altered to tolerate varying amounts of sequence mismatch. Included are regions of counterpart, target and capture nucleotide sequences 55% or more, 56% or more, 57% or more, 10 58% or more, 59% or more, 60% or more, 61% or more, 62% or more, 63% or more, 64% or more, 65% or more, 66% or more, 67% or more, 68% or more, 69% or more, 70% or more, 71% or more, 72% or more, 73% or more, 74% or more, 75% or more, 76% or more, 77% or more, 78% or more, 79% or more, 80% or more, 81% or more, 82% or more, 83% or more, 84% or more, 85% or more, 86% or more, 87% or more, 88% or more, 89% or more, 90% or more, 91% 15 or more, 92% or more, 93% or more, 94% or more, 95% or more, 96% or more, 97% or more, 98% or more or 99% or more complementary to each other.

Primers that are substantially complimentary to a target nucleic acid sequence are also substantially identical to the compliment of the target nucleic acid sequence. That is, primers can 20 be substantially identical to the anti-sense strand of the nucleic acid. As referred to herein, "substantially identical" with respect to sequences refers to nucleotide sequences that are 55% or more, 56% or more, 57% or more, 58% or more, 59% or more, 60% or more, 61% or more, 62% or more, 63% or more, 64% or more, 65% or more, 66% or more, 67% or more, 68% or more, 69% or more, 70% or more, 71% or more, 72% or more, 73% or more, 74% or more, 75% or 25 more, 76% or more, 77% or more, 78% or more, 79% or more, 80% or more, 81% or more, 82% or more, 83% or more, 84% or more, 85% or more, 86% or more, 87% or more, 88% or more, 89% or more, 90% or more, 91% or more, 92% or more, 93% or more, 94% or more, 95% or more, 96% or more, 97% or more, 98% or more or 99% or more identical to each other. One test for determining whether two nucleotide sequences are substantially identical is to determine the 30 percent of identical nucleotide sequences shared.

Primer sequences and length may affect hybridization to target nucleic acid sequences. Depending on the degree of mismatch between the primer and target nucleic acid, low, medium or high stringency conditions may be used to effect primer/target annealing. As used herein, the 35 term "stringent conditions" refers to conditions for hybridization and washing. Methods for

hybridization reaction temperature condition optimization are known to those of skill in the art, and may be found in Current Protocols in Molecular Biology, John Wiley & Sons, N.Y., 6.3.1-6.3.6 (1989). Aqueous and non-aqueous methods are described in that reference and either can be used. Non-limiting examples of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 50°C. Another example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 55°C. A further example of stringent hybridization conditions is hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 60°C. Often, stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 65°C. More often, stringency conditions are 0.5M sodium phosphate, 7% SDS at 65°C, followed by one or more washes at 0.2X SSC, 1% SDS at 65°C. Stringent hybridization temperatures can also be altered (i.e. lowered) with the addition of certain organic solvents, formamide for example. Organic solvents, like formamide, reduce the thermal stability of double-stranded polynucleotides, so that hybridization can be performed at lower temperatures, while still maintaining stringent conditions and extending the useful life of nucleic acids that may be heat labile.

As used herein, the phrase "hybridizing" or grammatical variations thereof, refers to binding of a first nucleic acid molecule to a second nucleic acid molecule under low, medium or high stringency conditions, or under nucleic acid synthesis conditions. Hybridizing can include instances where a first nucleic acid molecule binds to a second nucleic acid molecule, where the first and second nucleic acid molecules are complementary. As used herein, "specifically hybridizes" refers to preferential hybridization under nucleic acid synthesis conditions of a primer, to a nucleic acid molecule having a sequence complementary to the primer compared to hybridization to a nucleic acid molecule not having a complementary sequence. For example, specific hybridization includes the hybridization of a primer to a target nucleic acid sequence that is complementary to the primer.

In some embodiments primers can include a nucleotide subsequence that may be complementary to a solid phase nucleic acid primer hybridization sequence or substantially complementary to a solid phase nucleic acid primer hybridization sequence (e.g., about 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or greater than 99% identical to the primer hybridization

sequence complement when aligned). A primer may contain a nucleotide subsequence not complementary to or not substantially complementary to a solid phase nucleic acid primer hybridization sequence (e.g., at the 3' or 5' end of the nucleotide subsequence in the primer complementary to or substantially complementary to the solid phase primer hybridization  
5 sequence).

A primer, in certain embodiments, may contain a detectable molecule or entity (e.g., a fluorophore, radioisotope, colorimetric agent, particle, enzyme and the like). When desired, the nucleic acid can be modified to include a detectable label using any method known to one of skill  
10 in the art. The label may be incorporated as part of the synthesis, or added on prior to using the primer in any of the processes described herein. Incorporation of label may be performed either in liquid phase or on solid phase. In some embodiments the detectable label may be useful for detection of targets. In some embodiments the detectable label may be useful for the quantification target nucleic acids (e.g., determining copy number of a particular sequence or  
15 species of nucleic acid). Any detectable label suitable for detection of an interaction or biological activity in a system can be appropriately selected and utilized by the artisan. Examples of detectable labels are fluorescent labels such as fluorescein, rhodamine, and others (e.g., Anantha, et al., *Biochemistry* (1998) 37:2709 2714; and Qu & Chaires, *Methods Enzymol.* (2000) 321:353 369); radioactive isotopes (e.g., <sup>125</sup>I, <sup>131</sup>I, <sup>35</sup>S, <sup>31</sup>P, <sup>32</sup>P, <sup>33</sup>P, <sup>14</sup>C, <sup>3</sup>H, <sup>7</sup>Be, <sup>28</sup>Mg, <sup>57</sup>Co, <sup>65</sup>Zn, <sup>67</sup>Cu, <sup>68</sup>Ge, <sup>82</sup>Sr, <sup>83</sup>Rb, <sup>95</sup>Tc, <sup>96</sup>Tc, <sup>103</sup>Pd, <sup>109</sup>Cd, and <sup>127</sup>Xe); light scattering  
20 labels (e.g., U.S. Patent No. 6,214,560, and commercially available from Genicon Sciences Corporation, CA); chemiluminescent labels and enzyme substrates (e.g., dioxetanes and acridinium esters), enzymic or protein labels (e.g., green fluorescence protein (GFP) or color variant thereof, luciferase, peroxidase); other chromogenic labels or dyes (e.g., cyanine), and  
25 other cofactors or biomolecules such as digoxigenin, streptavidin, biotin (e.g., members of a binding pair such as biotin and avidin for example), affinity capture moieties and the like. In some embodiments a primer may be labeled with an affinity capture moiety. Also included in detectable labels are those labels useful for mass modification for detection with mass spectrometry (e.g., matrix-assisted laser desorption ionization (MALDI) mass spectrometry and electrospray (ES)  
30 mass spectrometry).

A primer also may refer to a polynucleotide sequence that hybridizes to a subsequence of a target nucleic acid or another primer and facilitates the detection of a primer, a target nucleic acid or both, and amplification products or extension products, as with molecular beacons, for  
35 example. The term "molecular beacon" as used herein refers to detectable molecule, where the

detectable property of the molecule is detectable only under certain specific conditions, thereby enabling it to function as a specific and informative signal. Non-limiting examples of detectable properties are, optical properties, electrical properties, magnetic properties, chemical properties and time or speed through an opening of known size.

5

In some embodiments a molecular beacon can be a single-stranded oligonucleotide capable of forming a stem-loop structure, where the loop sequence may be complementary to a target nucleic acid sequence of interest and is flanked by short complementary arms that can form a stem. The oligonucleotide may be labeled at one end with a fluorophore and at the other end  
10 with a quencher molecule. In the stem-loop conformation, energy from the excited fluorophore is transferred to the quencher, through long-range dipole-dipole coupling similar to that seen in fluorescence resonance energy transfer, or FRET, and released as heat instead of light. When the loop sequence is hybridized to a specific target sequence, the two ends of the molecule are separated and the energy from the excited fluorophore is emitted as light, generating a detectable  
15 signal. Molecular beacons offer the added advantage that removal of excess probe is unnecessary due to the self-quenching nature of the unhybridized probe. In some embodiments molecular beacon probes can be designed to either discriminate or tolerate mismatches between the loop and target sequences by modulating the relative strengths of the loop-target hybridization and stem formation. As referred to herein, the term "mismatched nucleotide" or a  
20 "mismatch" refers to a nucleotide that is not complementary to the target sequence at that position or positions. A probe may have at least one mismatch, but can also have 2, 3, 4, 5, 6 or 7 or more mismatched nucleotides.

### Detection

25

Polymorphisms, polynucleotide sequences generated, amplified nucleic acid species (e.g. amplicons or amplification products) or detectable products (e.g., extension products), prepared from the foregoing, can be detected by a suitable detection process. Non limiting examples of methods of detection, quantification, sequencing and the like are; mass detection of mass  
30 modified amplicons (e.g., matrix-assisted laser desorption ionization (MALDI) mass spectrometry and electrospray (ES) mass spectrometry), a primer extension method (e.g., iPLEX™; Sequenom, Inc.), microsequencing methods (e.g., a modification of primer extension methodology), ligase sequence determination methods (e.g., U.S. Pat. Nos. 5,679,524 and 5,952,174, and WO 01/27326), mismatch sequence determination methods (e.g., U.S. Pat. Nos.  
35 5,851,770; 5,958,692; 6,110,684; and 6,183,958), direct DNA sequencing, restriction fragment

length polymorphism (RFLP analysis), allele specific oligonucleotide (ASO) analysis, methylation-specific PCR (MSPCR), pyrosequencing analysis, acycloprime analysis, Reverse dot blot, GeneChip microarrays, Dynamic allele-specific hybridization (DASH), Peptide nucleic acid (PNA) and locked nucleic acids (LNA) probes, TaqMan, Molecular Beacons, Intercalating dye, FRET primers, AlphaScreen, SNPstream, genetic bit analysis (GBA), Multiplex minisequencing, SNaPshot, GOOD assay, Microarray miniseq, arrayed primer extension (APEX), Microarray primer extension (e.g., microarray sequence determination methods), Tag arrays, Coded microspheres, Template-directed incorporation (TDI), fluorescence polarization, Colorimetric oligonucleotide ligation assay (OLA), Sequence-coded OLA, Microarray ligation, Ligase chain reaction, Padlock probes, Invader assay, hybridization methods (e.g., hybridization using at least one probe, hybridization using at least one fluorescently labeled probe, and the like), conventional dot blot analyses, single strand conformational polymorphism analysis (SSCP, e.g., U.S. Patent Nos. 5,891,625 and 6,013,499; Orita et al., Proc. Natl. Acad. Sci. U.S.A 86: 27776-2770 (1989)), denaturing gradient gel electrophoresis (DGGE), heteroduplex analysis, mismatch cleavage detection, and techniques described in Sheffield et al., Proc. Natl. Acad. Sci. USA 49: 699-706 (1991), White et al., Genomics 12: 301-306 (1992), Grompe et al., Proc. Natl. Acad. Sci. USA 86: 5855-5892 (1989), and Grompe, Nature Genetics 5: 111-117 (1993), cloning and sequencing, electrophoresis, the use of hybridization probes and quantitative real time polymerase chain reaction (QRT-PCR), digital PCR, nanopore sequencing, chips and combinations thereof. The detection and quantification of alleles or paralogs can be carried out using the "closed-tube" methods described in U.S. Patent Application 11/950,395, which was filed December 4, 2007. In some embodiments the amount of each amplified nucleic acid species is determined by mass spectrometry, primer extension, sequencing (e.g., any suitable method, for example nanopore or pyrosequencing), Quantitative PCR (Q-PCR or QRT-PCR), digital PCR, combinations thereof, and the like.

In addition to the methods of detection listed above, the following detection methods may also be used to detect amplified nucleic acid species (e.g., target sequences). In some embodiments, the amplified nucleic acid species, can be sequenced directly using any suitable nucleic acid sequencing method. Non-limiting examples of nucleic acid sequencing methods useful for process described herein are; pyrosequencing, nanopore based sequencing methods (e.g., sequencing by synthesis), sequencing by ligation, sequencing by hybridization, microsequencing (primer extension based polymorphism detection), and conventional nucleotide sequencing (e.g., dideoxy sequencing using conventional methods).

35

In some embodiments, the amplified sequence(s) may be cloned prior to sequence analysis. That is, the amplified nucleic acid species may be ligated into a nucleic acid cloning vector by any process known to one of skill in the art. Cloning of the amplified nucleic acid species may be performed by including unique restriction sites in primer subsequences, which can be used to generate a fragment flanked by restriction sites useful for cloning into an appropriately prepared vector, in some embodiments. In certain embodiments blunt-ended cloning can be used to clone amplified nucleic acid species into an appropriately prepared cloning vector. Cloning of the amplified nucleic acid species may be useful for further manipulation, modification, storage, and analysis of the target sequence of interest. In some embodiments, primers may be designed to overlap an SNP site to allow analysis by allele-specific PCR. Allele-specific PCR may be used to discriminate between Y-chromosome loci described herein, because only the correctly hybridized primers will be amplified. In some embodiments, the amplified nucleic acid species may be further analyzed by hybridization (e.g., liquid or solid phase hybridization using sequence specific probes, for example).

Amplified nucleic acids (including amplified nucleic acids that result from reverse transcription) may be modified nucleic acids. Reverse transcribed nucleic acids also may be modified nucleic acids. Modified nucleic acids can include nucleotide analogs, and in certain embodiments include a detectable label and/or a capture agent (e.g., biomolecules or members of a binding pair, as listed below). Modified nucleic acids can be detected by detecting a detectable label or "signal-generating moiety" in some embodiments. The term "signal-generating" as used herein refers to any atom or molecule that can provide a detectable or quantifiable effect, and that can be attached to a nucleic acid. In certain embodiments, a detectable label generates a unique light signal, a fluorescent signal, a luminescent signal, an electrical property, a chemical property, a magnetic property and the like.

Detectable labels include, but are not limited to, nucleotides (labeled or unlabelled), compomers, sugars, peptides, proteins, antibodies, chemical compounds, conducting polymers, binding moieties such as biotin, mass tags, colorimetric agents, light emitting agents, chemiluminescent agents, light scattering agents, fluorescent tags, radioactive tags, charge tags (electrical or magnetic charge), volatile tags and hydrophobic tags, biomolecules (e.g., members of a binding pair antibody/antigen, antibody/antibody, antibody/antibody fragment, antibody/antibody receptor, antibody/protein A or protein G, hapten/anti-hapten, biotin/avidin, biotin/streptavidin, folic acid/folate binding protein, vitamin B12/intrinsic factor, chemical reactive group/complementary chemical reactive group (e.g., sulfhydryl/maleimide, sulfhydryl/haloacetyl derivative,

amine/isotriocyanate, amine/succinimidyl ester, and amine/sulfonyl halides) and the like, some of which are further described below. In some embodiments a probe may contain a signal-generating moiety that hybridizes to a target and alters the passage of the target nucleic acid through a nanopore, and can generate a signal when released from the target nucleic acid when it passes through the nanopore (e.g., alters the speed or time through a pore of known size).

A solution containing amplicons produced by an amplification process, or a solution containing extension products produced by an extension process, can be subjected to further processing. For example, a solution can be contacted with an agent that removes phosphate moieties from free nucleotides that have not been incorporated into an amplicon or extension product. An example of such an agent is a phosphatase (e.g., alkaline phosphatase). Amplicons and extension products also may be associated with a solid phase, may be washed, may be contacted with an agent that removes a terminal phosphate (e.g., exposure to a phosphatase), may be contacted with an agent that removes a terminal nucleotide (e.g., exonuclease), may be contacted with an agent that cleaves (e.g., endonuclease, ribonuclease), and the like.

The term "solid support" or "solid phase" as used herein refers to an insoluble material with which nucleic acid can be associated. Examples of solid supports for use with processes described herein include, without limitation, arrays, beads (e.g., paramagnetic beads, magnetic beads, microbeads, nanobeads) and particles (e.g., microparticles, nanoparticles). Particles or beads having a nominal, average or mean diameter of about 1 nanometer to about 500 micrometers can be utilized, such as those having a nominal, mean or average diameter, for example, of about 10 nanometers to about 100 micrometers; about 100 nanometers to about 100 micrometers; about 1 micrometer to about 100 micrometers; about 10 micrometers to about 50 micrometers; about 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 200, 300, 400, 500, 600, 700, 800 or 900 nanometers; or about 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 200, 300, 400, 500 micrometers.

A solid support can comprise virtually any insoluble or solid material, and often a solid support composition is selected that is insoluble in water. For example, a solid support can comprise or consist essentially of silica gel, glass (e.g. controlled-pore glass (CPG)), nylon, Sephadex®, Sepharose®, cellulose, a metal surface (e.g. steel, gold, silver, aluminum, silicon and copper), a magnetic material, a plastic material (e.g., polyethylene, polypropylene, polyamide, polyester, polyvinylidenedifluoride (PVDF)) and the like. Beads or particles may be swellable (e.g., polymeric beads such as Wang resin) or non-swellable (e.g., CPG). Commercially available



examples of beads include without limitation Wang resin, Merrifield resin and Dynabeads® and SoluLink.

5 A solid support may be provided in a collection of solid supports. A solid support collection comprises two or more different solid support species. The term "solid support species" as used herein refers to a solid support in association with one particular solid phase nucleic acid species or a particular combination of different solid phase nucleic acid species. In certain embodiments, a solid support collection comprises 2 to 10,000 solid support species, 10 to 1,000 solid support species or about 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85,  
10 90, 95, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000 or 10000 unique solid support species. The solid supports (e.g., beads) in the collection of solid supports may be homogeneous (e.g., all are Wang resin beads) or heterogeneous (e.g., some are Wang resin beads and some are magnetic beads). Each solid support species in a collection of solid supports sometimes is labeled with a specific identification  
15 tag. An identification tag for a particular solid support species sometimes is a nucleic acid (e.g., "solid phase nucleic acid") having a unique sequence in certain embodiments. An identification tag can be any molecule that is detectable and distinguishable from identification tags on other solid support species.

20 Mass spectrometry is a particularly effective method for the detection of nucleic acids (e.g., PCR amplicon, primer extension product, detector probe cleaved from a target nucleic acid). Presence of a target nucleic acid is verified by comparing the mass of the detected signal with the expected mass of the target nucleic acid. The relative signal strength, e.g., mass peak on a spectra, for a particular target nucleic acid indicates the relative population of the target nucleic acid amongst  
25 other nucleic acids, thus enabling calculation of a ratio of target to other nucleic acid or sequence copy number directly from the data. For a review of genotyping methods using Sequenom® standard iPLEX™ assay and MassARRAY® technology, see Jurinke, C., Oeth, P., van den Boom, D., "MALDI-TOF mass spectrometry: a versatile tool for high-performance DNA analysis." Mol. Biotechnol. 26, 147-164 (2004); and Oeth, P. et al., "iPLEX™ Assay: Increased Plexing  
30 Efficiency and Flexibility for MassARRAY® System through single base primer extension with mass-modified Terminators." SEQUENOM Application Note (2005). For a review of detecting and quantifying target nucleic using cleavable detector probes that are cleaved during the amplification process and detected by mass spectrometry, see US Patent Application Number 11/950,395, which was filed December 4, 2007, and is hereby incorporated by reference. Such

approaches may be adapted to detection of chromosome abnormalities by methods described herein.

5 In some embodiments, amplified nucleic acid species may be detected by (a) contacting the amplified nucleic acid species (e.g., amplicons) with extension primers (e.g., detection or detector primers), (b) preparing extended extension primers, and (c) determining the relative amount of the one or more mismatch nucleotides (e.g., SNP that exist between SNP-alleles or paralogous sequences) by analyzing the extended detection primers (e.g., extension primers, or detection of extension products). In certain embodiments one or more mismatch nucleotides may be  
10 analyzed by mass spectrometry. In some embodiments amplification, using methods described herein, may generate between about 1 to about 100 amplicon sets, about 2 to about 80 amplicon sets, about 4 to about 60 amplicon sets, about 6 to about 40 amplicon sets, and about 8 to about 20 amplicon sets (e.g., about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, or about 100 amplicon sets).

15 An example using mass spectrometry for detection of amplicon sets (e.g., sets of amplification products) is presented herein. Amplicons may be contacted (in solution or on solid phase) with a set of oligonucleotides (the same primers used for amplification or different primers representative of subsequences in the primer or target nucleic acid) under hybridization  
20 conditions, where: (1) each oligonucleotide in the set comprises a hybridization sequence capable of specifically hybridizing to one amplicon under the hybridization conditions when the amplicon is present in the solution, (2) each oligonucleotide in the set comprises a distinguishable tag located 5' of the hybridization sequence, (3) a feature of the distinguishable tag of one oligonucleotide detectably differs from the features of distinguishable tags of other  
25 oligonucleotides in the set; and (4) each distinguishable tag specifically corresponds to a specific amplicon and thereby specifically corresponds to a specific target nucleic acid. The hybridized amplicon and "detection" primer are subjected to nucleotide synthesis conditions that allow extension of the detection primer by one or more nucleotides (labeled with a detectable entity or moiety, or unlabeled), where one of the one or more nucleotides can be a terminating nucleotide.  
30 In some embodiments one or more of the nucleotides added to the primer may comprises a capture agent. In embodiments where hybridization occurred in solution, capture of the primer/amplicon to solid support may be desirable. The detectable moieties or entities can be released from the extended detection primer, and detection of the moiety determines the presence, absence or copy number of the nucleotide sequence of interest. In certain  
35 embodiments, the extension may be performed once yielding one extended oligonucleotide. In

some embodiments, the extension may be performed multiple times (e.g., under amplification conditions) yielding multiple copies of the extended oligonucleotide. In some embodiments performing the extension multiple times can produce a sufficient number of copies such that interpretation of signals, representing copy number of a particular sequence, can be made with a confidence level of 95% or more (e.g., confidence level of 95% or more, 96% or more, 97% or more, 98% or more, 99% or more, or a confidence level of 99.5% or more). In some  
5       embodiments, the method for detecting amplicon sets can be used to detect extension products.

Methods provided herein allow for high-throughput detection of nucleic acid species in a plurality  
10       of nucleic acids (e.g., nucleotide sequence species, amplified nucleic acid species and detectable products generated from the foregoing). Multiplexing refers to the simultaneous detection of more than one nucleic acid species. General methods for performing multiplexed reactions in conjunction with mass spectrometry, are known (see, e.g., U.S. Pat. Nos. 6,043,031, 5,547,835 and International PCT application No. WO 97/37041). Multiplexing provides an advantage that a  
15       plurality of nucleic acid species (e.g., some having different sequence variations) can be identified in as few as a single mass spectrum, as compared to having to perform a separate mass spectrometry analysis for each individual target nucleic acid species. Methods provided herein lend themselves to high-throughput, highly-automated processes for analyzing sequence variations with high speed and accuracy, in some embodiments. In some embodiments, methods  
20       herein may be multiplexed at high levels in a single reaction.

Microarrays may be adapted for use in embodiments described herein. A microarray can be utilized for determining whether a polymorphic variant is present or absent in a nucleic acid sample. A microarray may include any oligonucleotides or primers described herein, and  
25       methods for making and using oligonucleotide microarrays suitable for prognostic use are disclosed in U.S. Pat. Nos. 5,492,806; 5,525,464; 5,589,330; 5,695,940; 5,849,483; 6,018,041; 6,045,996; 6,136,541; 6,142,681; 6,156,501; 6,197,506; 6,223,127; 6,225,625; 6,229,911; 6,239,273; WO 00/52625; WO 01/25485; and WO 01/29259. The microarray typically comprises a solid support and the oligonucleotides may be linked to this solid support by covalent bonds or  
30       by non-covalent interactions. The oligonucleotides may also be linked to the solid support directly or by a spacer molecule. A microarray may comprise one or more oligonucleotides complementary to a polymorphic site within a nucleotide sequence in Tables 6, 7 or 8. Microarrays may be used with multiplexed protocols described herein.

