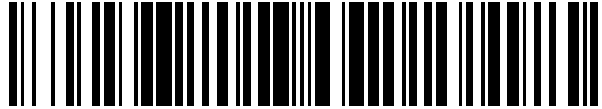


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 560 929**

51 Int. Cl.:

C12Q 1/68 (2006.01)

C12P 19/34 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **01.12.2010 E 10844163 (5)**

97 Fecha y número de publicación de la concesión europea: **09.09.2015 EP 2513339**

54 Título: **Métodos de determinación de la fracción de ácido nucleico fetal en muestras maternas**

30 Prioridad:

26.10.2010 US 455849 P

19.01.2010 US 296358 P

01.07.2010 US 360837 P

26.10.2010 US 407017 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:

23.02.2016

73 Titular/es:

VERINATA HEALTH, INC. (100.0%)

800 Saginaw Drive

Redwood City CA 94063, US

72 Inventor/es:

RAVA, RICHARD P.;

CHUU, YUE-JEN;

CHINNAPPA, MANJULA;

COMSTOCK, DAVID A.;

HEILEK, GABRIELLE y

HUNKAPILLER, MICHAEL

74 Agente/Representante:

IZQUIERDO BLANCO, María Alicia

ES 2 560 929 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

Métodos de determinación de la fracción de ácido nucleico fetal en muestras maternas**DESCRIPCIÓN****5 CAMPO DE LA INVENCION**

La presente invención se refiere a métodos y composiciones para detectar ácidos nucleicos fetales en una muestra materna y determinar la fracción de ácido nucleico fetal libre células que circula en una muestra materna.

10 ANTECEDENTES DE LA INVENCION

Las pruebas prenatales invasivas son potencialmente peligrosas para la madre y el feto. Por tanto, existe la necesidad del desarrollo de pruebas prenatales no invasivas. La sangre materna puede contener células fetales (véase, por ejemplo, Huang et al. (2008), Methods in Molecular Biology, 444:203-208). Mientras que las células fetales circulantes presentan una atractiva diana para diagnósticos prenatales no invasivos, particularmente para el diagnóstico del sexo fetal y anomalías cromosómicas por simple cariotipado, la escasez de células fetales intactas en la circulación materna (aproximadamente una célula por ml de sangre materna), la baja eficiencia del enriquecimiento (Bianchi et al., Am J Hum Genet 61:822-829 [1997]) y las dificultades con los análisis cromosómicos asociados a núcleos anormalmente densos en algunas células (Babochkina et al., Haematologica 90:740-745 [2005]) han favorecido la investigación de ADN libre de células.

El establecimiento de concentraciones de ADN fetal libre de células (ADNlc) en plasma materno en mujeres embarazadas sanas ha formado la plataforma sobre la que pueden estudiarse las anomalías del ADN fetal en trastornos asociados al embarazo. Se ha mostrado que el hallazgo de un aumento gradual en la concentración de ADN fetal en suero materno a medida que avanza la gestación precede a complicaciones asociadas al parto prematuro. También se ha encontrado un aumento de cinco veces en la concentración de ADN fetal en el suero obtenido de mujeres afectadas por preeclampsia. Otros trastornos del embarazo relacionados que se han asociado a un elevada concentración de ADNlc incluyen hiperémesis gravídica (graves náuseas del embarazo), placentación invasiva (en la que la placenta se pone en contacto con la circulación sanguínea materna), restricción del crecimiento intrauterino, hemorragia feto-materna y polihidramnios (Wright C.F. y Burton H., Human Reproduction Update 15(1):139-151 [2009]).

El análisis cuantitativo de ADN libre de células por estrategias de PCR en tiempo real también ha indicado que las concentraciones de ADN fetal circulatorio son elevadas en embarazos con aneuploidías fetales, más en particular trisomía 21 (Lo et al., Clin Chem 45:1747-1751 [1999]). Sin embargo, la fracción de ADN fetal en ADN de plasma sin células maternas se determina normalmente comparando la cantidad de locus específicos del feto (tales como el locus SRY en el cromosoma Y en embarazos masculinos) con la de un locus sobre un autosoma que es común a tanto la madre como al feto usando PCR cuantitativa en tiempo real (Dahllan et al., Lancet 369:474-481 [2007]; Li et al., Clin Chem 1002-1011 [2004]; Fan et al., Proc Natl Acad Sci 105:16266-16271 [2008]).

Así, existe la necesidad de métodos adicionales que permitirían la determinación de la fracción de ácido nucleico fetal en tanto embarazos masculinos como femeninos.

El método de la invención satisface la necesidad proporcionando los medios para determinar la fracción fetal que es independiente del sexo del feto. El método puede aplicarse para determinar simultáneamente la presencia o ausencia de una aneuploidía cromosómica u otra variación del número de copias, y puede usarse conjuntamente con métodos conocidos que se usan para determinar aneuploidías en muestra materna.

50 RESUMEN DE LA INVENCION

La invención se refiere a composiciones y métodos para determinar la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ácidos nucleicos fetales y maternos. La fracción de ácidos nucleicos fetales puede usarse en determinar la presencia o ausencia de aneuploidía fetal.

La invención proporciona un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna de sangre que comprende una mezcla de ADN fetal y genómico, en la que dicho ADN fetal y genómico es ADN libre de células (ADNlc), comprendiendo dicho método: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en dicha mezcla, en la que cada ácido nucleico diana polimórfico amplificado comprende al menos un sitio polimórfico, usando pares de cebadores cada uno capaz de amplificar una secuencia de ácidos nucleicos diana que comprende un sitio polimórfico en una reacción de PCR de múltiplex, para generar un panel de sitios polimórficos amplificados que contiene un número suficiente de sitios polimórficos tal que al menos 25 sean sitios polimórficos informativos; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que dicha secuenciación comprende proporcionar una pluralidad de marcas de secuencia para dicho panel de sitios polimórficos amplificados; y (c) usar dicha pluralidad de marcas de secuencia para: (i) identificar al menos un sitio polimórfico informativo en dicho panel de sitios polimórficos amplificados, en el

que dicho al menos un sitio polimórfico informativo se identifica por la diferencia en las secuencias alélicas en cada sitio polimórfico y el número de marcas de secuencia mapeadas en cada uno de los posibles alelos en cada sitio polimórfico, y (ii) determinar dicha fracción de ácidos nucleicos fetales usando el número total de marcas que se mapean en un primer alelo en un sitio polimórfico identificado y el número total de marcas que se mapean en un segundo alelo en el sitio polimórfico identificado informativo para calcular dicha fracción de ácidos nucleicos fetales.

En una realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende proporcionar una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción. La pluralidad de ácidos nucleicos polimórficos se localiza en una pluralidad de cromosomas diferentes. En una realización, la pluralidad de ácidos nucleicos polimórficos puede localizarse en los cromosomas 1-22. En otra realización, la pluralidad de ácidos nucleicos polimórficos puede localizarse en diferentes autosomas distintos de los cromosomas 13, 18 y 21. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

En una realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende una amplificación y proporciona una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

En otra realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende proporcionar una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción. La pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un único nucleótido (SNP). Alternativamente, la pluralidad de ácidos nucleicos diana polimórficos comprende al menos una repetición corta en tándem (STR). La pluralidad de ácidos nucleicos polimórficos se localiza en una pluralidad de cromosomas diferentes. Por ejemplo, la pluralidad de ácidos nucleicos polimórficos puede localizarse en una pluralidad de cromosomas diferentes distintos de los cromosomas 13, 18, 21, X o Y. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

En otra realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende proporcionar una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción. La pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un único nucleótido (SNP). Alternativamente, la pluralidad de ácidos nucleicos diana polimórficos comprende al menos una repetición corta en tándem (STR). La pluralidad de ácidos nucleicos polimórficos se localiza en una pluralidad de cromosomas diferentes. En una realización, la pluralidad de ácidos nucleicos polimórficos puede localizarse en los cromosomas 1-22. Por ejemplo, la pluralidad de ácidos nucleicos polimórficos puede localizarse en una pluralidad de cromosomas diferentes distintos de los cromosomas 13, 18, 21, X o Y. La secuenciación masivamente paralela es secuenciación por síntesis con terminadores de colorante reversible. Alternativamente, la secuenciación masivamente paralela es secuenciación por ligación o secuenciación de una molécula única. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

En otra realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende una amplificación y proporciona una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción. La pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un único nucleótido (SNP). Alternativamente, la pluralidad de ácidos nucleicos diana polimórficos comprende al menos una repetición corta en tándem (STR). La pluralidad de ácidos nucleicos polimórficos se localiza en una pluralidad de cromosomas diferentes. En una realización, la pluralidad de ácidos nucleicos polimórficos puede localizarse en los cromosomas 1-22. Por ejemplo, la pluralidad de ácidos nucleicos polimórficos puede localizarse en una pluralidad de cromosomas diferentes distintos de los cromosomas 13, 18, 21, X o Y. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva.

Preferentemente, la muestra materna es una muestra de plasma.

En otra realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende una amplificación y proporciona una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción. La pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un único nucleótido (SNP). Alternativamente, la pluralidad de ácidos nucleicos diana polimórficos comprende al menos una repetición corta en tándem (STR). La pluralidad de ácidos nucleicos polimórficos se localiza en una pluralidad de cromosomas diferentes. En una realización, la pluralidad de ácidos nucleicos polimórficos puede localizarse en los cromosomas 1-22. Por ejemplo, la pluralidad de ácidos nucleicos polimórficos puede localizarse en una pluralidad de cromosomas diferentes distintos de los cromosomas 13, 18, 21, X o Y. La secuenciación masivamente paralela es secuenciación por síntesis por terminadores de colorante reversible. Alternativamente, la secuenciación masivamente paralela es secuenciación por ligación o secuenciación de una molécula única. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

En una realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, por ejemplo, ADN libre de células, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende proporcionar una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción. La pluralidad de ácidos nucleicos polimórficos se localiza en una pluralidad de cromosomas diferentes. En una realización, la pluralidad de ácidos nucleicos polimórficos puede localizarse en los cromosomas 1-22. Por ejemplo, la pluralidad de ácidos nucleicos polimórficos puede localizarse en una pluralidad de cromosomas diferentes distintos de los cromosomas 13, 18, 21, X o Y. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

En una realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, por ejemplo, ADN libre de células, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende una amplificación y proporciona una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción. La pluralidad de ácidos nucleicos polimórficos se localiza en una pluralidad de cromosomas diferentes. En una realización, la pluralidad de ácidos nucleicos polimórficos puede localizarse en los cromosomas 1-22. Por ejemplo, la pluralidad de ácidos nucleicos polimórficos puede localizarse en una pluralidad de cromosomas diferentes distintos de los cromosomas 13, 18, 21, X o Y. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

En otra realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, por ejemplo, ADN libre de células, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende proporcionar una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción. La pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un único nucleótido (SNP). Alternativamente, la pluralidad de ácidos nucleicos diana polimórficos comprende al menos una repetición corta en tándem (STR). La pluralidad de ácidos nucleicos polimórficos se localiza en una pluralidad de cromosomas diferentes. En una realización, la pluralidad de ácidos nucleicos polimórficos puede localizarse en los cromosomas 1-22. Por ejemplo, la pluralidad de ácidos nucleicos polimórficos puede localizarse en una pluralidad de cromosomas diferentes distintos de los cromosomas 13, 18, 21, X o Y. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

En otra realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, por ejemplo, ADN libre de células, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende proporcionar una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción. La pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un único nucleótido (SNP). Alternativamente, la pluralidad de ácidos nucleicos diana polimórficos comprende al menos una repetición corta en tándem (STR). La pluralidad de ácidos nucleicos polimórficos se localiza en una pluralidad de cromosomas diferentes. En una realización, la pluralidad de ácidos nucleicos polimórficos puede localizarse en los cromosomas 1-

22. Por ejemplo, la pluralidad de ácidos nucleicos polimórficos puede localizarse en una pluralidad de cromosomas diferentes distintos de los cromosomas 13, 18, 21, X o Y. La secuenciación masivamente paralela es secuenciación por síntesis con terminadores de colorante reversible. Alternativamente, la secuenciación masivamente paralela es secuenciación por ligación o secuenciación de una molécula única. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

En otra realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, por ejemplo, ADN libre de células, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende una amplificación y proporciona una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción. La pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un único nucleótido (SNP). Alternativamente, la pluralidad de ácidos nucleicos diana polimórficos comprende al menos una repetición corta en tándem (STR). La pluralidad de ácidos nucleicos polimórficos se localiza en una pluralidad de cromosomas diferentes. En una realización, la pluralidad de ácidos nucleicos polimórficos puede localizarse en los cromosomas 1-22. Por ejemplo, la pluralidad de ácidos nucleicos polimórficos puede localizarse en una pluralidad de cromosomas diferentes distintos de los cromosomas 13, 18, 21, X o Y. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

En otra realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, por ejemplo, ADN libre de células, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende una amplificación y proporciona una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción. La pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un único nucleótido (SNP). Alternativamente, la pluralidad de ácidos nucleicos diana polimórficos comprende al menos una repetición corta en tándem (STR). La pluralidad de ácidos nucleicos polimórficos se localiza en una pluralidad de cromosomas diferentes. En una realización, la pluralidad de ácidos nucleicos polimórficos puede localizarse en los cromosomas 1-22. Por ejemplo, la pluralidad de ácidos nucleicos polimórficos puede localizarse en una pluralidad de cromosomas diferentes distintos de los cromosomas 13, 18, 21, X o Y. La secuenciación masivamente paralela es secuenciación por síntesis con terminadores de colorante reversible. Alternativamente, la secuenciación masivamente paralela es secuenciación por ligación o secuenciación de una molécula única. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

En una realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende proporcionar una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción, en la que determinar la fracción comprende determinar el número de marcas de secuencia fetales y maternas mapeadas en un genoma de referencia que comprende al menos un ácido nucleico polimórfico. La pluralidad de ácidos nucleicos polimórficos se localiza en una pluralidad de cromosomas diferentes. En una realización, la pluralidad de ácidos nucleicos polimórficos puede localizarse en los cromosomas 1-22. Por ejemplo, la pluralidad de ácidos nucleicos polimórficos puede localizarse en una pluralidad de cromosomas diferentes distintos de los cromosomas 13, 18, 21, X o Y. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

En una realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende una amplificación y proporciona una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción, en la que determinar la fracción comprende determinar el número de marcas de secuencia fetales y maternas mapeadas en un genoma de referencia que comprende al menos un ácido nucleico polimórfico. La pluralidad de ácidos nucleicos polimórficos se localiza en una pluralidad de cromosomas diferentes. En una realización, la pluralidad de ácidos nucleicos polimórficos puede localizarse en los cromosomas 1-22. Por ejemplo, la pluralidad de ácidos nucleicos polimórficos puede localizarse en una pluralidad de cromosomas diferentes distintos de los cromosomas 13, 18, 21, X o Y. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

En otra realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, en la que el método

comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende proporcionar una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción, en la que determinar la fracción comprende determinar el número de marcas de secuencia fetales y maternas mapeadas en un genoma de referencia que comprende al menos un ácido nucleico polimórfico. La pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un único nucleótido (SNP). Alternativamente, la pluralidad de ácidos nucleicos diana polimórficos comprende al menos una repetición corta en tándem (STR). La pluralidad de ácidos nucleicos polimórficos se localiza en una pluralidad de cromosomas diferentes. En una realización, la pluralidad de ácidos nucleicos polimórficos puede localizarse en los cromosomas 1-22. Por ejemplo, la pluralidad de ácidos nucleicos polimórficos puede localizarse en una pluralidad de cromosomas diferentes distintos de los cromosomas 13, 18, 21, X o Y. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

En otra realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende proporcionar una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción, en la que determinar la fracción comprende determinar el número de marcas de secuencia fetales y maternas mapeadas en un genoma de referencia que comprende al menos un ácido nucleico polimórfico. La pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un único nucleótido (SNP). Alternativamente, la pluralidad de ácidos nucleicos diana polimórficos comprende al menos una repetición corta en tándem (STR). La pluralidad de ácidos nucleicos polimórficos se localiza en una pluralidad de cromosomas diferentes. En una realización, la pluralidad de ácidos nucleicos polimórficos puede localizarse en los cromosomas 1-22. Por ejemplo, la pluralidad de ácidos nucleicos polimórficos puede localizarse en una pluralidad de cromosomas diferentes distintos de los cromosomas 13, 18, 21, X o Y. La secuenciación masivamente paralela es secuenciación por síntesis con terminadores de colorante reversible. Alternativamente, la secuenciación masivamente paralela es secuenciación por ligación o secuenciación de una molécula única. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

En otra realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende una amplificación y proporciona una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción, en la que determinar la fracción comprende determinar el número de marcas de secuencia fetales y maternas mapeadas en un genoma de referencia que comprende al menos un ácido nucleico polimórfico. La pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un único nucleótido (SNP). Alternativamente, la pluralidad de ácidos nucleicos diana polimórficos comprende al menos una repetición corta en tándem (STR). La pluralidad de ácidos nucleicos polimórficos se localiza en una pluralidad de cromosomas diferentes. En una realización, la pluralidad de ácidos nucleicos polimórficos puede localizarse en los cromosomas 1-22. Por ejemplo, la pluralidad de ácidos nucleicos polimórficos puede localizarse en una pluralidad de cromosomas diferentes distintos de los cromosomas 13, 18, 21, X o Y. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

En otra realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende una amplificación y proporciona una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción, en la que determinar la fracción comprende determinar el número de marcas de secuencia fetales y maternas mapeadas en un genoma de referencia que comprende al menos un ácido nucleico polimórfico. La pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un único nucleótido (SNP). Alternativamente, la pluralidad de ácidos nucleicos diana polimórficos comprende al menos una repetición corta en tándem (STR). La pluralidad de ácidos nucleicos polimórficos se localiza en una pluralidad de cromosomas diferentes. En una realización, la pluralidad de ácidos nucleicos polimórficos puede localizarse en los cromosomas 1-22. Por ejemplo, la pluralidad de ácidos nucleicos polimórficos puede localizarse en una pluralidad de cromosomas diferentes distintos de los cromosomas 13, 18, 21, X o Y. La secuenciación masivamente paralela es secuenciación por síntesis con terminadores de colorante reversible. Alternativamente, la secuenciación masivamente paralela es secuenciación por ligación o secuenciación de una molécula única. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

En una realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, por ejemplo, ADN libre de células, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende proporcionar una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción, en la que determinar la fracción comprende determinar el número de marcas de secuencia fetales y maternas mapeadas en un genoma de referencia que comprende al menos un ácido nucleico polimórfico. La pluralidad de ácidos nucleicos polimórficos se localiza en una pluralidad de cromosomas diferentes. En una realización, la pluralidad de ácidos nucleicos polimórficos puede localizarse en los cromosomas 1-22. Por ejemplo, la pluralidad de ácidos nucleicos polimórficos puede localizarse en una pluralidad de cromosomas diferentes distintos de los cromosomas 13, 18, 21, X o Y. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

En una realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, por ejemplo, ADN libre de células, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende una amplificación y proporciona una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción, en la que determinar la fracción comprende determinar el número de marcas de secuencia fetales y maternas mapeadas en un genoma de referencia que comprende al menos un ácido nucleico polimórfico. La pluralidad de ácidos nucleicos polimórficos se localiza en una pluralidad de cromosomas diferentes. En una realización, la pluralidad de ácidos nucleicos polimórficos puede localizarse en los cromosomas 1-22. Por ejemplo, la pluralidad de ácidos nucleicos polimórficos puede localizarse en una pluralidad de cromosomas diferentes distintos de los cromosomas 13, 18, 21, X o Y. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

En otra realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, por ejemplo, ADN libre de células, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende proporcionar una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción, en la que determinar la fracción comprende determinar el número de marcas de secuencia fetales y maternas mapeadas en un genoma de referencia que comprende al menos un ácido nucleico polimórfico. La pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un único nucleótido (SNP). Alternativamente, la pluralidad de ácidos nucleicos diana polimórficos comprende al menos una repetición corta en tándem (STR). La pluralidad de ácidos nucleicos polimórficos se localiza en una pluralidad de cromosomas diferentes. En una realización, la pluralidad de ácidos nucleicos polimórficos puede localizarse en los cromosomas 1-22. Por ejemplo, la pluralidad de ácidos nucleicos polimórficos puede localizarse en una pluralidad de cromosomas diferentes distintos de los cromosomas 13, 18, 21, X o Y. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

En otra realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, por ejemplo, ADN libre de células, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende proporcionar una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción, en la que determinar la fracción comprende determinar el número de marcas de secuencia fetales y maternas mapeadas en un genoma de referencia que comprende al menos un ácido nucleico polimórfico. La pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un único nucleótido (SNP). Alternativamente, la pluralidad de ácidos nucleicos diana polimórficos comprende al menos una repetición corta en tándem (STR). La secuenciación masivamente paralela es secuenciación por síntesis con terminadores de colorante reversible. Alternativamente, la secuenciación masivamente paralela es secuenciación por ligación o secuenciación de una molécula única. La pluralidad de ácidos nucleicos polimórficos se localiza en una pluralidad de cromosomas diferentes. En una realización, la pluralidad de ácidos nucleicos polimórficos puede localizarse en los cromosomas 1-22. Por ejemplo, la pluralidad de ácidos nucleicos polimórficos puede localizarse en una pluralidad de cromosomas diferentes distintos de los cromosomas 13, 18, 21, X o Y. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

En otra realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, por ejemplo, ADN libre de células, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al

menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende una amplificación y proporciona una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción, en la que determinar la fracción comprende determinar el número de marcas de secuencia fetales y maternas mapeadas en un genoma de referencia que comprende al menos un ácido nucleico polimórfico. La pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un único nucleótido (SNP). Alternativamente, la pluralidad de ácidos nucleicos diana polimórficos comprende al menos una repetición corta en tándem (STR). La pluralidad de ácidos nucleicos polimórficos se localiza en una pluralidad de cromosomas diferentes. En una realización, la pluralidad de ácidos nucleicos polimórficos puede localizarse en los cromosomas 1-22. Por ejemplo, la pluralidad de ácidos nucleicos polimórficos puede localizarse en una pluralidad de cromosomas diferentes distintos de los cromosomas 13, 18, 21, X o Y. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

En otra realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, por ejemplo, ADN libre de células, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende una amplificación y proporciona una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción, en la que determinar la fracción comprende determinar el número de marcas de secuencia fetales y maternas mapeadas en un genoma de referencia que comprende al menos un ácido nucleico polimórfico. La pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un único nucleótido (SNP). Alternativamente, la pluralidad de ácidos nucleicos diana polimórficos comprende al menos una repetición corta en tándem (STR). La secuenciación masivamente paralela es secuenciación por síntesis con terminadores de colorante reversible. Alternativamente, la secuenciación masivamente paralela es secuenciación por ligación o secuenciación de una molécula única. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

En otra realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, por ejemplo, ADN libre de células, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende proporcionar una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción. La pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un único nucleótido (SNP). Alternativamente, la pluralidad de ácidos nucleicos diana polimórficos comprende al menos una repetición corta en tándem (STR). En algunas realizaciones, el al menos un SNP es un SNP único seleccionado de rs560681, rs1109037, rs9866013, rs13182883, rs13218440, rs7041158, rs740598, rs10773760, rs4530059, rs7205345, rs8078417, rs576261, rs2567608, rs430046, rs9951171, rs338882, rs10776839, rs9905977, rs1277284, rs258684, rs1347696, rs508485, rs9788670, rs8137254, rs3143, rs2182957, rs3739005 y rs530022. En otras realizaciones, el al menos un SNP es un SNP en tándem seleccionado de los pares de SNP en tándem rs7277033-rs2110153; rs2822654-rs1882882; rs368657-rs376635; rs2822731-rs2822732; rs1475881-rs7275487; rs1735976-rs2827016; rs447340-rs2824097; rs418989- rs13047336; rs987980- rs987981; rs4143392-rs4143391; rs1691324-rs13050434; rs11909758-rs9980111; rs2826842-rs232414; rs1980969-rs1980970; rs9978999-rs9979175; rs1034346-rs12481852; rs7509629-rs2828358; rs4817013-rs7277036; rs9981121-rs2829696; rs455921-rs2898102; rs2898102-rs458848; rs961301-rs2830208; rs2174536-rs458076; rs1088023-rs11088024; rs1011734-rs1011733; rs2831244-rs9789838; rs8132769-rs2831440; rs8134080-rs2831524; rs4817219-rs4817220; rs2250911-rs2250997; rs2831899-rs2831900; rs2831902-rs2831903; rs11088086-rs2251447; rs2832040-rs11088088; rs2832141-rs2246777; rs2832959-rs9980934; rs2833734-rs2833735; rs933121-rs933122; rs2834140-rs12626953; rs2834485-rs3453; rs9974986-rs2834703; rs2776266-rs2835001; rs1984014-rs1984015; rs7281674-rs2835316; rs13047304-rs13047322; rs2835545-rs4816551; rs2835735-rs2835736; rs13047608-rs2835826; rs2836550-rs2212596; rs2836660-rs2836661; rs465612-rs8131220; rs9980072-rs8130031; rs418359-rs2836926; rs7278447-rs7278858; rs385787-rs367001; rs367001-rs386095; rs2837296-rs2837297; y rs2837381-rs4816672. La al menos una STR está seleccionada de CSF1PO, FGA, TH01, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, D2S1338, Penta D, Penta E, D22S1045, D20S1082, D20S482, D18S853, D17S1301, D17S974, D14S1434, D12ATA63, D11S4463, D10S1435, D10S1248, D9S2157, D9S1122, D8S1115, D6S1017, D6S474, D5S2500, D4S2408, D4S2364, D3S4529, D3S3053, D2S1776, D2S441, D1S1677, D1S1627 y D1GATA113. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

En otra realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, por ejemplo, ADN libre de células, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende proporcionar una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción. La pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un único nucleótido (SNP).