In certain embodiments, the number of nucleic acid species multiplexed include, without limitation, about 1 to about 500 (e.g., about 1-3, 3-5, 5-7, 7-9, 9-11, 11-13, 13-15, 15-17, 17-19, 19-21, 21-23, 23-25, 25-27, 27-29, 29-31, 31-33, 33-35, 35-37, 37-39, 39-41, 41-43, 43-45, 45-47, 47-49, 49-51, 51-53, 53-55, 55-57, 57-59, 59-61, 61-63, 63-65, 65-67, 67-69, 69-71, 71-73, 73-75, 75-77, 77-79, 79-81, 81-83, 83-85, 85-87, 87-89, 89-91, 91-93, 93-95, 95-97, 97-101, 101-103, 103-105, 105-107, 107-109, 109-111, 111-113, 113-115, 115-117, 117-119, 121-123, 123-125, 125-127, 127-129, 129-131, 131-133, 133-135, 135-137, 137-139, 139-141, 141-143, 143-145, 145-147, 147-149, 149-151, 151-153, 153-155, 155-157, 157-159, 159-161, 161-163, 163-165, 165-167, 167-169, 169-171, 171-173, 173-175, 175-177, 177-179, 179-181, 181-183, 183-185, 185-187, 187-189, 189-191, 191-193, 193-195, 195-197, 197-199, 199-201, 201-203, 203-205, 205-207, 207-209, 209-211, 211-213, 213-215, 215-217, 217-219, 219-221, 221-223, 223-225, 225-227, 227-229, 229-231, 231-233, 233-235, 235-237, 237-239, 239-241, 241-243, 243-245, 245-247, 247-249, 249-251, 251-253, 253-255, 255-257, 257-259, 259-261, 261-263, 263-265, 265-267, 267-269, 269-271, 271-273, 273-275, 275-277, 277-279, 279-281, 281-283, 283-285, 285-287, 287-289, 289-291, 291-293, 293-295, 295-297, 297-299, 299-301, 301-303, 303-305, 305-307, 307-309, 309-311, 311-313, 313-315, 315-317, 317-319, 319-321, 321-323, 323-325, 325-327, 327-329, 329-331, 331-333, 333-335, 335-337, 337-339, 339-341, 341-343, 343-345, 345-347, 347-349, 349-351, 351-353, 353-355, 355-357, 357-359, 359-361, 361-363, 363-365, 365-367, 367-369, 369-371, 371-373, 373-375, 375-377, 377-379, 379-381, 381-383, 383-385, 385-387, 387-389, 389-391, 391-393, 393-395, 395-397, 397-401, 401-403, 403-405, 405-407, 407-409, 409-411, 411-413, 413-415, 415-417, 417-419, 419-421, 421-423, 423-425, 425-427, 427-429, 429-431, 431-433, 433-435, 435-437, 437-439, 439-441, 441-443, 443-445, 445-447, 447-449, 449-451, 451-453, 453-455, 455-457, 457-459, 459-461, 461-463, 463-465, 465-467, 467-469, 469-471, 471-473, 473-475, 475-477, 477-479, 479-481, 481-483, 483-485, 485-487, 487-489, 489-491, 491-493, 493-495, 495-497, 497-501).

Design methods for achieving resolved mass spectra with multiplexed assays can include primer and oligonucleotide design methods and reaction design methods. For primer and oligonucleotide design in multiplexed assays, the same general guidelines for primer design applies for uniplexed reactions, such as avoiding false priming and primer dimers, only more primers are involved for multiplex reactions. For mass spectrometry applications, analyte peaks in the mass spectra for one assay are sufficiently resolved from a product of any assay with which that assay is multiplexed, including pausing peaks and any other by-product peaks. Also, analyte peaks optimally fall within a user-specified mass window, for example, within a range of 5,000-8,500 Da. In some embodiments multiplex analysis may be adapted to mass spectrometric

detection of chromosome abnormalities, for example. In certain embodiments multiplex analysis may be adapted to various single nucleotide or nanopore based sequencing methods described herein. Commercially produced micro-reaction chambers or devices or arrays or chips may be used to facilitate multiplex analysis, and are commercially available.

5

Nucleotide sequence species, amplified nucleic acid species, or detectable products generated from the foregoing may be subject to sequence analysis. The term "sequence analysis" as used herein refers to determining a nucleotide sequence of an amplification product. The entire sequence or a partial sequence of an amplification product can be determined, and the  
10 determined nucleotide sequence is referred to herein as a "read." For example, linear amplification products may be analyzed directly without further amplification in some embodiments (e.g., by using single-molecule sequencing methodology (described in greater detail hereafter)). In certain embodiments, linear amplification products may be subject to further amplification and then analyzed (e.g., using sequencing by ligation or pyrosequencing  
15 methodology (described in greater detail hereafter)). Reads may be subject to different types of sequence analysis. Any suitable sequencing method can be utilized to detect, and determine the amount of, nucleotide sequence species, amplified nucleic acid species, or detectable products generated from the foregoing. Examples of certain sequencing methods are described hereafter.

20 The terms "sequence analysis apparatus" and "sequence analysis component(s)" used herein refer to apparatus, and one or more components used in conjunction with such apparatus, that can be used by a person of ordinary skill to determine a nucleotide sequence from amplification products resulting from processes described herein (e.g., linear and/or exponential amplification products). Examples of sequencing platforms include, without limitation, the 454 platform  
25 (Roche) (Margulies, M. et al. 2005 Nature 437, 376-380), Illumina Genomic Analyzer (or Solexa platform) or SOLID System (Applied Biosystems) or the Helicos True Single Molecule DNA sequencing technology (Harris TD et al. 2008 Science, 320, 106-109), the single molecule, real-time (SMRTTM) technology of Pacific Biosciences, and nanopore sequencing (Soni GV and Meller A. 2007 Clin Chem 53: 1996-2001). Such platforms allow sequencing of many nucleic  
30 acid molecules isolated from a specimen at high orders of multiplexing in a parallel manner (Dear Brief Funct Genomic Proteomic 2003; 1: 397-416). Each of these platforms allows sequencing of clonally expanded or non-amplified single molecules of nucleic acid fragments. Certain platforms involve, for example, (i) sequencing by ligation of dye-modified probes (including cyclic ligation and cleavage), (ii) pyrosequencing, and (iii) single-molecule sequencing. Nucleotide sequence  
35 species, amplification nucleic acid species and detectable products generated there from can be

considered a “study nucleic acid” for purposes of analyzing a nucleotide sequence by such sequence analysis platforms.

5 Sequencing by ligation is a nucleic acid sequencing method that relies on the sensitivity of DNA ligase to base-pairing mismatch. DNA ligase joins together ends of DNA that are correctly base paired. Combining the ability of DNA ligase to join together only correctly base paired DNA ends, with mixed pools of fluorescently labeled oligonucleotides or primers, enables sequence determination by fluorescence detection. Longer sequence reads may be obtained by including primers containing cleavable linkages that can be cleaved after label identification. Cleavage at 10 the linker removes the label and regenerates the 5' phosphate on the end of the ligated primer, preparing the primer for another round of ligation. In some embodiments primers may be labeled with more than one fluorescent label (e.g., 1 fluorescent label, 2, 3, or 4 fluorescent labels).

15 An example of a system that can be used by a person of ordinary skill based on sequencing by ligation generally involves the following steps. Clonal bead populations can be prepared in emulsion microreactors containing target nucleic acid sequences (“template”), amplification reaction components, beads and primers. After amplification, templates are denatured and bead enrichment is performed to separate beads with extended templates from undesired beads (e.g., beads with no extended templates). The template on the selected beads undergoes a 3' 20 modification to allow covalent bonding to the slide, and modified beads can be deposited onto a glass slide. Deposition chambers offer the ability to segment a slide into one, four or eight chambers during the bead loading process. For sequence analysis, primers hybridize to the adapter sequence. A set of four color dye-labeled probes competes for ligation to the sequencing primer. Specificity of probe ligation is achieved by interrogating every 4th and 5th base during the 25 ligation series. Five to seven rounds of ligation, detection and cleavage record the color at every 5th position with the number of rounds determined by the type of library used. Following each round of ligation, a new complimentary primer offset by one base in the 5' direction is laid down for another series of ligations. Primer reset and ligation rounds (5–7 ligation cycles per round) are repeated sequentially five times to generate 25–35 base pairs of sequence for a single tag. 30 With mate-paired sequencing, this process is repeated for a second tag. Such a system can be used to exponentially amplify amplification products generated by a process described herein, e.g., by ligating a heterologous nucleic acid to the first amplification product generated by a process described herein and performing emulsion amplification using the same or a different solid support originally used to generate the first amplification product. Such a system also may 35 be used to analyze amplification products directly generated by a process described herein by

bypassing an exponential amplification process and directly sorting the solid supports described herein on the glass slide.

5 Pyrosequencing is a nucleic acid sequencing method based on sequencing by synthesis, which relies on detection of a pyrophosphate released on nucleotide incorporation. Generally, sequencing by synthesis involves synthesizing, one nucleotide at a time, a DNA strand complimentary to the strand whose sequence is being sought. Target nucleic acids may be immobilized to a solid support, hybridized with a sequencing primer, incubated with DNA polymerase, ATP sulfurylase, luciferase, apyrase, adenosine 5' phosphosulfate and luciferin. 10 Nucleotide solutions are sequentially added and removed. Correct incorporation of a nucleotide releases a pyrophosphate, which interacts with ATP sulfurylase and produces ATP in the presence of adenosine 5' phosphosulfate, fueling the luciferin reaction, which produces a chemiluminescent signal allowing sequence determination. The amount of light generated is proportional to the number of bases added. Accordingly, the sequence downstream of the 15 sequencing primer can be determined.

An example of a system that can be used by a person of ordinary skill based on pyrosequencing generally involves the following steps: ligating an adaptor nucleic acid to a study nucleic acid and hybridizing the study nucleic acid to a bead; amplifying a nucleotide sequence in the study nucleic 20 acid in an emulsion; sorting beads using a picoliter multiwell solid support; and sequencing amplified nucleotide sequences by pyrosequencing methodology (e.g., Nakano et al., "Single-molecule PCR using water-in-oil emulsion;" Journal of Biotechnology 102: 117-124 (2003)). Such a system can be used to exponentially amplify amplification products generated by a process described herein, e.g., by ligating a heterologous nucleic acid to the first amplification product 25 generated by a process described herein.

Certain single-molecule sequencing embodiments are based on the principal of sequencing by synthesis, and utilize single-pair Fluorescence Resonance Energy Transfer (single pair FRET) as a mechanism by which photons are emitted as a result of successful nucleotide incorporation. 30 The emitted photons often are detected using intensified or high sensitivity cooled charge-coupled devices in conjunction with total internal reflection microscopy (TIRM). Photons are only emitted when the introduced reaction solution contains the correct nucleotide for incorporation into the growing nucleic acid chain that is synthesized as a result of the sequencing process. In FRET based single-molecule sequencing, energy is transferred between two fluorescent dyes, 35 sometimes polymethine cyanine dyes Cy3 and Cy5, through long-range dipole interactions. The

donor is excited at its specific excitation wavelength and the excited state energy is transferred, non-radiatively to the acceptor dye, which in turn becomes excited. The acceptor dye eventually returns to the ground state by radiative emission of a photon. The two dyes used in the energy transfer process represent the “single pair”, in single pair FRET. Cy3 often is used as the donor fluorophore and often is incorporated as the first labeled nucleotide. Cy5 often is used as the acceptor fluorophore and is used as the nucleotide label for successive nucleotide additions after incorporation of a first Cy3 labeled nucleotide. The fluorophores generally are within 10 nanometers of each for energy transfer to occur successfully.

10 An example of a system that can be used based on single-molecule sequencing generally involves hybridizing a primer to a target nucleic acid sequence to generate a complex; associating the complex with a solid phase; iteratively extending the primer by a nucleotide tagged with a fluorescent molecule; and capturing an image of fluorescence resonance energy transfer signals after each iteration (e.g., U.S. Patent No. 7,169,314; Braslavsky et al., PNAS 15 100(7): 3960-3964 (2003)). Such a system can be used to directly sequence amplification products (linearly or exponentially amplified products) generated by processes described herein. In some embodiments the amplification products can be hybridized to a primer that contains sequences complementary to immobilized capture sequences present on a solid support, a bead or glass slide for example. Hybridization of the primer-amplification product complexes with the 20 immobilized capture sequences, immobilizes amplification products to solid supports for single pair FRET based sequencing by synthesis. The primer often is fluorescent, so that an initial reference image of the surface of the slide with immobilized nucleic acids can be generated. The initial reference image is useful for determining locations at which true nucleotide incorporation is occurring. Fluorescence signals detected in array locations not initially identified in the “primer only” reference image are discarded as non-specific fluorescence. Following immobilization of the primer-amplification product complexes, the bound nucleic acids often are sequenced in parallel by the iterative steps of, a) polymerase extension in the presence of one fluorescently labeled nucleotide, b) detection of fluorescence using appropriate microscopy, TIRM for example, c) removal of fluorescent nucleotide, and d) return to step a with a different fluorescently labeled 25 nucleotide.

In some embodiments, nucleotide sequencing may be by solid phase single nucleotide sequencing methods and processes. Solid phase single nucleotide sequencing methods involve contacting target nucleic acid and solid support under conditions in which a single molecule of 35 sample nucleic acid hybridizes to a single molecule of a solid support. Such conditions can



include providing the solid support molecules and a single molecule of target nucleic acid in a "microreactor." Such conditions also can include providing a mixture in which the target nucleic acid molecule can hybridize to solid phase nucleic acid on the solid support. Single nucleotide sequencing methods useful in the embodiments described herein are described in United States  
5 Provisional Patent Application Serial Number 61/021,871 filed January 17, 2008.

In certain embodiments, nanopore sequencing detection methods include (a) contacting a target nucleic acid for sequencing ("base nucleic acid," e.g., linked probe molecule) with sequence-specific detectors, under conditions in which the detectors specifically hybridize to substantially  
10 complementary subsequences of the base nucleic acid; (b) detecting signals from the detectors and (c) determining the sequence of the base nucleic acid according to the signals detected. In certain embodiments, the detectors hybridized to the base nucleic acid are disassociated from the base nucleic acid (e.g., sequentially dissociated) when the detectors interfere with a nanopore structure as the base nucleic acid passes through a pore, and the detectors disassociated from  
15 the base sequence are detected. In some embodiments, a detector disassociated from a base nucleic acid emits a detectable signal, and the detector hybridized to the base nucleic acid emits a different detectable signal or no detectable signal. In certain embodiments, nucleotides in a nucleic acid (e.g., linked probe molecule) are substituted with specific nucleotide sequences corresponding to specific nucleotides ("nucleotide representatives"), thereby giving rise to an  
20 expanded nucleic acid (e.g., U.S. Patent No. 6,723,513), and the detectors hybridize to the nucleotide representatives in the expanded nucleic acid, which serves as a base nucleic acid. In such embodiments, nucleotide representatives may be arranged in a binary or higher order arrangement (e.g., Soni and Meller, *Clinical Chemistry* 53(11): 1996-2001 (2007)). In some embodiments, a nucleic acid is not expanded, does not give rise to an expanded nucleic acid,  
25 and directly serves a base nucleic acid (e.g., a linked probe molecule serves as a non-expanded base nucleic acid), and detectors are directly contacted with the base nucleic acid. For example, a first detector may hybridize to a first subsequence and a second detector may hybridize to a second subsequence, where the first detector and second detector each have detectable labels that can be distinguished from one another, and where the signals from the first detector and  
30 second detector can be distinguished from one another when the detectors are disassociated from the base nucleic acid. In certain embodiments, detectors include a region that hybridizes to the base nucleic acid (e.g., two regions), which can be about 3 to about 100 nucleotides in length (e.g., about 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 50, 55, 60, 65, 70, 75, 80, 85, 90, or 95 nucleotides in length). A detector also may include one or more  
35 regions of nucleotides that do not hybridize to the base nucleic acid. In some embodiments, a

detector is a molecular beacon. A detector often comprises one or more detectable labels independently selected from those described herein. Each detectable label can be detected by any convenient detection process capable of detecting a signal generated by each label (e.g., magnetic, electric, chemical, optical and the like). For example, a CD camera can be used to detect signals from one or more distinguishable quantum dots linked to a detector.

In certain sequence analysis embodiments, reads may be used to construct a larger nucleotide sequence, which can be facilitated by identifying overlapping sequences in different reads and by using identification sequences in the reads. Such sequence analysis methods and software for constructing larger sequences from reads are known to the person of ordinary skill (e.g., Venter et al., *Science* 291: 1304-1351 (2001)). Specific reads, partial nucleotide sequence constructs, and full nucleotide sequence constructs may be compared between nucleotide sequences within a sample nucleic acid (i.e., internal comparison) or may be compared with a reference sequence (i.e., reference comparison) in certain sequence analysis embodiments. Internal comparisons sometimes are performed in situations where a sample nucleic acid is prepared from multiple samples or from a single sample source that contains sequence variations. Reference comparisons sometimes are performed when a reference nucleotide sequence is known and an objective is to determine whether a sample nucleic acid contains a nucleotide sequence that is substantially similar or the same, or different, than a reference nucleotide sequence. Sequence analysis can be facilitated by the use of sequence analysis apparatus and components described above.

Y-chromosome loci described herein can also be detected using standard electrophoretic techniques. Although the detection step can sometimes be preceded by an amplification step, amplification is not required in the embodiments described herein. Examples of methods for detection and quantification of the Y-chromosome loci described herein using electrophoretic techniques can be found in the art. A non-limiting example is presented herein. After running a sample (e.g., mixed nucleic acid sample isolated from maternal serum, or amplification nucleic acid species, for example) in an agarose or polyacrylamide gel, the gel may be labeled (e.g., stained) with ethidium bromide (see, Sambrook and Russell, *Molecular Cloning: A Laboratory Manual* 3d ed., 2001). The presence of a band of the same size as the standard control is an indication of the presence of a target nucleic acid sequence, the amount of which may then be compared to the control based on the intensity of the band, thus detecting and quantifying the target sequence of interest. In some embodiments, restriction enzymes capable of distinguishing between maternal and paternal alleles may be used to detect and quantify target nucleic acid

species. In certain embodiments, oligonucleotide probes specific to Y-chromosome loci described herein (see Table 1) can be used to detect the presence of the target sequence of interest. The oligonucleotides can also be used to indicate the amount of the target nucleic acid molecules in comparison to the standard control, based on the intensity of signal imparted by the probe.

Sequence-specific probe hybridization can be used to detect a particular nucleic acid in a mixture or mixed population comprising other species of nucleic acids. Under sufficiently stringent hybridization conditions, the probes hybridize specifically only to substantially complementary sequences. The stringency of the hybridization conditions can be relaxed to tolerate varying amounts of sequence mismatch. A number of hybridization formats are known in the art, which include but are not limited to, solution phase, solid phase, or mixed phase hybridization assays. The following articles provide an overview of the various hybridization assay formats: Singer et al., *Biotechniques* 4:230, 1986; Haase et al., *Methods in Virology*, pp. 189-226, 1984; Wilkinson, *In situ Hybridization*, Wilkinson ed., IRL Press, Oxford University Press, Oxford; and Hames and Higgins eds., *Nucleic Acid Hybridization: A Practical Approach*, IRL Press, 1987.

Hybridization complexes can be detected by techniques known in the art. Nucleic acid probes capable of specifically hybridizing to a target nucleic acid (e.g., the mRNA or the amplified DNA) can be labeled by any suitable method, and the labeled probe used to detect the presence of hybridized nucleic acids. One commonly used method of detection is autoradiography, using probes labeled with  $^3\text{H}$ ,  $^{125}\text{I}$ ,  $^{35}\text{S}$ ,  $^{14}\text{C}$ ,  $^{32}\text{P}$ ,  $^{33}\text{P}$ , or the like. The choice of radioactive isotope depends on research preferences due to ease of synthesis, stability, and half-lives of the selected isotopes. Other labels include compounds (e.g., biotin and digoxigenin), which bind to antigens or antibodies labeled with fluorophores, chemiluminescent agents, and enzymes. In some embodiments, probes can be conjugated directly with labels such as fluorophores, chemiluminescent agents or enzymes. The choice of label depends on sensitivity required, ease of conjugation with the probe, stability requirements, and available instrumentation.

Primer extension polymorphism detection methods, also referred to herein as "microsequencing" methods, typically are carried out by hybridizing a complementary oligonucleotide to a nucleic acid carrying the polymorphic site. In these methods, the oligonucleotide typically hybridizes adjacent to the polymorphic site. The term "adjacent" as used in reference to "microsequencing" methods, refers to the 3' end of the extension oligonucleotide being sometimes 1 nucleotide from the 5' end of the polymorphic site, often 2 or 3, and at times 4, 5, 6, 7, 8, 9, or 10 nucleotides from

the 5' end of the polymorphic site, in the nucleic acid when the extension oligonucleotide is hybridized to the nucleic acid. The extension oligonucleotide then is extended by one or more nucleotides, often 1, 2, or 3 nucleotides, and the number and/or type of nucleotides that are added to the extension oligonucleotide determine which polymorphic variant or variants are present. Oligonucleotide extension methods are disclosed, for example, in U.S. Patent Nos. 4,656,127; 4,851,331; 5,679,524; 5,834,189; 5,876,934; 5,908,755; 5,912,118; 5,976,802; 5,981,186; 6,004,744; 6,013,431; 6,017,702; 6,046,005; 6,087,095; 6,210,891; and WO 01/20039. The extension products can be detected in any manner, such as by fluorescence methods (see, e.g., Chen & Kwok, *Nucleic Acids Research* 25: 347-353 (1997) and Chen et al., *Proc. Natl. Acad. Sci. USA* 94/20: 10756-10761 (1997)) or by mass spectrometric methods (e.g., MALDI-TOF mass spectrometry) and other methods described herein. Oligonucleotide extension methods using mass spectrometry are described, for example, in U.S. Patent Nos. 5,547,835; 5,605,798; 5,691,141; 5,849,542; 5,869,242; 5,928,906; 6,043,031; 6,194,144; and 6,258,538.

Microsequencing detection methods often incorporate an amplification process that proceeds the extension step. The amplification process typically amplifies a region from a nucleic acid sample that comprises the polymorphic site. Amplification can be carried out utilizing methods described above, or for example using a pair of oligonucleotide primers in a polymerase chain reaction (PCR), in which one oligonucleotide primer typically is complementary to a region 3' of the polymorphism and the other typically is complementary to a region 5' of the polymorphism. A PCR primer pair may be used in methods disclosed in U.S. Patent Nos. 4,683,195; 4,683,202, 4,965,188; 5,656,493; 5,998,143; 6,140,054; WO 01/27327; and WO 01/27329 for example. PCR primer pairs may also be used in any commercially available machines that perform PCR, such as any of the GeneAmp® Systems available from Applied Biosystems.

Whole genome sequencing may also be utilized for discriminating alleles of RNA transcripts, in some embodiments. Examples of whole genome sequencing methods include, but are not limited to, nanopore-based sequencing methods, sequencing by synthesis and sequencing by ligation, as described above.

Rhd Assay

Methods and compositions described herein may be combined with other assays to determine fetal blood type and/or blood compatibility. For example, the present invention may be combined with nucleic acid-based assays that are useful for RhD typing. More specifically, RhD typing may

include compositions and methods for determining the presence or absence of any one of exon 4, exon 5, exon 7 or exon 10 in the RhD gene, or the presence or absence of the RhD pseudogene  $\psi$  (psi).

## 5 Fetal Identifiers

Cell-free fetal DNA constitutes only a minor fraction of the total DNA found in maternal plasma. The amount of fetal DNA in maternal plasma is dependent on the gestational age and is estimated at 3-6% in certain embodiments, and sometimes is about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34 or 35% of extracellular nucleic acid in a sample. Therefore, it sometimes is helpful to ascertain that the diagnostic result is true and not caused by insufficient amount of circulating fetal DNA or loss of the fetal DNA during sample processing.

15 The use of polymorphisms provide a means to confirm the presence of fetal DNA and therefore complete the analysis of negative, and otherwise inconclusive, test result in non-invasive prenatal diagnostics. The use of single nucleotide polymorphisms (SNPs), the most abundant type of polymorphism in the human genome, or insertion/deletion (Ins/Del) polymorphisms may serve as fetal identifiers to determine the presence of fetal DNA in a processed sample (Li, Y., Wenzel, F., Holzgreve, W., Hahn, S., Genotyping fetal paternally inherited SNPs by MALDI-TOF MS using cell-free fetal DNA in maternal plasma: Influence of size fractionation. *Electrophoresis* 27, 3889-3896 (2006); Van der Schoot, C.E., Rijnders, R.J., Bossers, B., de Haas, M., Christiaens, G.C., Dee, R. Real-time PCR of bi-allelic insertion/deletion polymorphisms can serve as a reliable positive control for cell-free fetal DNA in non-invasive prenatal genotyping [abstract] *Blood* 102, 93a (2003); and Chow, K.C., Chiu, R.W., Tsui, N.B., Ding, C., Lau, T.K., Leung, T.N., Lo, Y.M., Mass Spectrometric detection of a SNP panel as an internal positive control for fetal DNA analysis in maternal plasma. *Clin. Chem.* 53, 141-142 (2007), all of which are hereby incorporated by reference).

30 A SNP is considered informative for the determination of the presence of fetal DNA, if the mother is homozygous and the fetus inherited the opposite allele from the father, rendering the genotype of the fetus heterozygous.

To ensure a high probability that the presence of fetal DNA can be confirmed by the presence of the paternally-inherited allele in at least 1 SNP, a sufficient number of SNPs or Ins/Dels with a

high population frequency ( $>0.4$  for the minor frequent allele) has to be analyzed. A scheme exemplifying the concept of using SNPs to confirm the presence of fetal DNA in maternal plasma is depicted in Figure 5.

- 5 Analysis of multiple polymorphisms in DNA extracted from maternal plasma creates a two-fold challenge: firstly, the paternally-inherited allele needs to be detected in the background of the maternal DNA; secondly, the high number of polymorphisms require significant sample material and a significant number of reactions before a conclusive test result is achieved.

## 10 Data Processing

The term "detection" or "detecting" or "identifying" Y-chromosome nucleic acid or the sex of a fetus as used herein refers to identification of the presence or absence of Y-chromosome nucleic acid by processing data arising from detecting one or more sets of amplified nucleic acid species, nucleotide sequence species, or a detectable product generated from the foregoing (collectively "detectable product"). Any suitable detection device and method can be used to distinguish one or more sets of detectable products, as addressed herein. An outcome pertaining to the presence or absence of Y-chromosome nucleic acid can be expressed in any suitable form, including, without limitation, ratio, deviation in ratio, frequency, distribution, probability (e.g., odds ratio, p-value), likelihood, percentage, value over a threshold, or risk factor, associated with the presence of Y-chromosome nucleic acid for a subject or sample. An outcome may be provided with one or more of sensitivity, specificity, standard deviation, coefficient of variation (CV) and/or confidence level, or combinations of the foregoing, in certain embodiments.

25 An outcome pertaining to presence or absence of Y-chromosome nucleic acid may be determined for all samples tested, and in some embodiments, an outcome pertaining to presence or absence of Y-chromosome nucleic acid is determined in a subset of the samples (e.g., samples from individual pregnant females). In certain embodiments, an outcome is determined for about 60, 65, 70, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 30 95, 96, 97, 98, 99%, or greater than 99%, of samples analyzed in a set. A set of samples can include any suitable number of samples, and in some embodiments, a set has about 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900 or 1000 samples, or more than 1000 samples. The set may be considered with respect to samples tested in a particular period of time, and/or at a particular location. The set may be otherwise 35 defined by, for example, gestational age and/or ethnicity. The set may be comprised of a sample

which is subdivided into subsamples or replicates all or some of which may be tested. The set may comprise a sample from the same subject collected at two different times. In certain embodiments, an outcome is determined about 60% or more of the time for a given sample analyzed (e.g., about 65, 70, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 5 93, 94, 95, 96, 97, 98, 99%, or more than 99% of the time for a given sample). In certain embodiments, analyzing a higher number of characteristics (e.g., sequence variations) that discriminate alleles can increase the percentage of outcomes determined for the samples (e.g., discriminated in a multiplex analysis). In some embodiments, one or more tissue or fluid samples (e.g., one or more blood samples) are provided by a subject (e.g., pregnant female). In certain 10 embodiments, one or more nucleic acid samples, or two or more replicate nucleic acid samples, are isolated from a single tissue or fluid sample, and analyzed by methods described herein.

In certain multiplex embodiments for determining the sex of a fetus, the presence of a male fetus is the outcome called where a detectable amount of amplification product for 100% of all Y- 15 chromosome targets assayed in the multiplex assay is detected. In some embodiments, the presence of a male fetus is the outcome called where a detectable amount of amplification product for about 90% to about 99% (e.g., about 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or more than 99%) of all Y-chromosome targets assayed is detected. In some embodiments, the presence of a male fetus is the outcome called where a detectable amount of amplification 20 product for all but one, or all but two, of all Y-chromosome targets assayed is detected. In certain embodiments, the presence of a female fetus is the outcome called where a detectable amount of amplification product for none of the Y-chromosome targets assayed is detected. In some embodiments, the presence of a female fetus is the outcome called where a detectable amount of amplification product for about 1% to about 25% (e.g., less than about 1%, or about 1, 2, 3, 4, 5, 25 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 or 25%) of all Y-chromosome targets assayed is detected. In some embodiments, the presence of a female fetus is the outcome called where a detectable amount of amplification product for only one, or only two, of all Y-chromosome targets assayed is detected.

30 Detection of presence or absence of Y-chromosome nucleic acid based on one or more sets of detectable products may be identified based on one or more calculated variables, including, but not limited to, ratio, distribution, frequency, sensitivity, specificity, standard deviation, coefficient of variation (CV), a threshold, confidence level, score, probability and/or a combination thereof. In some embodiments, (i) the number of sets selected for a diagnostic method, and/or (ii) the

particular nucleotide sequence species of each set selected for a diagnostic method, is determined in part or in full according to one or more of such calculated variables.

5 In certain embodiments, one or more of ratio, sensitivity, specificity and/or confidence level are expressed as a percentage. In some embodiments, the percentage, independently for each variable, is greater than about 90% (e.g., about 90, 91, 92, 93, 94, 95, 96, 97, 98 or 99%, or greater than 99% (e.g., about 99.5%, or greater, about 99.9% or greater, about 99.95% or greater, about 99.99% or greater)). Coefficient of variation (CV) in some embodiments is expressed as a percentage, and sometimes the percentage is about 10% or less (e.g., about 10,  
10 9, 8, 7, 6, 5, 4, 3, 2 or 1%, or less than 1% (e.g., about 0.5% or less, about 0.1% or less, about 0.05% or less, about 0.01% or less)). A probability (e.g., that a particular outcome determined by an algorithm is not due to chance) in certain embodiments is expressed as a p-value, and sometimes the p-value is about 0.05 or less (e.g., about 0.05, 0.04, 0.03, 0.02 or 0.01, or less than 0.01 (e.g., about 0.001 or less, about 0.0001 or less, about 0.00001 or less, about 0.000001  
15 or less)).

For example, scoring or a score may refer to calculating the probability that Y-chromosome nucleic acid is present or absent in a subject/sample. The value of a score may be used to determine for example the variation, difference, or ratio of amplified nucleic detectable product  
20 that may correspond to the actual presence or absence of Y-chromosome nucleic acid. For example, calculating a positive score from detectable products can lead to an identification of presence or absence of Y-chromosome nucleic acid, which is particularly relevant to analysis of single samples.

25 In certain embodiments, simulated (or simulation) data can aid data processing for example by training an algorithm or testing an algorithm. Simulated data may for instance involve hypothetical various samples of different concentrations of fetal and maternal nucleic acid in serum, plasma and the like. Simulated data may be based on what might be expected from a real population or may be skewed to test an algorithm and/or to assign a correct classification  
30 based on a simulated data set. Simulated data also is referred to herein as "virtual" data. Fetal/maternal contributions within a sample can be simulated as a table or array of numbers (for example, as a list of peaks corresponding to the mass signals of cleavage products of a reference biomolecule or amplified nucleic acid sequence), as a mass spectrum, as a pattern of bands on a gel, label intensity, or as a representation of any technique that measures mass distribution.  
35 Simulations can be performed in most instances by a computer program. One possible step in



using a simulated data set is to evaluate the confidence of the identified results, i.e. how well the selected positives/negatives match the sample and whether there are additional variations. A common approach is to calculate the probability value (p-value) which estimates the probability of a random sample having better score than the selected one. As p-value calculations can be prohibitive in certain circumstances, an empirical model may be assessed, in which it is assumed that at least one sample matches a reference sample (with or without resolved variations). Other distributions such as Poisson distribution can be used to describe the probability distribution.

In certain embodiments, an algorithm can assign a confidence value to the true positives, true negatives, false positives and false negatives calculated. The assignment of a likelihood of the occurrence of Y-chromosome nucleic acid can also be based on a certain probability model.

Simulated data often is generated in an in silico process. As used herein, the term "in silico" refers to research and experiments performed using a computer. In silico methods include, but are not limited to, molecular modeling studies, karyotyping, genetic calculations, and virtual representations of molecular structures and/or processes, such as genetic information.

As used herein, a "data processing routine" refers to a process, that can be embodied in software, that determines the biological significance of acquired data (i.e., the ultimate results of an assay). For example, a data processing routine can determine the amount of each nucleotide sequence species based upon the data collected. A data processing routine also may control an instrument and/or a data collection routine based upon results determined. A data processing routine and a data collection routine often are integrated and provide feedback to operate data acquisition by the instrument, and hence provide assay-based judging methods provided herein.

As used herein, software refers to computer readable program instructions that, when executed by a computer, perform computer operations. Typically, software is provided on a program product containing program instructions recorded on a computer readable medium, including, but not limited to, magnetic media including floppy disks, hard disks, and magnetic tape; and optical media including CD-ROM discs, DVD discs, magneto-optical discs, and other such media on which the program instructions can be recorded.

Different methods of predicting presence or absence of Y-chromosome nucleic acid can produce different types of results. For any given prediction, there are four possible types of outcomes: true positive, true negative, false positive, or false negative. The term "true positive" as used

herein refers to a subject correctly diagnosed as having Y-chromosome nucleic acid. The term “false positive” as used herein refers to a subject wrongly identified as having Y-chromosome nucleic acid. The term “true negative” as used herein refers to a subject correctly identified as not having Y-chromosome nucleic acid. The term “false negative” as used herein refers to a subject  
5 wrongly identified as not having Y-chromosome nucleic acid. Two measures of performance for any given method can be calculated based on the ratios of these occurrences: (i) a sensitivity value, the fraction of predicted positives that are correctly identified as being positives (e.g., the fraction of nucleotide sequence sets correctly identified by level comparison  
10 detection/determination as indicative of Y-chromosome nucleic acid, relative to all nucleotide sequence sets identified as such, correctly or incorrectly), thereby reflecting the accuracy of the results in detecting the Y-chromosome nucleic acid; and (ii) a specificity value, the fraction of predicted negatives correctly identified as being negative (the fraction of nucleotide sequence sets correctly identified by level comparison detection/determination as indicative of absence of Y-chromosome nucleic acid, relative to all nucleotide sequence sets identified as such, correctly  
15 or incorrectly), thereby reflecting accuracy of the results in detecting the absence of Y-chromosome nucleic acid.

The term “sensitivity” as used herein refers to the number of true positives divided by the number of true positives plus the number of false negatives, where sensitivity (sens) may be within the  
20 range of  $0 \leq \text{sens} \leq 1$ . Ideally, method embodiments herein have the number of false negatives equaling zero or close to equaling zero, so that no subject is wrongly identified as not having Y-chromosome nucleic acid when they indeed have Y-chromosome nucleic acid. Conversely, an assessment often is made of the ability of a prediction algorithm to classify negatives correctly, a complementary measurement to sensitivity. The term “specificity” as used herein refers to the  
25 number of true negatives divided by the number of true negatives plus the number of false positives, where sensitivity (spec) may be within the range of  $0 \leq \text{spec} \leq 1$ . Ideally, methods embodiments herein have the number of false positives equaling zero or close to equaling zero, so that no subject wrongly identified as having Y-chromosome nucleic acid when they do not have the Y-chromosome nucleic acid. Hence, a method that has sensitivity and specificity  
30 equaling one, or 100%, sometimes is selected.

One or more prediction algorithms may be used to determine significance or give meaning to the detection data collected under variable conditions that may be weighed independently of or  
35 dependently on each other. The term “variable” as used herein refers to a factor, quantity, or function of an algorithm that has a value or set of values. For example, a variable may be the

design of a set of amplified nucleic acid species, the number of sets of amplified nucleic acid species, percent fetal genetic contribution tested, percent maternal genetic contribution tested, type of sex-linked abnormalities assayed, the age of the mother and the like. The term “independent” as used herein refers to not being influenced or not being controlled by another.

5 The term “dependent” as used herein refers to being influenced or controlled by another. For example, presence of Y-chromosome nucleic acid and male sex for a fetus are variables that are dependent upon each other.

10 Any suitable type of method or prediction algorithm may be utilized to give significance to the data of the present invention within an acceptable sensitivity and/or specificity. For example, prediction algorithms such as Mann-Whitney U Test, binomial test, log odds ratio, Chi-squared test, z-test, t-test, ANOVA (analysis of variance), regression analysis, neural nets, fuzzy logic, Hidden Markov Models, multiple model state estimation, and the like may be used. One or more methods or prediction algorithms may be determined to give significance to the data having different  
15 independent and/or dependent variables of the present invention. And one or more methods or prediction algorithms may be determined not to give significance to the data having different independent and/or dependent variables of the present invention. One may design or change parameters of the different variables of methods described herein based on results of one or more prediction algorithms (e.g., number of sets analyzed, types of nucleotide species in each  
20 set).

In certain embodiments, several algorithms may be chosen to be tested. These algorithms then can be trained with raw data. For each new raw data sample, the trained algorithms will assign a classification to that sample (i.e. male or female fetus). Based on the classifications of the new  
25 raw data samples, the trained algorithms’ performance may be assessed based on sensitivity and specificity. Finally, an algorithm with the highest sensitivity and/or specificity or combination thereof may be identified.

As noted above, algorithms, software, processors and/or machines, for example, can be utilized  
30 to (i) process detection data pertaining to nucleotide sequence species and/or amplified nucleic acid species of sets, and/or (ii) identify the presence or absence of Y-chromosome nucleic acid.

In certain embodiments, provided are methods for determining fetal sex in a pregnant woman that comprise: (a) providing a system, where the system comprises distinct software modules, and  
35 where the distinct software modules comprise a signal detection module, a logic processing

module, and a data display organization module; (b) detecting signal information indicating the presence or absence of Y chromosome nucleic acid in a sample from the pregnant female; (c) receiving, by the logic processing module, the signal information; (d) calling the fetal sex by the logic processing module, where the presence of Y chromosome nucleic acid is indicative of a male fetus, and the absence of Y chromosome nucleic acid is indicative of a female fetus; and (e) organizing, by the data display organization model in response to being called by the logic processing module, a data display indicating the sex of the fetus. The Y chromosome nucleic acid may be, for example, a nucleic acid corresponding to one or more Y chromosome loci described herein. The Y chromosome nucleic acid may, for example, be detected, may correspond to, or may be complementary to, a Y chromosome primer described herein.

In certain embodiments, provided are methods for determining the amount of fetal DNA in a sample, where the sample is obtained from a pregnant woman carrying a male fetus, which comprise: (a) providing a system, where the system comprises distinct software modules, and where the distinct software modules comprise a signal detection module, a logic processing module, and a data display organization module; (b) detecting signal information indicating the amount of Y chromosome nucleic acid in the sample; (c) receiving, by the logic processing module, the signal information; (d) calling the amount of fetal DNA in the sample by the logic processing module, where the amount of Y chromosome nucleic acid is indicative of the amount of fetal DNA in the sample; and (e) organizing, by the data display organization model in response to being called by the logic processing module, a data display indicating the amount of fetal DNA in the sample. The Y chromosome nucleic acid may be, for example, a nucleic acid corresponding to one or more Y chromosome loci described herein. The Y chromosome nucleic acid may, for example, be detected, may correspond to, or may be complementary to, a Y chromosome primer described herein.

Provided also are methods for determining fetal sex in a pregnant woman, which comprise providing signal information indicating the presence or absence of Y chromosome nucleic acid in a sample from the pregnant female; providing a system, where the system comprises distinct software modules, and where the distinct software modules comprise a signal detection module, a logic processing module, and a data display organization module; receiving, by the logic processing module, the signal information; calling the fetal sex by the logic processing module where the presence of Y chromosome nucleic acid is indicative of a male fetus, and the absence of Y chromosome nucleic acid is indicative of a female fetus; and organizing, by the data display organization model in response to being called by the logic processing module, a data display

indicating the sex of the fetus. The Y chromosome nucleic acid may be, for example, a nucleic acid corresponding to one or more Y chromosome loci described herein. The Y chromosome nucleic acid may, for example, be detected, may correspond to, or may be complementary to, a Y chromosome primer described herein.

5

Provided also are methods for determining the amount of fetal DNA in a sample, where the sample is obtained from a pregnant woman carrying a male fetus, which comprise providing signal information indicating the amount of Y chromosome nucleic acid in the sample; providing a system, where the system comprises distinct software modules, and where the distinct software  
10 modules comprise a signal detection module, a logic processing module, and a data display organization module; receiving, by the logic processing module, the signal information; calling the amount of fetal DNA in the sample by the logic processing module where the amount of Y chromosome nucleic acid is indicative of the amount of fetal DNA in the sample; and organizing,  
15 by the data display organization model in response to being called by the logic processing module, a data display indicating amount of fetal DNA in the sample. The Y chromosome nucleic acid may be, for example, a nucleic acid corresponding to one or more Y chromosome loci described herein. The Y chromosome nucleic acid may, for example, be detected, may correspond to, or may be complementary to, a Y chromosome primer described herein.

20 Provided also are methods for determining fetal sex in a pregnant woman, which comprise providing a system, where the system comprises distinct software modules, and where the distinct software modules comprise a signal detection module, a logic processing module, and a data display organization module; receiving, by the logic processing module, signal information indicating the presence or absence of Y chromosome nucleic acid in a sample from the pregnant  
25 female; calling the fetal sex by the logic processing module where the presence of Y chromosome nucleic acid is indicative of a male fetus, and the absence of Y chromosome nucleic acid is indicative of a female fetus; and organizing, by the data display organization model in response to being called by the logic processing module, a data display indicating the sex of the fetus. The Y chromosome nucleic acid may be, for example, a nucleic acid corresponding to one  
30 or more Y chromosome loci described herein. The Y chromosome nucleic acid may, for example, be detected, may correspond to, or may be complementary to, a Y chromosome primer described herein.

35 Provided also are methods for determining the amount of fetal DNA in a sample, where the sample is obtained from a pregnant woman carrying a male fetus, which comprise providing a

system, where the system comprises distinct software modules, and where the distinct software modules comprise a signal detection module, a logic processing module, and a data display organization module; receiving, by the logic processing module, signal information indicating the amount of Y chromosome nucleic acid in the sample; calling the amount of fetal DNA in the sample by the logic processing module, where the amount of Y chromosome nucleic acid is indicative of the amount of fetal DNA in the sample; and organizing, by the data display organization model in response to being called by the logic processing module, a data display indicating the amount of fetal DNA in the sample. The Y chromosome nucleic acid may be, for example, a nucleic acid corresponding to one or more Y chromosome loci described herein. The Y chromosome nucleic acid may, for example, be detected, may correspond to, or may be complementary to, a Y chromosome primer described herein.

By "providing signal information" is meant any manner of providing the information, including, for example, computer communication means from a local, or remote site, human data entry, or any other method of transmitting signal information. The signal information may be generated in one location and provided to another location.

By "obtaining" or "receiving" signal information is meant receiving the signal information by computer communication means from a local, or remote site, human data entry, or any other method of receiving signal information. The signal information may be generated in the same location at which it is received, or it may be generated in a different location and transmitted to the receiving location.

By "indicating" or "representing" the amount is meant that the signal information is related to, or correlates with, for example, the presence of Y chromosome nucleic acid. The information may be, for example, the calculated data associated with the presence of the nucleic acid as obtained, for example, after converting raw data obtained by mass spectrometry.

Also provided are computer program products, such as, for example, a computer program products comprising a computer usable medium having a computer readable program code embodied therein, the computer readable program code adapted to be executed to implement a method for determining fetal sex in a pregnant woman, which comprises (a) providing a system, where the system comprises distinct software modules, and where the distinct software modules comprise a signal detection module, a logic processing module, and a data display organization module; (b) detecting signal information indicating the presence or absence of Y chromosome

nucleic acid in a sample from the pregnant female; (c) receiving, by the logic processing module, the signal information; (d) calling the fetal sex by the logic processing module where the presence of Y chromosome nucleic acid is indicative of a male fetus, and the absence of Y chromosome nucleic acid is indicative of a female fetus; and organizing, by the data display organization model  
5 in response to being called by the logic processing module, a data display indicating the sex of the fetus. The Y chromosome nucleic acid may be, for example, a nucleic acid corresponding to one or more Y chromosome loci described herein. The Y chromosome nucleic acid may, for example, be detected, may correspond to, or may be complementary to, a Y chromosome primer described herein.

10  
Also provided are computer program products, such as, for example, a computer program products comprising a computer usable medium having a computer readable program code embodied therein, the computer readable program code adapted to be executed to implement a method for determining the amount of fetal DNA in a sample, where the sample is obtained from  
15 a pregnant woman carrying a male fetus, which comprises (a) providing a system, where the system comprises distinct software modules, and where the distinct software modules comprise a signal detection module, a logic processing module, and a data display organization module; (b) detecting signal information indicating the amount of Y chromosome nucleic acid in the sample; (c) receiving, by the logic processing module, the signal information; (d) calling the amount of  
20 fetal DNA in the sample by the logic processing module, where the amount of Y chromosome nucleic acid is indicative of the amount of fetal DNA in the sample; and (e) organizing, by the data display organization model in response to being called by the logic processing module, a data display indicating the amount of fetal DNA in the sample. The Y chromosome nucleic acid may be, for example, a nucleic acid corresponding to one or more Y chromosome loci described  
25 herein. The Y chromosome nucleic acid may, for example, be detected, may correspond to, or may be complementary to, a Y chromosome primer described herein.

Also provided are computer program products, such as, for example, computer program products comprising a computer usable medium having a computer readable program code embodied  
30 therein, the computer readable program code adapted to be executed to implement a method for determining fetal sex in a pregnant woman, which comprises providing a system, where the system comprises distinct software modules, and where the distinct software modules comprise a signal detection module, a logic processing module, and a data display organization module; receiving signal information indicating the presence or absence of Y chromosome nucleic acid in  
35 a sample from the pregnant female; calling the fetal sex by the logic processing module where

the presence of Y chromosome nucleic acid is indicative of a male fetus, and the absence of Y chromosome nucleic acid is indicative of a female fetus; and organizing, by the data display organization model in response to being called by the logic processing module, a data display indicating the sex of the fetus. The Y chromosome nucleic acid may be, for example, a nucleic acid corresponding to one or more Y chromosome loci described herein. The Y chromosome nucleic acid may, for example, be detected, may correspond to, or may be complementary to, a Y chromosome primer described herein.