Alternativamente, la pluralidad de ácidos nucleicos diana polimórficos comprende al menos una repetición corta en tándem (STR). En algunas realizaciones, el al menos un SNP es un SNP único seleccionado de rs560681, rs1109037, rs9866013, rs13182883, rs13218440, rs7041158, rs740598, rs10773760, rs4530059, rs7205345, rs8078417, rs576261, rs2567608, rs430046, rs9951171, rs338882, rs10776839, rs9905977, rs1277284, rs258684, rs1347696, rs508485, rs9788670, rs8137254, rs3143, rs2182957, rs3739005 y rs530022. En otras realizaciones, el al menos un SNP es un SNP en tándem seleccionado de los pares de SNP en tándem rs7277033-rs2110153; rs2822654-rs1882882; rs368657-rs376635; rs2822731-rs2822732; rs1475881-rs7275487; rs1735976-rs2827016; rs447340-rs2824097; rs418989- rs13047336; rs987980- rs987981; rs4143392-rs4143391; rs1691324-rs13050434; rs11909758-rs9980111; rs2826842-rs232414; rs1980969-rs1980970; rs9978999-rs9979175; rs1034346-rs12481852; rs7509629-rs2828358; rs4817013-rs7277036; rs9981121-rs2829696; rs455921-rs2898102; rs2898102-rs458848; rs961301-rs2830208; rs2174536-rs458076; rs11088023-rs11088024; rs1011734-rs1011733; rs2831244-rs9789838; rs8132769-rs2831440; rs8134080-rs2831524; rs4817219-rs4817220; rs2250911-rs2250997; rs2831899-rs2831900; rs2831902-rs2831903; rs11088086-rs2251447; rs2832040-rs11088088; rs2832141-rs2246777; rs2832959-rs9980934; rs2833734-rs2833735; rs933121-rs933122; rs2834140-rs12626953; rs2834485-rs3453; rs9974986-rs2834703; rs2776266-rs2835001; rs1984014-rs1984015; rs7281674-rs2835316; rs13047304-rs13047322; rs2835545-rs4816551; rs2835735-rs2835736; rs13047608-rs2835826; rs2836550-rs2212596; rs2836660-rs2836661; rs465612-rs8131220; rs9980072-rs8130031; rs418359-rs2836926; rs7278447-rs7278858; rs385787-rs367001; rs367001-rs386095; rs2837296-rs2837297; y rs2837381-rs4816672. La al menos una STR está seleccionada de CSF1PO, FGA, TH01, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, D2S1338, Penta D, Penta E, D22S1045, D20S1082, D20S482, D18S853, D17S1301, D17S974, D14S1434, D12ATA63, D11S4463, D10S1435, D10S1248, D9S2157, D9S1122, D8S1115, D6S1017, D6S474, D5S2500, D4S2408, D4S2364, D3S4529, D3S3053, D2S1776, D2S441, D1S1677, D1S1627 y D1GATA113. La secuenciación masivamente paralela es secuenciación por síntesis con terminadores de colorante reversible. Alternativamente, la secuenciación masivamente paralela es secuenciación por ligación o secuenciación de una molécula única. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

En otra realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, por ejemplo, ADN libre de células, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende una amplificación y proporciona una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción. La pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un único nucleótido (SNP). Alternativamente, la pluralidad de ácidos nucleicos diana polimórficos comprende al menos una repetición corta en tándem (STR). En algunas realizaciones, el al menos un SNP es un SNP único seleccionado de rs560681, rs1109037, rs9866013, rs13182883, rs13218440, rs7041158, rs740598, rs10773760, rs4530059, rs7205345, rs8078417, rs576261, rs2567608, rs430046, rs9951171, rs338882, rs10776839, rs9905977, rs1277284, rs258684, rs1347696, rs508485, rs9788670, rs8137254, rs3143, rs2182957, rs3739005 y rs530022. En otras realizaciones, el al menos un SNP es un SNP en tándem seleccionado de los pares de SNP en tándem rs7277033-rs2110153; rs2822654-rs1882882; rs368657-rs376635; rs2822731-rs2822732; rs1475881-rs7275487; rs1735976-rs2827016; rs447340-rs2824097; rs418989-rs13047336; rs987980-rs987981; rs4143392-rs4143391; rs1691324-rs13050434; rs11909758-rs9980111; rs2826842-rs232414; rs1980969-rs1980970; rs9978999-rs9979175; rs1034346-rs12481852; rs7509629-rs2828358; rs4817013-rs7277036; rs9981121-rs2829696; rs455921-rs2898102; rs2898102-rs458848; rs961301-rs2830208; rs2174536-rs458076; rs11088023-rs11088024; rs1011734-rs1011733; rs2831244-rs9789838; rs8132769-rs2831440; rs8134080-rs2831524; rs4817219-rs4817220; rs2250911-rs2250997; rs2831899-rs2831900; rs2831902-rs2831903; rs11088086-rs2251447; rs2832040-rs11088088; rs2832141-rs2246777; rs2832959-rs9980934; rs2833734-rs2833735; rs933121-rs933122; rs2834140-rs12626953; rs2834485-rs3453; rs9974986-rs2834703; rs2776266-rs2835001; rs1984014-rs1984015; rs7281674-rs2835316; rs13047304-rs13047322; rs2835545-rs4816551; rs2835735-rs2835736; rs13047608-rs2835826; rs2836550-rs2212596; rs2836660-rs2836661; rs465612-rs8131220; rs9980072-rs8130031; rs418359-rs2836926; rs7278447-rs7278858; rs385787-rs367001; rs367001-rs386095; rs2837296-rs2837297; y rs2837381-rs4816672. La al menos una STR está seleccionada de CSF1PO, FGA, TH01, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, D2S1338, Penta D, Penta E, D22S1045, D20S1082, D20S482, D18S853, D17S1301, D17S974, D14S1434, D12ATA63, D11S4463, D10S1435, D10S1248, D9S2157, D9S1122, D8S1115, D6S1017, D6S474, D5S2500, D4S2408, D4S2364, D3S4529, D3S3053, D2S1776, D2S441, D1S1677, D1S1627 y D1GATA113. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

En otra realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, por ejemplo, ADN libre de células, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende una amplificación y proporciona una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción. La pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un único

nucleótido (SNP). Alternativamente, la pluralidad de ácidos nucleicos diana polimórficos comprende al menos una repetición corta en tándem (STR). En algunas realizaciones, el al menos un SNP es un SNP único seleccionado de rs560681, rs1109037, rs9866013, rs13182883, rs13218440, rs7041158, rs740598, rs10773760, rs4530059, rs7205345, rs8078417, rs576261, rs2567608, rs430046, rs9951171, rs338882, rs10776839, rs9905977, rs1277284, rs258684, rs1347696, rs508485, rs9788670, rs8137254, rs3143, rs2182957, rs3739005 y rs530022. En otras realizaciones, el al menos un SNP es un SNP en tándem seleccionado de los pares de SNP en tándem rs7277033-rs2110153; rs2822654-rs1882882; rs368657-rs376635; rs2822731-rs2822732; rs1475881-rs7275487; rs1735976-rs2827016; rs447340-rs2824097; rs418989-rs13047336; rs987980-rs987981; rs4143392-rs4143391; rs1691324-rs13050434; rs11909758-rs9980111; rs2826842-rs232414; rs1980969-rs1980970; rs9978999-rs9979175; rs1034346-rs12481852; rs7509629-rs2828358; rs4817013-rs7277036; rs9981121-rs2829696; rs455921-rs2898102; rs2898102-rs458848; rs961301-rs2830208; rs2174536-rs458076; rs11088023-rs11088024; rs1011734-rs1011733; rs2831244-rs9789838; rs8132769-rs2831440; rs8134080-rs2831524; rs4817219-rs4817220; rs2250911-rs2250997; rs2831899-rs2831900; rs2831902-rs2831903; rs11088086-rs2251447; rs2832040-rs11088088; rs2832141-rs2246777; rs2832959-rs9980934; rs2833734-rs2833735; rs933121-rs933122; rs2834140-rs12626953; rs2834485-rs3453; rs9974986-rs2834703; rs2776266-rs2835001; rs1984014-rs1984015; rs7281674-rs2835316; rs13047304-rs13047322; rs2835545-rs4816551; rs2835735-rs2835736; rs13047608-rs2835826; rs2836550-rs2212596; rs2836660-rs2836661; rs465612-rs8131220; rs9980072-rs8130031; rs418359-rs2836926; rs7278447-rs7278858; rs385787-rs367001; rs367001-rs386095; rs2837296-rs2837297; y rs2837381-rs4816672. La al menos una STR está seleccionada de CSF1PO, FGA, TH01, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, D2S1338, Penta D, Penta E, D22S1045, D20S1082, D20S482, D18S853, D17S1301, D17S974, D14S1434, D12ATA63, D11S4463, D10S1435, D10S1248, D9S2157, D9S1122, D8S1115, D6S1017, D6S474, D5S2500, D4S2408, D4S2364, D3S4529, D3S3053, D2S1776, D2S441, D1S1677, D1S1627 y D1GATA113. La secuenciación masivamente paralela es secuenciación por síntesis con terminadores de colorante reversible. Alternativamente, la secuenciación masivamente paralela es secuenciación por ligación o secuenciación de una molécula única. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

En una realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende proporcionar una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción, en la que determinar la fracción comprende determinar el número de marcas de secuencia fetales y maternas mapeadas en un genoma de referencia que comprende al menos un ácido nucleico polimórfico. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

En otra realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende proporcionar una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción, en la que determinar la fracción comprende determinar el número de marcas de secuencia fetales y maternas mapeadas en un genoma de referencia que comprende al menos un ácido nucleico polimórfico. La pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un único nucleótido (SNP). Alternativamente, la pluralidad de ácidos nucleicos diana polimórficos comprende al menos una repetición corta en tándem (STR). En algunas realizaciones, el al menos un SNP es un SNP único seleccionado de rs560681, rs109037, rs9866013, rs13182883, rs13218440, rs7041158, rs740598, rs10773760, rs4530059, rs7205345, rs8078417, rs576261, rs2567608, rs430046, rs9951171, rs338882, rs10776839, rs9905977, rs1277284, rs258684, rs1347696, rs508485, rs9788670, rs8137254, rs3143, rs2182957, rs3739005 y rs530022. En otras realizaciones, el al menos un SNP es un SNP en tándem seleccionado de los pares de SNP en tándem rs7277033-rs2110153; rs2822654-rs1882882; rs368657-rs376635; rs2822731-rs2822732; rs1475881-rs7275487; rs1735976-rs2827016; rs447340-rs2824097; rs418989-rs13047336; rs987980-rs987981; rs4143392-rs4143391; rs1691324-rs13050434; rs11909758-rs9980111; rs2826842-rs232414; rs1980969-rs1980970; rs9978999-rs9979175; rs1034346-rs12481852; rs7509629-rs2828358; rs4817013-rs7277036; rs9981121-rs2829696; rs455921-rs2898102; rs2898102-rs458848; rs961301-rs2830208; rs2174536-rs458076; rs11088023-rs11088024; rs1011734-rs1011733; rs2831244-rs9789838; rs8132769-rs2831440; rs8134080-rs2831524; rs4817219-rs4817220; rs2250911-rs2250997; rs2831899-rs2831900; rs2831902-rs2831903; rs11088086-rs2251447; rs2832040-rs11088088; rs2832141-rs2246777; rs2832959-rs9980934; rs2833734-rs2833735; rs933121-rs933122; rs2834140-rs12626953; rs2834485-rs3453; rs9974986-rs2834703; rs2776266-rs2835001; rs1984014-rs1984015; rs7281674-rs2835316; rs13047304-rs13047322; rs2835545-rs4816551; rs2835735-rs2835736; rs13047608-rs2835826; rs2836550-rs2212596; rs2836660-rs2836661; rs465612-rs8131220; rs9980072-rs8130031; rs418359-rs2836926; rs7278447-rs7278858; rs385787-rs367001; rs367001-rs386095; rs2837296-rs2837297; y rs2837381-rs4816672. La al menos una STR está seleccionada de CSF1PO, FGA, TH01, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, D2S1338, Penta D, Penta E, D22S1045, D20S1082, D20S482, D18S853, D17S1301, D17S974, D14S1434, D12ATA63, D11S4463, D10S1435, D10S1248, D9S2157, D9S1122, D8S1115,

D6S1017, D6S474, D5S2500, D4S2408, D4S2364, D3S4529, D3S3053, D2S1776, D2S441, D1S1677, D1S1627 y D1GATA113. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

5 En otra realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende proporcionar una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción, en la que determinar la fracción comprende determinar el número de marcas de secuencia fetales y maternas mapeadas en un genoma de referencia que comprende al menos un ácido nucleico polimórfico. La pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un único nucleótido (SNP). Alternativamente, la pluralidad de ácidos nucleicos diana polimórficos comprende al menos una repetición corta en tándem (STR). En algunas realizaciones, el al menos un SNP es un SNP único seleccionado de rs560681, rs1109037, rs9866013, rs13182883, rs13218440, rs7041158, rs740598, rs10773760, rs4530059, rs7205345, rs8078417, rs576261, rs2567608, rs430046, rs9951171, rs338882, rs10776839, rs9905977, rs1277284, rs258684, rs1347696, rs508485, rs9788670, rs8137254, rs3143, rs2182957, rs3739005 y rs530022. En otras realizaciones, el al menos un SNP es un SNP en tándem seleccionado de los pares de SNP en tándem rs7277033-rs2110153; rs2822654-rs1882882; rs368657-rs376635; rs2822731-rs2822732; rs1475881-rs7275487; rs1735976-rs2827016; rs447340-rs2824097; rs418989-rs13047336; rs987980-rs987981; rs4143392-rs4143391; rs1691324-rs13050434; rs11909758-rs9980111; rs2826842-rs232414; rs1980969-rs1980970; rs9978999-rs9979175; rs1034346-rs12481852; rs7509629-rs2828358; rs4817013-rs7277036; rs9981121-rs2829696; rs455921-rs2898102; rs2898102-rs458848; rs961301-rs2830208; rs2174536-rs458076; rs11088023-rs11088024; rs1011734-rs1011733; rs2831244-rs9789838; rs8132769-rs2831440; rs8134080-rs2831524; rs4817219-rs4817220; rs2250911-rs2250997; rs2831899-rs2831900; rs2831902-rs2831903; rs11088086-rs2251447; rs2832040-rs11088088; rs2832141-rs2246777; rs2832959-rs9980934; rs2833734-rs2833735; rs933121-rs933122; rs2834140-rs12626953; rs2834485-rs3453; rs9974986-rs2834703; rs2776266-rs2835001; rs1984014-rs1984015; rs7281674-rs2835316; rs13047304-rs13047322; rs2835545-rs4816551; rs2835735-rs2835736; rs13047608-rs2835826; rs2836550-rs2212596; rs2836660-rs2836661; rs465612-rs8131220; rs9980072-rs8130031; rs418359-rs2836926; rs7278447-rs7278858; rs385787-rs367001; rs367001-rs386095; rs2837296-rs2837297; y rs2837381-rs4816672. La al menos una STR está seleccionada de CSF1PO, FGA, TH01, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, D2S1338, Penta D, Penta E, D22S1045, D20S1082, D20S482, D18S853, D17S1301, D17S974, D14S1434, D12ATA63, D11S4463, D10S1435, D10S1248, D9S2157, D9S1122, D8S1115, D6S1017, D6S474, D5S2500, D4S2408, D4S2364, D3S4529, D3S3053, D2S1776, D2S441, D1S1677, D1S1627 y D1GATA113. La secuenciación masivamente paralela es secuenciación por síntesis con terminadores de colorante reversible. Alternativamente, la secuenciación masivamente paralela es secuenciación por ligación o secuenciación de una molécula única. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

40 En otra realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende una amplificación y proporciona una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción, en la que determinar la fracción comprende determinar el número de marcas de secuencia fetales y maternas mapeadas en un genoma de referencia que comprende al menos un ácido nucleico polimórfico. La pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un único nucleótido (SNP). Alternativamente, la pluralidad de ácidos nucleicos diana polimórficos comprende al menos una repetición corta en tándem (STR). En algunas realizaciones, el al menos un SNP es un SNP único seleccionado de rs560681, rs1109037, rs9866013, rs13182883, rs13218440, rs7041158, rs740598, rs10773760, rs4530059, rs7205345, rs8078417, rs576261, rs2567608, rs430046, rs9951171, rs338882, rs10776839, rs9905977, rs1277284, rs258684, rs1347696, rs508485, rs9788670, rs8137254, rs3143, rs2182957, rs3739005 y rs530022. En otras realizaciones, el al menos un SNP es un SNP en tándem seleccionado de los pares de SNP en tándem rs7277033-rs2110153; rs2822654-rs1882882; rs368657-rs376635; rs2822731-rs2822732; rs1475881-rs7275487; rs1735976-rs2827016; rs447340-rs2824097; rs418989-rs13047336; rs987980-rs987981; rs4143392-rs4143391; rs1691324-rs13050434; rs11909758-rs9980111; rs2826842-rs232414; rs1980969-rs1980970; rs9978999-rs9979175; rs1034346-rs12481852; rs7509629-rs2828358; rs4817013-rs7277036; rs9981121-rs2829696; rs455921-rs2898102; rs2898102-rs458848; rs961301-rs2830208; rs2174536-rs458076; rs11088023-rs11088024; rs1011734-rs1011733; rs2831244-rs9789838; rs8132769-rs2831440; rs8134080-rs2831524; rs4817219-rs4817220; rs2250911-rs2250997; rs2831899-rs2831900; rs2831902-rs2831903; rs11088086-rs2251447; rs2832040-rs11088088; rs2832141-rs2246777; rs2832959-rs9980934; rs2833734-rs2833735; rs933121-rs933122; rs2834140-rs12626953; rs2834485-rs3453; rs9974986-rs2834703; rs2776266-rs2835001; rs1984014-rs1984015; rs7281674-rs2835316; rs13047304-rs13047322; rs2835545-rs4816551; rs2835735-rs2835736; rs13047608-rs2835826; rs2836550-rs2212596; rs2836660-rs2836661; rs465612-rs8131220; rs9980072-rs8130031; rs418359-rs2836926; rs7278447-rs7278858; rs385787-rs367001; rs367001-rs386095; rs2837296-rs2837297; y rs2837381-rs4816672. La al menos una STR está seleccionada de CSF1PO, FGA, TH01, vWA, D3S1358, D5S818, D7S820, D8S1179,

D13S317, D16S539, D18S51, D21S11, D2S1338, Penta D, Penta E, D22S1045, D20S1082, D20S482, D18S853, D17S1301, D17S974, D14S1434, D12ATA63, D11S4463, D10S1435, D10S1248, D9S2157, D9S1122, D8S1115, D6S1017, D6S474, D5S2500, D4S2408, D4S2364, D3S4529, D3S3053, D2S1776, D2S441, D1S1677, D1S1627 y D1GATA113. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

En otra realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende una amplificación y proporciona una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción, en la que determinar la fracción comprende determinar el número de marcas de secuencia fetales y maternas mapeadas en un genoma de referencia que comprende al menos un ácido nucleico polimórfico. La pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un único nucleótido (SNP). Alternativamente, la pluralidad de ácidos nucleicos diana polimórficos comprende al menos una repetición corta en tándem (STR). En algunas realizaciones, el al menos un SNP es un SNP único seleccionado de rs560681, rs1109037, rs9866013, rs13182883, rs13218440, rs7041158, rs740598, rs10773760, rs4530059, rs7205345, rs8078417, rs576261, rs2567608, rs430046, rs9951171, rs338882, rs10776839, rs9905977, rs1277284, rs258684, rs1347696, rs508485, rs9788670, rs8137254, rs3143, rs2182957, rs3739005 y rs530022. En otras realizaciones, el al menos un SNP es un SNP en tándem seleccionado de los pares de SNP en tándem rs7277033-rs2110153; rs2822654-rs1882882; rs368657-rs376635; rs2822731-rs2822732; rs1475881-rs7275487; rs1735976-rs2827016; rs447340-rs2824097; rs418989- rs13047336; rs987980- rs987981; rs4143392-rs4143391; rs1691324-rs13050434; rs11909758-rs9980111; rs2826842-rs232414; rs1980969-rs1980970; rs9978999-rs9979175; rs1034346-rs12481852; rs7509629-rs2828358; rs4817013-rs7277036; rs9981121-rs2829696; rs455921-rs2898102; rs2898102-rs458848; rs961301-rs2830208; rs2174536-rs458076; rs11088023-rs11088024; rs1011734-rs1011733; rs2831244-rs9789838; rs8132769-rs2831440; rs8134080-rs2831524; rs4817219-rs4817220; rs2250911-rs2250997; rs2831899-rs2831900; rs2831902-rs2831903; rs11088086-rs2251447; rs2832040-rs11088088;rs2832141-rs2246777; rs2832959-rs9980934; rs2833734-rs2833735; rs933121-rs933122; rs2834140-rs12626953; rs2834485-rs3453; rs9974986-rs2834703; rs2776266-rs2835001; rs1984014-rs1984015; rs7281674-rs2835316; rs13047304-rs13047322; rs2835545-rs4816551; rs2835735-rs2835736; rs13047608-rs2835826; rs2836550-rs2212596; rs2836660-rs2836661; rs465612-rs8131220; rs9980072-rs8130031; rs418359-rs2836926; rs7278447-rs7278858; rs385787-rs367001; rs367001-rs386095; rs2837296-rs2837297; y rs2837381-rs4816672. La al menos una STR está seleccionada de CSF1PO, FGA, TH01, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, D2S1338, Penta D, Penta E, D22S1045, D20S1082, D20S482, D18S853, D17S1301, D17S974, D14S1434, D12ATA63, D11S4463, D10S1435, D10S1248, D9S2157, D9S1122, D8S1115, D6S1017, D6S474, D5S2500, D4S2408, D4S2364, D3S4529, D3S3053, D2S1776, D2S441, D1S1677, D1S1627 y D1GATA113. La secuenciación masivamente paralela es secuenciación por síntesis con terminadores de colorante reversible. Alternativamente, la secuenciación masivamente paralela es secuenciación por ligación o secuenciación de una molécula única. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

En una realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, por ejemplo, ADN libre de células, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende proporcionar una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción, en la que determinar la fracción comprende determinar el número de marcas de secuencia fetales y maternas mapeadas en un genoma de referencia que comprende al menos un ácido nucleico polimórfico. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

En una realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, por ejemplo, ADN libre de células, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende una amplificación y proporciona una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción, en la que determinar la fracción comprende determinar el número de marcas de secuencia fetales y maternas mapeadas en un genoma de referencia que comprende al menos un ácido nucleico polimórfico. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

En otra realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, por ejemplo, ADN libre de células, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos

diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende proporcionar una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción, en la que determinar la fracción comprende determinar el número de marcas de secuencia fetales y maternas mapeadas en un genoma de referencia que comprende al menos un ácido nucleico polimórfico. La pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un único nucleótido (SNP). Alternativamente, la pluralidad de ácidos nucleicos diana polimórficos comprende al menos una repetición corta en tándem (STR). En algunas realizaciones, el al menos un SNP es un SNP único seleccionado de rs560681, rs1109037, rs9866013, rs13182883, rs13218440, rs7041158, rs740598, rs10773760, rs4530059, rs7205345, rs8078417, rs576261, rs2567608, rs430046, rs9951171, rs338882, rs10776839, rs9905977, rs1277284, rs258684, rs1347696, rs508485, rs9788670, rs8137254, rs3143, rs2182957, rs3739005 y rs530022. En otras realizaciones, el al menos un SNP es un SNP en tándem seleccionado de los pares de SNP en tándem rs7277033-rs2110153; rs2822654-rs1882882; rs368657-rs376635; rs2822731-rs2822732; rs1475881-rs7275487; rs1735976-rs2827016; rs447340-rs2824097; rs418989- rs13047336; rs987980- rs987981; rs4143392-rs4143391; rs1691324-rs13050434; rs11909758-rs9980111; rs2826842-rs232414; rs1980969-rs1980970; rs9978999-rs9979175; rs1034346-rs12481852; rs7509629-rs2828358; rs4817013-rs7277036; rs9981121-rs2829696; rs455921-rs2898102; rs2898102-rs458848; rs961301-rs2830208; rs2174536-rs458076; rs11088023-rs11088024; rs1011734-rs1011733; rs2831244-rs9789838; rs8132769-rs2831440; rs8134080-rs2831524; rs4817219-rs4817220; rs2250911-rs2250997; rs2831899-rs2831900; rs2831902-rs2831903; rs11088086-rs2251447; rs2832040-rs11088088; rs2832141-rs2246777; rs2832959-rs9980934; rs2833734-rs2833735; rs933121-rs933122; rs2834140-rs12626953; rs2834485-rs3453; rs9974986-rs2834703; rs2776266-rs2835001; rs1984014-rs1984015; rs7281674-rs2835316; rs13047304-rs13047322; rs2835545-rs4816551; rs2835735-rs2835736; rs13047608-rs2835826; rs2836550-rs2212596; rs2836660-rs2836661; rs465612-rs8131220; rs9980072-rs8130031; rs418359-rs2836926; rs7278447-rs7278858; rs385787-rs367001; rs367001-rs386095; rs2837296-rs2837297; y rs2837381-rs4816672. La al menos una STR está seleccionada de CSF1PO, FGA, TH01, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, D2S1338, Penta D, Penta E, D22S1045, D20S1082, D20S482, D18S853, D17S1301, D17S974, D14S1434, D12ATA63, D11S4463, D10S1435, D10S1248, D9S2157, D9S1122, D8S1115, D6S1017, D6S474, D5S2500, D4S2408, D4S2364, D3S4529, D3S3053, D2S1776, D2S441, D1S1677, D1S1627 y D1GATA113. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