Also provided are computer program products, such as, for example, computer program products comprising a computer usable medium having a computer readable program code embodied therein, the computer readable program code adapted to be executed to implement a method for determining the amount of fetal DNA in a sample, where the sample is obtained from a pregnant woman carrying a male fetus, which comprises providing a system, where the system comprises distinct software modules, and where the distinct software modules comprise a signal detection module, a logic processing module, and a data display organization module; receiving signal information indicating the amount of Y chromosome nucleic acid in the sample; calling the amount of fetal DNA in the sample by the logic processing module, where the amount of Y chromosome nucleic acid is indicative of the amount of fetal DNA in the sample; and organizing, by the data display organization model in response to being called by the logic processing module, a data display indicating the amount of fetal DNA in the sample. The Y chromosome nucleic acid may be, for example, a nucleic acid corresponding to one or more Y chromosome loci described herein. The Y chromosome nucleic acid may, for example, be detected, may correspond to, or may be complementary to, a Y chromosome primer described herein.

Signal information may be, for example, mass spectrometry data obtained from mass spectrometry of RNA, or of amplified nucleic acid. As the RNA may be amplified into a nucleic acid that is detected, the signal information may be detection information, such as mass spectrometry data, obtained from stoichiometrically-produced nucleic acid from the RNA. The mass spectrometry data may be raw data, such as, for example, a set of numbers, or, for example, a two dimensional display of the mass spectrum. The signal information may be converted or transformed to any form of data that may be provided to, or received by, a computer system. The signal information may also, for example, be converted, or transformed to identification data or information representing the fetal sex. The signal information may also, for example, be converted or transformed to identification data or information representing the amount of fetal DNA.



Also provided is a machine for determining fetal sex in a pregnant woman, where the machine comprises a computer system having distinct software modules, and where the distinct software modules comprise a signal detection module, a logic processing module, and a data display organization module, where the software modules are adapted to be executed to implement a method for determining fetal sex in a pregnant woman, which comprises (a) detecting signal information indicating the presence or absence of Y chromosome nucleic acid in a sample from the pregnant female; (b) receiving, by the logic processing module, the signal information; (c) calling the fetal sex by the logic processing module, where the presence of Y chromosome nucleic acid is indicative of a male fetus, and the absence of Y chromosome nucleic acid is indicative of a female fetus; and organizing, by the data display organization model in response to being called by the logic processing module, a data display indicating the sex of the fetus. The Y chromosome nucleic acid may be, for example, a nucleic acid corresponding to one or more Y chromosome loci described herein. The Y chromosome nucleic acid may, for example, be detected, may correspond to, or may be complementary to, a Y chromosome primer described herein.

The machine may further comprise a memory module for storing signal information or data indicating fetal sex. Also provided are methods for determining fetal sex in a pregnant woman, where the methods comprise the use of a machine for determining fetal sex in a pregnant woman.

Also provided is a machine for determining the amount of fetal DNA in a sample, where the sample is obtained from a pregnant woman carrying a male fetus, where the machine comprises a computer system having distinct software modules, and where the distinct software modules comprise a signal detection module, a logic processing module, and a data display organization module, where the software modules are adapted to be executed to implement a method for determining the amount of fetal DNA in a sample, which comprises (a) detecting signal information indicating the amount of Y chromosome nucleic acid in the sample; (b) receiving, by the logic processing module, the signal information; (c) calling the amount of fetal DNA in the sample by the logic processing module, where the amount of Y chromosome nucleic acid is indicative of the amount of fetal DNA in the sample; and (e) organizing, by the data display organization model in response to being called by the logic processing module, a data display indicating the amount of fetal DNA in the sample. The Y chromosome nucleic acid may be, for example, a nucleic acid corresponding to one or more Y chromosome loci described herein. The

Y chromosome nucleic acid may, for example, be detected, may correspond to, or may be complementary to, a Y chromosome primer described herein.

5 The machine may further comprise a memory module for storing signal information or data indicating the amount of fetal DNA. Also provided are methods for determining the amount of fetal DNA in a sample, where the sample is obtained from a pregnant woman carrying a male fetus, where the methods comprise the use of a machine for determining the amount of fetal DNA in a sample.

10 Also provided are methods for determining fetal sex in a pregnant woman that comprise: (a) detecting signal information, where the signal information indicates the presence or absence of Y chromosome nucleic acid in a sample from the pregnant female; (b) transforming the signal information representing presence or absence of Y chromosome nucleic acid into identification data, where the identification data represents the fetal sex. The Y chromosome nucleic acid may  
15 be, for example, a nucleic acid corresponding to one or more Y chromosome loci described herein. The Y chromosome nucleic acid may, for example, be detected, may correspond to, or may be complementary to, a Y chromosome primer described herein.

Also provided are methods for determining the amount of fetal DNA in a sample, where the  
20 sample is obtained from a pregnant woman carrying a male fetus, that comprise: (a) detecting signal information, where the signal information indicates the amount of Y chromosome nucleic acid in the sample; (b) transforming the signal information representing the amount of Y chromosome nucleic acid into identification data, where the identification data represents the amount of fetal DNA in the sample. The Y chromosome nucleic acid may be, for example, a  
25 nucleic acid corresponding to one or more Y chromosome loci described herein. The Y chromosome nucleic acid may, for example, be detected, may correspond to, or may be complementary to, a Y chromosome primer described herein.

Also provided are methods for determining fetal sex in a pregnant woman that comprise: (a)  
30 providing signal information indicating the presence or absence of Y chromosome nucleic acid in a sample from the pregnant female; (b) transforming the signal information representing the presence or absence of Y chromosome nucleic acid into identification data, where the identification data represents the fetal sex, whereby the presence of Y chromosome nucleic acid is indicative of a male fetus, and the absence of Y chromosome nucleic acid is indicative of a  
35 female fetus; and (c) displaying the identification data. The Y chromosome nucleic acid may be,

for example, a nucleic acid corresponding to one or more Y chromosome loci described herein. The Y chromosome nucleic acid may, for example, be detected, may correspond to, or may be complementary to, a Y chromosome primer described herein.

5 Also provided are methods for determining the amount of fetal DNA in a sample, where the sample is obtained from a pregnant woman carrying a male fetus, that comprise: (a) providing signal information indicating the amount of Y chromosome nucleic acid in the sample; (b) transforming the signal information representing the amount of Y chromosome nucleic acid into identification data, where the identification data represents the amount of fetal DNA in the  
10 sample; and (c) displaying the identification data. The Y chromosome nucleic acid may be, for example, a nucleic acid corresponding to one or more Y chromosome loci described herein. The Y chromosome nucleic acid may, for example, be detected, may correspond to, or may be complementary to, a Y chromosome primer described herein.

15 Also provided are methods for determining fetal sex in a pregnant woman that comprise: (a) receiving signal information indicating the presence or absence of Y chromosome nucleic acid in a sample from the pregnant female; (b) transforming the signal information representing the presence or absence of Y chromosome nucleic acid into identification data, where the identification data represents the fetal sex, whereby the presence of Y chromosome nucleic acid  
20 is indicative of a male fetus, and the absence of Y chromosome nucleic acid is indicative of a female fetus; and (c) displaying the identification data. The Y chromosome nucleic acid may be, for example, a nucleic acid corresponding to one or more Y chromosome loci described herein. The Y chromosome nucleic acid may, for example, be detected, may correspond to, or may be complementary to, a Y chromosome primer described herein.

25 Also provided are methods for determining the amount of fetal DNA in a sample, where the sample is obtained from a pregnant woman carrying a male fetus that comprise: (a) receiving signal information indicating the amount of Y chromosome nucleic acid in the sample; (b) transforming the signal information representing the amount of Y chromosome nucleic acid into  
30 identification data, where the identification data represents the amount of fetal DNA in the sample; and (c) displaying the identification data. The Y chromosome nucleic acid may be, for example, a nucleic acid corresponding to one or more Y chromosome loci described herein. The Y chromosome nucleic acid may, for example, be detected, may correspond to, or may be complementary to, a Y chromosome primer described herein.

35

For purposes of these, and similar embodiments, the term "signal information" indicates information readable by any electronic media, including, for example, computers that represent data derived using the present methods. For example, "signal information" can represent the amount of Y chromosome nucleic acid; a ratio of Y chromosome nucleic acid to female nucleic acid, an amount of amplified nucleic acid that corresponds to, or is complementary to, a Y chromosome locus, or an amount of amplified nucleic acid that corresponds to, or is complementary to, more than one Y chromosome locus. Signal information, such as in these examples, that represents physical substances may be transformed into identification data, such as a visual display that represents other physical substances, such as, for example, fetal sex. Identification data may be displayed in any appropriate manner, including, but not limited to, in a computer visual display, by encoding the identification data into computer readable media that may, for example, be transferred to another electronic device (e.g., electronic record), or by creating a hard copy of the display, such as a print out or physical record of information. The information may also be displayed by auditory signal or any other means of information communication. In some embodiments, the signal information may be detection data obtained using methods to detect the Y chromosome nucleic acid.

Once the signal information is detected, it may be forwarded to the logic processing module. The logic processing module may "call" or "identify" the fetal sex.

Provided also are methods for transmitting prenatal genetic information to a human pregnant female subject, which comprise determining the fetal sex in a pregnant woman where the fetal sex has been determined from the presence or absence of Y chromosome nucleic acid in a sample from the pregnant female; and transmitting the fetal sex to the pregnant female subject. The Y chromosome nucleic acid may be, for example, a nucleic acid corresponding to one or more Y chromosome loci described herein. The Y chromosome nucleic acid may, for example, be detected, may correspond to, or may be complementary to, a Y chromosome primer described herein.

Provided also are methods for transmitting prenatal genetic information to a human pregnant female subject carrying a male fetus, which comprise determining the amount of fetal DNA in a sample obtained from the pregnant woman, and where the amount of fetal DNA has been determined from the amount of Y chromosome nucleic acid in the sample; and transmitting the amount of fetal DNA to the pregnant female subject. The Y chromosome nucleic acid may be, for example, a nucleic acid corresponding to one or more Y chromosome loci described herein. The

Y chromosome nucleic acid may, for example, be detected, may correspond to, or may be complementary to, a Y chromosome primer described herein.

5 Provided also are methods for transmitting prenatal genetic information to a human pregnant female subject, which comprise determining the fetal sex in a pregnant woman where the fetal sex has been determined from the presence or absence of Y chromosome nucleic acid in a sample from the pregnant female; and transmitting prenatal genetic information representing the fetal sex to the pregnant female subject. The Y chromosome nucleic acid may be, for example, a nucleic acid corresponding to one or more Y chromosome loci described herein. The Y  
10 chromosome nucleic acid may, for example, be detected, may correspond to, or may be complementary to, a Y chromosome primer described herein.

15 Provided also are methods for transmitting prenatal genetic information to a human pregnant female subject carrying a male fetus, which comprise determining the amount of fetal DNA in a sample is obtained from the pregnant woman, and where the amount of fetal DNA has been determined from the amount of Y chromosome nucleic acid in the sample; and transmitting prenatal genetic information representing the amount of fetal DNA in the sample to the pregnant female subject. The Y chromosome nucleic acid may be, for example, a nucleic acid corresponding to one or more Y chromosome loci described herein. The Y chromosome nucleic  
20 acid may, for example, be detected, may correspond to, or may be complementary to, a Y chromosome primer described herein.

The terms “determining the fetal sex” or “determining the amount of fetal DNA” as used herein refers to any method for obtaining such information, including, without limitation, obtaining the  
25 information from a laboratory file. A laboratory file can be generated by a laboratory that carried out an assay to determine the fetal sex in a pregnant woman or the amount of fetal DNA in a sample. The laboratory may be in the same location or different location (e.g., in another country) as the personnel identifying the fetal sex or the amount of fetal DNA from the laboratory file. For example, the laboratory file can be generated in one location and transmitted to another location  
30 in which the information therein will be transmitted to the pregnant female subject. The laboratory file may be in tangible form or electronic form (e.g., computer readable form), in certain embodiments.

The term “transmitting the fetal sex to the pregnant female subject” or “transmitting the amount of  
35 fetal DNA in the sample” or any other information transmitted as used herein refers to

communicating the information to the female subject, or family member, guardian or designee thereof, in a suitable medium, including, without limitation, in verbal, document, or file form.

Also provided are methods for providing to a human pregnant female subject a medical  
5 prescription based on prenatal genetic information, which comprise determining the fetal sex in a pregnant female, where the fetal sex has been determined from the presence or absence of Y chromosome nucleic acid in a sample from the pregnant female; and providing a medical prescription based on the fetal sex. The Y chromosome nucleic acid may be, for example, a  
10 nucleic acid corresponding to one or more Y chromosome loci described herein. The Y chromosome nucleic acid may, for example, be detected, may correspond to, or may be complementary to, a Y chromosome primer described herein.

Also provided are methods for providing to a human pregnant female subject carrying a male fetus a medical prescription based on prenatal genetic information, which comprise determining  
15 the amount of fetal DNA in a sample, obtained from the pregnant woman, where the amount of fetal DNA has been determined from the amount of Y chromosome nucleic acid in the sample; and providing a medical prescription based on the amount of fetal DNA in the sample. The Y chromosome nucleic acid may be, for example, a nucleic acid corresponding to one or more Y chromosome loci described herein. The Y chromosome nucleic acid may, for example, be  
20 detected, may correspond to, or may be complementary to, a Y chromosome primer described herein.

The term "providing a medical prescription based on prenatal genetic information" refers to communicating the prescription to the female subject, or family member, guardian or designee  
25 thereof, in a suitable medium, including, without limitation, in verbal, document or file form.

Also provided are methods for providing to a human pregnant female subject a medical prescription based on prenatal genetic information, which comprise reporting to the pregnant female subject the fetal sex in the pregnant female, where the fetal sex has been determined  
30 from the presence or absence of Y chromosome nucleic acid in a sample from the pregnant female; and providing a medical prescription based on fetal sex to the pregnant female subject. The Y chromosome nucleic acid may be, for example, a nucleic acid corresponding to one or more Y chromosome loci described herein. The Y chromosome nucleic acid may, for example, be detected, may correspond to, or may be complementary to, a Y chromosome primer described  
35 herein.

Also provided are methods for providing to a human pregnant female subject carrying a male fetus, a medical prescription based on prenatal genetic information, which comprise reporting to the pregnant female subject the amount of fetal DNA in a sample obtained from the pregnant  
5 woman, where the amount of fetal DNA has been determined from the amount of Y chromosome nucleic acid in the sample; and providing a medical prescription based on the amount of fetal DNA in the sample to the pregnant female subject. The Y chromosome nucleic acid may be, for example, a nucleic acid corresponding to one or more Y chromosome loci described herein. The Y chromosome nucleic acid may, for example, be detected, may correspond to, or may be  
10 complementary to, a Y chromosome primer described herein.

The medical prescription may be for any course of action determined by, for example, a medical professional upon reviewing the prenatal genetic information. For example, the prescription may be for the pregnant female subject to undergo an amniocentesis procedure. Or, in another  
15 example, the medical prescription may be for the pregnant female subject to undergo another genetic test. In yet another example, the medical prescription may be medical advice to not undergo further genetic testing.

Also provided are files, such as, for example, a file comprising the determination of fetal sex in a  
20 pregnant female subject, where the fetal sex has been determined from the presence or absence of Y chromosome nucleic acid in a sample from the pregnant female.

Also provided are files, such as, for example, a file comprising the determination of the amount of fetal DNA in a sample, where the sample is obtained from a pregnant woman carrying a male  
25 fetus, where the amount of fetal DNA has been determined from the amount of Y chromosome nucleic acid in the sample.

The file may be, for example, but not limited to, a computer readable file, a paper file, or a  
30 medical record file.

Computer program products include, for example, any electronic storage medium that may be used to provide instructions to a computer, such as, for example, a removable storage device, CD-ROMS, a hard disk installed in hard disk drive, signals, magnetic tape, DVDs, optical disks, flash drives, RAM or floppy disk, and the like.  
35

The systems discussed herein may further comprise general components of computer systems, such as, for example, network servers, laptop systems, desktop systems, handheld systems, personal digital assistants, computing kiosks, and the like. The computer system may comprise one or more input means such as a keyboard, touch screen, mouse, voice recognition or other  
5 means to allow the user to enter data into the system. The system may further comprise one or more output means such as a CRT or LCD display screen, speaker, FAX machine, impact printer, inkjet printer, black and white or color laser printer or other means of providing visual, auditory or hardcopy output of information.

10 The input and output means may be connected to a central processing unit which may comprise among other components, a microprocessor for executing program instructions and memory for storing program code and data. In some embodiments the methods may be implemented as a single user system located in a single geographical site. In other embodiments methods may be implemented as a multi-user system. In the case of a multi-user implementation, multiple central  
15 processing units may be connected by means of a network. The network may be local, encompassing a single department in one portion of a building, an entire building, span multiple buildings, span a region, span an entire country or be worldwide. The network may be private, being owned and controlled by the provider or it may be implemented as an internet based service where the user accesses a web page to enter and retrieve information.

20 The various software modules associated with the implementation of the present products and methods can be suitably loaded into the a computer system as desired, or the software code can be stored on a computer-readable medium such as a floppy disk, magnetic tape, or an optical disk, or the like. In an online implementation, a server and web site maintained by an  
25 organization can be configured to provide software downloads to remote users. As used herein, "module," including grammatical variations thereof, means, a self-contained functional unit which is used with a larger system. For example, a software module is a part of a program that performs a particular task. Thus, provided herein is a machine comprising one or more software modules described herein, where the machine can be, but is not limited to, a computer (e.g.,  
30 server) having a storage device such as floppy disk, magnetic tape, optical disk, random access memory and/or hard disk drive, for example.

The present methods may be implemented using hardware, software or a combination thereof and may be implemented in a computer system or other processing system. An example  
35 computer system may include one or more processors. A processor can be connected to a



communication bus. The computer system may include a main memory, sometimes random access memory (RAM), and can also include a secondary memory. The secondary memory can include, for example, a hard disk drive and/or a removable storage drive, representing a floppy disk drive, a magnetic tape drive, an optical disk drive, memory card etc. The removable storage drive reads from and/or writes to a removable storage unit in a well-known manner. A removable storage unit includes, but is not limited to, a floppy disk, magnetic tape, optical disk, etc. which is read by and written to by, for example, a removable storage drive. As will be appreciated, the removable storage unit includes a computer usable storage medium having stored therein computer software and/or data.

10

In alternative embodiments, secondary memory may include other similar means for allowing computer programs or other instructions to be loaded into a computer system. Such means can include, for example, a removable storage unit and an interface device. Examples of such can include a program cartridge and cartridge interface (such as that found in video game devices), a removable memory chip (such as an EPROM, or PROM) and associated socket, and other removable storage units and interfaces which allow software and data to be transferred from the removable storage unit to a computer system.

15

The computer system may also include a communications interface. A communications interface allows software and data to be transferred between the computer system and external devices. Examples of communications interface can include a modem, a network interface (such as an Ethernet card), a communications port, a PCMCIA slot and card, etc. Software and data transferred via communications interface are in the form of signals, which can be electronic, electromagnetic, optical or other signals capable of being received by communications interface. These signals are provided to communications interface via a channel. This channel carries signals and can be implemented using wire or cable, fiber optics, a phone line, a cellular phone link, an RF link and other communications channels. Thus, in one example, a communications interface may be used to receive signal information to be detected by the signal detection module.

25

30

In a related aspect, the signal information may be input by a variety of means, including but not limited to, manual input devices or direct data entry devices (DDEs). For example, manual devices may include, keyboards, concept keyboards, touch sensitive screens, light pens, mouse, tracker balls, joysticks, graphic tablets, scanners, digital cameras, video digitizers and voice recognition devices. DDEs may include, for example, bar code readers, magnetic strip codes,

35

smart cards, magnetic ink character recognition, optical character recognition, optical mark recognition, and turnaround documents. In one embodiment, an output from a gene or chip reader may serve as an input signal.

5 In some embodiments, provided is a kit for determining the fetal sex in a pregnant woman. One component of the kit is primers for amplifying the region of interest, such as, for example, primers directed to one or more loci described herein. Another component of the kit may be, for example, a standard control primer representing chromosome nucleic acid that is not Y-chromosome specific.

10 In some embodiments, provided is a kit for determining the amount of fetal DNA in a sample, where the sample is obtained from a pregnant woman carrying a male fetus. One component of the kit is primers for amplifying the region of interest, such as, for example, primers directed to one or more loci described herein. Another component of the kit may be, for example, a standard control primer representing chromosome nucleic acid that is not Y-chromosome specific.

### Kits

Furthermore, the invention in part provides kits comprising compositions described herein and optionally instructions for carrying out methods described herein. Parts of the kit can be packaged individually in vials or in combination in containers or multicontainer units. Kits may be advantageously used for carrying out methods described herein and can be, inter alia, employed in a variety of applications referred herein. Manufacture of kits often are in accordance with standard procedures known to persons skilled in the art.

25

### Examples

The following examples illustrate certain embodiments of the invention and are not limiting. Those of skill in the art will readily recognize a variety of non-critical parameters that can be changed or modified to yield essentially similar results.

30

Example 1: Fetal Sex Assay

Determination of fetal sex was enabled by multiplex PCR followed by multiplexed primer extension analysis using MALDI-TOF MS. Initial testing on male and female genomic DNA showed the ability of the assay system to discriminate the presence or absence of Y-chromosomal DNA. Secondary studies were performed using DNA mixtures with as low as 20 genomic copies of either male or female DNA mixed with a 50-fold excess of female genomic DNA.

- 5
- 10 Testing of the assay using maternal plasma and non-pregnant female plasma samples was performed and shows a high degree of reproducibility between replicate sample aliquots of maternal plasma and high specificity as determined by the lack of Y-chromosomal target detection in plasma DNA isolated from non-pregnant female donors.

15 *Detailed steps for the fetal sex determination assay protocol*

The protocol provided below does not limit the scope of the invention. Instead it provides an representative protocol for practicing a part of the invention.

- 20 1. Prepare multiplex PCR primer mix using a set of primers provided in Table 3 (MP set 1, 2, 3 or 4). Mix 100  $\mu$ l primer (5.0  $\mu$ l volume) with 910  $\mu$ l water for a final concentration of 0.5  $\mu$ M.