En otra realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, por ejemplo, ADN libre de células, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende proporcionar una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción, en la que determinar la fracción comprende determinar el número de marcas de secuencia fetales y maternas mapeadas en un genoma de referencia que comprende al menos un ácido nucleico polimórfico. La pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un único nucleótido (SNP). Alternativamente, la pluralidad de ácidos nucleicos diana polimórficos comprende al menos una repetición corta en tándem (STR). En algunas realizaciones, el al menos un SNP es un SNP único seleccionado de rs560681, rs1109037, rs9866013, rs13182883, rs13218440, rs7041158, rs740598, rs10773760, rs4530059, rs7205345, rs8078417, rs576261, rs2567608, rs430046, rs9951171, rs338882, rs10776839, rs9905977, rs1277284, rs258684, rs1347696, rs508485, rs9788670, rs8137254, rs3143, rs2182957, rs3739005 y rs530022. En otras realizaciones, el al menos un SNP es un SNP en tándem seleccionado de los pares de SNP en tándem rs7277033-rs2110153; rs2822654-rs1882882; rs368657-rs376635; rs2822731-rs2822732; rs1475881-rs7275487; rs1735976-rs2827016; rs447340-rs2824097; rs418989- rs13047336; rs987980- rs987981; rs4143392-rs4143391; rs1691324-rs13050434; rs11909758-rs9980111; rs2826842-rs232414; rs1980969-rs1980970; rs9978999-rs9979175; rs1034346-rs12481852; rs7509629-rs2828358; rs4817013-rs7277036; rs9981121-rs2829696; rs455921-rs2898102; rs2898102-rs458848; rs961301-rs2830208; rs2174536-rs458076; rs11088023-rs11088024; rs1011734-rs1011733; rs2831244-rs9789838; rs8132769-rs2831440; rs8134080-rs2831524; rs4817219-rs4817220; rs2250911-rs2250997; rs2831899-rs2831900; rs2831902-rs2831903; rs11088086-rs2251447; rs2832040-rs11088088; rs2832141-rs2246777; rs2832959-rs9980934; rs2833734-rs2833735; rs933121-rs933122; rs2834140-rs12626953; rs2834485-rs3453; rs9974986-rs2834703; rs2776266-rs2835001; rs1984014-rs1984015; rs7281674-rs2835316; rs13047304-rs13047322; rs2835545-rs4816551; rs2835735-rs2835736; rs13047608-rs2835826; rs2836550-rs2212596; rs2836660-rs2836661; rs465612-rs8131220; rs9980072-rs8130031; rs418359-rs2836926; rs7278447-rs7278858; rs385787-rs367001; rs367001-rs386095; rs2837296-rs2837297; y rs2837381-rs4816672. La al menos una STR está seleccionada de CSF1PO, FGA, TH01, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, D2S1338, Penta D, Penta E, D22S1045, D20S1082, D20S482, D18S853, D17S1301, D17S974, D14S1434, D12ATA63, D11S4463, D10S1435, D10S1248, D9S2157, D9S1122, D8S1115, D6S1017, D6S474, D5S2500, D4S2408, D4S2364, D3S4529, D3S3053, D2S1776, D2S441, D1S1677, D1S1627 y D1GATA113. La secuenciación masivamente paralela es secuenciación por síntesis con terminadores de colorante reversible. Alternativamente, la secuenciación masivamente paralela es secuenciación por ligación o secuenciación de una molécula única. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

En otra realización, en el presente documento se desvela un método de determinación de la fracción de ácidos

nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, por ejemplo, ADN libre de células, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende una amplificación y proporciona una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción, en la que determinar la fracción comprende determinar el número de marcas de secuencia fetales y maternas mapeadas en un genoma de referencia que comprende al menos un ácido nucleico polimórfico. La pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un único nucleótido (SNP). Alternativamente, la pluralidad de ácidos nucleicos diana polimórficos comprende al menos una repetición corta en tándem (STR). En algunas realizaciones, el al menos un SNP es un SNP único seleccionado de rs560681, rs1109037, rs9866013, rs13182883, rs13218440, rs7041158, rs740598, rs10773760, rs4530059, rs7205345, rs8078417, rs576261, rs2567608, rs430046, rs9951171, rs338882, rs10776839, rs9905977, rs1277284, rs258684, rs1347696, rs508485, rs9788670, rs8137254, rs3143, rs2182957, rs3739005 y rs530022. En otras realizaciones, el al menos un SNP es un SNP en tándem seleccionado de los pares de SNP en tándem rs7277033-rs2110153; rs2822654-rs1882882; rs368657-rs376635; rs2822731-rs2822732; rs1475881-rs7275487; rs1735976-rs2827016; rs447340-rs2824097; rs418989- rs13047336; rs987980- rs987981; rs4143392-rs4143391; rs1691324-rs13050434; rs11909758-rs9980111; rs2826842-rs232414; rs1980969-rs1980970; rs9978999-rs9979175; rs1034346-rs12481852; rs7509629-rs2828358; rs4817013-rs7277036; rs9981121-rs2829696; rs455921-rs2898102; rs2898102-rs458848; rs961301-rs2830208; rs2174536-rs458076; rs11088023-rs11088024; rs1011734-rs1011733; rs2831244-rs9789838; rs8132769-rs2831440; rs8134080-rs2831524; rs4817219-rs4817220; rs2250911-rs2250997; rs2831899-rs2831900; rs2831902-rs2831903; rs11088086-rs2251447; rs2832040-rs11088088; rs2832141-rs2246777; rs2832959-rs9980934; rs2833734-rs2833735; rs933121-rs933122; rs2834140-rs12626953; rs2834485-rs3453; rs9974986-rs2834703; rs2776266-rs2835001; rs1984014-rs1984015; rs7281674-rs2835316; rs13047304-rs13047322; rs2835545-rs4816551; rs2835735-rs2835736; rs13047608-rs2835826; rs2836550-rs2212596; rs2836660-rs2836661; rs465612-rs8131220; rs9980072-rs8130031; rs418359-rs2836926; rs7278447-rs7278858; rs385787-rs367001; rs367001-rs386095; rs2837296-rs2837297; y rs2837381-rs4816672. La al menos una STR está seleccionada de CSF1PO, FGA, TH01, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, D2S1338, Penta D, Penta E, D22S1045, D20S1082, D20S482, D18S853, D17S1301, D17S974, D14S1434, D12ATA63, D11S4463, D10S1435, D10S1248, D9S2157, D9S1122, D8S1115, D6S1017, D6S474, D5S2500, D4S2408, D4S2364, D3S4529, D3S3053, D2S1776, D2S441, D1S1677, D1S1627 y D1GATA113. La muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

En otra realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ADN fetal y genómico, por ejemplo, ADN libre de células, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en la mezcla de ADN fetal y genómico; (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que la secuenciación comprende una amplificación y proporciona una pluralidad de marcas de secuencia; y (c) basándose en la secuenciación, determinar la fracción, en la que determinar la fracción comprende determinar el número de marcas de secuencia fetales y maternas mapeadas en un genoma de referencia que comprende al menos un ácido nucleico polimórfico. La pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un único nucleótido (SNP). Alternativamente, la pluralidad de ácidos nucleicos diana polimórficos comprende al menos una repetición corta en tándem (STR). En algunas realizaciones, el al menos un SNP es un SNP único seleccionado de rs560681, rs1109037, rs9866013, rs13182883, rs13218440, rs7041158, rs740598, rs10773760, rs4530059, rs7205345, rs8078417, rs576261, rs2567608, rs430046, rs9951171, rs338882, rs10776839, rs9905977, rs1277284, rs258684, rs1347696, rs508485, rs9788670, rs8137254, rs3143, rs2182957, rs3739005 y rs530022. En otras realizaciones, el al menos un SNP es un SNP en tándem seleccionado de los pares de SNP en tándem rs7277033-rs2110153; rs2822654-rs1882882; rs368657-rs376635; rs2822731-rs2822732; rs1475881-rs7275487; rs1735976-rs2827016; rs447340-rs2824097; rs418989- rs13047336; rs987980- rs987981; rs4143392-rs4143391; rs1691324-rs13050434; rs11909758-rs9980111; rs2826842-rs232414; rs1980969-rs1980970; rs9978999-rs9979175; rs1034346-rs12481852; rs7509629-rs2828358; rs4817013-rs7277036; rs9981121-rs2829696; rs455921-rs2898102; rs2898102-rs458848; rs961301-rs2830208; rs2174536-rs458076; rs11088023-rs11088024; rs1011734-rs1011733; rs2831244-rs9789838; rs8132769-rs2831440; rs8134080-rs2831524; rs4817219-rs4817220; rs2250911-rs2250997; rs2831899-rs2831900; rs2831902-rs2831903; rs11088086-rs2251447; rs2832040-rs11088088; rs2832141-rs2246777; rs2832959-rs9980934; rs2833734-rs2833735; rs933121-rs933122; rs2834140-rs12626953; rs2834485-rs3453; rs9974986-rs2834703; rs2776266-rs2835001; rs1984014-rs1984015; rs7281674-rs2835316; rs13047304-rs13047322; rs2835545-rs4816551; rs2835735-rs2835736; rs13047608-rs2835826; rs2836550-rs2212596; rs2836660-rs2836661; rs465612-rs8131220; rs9980072-rs8130031; rs418359-rs2836926; rs7278447-rs7278858; rs385787-rs367001; rs367001-rs386095; rs2837296-rs2837297; y rs2837381-rs4816672. La al menos una STR está seleccionada de CSF1PO, FGA, TH01, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, D2S1338, Penta D, Penta E, D22S1045, D20S1082, D20S482, D18S853, D17S1301, D17S974, D14S1434, D12ATA63, D11S4463, D10S1435, D10S1248, D9S2157, D9S1122, D8S1115, D6S1017, D6S474, D5S2500, D4S2408, D4S2364, D3S4529, D3S3053, D2S1776, D2S441, D1S1677, D1S1627 y D1GATA113. La secuenciación masivamente paralela es secuenciación por síntesis con terminadores de colorante reversible. Alternativamente, la secuenciación masivamente paralela es secuenciación por ligación o secuenciación de una molécula única. La

muestra materna está seleccionada de sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma.

5 En otra realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra de plasma materno que comprende una mezcla de ADN fetal y genómico, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos polimórficos en dicha mezcla de ácidos nucleicos fetales y maternos, en la que cada uno de dicho al menos un ácido nucleico polimórfico comprende una STR; (b) determinar la cantidad de alelos de STR fetales y maternos de al menos un ácido nucleico polimórfico; y (c) determinar dicha fracción usando dicha cantidad de alelos de STR fetales y maternos. La pluralidad de ácidos nucleicos polimórficos se localiza en una pluralidad de cromosomas diferentes. En una realización, la pluralidad de ácidos nucleicos polimórficos puede localizarse en los cromosomas 1-22. Por ejemplo, la pluralidad de ácidos nucleicos polimórficos puede localizarse en una pluralidad de cromosomas diferentes distintos de los cromosomas 13, 18, 21, X o Y. El método es un método independiente del sexo del feto.

15 En otra realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra de plasma materno que comprende una mezcla de ADN fetal y genómico, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos polimórficos en dicha mezcla de ácidos nucleicos fetales y maternos, en la que cada uno de dicho al menos un ácido nucleico polimórfico comprende una STR; (b) determinar la cantidad de alelos de STR fetales y maternos de al menos un ácido nucleico polimórfico; y (c) determinar dicha fracción usando dicha cantidad de alelos de STR fetales y maternos. La pluralidad de ácidos nucleicos polimórficos se localiza en una pluralidad de cromosomas diferentes. En una realización, la pluralidad de ácidos nucleicos polimórficos puede localizarse en los cromosomas 1-22. Por ejemplo, la pluralidad de ácidos nucleicos polimórficos se localiza en una pluralidad de cromosomas diferentes distintos de los cromosomas 13, 18, 21, X o Y. El método comprende además preamplificar la mezcla de ácidos nucleicos fetales y maternos. El método es un método independiente del sexo del feto.

30 En otra realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra de plasma materno que comprende una mezcla de ADN fetal y genómico, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos polimórficos en dicha mezcla de ácidos nucleicos fetales y maternos, en la que cada uno de dicho al menos un ácido nucleico polimórfico comprende una STR; (b) determinar la cantidad de alelos de STR fetales y maternos de al menos un ácido nucleico polimórfico; y (c) determinar dicha fracción usando dicha cantidad de alelos de STR fetales y maternos. La pluralidad de ácidos nucleicos polimórficos se localiza en una pluralidad de cromosomas diferentes. En una realización, la pluralidad de ácidos nucleicos polimórficos puede localizarse en los cromosomas 1-22. Por ejemplo, la pluralidad de ácidos nucleicos polimórficos puede localizarse en una pluralidad de cromosomas diferentes distintos de los cromosomas 13, 18, 21, X o Y. El método comprende además resolver el tamaño de las STR usando electroforesis capilar. El método es un método independiente del sexo del feto.

40 En otra realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra de plasma materno que comprende una mezcla de ADN fetal y genómico, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos polimórficos en dicha mezcla de ácidos nucleicos fetales y maternos, en la que cada uno de dicho al menos un ácido nucleico polimórfico comprende una STR; (b) determinar la cantidad de alelos de STR fetales y maternos de al menos un ácido nucleico polimórfico; y (c) determinar dicha fracción usando dicha cantidad de alelos de STR fetales y maternos. La pluralidad de ácidos nucleicos polimórficos se localiza en una pluralidad de cromosomas diferentes. En una realización, la pluralidad de ácidos nucleicos polimórficos puede localizarse en los cromosomas 1-22. Por ejemplo, la pluralidad de ácidos nucleicos polimórficos puede localizarse en una pluralidad de cromosomas diferentes distintos de los cromosomas 13, 18, 21, X o Y. El método comprende además preamplificar la mezcla de ácidos nucleicos fetales y maternos, y resolver el tamaño de las STR usando electroforesis capilar. El método es un método independiente del sexo del feto.

55 En otra realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra de plasma materno que comprende una mezcla de ADN fetal y genómico, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos polimórficos en dicha mezcla de ácidos nucleicos fetales y maternos, en la que cada uno de dicho al menos un ácido nucleico polimórfico comprende una STR; (b) determinar la cantidad de alelos de STR fetales y maternos de al menos un ácido nucleico polimórfico; y (c) determinar dicha fracción usando dicha cantidad de alelos de STR fetales y maternos. La pluralidad de ácidos nucleicos polimórficos puede localizarse en una pluralidad de cromosomas diferentes. Por ejemplo, la pluralidad de ácidos nucleicos polimórficos se localiza en una pluralidad de cromosomas diferentes distintos de los cromosomas 13, 18, 21, X o Y. La al menos una STR está seleccionada de CSF1PO, FGA, TH01, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D2S1338, Penta D, Penta E, D22S1045, D20S1082, D20S482, D18S853, D17S1301, D17S974, D14S1434, D12ATA63, D11S4463, D10S1435, D10S1248, D9S2157, D9S1122, D8S1115, D6S1017, D6S474, D5S2500, D4S2408, D4S2364, D3S4529, D3S3053, D2S1776, D2S441, D1S1677, D1S1627 y D1GATA113. Opcionalmente, la al menos una STR puede ser el panel de STR que comprende CSF1PO, D13S317, D16S539, D18S51, D21S11, D2S1338, D7S820 y FGA. El método es un método independiente del sexo del feto.

En otra realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra de plasma materno que comprende una mezcla de ADN fetal y genómico, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos polimórficos en dicha mezcla de ácidos nucleicos fetales y maternos, en la que cada uno de dicho al menos un ácido nucleico polimórfico comprende una STR; (b) determinar la cantidad de alelos de STR fetales y maternos de al menos un ácido nucleico polimórfico; y (c) determinar dicha fracción usando dicha cantidad de alelos de STR fetales y maternos. La pluralidad de ácidos nucleicos polimórficos se localiza en una pluralidad de cromosomas diferentes. En una realización, la pluralidad de ácidos nucleicos polimórficos puede localizarse en los cromosomas 1-22. Por ejemplo, la pluralidad de ácidos nucleicos polimórficos puede localizarse en una pluralidad de cromosomas diferentes distintos de los cromosomas 13, 18, 21, X o Y. El método comprende además preamplificar la mezcla de ácidos nucleicos fetales y maternos. La al menos una STR está seleccionada de CSF1PO, FGA, TH01, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D2S1338, Penta D, Penta E, D22S1045, D20S1082, D20S482, D18S853, D17S1301, D17S974, D14S1434, D12ATA63, D11S4463, D10S1435, D10S1248, D9S2157, D9S1122, D8S1115, D6S1017, D6S474, D5S2500, D4S2408, D4S2364, D3S4529, D3S3053, D2S1776, D2S441, D1S1677, D1S1627 y D1GATA113. Opcionalmente, la al menos una STR puede ser el panel de STR que comprende CSF1PO, D13S317, D16S539, D18S51, D21S11, D2S1338, D7S820 y FGA. El método es un método independiente del sexo del feto.

En otra realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra de plasma materno que comprende una mezcla de ADN fetal y genómico, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos polimórficos en dicha mezcla de ácidos nucleicos fetales y maternos, en la que cada uno de dicho al menos un ácido nucleico polimórfico comprende una STR; (b) determinar la cantidad de alelos de STR fetales y maternos de al menos un ácido nucleico polimórfico; y (c) determinar dicha fracción usando dicha cantidad de alelos de STR fetales y maternos. La pluralidad de ácidos nucleicos polimórficos se localiza en una pluralidad de cromosomas diferentes. En una realización, la pluralidad de ácidos nucleicos polimórficos puede localizarse en los cromosomas 1-22. Por ejemplo, la pluralidad de ácidos nucleicos polimórficos se localiza en una pluralidad de cromosomas diferentes distintos de los cromosomas 13, 18, 21, X o Y. El método comprende además resolver el tamaño de las STR usando electroforesis capilar. La al menos una STR está seleccionada de CSF1PO, FGA, TH01, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D2S1338, Penta D, Penta E, D22S1045, D20S1082, D20S482, D18S853, D17S1301, D17S974, D14S1434, D12ATA63, D11S4463, D10S1435, D10S1248, D9S2157, D9S1122, D8S1115, D6S1017, D6S474, D5S2500, D4S2408, D4S2364, D3S4529, D3S3053, D2S1776, D2S441, D1S1677, D1S1627 y D1GATA113. Opcionalmente, la al menos una STR puede ser el panel de STR que comprende CSF1PO, D13S317, D16S539, D18S51, D21S11, D2S1338, D7S820 y FGA.

En otra realización, en el presente documento se desvela un método de determinación de la fracción de ácidos nucleicos fetales en una muestra de plasma materno que comprende una mezcla de ADN fetal y genómico, en la que el método comprende las etapas de: (a) amplificar una pluralidad de ácidos nucleicos polimórficos en dicha mezcla de ácidos nucleicos fetales y maternos, en la que cada uno de dicho al menos un ácido nucleico polimórfico comprende una STR; (b) determinar la cantidad de alelos de STR fetales y maternos de al menos un ácido nucleico polimórfico; y (c) determinar dicha fracción usando dicha cantidad de alelos de STR fetales y maternos. La pluralidad de ácidos nucleicos polimórficos se localiza en una pluralidad de cromosomas diferentes. En una realización, la pluralidad de ácidos nucleicos polimórficos puede localizarse en los cromosomas 1-22. Por ejemplo, la pluralidad de ácidos nucleicos polimórficos puede localizarse en una pluralidad de cromosomas diferentes distintos de los cromosomas 13, 18, 21, X o Y. El método comprende además preamplificar la mezcla de ácidos nucleicos fetales y maternos, y resolver el tamaño de las STR usando electroforesis capilar. La al menos una STR está seleccionada de CSF1PO, FGA, TH01, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D2S1338, Penta D, Penta E, D22S1045, D20S1082, D20S482, D18S853, D17S1301, D17S974, D14S1434, D12ATA63, D11S4463, D10S1435, D10S1248, D9S2157, D9S1122, D8S1115, D6S1017, D6S474, D5S2500, D4S2408, D4S2364, D3S4529, D3S3053, D2S1776, D2S441, D1S1677, D1S1627 y D1GATA113. Opcionalmente, la al menos una STR puede ser el panel de STR que comprende CSF1PO, D13S317, D16S539, D18S51, D21S11, D2S1338, D7S820 y FGA.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

Las novedosas características de la invención se exponen con particularidad en las reivindicaciones adjuntas. Se obtendrá un mejor entendimiento de las características y ventajas de la presente invención por referencia a la siguiente descripción detallada que expone realizaciones ilustrativas, en las que se utilizan los principios de la invención, y los dibujos adjuntos de los que:

La **Figura 1** es un diagrama de flujo de un método 100 para determinar la fracción fetal en una muestra materna de prueba que comprende una mezcla de ácidos nucleicos fetales y maternos usando métodos de secuenciación masivamente paralela o separación por tamaño de secuencias de ácidos nucleicos polimórficas.

La **Figura 2** es un diagrama de barras que muestra la identificación de secuencias polimórficas fetales y maternas (SNP) usadas para determinar la fracción fetal en una muestra de prueba. Se muestran el número total de lecturas de secuencia (Eje y) mapeadas en las secuencias de SNP identificadas por números rs

(Eje x) y el nivel relativo de ácidos nucleicos fetales (*).

La **Figura 3** es un diagrama de flujo que explica resumidamente realizaciones alternativas del método de determinación de la fracción fetal en la secuenciación masivamente paralela mostrada en la Figura 1.

La **Figura 4** ilustra marcadores de STR usados en el kit de amplificación por PCR AmpF!STR® Identifier®.

La **Figura 5** ilustra marcadores de STR usados en el kit de amplificación por PCR AmpF!STR® MiniFiler®.

La **Figura 6** ilustra la correlación de la fracción fetal determinada en la separación por tamaño por secuenciación masivamente paralela de secuencias polimórficas que comprenden SNP y STR.

La **Figura 7** ilustra una realización de uso de la fracción fetal para determinar umbrales de corte para la detección de aneuploidías.

DESCRIPCIÓN DETALLADA DE LA INVENCION

La invención se refiere a métodos de determinación de la fracción de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de ácidos nucleicos fetales y maternos. La fracción de ácidos nucleicos fetales pueden usarse en determinar la presencia o ausencia de aneuploidía fetal.

A menos que se indique lo contrario, la práctica de la presente invención implica técnicas convencionales comúnmente usadas en los campos de la biología molecular, microbiología, purificación de proteínas, ingeniería de proteínas, secuenciación de proteína y de ADN, y de ADN recombinante, que están dentro de la experiencia de la materia. Tales técnicas son conocidas para aquellos expertos en la materia y se describen en numerosos textos y trabajos de referencia estándar.

Los intervalos numéricos son inclusivos de los números que definen el intervalo. Se pretende que cada limitación numérica máxima dada en toda esta memoria descriptiva incluya cada limitación numérica inferior, como si tales limitaciones numéricas inferiores estuvieran expresamente escritas en el presente documento. Cada limitación numérica mínima dada en toda esta memoria descriptiva incluirá cada limitación numérica superior, como si tales limitaciones numéricas superiores estuvieran expresamente escritas en el presente documento. Cada intervalo numérico dado en toda esta memoria descriptiva incluirá cada intervalo numérico más estrecho que entra dentro de tal intervalo numérico más ancho, como si tales intervalos numéricos más estrechos estuvieran todos expresamente escritos en el presente documento.

Los títulos provistos en el presente documento no son limitaciones de los diversos aspectos o realizaciones de la invención que pueden ser tomados por referencia a la memoria descriptiva en conjunto. Por consiguiente, como se indica anteriormente, los términos definidos inmediatamente a continuación se definen más completamente por referencia a la memoria descriptiva en conjunto.

A menos que se defina de otro modo en el presente documento, todos los términos técnicos y científicos usados en el presente documento tienen el mismo significado que comúnmente es entendido por un experto habitual en la materia a la que pertenece la presente invención. Diversos diccionarios científicos que incluyen los términos incluidos en el presente documento son muy conocidos y están disponibles para aquellos en la materia. Aunque cualquier método y material similar o equivalente a aquellos descritos en el presente documento encuentre uso en la práctica o prueba de la presente invención, se describen algunos métodos preferidos y materiales. Por consiguiente, los términos definidos inmediatamente a continuación se describen más completamente por referencia a la memoria descriptiva en conjunto. Debe entenderse que la presente invención no se limita a la metodología particular, protocolos y reactivos descritos, ya que éstos pueden variar, dependiendo del contexto en el que se usen por aquellos expertos en la materia.

DEFINICIONES

Como se usa en el presente documento, los términos en singular “un”, “una”, “el” y “la” incluyen la referencia en plural, a menos que el contexto indique claramente de otro modo. A menos que se indique lo contrario, los ácidos nucleicos se escriben de izquierda a derecha en la orientación de 5' a 3' y las secuencias de aminoácidos se escriben de izquierda a derecha en la orientación de amino a carboxi, respectivamente.

Como se usa en el presente documento, los términos en singular “un”, “una”, “el” y “la” incluyen la referencia en plural, a menos que el contexto indique claramente de otro modo. A menos que se indique lo contrario, los ácidos nucleicos se escriben de izquierda a derecha en la orientación de 5' a 3' y las secuencias de aminoácidos se escriben de izquierda a derecha en la orientación de amino a carboxi, respectivamente.

El término “porción”, cuando se usa en referencia a la cantidad de información de secuencia de moléculas de ácidos nucleicos fetales y maternos en una muestra biológica en el presente documento, se refiere a la cantidad de información de secuencia de moléculas de ácidos nucleicos fetales y maternos en una muestra biológica que en suma equivale a menos de la información de secuencia de <1 genoma humano.

Los términos “polinucleótido”, “ácido nucleico” y “moléculas de ácidos nucleicos” se usan indistintamente y se refieren a una secuencia de nucleótidos covalentemente enlazada (es decir, ribonucleótidos para ARN y

desoxirribonucleótidos para ADN) en la que la posición 3' de la pentosa de un nucleótido está unida por un grupo fosfodiéster a la posición 5' de la pentosa del siguiente, incluyen secuencias de cualquier forma de ácido nucleico, que incluyen, pero no se limitan a, moléculas de ARN, ADN y ADNlc. El término "polinucleótido" incluye, sin limitación, polinucleótido mono y bicatenario.

El término "variación del número de copias" en el presente documento se refiere a la variación en el número de copias de una secuencia de ácidos nucleicos que tiene 1 kb o más presente en una muestra de prueba en comparación con el número de copias de la secuencia de ácidos nucleicos presente en una muestra clasificada. Una "variante del número de copias" se refiere a la secuencia de 1 kb o mayor de ácido nucleico en la que las diferencias en el número de copias se encuentran por comparación de una secuencia de interés en la muestra de prueba con la presente en una muestra clasificada. Las variantes/variaciones del número de copias incluyen deleciones, que incluyen microdeleciones, inserciones, que incluyen microinserciones, duplicaciones, multiplicaciones, inversiones, translocalizaciones y variantes multisitio complejas. Las CNV engloban aneuploidías cromosómicas y aneuploidías parciales.

Como se usa en el presente documento, el término "fracción fetal" se usa indistintamente con "fracción de ácido nucleico fetal", que se refiere a la fracción de ácido nucleico fetal en una muestra que comprende ácido nucleico fetal y materno. Similarmente, el término "fracción secundaria" o "componente secundario" en el presente documento se refiere a la menor fracción del material genético total que está presente en una muestra que contiene material genético derivado de fuentes separadas, por ejemplo, individuos.

Como se usa en el presente documento, el término "alelo" se refiere a una forma específica de una secuencia genética (tal como un gen) dentro de una célula, una muestra, un individuo o dentro de una población, diferenciándose la forma específica de otras formas del mismo gen en la secuencia de al menos uno, y frecuentemente más de uno, sitio de variante dentro de la secuencia del gen. Las secuencias en estos sitios de variante que se diferencian entre diferentes alelos se llaman "variantes", "polimorfismos" o "mutaciones". En general, se usa polimorfismo para referirse a variantes que tienen una frecuencia de al menos el 1 % en una población, mientras que el término mutación se usa generalmente para variantes que se producen a una frecuencia de menos del 1 % en una población. En organismos diploides tales como los seres humanos, en cada localización cromosómica específica autosómica o "locus" un individuo posee dos alelos, un primero heredado de un progenitor y un segundo heredado del otro progenitor, por ejemplo, uno de la madre y uno del padre. Un individuo es "heterocigoto" en un locus si tiene dos alelos diferentes en el locus. Un individuo es "homocigoto" en un locus si tiene dos alelos idénticos en ese locus.

El término "enriquecer" en el presente documento se refiere al proceso de amplificar ácidos nucleicos diana polimórficos contenidos en una porción de una muestra materna, y combinar el producto amplificado con el resto de la muestra materna de la que se eliminó la porción.

Como se usa en el presente documento, el término "genotipar" se refiere a la determinación de la información genética que un individuo lleva en una o más posiciones en el genoma. Por ejemplo, el genotipado puede comprender la determinación de qué alelo o alelos lleva un individuo para un SNP único o la determinación de qué alelo o alelos lleva un individuo para una pluralidad de SNP. Por ejemplo, un nucleótido particular en un genoma puede ser una T en algunos individuos y una C en otros individuos. Aquellos individuos que tienen una T en la posición tienen el alelo T y aquellos que tienen una C tienen el alelo C. En un organismo diploide, el individuo tendrá dos copias de la secuencia que contiene la posición polimórfica, de manera que el individuo puede tener un alelo T y un alelo C o alternativamente dos copias del alelo T o dos copias del alelo C. Aquellos individuos que tienen dos copias del alelo C son homocigotos para el alelo C, aquellos individuos que tienen dos copias del alelo T son homocigotos para el alelo T, y aquellos individuos que tienen una copia de cada alelo son heterocigotos. Los alelos se denominan frecuentemente el alelo A, frecuentemente el alelo principal, y el alelo B, frecuentemente el alelo secundario. Los genotipos pueden ser AA (homocigoto A), BB (homocigoto B) o AB (heterocigoto). Los métodos de genotipado generalmente proporcionan la identificación de la muestra como AA, BB o AB.

Como se usa en el presente documento, el término "cromosoma" se refiere al portador del gen que lleva la herencia de una célula viva que se deriva de cromatina y que comprende ADN y componentes de proteína (especialmente histonas). En el presente documento se emplea el sistema de numeración de cromosomas del genoma humano individual internacionalmente reconocido convencional. El tamaño de un cromosoma individual puede variar de un tipo a otro con un genoma multi-cromosómico dado y de un genoma a otro. En el caso del genoma humano, la masa del ADN entero de un cromosoma dado es normalmente superior a aproximadamente 100.000.000 pb. Por ejemplo, el tamaño del genoma humano entero es aproximadamente 3×10^9 pb. El cromosoma más grande, el cromosoma nº 1, contiene aproximadamente $2,4 \times 10^8$ mientras que el cromosoma más pequeño, el cromosoma nº 22, contiene aproximadamente $5,3 \times 10^7$ pb.

El término "aneuploidía" en el presente documento se refiere a la aparición de uno o más cromosomas adicionales o ausentes.

Como se usa en el presente documento, el término "región cromosómica" es una porción de un cromosoma. El

tamaño o grado físico actual de cualquier región cromosómica individual puede variar enormemente. El término "región" no es necesariamente definitivo de uno o más genes particulares, debido a que una región no necesita tener específicamente en cuenta los segmentos codificantes particulares (exones) de un gen individual.

5 Como se usa en el presente documento, el término "marcador genético" se refiere a una secuencia de ADN que tiene una localización específica en un cromosoma que puede medirse en un laboratorio. El término "marcador genético" también puede usarse para referirse a, por ejemplo, un ADNc y/o un ARNm codificado por una secuencia genómica, además de a esa secuencia genómica. Para ser útil, un marcador necesita tener dos o más alelos o variantes. Los marcadores pueden ser tanto directos, es decir, estar localizados dentro del gen o locus de interés (es decir, gen candidato), como indirectos, es decir, estrechamente asociados al gen o locus de interés (supuestamente debido a una localización que está próxima a, pero no dentro del gen o locus de interés). Además, los marcadores también pueden incluir secuencias que tanto sí como no modifican la secuencia de aminoácidos de un gen.

15 Como se usa en el presente documento, el término "muestra materna" se refiere a una muestra biológica obtenida de un sujeto embarazado, y comprende una mezcla de ácidos nucleicos fetales y maternos. Un "sujeto embarazado" no se limita a un ser humano, sino que también puede incluir otros organismos que incluyen, pero no se limitan a, mamíferos, plantas, bacterias o células derivadas de cualquiera de los anteriores.

20 El término "amplificación del genoma completo" o "WGA", como se usa en el presente documento, generalmente se refiere a un método de amplificación de una muestra de ADN limitada de una manera no específica, con el fin de generar una nueva muestra que es indistinguible de la original, pero con una mayor concentración de ADN. La técnica de amplificación del genoma completo ideal amplificaría una muestra hasta un nivel de microgramo, mientras que mantiene la representación de secuencia original. El ADN de la muestra puede incluir un genoma entero o una porción del mismo. La PCR cebada con oligonucleótidos degenerados (DOP), la técnica de PCR de extensión con cebadores (PEP) que incluye preamplificación de la extensión de cebadores mejorada modificada (miPEP), y la amplificación de múltiples desplazamientos (MDA), son ejemplos de métodos de amplificación del genoma completo.

25 El término "repetición corta en tándem" o "STR", como se usa en el presente documento, se refiere a una clase de polimorfismos que se produce cuando se repite un patrón de dos o más nucleótidos y las secuencias repetidas están directamente adyacentes entre sí. El patrón puede oscilar en longitud de 2 a 10 pares de bases (pb) (por ejemplo (CATG)_n en una región genómica) y normalmente está en la región de intrón no codificante. Examinando varios loci de STR y contando cuántas repeticiones de una secuencia de STR específica hay en un locus dado es posible crear un perfil genético único en un individuo.

30 El término "cebador", como se usa en el presente documento, se refiere a un oligonucleótido aislado que es capaz de actuar de punto de iniciación de la síntesis cuando se pone en condiciones en las que se induce la síntesis de un producto de extensión del cebador que es complementario a una hebra de ácido nucleico (es decir, en presencia de nucleótidos y un agente inductor tal como ADN polimerasa y a una temperatura y pH adecuados). El cebador es preferentemente monocatenario para la máxima eficiencia en la amplificación, pero alternativamente puede ser bicatenario. Si es bicatenario, el cebador se trata primero para separar sus hebras antes de usarse para preparar productos de extensión. Preferentemente, el cebador es un oligodesoxirribonucleótido. El cebador debe ser suficientemente largo para cebar la síntesis de productos de extensión en presencia del agente inductor. Las longitudes exactas de los cebadores dependerán de muchos factores, que incluyen temperatura, fuente del cebador, uso del método y los parámetros usados para el diseño del cebador, como se desvela en el presente documento.

35 El término "par de cebadores" o "conjunto de cebadores" se refiere a un conjunto de cebadores que incluyen un "cebador aguas arriba" 5' o "cebador directo" que se hibrida con el complemento del extremo 5' de la secuencia de ADN que va a amplificarse y un "cebador aguas abajo" 3' o "cebador inverso" que se hibrida con el extremo 3' de la secuencia que va a amplificarse. Como se reconocerá por aquellos expertos en la materia, los términos "aguas arriba" y "aguas abajo" o "directo" e "inverso" no pretenden ser limitantes, sino que proporcionan orientación ilustrativa en realizaciones particulares. Se dice que un par de cebadores es "único" si puede emplearse para amplificar específicamente una secuencia de nucleótidos diana particular en una mezcla de amplificación dada.

40 El término "marcador polimórfico" o "sitio polimórfico" es un locus en el que se produce divergencia de la secuencia de nucleótidos. El locus puede ser tan pequeño como una par de bases. Marcadores ilustrativos tienen al menos dos alelos, cada uno de los cuales se produce a una frecuencia superior al 1 %, y más normalmente superior al 10 % o 20 % de una población seleccionada. Un sitio polimórfico puede ser tan pequeño como una par de bases. Los marcadores polimórficos incluyen polimorfismo en la longitud del fragmento de restricción (RFLP), número variable de repeticiones en tándem (VNTR), regiones hipervariables, minisatélites, repeticiones de dinucleótidos, repeticiones de trinucleótidos, repeticiones de tetranucleótidos, repeticiones de secuencias simples, deleciones y elementos de inserción tales como Alu. La primera forma alélica identificada se designa arbitrariamente la forma de referencia y otras formas alélicas se designan alelos alternativos o de variante. La forma alélica que se produce más frecuentemente en una población seleccionada se denomina algunas veces la forma no mutante. Los organismos diploides pueden ser homocigotos o heterocigotos para las formas alélicas. Un polimorfismo dialélico tiene dos formas. Un polimorfismo trialélico tiene tres formas. Un polimorfismo entre dos ácidos nucleicos puede producirse naturalmente, o producirse por exposición a o contacto con productos químicos, enzimas u otros agentes, o

exposición a agentes que producen daño a los ácidos nucleicos, por ejemplo, radiación ultravioleta, mutágenos o carcinógenos. Los términos "locus polimórfico" y "sitio polimórfico" se usan en el presente documento indistintamente.

5 Los términos "ácido nucleico diana polimórfico", "secuencia polimórfica", "secuencia de ácidos nucleicos diana polimórfica" y "ácido nucleico polimórfico" se usan indistintamente en el presente documento para referirse a una secuencia de ácidos nucleicos, por ejemplo, una secuencia de ADN, que comprende uno o más sitios polimórficos, por ejemplo, un SNP o un SNP en tándem. Las secuencias polimórficas según la presente tecnología pueden usarse para diferenciar específicamente entre alelos maternos y no maternos en la muestra materna que comprende una
10 mezcla de ácidos nucleicos fetales y maternos.

Un "polimorfismo de un único nucleótido" (SNP) se produce en un sitio polimórfico ocupado por un único nucleótido, que es el sitio de variación entre las secuencias alélicas. El sitio está normalmente precedido por y seguido de secuencias altamente conservadas del alelo (por ejemplo, secuencias que varían en menos de 1/100 o 1/1000 de miembros de las poblaciones). Un SNP normalmente se produce debido a la sustitución de un nucleótido por otro en el sitio polimórfico. Una transición es la sustitución de una purina por otra purina o una pirimidina por otra pirimidina. Una transversión es la sustitución de una purina por una pirimidina o viceversa. Los SNP también pueden surgir de una delección de un nucleótido o una inserción de un nucleótido con respecto a un alelo de referencia. Los polimorfismos de un único nucleótido (SNP) son posiciones en las que dos bases alternativas se producen a una frecuencia apreciable (>1%) en la población humana, y son el tipo más común de variación genética humana.
15
20

Como se usa en el presente documento, el término "repetición corta en tándem" o "STR", como se usa en el presente documento, se refiere a una clase de polimorfismos que se produce cuando se repite un patrón de dos o más nucleótidos y las secuencias repetidas están directamente adyacentes entre sí. El patrón puede oscilar en longitud de 2 a 10 pares de bases (pb) (por ejemplo (CATG)_n en una región genómica) y normalmente está en la región de intrón no codificante. Examinando varios loci de STR y contando cuántas repeticiones de una secuencia de STR específica hay en un locus dado es posible crear un perfil genético único de un individuo.
25

Como se usa en el presente documento, el término "miniSTR" en el presente documento se refiere a la repetición en tándem de cuatro o más pares de bases que abarcan menos de aproximadamente 300 pares de bases, menos de aproximadamente 250 pares de bases, menos de aproximadamente 200 pares de bases, menos de aproximadamente 150 pares de bases, menos de aproximadamente 100 pares de bases, menos de aproximadamente 50 pares de bases, o menos de aproximadamente 25 pares de bases. Las "miniSTR" son STR que son amplificables a partir de moldes de ADNlc.
30
35

El término "SNP en tándem" en el presente documento se refiere a dos o más SNP que están presentes dentro de una secuencia de ácidos nucleicos diana polimórfica.

Los términos "pluralidad de ácidos nucleicos diana polimórficos", "ácidos nucleicos polimórficos" y "secuencias polimórficas" se usan indistintamente en el presente documento y se refieren a varias secuencias de ácidos nucleicos que comprenden cada una al menos un sitio polimórfico, por ejemplo, un SNP, tal que al menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30 o 40 o más sitios polimórficos diferentes se amplifiquen a partir de los ácidos nucleicos diana polimórficos para identificar y/o cuantificar alelos presentes en muestras maternas que comprenden ácidos nucleicos fetales y maternos.
40
45

Como se usa en el presente documento, el término "sustancialmente libre de células" engloba preparaciones de la muestra deseada de la que se eliminan componentes que normalmente están asociados a ella. Por ejemplo, una muestra de plasma se vuelve esencialmente libre de células eliminando glóbulos sanguíneos, por ejemplo, glóbulos rojos, que normalmente están asociados a ella. En algunas realizaciones, muestras sustancialmente libres se procesan para eliminar células que de otro modo contribuirían al material genético deseado que va a probarse para una anomalía.
50

Como se usa en el presente documento, el término "cromosoma" se refiere al portador del gen que lleva la herencia de una célula viva que se deriva de cromatina y que comprende ADN y componentes de proteína (especialmente histonas). En el presente documento se emplea el sistema de numeración de cromosomas del genoma humano individual internacionalmente reconocido convencional. El tamaño de un cromosoma individual puede variar de un tipo a otro con un genoma multi-cromosómico dado y de un genoma a otro. En el caso del genoma humano, la masa del ADN entero de un cromosoma dado es normalmente superior a aproximadamente 100.000.000 pb. Por ejemplo, el tamaño del genoma humano entero es aproximadamente 3×10^9 pb. El cromosoma más grande, el cromosoma nº 1, contiene aproximadamente $2,4 \times 10^8$ mientras que el cromosoma más pequeño, el cromosoma nº 22, contiene aproximadamente $5,3 \times 10^7$ pb.
55
60

El término "oligonucleótido" se usa para referirse a un ácido nucleico que es relativamente corto, generalmente más corto de 200 nucleótidos, más particularmente, más corto de 100 nucleótidos, lo más particularmente, más corto de 50 nucleótidos. Normalmente, los oligonucleótidos son moléculas de ADN monocatenario.
65

El término “cebador” se refiere a un oligonucleótido que es capaz de hibridarse con un ácido nucleico y servir de sitio de iniciación para la polimerización del nucleótido (ARN o ADN) bajo condiciones apropiadas (es decir, en presencia de cuatro trifosfatos de nucleósido diferentes y un agente para la polimerización, tal como ADN o ARN polimerasa o transcriptasa inversa) en un tampón apropiado y a una temperatura adecuada. La longitud apropiada de un cebador depende del uso previsto del cebador, pero los cebadores normalmente tienen al menos 7 nucleótidos de longitud y, más normalmente oscilan de 10 a 30 nucleótidos, o incluso más normalmente de 15 a 30 nucleótidos, de longitud. Otros cebadores pueden ser algo más largos, por ejemplo, 30 a 50 nucleótidos de longitud.

El término “llamada de alelos”, como se usa en el presente documento, se refiere a la caracterización satisfactoria de un alelo por un método de análisis dado. Si el análisis proporciona caracterización satisfactoria de ambos alelos de un locus de gen de una muestra de ADN, se dice que se han hecho dos llamadas de alelos. Si un alelo está caracterizado, mientras que el otro alelo no está caracterizado, se dice que se hace una llamada de alelos. Si ninguno de los dos alelos está satisfactoriamente caracterizado, no se hacen llamadas de alelos.

El término “alelo”, como se usa en el presente documento, es una cualquiera de varias codificaciones de ADN viables que ocupan un locus dado (posición) en un cromosoma. Normalmente, los alelos son secuencias de ADN (ácido desoxirribonucleico) que codifican un gen, pero algunas veces el término se usa para referirse a una secuencia no de genes. El genotipo de un individuo para ese gen es el conjunto de alelos que resulta poseer. En un organismo diploide, uno que tiene dos copias de cada cromosoma, dos alelos constituyen el genotipo del individuo.

El término “mezcla de reacción”, como se usa en el presente documento, se refiere a una mezcla que contiene componentes suficientes para llevar a cabo una reacción de amplificación.

El término “densidad de marcas de secuencia” en el presente documento se refiere al número de lecturas de secuencia que son mapeadas en una secuencia de genoma de referencia, por ejemplo, la densidad de marcas de secuencia para el cromosoma 21 es el número de lecturas de secuencia generadas por el método de secuenciación que son mapeadas en el cromosoma 21 del genoma de referencia. El término “relación de densidades de las marcas de secuencia” en el presente documento se refiere a la relación del número de marcas de secuencia que son mapeadas en un cromosoma del genoma de referencia, por ejemplo, el cromosoma 21, con respecto a la longitud del cromosoma del genoma de referencia 21.

Los términos “valor umbral” y “valor umbral clasificado” en el presente documento se refieren a cualquier número que se calcula usando un conjunto de datos de clasificación y sirve de límite de diagnóstico de una variación del número de copias, por ejemplo, una aneuploidía, en un organismo. Si se supera un umbral por resultados obtenidos de la práctica de la invención, un sujeto puede diagnosticarse con una variación del número de copias, por ejemplo, trisomía 21.

El término “lectura” se refiere a una secuencia de ADN de longitud suficiente (por ejemplo, al menos aproximadamente 30 pb) que puede usarse para identificar una secuencia o región mayor, por ejemplo, que puede alinearse y asignarse específicamente a un cromosoma o región genómica o gen.

El término “marca de secuencia” se usa en el presente documento indistintamente con el término “marca de secuencia mapeada” para referirse a una lectura de secuencia que se ha asignado específicamente, es decir, mapeado, con una secuencia mayor, por ejemplo, un genoma de referencia, por alineamiento. Las marcas de secuencia mapeadas se mapean únicamente en un genoma de referencia, es decir, se asignan a una única localización al genoma de referencia. Marcas que pueden ser mapeadas en más de una localización en un genoma de referencia, es decir, marcas que no se mapean únicamente, no se incluyen en el análisis.

Los términos “alineado”, “alineamiento” o “alinearse” se refieren a una o más secuencias que se identifican como una coincidencia en términos del orden de sus moléculas de ácidos nucleicos con una secuencia conocida de un genoma de referencia. Tal alineamiento puede hacerse manualmente o por un algoritmo informático, incluyendo los ejemplos el programa informático Efficient Local Alignment of Nucleotide Data (ELAND) distribuido como parte del canal Illumina Genomics Analysis. La coincidencia de una lectura de secuencia en la alineación puede ser una coincidencia de secuencia del 100 % o inferior al 100 % (coincidencia no perfecta).

El término “genoma de referencia” se refiere a cualquier secuencia del genoma conocida particular, tanto parcial como completa, de cualquier organismo o virus que puede usarse para hacer referencia secuencias identificadas de un sujeto. Por ejemplo, un genoma de referencia usado para sujetos humanos, además de muchos otros organismos, se encuentra en el Centro Nacional para Información Biotecnológica en www.ncbi.nlm.nih.gov. Un “genoma” se refiere a la información genética completa de un organismo o virus, expresado en secuencias de ácidos nucleicos.

El término “genoma de secuencias diana artificial” en el presente documento se refiere a una agrupación de secuencias conocidas que engloban alelos de sitios polimórficos conocidos. Por ejemplo, un “genoma de referencia de SNP” es un genoma de secuencias diana artificial que comprende una agrupación de secuencias que engloban alelos de SNP conocidos.

El término “secuencia clínicamente relevante” en el presente documento se refiere a una secuencia de ácidos nucleicos que se sabe o que se sospecha que va estar asociada o participar en una condición genética o de enfermedad. La determinación de la ausencia o presencia de una secuencia clínicamente relevante puede ser útil en determinar un diagnóstico o confirmar un diagnóstico de una afección médica, o proporcionar un pronóstico para el desarrollo de una enfermedad.

El término “muestra mixta” en el presente documento se refiere a una muestra que contiene una mezcla de ácidos nucleicos, que se derivan de diferentes genomas.

El término “muestra materna original” en el presente documento se refiere a una muestra biológica obtenida de un sujeto embarazado, por ejemplo, una mujer, que sirve de fuente de la que se extrae una porción para amplificar ácidos nucleicos diana polimórficos. La “muestra original” puede ser cualquier muestra obtenida de un sujeto embarazado, y las fracciones procesadas de la misma, por ejemplo, una muestra de ADNlc purificado extraída de una muestra de plasma materno. El término “muestra materna original” en el presente documento se refiere a una muestra biológica obtenida de un sujeto embarazado, por ejemplo, una mujer, que sirve de fuente de la que se extrae una porción para amplificar ácidos nucleicos diana polimórficos. La “muestra original” puede ser cualquier muestra obtenida de un sujeto embarazado, y las fracciones procesadas de la misma, por ejemplo, una muestra de ADNlc purificado extraída de una muestra de plasma materno.

El término “fluido biológico” en el presente documento se refiere a un líquido tomado de una fuente biológica e incluye, por ejemplo, sangre, suero, plasma, esputo, líquido de lavado, líquido cefalorraquídeo, orina, semen, sudor, lágrimas, saliva y similares. Como se usa en el presente documento, los términos “sangre”, “plasma” y “suero” engloban expresamente fracciones o porciones procesadas de los mismos. Similarmente, si se toma una muestra de una biopsia, hisopo, frotis, etc., la “muestra” engloba expresamente una fracción procesada o porción derivada de la biopsia, hisopo, frotis, etc.

Los términos “ácidos nucleicos maternos” y “ácidos nucleicos fetales” en el presente documento se refieren a los ácidos nucleicos de un sujeto femenino embarazado y los ácidos nucleicos del feto que es llevado por la hembra embarazada, respectivamente.

El término “correspondiente a” en el presente documento se refiere a una secuencia de ácidos nucleicos, por ejemplo, un gen o un cromosoma, que está presente en el genoma de diferentes sujetos, y que no tiene necesariamente la misma secuencia en todos los genomas, pero sirve para proporcionar la identidad en vez de la información genética de una secuencia de interés, por ejemplo, un gen o cromosoma.

El término “grupo de cromosomas” en el presente documento se refiere a dos o más cromosomas.

El término “sujeto” en el presente documento se refiere a un sujeto humano, además de un sujeto no humano, tal como un mamífero, un invertebrado, un vertebrado, un hongo, una levadura, una bacteria y un virus. Aunque los ejemplos en el presente documento se refieren a células humanas y el lenguaje se refiere principalmente a asuntos humanos, el concepto de la presente invención es aplicable a genomas de cualquier planta o animal, y es útil en los campos de la veterinaria y medicina, ciencias animales, laboratorios de investigación y demás.

DESCRIPCIÓN

Los métodos descritos en el presente documento permiten la determinación de la fracción del componente de ácido nucleico fetal secundario en una muestra que comprende una mezcla de ácidos nucleicos fetales y maternos. En particular, el método permite la determinación de la fracción de ADNlc contribuida por un feto a la mezcla de ADNlc fetal y materno en una muestra materna, por ejemplo, una muestra de plasma. La diferencia entre la fracción materna y la fetal se determina por la contribución relativa de un alelo polimórfico derivado del genoma fetal a la contribución del alelo polimórfico correspondiente derivado del genoma materno. Pueden usarse secuencias polimórficas conjuntamente con pruebas de diagnóstico clínicamente relevantes como control positivo para la presencia de ADNlc con el fin de recalcar los resultados negativos falsos o positivos falsos que son el resultado de bajos niveles de ADNlc por debajo del límite de identificación. Los métodos descritos son independientes del sexo del feto, y son útiles a través de un intervalo de edades gestacionales.