TABLE 3

MP Set	ASSAY_ID	2nd-PCR Primer	1st-PCR Primer	Extend Primer
1	SRY-4-i-2	<u>ACGTTGGATGGCATT</u> TTCCACTGGTATCCC	<u>ACGTTGGATGAGATGGCTCTAGAGAA</u> TCCC	CCAGAA <u>TGCGAA</u> ACTC
	HSFY-1	<u>ACGTTGGATGAAAGTTGCCT</u> TTCTGCCTGC	<u>ACGTTGGATGTTCCCTCTCTCTCC</u> ACCCC	CCCATCTCTTCTCAATCC
	RBM2-1	<u>ACGTTGGATGGAAGTTGGAGAGT</u> TACTCG	<u>ACGTTGGATGAGAGAAGGCGGAT</u> TCCTTTG	AAACATGCTCACGATCAC
	ALB-2-i	<u>ACGTTGGATGGCAAATTCAGTT</u> ACTTCATTC	<u>ACGTTGGATGCAGTATCTTCAGCAG</u> TGTCC	GCAGTGTCCATTTGAAGAT
	TTY22-1	<u>ACGTTGGATGCTCTGGCTAGCAT</u> CACAATG	<u>ACGTTGGATGACGAAGGGCTACT</u> CTCTAC	GCAGGGACTTGTGCGTAGG
	RBM1A1-1	<u>ACGTTGGATGTCTATTCTTCC</u> CGAGAGACC	<u>ACGTTGGATGGAACCTCAGGCT</u> CTTTGTCC	AAAAAGCGCAGGTATTTTCTG
	XKRY-1	<u>ACGTTGGATGCCCTCTGTGGT</u> ACAGAAAC	<u>ACGTTGGATGATGACCCCAAAA</u> GCACAGAC	CAATTGGATCACATTACATCAAA
	TTY16-1	<u>ACGTTGGATGTCGAAATTTGAT</u> TCCCAGAGG	<u>ACGTTGGATGCCAGACTGAAAT</u> TCCCAGG	ggggTGAAATCCCAGACAATGGA
	CDY1-1	<u>ACGTTGGATGGTCAGGAGATC</u> GAGACAATC	<u>ACGTTGGATGTGCCCTACGAA</u> TTTGTG	GGCCCTACGAAATTTGTGTATTTT
	PRY-2	<u>ACGTTGGATGCTTGAGTCTGGG</u> AGTTTGAG	<u>ACGTTGGATGACCACAGCTGG</u> CTAATTTTG	GTCACATTTTGTGGCCCT
	TTY22-2	<u>ACGTTGGATGACTAGGCACCT</u> CAATTCAG	<u>ACGTTGGATGTCCAGTAGTGT</u> CTGTTGAGAG	GGACTCTCAATGAAAGCA
	ALB-2-i [1]	<u>ACGTTGGATGGCAAATTCAGT</u> TACTTCATTC	<u>ACGTTGGATGCAGTATCTTCAG</u> CAGTGTCC	GCAGTGTCCATTTGAAGAT
	XKRY-2	<u>ACGTTGGATGGCATTCTGTT</u> GGTATTATC	<u>ACGTTGGATGAGTAAAGCT</u> CCCCCTGTTTC	IGGAAGAAATGCCAGAGTCA
2	TTY16-2	<u>ACGTTGGATGCCTCATGAAG</u> GGATATGTGC	<u>ACGTTGGATGGTGGCATTGAC</u> TGTGTGAC	TAGCCATAGTGCACATCTCA
	RBM2-2	<u>ACGTTGGATGTTCTTTGGCT</u> GAGAAAGGAC	<u>ACGTTGGATGAGTGACACAG</u> GGGAAAACACG	CTCTTCACTGATTTTAAAGTTT
	BPY22-2	<u>ACGTTGGATGGTAGAGTAGA</u> GTGACACTCC	<u>ACGTTGGATGTGATGTGACAG</u> CCCTCAAAG	TGTCACCAAGCACACTATTGCCAGG
	RBM1A1-2	<u>ACGTTGGATGGAGTCTTTC</u> CTTGTGCG	<u>ACGTTGGATGCTTTTAGAGC</u> GTAGACAAAAC	TAGACGTAGACAAAACCTGGATAGACA
	CDY1-2	<u>ACGTTGGATGTGGAGGCAG</u> AAAGAGTGTC	<u>ACGTTGGATGAAAAGACACG</u> GGATTGGAAC	cccagTGGAAACAGTTGCTGTGCCACCT
	SRY-4-i-2 [1]	<u>ACGTTGGATGGCATT</u> TTCCACTGGTATCCC	<u>ACGTTGGATGAGATGGCTCT</u> AGAGAAATCCC	CCAGAA <u>TGCGAA</u> ACTC
	CYORF14-3	<u>ACGTTGGATGATCAACA</u> AAACAAGGGGCTTC	<u>ACGTTGGATGGCTACTGGT</u> CTAGCCCTTAT	GACATGAAGTCAATTTGCT

	ALB-3-i	<u>ACGTTGGATGGCAAATTCAGTTACTTCAATC</u>	<u>ACGTTGGATGCAGTATCTTTCAGCAGTGTCC</u>	cGCAGTGTCACATTTGAAGAT
	TTY22-3	<u>ACGTTGGATGACCCCAATAGAAGTGATAG</u>	<u>ACGTTGGATGTGGCTGGACAGCTTTCATAC</u>	AATTCGAGTTACAGCCACCG
	RBM1A1-3	<u>ACGTTGGATGTCATACGTAGAGTATCGGTG</u>	<u>ACGTTGGATGTGGAATGTGTGGCTTGGG</u>	GTATCCTGAAGCCCAATAAATAC
	XKRY-3	<u>ACGTTGGATGAGGCCAAATAGCTTTACTC</u>	<u>ACGTTGGATGCCCTGAAGGAATAAATGGAC</u>	GAAGGAATAAATGGACTCTCGAT
	HSFY-3	<u>ACGTTGGATGGTCACTGCAC TAGGCATTC</u>	<u>ACGTTGGATGTACCAAGGCATTTGGACTCTG</u>	CTGCACATGAGATACATATCTTCC
	PRY-3	<u>ACGTTGGATGAATTCAGAGCCTGACCCCAAG</u>	<u>ACGTTGGATGTGGACCCCAAGGATAAAC</u>	GTGGACCCCAAGGATAAACAATAA
	CDY1-3	<u>ACGTTGGATGGCATGCCCTCCCTTGCTGTC</u>	<u>ACGTTGGATGGGTAAAGCATCTGCCAATG</u>	CTGCCAATGAATGTTAATTGCTGGGC
	HSFY-4	<u>ACGTTGGATGGCTCTCTACTTTACCTTCCC</u>	<u>ACGTTGGATGTGGAACATAGAGAGCACCCAG</u>	AGCAACCCCAACCCCTCTGC
	TTY16-4	<u>ACGTTGGATGTCAAGACTGTGAGGTGGTTG</u>	<u>ACGTTGGATGGTGGTCTCTGATTCTTGGCAC</u>	GGCTCTTTCAGGAATGGA
	XKRY-4	<u>ACGTTGGATGTGGAAGTTACAGGCCTTGAG</u>	<u>ACGTTGGATGTGACAAAAGTCGACTCAGTGC</u>	TTGCACCTGTATCACACAGT
	RBM1A1-4	<u>ACGTTGGATGGACCATAGTCTCAGTATGCC</u>	<u>ACGTTGGATGGCACACACAAAATCATCCAAG</u>	GTTCCCGAGAAAAC TAGGGGATT
4	TTY22-4	<u>ACGTTGGATGGGAAGTTGTATAATTGCTCC</u>	<u>ACGTTGGATGCAATGTTCACTGCCCATTC</u>	CCATTGACTGCTCAAAATTTACA
	ALB-4-i	<u>ACGTTGGATGGCAAATTCAGTTACTTCAATC</u>	<u>ACGTTGGATGCAGTATCTTTCAGCAGTGTCC</u>	tcCGCAGTGTCCATTTGAAGAT
	BPY22-4	<u>ACGTTGGATGTCAACCAAGGGATGAAAGCC</u>	<u>ACGTTGGATGAATGATGCTTTCAGTCCCACC</u>	GCCAGAGCCACAGAGGGGCATTTT
	CDY1-4	<u>ACGTTGGATGGCTTTTGCATAACTGAGCAC</u>	<u>ACGTTGGATGCCCGCTACACTTTGTATGACC</u>	ccccgGTCTCTTTTATGAGCACA
	RBM1A2-4	<u>ACGTTGGATGGTCAGGTACATACTAAAATGC</u>	<u>ACGTTGGATGTGTAACACTCAGAGCAAAGG</u>	GTAACACTCAGAGCAAAGGGGCAC
5	DBY	<u>ACGTTGGATGCTGAATGAAAGTTCAACCGTC</u>	<u>ACGTTGGATGGATACTATGCCACCAAAGGG</u>	gCAGTACCTGTATTTCCCTTAG
6	TTY1	<u>ACGTTGGATGTACCTTATGGCAGGGACTTGG</u>	<u>ACGTTGGATGGTCTGGCTAGCATCACAAATG</u>	gTATGGCAGGGACTTGTCCGCTAGG

PCR primers and extend primers are provided in Table 3. Designations MP1- MP4 correspond to multiplex (MP) primer sets 1-4. The underlined portion of the primer is a universal tag sequence for improved amplification reactions. Lower case nucleotides in the extend primer sequence represent non-template nucleotides that increase the mass of the oligonucleotide.

In Table 3, MP5 provides PCR primers and an extend primer directed to exon 6 of the dead box, Y isoform gene (NCBI reference mRNA sequence NM\_004660 (DBY); chromosome positions chrY:13537002-13537120 from NCBI Build 36.1)

5 MP6 provides PCR primers and an extend primer directed to two sequences on the Y chromosome: testis-specific transcript, Y-linked 22 - 5'UTR (NCBI reference mRNA sequence AF527834 (TTY22); chromosome positions chrY:10248638-10248718 from NCBI Build 36.1) and an unnamed sequence that is 26kb upstream of TTTY1 (chrY:6291654-6291734 from NCBI Build 36.1). The TTTY assay provides an ultra-  
 10 sensitive test for gender determination since two Y-chromosome sequences are detected.

The SRY assay provides PCR primers and an extend primer directed to exon 1 of the sex determining region Y (NCBI reference mRNA sequence NM\_003140.1 (SRY); chromosome positions chrY:2715334-2715425 from NCBI Build 36.1). This assay may  
 15 be used alone or in combination with other assays to determine the sex of a fetus, or to determine the amount of fetal nucleic acid present in a sample.

2. Prepare PCR cocktail as in Table 4.

20

<b>Table 4. PCR cocktail preparation with UNG</b>		
Reagent	Conc. in 50 ul rxn	Volume reagent for 50ul reaction (ul)
Water, HPLC grade	N/A	7.55
10xPCR Buffer (contains 15mM MgCl <sub>2</sub> , Tris-Cl, KCl, (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> , pH 8.7 (20°C))	1.25x	6.25
25mM MgCl <sub>2</sub> (Qiagen)	1.625mM	3.25
PCR Nucleotide Mix PLUS (10 mM dATP, dCTP, dGTP/30mM dUTP)	200µM	1
0.5uM PCR primer mix - F, -R	0.1 µM	10
1U/µl Uracil-DNA-Glycosylase (UNG)(NEB)	1.25U/rxn	1.25
5U/µl HotStar Taq (Qiagen)	3.5U/rxn	0.7

DNA (plasma) Added Separately	--	20
Total		50

3. To the DNAs, add 30 uL PCR cocktail per well, mixing well.
4. Seal plate and centrifuge.
5. Run PCR with \_PCR30-11 \_ program on PCR cycler (Table 5).

5

<b>Table 5. PCR 30-11 Thermocycling Conditions</b>		
PCR Cycle	Cycling Conditions	Number of Cycles
UDG Incubation	30°C for 10 minutes	1 Cycle
Initial Denaturation	94°C for 15 minutes	1 Cycle
Cycled Template Denaturation	94°C for 20 seconds	30 Cycles
Cycled primer Annealing	58°C for 30 seconds	
Cycled primer Extension	72°C for 1 minute	
Cycled Template Denaturation	94°C for 20 seconds	11 Cycles
Cycled primer Annealing	62°C for 30 seconds	
Cycled primer Extension	72°C for 1 minute	
Final Extension	72°C for 3 minute	1 Cycle
Hold	4°C	

6. Prepare SAP cocktail as in Table 6.

<b>Table 6. SAP Cocktail Mix</b>	
SAP Mix Reagent	Volume for 1 rxn (uL)
Nanopure Water	2.95 uL
SAP Buffer	0.34 uL
Shrimp Alkaline Phosphatase (SAP) (1.7U/uL)	0.71 uL
Total Volume	4.00 uL

7. From the PCR plate, aliquot 10 uL PCR into a new 96-well plate.
8. Dispense 4 uL SAP cocktail mix into the PCR aliquoted plate.
9. Seal plate, centrifuge, and cycle using the SAP 40-5 conditions in Table 7.

<b>Table 7. SAP Thermal Cycling Conditions – SAP_40_5</b>		
SAP Cycle	Cycling Conditions	Number of Cycles
Initial Incubation	37°C for 40 minutes	1 Cycle
Cycled Template Denaturation	85°C for 5 minutes	1 Cycle
Hold	4°C	

10. Prepare a multiplexed extend primer mix according to Table 8.

<b>Table 8. Extend Primer Mix Preparation</b>				
200uM Extend primer	Conc. in mix	Volume primer	Volume Water	Final conc. primer in 18 uL rxn
SRY-4-i-2	7uM	17.5	272.5	0.625uM
HSFY-1	7uM	17.5		
RBM2-1	7uM	17.5		
ALB-2-i	7uM	17.5		
TTY22-1	7uM	17.5		
RBM1A1-1	14uM	35		1.25uM
XKRY-1	14uM	35		
TTY16-1	14uM	35		
CDY1-1	14uM	35		
	Total	500 ul		

5 11. Prepare EXTEND cocktail as in Table 9 using the corresponding multiplex set of Extend Primers provided in Table 3.

TABLE 9

<b>Extend Reagent</b>	<b>Volume for 1 rxn</b>
Water (HPLC grade)	1.238 uL
iPLEX Buffer Plus (10x)	0.4 uL
iPLEX termination Mix	0.4 uL
Extend Primer Mix, 7/14 uM	1.88 uL
Thermostable DNA Polymerase (32 U/uL)	0.082 uL
Total Volume	4.0 uL

10 12. Dispense 4 uL of the Extend cocktail mix into corresponding wells of the plate from the SAP incubation step.



13. Seal plate, centrifuge, and cycle with the following conditions in Table 10.

Table 10. 200 step EXTEND cycling		
PCR Cycle	Cycling Conditions	Number of Cycles
Initial Denaturation	94°C for 30 seconds	1 Cycle
Cycled Template Denaturation	94°C for 5 seconds	} 5 cycles }40 cycles
Cycled primer Annealing	52°C for 5 seconds	
Cycled primer Extension	80°C for 5 seconds	
Final Extension	72°C for 3 minutes	
Hold	4°C	

- 5 14. Sample Conditioning, Dispensing, and Acquiring Spectra:
- a. Add 32ul water per well.
  - b. Add 15mg Clean Resin, per well.
  - c. Rotate 360 degrees for 10 min. Centrifuge at 4000 rpm for 5 minutes

- 10 15. Spotting to SpectroCHIP® solid support and MALDI run:
- a. Pintool: Samsung 6-pin Nanodispenser
  - b. pintool settings:
    - i. Calibrant: 96 to 96, non-skirted plate in adapter, 120mm/sec dispense speed
    - 15 ii. Analyte: 96 to 96, non-skirted plate in adapter, 20-50 mm/sec dispense speed (dependent upon volume check)

Steps 10-14 are further described hereafter. Following genomic amplification, the assay interrogates amplified regions through the use of specific primers that are designed to hybridize directly adjacent to the site of interest. These DNA oligonucleotides are referred to as iPLEX® MassEXTEND® primers. In the extension reaction, the iPLEX primers are hybridized to the complementary DNA templates and extended with a DNA polymerase. Special termination mixtures that contain different combinations of deoxy- and dideoxynucleotide triphosphates along with enzyme and buffer, direct limited extension of the iPLEX primers. Primer extension occurs until a complementary dideoxynucleotide is incorporated.

The extension reaction generates primer products of varying length, each with a unique molecular weight. As a result, the primer extension products can be simultaneously separated and detected using Matrix Assisted Laser Desorption/Ionization, Time-Of-Flight(MALDI-TOF) mass spectrometry on the Sequenom MassARRAY® Analyzer Compact System.

### *Assay Performance*

Initial development of XLR (X-linked, Lymphocyte Regulated) assay was performed by designing assay multiplexes that each amplified 11 Y-chromosomal sequences and the ALB sequence. These were tested for their ability to generate assay calls and successful primer extension when amplifying ~1000 genomic copies of male DNA. At high DNA copy number, assays within each multiplex gave expected calls but variable extend rates with 35-40 PCR cycles and a 200 step extend cycling reaction. Further testing was performed using a dilution series of template DNA with 50-800 genomic copies of male DNA per reaction. As indicated by decreased call rates and extend rates, these studies identified individual Y-chromosome assays within the multiplexes that showed decreased performance when using lower amounts of DNA template. The three poorest performing Y-chromosome target assays from each multiplex were removed to give four multiplexes designed to amplify eight Y-chromosomal sequences and ALB. See Table 3.

Performance of each of these resulting 9-plexes was tested using between 12-400 copies of male genomic DNA, with results showing equivalently high call rates in assays of multiplexes 1, 2 and 4 at all levels of male genomic DNA, but with variable extension rates. Also in these studies, multiplex 3 showed consistent call drop out of one Y-chromosome assay (PRY-4) when using low copy numbers of male DNA. These resulting nine-plexes were carried forward for further testing using the DNA mixture model system described below.

### *Multiplexing Results*

A mixture model was developed to test the ability of the four XLR gender multiplexes to detect low copy numbers of Y-chromosomal target sequences in a high copy number background of female genomic DNA. This model system was designed to mimic the mixtures of fetal and maternal DNA in a maternal plasma sample. In this model, mixtures were prepared so that each PCR would receive 20 genomic copies ('fetal') of either a male or female CEPH genomic DNA combined with 980 genomic copies ('maternal') of a female CEPH DNA to generate a 2% mixture

of 'fetal'/'maternal' DNA. Five female 'fetal' DNAs and six 'fetal' male DNAs were combined with 8 female 'maternal' DNAs, to generate 88 unique DNA samples for analysis.

5 Specific PCR and extend primer sequences used for the four individual multiplexes are listed in Table 3.

The four XLR multiplexes containing ALB and 8 Y-chromosomal assays were tested for their ability to detect Y-chromosomal targets in the 2% DNA mixtures described above. Standard PCR 40 58-2s-62 cycling conditions were used for these studies.

10 Each of the multiplexes showed some degree of spurious Y-target calls in samples containing only female DNA. MP1, MP2, MP3, and MP4 showed 7, 3, 7, and 8 spurious Y-target calls, respectively, when using female/female DNA mixtures. These spurious calls are possibly associated with the method of PCR set up for the DNA mixtures, with the use of a prePCR liquid  
15 handler (Matrix) to add DNA to the PCR plate containing aliquoted PCR cocktail. MP1 was the only multiplex that showed all Y-target calls for samples containing 2% male genomic DNA. MP2 showed notable dropout for Y-target assays PRY-2 and TTTY22-2 in samples containing 2% male genomic DNA. MP3 showed notable dropout for TTTY22-3 and no detection of the PRY-3 target in samples containing 2% male genomic DNA. MP4 showed notable dropouts for TTTY22-  
20 4 target detection in samples containing 2% male genomic DNA. The assay failure involving the TTTY22 loci in multiplexes 2, 3, and 4 occurred only in DNA mixtures containing the male DNA NA04477. One potential explanation might be DNA specific alteration of the TTTY22 region that interferes with amplification using the specific amplicon designs for TTTY22-2, TTTY22-3, and TTTY22-4, but not the TTTY22-1 assay design in MP1. Because MP1 was the only multiplex  
25 without dropout of Y-target calls in samples containing Y-chromosomal template and MP1 showed similar levels of spurious calls to the other multiplexes, MP1 was carried forward to further testing. However, MP2, MP3 and MP4 are still considered good tests for the determination of fetal sex.

### 30 *Replicate Analysis*

Replicate analysis of the 2% male DNA mixtures was used to test the reproducibility of Y-target detection in samples containing Y-chromosomal DNA template and specificity of Y-target detection in samples containing no Y-chromosomal template. PCR-40 58-2s-62 cycling was used  
35 in these studies. All Y-target assays were successfully called in samples containing male DNA

(XY/XX samples on right side of the panel). Multiple spurious Y-target calls were made in female DNA only samples. While there was no correlation between specific DNAs and spurious calls, spurious Y-target call rates were highest in assays RBMY1a1 and RBMY2. The RBMY gene family has multiple copies on the Y-chromosome and shares homology with RBMX gene on the X-chromosome. One or both of these factors may play a role in the higher degree of spurious calls for the RBMY assays. One explanation is that the higher copy numbers of the RBMY template on the Y-chromosome may lead to higher chances of successful amplification when there is a single copy contamination event with Y-chromosome fragments containing the RBMY gene loci. Alternatively, non-specific recognition of RBMX template by RBMY primers at early PCR cycle number may lead to amplification. However, this non-specific amplification appears unlikely as there were no candidate template sites identified when gene specific PCR primer sequences for both RBMY1a1 and RBMY2 assay were queried through the NCBI reverse e-PCR site.

15 *PCR Cycling Conditions*

In an effort to match assay conditions between RHD and XLR, PCR cycling conditions used for the RHD assay (PCR 30-11) were tested with XLR MP1. In addition, a higher stringency cycling condition (PCR 40-62) was also tested. These additional cycling conditions are summarized in Table 11.

With PCR 40-62 cycling conditions, no improvement was seen in RBMY1a1 and RBMY2 spurious call rates. Spurious calls rates may be improved in other assays using the PCR 40-62 conditions, however, this condition was tested on only one plate replicate.

Of note, the median extension rate of assay CDY1-1 was improved from ~75% in both the PCR 40 58-2s-62 and PCR 30-11 cycling conditions to nearly complete extension when using PCR 40-62 conditions. However, as there was no change in call rate with cycling conditions, it was decided that PCR 30-11 cycling would be used in future work in an effort to minimize differences between the XLR assay and RHD assays.

TABLE 11 - Alternate PCR cycling conditions

PCR 30-11	UNG Incubation	30°C 10 minutes	1 cycle
	Taq Activation	94°C 15 minutes	1 cycle
	Denature	94°C 20 seconds	30 cycles
	Anneal	58°C 30 seconds	

	Elongation	72°C 60 seconds	10 cycles
	Denature	94°C 20 seconds	
	Anneal	62°C 30 seconds	
	Elongation	72°C 60 seconds	
	Final Elongation	72°C 3 minutes	1 cycle
	Storage	4°C forever	1 cycle
PCR 40-62	UNG Incubation	30°C 10 minutes	1 cycle
	Taq Activation	94°C 15 minutes	1 cycle
	Denature	94°C 20 seconds	40 cycles
	Anneal	62°C 30 seconds	
	Elongation	72°C 60 seconds	
	Final Elongation	72°C 3 minutes	1 cycle
	Storage	4°C forever	1 cycle

*Analytical sensitivity*

5 In five replicate analyses of 48 DNA mixtures containing 20 genomic copies of male DNA in a background of 980 genomic copies female DNA, 8 Y-targets were detected in each case using the MP1 assay design with PCR 40 58-2s-62 cycling conditions. Using an assay calling rule of 8 positive Y-targets to determine the presence of male DNA, there was 100% sensitivity to detect the presence of male DNA in this model system.

10 In five replicate analyses of 40 DNA mixtures containing 1000 genomic copies of only female DNA, 0 Y-target calls in 169 of 200 samples containing only female DNA, 1 Y-target call in 26 of 200 female DNA samples, and 2 Y-target calls in 5 of 200 female DNA samples measured.

15 *Plasma samples from pregnant women*

Maternal plasma was collected using an IRB approved protocol with donor consent. 20 mL peripheral blood was collected in lavender top EDTA blood tubes. Following collection, samples were centrifuged and the plasma fraction transferred to individual 1 mL aliquots. Aliquoted plasma was stored at -80°C.

For all plasma samples, plasma DNA was isolated using the QIAamp Viral MinElute kit from Qiagen® with the vacuum protocol. Final elution was performed with 60 ul water, and 20 ul of this eluate was used for PCR unless otherwise indicated. PCR 30-11 cycling conditions were used unless indicated. TypePLEX extend reaction was performed with the 200 extension reaction cycling.

Any DNA isolation, extraction or enrichment method known in the art or not yet developed may be used that yields sufficient fetal nucleic acid.

5 Maternal plasma obtained from Precision Med protocol #3402 was used to test the ability of fetal gender assay MP1 to detect to the presence of Y-chromosomal DNA in maternal plasma. For this study, 2 separate aliquots of 14 second trimester maternal plasma samples were tested independently on different days. Of these samples, fetal gender phenotype was known in only 5 pregnancies.