La **Figura 1** proporciona un diagrama de flujo de una realización del método de la invención **100** para determinar la fracción de ácidos nucleicos fetales en una muestra materna biológica en secuenciación masivamente paralela de ácidos nucleicos diana polimórficos amplificados por PCR. En la etapa **110**, una muestra materna que comprende una mezcla de ácidos nucleicos fetales y maternos se obtiene de un sujeto. La muestra es una muestra materna que se obtiene de una hembra embarazada, por ejemplo, una mujer embarazada. Otras muestras maternas pueden ser de mamíferos, por ejemplo, vaca, caballo, perro o gato. Si el sujeto es humano, la muestra puede tomarse en el primer o segundo trimestre del embarazo. Puede usarse cualquier muestra materna biológica como fuente de ácidos nucleicos fetales y maternos que están contenidos en células o que están “libres de células”. En algunas realizaciones, es ventajoso obtener una muestra materna que comprende ácidos nucleicos libres de células, por ejemplo, ADNlc. Preferentemente, la muestra materna biológica es una muestra de fluido biológico. Un fluido

biológico incluye, como ejemplos no limitantes, sangre, plasma, suero, sudor, lágrimas, esputo, orina, esputo, flujo del oído, linfa, saliva, líquido cefalorraquídeo, lavados, suspensión de médula ósea, flujo vaginal, lavado transcervical, líquido cerebral, ascitis, leche, secreciones de las vías respiratorias, intestinales y genitourinarias, y muestras de leucoforesis. En algunas realizaciones, la muestra de fluido biológico es una muestra que es fácilmente obtenible por procedimientos no invasivos, por ejemplo, sangre, plasma, suero, sudor, lágrimas, esputo, orina, esputo, flujo del oído y saliva. En algunas realizaciones, la muestra biológica es una muestra de sangre periférica, o el plasma y/o las fracciones de suero del mismo. En otra realización, la muestra es una mezcla de dos o más muestras biológicas, por ejemplo, una muestra biológica puede comprender dos o más de una muestra de fluido biológico. Como se usa en el presente documento, los términos “sangre”, “plasma” y “suero” engloban expresamente fracciones o porciones procesadas de los mismos. En algunas realizaciones, la muestra biológica se procesa para obtener una fracción de muestra, por ejemplo, plasma, que contiene la mezcla de ácidos nucleicos fetales y maternos. Una muestra que puede usarse para determinar el genotipo de uno o más alelos fetales puede ser cualquier muestra que contenga células fetales o ácido nucleico fetal. Por ejemplo, puede usarse muestra de suero o plasma materno que comprende ácidos nucleicos libres de células fetales y maternas (por ejemplo, ADN o ARN) para determinar los genotipos de alelos fetales. En una realización, la muestra puede comprender una célula fetal, por ejemplo, un glóbulo rojo nucleado fetal o un trofoblasto.

En la etapa **120**, la mezcla de ácidos nucleicos fetales y maternos se procesa adicionalmente a partir de la fracción de muestra, por ejemplo, plasma, para obtener una muestra que comprende una mezcla purificada de ácidos nucleicos fetales y maternos, por ejemplo, ADNlc. Los ácidos nucleicos libres de células, que incluyen ADN libre de células, pueden obtenerse por diversos métodos conocidos en la técnica a partir de muestras biológicas que incluyen, pero no se limitan a, plasma, suero y orina (Fan et al., Proc Natl Acad Sci 105:16266-16271 [2008]; Koide et al., Prenatal Diagnosis 25:604-607 [2005]; Chen et al., Nature Med. 2: 1033-1035 [1996]; Lo et al., Lancet 350: 485-487 [1997]). Para separar el ADNlc de las células, pueden usarse fraccionamiento, centrifugación (por ejemplo, centrifugación en gradiente de densidad), precipitación específica de ADN o citometría de alta resolución y/o métodos de separación. Ejemplos de métodos para procesar muestras de fluido se han desvelado previamente, por ejemplo, las solicitudes de patente de EE.UU. Nº 20050282293, 20050224351 y 20050065735. Están disponibles kits comercialmente disponibles para la separación manual y automática de ADNlc (Roche Diagnostics, Indianápolis, IN, Qiagen, Valencia, CA, Macherey-Nagel, Duren, DE). En algunos casos, puede ser ventajoso fragmentar las moléculas de ácidos nucleicos en la muestra de ácido nucleico. La fragmentación puede ser aleatoria, o puede ser específica, como se consigue, por ejemplo, usando digestión con endonucleasas de restricción. Métodos para la fragmentación aleatoria son muy conocidos en la técnica e incluyen, por ejemplo, digestión limitada con DNasa, tratamiento con álcalis y cizallamiento físico. En una realización, los ácidos nucleicos de la muestra se obtienen como ADNlc, que no se somete a fragmentación. En otras realizaciones, los ácidos nucleicos de la muestra se obtienen como ADN genómico, que se somete a fragmentación en fragmentos de aproximadamente 500 o más pares de bases, y a los que pueden aplicársele fácilmente métodos de NGS.

En la etapa **130**, una porción de la mezcla purificada de ADNlc fetal y materno se usa para amplificar una pluralidad de ácidos nucleicos diana polimórficos que comprende cada uno un sitio polimórfico. En algunas realizaciones, los ácidos nucleicos diana comprenden cada uno un SNP. En otras realizaciones, cada uno de los ácidos nucleicos diana comprende un par de SNP en tándem. En todavía otras realizaciones, cada uno de los ácidos nucleicos diana comprende una STR. Los sitios polimórficos que están contenidos en los ácidos nucleicos diana incluyen, sin limitación, polimorfismos de un único nucleótido (SNP), SNP en tándem, delecciones o inserciones de múltiples bases a pequeña escala, llamados IN-DELS (también llamados polimorfismos de inserción/delección o DIP), polimorfismos de múltiples nucleótidos (MNP), repeticiones cortas en tándem (STR), polimorfismo en la longitud del fragmento de restricción (RFLP), o un polimorfismo que comprende cualquier otro cambio de secuencia en un cromosoma. En algunas realizaciones, los sitios polimórficos que están englobados por el método de la invención se localizan en cromosomas autosómicos, permitiendo así la determinación de la fracción fetal independientemente del sexo del feto. También pueden usarse polimorfismos asociados a cromosomas distintos del cromosoma 13, 18, 21 e Y en los métodos descritos en el presente documento.

Los polimorfismos pueden ser indicativos, informativos, o ambos. Los polimorfismos indicativos indican la presencia de ADN libre de células fetal en una muestra materna. Por ejemplo, cuanto más haya de una secuencia genética particular, por ejemplo, un SNP, más traducirá un método su presencia en una intensidad de color particular, densidad de color, o alguna otra propiedad que es detectable y medible e indicativa de la presencia, ausencia y cantidad de un fragmento particular de ADN y/o polimorfismo particular, por ejemplo, SNP del embrión. Los polimorfismos informativos dan información sobre el feto - por ejemplo, la presencia o ausencia de una enfermedad, anomalía genética, o cualquier otra información biológica tal como el estado de gestación o sexo. Con respecto a la presente invención, los métodos no se realizan usando todos los posibles SNP en un genoma, sino que se usan aquellos que se dice que son “informativos”. Los “SNP informativos” en este caso son aquellos que identifican diferencias en la secuencia de la madre y el feto. Puede usarse cualquier sitio polimórfico que pueda estar englobado por las lecturas generadas por los métodos de secuenciación descritos en el presente para determinar la fracción fetal.

En una realización, una porción de la mezcla de ácidos nucleicos fetales y maternos en la muestra, por ejemplo, ADNlc, se usa como molde para amplificar ácidos nucleicos diana que comprenden al menos un SNP. En algunas

realizaciones, cada ácido nucleico diana comprende un único SNP, es decir, un SNP. Están disponibles secuencias de ácidos nucleicos diana que comprenden SNP de bases de datos públicamente accesibles que incluyen, pero no se limitan a, la base de datos Human SNP en la dirección de la malla mundial wi.mit.edu, NCBI dbSNP Home Page en la dirección de la malla mundial ncbi.nlm.nih.gov, la dirección de la malla mundial lifesciences.perkinelmer.com, Applied Biosystems por Life Technologies™ (Carlsbad, CA) en la dirección de la malla mundial appliedbiosystems.com, la base de datos Celera Human SNP en la dirección de la malla mundial celera.com, la base de datos de SNP del Genome Analysis Group (GAN) en la dirección de la malla mundial gan.iarc.fr. En una realización, los SNP elegidos para enriquecer el ADN fetal y materno están seleccionados del grupo de 92 SNP de identificación individual (IISNP) descritos por Pakstis et al. (Pakstis et al. *Hum Genet* 127:315-324 [2010]), que se ha mostrado que tienen una variación muy pequeña en la frecuencia a través de las poblaciones ($F_{st} < 0,06$), y que son altamente informativos en todo el mundo teniendo una heterocigosidad promedio $\geq 0,4$. Los SNP que están englobados por el método de la invención incluyen SNP enlazados y sin enlazar. Otros SNP útiles aplicables o útiles para los métodos descritos en el presente documento se desvelan en las solicitudes de patente de EE.UU. N° 20080070792, 20090280492, 20080113358, 20080026390, 20080050739, 20080220422 y 20080138809. Cada ácido nucleico diana comprende al menos un sitio polimórfico, por ejemplo, un único SNP, que se diferencia de ese presente en otro ácido nucleico diana para generar un panel de sitios polimórficos, por ejemplo, SNP, que contienen un número suficiente de sitios polimórficos de los que al menos 1, al menos 2, al menos 3, al menos 4, al menos 5, al menos 6, al menos 7, al menos 8, al menos 9, al menos 10, al menos 11, al menos 12, al menos 13, al menos 14, al menos 15, al menos 16, al menos 17, al menos 18, al menos 19, al menos 20, al menos 25, al menos 30, al menos 40 o más son informativos. Por ejemplo, puede configurarse un panel de SNP que comprende al menos un SNP informativo. En una realización, los SNP que son dirigidos para la amplificación están seleccionados de rs560681, rs1109037, rs9866013, rs13182883, rs13218440, rs7041158, rs740598, rs10773760, rs4530059, rs7205345, rs8078417, rs576261, rs2567608, rs430046, rs9951171, rs338882, rs10776839, rs9905977, rs1277284, rs258684, rs1347696, rs508485, rs9788670, rs8137254, rs3143, rs2182957, rs3739005 y rs530022. En una realización, el panel de SNP comprende al menos 3, al menos 5, al menos 10, al menos 13, al menos 15, al menos 20, al menos 25, al menos 30 o más SNP. En una realización, el panel de SNP comprende rs560681, rs1109037, rs9866013, rs13182883, rs13218440, rs7041158, rs740598, rs10773760, rs4530059, rs7205345, rs8078417, rs576261 y rs2567608. Los ácidos nucleicos polimórficos que comprenden los SNP pueden amplificarse usando pares de cebadores a modo de ejemplo proporcionados en el Ejemplo 3, y desvelados como SEC ID N°: 57-112.

En otras realizaciones, cada ácido nucleico diana comprende dos o más SNP, es decir, cada ácido nucleico diana comprende SNP en tándem. Preferentemente, cada ácido nucleico diana comprende dos SNP en tándem. Los SNP en tándem se analizan como una única unidad como haplotipos cortos, y se proporcionan en el presente documento como conjuntos de dos SNP. Para identificar secuencias de SNP en tándem adecuadas puede buscarse en la base de datos International HapMap Consortium (The International HapMap Project, *Nature* 426:789-796 [2003]). La base de datos está disponible en la malla mundial en hapmap.org. En una realización, los SNP en tándem que son dirigidos para amplificación están seleccionados de los siguientes conjuntos de pares en tándem de SNP rs7277033-rs2110153; rs2822654-rs1882882; rs368657-rs376635; rs2822731-rs2822732; rs1475881-rs7275487; rs1735976-rs2827016; rs447340-rs2824097; rs418989-rs13047336; rs987980-rs987981; rs4143392-rs4143391; rs1691324-rs13050434; rs11909758-rs9980111; rs2826842-rs232414; rs1980969-rs1980970; rs9978999-rs9979175; rs1034346-rs12481852; rs7509629-rs2828358; rs4817013-rs7277036; rs9981121-rs2829696; rs455921-rs2898102; rs2898102-rs458848; rs961301-rs2830208; rs2174536-rs458076; rs11088023-rs11088024; rs1011734-rs1011733; rs2831244-rs9789838; rs8132769-rs2831440; rs8134080-rs2831524; rs4817219-rs4817220; rs2250911-rs2250997; rs2831899-rs2831900; rs2831902-rs2831903; rs11088086-rs2251447; rs2832040-rs11088088; rs2832141-rs2246777; rs2832959-rs9980934; rs2833734-rs2833735; rs933121-rs933122; rs2834140-rs12626953; rs2834485-rs3453; rs9974986-rs2834703; rs2776266-rs2835001; rs1984014-rs1984015; rs7281674-rs2835316; rs13047304-rs13047322; rs2835545-rs4816551; rs2835735-rs2835736; rs13047608-rs2835826; rs2836550-rs2212596; rs2836660-rs2836661; rs465612-rs8131220; rs9980072-rs8130031; rs418359-rs2836926; rs7278447-rs7278858; rs385787-rs367001; rs367001-rs386095; rs2837296-rs2837297; y rs2837381-rs4816672. Los ácidos nucleicos polimórficos que comprenden los SNP en tándem pueden amplificarse usando pares de cebadores que amplifican secuencias polimórficas que comprenden los SNP en tándem. Ejemplos de pares de cebadores que pueden usarse para amplificar los SNP en tándem en el presente documento se desvelan en SEC ID N°: 197-310 como se proporciona en el Ejemplo 8.

En una realización, una porción de la mezcla de ácidos nucleicos fetales y maternos en la muestra, por ejemplo, ADN fetal, se usa como molde para amplificar ácidos nucleicos diana que comprenden al menos una STR. En algunas realizaciones, cada ácido nucleico diana comprende una STR única, es decir, una STR. Los loci de STR se encuentran en casi cada cromosoma en el genoma y pueden amplificarse usando una variedad de cebadores de la reacción en cadena de la polimerasa (PCR). Se han preferido repeticiones de tetranucleótidos entre los científicos forenses debido a su fidelidad en la amplificación por PCR, aunque también están en uso algunas repeticiones de tri- y pentanucleótidos. Un listado exhaustivo de referencias, hechos e información de secuencia sobre STR, cebadores de PCR publicados, sistemas de múltiplex comunes y datos de población relacionados está compilado en STRBase, a la que puede accederse mediante la malla mundial en ibm4.carb.nist.gov:8800/DNA/home.htm. La información de secuencias de GenBank® (<http://www2.ncbi.nlm.nih.gov/cgi-bin/genbank>) para loci de STR comúnmente usados también está accesible mediante STRBase. Kits comerciales disponibles para el análisis de loci de STR generalmente proporcionan todos los componentes de reacción necesarios y controles requeridos para la

amplificación. Los sistemas de múltiplex de STR permiten la amplificación simultánea de múltiples loci que no se solapan en una única reacción, aumentando sustancialmente el rendimiento. Con detección fluorescente multicolor, incluso loci que se solapan pueden multiplexarse. La naturaleza polimórfica de las secuencias de ADN repetidas en tándem que están extendidas en todo el genoma humano los ha hecho marcadores genéticos importantes para el estudio del mapeo genético, análisis de enlaces y prueba de la identidad humana. Debido al alto polimorfismo de las STR, la mayoría de los individuos serán heterocigotos, es decir, la mayoría de las personas poseerán dos alelos (versiones), cada uno heredado de cada progenitor - con un número de repeticiones diferente. Los productos de PCR que comprenden las STR pueden separarse y detectarse usando métodos manuales, semi-automatizados o automatizados. Los sistemas semi-automatizados se basan en gel y combinan electroforesis, detección y análisis en una unidad. En un sistema semi-automatizado, el ensamblaje del gel y la carga de muestra son todavía procesos manuales; sin embargo, una vez se cargan las muestras sobre el gel, la electroforesis, detección y análisis proceden automáticamente. La recogida de datos se produce en "tiempo real", ya que los fragmentos fluorescentemente marcados migran hacia el detector en un punto fijo y pueden visualizarse a medida que se recogen. Como su nombre implica, la electroforesis capilar se lleva a cabo en un tubo microcapilar en vez de entre placas de vidrio. Una vez se cargan las muestras, el polímero de gel y el tampón en el instrumento, el capilar se llena del polímero de gel y la muestra se carga automáticamente. Por tanto, la secuencia de STR fetal heredada no maternamente se diferenciará en el número de repeticiones de la secuencia materna. La amplificación de estas secuencias de STR puede producir uno o dos productos de amplificación principales correspondientes a los alelos maternos (y el alelo fetal maternamente heredado) y un producto secundario correspondiente al alelo fetal no maternamente heredado. Esta técnica se informó por primera vez en 2000 (Pertl et al., *Human Genetics* 106:45-49 [2002]) y se ha desarrollado posteriormente usando identificación simultánea de múltiples regiones de STR diferentes usando PCR en tiempo real (Liu et al., *Acta Obstet Gyn Scand* 86:535-541 [2007]). Se han usado amplicones de PCR de diversos tamaños para distinguir las distribuciones de tamaño respectivas de especies de ADN fetal y materno circulante, y han mostrado que las moléculas de ADN fetal en el plasma de mujeres embarazadas son generalmente más cortas que las moléculas de ADN maternas (Chan et al., *Clin Chem* 50:8892 [2004]). El fraccionamiento por tamaño del ADN fetal circulante ha confirmado que la longitud promedio de los fragmentos de ADN fetal que circula es <300 pb, mientras que se ha estimado que el ADN materno tiene entre aproximadamente 0,5 y 1 Kb (Li et al., *Clin Chem*, 50: 1002-1011 [2004]). En el presente documento se desvela un método de determinación de la fracción de ácido nucleico fetal en una muestra materna que comprende determinar la cantidad de copias de al menos un alelo fetal y uno maternal en un sitio miniSTR polimórfico, que puede amplificarse para generar amplicones que son de longitudes de aproximadamente el tamaño de los fragmentos de ADN fetal circulante, por ejemplo, inferior a aproximadamente 250 pares de bases. En una realización, la fracción fetal puede determinarse por un método que comprende la secuenciación de al menos una porción de ácidos nucleicos diana polimórficos amplificados que comprende cada una una miniSTR. Los alelos fetales y maternos en un sitio de STR informativo se distinguen por sus diferentes longitudes, es decir, número de repeticiones, y la fracción fetal puede calcularse como una relación de porcentaje de la cantidad de alelos maternos fetales en ese sitio. El método puede usar uno o una combinación de cualquier número de miniSTR informativas para determinar la fracción de ácido nucleico fetal. Por ejemplo, uno cualquiera o una combinación de cualquier número de miniSTR, por ejemplo, pueden usarse las miniSTR desveladas en la Tabla 7 y Figuras 4 y 5. En una realización, la fracción de ácido nucleico fetal en una muestra materna se realiza usando un método que incluye determinar el número de copias del ácido nucleico materno y fetal presente en la muestra materna amplificando al menos una miniSTR autosómica elegida de CSF1PO, FGA, TH01, TPOX, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, Penta D, Penta E, D2S1338, D1S1677, D2S441, D4S2364, D10S1248, D14S1434, D22S1045, D22S1045, D20S1082, D20S482, D18S853, D17S1301, D17S974, D14S1434, D12ATA63, D11S4463, D10S1435, D10S1248, D9S2157, D9S1122, D8S1115, D6S1017, D6S474, D5S2500, D5S2500, D4S2408, D4S2364, D3S4529, D3S3053, D2S1776, D2S441, D1S1677, D1S1627 y D1GATA113. En otra realización, la al menos una miniSTR autosómica es el grupo de miniSTR CSF1PO, FGA, D13S317, D16S539, D18S51, D2S1338, D21S11, D2S1338 y D7S820. En una realización, el método comprende determinar el número de copias de al menos un alelo fetal y al menos uno materno al menos en una miniSTR polimórfica que se amplifica para generar amplicones que son inferiores a aproximadamente 300 pb, inferiores a aproximadamente 250 pb, inferiores a aproximadamente 200 pb, inferiores a aproximadamente 150 pb, inferiores a aproximadamente 100 pb, o inferiores a aproximadamente 50 pb. En otra realización, los amplicones que se generan amplificando las miniSTR son inferiores a aproximadamente 300 pb. En otra realización, los amplicones que se generan amplificando las miniSTR son inferiores a aproximadamente 250 pb. En otra realización, los amplicones que se generan amplificando las miniSTR son inferiores a aproximadamente 200 pb. La amplificación del alelo informativo incluye usar cebadores de miniSTR, que permiten la amplificación de amplicones de tamaño reducido para detectar alelos STR que son inferiores a aproximadamente 500 pb, inferiores a aproximadamente 450 pb, inferiores a aproximadamente 400 pb, inferiores a aproximadamente 350 pb, inferiores a aproximadamente 300 pares de bases (pb), inferiores a aproximadamente 250 pb, inferiores a aproximadamente 200 pb, inferiores a aproximadamente 150 pb, inferiores a aproximadamente 100 pb, o inferiores a aproximadamente 50 pb. Los amplicones de tamaño reducido generados usando los cebadores de miniSTR se conocen como miniSTR que se identifican según el nombre del marcador correspondiente al locus con el que se han mapeado. En una realización, los cebadores de miniSTR incluyen cebadores de mini STR que han permitido la máxima reducción de tamaño en el tamaño del amplicón para los 13 loci STR CODIS, además de D2S1338, Penta D y pentaE encontrados en kits de STR comercialmente disponibles (Butler et al., *J Forensic Sci* 48:1054-1064 [2003]), loci de miniSTR que no están asociados a los marcadores CODIS como se describe por Coble y Butler (Coble y Butler, *J Forensic Sci* 50:43-53 [2005]) y otras miniSTR que se han caracterizado en NIST. Información referente a las miniSTR caracterizadas en

NIST está accesible mediante la malla mundial en cstl.nist.gov/biotech/strbase/newSTRs.htm. Puede usarse un par cualquiera o una combinación de dos o más pares de cebadores de miniSTR para amplificar al menos un miniSTR.

5 En una realización, conjuntos de cebadores a modo de ejemplo que pueden usarse para amplificar STR en muestras de ADNlc materno incluyen los conjuntos de cebadores proporcionados en el Ejemplo 9 y desvelados como SEC ID N°:113-196.

10 La identificación del sexo (tipado del sexo) se realiza comúnmente conjuntamente con el tipado de STR usando productos de PCR generados a partir del gen amelogenina que se produce en tanto el cromosoma X como Y. La amelogenina no es un locus de STR, pero produce productos de PCR específicos de los cromosomas X e Y. Un conjunto de cebadores de PCR comúnmente usado publicado por primera vez por Sullivan et al. (1993) (Sullivan et al., *BioTechniques* 15:637-641 [1993]) se dirige a una delección de 6 pb que se produce en el cromosoma X, que permite distinguir entre sí amplicones generados a partir del cromosoma X e Y cuando se realiza separación electroforética para separar alelos de STR. Los kits de STR más comerciales utilizan los cebadores de Sullivan et al. (1993) o modificaciones menores. Como las hembras son X,X, solo se observa un único pico cuando se prueba ADN de hembra mientras que los machos, que poseen tanto cromosomas X como Y, presentan dos picos con una prueba de amelogenina estándar. En una realización, el método de determinación de la fracción de ácido nucleico fetal en una muestra materna comprende coamplificar amelogenina con al menos una miniSTR. En otra realización, el método no comprende coamplificar amelogenina con loci de miniSTR.

20 La amplificación de los ácidos nucleicos diana en la mezcla de ácido nucleico fetal y materna, por ejemplo, ADNlc, se lleva a cabo por cualquier método que use PCR o variaciones del método que incluyen, pero no se limitan a, PCR digital, PCR en tiempo real (RT-PCR), sistema TaqMan de PCR (Applied Biosystems), métodos SNPlex o GenPlex, PCR asimétrica, amplificación dependiente de helicasa, PCR de inicio caliente, qPCR, PCR en fase sólida y PCR de aterrizaje. Alternativamente, la replicación de secuencias de ácidos nucleicos diana puede obtenerse por métodos independientes de enzima, por ejemplo, síntesis química en fase sólida usando los fosforamiditos. La amplificación de las secuencias diana se lleva a cabo usando pares de cebadores cada uno capaz de amplificar una secuencia de ácidos nucleicos diana que comprende el sitio polimórfico, por ejemplo, SNP, en una reacción de PCR de múltiplex. Las reacciones de PCR de múltiplex incluyen combinar al menos 2, al menos tres, al menos 3, al menos 5, al menos 10, al menos 15, al menos 20, al menos 25, al menos 30, al menos 40 o más conjuntos de cebadores en la misma reacción para cuantificar los ácidos nucleicos diana amplificados que comprenden al menos dos, al menos tres, al menos 5, al menos 10, al menos 15, al menos 20, al menos 25, al menos 30, al menos 40 o más sitios polimórficos en la misma reacción de secuenciación. Puede configurarse cualquier panel de conjuntos de cebadores para amplificar al menos una secuencia polimórfica informativa.

35 En la etapa **140** del método **100 (Figura 1)**, una porción o todas las secuencias polimórficas amplificadas se usan para preparar una biblioteca de secuenciación para la secuenciación en un modo paralelo como se describe. En una realización, la biblioteca se prepara para la secuenciación por síntesis usando química de secuenciación basada en el terminador reversible de Illumina.

40 En la etapa **140**, la información de secuencia que se necesita para determinar la fracción fetal se obtiene usando cualquiera de los métodos de secuenciación de ADN conocidos. En una realización, el método descrito en el presente documento emplea la tecnología de secuenciación de nueva generación (NGS) en la que moldes de ADN clónicamente amplificados o moléculas de ADN individuales se secuencian en un modo masivamente paralelo dentro de una celda de flujo (por ejemplo, como se describe en Volkerding et al *Clin Chem* 55:641-658 [2009]; Metzker *M Nature Rev* 11:31-46 [2010]). Además de la información de secuencia de alta resolución, la NGS proporciona información cuantitativa digital, porque cada lectura de secuencia es una "marca de secuencia" contable que representa un ADN molde clónico individual o una molécula de ADN individual. Esta cuantificación permite que la NGS amplíe el concepto de PCR digital de contar moléculas de ADN libres de células (Fan et al., *Proc Natl Acad Sci U S A* 105:16266-16271 [2008]; Chiu et al., *Proc Natl Acad Sci U S A* 2008;105:20458-20463 [2008]). Las tecnologías de secuenciación de NGS incluyen pirosecuenciación, secuenciación por síntesis con terminadores de colorante reversible, secuenciación por ligación de sondas de oligonucleótidos y secuenciación en tiempo real.

55 Algunas de las tecnologías de secuenciación están disponibles comercialmente, tales como la plataforma de secuenciación por hibridación de Affymetrix Inc. (Sunnyvale, CA) y las plataformas de secuenciación por síntesis de 454 Life Sciences (Bradford, CT), Illumina/Solexa (Hayward, CA) y Helicos Biosciences (Cambridge, MA), y la plataforma de secuenciación por ligación de Applied Biosystems (Foster City, CA), como se describen más adelante. Además de la secuenciación de una molécula única realizada usando la secuenciación por síntesis de Helicos Biosciences, otras tecnologías de secuenciación de una molécula única están englobadas por el método de la invención e incluyen la tecnología SMRT™ de Pacific Biosciences, la tecnología Ion Torrent™ y la secuenciación de nanoporos que se desarrolla, por ejemplo, por Oxford Nanopore Technologies.

60 Aunque el método automatizado de Sanger se considera una tecnología de 'primera generación', la secuenciación de Sanger, que incluye la secuenciación de Sanger automatizada, también puede emplearse por el método de la invención. Métodos de secuenciación adicionales que comprenden el uso de tecnologías de obtención de imágenes de ácido nucleico en desarrollo, por ejemplo, microscopía de fuerza atómica (AFM) o microscopía electrónica de

transmisión (TEM), también están englobados por el método de la invención. Las tecnologías de secuenciación a modo de ejemplo se describen a continuación.

5 En una realización, la tecnología de secuenciación de ADN que se usa en el método de la invención es la secuenciación de una molécula única verdadera (tSMS) de Helicos (por ejemplo, como se describe en Harris T.D. et al., Science 320:106-109 [2008]). En la técnica de tSMS, una muestra de ADN se escinde en hebras de aproximadamente 100 a 200 nucleótidos, y se añade una secuencia de poliA al extremo 3' de cada hebra de ADN. Cada hebra se marca mediante la adición de un nucleótido de adenosina fluorescentemente marcado. Entonces, las hebras de ADN se hibridan con una celda de flujo, que contiene millones de sitios de captura de oligo-T que se inmovilizan en la superficie de la celda de flujo. Los moldes pueden tener una densidad de aproximadamente 100 millones de moldes/cm². Entonces, la celda de flujo se carga en un instrumento, por ejemplo, secuenciador HeliScope™, y un láser ilumina la superficie de la celda de flujo, revelando la posición de cada molde. Una cámara de CCD puede mapear la posición de los moldes sobre la superficie de la celda de flujo. Entonces, la marca fluorescente del molde se escinde y se lava. La reacción de secuenciación empieza introduciendo una ADN polimerasa y un nucleótido fluorescentemente marcado. El ácido nucleico de oligo-T sirve de cebador. La polimerasa incorpora los nucleótidos marcados al cebador en un modo dirigido al molde. Se eliminan la polimerasa y los nucleótidos sin incorporar. Los moldes que han dirigido la incorporación del nucleótido fluorescentemente marcado se distinguen por obtención de imágenes de la superficie de la celda de flujo. Después de la obtención de imágenes, una etapa de escisión elimina la marca fluorescente, y el proceso se repite con otros nucleótidos fluorescentemente marcados hasta que se logra la longitud de lectura deseada. La información de secuencia se recoge con cada etapa de adición de nucleótido.

25 En una realización, la tecnología de secuenciación de ADN que se usa en el método de la invención es la secuenciación 454 (Roche) (por ejemplo, como se describe en Margulies, M. et al. Nature 437:376-380 [2005]). La secuenciación 454 implica dos etapas. En la primera etapa, el ADN se corta en fragmentos de aproximadamente 300-800 pares de bases, y los fragmentos se hacen de extremos romos. Entonces se enlazan adaptadores de oligonucleótidos a los extremos de los fragmentos. Los adaptadores sirven de cebadores para la amplificación y secuenciación de los fragmentos. Los fragmentos pueden unirse a perlas de captura de ADN, por ejemplo, perlas recubiertas de estreptavidina usando, por ejemplo, el adaptador B, que contiene la marca 5'-biotina. Los fragmentos unidos a las perlas se amplifican por PCR dentro de gotitas de una emulsión de aceite-agua. El resultado es múltiples copias de fragmentos de ADN clónicamente amplificado sobre cada perla. En la segunda etapa, las perlas se capturan en pocillos (tamaño de pico-litros). Se realiza la pirosecuenciación en cada fragmento de ADN en paralelo. La adición de uno o más nucleótidos genera una señal de luz que se registra por una cámara de CCD en un instrumento de secuenciación. La intensidad de la señal es proporcional al número de nucleótidos incorporados. 35 La pirosecuenciación hace uso de pirofosfato (PPi) que es liberado tras la adición del nucleótido. El PPi se convierte en ATP por ATP sulfurilasa en presencia de adenosina 5' fosfosulfato. La luciferasa usa ATP para convertir la luciferina en oxiluciferina, y esta reacción genera luz que se distingue y se analiza.

40 En una realización, la tecnología de secuenciación de ADN que se usa en el método de la invención es la tecnología SOLiD (Applied Biosystems). En la secuenciación por ligación SOLiD, el ADN genómico se corta en fragmentos, y los adaptadores se unen a los extremos 5' y 3' de los fragmentos para generar una biblioteca de fragmentos. Alternativamente, pueden introducirse adaptadores internos a los extremos 5' y 3' de los fragmentos, circularizando los fragmentos, digiriendo el fragmento circularizado para generar un adaptador interno y uniendo los adaptadores a los extremos 5' y 3' de los fragmentos resultantes para generar una biblioteca emparejada de parejas. A 45 continuación, se preparan poblaciones de perlas clónicas en microrreactores que contienen perlas, cebadores, molde y componentes de PCR. Tras la PCR, los moldes se desnaturalizan y las perlas se enriquecen para separar las perlas con moldes extendidos. Los moldes en las perlas seleccionadas se someten a una modificación en 3' que permite la unión a un portaobjetos de vidrio. La secuencia puede determinarse por hibridación y ligación secuencial de oligonucleótidos parcialmente al azar con una base central determinada (o par de bases) que se identifica por un fluoróforo específico. Después de registrar un color, el oligonucleótido ligado se escinde y se elimina y entonces se repite el proceso.

55 En una realización, la tecnología de secuenciación de ADN que se usa en el método de la invención es la tecnología de secuenciación en tiempo real de una molécula individual (SMRT™) de Pacific Biosciences. En la secuenciación de SMRT se obtienen imágenes de la incorporación continua de nucleótidos marcados con colorante durante la síntesis de ADN. Las moléculas individuales de ADN polimerasa están unidas a la superficie inferior de identificadores de longitud de onda en modo cero individuales (identificadores de ZMW) que obtienen información de la secuencia mientras que los nucleótidos enlazados en fosfo están siendo incorporados en la hebra de cebador creciente. Un ZMW es una estructura de confinamiento que permite la observación de la incorporación de un único nucleótido por ADN polimerasa contra la referencia de nucleótidos fluorescentes que rápidamente difunden dentro y fuera del ZMW (en microsegundos). Dura varios milisegundos incorporar un nucleótido en una hebra en crecimiento. Durante este tiempo, la marca fluorescente se excita y produce una señal fluorescente, y la marca fluorescente se escinde. La identificación de la fluorescencia correspondiente del colorante indica qué base se incorporó. Se repite el proceso.

65 En una realización, la tecnología de secuenciación de ADN que se usa en el método de la invención es la

secuenciación en nanoporos (por ejemplo, como se describen en Soni GV y Meller A. Clin Chem 53: 1996-2001 [2007]). Las técnicas de análisis de ADN por secuenciación en nanoporos están siendo industrialmente desarrolladas por varias empresas, que incluyen Oxford Nanopore Technologies (Oxford, Reino Unido). La secuenciación de nanoporos es una tecnología de secuenciación de moléculas individuales por la cual una única molécula de ADN se secuencia directamente a medida que pasa a través de un nanoporo. Un nanoporo es un orificio pequeño, del orden de 1 nanómetro de diámetro. La inmersión de un nanoporo en un fluido conductor y la aplicación de un potencial (voltaje) a través de él produce una ligera corriente eléctrica debido a la conducción de iones a través del nanoporo. La cantidad de corriente que fluye es sensible al tamaño y forma del nanoporo. Como una molécula de ADN pasa a través de un nanoporo, cada nucleótido en la molécula de ADN obstruye el nanoporo a un grado diferente, cambiando la magnitud de la corriente a través del nanoporo en diferentes grados. Así, este cambio en la corriente a medida que la molécula de ADN pasa a través del nanoporo representa una lectura de la secuencia de ADN.

En una realización, la tecnología de secuenciación de ADN que se usa en el método de la invención es la matriz de transistores de efecto campo sensible a productos químicos (chemFET) (por ejemplo, como se describe en la publicación de solicitud de patente de EE.UU. N° 20090026082). En un ejemplo de la técnica, las moléculas de ADN pueden colocarse en cámaras de reacción, y las moléculas de molde pueden hibridarse con un cebador de secuenciación unido a una polimerasa. La incorporación de uno o más trifosfatos en una nueva hebra de ácido nucleico en el extremo 3' del cebador de secuenciación puede distinguirse por un cambio en la corriente por una a chemFET. Una matriz puede tener múltiples sensores chemFET. En otro ejemplo, ácidos nucleicos individuales pueden unirse a perlas, y los ácidos nucleicos pueden amplificarse en la perla, y las perlas individuales pueden transferirse a cámaras de reacción individuales en una matriz chemFET, teniendo cada cámara un sensor chemFET, y pueden secuenciarse los ácidos nucleicos.

En una realización, la tecnología de secuenciación de ADN que se usa en el método de la invención es el método de Halcyon Molecular que usa microscopía electrónica de transmisión (TEM). El método, llamado nanotransferencia rápida de colocación de moléculas individuales (IMPRNT), comprende utilizar la obtención de imágenes por microscopio electrónico de transmisión con resolución para un único átomo de ADN de alto peso molecular (150 kb o mayor) selectivamente marcado con marcadores de átomos pesados y disponer estas moléculas sobre películas ultra-finas en matrices paralelas ultra-densas (3 nm de hebra a hebra) con separación base a base coherente. El microscopio electrónico se usa para obtener imágenes de las moléculas en las películas para determinar la posición de los marcadores de átomos pesados y para extraer información de secuencias de bases del ADN. El método se describe adicionalmente en la publicación de patente PCT WO 2009/046445. El método permite secuenciar genomas humanos completos en menos de diez minutos.

En una realización, la tecnología de secuenciación de ADN es la secuenciación de una molécula única Ion Torrent, que empareja la tecnología de semiconductores con una simple química de secuenciación para traducir directamente la información químicamente codificada (A, C, G, T) en información digital (0, 1) en un chip semiconductor. En la naturaleza, cuando un nucleótido se incorpora en una hebra de ADN por una polimerasa, se libera un ión hidrógeno como subproducto. Ion Torrent usa un matriz de alta densidad de pocillos micro-mecanizados para realizar este proceso bioquímico de una forma masivamente paralela. Cada pocillo contiene una molécula de ADN diferente. Debajo de los pocillos está una capa sensible a los iones y debajo de ésta un sensor de iones. Cuando un nucleótido, por ejemplo, una C, se añade a un ADN molde y luego se incorpora en una hebra de ADN, se liberará un ión hidrógeno. La carga de ese ión cambiará el pH de la disolución, que puede identificarse por el sensor de iones de Ion Torrent. El secuenciador - el medidor de pH en estado sólido esencialmente más pequeño del mundo - llama a la base, pasando directamente de información química a información digital. Entonces, el secuenciador Ion personal Genoma Machine (PGM™) inunda secuencialmente el chip con un nucleótido después del otro. Si el siguiente nucleótido que inunda el chip no es una coincidencia, no se registrará cambio de voltaje y no se llamará a ninguna base. Si hay dos bases idénticas en la hebra de ADN, el voltaje se duplicará, y el chip registrará dos bases idénticas llamadas. La identificación directa permite el registro de incorporación de nucleótidos en segundos.

Otros métodos de secuenciación incluyen PCR digital y secuenciación por hibridación. La reacción en cadena de la polimerasa digital (PCR digital o dPCR) puede usarse para identificar y cuantificar directamente ácidos nucleicos en una muestra. La PCR digital puede realizarse en una emulsión. Se separan ácidos nucleicos individuales, por ejemplo, en un dispositivo de cámara microfluídica, y cada nucleico puede amplificarse individualmente por PCR. Los ácidos nucleicos pueden separarse tal que haya un promedio de aproximadamente 0,5 ácidos nucleicos/pocillo, o no más de un ácido nucleico/pocillo. Pueden usarse sondas diferentes para distinguir alelos fetales y alelos maternos. Los alelos pueden enumerarse para determinar el número de copias. En la secuenciación por hibridación, la hibridación comprende poner en contacto la pluralidad de secuencias de polinucleótidos con una pluralidad de sondas de polinucleótidos, en la que cada una de la pluralidad de sondas de polinucleótidos puede estar opcionalmente unida a un sustrato. El sustrato podría ser una superficie plana que comprende una matriz de secuencias de nucleótidos conocidas. El patrón de hibridación con la matriz puede usarse para determinar las secuencias de polinucleótidos presentes en la muestra. En otras realizaciones, cada sonda está unida a una perla, por ejemplo, una perla magnética o similares. La hibridación con las perlas puede identificarse y usarse para identificar la pluralidad de secuencias de polinucleótidos dentro de la muestra.

En una realización, el método emplea secuenciación masivamente paralela de millones de fragmentos de ADN usando la secuenciación por síntesis de Illumina y química de secuenciación basada en terminadores reversibles (por ejemplo, como se describe en Bentley et al., Nature 6:53-59 [2009]). El ADN molde puede ser ADN genómico, por ejemplo, ADNlc. En algunas realizaciones, el ADN genómico de células aisladas se usa como molde, y se fragmenta en longitudes de varios cientos de pares de bases. En otras realizaciones, se usa ADNlc como molde, y no se requiere fragmentación ya que el ADNlc existe como fragmentos cortos. Por ejemplo, el ADNlc fetal circula en la circulación sanguínea como fragmentos de <300 pb, y se ha estimado que el ADNlc materno circula como fragmentos de entre aproximadamente 0,5 y 1 Kb (Li et al., Clin Chem, 50: 1002-1011 [2004]). La tecnología de secuenciación de Illumina se basa en la unión de ADN genómico fragmentado a una superficie ópticamente transparente plana sobre la que se unen anclajes de oligonucleótido. El ADN molde se prepara en los extremos para generar extremos romos fosforilados en 5', y la actividad de polimerasa del fragmento de Klenow se usa para añadir una única base de A al extremo 3' de los fragmentos de ADN fosforilados romos. Esta adición prepara los fragmentos de ADN para la ligación a adaptadores de oligonucleótido, que tienen nucleótidos protuberantes en una única base de T en su extremo 3' para aumentar la eficiencia de ligación. Los oligonucleótidos del adaptador son complementarios a los anclajes de la celda de flujo. Bajo condiciones de dilución limitante, el ADN molde monocatenario modificado por el adaptador se añade a la celda de flujo y se inmoviliza por hibridación a los anclajes. Los fragmentos de ADN unidos se extienden y se amplifican los puentes para crear una celda de flujo de secuenciación de densidad ultra-alta con cientos de millones de agrupaciones, cada una de las cuales contiene ~1.000 copias del mismo molde. Las moléculas de ADN amplificado agrupadas se secuencian usando una robusta tecnología de secuenciación de ADN por síntesis de cuatro colores que emplea terminadores reversibles con colorantes fluorescentes eliminables. La identificación de fluorescencia de alta sensibilidad se logra usando excitación de láser y óptica de reflexión interna total. Las lecturas de secuencia corta de aproximadamente 20-40 pb, por ejemplo, 36 pb, están alineadas contra un genoma de referencia enmascarado en las repeticiones y las diferencias genéticas se llaman usando software de canal de análisis de datos especialmente desarrollado. Después de completarse la primera lectura, los moldes pueden regenerarse *in situ* para permitir una segunda lectura de los extremos opuestos de los fragmentos. Así, según el método se usa tanto la secuenciación de fragmentos de ADN de extremos individuales como de extremos apareados. Se realiza la secuenciación parcial de fragmentos de ADN presentes en la muestra, y se cuentan marcas de secuencia que comprenden lecturas de longitud predeterminada, por ejemplo, 36 pb, que son mapeadas en un genoma de referencia conocido.

La longitud de la lectura de secuencia está asociada a la tecnología de secuenciación particular. Los métodos de NGS proporcionan lecturas de secuencia que varían en tamaño de decenas a cientos de pares de bases. En algunas realizaciones del método descrito en el presente documento, las lecturas de secuencia son aproximadamente 20 pb, aproximadamente 25 pb, aproximadamente 30 pb, aproximadamente 35 pb, aproximadamente 40 pb, aproximadamente 45 pb, aproximadamente 50 pb, aproximadamente 55 pb, aproximadamente 60 pb, aproximadamente 65 pb, aproximadamente 70 pb, aproximadamente 75 pb, aproximadamente 80 pb, aproximadamente 85 pb, aproximadamente 90 pb, aproximadamente 95 pb, aproximadamente 100 pb, aproximadamente 110 pb, aproximadamente 120 pb, aproximadamente 130, aproximadamente 140 pb, aproximadamente 150 pb, aproximadamente 200 pb, aproximadamente 250 pb, aproximadamente 300 pb, aproximadamente 350 pb, aproximadamente 400 pb, aproximadamente 450 pb, aproximadamente 500 pb, aproximadamente 550 pb o aproximadamente 600 pb. Se espera que los avances tecnológicos permitan lecturas de extremos individuales de más de 500 pb, que permitan lecturas de más de aproximadamente 1000 pb cuando se generan lecturas de extremos apareados. En una realización, las lecturas de secuencia tienen 36 pb. Otros métodos de secuenciación que pueden emplearse por el método de la invención incluyen los métodos de secuenciación de una molécula única que pueden secuenciar moléculas de ácidos nucleicos >5000 pb. La masiva cantidad de salida de secuencias se transfiere por un canal de análisis que transforma la salida de imágenes primarias del secuenciador en cadenas de bases. Un paquete de algoritmos integrados realiza las principales etapas de transformación de datos primarios: análisis de imágenes, puntuación de la intensidad, llamada de bases y alineamiento.

En una realización se realiza la secuenciación parcial de ácidos nucleicos polimórficos diana amplificados, y se cuentan marcas de secuencia que comprenden lecturas de longitud predeterminada, por ejemplo, 36 pb, que se mapean en un genoma de referencia conocido. Solo lecturas de secuencia que se alinean únicamente con un genoma de referencia se cuentan como marcas de secuencia. En una realización, el genoma de referencia es un genoma de secuencias diana artificial que comprende las secuencias de los ácidos nucleicos diana polimórficos, por ejemplo, SNP. En una realización, el genoma de referencia es un genoma de referencia SNP artificial. En otra realización, el genoma de referencia es un genoma de referencia STR artificial. En otra realización más, el genoma de referencia es un genoma de referencia de STR en tándem artificial. Los genomas de referencia artificiales pueden compilarse usando las secuencias de los ácidos nucleicos polimórficos diana. Los genomas de referencia artificiales pueden comprender secuencias diana polimórficas que comprenden cada una uno o más tipos diferentes de secuencias polimórficas. Por ejemplo, un genoma de referencia artificial puede comprender secuencias polimórficas que comprenden alelos de SNP y/o STR. En una realización, el genoma de referencia es la secuencia NCBI36/hg18 del genoma de referencia humano, que está disponible en la malla mundial en genome.ucsc.edu/cgi-bin/hgGateway?org=Human&db=hg18&hgsid=166260105. Otras fuentes de información de secuencia pública incluyen GenBank, dbEST, dbSTS, EMBL (el Laboratorio Europeo de Biología Molecular) y DDBJ (la base de datos de ADN de Japón). En otra realización, el genoma de referencia comprende la secuencia NCBI36/hg18 del genoma

de referencia humano y un genoma de secuencias diana artificial, que incluye las secuencias polimórficas diana, por ejemplo, un genoma de SNP. El mapeo de las marcas de secuencia se logra comparando la secuencia de la marca mapeada con la secuencia del genoma de referencia para determinar el origen cromosómico de la molécula de ácido nucleico secuenciado (por ejemplo, ADNlc), y no se necesita información de secuencia genética específica. Están disponibles varios algoritmos informáticos para alinear secuencias, que incluyen, sin limitación BLAST (Altschul et al., 1990), BLITZ (MPsrch) (Sturrock & Collins, 1993), FASTA (Person & Lipman, 1988), BOWTIE (Langmead et al., Genome Biology 10:R25,1-R25,10 [2009]) o ELAND (Illumina, Inc., San Diego, CA, USA). En una realización, un extremo de las copias clónicamente expandidas de las moléculas de ADNlc de plasma se secuencian y se procesan por análisis de alineamiento bioinformático para el analizador de genoma de Illumina, que usa el software Efficient Large-Scale Alignment of Nucleotide Databases (ELAND). En realizaciones del método que comprenden determinar la presencia o ausencia de una aneuploidía y la fracción fetal usando métodos de secuenciación de NGS, la información del análisis de secuenciación para la determinación de aneuploidía puede permitir un pequeño grado de desapareamiento (0-2 desapareamientos por marca de secuencia) para explicar polimorfismos menores que pueden existir entre el genoma de referencia y los genomas en la muestra mezclada. La información del análisis de secuenciación para la determinación de la fracción fetal puede permitir un pequeño grado de desapareamiento dependiendo de la secuencia polimórfica. Por ejemplo, puede permitirse un pequeño grado de desapareamiento si la secuencia polimórfica es una STR. En casos cuando la secuencia polimórfica sea un SNP, se cuenta primero toda secuencia que se aparece exactamente con cualquiera de los dos alelos en el sitio SNP y se filtra de las lecturas restantes, para lo que puede permitirse un pequeño grado de desapareamiento. La cuantificación del número de lecturas de secuencia que se alinean con cada cromosoma para determinar aneuploidías cromosómicas puede determinarse como se describe en el presente documento, o usando análisis alternativos que emplean la normalización de la mediana del número de marcas de secuencia para un cromosoma de interés con la mediana del número de marcas para cada uno de los otros autosomas (Fan et al., Proc Natl Acad Sci 105:16266-16271 [2008]), o que comparan el número de lecturas únicas que se alinean con cada cromosoma con el número total de lecturas que se alinean con todos los cromosomas para derivar un porcentaje de representación genómica para cada cromosoma. Se genera una "puntuación z" para representar la diferencia entre el porcentaje de representación genómica del cromosoma de interés y la representación en porcentaje media para el mismo cromosoma entre un grupo de control euploide, dividido por la desviación estándar (Chiu et al., Clin Chem 56:459-463 [2010]). En otra realización, la información de secuenciación puede determinarse como se describe en la solicitud de patente provisional de EE.UU. titulada "Normalizing Biological Assays", expediente nº 32047-768,101, presentada el 19 de enero de 2010.

La información del análisis de secuenciación para la determinación de la fracción fetal puede permitir un pequeño grado de desapareamiento dependiendo de la secuencia polimórfica. Por ejemplo, puede permitirse un pequeño grado de desapareamiento si la secuencia polimórfica es una STR. En casos cuando la secuencia polimórfica sea un SNP, todas las secuencias que se aparecen exactamente con cualquiera de los dos alelos en el sitio de SNP se cuentan primero y se filtran de las lecturas restantes, para lo que puede permitirse un pequeño grado de desapareamiento. El presente método de determinación de la fracción fetal por secuenciación de ácidos nucleicos puede usarse en combinación con otros métodos.

En la etapa **160**, la fracción fetal se determina basándose en el número total de marcas que se mapean en el primer alelo y el número total de marcas que se mapean en el segundo alelo en un sitio polimórfico informativo, por ejemplo, un SNP, contenido en un genoma de referencia. Por ejemplo, el genoma de referencia es un genoma de secuencia diana artificial que engloba las secuencias polimórficas que comprenden los SNP rs560681, rs1109037, rs9866013, rs13182883, rs13218440, rs7041158, rs740598, rs10773760, rs4530059, rs7205345, rs8078417, rs576261, rs2567608, rs430046, rs9951171, rs338882, rs10776839, rs9905977, rs1277284, rs258684, rs1347696, rs508485, rs9788670, rs8137254, rs3143, rs2182957, rs3739005 y rs530022. En una realización, el genoma de referencia artificial incluye las secuencias polimórficas diana de SEC ID Nº: 1-56 (véase el Ejemplo-3).

En otra realización, el genoma de referencia artificial es un genoma de secuencia diana artificial que engloba secuencias polimórficas que comprenden los SNP en tándem rs7277033-rs2110153; rs2822654-rs1882882; rs368657-rs376635; rs2822731-rs2822732; rs1475881-rs7275487; rs1735976-rs2827016; rs447340-rs2824097; rs418989-rs13047336; rs987980-rs987981; rs4143392-rs4143391; rs1691324-rs13050434; rs11909758-rs9980111; rs2826842-rs232414; rs1980969-rs1980970; rs9978999-rs9979175; rs1034346-rs12481852; rs7509629-rs2828358; rs4817013-rs7277036; rs9981121-rs2829696; rs455921-rs2898102; rs2898102-rs458848; rs961301-rs2830208; rs2174536-rs458076; rs11088023-rs11088024; rs1011734-rs1011733; rs2831244-rs9789838; rs8132769-rs2831440; rs8134080-rs2831524; rs4817219-rs4817220; rs2250911-rs2250997; rs2831899-rs2831900; rs2831902-rs2831903; rs11088086-rs2251447; rs2832040-rs11088088; rs2832141-rs2246777; rs2832959-rs9980934; rs2833734-rs2833735; rs933121-rs933122; rs2834140-rs12626953; rs2834485-rs3453; rs9974986-rs2834703; rs2776266-rs2835001; rs1984014-rs1984015; rs7281674-rs2835316; rs13047304-rs13047322; rs2835545-rs4816551; rs2835735-rs2835736; rs13047608-rs2835826; rs2836550-rs2212596; rs2836660-rs2836661; rs465612-rs8131220; rs9980072-rs8130031; rs418359-rs2836926; rs7278447-rs7278858; rs385787-rs367001; rs367001-rs386095; rs2837296-rs2837297; y rs2837381-rs4816672.

En otra realización, el genoma de referencia artificial engloba secuencias polimórficas que comprenden las STR seleccionadas de CSF1PO, FGA, TH01, TPOX, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539,

D18S51, D21S11, Penta D, Penta E, D2S1338, D1S1677, D2S441, D4S2364, D10S1248, D14S1434, D22S1045, D22S1045, D20S1082, D20S482, D18S853, D17S1301, D17S974, D14S1434, D12ATA63, D11S4463, D10S1435, D10S1248, D9S2157, D9S1122, D8S1115, D6S1017, D6S474, D5S2500, D5S2500, D4S2408, D4S2364, D3S4529, D3S3053, D2S1776, D2S441, D1S1677, D1S1627 y D1GATA113. La composición del genoma de referencia artificial variará dependiendo de las secuencias polimórficas que se usan para determinar la fracción fetal. Por consiguiente, un genoma de secuencias dianas artificial no se limita a las secuencias de SNP, SNP en tándem o de STR ejemplificadas en el presente documento.