10 Shown in Figure 2 below is the summary of Y-target calls for the first replicate of each plasma aliquot from each assay date. Of the 14 samples, 6 donors (right most six samples) appear to carry a male fetus as judged by the presence of 8 Y-target calls in both analysis dates. Male fetal gender phenotype was confirmed in 2 of these pregnancies but no fetal gender phenotype information was available for the remaining 4 samples. Of the 8 samples on the left of the figure, 15 5 show no Y-target calls in either replicate analysis, while 3 show 1 Y-target call in either the 20070829 or 20071130 analysis but not both. Given the results from the DNA model testing of female only DNAs showing 0-2 Y-target calls, these results suggest these 8 donors carry a female fetus. Female fetal gender was confirmed in 3 of these pregnancies, but no fetal gender 20 phenotype information was available for the remaining 5 samples as of the preparation of this report. Of the results with apparent female fetus, spurious calls for samples #204, 218, and 251 were for RBMY1a1, TTTY16, and RBMY2, respectively.

*Plasma samples from non-pregnant female*

25 Replicate plasma aliquots obtained from a non-pregnant female donor were tested in parallel as a negative control for the later Precision Med 3402 sample analysis (20071130). As seen in Figure 3, of 10 non-pregnant female plasma DNA preparations, only one aliquot showed 1 Y-target call. Inclusion of 23 NTC PCRs (water only) showed 7 NTC samples each with one Y-target call. The 30 remaining 16 NTC samples showed no Y-target calls. These results correspond to those seen in DNA mixture model testing.

*Additional plasma samples*

In addition to the plasma samples above, plasma samples from Hemacare® protocols #0049 and #0071 were assayed. In total, 57 plasma aliquots obtained from 25 donors at varying time points during pregnancy were assayed. These included 8 donors with plasma obtained from each trimester, 5 donors with plasma obtained from both 1st and 2nd trimester, and 12 donors with plasma collected from 1st trimester only. No fetal gender information was available for these samples at the time of testing.

50 Fifty-two PCR analyses showed 8 Y-target calls. These plasma DNA isolates were derived from 12 donors and show 100% concordance for 8 Y-target calls between both plasma DNA PCR replicate testing and sampling at various trimesters. This demonstrates the high degree of reproducibility of results when all Y-targets are called. Of the remaining 62 plasma DNA PCR replicates, 20 show 0 Y-target calls and 42 show 1-4 Y-target calls. Samples #004, #011 and #063 (with 3, 3, and 4 Y-target calls in at least one PCR replicate, respectively) are from donor very likely carrying female fetuses. This conclusion can be reached by comparing Y-target calls from individual plasma DNA replicate preparations between plasma samples collected in the 1st, 2nd, and 3rd trimesters of pregnancy for these donors. Of note in the samples with 1-4 Y-target calls is the minimal concordance of Y-target calls that are made between PCR replicates, indicating a high degree of variability in these calls and that the variability arises during the PCR amplification. That the spurious calls originate in the PCR is consistent with knowledge gained in prior development of RHD and other assays.

In total, there were 69 Y-target calls in all Hemacare plasma samples that showed less than 8 positive Y-targets. All assays except RBMY1a1 and RBMY2 gave between 3-6 calls per Y-target assay in these samples. However, Y-target assays for RBMY1a1 and RBMY2 gave 26 and 16 calls, respectively in these samples.

Using the remaining portion of plasma DNA isolated from the original plasma aliquot, a third PCR replicate was generated for samples 001 (as a negative control), 003 (as a positive control), 004, 011, and 063. The DNA volume used in these PCRs was only 12-17 ul instead of the 20 ul used in the first 2 PCR replicates due to limited sample volume. The resulting PCR product was assayed in duplicate in the TypePLEX extend reaction.

*Overall Performance Assessment*

The 2% DNA mixture models studies demonstrated the ability of XLR MP1 to detect male DNA in a sample at concentrations as low as 20 genomic copies in a background of female genomic DNA as high as 980 genomic copies. In the replicate studies of MP1, XLR MP1 gave 8 Y-target calls in 240 of 240 samples analyses containing male DNA. In parallel, the XLR MP1 assay gave 0 Y-target calls in 169 of 200 samples containing only female DNA, 1 Y-target call in 26 of 200 female DNA samples, and 2 Y-target calls in 5 of 200 female DNA samples measured. These results demonstrate a distinct grouping pattern whereby the presence of male DNA in a sample, even at concentrations as low as 20 genomic copies, will give 8 Y-target calls and the absence of male DNA in the sample will give less than 2 Y-targets calls.

Representative criteria for assay calling are provided in Table 12.

15 TABLE 12

Y-target calls	Assay result call
8	Male
3-7	Inconclusive
0-2	Female

Example 2: Examples of Embodiments

20 Described hereafter are non-limiting examples of certain embodiments of the invention.

A1. A method for identifying the presence or absence of Y-chromosome nucleic acid in extracellular nucleic acid from a pregnant female, which comprises:

(a) contacting under amplification conditions extracellular nucleic acid from a pregnant female with one or more primer sets selected from the group consisting of:

**first primer**

AGATGGCTCTAGAGAATCCC  
 TTTCTCTCTCTCCACCCC  
 AGAGAAGGCGGATTCTTTG  
 ACGAAGGGCTACTTCTCTAC  
 GAACCTCAGGCTCTTTGTCC  
 ATGACCCCAAAGCACAGAC

**corresponding second primer**

GCATTTTCCACTGGTATCCC  
 AAAGTTGCCTTTCTGCCTGC  
 GAAGTTGGGAGAGTTACTCG  
 CTCTGGCTAGCATCACAATG  
 TCTATTCTTGCCGAGAGACC  
 TCCCTCTGTGGTACAGAAAC



TCCCAGACTGAAATCCCAAG	TCGAATTTGATTCCCAGAGG
CTGGCCCTACGAATTTGTTG	GTCAGGAGATCGAGACAATC
ACCACAGCTGGCTAATTTTG	CTTGAGTCTGGGAGTTTGAG
TCCAGTAGTGCTGTTGAGAG	ACTAGGCACCTCATTCTCAG
AGTGTAAGCTCCCCTGTTTC	GGCATTTCGTTGGTGATTATC
ATGGTGGCATTGACTGTGAC	CCTCATGAAGGGATATGTGC
AGTGACACAGGGAAAACACG	TTCTTTGGCTGAGAAAGGAC
GTGATGTGACAGCCTCAAAG	GTAGAGTAGAGTGACACTCC
CTTTTAGAGCGTAGACAAAC	GAGTCTTTCCTTCTTGTCG
GAAAGACACGGGATTGGAAC	TGGAGGCAGAAAGAAGTGTC
AGATGGCTCTAGAGAATCCC	GCATTTTCCACTGGTATCCC
GCTACTGGGTCTAGCCTTAT	ATCAACAAACAAGGGGCTTC
TGGCTGGACAGCTTTCATAC	ACCCCAATAGAAGTGATAG
TGGAATGTGTTGGCTTGGG	TCATACGTAGAGTATCGGTG
CCCTGAAGGAATAAATGGAC	AGGCCAAATAGTCTTTACTC
TACCAAGGCATTGGACTCTG	GTCATCTGCACTAGGCATTC
TGTGGACCCCAGGATATAAC	AATTCAGAGCCTGACCCAAG
GGGTAAAGCATCTGCCAATG	GCATGCCTCCCTTGCTGTC
TGGAACATAGAGAGCACCAG	GCTCTCTACTTTACCTTCCC
ATGGGTCCTGATTCTTGCAC	TCAAGACTGTGAGGTGGTTG
TGACAAAGTCGACTCAGTGC	TGGAAGTTACAGGCCTTGAG
GCACACACAAATCATCCAAG	GACCATAGTCTCAGTATGCC
CAATGTTCACTGCCCATTCC	GGAAGTTTGTATAATTGCTCC
AATGATGCTTCAGTCCCACC	TCAACCAAGGGATGAAAGCC
CCGCTACACTTTGTATGACC	GCTTTTGCATAACTGAGCAC
TGTAACACTCAGAGCAAAGG	GTCAGGTACATACTAAAATGC
GATACTATGCCACCAAAGGG	CTGAATGAAAGTTCAAACGTC
CTCTGGCTAGCATCACAATG	TACCTTATGGCAGGGACTTG.

wherein (i) the nucleic acid comprises maternal nucleic acid and fetal nucleic acid, (ii) each primer of each primer set hybridizes to Y-chromosome nucleic acid, and (iii) each primer set consists of a first primer and a corresponding second primer; and

5 (b) detecting the presence or absence of amplification product from each primer set, whereby the presence of Y-chromosome nucleic acid is based on detecting the presence of the amplification product.

10 A2. The method of embodiment A1, wherein the extracellular nucleic acid is deoxyribonucleic acid (DNA).

A3. The method of embodiment A1, wherein the extracellular nucleic acid is ribonucleic acid (RNA).

5 A4. The method of embodiment A1, wherein the extracellular nucleic acid is contacted with one or more primer sets selected from the group consisting of

**first primer**

AGATGGCTCTAGAGAATCCC  
 GATACTATGCCACCAAAGGG  
 CTCTGGCTAGCATCACAATG

**corresponding second primer**

GCATTTTCCACTGGTATCCC  
 CTGAATGAAAGTTCAAACGTC  
 TACCTTATGGCAGGGACTTG.

10 A5. The method of embodiment A1, wherein detecting the presence or absence of the amplification product comprises contacting under extension conditions each amplification product with an extension oligonucleotide selected from the group consisting of

CCAGAATGCGAAACTC  
 CCCATCTCTTCTCAATCC  
 AAACATGCTCACGATCAC  
 GCAGGGACTTGTCGCTAGG  
 AAAAAGCGCAGGTATTTTCTG  
 CAATTGGATCACATTACATCAAA  
 TGAAATCCCAAGACAATGGA  
 GGCCCTACGAATTTGTTGTATTTTT  
 GTCACATTTTGTGCCCT  
 GGACTCTCAATGAAAGCA  
 GGAAGAATGCCAGAGTCA  
 TAGCCATAGTGCACATCTCA  
 CTCCTCACTGATTTTAAAGTTT  
 TGTCACCAAGCACACTATTGCCAGG  
 TAGAGCGTAGACAAACTGGATAGACA  
 TGGAACAGTTGCTGTGCCACCT  
 CCAGAATGCGAAACTC  
 GACATGAAGTCATTTGCT  
 AATTCGAGTTACAGCCACCG  
 GTATCCTGAAGCCAATAAATAC  
 GAAGGAATAAATGGACTCTCGAT

CTGCACATGAGATACATATCTTCC  
 GTGGACCCCAGGATATAACAAATTA  
 CTGCCAATGAAATGTTAATTGCTGGGC  
 AGCAACCCAACCCTCTGC  
 GGCTCTTTCAGGAATGGA  
 TTGCACCTGTATCACACAGT  
 GTTCCCGAGAACTAGGGATT  
 CCATTGACTGCTCAAATTTACA  
 GCCAGAGCCACAGAGGGCATTTT  
 GTCTCTTTTTATGAGACA  
 GTAACACTCAGAGCAAAGGGGCACT  
 ACAGTACCTGTATTTCTTAG  
 TATGGCAGGGACTTGTCGCTAGG

and detecting extended extension oligonucleotide.

5 A6. The method of embodiment A5, wherein the extension oligonucleotide is selected from the group consisting of

CCAGAATGCGAAACTC  
 ACAGTACCTGTATTTCTTAG  
 TATGGCAGGGACTTGTCGCTAGG.

10 A7. The method of embodiment A5 or embodiment A6, wherein the extended extension oligonucleotide is detected by mass spectrometry.

A8. The method of any one of embodiments A1-A7, wherein the extracellular nucleic acid is from blood serum.

15 A9. The method of any one of embodiments A1-A7, wherein the extracellular nucleic acid is from blood plasma.

20 A10. The method of any one of embodiments A1-A9, which further comprises determining the sex of the fetus based on the presence or absence of Y-chromosome nucleic acid, whereby the presence of Y-chromosome nucleic acid determines the fetus is male.

A11. The method of any one of embodiments A1-A10, wherein one or both primers in a primer set include a universal tag sequence.

5 A12. The method of any one of embodiments A1-A11, which further comprises contacting the extracellular nucleic acid with a control primer set under amplification conditions and detecting the presence or absence of a control amplification product from the control primer set, wherein the control primers of the control primer set can hybridize to the maternal nucleic acid.

10 A13. The method of embodiment A12, wherein the control primers of the control primer set hybridize to a nucleotide sequence that encodes or controls the expression of an albumin protein.

A14. The method of embodiment A13, wherein the control primer set comprises the following control primers:

<u>first primer</u>	<u>corresponding second primer</u>
CAGTATCTTCAGCAGTGTCC	GCAAATTCAGTACTTCATTC.

15 A15. The method of any one of embodiments A12-A14, wherein detecting the presence or absence of the control amplification product comprises contacting under extension conditions each control amplification product with a control extension oligonucleotide and detecting extended control extension oligonucleotide.

20 A16. The method of embodiment A15, wherein the control extension oligonucleotide has the nucleotide sequence GCAGTGTCCATTTGAAGAT.

A17. The method of any one of embodiments A1-A16, wherein the extracellular nucleic acid is contacted with two or more primer sets in a multiplex reaction.

25 A18. The method of any one of embodiments A1-A3, A5 and A7-A17, wherein the extracellular nucleic acid is contacted with the following primer sets

<u>first primer</u>	<u>corresponding second primer</u>
AGATGGCTCTAGAGAATCCC	GCATTTTCCACTGGTATCCC
TTTCCTCTCTCTCCACCCC	AAAGTTGCCTTTCTGCCTGC
AGAGAAGGCGGATTCTTTG	GAAGTTGGGAGAGTTACTCG
ACGAAGGGCTACTTCTCTAC	CTCTGGCTAGCATCACAATG
GAACCTCAGGCTCTTTGTCC	TCTATTCTTGCCGAGAGACC

ATGACCCCAAAAGCACAGAC  
 TCCAGACTGAAATCCAAG  
 CTGGCCCTACGAATTTGTTG

TCCCTCTGTGGTACAGAAAC  
 TCGAATTTGATTCCCAGAGG  
 GTCAGGAGATCGAGACAATC.

B1. A method for determining the presence or absence of Y-chromosome nucleic acid in a sample from a pregnant female, the method comprising amplifying nucleic acid from the sample from the pregnant female with at least one primer pair selected from Table 3, wherein each primer in the primer pair may comprise the entire sequence shown in the table or only the non-  
 5 underlined sequence-specific portion, wherein the presence of amplification products indicates the presence of Y-chromosome nucleic acid.

B2. A method for determining the presence or absence of Y-chromosome nucleic acid in a sample from a pregnant female, the method comprising:  
 10 a) contacting nucleic acid from a pregnant female with at least one primer pair selected from Table 3, wherein each primer of the primer pair may comprise the entire sequence shown in the table or only the non-underlined sequence-specific portion;  
 and b) contacting nucleic acid amplification products of step a) with at least one  
 15 corresponding extend primer from Table 3, wherein the presence of extend products indicates the presence of Y-chromosome nucleic acid.

B3. The method of embodiment B1 or embodiment B2, which is further limited by one or more applicable embodiments presented in any one of embodiments A1-A19.

20 C1. A method for determining the amount of fetal nucleic acid in a sample of extracellular nucleic acid from a pregnant female, which comprises:

(a) contacting under amplification conditions extracellular nucleic acid from a pregnant female with one or more primer sets selected from the group consisting of:

**first primer**

AGATGGCTCTAGAGAATCCC  
 TTTCTCTCTCTCCACCCC  
 AGAGAAGGCGGATTCTTTG  
 ACGAAGGGCTACTTCTCTAC  
 GAACCTCAGGCTCTTTGTCC  
 ATGACCCCAAAAGCACAGAC  
 TCCAGACTGAAATCCAAG

**corresponding second primer**

GCATTTTCCACTGGTATCCC  
 AAAGTTGCCTTTCTGCCTGC  
 GAAGTTGGGAGAGTTACTCG  
 CTCTGGCTAGCATCACAATG  
 TCTATTCTTGCCGAGAGACC  
 TCCCTCTGTGGTACAGAAAC  
 TCGAATTTGATTCCCAGAGG

CTGGCCCTACGAATTTGTTG	GTCAGGAGATCGAGACAATC
ACCACAGCTGGCTAATTTTG	CTTGAGTCTGGGAGTTTGAG
TCCAGTAGTGCTGTTGAGAG	ACTAGGCACCTCATTCTCAG
AGTGTAAGCTCCCCTGTTTC	GGCATTTCGTTGGTGATTATC
ATGGTGGCATTGACTGTGAC	CCTCATGAAGGGATATGTGC
AGTGACACAGGGAAAACACG	TTCTTTGGCTGAGAAAGGAC
GTGATGTGACAGCCTCAAAG	GTAGAGTAGAGTGACACTCC
CTTTTAGAGCGTAGACAAAC	GAGTCTTTCCTTCTTGTCG
GAAAGACACGGGATTGGAAC	TGGAGGCAGAAAGAAGTGTC
AGATGGCTCTAGAGAATCCC	GCATTTTCCACTGGTATCCC
GCTACTGGGTCTAGCCTTAT	ATCAACAAACAAGGGGCTTC
TGGCTGGACAGCTTTCATAC	ACCCCAATAGAAGTGATAG
TGGAAATGTGTTGGCTTGGG	TCATACGTAGAGTATCGGTG
CCCTGAAGGAATAAATGGAC	AGGCCAAATAGTCTTTACTC
TACCAAGGCATTGGACTCTG	GTCATCTGCACTAGGCATTG
TGTGGACCCCAGGATATAAC	AATTCAGAGCCTGACCCAAG
GGGTAAAGCATCTGCCAATG	GCATGCCTCCCTTGCTGTC
TGGAACATAGAGAGCACCAG	GCTCTCTACTTTACCTTCCC
ATGGGTCTGATTCTTGCAC	TCAAGACTGTGAGGTGGTTG
TGACAAAGTCGACTCAGTGC	TGGAAGTTACAGGCCTTGAG
GCACACACAAATCATCCAAG	GACCATAGTCTCAGTATGCC
CAATGTTCACTGCCATTCC	GGAAGTTTGTATAATTGCTCC
AATGATGCTTCAGTCCCACC	TCAACCAAGGGATGAAAGCC
CCGCTACACTTTGTATGACC	GCTTTTGCATAACTGAGCAC
TGTAACACTCAGAGCAAAGG	GTCAGGTACATACTAAAATGC
GATACTATGCCACCAAAGGG	CTGAATGAAAGTTCAAACGTC
CTCTGGCTAGCATCACAATG	TACCTTATGGCAGGGACTTG.

wherein (i) the nucleic acid comprises maternal nucleic acid and fetal nucleic acid, (ii) each primer of each primer set hybridizes to Y-chromosome nucleic acid, and (iii) each primer set consists of a first primer and a corresponding second primer; and

5 (b) determining the amount of amplification product from each primer set, whereby the amount of fetal nucleic acid in the extracellular nucleic acid is determined based on the amount of the amplification product.

10 C2. The method of embodiment C1, wherein the extracellular nucleic acid is deoxyribonucleic acid (DNA).

C3. The method of embodiment C1, wherein the extracellular nucleic acid is ribonucleic acid (RNA).

5 C4. The method of embodiment C1, wherein the extracellular nucleic acid is contacted with one or more primer sets selected from the group consisting of

**first primer**

**corresponding second primer**

AGATGGCTCTAGAGAATCCC  
GATACTATGCCACCAAAGGG  
CTCTGGCTAGCATCACAATG

GCATTTTCCACTGGTATCCC  
CTGAATGAAAGTTCAAACGTC  
TACCTTATGGCAGGGACTTG.

10 C5. The method of embodiment C1, wherein detecting the presence or absence of the amplification product comprises contacting under extension conditions each amplification product with an extension oligonucleotide selected from the group consisting of

CCAGAATGCGAAACTC  
CCCATCTCTTCTCAATCC  
AAACATGCTCACGATCAC  
GCAGGGACTTGTCGCTAGG  
AAAAAGCGCAGGTATTTTCTG  
CAATTGGATCACATTACATCAAA  
TGAAATCCCAAGACAATGGA  
GGCCCTACGAATTTGTTGTATTTTT  
GTCACATTTTGTGCCCT  
GGACTCTCAATGAAAGCA  
GGAAGAATGCCAGAGTCA  
TAGCCATAGTGACATCTCA  
CTCTTCACTGATTTTAAAGTTT  
TGTCACCAAGCACACTATTGCCAGG  
TAGAGCGTAGACAAACTGGATAGACA  
TGGAACAGTTGCTGTGCCACCT  
CCAGAATGCGAAACTC  
GACATGAAGTCATTTGCT  
AATTCGAGTTACAGCCACCG  
GTATCCTGAAGCCAATAAATAC  
GAAGGAATAAATGGACTCTCGAT  
CTGCACATGAGATACATATCTTCC

GTGGACCCCAGGATATAACAAATTA  
 CTGCCAATGAAATGTTAATTGCTGGGC  
 AGCAACCCAACCCTCTGC  
 GGCTCTTTCAGGAATGGA  
 TTGCACCTGTATCACACAGT  
 GTTCCCGAGAACTAGGGATT  
 CCATTGACTGCTCAAATTTACA  
 GCCAGAGCCACAGAGGGCATTTT  
 GTCTCTTTTATGAGCACA  
 GTAACACTCAGAGCAAAGGGGCACT  
 ACAGTACCTGTATTTCCCTTAG  
 TATGGCAGGGACTTGTCGCTAGG

and detecting extended extension oligonucleotide.

5 C6. The method of embodiment C5, wherein the extension oligonucleotide is selected from the group consisting of

CCAGAATGCGAAACTC  
 ACAGTACCTGTATTTCCCTTAG  
 TATGGCAGGGACTTGTCGCTAGG.

10 C7. The method of embodiment C5 or embodiment C6, wherein the extended extension oligonucleotide is detected by mass spectrometry.

C8. The method of any one of embodiments C1-C7, wherein the extracellular nucleic acid is from blood serum.

15 C9. The method of any one of embodiments C1-C7, wherein the extracellular nucleic acid is from blood plasma.

C10. The method of any one of embodiments C1-C9, wherein the amount of the fetal nucleic acid is expressed as a percentage of fetal nucleic acid in the extracellular nucleic acid.

20 C11. The method of any one of embodiments C1-C10, wherein one or both primers in a primer set include a universal tag sequence.



5 C12. The method of any one of embodiments C1-C11, which further comprises contacting the extracellular nucleic acid with a control primer set under amplification conditions and detecting the presence or absence of a control amplification product from the control primer set, wherein the control primers of the control primer set can hybridize to the maternal nucleic acid.

C13. The method of embodiment C12, wherein the control primers of the control primer set hybridize to a nucleotide sequence that encodes or controls the expression of an albumin protein.

10 C14. The method of embodiment C13, wherein the control primer set comprises the following control primers:

<u>first primer</u>	<u>corresponding second primer</u>
CAGTATCTTCAGCAGTGCC	GCAAATTCAGTTACTTCATTC.

15 C15. The method of any one of embodiments C12-C14, wherein detecting the presence or absence of the control amplification product comprises contacting under extension conditions each control amplification product with a control extension oligonucleotide and detecting extended control extension oligonucleotide.

20 C16. The method of embodiment C15, wherein the control extension oligonucleotide has the nucleotide sequence GCAGTGTCCATTTGAAGAT.

C17. The method of any one of embodiments C1-C16, wherein the extracellular nucleic acid is contacted with two or more primer sets in a multiplex reaction.

25 C18. The method of any one of embodiments C1-C3, C5 and C7-C17, wherein the extracellular nucleic acid is contacted with the following primer sets

<u>first primer</u>	<u>corresponding second primer</u>
AGATGGCTCTAGAGAATCCC	GCATTTTCCACTGGTATCCC
TTTCCTCTCTCTCCACCCC	AAAGTTGCCTTTCTGCCTGC
AGAGAAGGCGGATTCCCTTG	GAAGTTGGGAGAGTTACTCG
ACGAAGGGCTACTTCTCTAC	CTCTGGCTAGCATCACAATG
GAACCTCAGGCTCTTTGTCC	TCTATTCTTGCCGAGAGACC
ATGACCCCAAAAGCACAGAC	TCCCTCTGTGGTACAGAAAC
TCCCAGACTGAAATCCCAAG	TCGAATTTGATTCCAGAGG

CTGGCCCTACGAATTTGTTG

GTCAGGAGATCGAGACAATC.