El sitio polimórfico informativo, por ejemplo, SNP, se identifica por la diferencia en las secuencias alélicas y la cantidad de cada uno de los posible alelos. El ADNlc fetal está presente a una concentración que es <10 % del ADNlc materno. Así, la presencia de una contribución menor de un alelo a la mezcla de ácidos nucleicos fetales y maternos con respecto a la contribución principal del alelo materno puede asignarse al feto. Los alelos que se derivan del genoma materno se denominan en el presente documento alelos principales, y los alelos que se derivan del genoma fetal se denominan en el presente documento alelos secundarios. Los alelos que se representan por niveles similares de marcas de secuencia mapeadas representan alelos maternos. Los resultados de una amplificación de ácidos nucleicos diana múltiplex a modo de ejemplo que comprende SNP y se derivan de una muestra de plasma materno se muestra en la Figura 2. Los SNP informativos se distinguen del cambio de un único nucleótido en un sitio polimórfico, y los alelos fetales se distinguen por su contribución menor relativa a la mezcla de ácidos nucleicos fetales y maternos en la muestra cuando se compara con la contribución principal a la mezcla por los ácidos nucleicos maternos. Por consiguiente, la abundancia relativa de ADNlc fetal en la muestra materna se determina como un parámetro del número total de marcas de secuencia únicas mapeadas en la secuencia de ácidos nucleicos diana en un genoma de referencia para cada uno de los dos alelos del sitio polimórfico predeterminado. En una realización, la fracción de ácidos nucleicos fetales en la mezcla de ácidos nucleicos fetales y maternos se calcula para cada uno del alelo informativo (alelo_x) del siguiente modo:

$$\% \text{ fracción fetal alelo}_x = \left(\frac{\sum \text{secuencias de etiquetas fetales para alelo}_{xy}}{\sum \text{secuencias de etiquetas maternas para alelo}_x} \right) \times 100$$

y la fracción fetal para la muestra se calcula como el promedio de la fracción fetal de todos los alelos informativos. Opcionalmente, la fracción de ácidos nucleicos fetales en la mezcla de ácidos nucleicos fetales y maternos se calcula para cada alelo informativo (alelo_x) del siguiente modo:

$$\% \text{ fracción fetal alelo}_x = \left(\frac{2 \times \sum \text{secuencias de etiquetas fetales para alelo}_{xy}}{\sum \text{secuencias de etiquetas maternas para alelo}_x} \right) \times 100,$$

para compensar la presencia de 2 alelos fetales, estando uno enmascarado por el fondo materno.

El porcentaje de fracción fetal se calcula para al menos 1, al menos 2, al menos 3, al menos 4, al menos 5, al menos 6, al menos 7, al menos 8, al menos 9, al menos 10, al menos 11, al menos 12, al menos 13, al menos 14, al menos 15, al menos 16, al menos 17, al menos 18, al menos 19, al menos 20, al menos 25, al menos 30, al menos 40 o más alelos informativos. En una realización, la fracción fetal es el promedio la fracción fetal determinada para al menos 3 alelos informativos.

La **Figura 3** muestra un diagrama de flujo de métodos alternativos por los cuales la fracción fetal puede determinarse a partir de ácidos nucleicos diana amplificados que se han combinado con una muestra de ADNlc fetal y materno amplificado para permitir la determinación simultánea de la fracción fetal y la presencia o ausencia de aneuploidía fetal enriqueciendo la muestra materna que comprende ácidos nucleicos fetales y maternos para ácidos nucleicos diana polimórficos. En una realización, la muestra que se enriquece es la fracción de plasma de una muestra de sangre (a). Por ejemplo, se usa una porción de una muestra de plasma materno original para amplificar secuencias de ácidos nucleicos diana. Posteriormente, algo o todo el producto amplificado se combina con la muestra de plasma original sin amplificar restante, enriqueciéndola así (véase el Ejemplo 7). En otra realización, la muestra que se enriquece es la muestra de ADNlc purificado que se extrae del plasma (b). Por ejemplo, el enriquecimiento comprende amplificar los ácidos nucleicos diana que están contenidos en una porción de una muestra original de mezcla purificada de ácidos nucleicos fetales y maternos, por ejemplo, ADNlc que se ha purificado de una muestra de plasma materno, y posteriormente combinar algo o todo el producto amplificado con la restante muestra purificada original sin amplificar (véase el Ejemplo 6). En otra realización más, la muestra que se enriquece es una muestra de biblioteca de secuenciación preparada a partir de una mezcla purificada de ácidos nucleicos fetales y maternos (c). Por ejemplo, el enriquecimiento comprende amplificar los ácidos nucleicos diana que están contenidos en una porción de una muestra original de mezcla purificada de ácidos nucleicos fetales y maternos, por ejemplo, ADNlc que se ha purificado de una muestra de plasma materno, preparar una primera biblioteca de secuenciación de secuencias de ácidos nucleicos sin amplificar, preparar una segunda biblioteca de secuenciación de ácidos nucleicos diana polimórficos amplificados y posteriormente combinar algo o toda de la segunda biblioteca de secuenciación con algo o todo de la primera biblioteca de secuenciación (véase el Ejemplo 5). La cantidad de producto amplificado que se usa para enriquecer la muestra original se selecciona para obtener información de secuenciación suficiente para determinar tanto la presencia como la ausencia de aneuploidía y la fracción fetal del mismo ciclo de secuenciación. Se mapean al menos aproximadamente el 3 %, al menos aproximadamente el 5 %, al menos aproximadamente el 7 %, al menos aproximadamente el 10 %, al menos

aproximadamente el 15 %, al menos aproximadamente el 20 %, al menos aproximadamente el 25 %, al menos aproximadamente el 30% o más del número total de marcas de secuencia obtenidas de secuenciación para determinar la fracción fetal. La secuenciación de la biblioteca generada tras uno cualquiera de los métodos representados en la Figura 3 proporciona marcas de secuencia derivadas de los ácidos nucleicos diana amplificados y marcas derivadas de la muestra materna sin amplificar original. La fracción fetal se calcula a partir del número de marcas mapeadas en un genoma de referencia artificial, y la presencia o ausencia de aneuploidía se determina a partir del número de marcas que se mapean con el genoma del sujeto, por ejemplo, genoma humano.

Un método alternativo para determinar la fracción fetal de ácidos nucleicos diana polimórficos amplificados en la etapa 130 de la Figura 1 usa separación por tamaño de ácidos nucleicos diana polimórficos amplificados que comprenden STR (etapa 150 de la Figura 1). Como se ha descrito anteriormente, el carácter polimórfico de un locus de STR es debido a la variación en el número de unidades repetidas en tándem entre alelos. Debido al alto polimorfismo de las STR, la mayoría de los individuos serán heterocigotos para STR. La amplificación de una STR producirá uno o dos productos de PCR en la mayoría de las muestras. En muestras obtenidas de mujeres embarazadas, por ejemplo, muestras de plasma, la amplificación de una STR producirá uno o dos productos de PCR importantes, que se corresponden con el uno o los dos alelos maternos que incluyen un alelo maternamente heredado fetal, y un tercer alelo paternamente heredado fetal que se detecta en una STR informativa.

Las STR que son dirigidas para amplificación son miniSTR como se describen en el presente documento que son inferiores a aproximadamente 300 pares de bases y que se amplifican en una reacción de PCR de múltiplex, que permite la amplificación simultánea de múltiples loci en una única reacción. Los cebadores se marcan con diferentes colorantes fluorescentes que emite cada uno fluorescencia a una longitud de onda diferente, por ejemplo, 6FAM™, VIC™, NED™ y PET™, y el número de unidades de repetición para cada STR fluorescentemente marcada en los productos de PCR resultantes se detecta tras su separación y dimensionamiento preciso que se logra por electroforesis en placa o capilar. En una realización se usa electroforesis capilar, y puede realizarse en canales microfabricados o matrices capilares. Alternativamente, se usan métodos que utilizan espectrometría de masas y tecnología de micromatrices. Puede realizarse análisis de STR múltiplex para determinar la fracción fetal usando kits comercialmente disponibles, por ejemplo, kit de amplificación por PCR AmpFESTR® Identifiler® (Figura 4) y kit de amplificación por PCR AmpFtSTR® MiniFiler® (Figura 5) (Applied Biosystems, Foster City, Calif.). El kit de amplificación por PCR AmpFtSTR® MiniFiler® se diseñó para amplificar como miniSTR ocho de los loci de mayor tamaño en el kit de amplificación por PCR AmpFtSTR® Identifiler®. Junto con el locus de identificación del sexo amelogenina, el múltiplex de nueve locus permite la amplificación simultánea de loci de muestras de ADNlc (véase el Ejemplo 10).

En una realización, el análisis de STR de múltiplex para determinar la fracción fetal se realiza amplificando ácidos nucleicos diana polimórficos presentes en una muestra de plasma materno que comprende cada una una miniSTR seleccionada de CSF1PO, FGA, TH01, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D2S1338, Penta D, Penta E, D22S1045, D20S1082, D20S482, D18S853, D17S1301, D17S974, D14S1434, D12ATA63, D11S4463, D10S1435, D10S1248, D9S2157, D9S1122, D8S1115, D6S1017, D6S474, D5S2500, D4S2408, D4S2364, D3S4529, D3S3053, D2S1776, D2S441, D1S1677, D1S1627 y D1GATA113. En otra realización, el análisis de STR de múltiplex para determinar la fracción fetal se realiza amplificando ácidos nucleicos diana polimórficos presentes en una muestra de plasma materno para el panel de miniSTR: CSF1PO, D13S317, D16S539, D18S51, D21S11, D2S1338, D7S820 y FGA. Las miniSTR pueden localizarse en el mismo cromosoma o en cromosomas diferentes. El método es un método independiente del sexo del feto. Por tanto, en algunas realizaciones, las miniSTR se localizan en cromosomas distintos del cromosoma Y. En otras realizaciones, las miniSTR se localizan en cromosomas distintos de los cromosomas 13, 18, 21 o X, es decir, cromosomas que podrían estar implicados en una aneuploidía.

Las muestras de plasma materno frecuentemente contienen menos de 100 pg de ADNlc. Las muestras de ADN de bajo número de copias pueden encontrarse por debajo de las limitaciones de sensibilidad de los métodos de análisis de STR. Las muestras no tratables pueden hacerse susceptibles aumentando el número de ADNlc de partida disponible para el posterior análisis de STR por una estrategia de amplificación de genoma completo. En una realización, la mezcla de ácidos nucleicos fetales y maternos puede preamplificarse antes de que los alelos se detecten o cuantifiquen. Por ejemplo, puede amplificarse ADN libre de células molde por PCR. El ácido nucleico puede amplificarse durante aproximadamente 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, 35, 36, 37, 38, 39 o 40 ciclos. El ácido nucleico puede amplificarse durante aproximadamente 1-10 ciclos, aproximadamente 1-20 ciclos, aproximadamente 1-30 ciclos, aproximadamente 1-40 ciclos, aproximadamente 5-15 ciclos, aproximadamente 5-20 ciclos, aproximadamente 5-30 ciclos, aproximadamente 5-40 ciclos, aproximadamente 10-15 ciclos, aproximadamente 10-20 ciclos, aproximadamente 10-30 ciclos, aproximadamente 10-40 ciclos, aproximadamente 20-30 ciclos, aproximadamente 20-40 ciclos, o aproximadamente 30-40 ciclos. La cantidad de ácido nucleico de molde que puede amplificarse puede ser aproximadamente 10-1000 pg, 25-1000 pg, 50-1000 pg, 100-1000 pg, 200-1000 pg, 300-1000 pg, 400-1000 pg, 500-1000 pg, 600-1000 pg, 700-1000 pg o 800-1000 pg. Tras la preamplificación, los ácidos nucleicos pueden diluirse antes de detectar o cuantificar los alelos. La preamplificación puede usarse para aumentar la sensibilidad de detección de los alelos en una muestra, por ejemplo, una muestra materna (Ejemplo 11). En otra realización, el genotipar un polimorfismo no necesita requerir una etapa de pre-amplificación. Puede usarse

cualquier método de amplificación basado en PCR para preamplificar el ADNlc. Los métodos de amplificación incluyen, pero no se limitan a, estrategias de amplificación del genoma completo que incluyen métodos tales como preamplificación por extensión de cebadores, PCR cebada con oligonucleótidos degenerados (DOP-PCR), fragmentos bajos de bajas cantidades de DOP-PCR, PCR de preamplificación por extensión de cebadores mejorada (IPEP PCR) y preamplificación por extensión de cebadores mejorada modificada (mIPEP). Así, en una realización, el método que se usa para determinar la fracción fetal en una muestra materna, por ejemplo, muestra de plasma, comprende preamplificar la mezcla de ácidos nucleicos fetales y maternos presente en la muestra de ADNlc de plasma usando un método de amplificación del genoma completo, amplificar una pluralidad de ácidos nucleicos polimórficos en dicha mezcla de ácidos nucleicos fetales y maternos, en la que cada uno de dicho al menos un ácido nucleico polimórfico comprende una STR; determinar la cantidad de alelos de STR fetales y maternos de al menos un ácido nucleico polimórfico; y calcular la fracción fetal a partir de la cantidad de alelos de STR fetales y maternos. Tras la preamplificación, se realiza análisis de STR de múltiplex para determinar la fracción fetal amplificando ácidos nucleicos diana polimórficos presentes en una muestra de plasma materno que comprende cada uno una miniSTR seleccionada de CSF1PO, FGA, TH01, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D2S1338, Penta D, Penta E, D22S1045, D20S1082, D20S482, D18S853, D17S1301, D17S974, D14S1434, D12ATA63, D11S4463, D10S1435, D10S1248, D9S2157, D9S1122, D8S1115, D6S1017, D6S474, D5S2500, D4S2408, D4S2364, D3S4529, D3S3053, D2S1776, D2S441, D1S1677, D1S1627 y D1GATA113. Alternativamente, se realiza análisis de STR múltiplex para determinar la fracción fetal amplificando ácidos nucleicos diana polimórficos para el panel de miniSTR: CSF1PO, D13S317, D16S539, D18S51, D21S11, D2S1338, D7S820 y FGA.

Aplicaciones

Los métodos descritos en el presente documento son aplicables al diagnóstico o pronóstico de diversas condiciones de enfermedad que incluyen, pero no se limitan a, cáncer, trastornos genéticos e infección. La fracción fetal de ácido nucleico en una muestra materna puede usarse para determinar una anomalía cromosómica. Ejemplos de anomalías cromosómicas incluyen, por ejemplo, aneuploidía, monosomía, trisomía, duplicación, inversión, deleción, poliploidía, deleción de una parte de un cromosoma, adición, adición de una parte de cromosoma, inserción, un fragmento de un cromosoma, una región de un cromosoma, transposición cromosómica y translocalización. Por ejemplo, aneuploidía puede referirse a la aparición de uno o más cromosomas adicionales o ausentes en una muestra.

Ejemplos de condiciones fetales que pueden determinarse usando los métodos de la invención proporcionada incluyen, por ejemplo, síndrome de Angleman (15q11.2-q13), síndrome del maullido (5p-), síndrome de DiGeorge y síndrome velocardiocfacial (22q11.2), síndrome de Miller-Dieker (17 p13.3), síndrome de Prader-Willi (15q11.2-q13), retinoblastoma (13q14), síndrome de Smith-Magenis (17 p11.2), trisomía 13, trisomía 16, trisomía 18, trisomía 21 (síndrome de Down), triploidía, síndrome de Williams (7q 11.23) y síndrome de Wolf-Hirschhorn (4p-). Ejemplos de anomalías de los cromosomas sexuales que pueden detectarse por los métodos descritos en el presente documento incluyen, pero no se limitan a, síndrome de Kallman (Xp22.3), deficiencia de esteroide sulfatasa (STS) (Xp22.3), ictiosis ligada al cromosoma X (Xp22.3), síndrome de Klinefelter (XXY), síndrome del X frágil, síndrome de Turner, superhembra o trisomía X, y monosomía X.

En una realización, la información de la fracción fetal puede usarse para establecer umbrales y estimar el tamaño de muestra mínimo en la detección de aneuploidías. Tal uso se describe en el Ejemplo 7 más adelante. La información de la fracción fetal puede usarse conjuntamente con la información de secuenciación. Por ejemplo, pueden usarse ácidos nucleicos de una muestra libre de células, por ejemplo, una muestra de plasma materno o de suero, para enumerar secuencias en una muestra. Las secuencias pueden enumerarse usando cualquiera de las técnicas de secuenciación descritas anteriormente. El conocimiento de la fracción fetal puede usarse para establecer umbrales de "corte" para llamar estados de "aneuploidía", "normales" o "marginales/no llamados" (incierto). Entonces, pueden realizarse cálculos para estimar el número mínimo de secuencias requerido para lograr la sensibilidad adecuada (es decir, probabilidad de identificar correctamente un estado de aneuploidía).

La determinación de la fracción fetal según el método de la invención puede ponerse en práctica en combinación con cualquier método usado para determinar la presencia de ausencia de aneuploidía fetal en una muestra de plasma materno. Además del método descrito en el presente documento para la determinación de aneuploidía, la determinación de la fracción fetal en secuenciación masivamente paralela puede usarse conjuntamente con otros métodos para determinar aneuploidía fetal, por ejemplo, según los métodos descritos en las publicaciones de solicitud de patente de EE.UU. N° US 2007/0202525A1; US2010/0112575A1, US 2009/0087847A1; US2009/0029377A1; US 2008/0220422A1; US2008/0138809A1, US2008/0153090A1 y la patente de EE.UU. 7.645.576. Los métodos también puede combinarse con ensayos para determinar otras condiciones prenatales asociadas a la madre y/o el feto. Por ejemplo, el método puede usarse conjuntamente con análisis prenatales, por ejemplo, como se describe en las publicaciones de solicitud de patente de EE.UU. N° US2010/0112590A1, US2009/0162842A1, US2007/0207466A1 y US2001/0051341A1.

Los métodos descritos pueden aplicarse a determinar la fracción de una población cualquiera de ácidos nucleicos en una mezcla de ácidos nucleicos contribuida por diferentes genomas. Además de determinar la fracción contribuida a una muestra por dos individuos, por ejemplo, los diferentes genomas están contribuidos por el feto y la madre que

lleva el feto, los métodos pueden usarse para determinar la fracción de un genoma en una mezcla derivada de dos células diferentes de un individuo, por ejemplo, los genomas son contribuidos a la muestra por células cancerosas aneuploides y células euploides normales del mismo sujeto.

5 Composiciones y kits

La presente invención también está relacionada con composiciones y kits o sistemas de reactivos útiles para poner en práctica los métodos descritos en el presente documento.

10 Las composiciones desveladas en el presente documento pueden incluirse en kits para la secuenciación masivamente paralela de mezclas de moléculas de ácidos nucleicos fetales y maternas, por ejemplo, ADNic, presentes en una muestra materna, por ejemplo, una muestra de plasma. Los kits comprenden una composición que comprende al menos un conjunto de cebadores para amplificar al menos un ácido nucleico diana polimórfico en dichas moléculas de ácidos nucleicos fetales y maternas. Los ácidos nucleicos diana polimórficos pueden comprender sin limitación polimorfismos de un único nucleótido (SNPs), SNP en tándem, deleciones o inserciones de múltiples bases a pequeña escala, llamados IN-DELS (también llamados polimorfismos de inserción/delección o DIP), polimorfismos de múltiples nucleótidos (MNP), repeticiones cortas en tándem (STR), polimorfismo en la longitud del fragmento de restricción (RFLP), o un polimorfismo que comprende cualquier otro cambio de secuencia en un cromosoma. Métodos de secuenciación que utilizan las composiciones desveladas en el presente documento son métodos de NGS de moléculas de ácidos nucleicos individuales o moléculas de ácidos nucleicos clónicamente amplificada como se describe en el presente documento. La métodos de secuenciación masivamente paralela de NGS incluyen pirosecuenciación, secuenciación por síntesis con terminadores de colorante reversible, secuenciación en tiempo real o secuenciación por ligación de sondas de oligonucleótidos.

25 En una realización, las composiciones incluyen cebadores para amplificar ácidos nucleicos diana polimórficos que comprenden cada uno al menos un SNP. En una realización, el al menos un SNP está seleccionado de los SNP rs560681, rs1109037, rs9866013, rs13182883, rs13218440, rs7041158, rs740598, rs10773760, rs4530059, rs7205345, rs8078417, rs576261, rs2567608, rs430046, rs9951171, rs338882, rs10776839, rs9905977, rs1277284, rs258684, rs1347696, rs508485, rs9788670, rs8137254, rs3143, rs2182957, rs3739005, y rs530022. Conjuntos de cebadores correspondientes a modo de ejemplo para amplificar los SNP se proporcionan como SEC ID N°:57-112.

30 En otra realización, la composición incluye cebadores para amplificar ácidos nucleicos diana polimórficos que comprenden cada uno al menos un SNP en tándem. En una realización, la composición incluye cebadores para amplificar SNP en tándem. En una realización, la composición incluye cebadores para amplificar los SNP en tándem desvelados en el presente documento, y la composición comprende los cebadores a modo de ejemplo correspondientes de SEC ID N°: 197-310.

35 En otra realización, la composición incluye cebadores para amplificar ácidos nucleicos diana polimórficos que comprenden cada uno al menos una STR. STR a modo de ejemplo incluyen CSF1PO, FGA, TH01, TPOX, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, D2S1338, Penta D, Penta E, D22S1045, D20S1082, D20S482, D18S853, D17S1301, D17S974, D14S1434, D12ATA63, D11S4463, D10S1435, D10S1248, D9S2157, D9S1122, D8S1115, D6S1017, D6S474, D5S2500, D4S2408, D4S2364, D3S4529, D3S3053, D2S1776, D2S441, D1S1677, D1S1627 y D1GATA113, que pueden amplificarse por los conjuntos correspondientes de cebadores de SEC ID N°: 1 13-196.

40 Los kits pueden contener una combinación de reactivos que incluye los elementos requeridos para realizar un ensayo según los métodos desvelados en el presente documento. El sistema de reactivos se presenta en una forma comercialmente envasada, como una composición o mezcla si lo permite la compatibilidad de los reactivos, en una configuración de dispositivo de prueba, o más normalmente como un kit de prueba, es decir, una combinación envasada de uno o más recipientes, dispositivos o similares que contienen los reactivos necesarios, y preferentemente que incluye instrucciones escritas para la realización de los ensayos. El kit desvelado en el presente documento puede adaptarse para cualquier configuración de ensayo y puede incluir composiciones para realizar cualquiera de los diversos formatos de ensayo descritos en el presente documento. En el presente documento se desvelan kits para determinar la fracción fetal comprenden composiciones que incluyen conjuntos de cebadores para amplificar ácidos nucleicos polimórficos presentes en una muestra materna como se ha descrito y, si procede, reactivos para purificar el ADNic. En una realización, un kit diseñado para permitir la cuantificación de secuencias polimórficas fetales y maternas, por ejemplo, STR y/o SNP en tándem, en una muestra de plasma de ADNic, incluye al menos un conjunto de oligonucleótidos específicos de alelo específicos para un SNP seleccionado y/o región de repeticiones en tándem. Preferentemente, el kit incluye una pluralidad de conjuntos de cebadores para amplificar un panel de secuencias polimórficas. Un kit puede comprender otros reactivos y/o información para genotipar o cuantificar alelos en una muestra (por ejemplo, tampones, nucleótidos, instrucciones). Los kits también incluyen una pluralidad de recipientes de tampones y reactivos apropiados.

45 La presente invención se describe en más detalle en los siguientes ejemplos que de ninguna forma tienen previsto limitar el alcance de la invención como se ha reivindicado. Las figuras adjuntas pretenden ser consideradas partes integrales de la memoria descriptiva y la descripción de la invención. Los siguientes ejemplos se ofrecen para

ilustrar, pero no para limitar, la invención reivindicada.

PARTE EXPERIMENTAL

5 Ejemplo 1

Determinación de la fracción fetal usando secuenciación masivamente paralela: Procesamiento de muestras y extracción de ADNlc

10 Se recogieron muestras de sangre periférica de mujeres embarazadas en su primer o segundo trimestre de embarazo y que se consideraron en riesgo de aneuploidía fetal. Se obtuvo el consentimiento informado de cada participante antes de tomar la sangre. Se recogió sangre antes de la amniocentesis o muestreo de vello coriónico. Se realizó el análisis del cariotipo usando muestras de vello coriónico o de amniocentesis para confirmar el cariotipo fetal.

15 La sangre periférica extraída de cada sujeto se recogió en tubos de ACD. Se transfirió un tubo de muestra de sangre (aproximadamente 6-9 ml/tubo) a un tubo de centrifuga de baja velocidad de 15 ml. La sangre se centrifugó a 2640 rpm, 4 °C durante 10 min usando la centrifuga Beckman Allegra 6 R y el rotor modelo GA 3.8.

20 Para la extracción de plasma libre de células, la capa de plasma superior se transfirió a un tubo de centrifuga de alta velocidad de 15 ml y se centrifugó a 16000 x g, 4 °C durante 10 min usando la centrifuga Beckman Coulter Avanti J-E, y el rotor JA-14. Las dos etapas de centrifugación se realizaron en el plazo de 72 h después de la extracción de sangre. El plasma libre de células que comprende ADNlc se almacenó a -80 °C y se congeló solo una vez antes de la amplificación del ADNlc de plasma o para la purificación de ADNlc.