C19. The method of any one of embodiments C1-C18, wherein the fetal nucleic acid is from a male fetus.

5 D1. A kit that comprises one or more primer sets selected from the group consisting of:

**first primer**

AGATGGCTCTAGAGAATCCC  
 TTTCTCTCTCTCCACCCC  
 AGAGAAGGCGGATTCTTTG  
 ACGAAGGGCTACTTCTCTAC  
 GAACCTCAGGCTCTTTGTCC  
 ATGACCCCAAAAGCACAGAC  
 TCCCAGACTGAAATCCCAAG  
 CTGGCCCTACGAATTTGTTG  
 ACCACAGCTGGCTAATTTTG  
 TCCAGTAGTGCTGTTGAGAG  
 AGTGTAAAGCTCCCCTGTTTC  
 ATGGTGGCATTGACTGTGAC  
 AGTGACACAGGGAAAACACG  
 GTGATGTGACAGCCTCAAAG  
 CTTTtagAGCGTAGACAAAC  
 GAAAGACACGGGATTGGAAC  
 AGATGGCTCTAGAGAATCCC  
 GCTACTGGGTCTAGCCTTAT  
 TGGCTGGACAGCTTTCATAC  
 TGGAAATGTGTTGGCTTGGG  
 CCCTGAAGGAATAAATGGAC  
 TACCAAGGCATTGGACTCTG  
 TGTGGACCCCAGGATATAAC  
 GGGTAAAGCATCTGCCAATG  
 TGGACATAGAGAGCACCAG  
 ATGGGTCTGATTCTTGCAC  
 TGACAAAGTCGACTCAGTGC  
 GCACACACAAATCATCCAAG  
 CAATGTTCACTGCCATTCC  
 AATGATGCTTCAGTCCCACC

**corresponding second primer**

GCATTTTCCACTGGTATCCC  
 AAAGTTGCCTTTCTGCCTGC  
 GAAGTTGGGAGAGTTACTCG  
 CTCTGGCTAGCATCACAATG  
 TCTATTCTTGCCGAGAGACC  
 TCCCTCTGTGGTACAGAAAC  
 TCGAATTTGATTCCCAGAGG  
 GTCAGGAGATCGAGACAATC  
 CTTGAGTCTGGGAGTTTGAG  
 ACTAGGCACCTCATTCTCAG  
 GGCATTCGTTGGTGATTATC  
 CCTCATGAAGGGATATGTGC  
 TTCTTTGGCTGAGAAAGGAC  
 GTAGAGTAGAGTGACACTCC  
 GAGTCTTTCCTTCTGTGCG  
 TGGAGGCAGAAAGAAGTGTC  
 GCATTTTCCACTGGTATCCC  
 ATCAACAAACAAGGGGCTTC  
 ACCCCCAATAGAAGTGATAG  
 TCATACGTAGAGTATCGGTG  
 AGGCCAAATAGTCTTACTC  
 GTCATCTGCACTAGGCATTG  
 AATTCAGAGCCTGACCCAAG  
 GCATGCCTCCCTTGCTGTC  
 GCTCTCTACTTTACCTTCCC  
 TCAAGACTGTGAGGTGGTTG  
 TGGAAAGTTACAGGCCTTGAG  
 GACCATAGTCTCAGTATGCC  
 GGAAGTTTGTATAATTGCTCC  
 TCAACCAAGGGATGAAAGCC

CCGCTACACTTTGTATGACC	GCTTTTGCATAACTGAGCAC
TGTAACACTCAGAGCAAAGG	GTCAGGTACATACTAAAATGC
GATACTATGCCACCAAAGGG	CTGAATGAAAGTTCAAACGTC
CTCTGGCTAGCATCACAATG	TACCTTATGGCAGGGACTTG

wherein each primer set consists of a first primer and a corresponding second primer.

5 D2. The kit of embodiment D1, wherein the one or more primer sets are selected from the group consisting of

<u>first primer</u>	<u>corresponding second primer</u>
AGATGGCTCTAGAGAATCCC	GCATTTTCCACTGGTATCCC
GATACTATGCCACCAAAGGG	CTGAATGAAAGTTCAAACGTC
CTCTGGCTAGCATCACAATG	TACCTTATGGCAGGGACTTG.

D3. The kit of embodiment D1 or D2, which further comprises one or more extension oligonucleotides selected from the group consisting of:

CCAGAATGCGAAACTC  
 CCCATCTCTTCTCAATCC  
 AAACATGCTCACGATCAC  
 GCAGGGACTTGTCGCTAGG  
 AAAAAGCGCAGGTATTTTCTG  
 CAATTGGATCACATTACATCAAA  
 TGAATCCCAAGACAATGGA  
 GGCCCTACGAATTTGTTGTATTTTT  
 GTCACATTTTGTGGCCCT  
 GGAAGTCTCAATGAAAGCA  
 GGAAGAATGCCAGAGTCA  
 TAGCCATAGTGCACATCTCA  
 CTCTTCACTGATTTTAAAGTTT  
 TGTACCAAGCACACTATTGCCAGG  
 TAGAGCGTAGACAAACTGGATAGACA  
 TGGAACAGTTGCTGTGCCACCT  
 CCAGAATGCGAAACTC  
 GACATGAAGTCATTTGCT  
 AATTCGAGTTACAGCCACCG  
 GTATCCTGAAGCCAATAAATAC

GAAGGAATAAATGGACTCTCGAT  
 CTGCACATGAGATACATATCTTCC  
 GTGGACCCCAGGATATAACAAATTA  
 CTGCCAATGAAATGTTAATTGCTGGGC  
 AGCAACCCAACCCTCTGC  
 GGCTCTTTCAGGAATGGA  
 TTGCACCTGTATCACACAGT  
 GTTCCCGAGAACTAGGGATT  
 CCATTGACTGCTCAAATTTACA  
 GCCAGAGCCACAGAGGGCATTTT  
 GTCTCTTTTTATGAGCACA  
 GTAACACTCAGAGCAAAGGGGCACT  
 ACAGTACCTGTATTTCTTAG  
 TATGGCAGGGACTTGTCGCTAGG.

D4. The kit of embodiment D3, wherein the one or more extension oligonucleotides are selected from the group consisting of

5

CCAGAATGCGAAACTC  
 ACAGTACCTGTATTTCTTAG  
 TATGGCAGGGACTTGTCGCTAGG.

D5. The kit of any one of embodiments D1-D4, wherein one or both primers in a primer set include a universal tag sequence.

10 D6. The kit of any one of embodiments D1-D5, which further comprises one or more control primer sets, wherein the control primers of the one or more control primer sets can hybridize to maternal nucleic acid under hybridization conditions.

15 D7. The kit of embodiment D6, wherein control primers of the one or more control primer sets hybridize to a nucleotide sequence that encodes or controls the expression of an albumin protein.

D8. The kit of embodiment D7, wherein the one or more control primer sets comprises the following control primers:

first primer  
 CAGTATCTTCAGCAGTGTCC

corresponding second primer  
 GCAAATTCAGTTACTTCATTC.

D9. The kit of any one of embodiments D6-D8, which further comprises one or more control extension oligonucleotides.

5 D10. The kit of embodiment D9, wherein the one or more control extension oligonucleotides comprises an oligonucleotide having the nucleotide sequence GCAGTGTCCATTTGAAGAT.

D11. The kit of any one of embodiments D1, D3 and D5-D10, wherein the kit comprises the following primer sets

**first primer**

**corresponding second primer**

AGATGGCTCTAGAGAATCCC  
TTTCCTCTCTCTCCACCCC  
AGAGAAGGCGGATTCTTTG  
ACGAAGGGCTACTTCTCTAC  
GAACCTCAGGCTCTTTGTCC  
ATGACCCCAAAAGCACAGAC  
TCCAGACTGAAATCCCAAG  
CTGGCCCTACGAATTTGTTG

GCATTTTCCACTGGTATCCC  
AAAGTTGCCTTTCTGCCTGC  
GAAGTTGGGAGAGTTACTCG  
CTCTGGCTAGCATCACAATG  
TCTATTCTTGCCGAGAGACC  
TCCCTCTGTGGTACAGAAAC  
TCGAATTTGATTCCAGAGG  
GTCAGGAGATCGAGACAATC.

10

D12. The kit of any one of embodiments D1-D11, which further comprises an enzyme.

D13. The kit of embodiment D12, wherein the enzyme is a polymerase.

15 D14. The kit of any one of embodiments D1-D13, which further comprises amplification components.

D15. The kit of any one of embodiments D1-D14, which further comprises one or more components for extracting nucleic acid from blood serum or blood plasma.

20

D16. The kit of any one of embodiments D1-D15, which further comprises instructions, or directions for accessing instructions, for carrying out a method described herein using the kit.

25 D17. A system that comprises a mass spectrometer and a kit described in any one of embodiments D1-D16.

The entirety of each patent, patent application, publication and document referenced herein hereby is incorporated by reference. Citation of the above patents, patent applications, publications and documents is not an admission that any of the foregoing is pertinent prior art, nor does it constitute any admission as to the contents or date of these publications or documents.

5

Modifications may be made to the foregoing without departing from the basic aspects of the invention. Although the invention has been described in substantial detail with reference to one or more specific embodiments, those of ordinary skill in the art will recognize that changes may be made to the embodiments specifically disclosed in this application, yet these modifications and improvements are within the scope and spirit of the invention.

10

The invention illustratively described herein suitably may be practiced in the absence of any element(s) not specifically disclosed herein. Thus, for example, in each instance herein any of the terms "comprising," "consisting essentially of," and "consisting of" may be replaced with either of the other two terms. The terms and expressions which have been employed are used as terms of description and not of limitation, and use of such terms and expressions do not exclude any equivalents of the features shown and described or portions thereof, and various modifications are possible within the scope of the invention claimed. The term "a" or "an" can refer to one of or a plurality of the elements it modifies (e.g., "a primer" can mean one or more primers) unless it is contextually clear either one of the elements or more than one of the elements is described. The term "about" as used herein refers to a value sometimes within 10% of the underlying parameter (i.e., plus or minus 10%), a value sometimes within 5% of the underlying parameter (i.e., plus or minus 5%), a value sometimes within 2.5% of the underlying parameter (i.e., plus or minus 2.5%), or a value sometimes within 1% of the underlying parameter (i.e., plus or minus 1%), and sometimes refers to the parameter with no variation. For example, a length of "about 100 nucleotides" can include lengths between 90 nucleotides and 110 nucleotides. Further, when a listing of values is described herein (e.g., 50%, 60%, 70%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93% or 94%), the listing includes all intermediate values thereof (e.g., 62%, 67%). Thus, it should be understood that although the present invention has been specifically disclosed by representative embodiments and optional features, modification and variation of the concepts herein disclosed may be resorted to by those skilled in the art, and such modifications and variations are considered within the scope of this invention.

15

20

25

30

Embodiments of the invention are set forth in the claims that follow.

What is claimed is:

1. A method for identifying the presence or absence of Y-chromosome nucleic acid in extracellular nucleic acid from a pregnant female, which comprises:

(a) contacting under amplification conditions extracellular nucleic acid from a pregnant female with one or more primer sets selected from the group consisting of:

**first primer**

AGATGGCTCTAGAGAATCCC  
 TTTCTCTCTCTCCACCCC  
 AGAGAAGGCGGATTCTTTG  
 ACGAAGGGCTACTTCTCTAC  
 GAACCTCAGGCTCTTTGTCC  
 ATGACCCCAAAGCACAGAC  
 TCCCAGACTGAAATCCCAAG  
 CTGGCCCTACGAATTTGTTG  
 ACCACAGCTGGCTAATTTTG  
 TCCAGTAGTGCTGTTGAGAG  
 AGTGTAAAGCTCCCCTGTTTC  
 ATGGTGGCATTGACTGTGAC  
 AGTGACACAGGGAAAACACG  
 GTGATGTGACAGCCTCAAAG  
 CTTTTAGAGCGTAGACAAAC  
 GAAAGACACGGGATTGGAAC  
 AGATGGCTCTAGAGAATCCC  
 GCTACTGGGTCTAGCCTTAT  
 TGGCTGGACAGCTTTCATAC  
 TGGAAATGTGTTGGCTTGGG  
 CCCTGAAGGAATAAATGGAC  
 TACCAAGGCATTGGACTCTG  
 TGTGGACCCCAGGATATAAC  
 GGGTAAAGCATCTGCCAATG  
 TGGAACATAGAGAGCACCAG  
 ATGGGTCTGATTCTTGACAC  
 TGACAAAGTCGACTCAGTGC  
 GCACACACAAATCATCCAAG  
 CAATGTTCACTGCCATTCC  
 AATGATGCTTCAGTCCCACC

**corresponding second primer**

GCATTTTCCACTGGTATCCC  
 AAAGTTGCCTTTCTGCCTGC  
 GAAGTTGGGAGAGTTACTCG  
 CTCTGGCTAGCATCACAATG  
 TCTATTCTTGCCGAGAGACC  
 TCCCTCTGTGGTACAGAAAC  
 TCGAATTTGATTCCCAGAGG  
 GTCAGGAGATCGAGACAATC  
 CTTGAGTCTGGGAGTTTGAG  
 ACTAGGCACCTCATTCTCAG  
 GGCATTCGTTGGTGATTATC  
 CCTCATGAAGGGATATGTGC  
 TTCTTTGGCTGAGAAAGGAC  
 GTAGAGTAGAGTGACTCTCC  
 GAGTCTTTCCTTCTTGTCG  
 TGGAGGCAGAAAGAAGTGTC  
 GCATTTTCCACTGGTATCCC  
 ATCAACAAACAAGGGGCTTC  
 ACCCCCAATAGAAGTGATAG  
 TCATACGTAGAGTATCGGTG  
 AGGCCAAATAGTCTTACTC  
 GTCATCTGCACTAGGCATTC  
 AATTCAGAGCCTGACCCAAG  
 GCATGCCTCCCTTGCTGTC  
 GCTCTCTACTTTACCTTCCC  
 TCAAGACTGTGAGGTGGTTG  
 TGGAAGTTACAGGCCTTGAG  
 GACCATAGTCTCAGTATGCC  
 GGAAGTTTGATAATTGCTCC  
 TCAACCAAGGGATGAAAGCC

CCGCTACACTTTGTATGACC  
 TGTAACACTCAGAGCAAAGG  
 GATACTATGCCACCAAAGGG  
 CTCTGGCTAGCATCACAATG

GCTTTTGCATAACTGAGCAC  
 GTCAGGTACATACTAAAATGC  
 CTGAATGAAAGTTCAAACGTC  
 TACCTTATGGCAGGGACTTG.

wherein (i) the nucleic acid comprises maternal nucleic acid and fetal nucleic acid, (ii) each primer of each primer set hybridizes to Y-chromosome nucleic acid, and (iii) each primer set consists of a first primer and a corresponding second primer; and

(b) detecting the presence or absence of amplification product from each primer set, whereby the presence of Y-chromosome nucleic acid is based on detecting the presence of the amplification product.

2. The method of claim 1, wherein the extracellular nucleic acid is deoxyribonucleic acid (DNA).
3. The method of claim 1, wherein the extracellular nucleic acid is ribonucleic acid (RNA).
4. The method of claim 1, wherein the extracellular nucleic acid is contacted with one or more primer sets selected from the group consisting of

**first primer**

AGATGGCTCTAGAGAATCCC  
 GATACTATGCCACCAAAGGG  
 CTCTGGCTAGCATCACAATG

**corresponding second primer**

GCATTTTCCACTGGTATCCC  
 CTGAATGAAAGTTCAAACGTC  
 TACCTTATGGCAGGGACTTG.

5. The method of claim 1, wherein detecting the presence or absence of the amplification product comprises contacting under extension conditions each amplification product with an extension oligonucleotide selected from the group consisting of

CCAGAATGCGAAACTC  
 CCCATCTCTTCTCAATCC  
 AAACATGCTCACGATCAC  
 GCAGGGACTTGTCGCTAGG  
 AAAAAAGCGCAGGTATTTTCTG  
 CAATTGGATCACATTACATCAAA  
 TGAATCCCAAGACAATGGA  
 GGCCCTACGAATTTGTTGATTTTT



GTCACATTTTGTGGCCCT  
 GGAAGAAATGCCAGAGTCA  
 TAGCCATAGTGCACATCTCA  
 CTCTTCACTGATTTTAAAGTTT  
 TGTCACCAAGCACACTATTGCCAGG  
 TAGAGCGTAGACAAACTGGATAGACA  
 TGGAACAGTTGCTGTGCCACCT  
 CCAGAATGCGAAACTC  
 GACATGAAGTCATTTGCT  
 AATTCGAGTTACAGCCACCG  
 GTATCCTGAAGCCAATAAATAC  
 GAAGGAATAAATGGACTCTCGAT  
 CTGCACATGAGATACATATCTTCC  
 GTGGACCCCAGGATATAACAAATTA  
 CTGCCAATGAAATGTTAATTGCTGGGC  
 AGCAACCCAACCCTCTGC  
 GGCTCTTTCAGGAATGGA  
 TTGCACCTGTATCACACAGT  
 GTTCCCGAGAAACTAGGGATT  
 CCATTGACTGCTCAAATTTACA  
 GCCAGAGCCACAGAGGGCATT  
 GTCTCTTTTATGAGCACA  
 GTAACACTCAGAGCAAAGGGGCACT  
 ACAGTACCTGTATTTCTTAG  
 TATGGCAGGGACTTGTGCTAGG

and detecting extended extension oligonucleotide.

6. The method of claim 5, wherein the extension oligonucleotide is selected from the group consisting of

CCAGAATGCGAAACTC  
 ACAGTACCTGTATTTCTTAG  
 TATGGCAGGGACTTGTGCTAGG.

7. The method of claim 5 or claim 6, wherein the extended extension oligonucleotide is detected by mass spectrometry.

8. The method of any one of claims 1-7, wherein the extracellular nucleic acid is from blood serum.
9. The method of any one of claims 1-7, wherein the extracellular nucleic acid is from blood plasma.
10. The method of any one of claims 1-9, which further comprises determining the sex of the fetus based on the presence or absence of Y-chromosome nucleic acid, whereby the presence of Y-chromosome nucleic acid determines the fetus is male.
11. The method of any one of claims 1-10, wherein one or both primers in a primer set include a universal tag sequence.
12. The method of any one of claims 1-11, which further comprises contacting the extracellular nucleic acid with a control primer set under amplification conditions and detecting the presence or absence of a control amplification product from the control primer set, wherein the control primers of the control primer set can hybridize to the maternal nucleic acid.
13. The method of claim 12, wherein the control primers of the control primer set hybridize to a nucleotide sequence that encodes or controls the expression of an albumin protein.
14. The method of claim 13, wherein the control primer set comprises the following control primers:
- |                     |                                    |
|---------------------|------------------------------------|
| <u>first primer</u> | <u>corresponding second primer</u> |
| CAGTATCTTCAGCAGTGTC | GCAAATTCAGTACTTCATTC.              |
15. The method of any one of claims 12-14, wherein detecting the presence or absence of the control amplification product comprises contacting under extension conditions each control amplification product with a control extension oligonucleotide and detecting extended control extension oligonucleotide.
16. The method of claim 15, wherein the control extension oligonucleotide has the nucleotide sequence GCAGTGTCCATTTGAAGAT.

17. The method of any one of claims 1-16, wherein the extracellular nucleic acid is contacted with two or more primer sets in a multiplex reaction.

18. The method of any one of claims 1-3, 5 and 7-17, wherein the extracellular nucleic acid is contacted with the following primer sets

**first primer**

AGATGGCTCTAGAGAATCCC  
 TTTCTCTCTCTCCACCCC  
 AGAGAAGGCGGATTCTTTG  
 ACGAAGGGCTACTTCTCTAC  
 GAACCTCAGGCTCTTTGTCC  
 ATGACCCCAAAGCACAGAC  
 TCCCAGACTGAAATCCCAAG  
 CTGGCCCTACGAATTTGTTG

**corresponding second primer**

GCATTTTCCACTGGTATCCC  
 AAAGTTGCCTTTCTGCCTGC  
 GAAGTTGGGAGAGTTACTCG  
 CTCTGGCTAGCATCACAATG  
 TCTATTCTTGCCGAGAGACC  
 TCCCTCTGTGGTACAGAAAC  
 TCGAATTTGATTCCCAGAGG  
 GTCAGGAGATCGAGACAATC.

19. A method for determining the presence or absence of Y-chromosome nucleic acid in a sample from a pregnant female, the method comprising amplifying nucleic acid from the sample from the pregnant female with at least one primer pair selected from Table 3, wherein each primer in the primer pair may comprise the entire sequence shown in the table or only the non-underlined sequence-specific portion, wherein the presence of amplification products indicates the presence of Y-chromosome nucleic acid.

20. A method for determining the presence or absence of Y-chromosome nucleic acid in a sample from a pregnant female, the method comprising:

a) contacting nucleic acid from a pregnant female with at least one primer pair selected from Table 3, wherein each primer of the primer pair may comprise the entire sequence shown in the table or only the non-underlined sequence-specific portion;

and b) contacting nucleic acid amplification products of step a) with at least one corresponding extend primer from Table 3, wherein the presence of extend products indicates the presence of Y-chromosome nucleic acid.

21. The method of claim 19 or claim 20, which is further limited by one or more applicable claims presented in any one of claims 1-18.

22. A method for determining the amount of fetal nucleic acid in a sample of extracellular nucleic acid from a pregnant female, which comprises:

(a) contacting under amplification conditions extracellular nucleic acid from a pregnant female with one or more primer sets selected from the group consisting of:

**first primer**

AGATGGCTCTAGAGAATCCC  
TTTCCTCTCTCCACCCC  
AGAGAAGGCGGATTCTTTG  
ACGAAGGGCTACTTCTCTAC  
GAACCTCAGGCTCTTTGTCC  
ATGACCCCAAAGCACAGAC  
TCCCAGACTGAAATCCCAAG  
CTGGCCCTACGAATTTGTTG  
ACCACAGCTGGCTAATTTTG  
TCCAGTAGTGCTGTTGAGAG  
AGTGTAAAGCTCCCCTGTTTC  
ATGGTGGCATTGACTGTGAC  
AGTGACACAGGGAAAACACG  
GTGATGTGACAGCCTCAAAG  
CTTTTAGAGCGTAGACAAAC  
GAAAGACACGGGATTGGAAC  
AGATGGCTCTAGAGAATCCC  
GCTACTGGGTCTAGCCTTAT  
TGGCTGGACAGCTTTCATAC  
TGGAAATGTGTTGGCTTGGG  
CCCTGAAGGAATAAATGGAC  
TACCAAGGCATTGGACTCTG  
TGTGGACCCCAGGATATAAC  
GGGTAAAGCATCTGCCAATG  
TGGAACATAGAGAGCACCAG  
ATGGGTCTGATTCTTGACAC  
TGACAAAGTCGACTCAGTGC  
GCACACACAAATCATCCAAG  
CAATGTTCACTGCCATTCC  
AATGATGCTTCAGTCCCACC  
CCGCTACACTTTGTATGACC  
TGTAACACTCAGAGCAAAGG

**corresponding second primer**

GCATTTTCCACTGGTATCCC  
AAAGTTGCCTTTCTGCCTGC  
GAAGTTGGGAGAGTTACTCG  
CTCTGGCTAGCATCACAATG  
TCTATTCTTGCCGAGAGACC  
TCCCTCTGTGGTACAGAAAC  
TCGAATTTGATTCCCAGAGG  
GTCAGGAGATCGAGACAATC  
CTTGAGTCTGGGAGTTTGAG  
ACTAGGCACCTCATTCTCAG  
GGCATTTCGTTGGTGATTATC  
CCTCATGAAGGGATATGTGC  
TTCTTTGGCTGAGAAAGGAC  
GTAGAGTAGAGTGACACTCC  
GAGTCTTTCCTTCTGTGCG  
TGGAGGCAGAAAGAAGTGTC  
GCATTTTCCACTGGTATCCC  
ATCAACAAACAAGGGGCTTC  
ACCCCAATAGAAGTGATAG  
TCATACGTAGAGTATCGGTG  
AGGCCAAATAGTCTTACTC  
GTCATCTGCACTAGGCATTC  
AATTCAGAGCCTGACCCAAG  
GCATGCCTCCCTTGCTGTC  
GCTCTCTACTTTACCTTCCC  
TCAAGACTGTGAGGTGGTTG  
TGGAAGTTACAGGCCTTGAG  
GACCATAGTCTCAGTATGCC  
GGAAGTTTGTATAATTGCTCC  
TCAACCAAGGGATGAAAGCC  
GCTTTTGCATAACTGAGCAC  
GTCAGGTACATACTAAAATGC

GATACTATGCCACCAAAGGG  
CTCTGGCTAGCATCACAATG

CTGAATGAAAGTTCAAACGTC  
TACCTTATGGCAGGGACTTG.

wherein (i) the nucleic acid comprises maternal nucleic acid and fetal nucleic acid, (ii) each primer of each primer set hybridizes to Y-chromosome nucleic acid, and (iii) each primer set consists of a first primer and a corresponding second primer; and

(b) determining the amount of amplification product from each primer set, whereby the amount of fetal nucleic acid in the extracellular nucleic acid is determined based on the amount of the amplification product.

23. The method of claim 22, wherein the extracellular nucleic acid is deoxyribonucleic acid (DNA).

24. The method of claim 22, wherein the extracellular nucleic acid is ribonucleic acid (RNA).

25. The method of claim 22, wherein the extracellular nucleic acid is contacted with one or more primer sets selected from the group consisting of

**first primer**

AGATGGCTCTAGAGAATCCC  
GATACTATGCCACCAAAGGG  
CTCTGGCTAGCATCACAATG

**corresponding second primer**

GCATTTTCCACTGGTATCCC  
CTGAATGAAAGTTCAAACGTC  
TACCTTATGGCAGGGACTTG.

26. The method of claim 22, wherein detecting the presence or absence of the amplification product comprises contacting under extension conditions each amplification product with an extension oligonucleotide selected from the group consisting of

CCAGAATGCGAAACTC  
CCCATCTTTCTCAATCC  
AAACATGCTCACGATCAC  
GCAGGGACTTGTGCTAGG  
AAAAAGCGCAGGTATTTTCTG  
CAATTGGATCACATTACATCAA  
TGAAATCCCAAGACAATGGA  
GGCCCTACGAATTTGTTGTATTTT  
GTCACATTTTGTGCCCT

GGACTCTCAATGAAAGCA  
GGAAGAATGCCAGAGTCA  
TAGCCATAGTGCACATCTCA  
CTCTTCACTGATTTTAAAGTTT  
TGTCACCAAGCACACTATTGCCAGG  
TAGAGCGTAGACAAACTGGATAGACA  
TGGAACAGTTGCTGTGCCACCT  
CCAGAATGCGAAACTC  
GACATGAAGTCATTTGCT  
AATTCGAGTTACAGCCACCG  
GTATCCTGAAGCCAATAAATAC  
GAAGGAATAAATGGACTCTCGAT  
CTGCACATGAGATACATATCTTCC  
GTGGACCCAGGATATAACAAATTA  
CTGCCAATGAAATGTTAATTGCTGGGC  
AGCAACCCAACCCTCTGC  
GGCTCTTTCAGGAATGGA  
TTGCACCTGTATCACACAGT  
GTTCCCGAGAACTAGGGATT  
CCATTGACTGCTCAAATTTACA  
GCCAGAGCCACAGAGGGCATTTT  
GTCTCTTTTATGAGCACA  
GTAACACTCAGAGCAAAGGGGCACT  
ACAGTACCTGTATTTCTTAG  
TATGGCAGGGACTTGTCGCTAGG

and detecting extended extension oligonucleotide.

27. The method of claim 26, wherein the extension oligonucleotide is selected from the group consisting of

CCAGAATGCGAAACTC  
ACAGTACCTGTATTTCTTAG  
TATGGCAGGGACTTGTCGCTAGG.

28. The method of claim 26 or claim 27, wherein the extended extension oligonucleotide is detected by mass spectrometry.

29. The method of any one of claims 22-28, wherein the extracellular nucleic acid is from blood serum.

30. The method of any one of claims 22-28, wherein the extracellular nucleic acid is from blood plasma.

31. The method of any one of claims 22-30, wherein the amount of the fetal nucleic acid is expressed as a percentage of fetal nucleic acid in the extracellular nucleic acid.

32. The method of any one of claims 22-31, wherein one or both primers in a primer set include a universal tag sequence.

33. The method of any one of claims 22-32, which further comprises contacting the extracellular nucleic acid with a control primer set under amplification conditions and detecting the presence or absence of a control amplification product from the control primer set, wherein the control primers of the control primer set can hybridize to the maternal nucleic acid.

34. The method of claim 33, wherein the control primers of the control primer set hybridize to a nucleotide sequence that encodes or controls the expression of an albumin protein.

35. The method of claim 34, wherein the control primer set comprises the following control primers:

first primer

CAGTATCTTCAGCAGTGTC

corresponding second primer

GCAAATTCAGTTACTTCATTC.

36. The method of any one of claims 22-35, wherein detecting the presence or absence of the control amplification product comprises contacting under extension conditions each control amplification product with a control extension oligonucleotide and detecting extended control extension oligonucleotide.

37. The method of claim 36, wherein the control extension oligonucleotide has the nucleotide sequence GCAGTGTCCATTTGAAGAT.

38. The method of any one of claims 22-37, wherein the extracellular nucleic acid is contacted with two or more primer sets in a multiplex reaction.

39. The method of any one of claims 22-24, 26 and 28-38, wherein the extracellular nucleic acid is contacted with the following primer sets

**first primer**

AGATGGCTCTAGAGAATCCC  
 TTTCTCTCTCTCCACCCC  
 AGAGAAGGCGGATTCTTTG  
 ACGAAGGGCTACTTCTCTAC  
 GAACCTCAGGCTCTTTGTCC  
 ATGACCCCAAAAGCACAGAC  
 TCCCAGACTGAAATCCCAAG  
 CTGGCCCTACGAATTTGTTG

**corresponding second primer**

GCATTTTCCACTGGTATCCC  
 AAAGTTGCCTTTCTGCCTGC  
 GAAGTTGGGAGAGTTACTCG  
 CTCTGGCTAGCATCACAATG  
 TCTATTCTTGCCGAGAGACC  
 TCCCTCTGTGGTACAGAAAC  
 TCGAATTTGATTCCCAGAGG  
 GTCAGGAGATCGAGACAATC.

40. A kit that comprises one or more primer sets selected from the group consisting of:

**first primer**

AGATGGCTCTAGAGAATCCC  
 TTTCTCTCTCTCCACCCC  
 AGAGAAGGCGGATTCTTTG  
 ACGAAGGGCTACTTCTCTAC  
 GAACCTCAGGCTCTTTGTCC  
 ATGACCCCAAAAGCACAGAC  
 TCCCAGACTGAAATCCCAAG  
 CTGGCCCTACGAATTTGTTG  
 ACCACAGCTGGCTAATTTTG  
 TCCAGTAGTGCTGTTGAGAG  
 AGTGAAGCTCCCCTGTTTC  
 ATGGTGGCATTGACTGTGAC  
 AGTGACACAGGGAAAACACG  
 GTGATGTGACAGCCTCAAAG  
 CTTTTAGAGCGTAGACAAAC  
 GAAAGACACGGGATTGGAAC  
 AGATGGCTCTAGAGAATCCC  
 GCTACTGGGTCTAGCCTTAT  
 TGGCTGGACAGCTTTCATAC  
 TGGAAATGTGTTGGCTTGGG  
 CCCTGAAGGAATAAATGGAC

**corresponding second primer**

GCATTTTCCACTGGTATCCC  
 AAAGTTGCCTTTCTGCCTGC  
 GAAGTTGGGAGAGTTACTCG  
 CTCTGGCTAGCATCACAATG  
 TCTATTCTTGCCGAGAGACC  
 TCCCTCTGTGGTACAGAAAC  
 TCGAATTTGATTCCCAGAGG  
 GTCAGGAGATCGAGACAATC  
 CTTGAGTCTGGGAGTTTGTG  
 ACTAGGCACCTCATTCTCAG  
 GGCATTCGTTGGTGATTATC  
 CCTCATGAAGGGATATGTGC  
 TTCTTTGGCTGAGAAAGGAC  
 GTAGAGTAGAGTGACTCTCC  
 GAGTCTTTCCTTCTTGCGG  
 TGGAGGCAGAAAGAAGTGTC  
 GCATTTTCCACTGGTATCCC  
 ATCAACAACAAGGGGCTTC  
 ACCCCCAATAGAAGTGATAG  
 TCATACGTAGAGTATCGGTG  
 AGGCCAAATAGCTTTACTC



TACCAAGGCATTGGACTCTG	GTCATCTGCACTAGGCATTC
TGTGGACCCCAGGATATAAC	AATTCAGAGCCTGACCCAAG
GGGTAAAGCATCTGCCAATG	GCATGCCTCCCTTGCTGTC
TGGAACATAGAGAGCACCAG	GCTCTCTACTTTACCTTCCC
ATGGGTCTTGATTCTTGAC	TCAAGACTGTGAGGTGGTTG
TGACAAAGTCGACTCAGTGC	TGGAAGTTACAGGCCTTGAG
GCACACACAAATCATCCAAG	GACCATAGTCTCAGTATGCC
CAATGTTCACTGCCATTCC	GGAAGTTTGTATAATTGCTCC
AATGATGCTTCAGTCCCACC	TCAACCAAGGGATGAAAGCC
CCGCTACACTTTGTATGACC	GCTTTTGCATAACTGAGCAC
TGTAACACTCAGAGCAAAGG	GTCAGGTACATACTAAAATGC
GATACTATGCCACCAAAGGG	CTGAATGAAAGTTCAAACGTC
CTCTGGCTAGCATCACAATG	TACCTTATGGCAGGGACTTG

wherein each primer set consists of a first primer and a corresponding second primer.

41. The kit of claim 40, wherein the one or more primer sets are selected from the group consisting of

**first primer**

AGATGGCTCTAGAGAATCCC  
 GATACTATGCCACCAAAGGG  
 CTCTGGCTAGCATCACAATG

**corresponding second primer**

GCATTTTCCACTGGTATCCC  
 CTGAATGAAAGTTCAAACGTC  
 TACCTTATGGCAGGGACTTG.

42. The kit of claim 40 or 41, which further comprises one or more extension oligonucleotides selected from the group consisting of:

CCAGAATGCGAAACTC  
 CCCATCTCTTCTCAATCC  
 AAACATGCTCACGATCAC  
 GCAGGGACTTGTCGCTAGG  
 AAAAAAGCGCAGGTATTTTCTG  
 CAATTGGATCACATTACATCAA  
 TGAAATCCCAAGACAATGGA  
 GGCCCTACGAATTTGTTGATTTTT  
 GTCACATTTTGTGCCCCT  
 GGACTCTCAATGAAAGCA  
 GGAAGAATGCCAGAGTCA

TAGCCATAGTGCACATCTCA  
 CTCTTCACTGATTTTAAAGTTT  
 TGTCACCAAGCACACTATTGCCAGG  
 TAGAGCGTAGACAACTGGATAGACA  
 TGGAACAGTTGCTGTGCCACCT  
 CCAGAATGCGAAACTC  
 GACATGAAGTCATTTGCT  
 AATTCGAGTTACAGCCACCG  
 GTATCCTGAAGCCAATAAATAC  
 GAAGGAATAAATGGACTCTCGAT  
 CTGCACATGAGATACATATCTTCC  
 GTGGACCCCAGGATATAACAAATTA  
 CTGCCAATGAAATGTTAATTGCTGGGC  
 AGCAACCCAACCCTCTGC  
 GGCTCTTTCAGGAATGGA  
 TTGCACCTGTATCACACAGT  
 GTTCCCAGAACTAGGGATT  
 CCATTGACTGCTCAAATTTACA  
 GCCAGAGCCACAGAGGGCATTTT  
 GTCTCTTTTTATGAGCACA  
 GTAACACTCAGAGCAAAGGGGCACT  
 ACAGTACCTGTATTTCCCTTAG  
 TATGGCAGGGACTTGTCGCTAGG.

43. The kit of claim 42, wherein the one or more extension oligonucleotides are selected from the group consisting of

CCAGAATGCGAAACTC  
 ACAGTACCTGTATTTCCCTTAG  
 TATGGCAGGGACTTGTCGCTAGG.

44. The kit of any one of claims 40-43, wherein one or both primers in a primer set include a universal tag sequence.

45. The kit of any one of claims 40-44, which further comprises one or more control primer sets, wherein the control primers of the one or more control primer sets can hybridize to maternal nucleic acid under hybridization conditions.

46. The kit of claim 45, wherein control primers of the one or more control primer sets hybridize to a nucleotide sequence that encodes or controls the expression of an albumin protein.

47. The kit of claim 46, wherein the one or more control primer sets comprises the following control primers:

<u>first primer</u>	<u>corresponding second primer</u>
CAGTATCTTCAGCAGTGTC	GCAAATTCAGTTACTTCATTC.

48. The kit of any one of claims 45-47, which further comprises one or more control extension oligonucleotides.

49. The kit of claim 48, wherein the one or more control extension oligonucleotides comprises an oligonucleotide having the nucleotide sequence GCAGTGTCCATTTGAAGAT.

50. The kit of any one of claims 40, 42 and 44-49, wherein the kit comprises the following primer sets

<b><u>first primer</u></b>	<b><u>corresponding second primer</u></b>
AGATGGCTCTAGAGAATCCC	GCATTTTCCACTGGTATCCC
TTTCCTCTCTCTCCACCCC	AAAGTTGCCTTTCTGCCTGC
AGAGAAGGCGGATTCTTTG	GAAGTTGGGAGAGTTACTCG
ACGAAGGGCTACTTCTCTAC	CTCTGGCTAGCATCACAATG
GAACCTCAGGCTCTTTGTCC	TCTATTCTTGCCGAGAGACC
ATGACCCCAAAAGCACAGAC	TCCCTCTGTGGTACAGAAAC
TCCCAGACTGAAATCCCAAG	TCGAATTTGATTCCCAGAGG
CTGGCCCTACGAATTTGTTG	GTCAGGAGATCGAGACAATC.

51. The kit of any one of claims 40-50, which further comprises an enzyme.

52. The kit of claim 51, wherein the enzyme is a polymerase.

53. The kit of any one of claims 40-52, which further comprises amplification components.

55. The kit of any one of claims 40-53, which further comprises components for extracting nucleic acid from blood serum or blood plasma.

56. The kit of any one of claims 40-55, which further comprises instructions, or directions for accessing instructions, for carrying out a method described herein using the kit.

57. A system that comprises a mass spectrometer and a kit described in any one of claims 40-56.

## FIGURE 1A

### CDY1-1

TATGGAGTACTTCAGGGTAGAATGGACAGATACTCAAAGCAGTATGAACAAACAAAGATTAAGGTAAAGATAAATACATGAACAA  
ATATGTAAGATAAAAAATTTGGCTGGGCTCAGTGGCTCACACCTGTAATCCCAGCATTTTGGGAGGCTGAGGTGCTTGATCACGA  
GGTCAGGAGATCGAGACAATCCTGGCTAACATGGTGAAACCCCGTCTCTAC [**T/\***]AAAAATACAACAAATTCGTAGGGCCAGGTG  
GCAGGTGCCTGTAGTCTCAGCTACTCGGGAGGCTGAGACAGGAGAAT

Amplicon length: 98 bp

Forward Primer: ACGTTGGATGGTCAGGAGATCGAGACAATC

Reverse Primer: ACGTTGGATGCTGGCCCTACGAATTTGTTG

Extend Primer: GGCCCTACGAATTTGTTGTATTTTT

Extend Direction: R

Extend Allele: T

## FIGURE 1B

### HSFY-1

CCTTTGCCTTTGTGTCCACAGAGGTTTCTGTATTCACCGTGCAGGTGCAAAACATACACCAGAGCAGAAAAGCAGTTTGTCTTC  
CTCTTTGTTTCTAAATATAGAGGTGCTTAAACAACATCCCTATTTGAGATTACTATTCATAAAAGAAAACAATCTCATCTTAGA  
GTAACCTTCACTTCTGTCTAGCAAAGAGAAAGTTGCCTTTCTGCCTGCAGAAATTATCAAACGCCTCTTTTGTATTTAATTTT  
ATGAGA [**A/\***]GGATTGAGAAGAGATGGGGGTGGAGAGAGAGGAAA

Amplicon length: 119 bp

Forward Primer: ACGTTGGATGAAAGTTGCCTTTCTGCCTGC

Reverse Primer: ACGTTGGATGTTTCCTCTCTCTCCACCCC

Extend Primer: CCCATCTCTTCTCAATCC

Extend Direction: R

Extend Allele: A

**FIGURE 1C**

RBM1A1-1

GGCCAGGCGAACCTCAGGCTCTTTGTCTACTAAAAAGCGCAGGTATTTTCTG [**T/\***] TTCTCTGGACAGCTGGGTCTCTCGGCAA  
 GAATAGAAAGCGAAGGTTTGGGATTTTGTCTATAAAAAGGGATGGGTTTTCTATGTGTGGGTGTTGAATTACGGGAGGAGTCAGTG  
 GGGAAAGAACTCCTCAGTGCTATTAAGAGACTCACTTTCGTAAACTCATTGATTTTCTGAGGATTCTACCTTTAACTGCCTAA  
 TGTGTCCGACTAGTTGTGGGAGATGGTGCTAAGCCGCCATTGGTTTT

Amplicon length: 101

Forward Primer: ACGTTGGATGGAACCTCAGGCTCTTTGTCC

Reverse Primer: ACGTTGGATGTCTATTCTTGCCGAGAGACC

Extend Primer: AAAAAGCGCAGGTATTTTCTG

Extend Direction: F

Extend Allele: T

**FIGURE 1D**

RBM2-1

CAATGCAGAAATAACATTTCAATTTTGGATTTGCAAAACAAGGATTGGTATGCAATAACTATTATTTTCAATGCTTGCTTTAATATC  
 TGCTCGAGTCTCCTTTTTCAGATCGACTCTCCCCACCATCTACTATAGATGCCACATAACTTGAGCTACCATATGCTTCACGAGGA  
 TCAGGGAGCACCTACCCAGAGAAGGCGGATTCCTTTGGTCTTTTCTGCAAACATGCTCACGATCAC [**A/\***] ATAATGAAAATCAC  
 CACAGCTCGAGTAACTCTCCCAACTTCTGCCATATCTATCTCGTGTA

Amplicon length: 111

Forward Primer: ACGTTGGATGAGAGAAGGCGGATTCCTTTG

Reverse Primer: ACGTTGGATGGAAGTTGGGAGAGTTACTCG

Extend Primer: AAACATGCTCACGATCAC

Extend Direction: F

Extend Allele: A

**FIGURE 1E**

TTY16-1

GCCACTCTTCTAGAAAGAACAGAAGAACCACAACCAAGAACAGAAATATACCAGTGTTTAATTGTCTTTCAGCCAATCCAAGGA  
 GAGACACTATCATTCATCTCTACCGGGCCTATTAAATTTACCTCGAATTTGATTCCCAGAGGAGTTGGTGCTTCACATCATCAGGG  
 GGAACCTT [C/\*]TCCATTGTCTTGGGATTTTCAGTCTGGGATAGAGACTTTGAACAGCAATAAGGTAATAAGGTCAGATAGGGGTGG  
 GGATACCCCTCTGGTGAGGGGTGGATGCCATGCTGTACCTTCACCTG

Amplicon length: 100

Forward Primer: ACGTTGGATGTCGAATTTGATTCCCAGAGG

Reverse Primer: ACGTTGGATGTCCCAGACTGAAATCCCAAG

Extend Primer: ggggTGAAATCCCAAGACAATGGA

Extend Direction: R

Extend Allele: C

**FIGURE 1F**

TTY22-1

ACCAATGCAATGCACTAGTCTCAGGGCACCAGGCCTGAGCTGTGAGCTCTGGCTAGCATCACAATGAATGCCACCATT [G/\*]CCT  
 AGCGACAAGTCCCTGCCATAAGGTAGAGAAGTAGCCCTTCGTGGAGGTGCACAGACCATGGATTCTTGCCTTCTTCTGTGTGGGT  
 CCACAGGATATTCCCATAGTTCTAGGAGAGGACAGACATTAGCCTGCCTAAAGAAAACATCAAGCAGAGCCCCAGGAATAAACTGTG  
 GAATTCCTAACGATTCAAAAAGATTTGCAGGATGCATCAGACCTGCC

Amplicon length: 98

Forward Primer: ACGTTGGATGCTCTGGCTAGCATCACAATG

Reverse Primer: ACGTTGGATGACGAAGGGCTACTTCTCTAC

Extend Primer: GCAGGGACTTGTGCTAGG

Extend Direction: R

Extend Allele: G

## FIGURE 1G

SRY\_4-i

ACGTCCAGGATAGAGTGAAGCGACCCATGAACGCATTCATCGTGTGGTCTCGCGATCAGAGGCGCAAGATGGCTCTAGAGAATCCC  
AGAATGCGAAACTC [A/\*] GAGATCAGCAAGCAGCTGGGATACCAGTGGAAAATGCTTACTGAAGCCGAAAAATGGCCATTCTTCC  
AGGAGGCACAGAAATTACAGGCCATGCACA

Amplicon length: 92

Forward Primer: ACGTTGGATGAGATGGCTCTAGAGAATCCC

Reverse Primer: ACGTTGGATGGCATTCTTCCACTGGTATCCC

Extend Primer: CCAGAATGCGAAACTC

Extend Direction: F

Extend Allele: A

## FIGURE 1H

XKRY-1

TATATCCTTTGAGTACGTATCTGGAAGTAGAGTAGCTGGATTATGTGGTAATTCCTATTTTTAATTTAACTGTATTTTGAACACT  
CAATTCTATGACCCCAAAGCACAGACTAGAAAAGCACAACAAAAACAATTGGATCACATTACATCAAA [C/\*] TAAAATGTTTC  
TGTACCACAGAGGGAAAAGGGTAAAAGCAAATGTTCACTGATAGTGTATATGCTATATTACAAGCACTTATTAATTAAGGTAT  
AGGTTCCAAAATGTACTAGGTGAAAAATGTATTATGGTTATTTTTTC

Amplicon length: 110

Forward Primer: ACGTTGGATGATGACCCCAAAGCACAGAC

Reverse Primer: ACGTTGGATGTCCCTCTGTGGTACAGAAAC

Extend Primer: CAATTGGATCACATTACATCAAA

Extend Direction: F

Extend Allele: C



## FIGURE 11

Alb-2-i

GCTCAGTATCTTCAGCAGTGTCCATTTGAAGAT [C/\*]ATGTAAAATTAGTGAATGAAGTAACTGAATTTGC

Amplicon length: 85bp

Forward Primer: ACGTTGGATGCAGTATCTTCAGCAGTGTCC

Reverse Primer: ACGTTGGATGGCAAATTCAGTACTTCATTC

Extend Primer: GCAGTGTCCATTTGAAGAT

Extend Direction: F

Extend Allele: C

FIGURE 2