25 Se extrajo ADN libre de células purificado (ADNlc) de plasma libre de células usando el kit QIAamp Blood DNA Mini (Qiagen) esencialmente según las instrucciones del fabricante. Se añadieron un mililitro de tampón AL y 100 µl de disolución de proteasa a 1 ml de plasma. La mezcla se incubó durante 15 minutos a 56 °C. Se añadió un mililitro de 100 % de etanol al digesto de plasma. La mezcla resultante se transfirió a columnas QIAamp mini que se ensamblaron con VacValves y VacConnectors proporcionados en el ensamblaje de columna QIAvac 24 Plus (Qiagen). Se aplicó vacío a las muestras, y el ADNlc retenido sobre los filtros de la columna se lavó a vacío con 750 µl de tampón AW1, seguido de un segundo lavado con 750 µl de tampón AW24. La columna se centrifugó a 14.000 rpm durante 5 minutos para eliminar cualquier tampón residual del filtro. El ADNlc se eluyó con tampón AE por centrifugación a 14.000 rpm, y la concentración se determinó usando la plataforma de cuantificación Qubit™ (Invitrogen).

Ejemplo 2

40 Determinación de la fracción fetal usando secuenciación masivamente paralela: Preparación de bibliotecas de secuenciación, secuenciación y análisis de datos de secuenciación

a. Preparación de bibliotecas de secuenciación

45 Todas las bibliotecas de secuenciación, es decir, bibliotecas diana, primarias y enriquecidas, se prepararon a partir de aproximadamente 2 ng de ADNlc purificado que se extrajo de plasma materno. La preparación de bibliotecas se realizó usando reactivos del conjunto de reactivos de ADN 1 NEBNext™ DNA Sample Prep (Parte N° E6000L; New England Biolabs, Ipswich, MA) para Illumina® del siguiente modo. Debido a que el ADN de plasma libre de células está fragmentado en la naturaleza, no se hizo fragmentación adicional por nebulización o sonicación en las muestras de ADN de plasma. Los nucleótidos protuberantes de aproximadamente 2 ng de fragmentos de ADNlc purificado

50 contenidos en 40 µl se convirtieron en extremos romos fosforilados según el módulo NEBNext® End Repair incubando en un tubo de microcentrifuga de 1,5 ml el ADNlc con 5 µl de 10X tampón de fosforilación, 2 µl de mezcla de disolución de desoxinucleótidos (10 mM de cada dNTP), 1 µl de una dilución 1:5 de ADN polimerasa I, 1 µl de ADN polimerasa T4 y 1 µl de polinucleótido cinasa T4 proporcionado en el conjunto de reactivos de ADN 1 NEBNext™ DNA Sample Prep durante 15 minutos a 20 °C. A continuación, las enzimas se inactivaron por calor

55 incubando la mezcla de reacción a 75 °C durante 5 minutos. La mezcla se enfrió a 4 °C, y se llevó a cabo prolongación de dA del ADN de extremos romos usando 10 µl de la mezcla maestra de prolongación con dA que contenía el fragmento de Klenow (3' a 5' exo menos) (conjunto de reactivos de ADN 1 NEBNext™ DNA Sample Prep), e incubando durante 15 minutos a 37 °C. Posteriormente, el fragmento de Klenow se inactivó por calor incubando la mezcla de reacción a 75 °C durante 5 minutos. Tras la inactivación del fragmento de Klenow se usó 1

60 µl de una dilución 1:5 de Illumina Genomic Adaptor Oligo Mix (Parte N° 1000521; Illumina Inc., Hayward, CA) para ligar los adaptadores de Illumina (Non-Index Y-Adaptors) al ADN prolongado con dA usando 4 µl de la ADN ligasa T4 proporcionada en el conjunto de reactivos de ADN 1 NEBNext™ DNA Sample Prep, incubando la mezcla de reacción durante 15 minutos a 25 °C. La mezcla se enfrió a 4 °C, y el ADNlc ligado al adaptador se purificó de adaptadores sin ligar, dímero de adaptador, y otros reactivos usando perlas magnéticas proporcionadas en el sistema de purificación por PCR Agencourt AMPure XP (Parte N° A63881; Beckman Coulter Genomics, Danvers, MA). Se realizaron dieciocho ciclos de PCR para enriquecer selectivamente el ADNlc ligado al adaptador usando la

mezcla maestra Phusion ® High-Fidelity (Finnzymes, Woburn, MA) y cebadores de PCR de Illumina complementarios a los adaptadores (Parte Nº 1000537 y 1000537). El ADN ligado al adaptador se sometió a PCR (98 °C durante 30 segundos; 18 ciclos de 98 °C durante 10 segundos, 65 °C durante 30 segundos y 72 °C durante 30 segundos; extensión final a 72 °C durante 5 minutos, y mantenimiento a 4 °C) usando los cebadores de PCR genómicos de Illumina (Partes Nº 100537 y 1000538) y la mezcla maestra Phusion HF PCR proporcionada en el conjunto de reactivos de ADN 1 NEBNext™ DNA Sample Prep, según las instrucciones del fabricante. El producto amplificado se purificó usando el sistema de purificación por PCR Agencourt AMPure XP (Agencourt Bioscience Corporation, Beverly, MA) según las instrucciones del fabricante disponibles en www.beckmangenomics.com/products/AMPureXPProtocol_000387v001.pdf. El producto purificado amplificado se eluyó en 40 µl de tampón Qiagen EB, y la concentración y la distribución de tamaño de las bibliotecas amplificadas se analizó usando el kit Agilent DNA 1000 para el bioanalizador 2100 (Agilent Technologies Inc., Santa Clara, CA).

b. Secuenciación

Se realizó secuenciación de ADN de biblioteca usando Genome Analyser II (Illumina Inc., San Diego, CA, EE.UU.) según los protocolos estándar del fabricante. Copias del protocolo para la secuenciación del genoma completo usando tecnología de Illumina/Solexa pueden encontrarse en BioTechniques.RTM. Protocol Guide 2007 publicado en diciembre de 2006: p 29, y en la malla mundial en biotechniques.com/default.asp?page=protocol&subsection=article_display&id=112378. La biblioteca de ADN se diluyó a 1 nM y se desnaturalizó. El ADN de la biblioteca (5 pM) se sometió a amplificación agrupada según el procedimiento descrito en Cluster Station User Guide y Cluster Station Operations Guide de Illumina, disponibles en la malla mundial en illumina.com/systems/genome_analyzer/cluster_station.ilmn. El ADN amplificado se secuenció usando Genome Analyser II de Illumina para obtener lecturas de extremos individuales de 36 pb. Solo se necesitan aproximadamente 30 pb de información de secuencia aleatoria para identificar una secuencia como que pertenece a un cromosoma humano específico. Secuencias más largas pueden identificar únicamente dianas más particulares. En el presente caso se obtuvieron un gran número de lecturas de 36 pb, cubriendo aproximadamente el 10 % del genoma.

c. Análisis de datos de secuenciación para la determinación de la fracción fetal

Tras completarse la secuenciación de la muestra, el “software de control del secuenciador” de Illumina transfirió la imagen y los archivos de llamada de bases a un servidor de Unix que ejecutaba el software de Illumina “Genome Analyser Pipeline” versión 1.51. Las lecturas de 36 pb se alinearon con un genoma de referencia artificial, por ejemplo, un genoma de SNP, usando el programa BOWTIE. El genoma de referencia artificial se identificó como la agrupación de las secuencias polimórficas de ADN que engloban los alelos comprendidos en las secuencias polimórficas diana. Por ejemplo, el genoma de referencia artificial es un genoma de SNP que comprende SEC ID Nº: 1-56. Solo las lecturas que se mapearon únicamente con el genoma artificial se usaron para el análisis de la fracción fetal. Las lecturas que coincidieron perfectamente con el genoma de SNP se contaron como marcas y se filtraron. De las restantes lecturas, solo las lecturas que tenían uno o dos desapareamientos se contaron como marcas y se incluyeron en el análisis. Se contaron las marcas mapeadas en cada uno de los alelos polimórficos, y la fracción fetal se determinó como un porcentaje de la relación del número de marcas mapeadas en el alelo principal, es decir, alelo materno, y el número de marcas mapeadas en el alelo secundario, es decir, alelo fetal.

Ejemplo 3

Selección de SNP autosómicos para la determinación de la fracción fetal

Se seleccionaron un conjunto de 28 SNP autosómicos de una lista de 92 SNP (Pakstis et al., Hum Genet 127:315-324 [2010]) y de Applied Biosystems por Life Technologies™ (Carlsbad, CA) en la dirección de la malla mundial appliedbiosystems.com, y se validaron para su uso en amplificación por PCR de múltiple. Se diseñaron cebadores para hibridarse con una secuencia próxima al sitio de SNP en el ADNlc para garantizar que se incluiría en la lectura de 36 pb generada a partir de la secuenciación masivamente paralela en el Analyser GII de Illumina, y para generar amplicones de longitud suficiente para someterse a amplificación de puentes durante la formación de agrupaciones. Así, se diseñaron cebadores para generar amplicones que tuvieran al menos 110 pb, que cuando se combinaron con los adaptadores universales (Illumina Inc., San Diego, CA) usados para la amplificación de agrupaciones, produjeron moléculas de ADN de al menos 200 pb. Se identificaron secuencias de cebadores, y se sintetizaron conjuntos de cebadores, es decir, cebadores directos e inversos, por Integrated DNA Technologies (San Diego, CA), y se guardaron como una disolución 1 µM que iba a usarse para amplificar secuencias polimórficas diana como se describe en los Ejemplos 4-7. La Tabla 1 proporciona los números ID de acceso de RefSNP (rs), los cebadores usados para amplificar la secuencia de ADNlc diana y las secuencias de los amplicones que comprenden los posibles alelos de SNP que se generarían usando los cebadores. Los SNP dados en la Tabla 1 se usaron para la amplificación simultánea de 13 secuencias diana en un ensayo multiplexado. El panel proporcionado en la Tabla 1 es un panel de SNP a modo de ejemplo. Pueden emplearse menos o más SNP para enriquecer el ADN fetal y para ácidos nucleicos diana polimórficos. Los SNP adicionales que pueden usarse incluyen los SNP dados en la Tabla 2. Los alelos de SNP se muestran en negrita y están subrayados. Otros SNP adicionales que pueden usarse para determinar la fracción fetal según el presente método incluyen rs315791, rs3780962, rs1410059, rs279844, rs38882,

ES 2 560 929 T3

rs9951171, rs214955, rs6444724, rs2503107, rs1019029, rs1413212, rs1031825, rs891700, rs1005533, rs2831700, rs354439, rs1979255, rs1454361, rs8037429 y rs1490413, que se han analizado para determinar la fracción fetal por PCR TaqMan, y se desvelan en las solicitudes provisionales de EE.UU. 61/296.358 y 61/360.837.

5

TABLA 1

SNP Panel para la determinación de la fracción fetal					
SNP ID	Chr	Amplicón: alelo 1	Amplicón: alelo 2	Cebador directo secuencia, nombre y SEQ ID NO.	Cebador inverso secuencia, nombre y SEQ ID NO.
rs560681	1	CACATGCACAGCCA	CACATGCACAGCCA	CACATGCA	CCCCAAGGTCC
		GCAACCCTGTCAGC	GCAACCCTGTCAGC	CAGCCAGC	TGTGACCTGAG
		AGGAGTTCCCACCA	AGGAGTTCCCACCA	AACCC	T
		GTTTCTTTCTGAGAA	GTTTCTTTCTGAGAA	(rs560681_C1	(rs560681_C1_1_R
		CATCTGTTCAGGTTT	CATCTGTTCAGGTTT	_1_F; SEQ ID	; SEQ ID NO:58)
		CTCTCCATCTCTATT	CTCTCCATCTCTGTT	NO:57)	
		TACTCAGGTCACAG	TACTCAGGTCACAG		
GACCTTGGGG (SEQ ID NO:1)	GACCTTGGGG (SEQ ID NO:2)				
rs109037	2	TGAGGAAGTGAGGC	TGAGGAAGTGAGGC	TGAGGAAG	TGCCAGTGCGA
		TCAGAGGGTAAGAA	TCAGAGGGTAAGAA	TGAGGCTC	GATGAAAGTCT
		ACTTTGTCACAGAGC	ACTTTGTCACAGAGC	AGAGGGT	TT
		TGGTGGTGAGGGTG	TGGTGGTGAGGGTG	(rs110937_C2	(rs110937_C2_1_R
		GAGATTTTACACTCC	GAGATTTTACACTCC	_1_F; SEQ ID	; SEQ ID NO:60)
		CTGCCTCCCACACCA	CTGCCTCCCACACCA	NO:59)	
		GTTTCTCCAGAGTGG	GTTTCTCCGAGTGG		
AAAGACTTTCATCTC	AAAGACTTTCATCTC				
GCACTGGCA (SEQ ID NO:3)	GCACTGGCA (SEQ ID NO:4)				

50

55

60

65

(continúa)

SNP Panel para la determinación de la fracción fetal					
SNP ID	Chr	Amplificón: alelo 1	Amplificón: alelo 2	Cebador directo secuencia, nombre y SEQ ID NO.	Cebador inverso secuencia, nombre y SEQ ID NO.
rs9866013	3	GTGCCTTCAGAACCT TTGAGATCTGATTCT ATTTTAAAGCTTCT TAGAAGAGAGATTG CAAAGTGGGTTGTTT CTCTAGCCAGACAG GGCAGGCAAATAGG GGTGGCTGGTGGGA TGGGA (SEQ ID NO:5)	GTGCCTTCAGAACCT TTGAGATCTGATTCT ATTTTAAAGCTTCT TAGAAGAGAGATTG CAAAGTGGGTTGTTT CTCTAGCCAGACAG GGCAGGTAATAGG GGTGGCTGGTGGGA TGGGA (SEQ ID NO:6)	GTGCCTTCA GAACCTTTG AGATCTGAT (rs9866013_C 3_1_F; SEQ ID NO:61)	TCCCATCCCAC CAGCCACCC (rs9866013_C3_1_ R; SEQ ID NO:62)
rs13182883	5	AGGTGTGTCTCTCTT TTGTGAGGGGAGGG GTCCCTTCTGGCCTA GTAGAGGGCCTGGC CTGCAGTGAGCATTG AAATCCTCAGGAA CAGGGTGGGAGGT GGGACAAAGG (SEQ ID NO:7)	AGGTGTGTCTCTCTT TTGTGAGGGGAGGG GTCCCTTCTGGCCTA GTAGAGGGCCTGGC CTGCAGTGAGCATTG AAATCCTCAGGAA CAGGGTGGGAGGT GGGACAAAGG (SEQ ID NO:8)	AGGTGTGTC TCTCTTTG TGAGGGG (rs13182883_ C5_1_F; SEQ ID NO:63)	CCTTTGTCCCAC CTCCCCACC (rs13182883_C5_1 _R; SEQ ID NO:64)
rs740598	10	GAAATGCCTTCTCAG GTAATGGAAGGTTA TCCAAATATTTTTCG TAAGTATTTCAAATA GCAATGGCTCGTCTA TGGTTAGTCTCAG CCACATTCTCAGAAC TGCTCAAACC (SEQ ID NO:13)	GAAATGCCTTCTCAG GTAATGGAAGGTTA TCCAAATATTTTTCG TAAGTATTTCAAATA GCAATGGCTCGTCTA TGGTTAGTCTCAG CCACATTCTCAGAAC TGCTCAAACC (SEQ ID NO:14)	GAAATGCC TTCTCAGGT AATGGAAG GT (SEQ ID NO:69)	GGTTTGAGCAG TTCTGAGAATG TGGCT (SEQ ID NO:70)
rs10773760	12	ACCCAAAACACTGG AGGGGCCTCTTCTCA TTTTCGGTAGACTGC AAGTGTTAGCCGTC GGGACCAGCTTCTGT CTGGAAGTTCGTCA AATTGCAGTTAGTCA CAAGTATGCCAAT AGCAGATAAGGG (SEQ ID NO:15)	ACCCAAAACACTGG AGGGGCCTCTTCTCA TTTTCGGTAGACTGC AAGTGTTAGCCGTC GGGACCAGCTTCTGT CTGGAAGTTCGTCA AATTGCAGTTAGTCA CAAGTATGCCAAT TAGCAGATAAGGG (SEQ ID NO:16)	ACCCAAAA CACTGGAG GGGCCT (SEQ ID NO:71)	CCCTTATCTGCT ATGTGGCATAAC TTGG (SEQ ID NO:72)

(continúa)

SNP Panel para la determinación de la fracción fetal					
SNP ID	Chr	Amplicón: alelo 1	Amplicón: alelo 2	Cebador directo secuencia, nombre y SEQ ID NO.	Cebador inverso secuencia, nombre y SEQ ID NO.
rs4530059	14	GCACCCAGAATTTAA ACAACGCTGACAAT AAATATGCAGTCGA TGATGACTTCCCAGA GCTCCAGAAGCAAC TCCAGCACACAGAG AGGCGCTGATGTGC CTGTCAGGTGC (SEQ ID NO:17)	GCACCAGAATTTAA ACAACGCTGACAAT AAATATGCAGTCGA TGATGACTTCCCAGA GCTCCAGAAGCAAC TCCAGCACACGGAG AGGCGCTGATGTGC CTGTCAGGTGC (SEQ ID NO:18)	GCACCAGA ATTTAAACA ACGCTGAC AA (SEQ ID NO:73)	GCACCTGACAG GCACATCAGCG (SEQ ID NO:74)
rs7205345	16	TGACTGTATAACCCCA GGTGCACCCCTGGGT CATCTCTATCATAGA ACTTATCTCACAGAG TATAAGAGCTGATTT CTGTGTCTGCCTCTC ACACTAGACTTCCAC ATCCTTAGTGC (SEQ ID NO:19)	TGACTGTATAACCCCA GGTGCACCCCTGGGT CATCTCTATCATAGA ACTTATCTCACAGAG TATAAGAGCTGATTT CTGTGTCTGCCTGTC ACACTAGACTTCCAC ATCCTTAGTGC (SEQ ID NO:20)	TGACTGTAT ACCCAGG TGCACCC (SEQ ID NO:75)	GCACCTAAGGAT GTGGAAGTCTA GTGTG (SEQ ID NO:76)
rs8078417	17	TGTACGTGGTCACCA GGGGACGCCTGGCG CTGCGAGGGAGGCC CCGAGCCTCGTGCCC CCGTGAAGCTTCAG CTCCCCTCCC ^U GGCT GTCCTTGAGGCTCTT CTCACACT (SEQ ID NO:21)	TGTACGTGGTCACCA GGGGACGCCTGGCG CTGCGAGGGAGGCC CCGAGCCTCGTGCCC CCGTGAAGCTTCAG CTCCCCTCCC ^U GGCT GTCCTTGAGGCTCTT CTCACACT (SEQ ID NO:22)	TGTACGTGG TCACCAGG GGACG (SEQ ID NO:77)	AGTGTGAGAAG AGCCTCAAGGA CAGC (SEQ ID NO:78)
rs576261	19	CAGTGGACCCTGCT GCACCTTTCCTCCCC TCCCATCAACCTCTT TTGTGCCTCCCCCTC CGTGTACCACCTTCT CTGTCACCA ^A CCCTG GCCTCACAACCTCTCT CCTTTGCCAC (SEQ ID NO:23)	CAGTGGACCCTGCT GCACCTTTCCTCCCC TCCCATCAACCTCTT TTGTGCCTCCCCCTC CGTGTACCACCTTCT CTGTCACCA ^C CCCTG GCCTCACAACCTCTCT CCTTTGCCAC (SEQ ID NO:24)	CAGTGGAC CCTGCTGCA CCTT (SEQ ID NO:79)	GTGGCAAAGGA GAGAGTTGTGA GG (SEQ ID NO:80)

(continúa)

SNP Panel para la determinación de la fracción fetal					
SNP ID	Chr	Amplicón: alelo 1	Amplicón: alelo 2	Cebador directo secuencia, nombre y SEQ ID NO.	Cebador inverso secuencia, nombre y SEQ ID NO.
rs2567608	20	CAGTGGCATAGTAG	CAGTGGCATAGTAG	CAGTGGCA	CCTCTCCGACA
		TCCAGGGGCTCCTCC	TCCAGGGGCTCCTCC	TAGTAGTCC	ACTTCCGCCG
		TCAGCACCTCCAGC	TCAGCACCTCCAGC	AGGGGCT	(SEQ ID NO:82)
		ACCTTCCAGGAGGC	ACCTTCCAGGAGGC	(SEQ ID	
		AGCAGCGCAGGCAG	AGCAGCGCAGGCAG	NO:81)	
		AGAACCCGCTGGAA	AGAACCCGCTGGAA		
		GAATCGGCGGAAGT	GGATCGGCGGAAGT		
		TGTCGGAGAGG (SEQ ID NO:25)	TGTCGGAGAGG (SEQ ID NO:26)		

25

30

35

40

45

50

55

60

65

TABLA 2

SNPs Adicionales para la determinación de la fracción fetal					
SNP ID	Chr	Amplicón: alelo 1	Amplicón: alelo 2	Cebador directo secuencia, nombre y SEQ ID NO.	Cebador inverso secuencia, nombre y SEQ ID NO.
rs430046	16	AGGTCTGGGGGCG GCTGAATGCCAAGC TGGGAATCTTAAAT GTAAAGGAACAAG GTCATAAATGAAT GGTGTGATGTAAAA GCTTGGGAGGTGAT TTCTGAGGGTAGGT GCTGGGTTAATGG GAGGA (SEQ ID NO:27)	AGGTCTGGGGGCCG TGAATGCCAAGCTGG GAATCTTAAATGTTA AGGAACAAGTCATA CAATGAATGGTGTGA TGTAAGGCTTGGGA GGTGATTTTGAGGG TAGGTGCTGGGTTA ATGGGAGGA (SEQ ID NO:28)	AGGTCTGG GGGCCGCT GAAT (rs430046_C1 _1_F; SEQ ID NO:83)	TCCTCCCATTA AACCCAGCACC T (rs430046_C1_1_R ; SEQ ID NO:84)
rs9951171	18	ACGGTTCCTGCTCG TAGGGGAGAAAAG TCCTCGTTGTTCCCT CTGGGATGCAACAT GAGAGAGCAGCAC ACTGAGGCTTTATG GATTGCCCTGCCAC AAGTGAACAGG (SEQ ID NO:29)	ACGGTTCCTGCTCGT AGGGGAGAAAAGTCC TCGTGTTCTCTGGG ATGCAACATGAGAGA GCAGCACACTGAGGC TTTATGGGTTGCCCT GCCACAAGTGAACAG G (SEQ ID NO:30)	ACGGTTCCTG TCCTGTAGG GGAGA (rs9951171_C 1_1_F; SEQ ID NO:85)	CCTGTTCACTTG TGGCAGGGCA (rs9951171_C1_1_ R; SEQ ID NO:86)
rs338882	5	GCGCAGTCAGATG GGCGTCTGGCGTC	GCGCAGTCAGATGGG CGTGCTGGCGTCTGT	GCGCAGTC AGATGGGC	TCCAGCCCTTG TCCCAAACGTG
		TGCTCTCTCTCTC CTGCTCTCGGCTT CAITTTCTCTCCTT CTGTCTCACCTTCT TTCGTGTGCCTGTG CACACACACGTTTG GGACAAGGG CTGGA (SEQ ID NO:31)	CTTCTCTCTCTCTG TCTCTGGCTTCATTTT TCTCTCCTCTGTCTC ACCTTCTTTCGTGTG CTGTGCAACACACG TTTGGGACAAGGG CTGGA (SEQ ID NO:32)	GTGC (rs338882_C1 _1_F; SEQ ID NO:87)	T (rs338882_C1_1_R ; SEQ ID NO:88)
rs10776839	9	GCCGGACCTGCGA AATCCAAAATGCC AAACATTCGCCCT CACATGATCCAGA GAGAGGGGACCCA GTGTTCCAGCTTG CAGCTGAGGAGCC CGAGTTGCGGTCA GATCAGAGCCCA GTTGCCCG (SEQ ID NO:33)	GCCGGACCTGCGAAA TCCAAAATGCCAAA CATTCCCGCTCACA TGATCCAGAGAGAG GGGACCCAGTGTTC CAGCTGAGCTGAG GAGCCCGAGTTGCC GTCAGATCAGAGCCC CAGTTGCCCG (SEQ ID NO:34)	GCCGGACC TGCGAAAT CCCAA (rs10776839 1_1_F; SEQ ID NO:89)	CGGGCAACTGG GGCTCTGATC (rs10776839_C1_1 _R; SEQ ID NO:90)

ES 2 560 929 T3

SNPs Adicionales para la determinación de la fracción fetal					
SNP ID	Chr	Amplicón: alelo 1	Amplicón: alelo 2	Cebador directo secuencia, nombre y SEQ ID NO.	Cebador inverso secuencia, nombre y SEQ ID NO.
rs9905977	17	AGCAGCCTCCCTCG ACTAGCTCACACTA CGATAAGGAAAAT TCATGAGCTGGTGT CCAAGGAGGGCTG GGTGACTCGTGGCT CAGTCAGCATCAAG ATTCTTTCGTCTTT CCCCTCTGCC (SEQ ID NO:35)	AGCAGCCTCCCTCGA CTAGCTCACACTACG ATAAGGAAAATTCAT GAGCTGGTGTCCAAG GAGGGCTGGGTGACT CGTGGCTCAGTCAGC GTCAAGATTCCTTTC GTCTTCCCCTCTGCC (SEQ ID NO:36)	AGCAGCCT CCCTCGACT AGCT (rs9905977_C 1_1_F; SEQ ID NO:91)	GGCAGAGGGGA AAGACGAAAGG A (rs9905977_C1_1_ R; SEQ ID NO:92)
rs1277284	4	TGGCATTGCCTGTA ATATACATAGCCAT GGTTTTTATAGGC AATTTAAGATGAAT AGCTTCTAAACTAT AGATAAGTTTCATT ACCCAGGAAGCT GAACTATAGCTACT TTACCCAAAATCAT TAGAATGGTGCTT (SEQ ID NO:37)	TGGCATTGCCTGTAA TATACATAGCCATGG TTTTTATAGGCAATT TAAGATGAATAGCTT CTAAACTATAGATAA GTTTCATTACCCAG GAAGCTGAACTATAG CTACTTTC C CCAAAA TCATTAGAATGGTGC TT (SEQ ID NO:38)	TGGCATTGC CTGTAATAT ACATAG (rs1277284_C 4_1_F; SEQ ID NO:93)	AAGCACCATTTC TAATGATTTTG G (rs1277284_C4_1_ R; SEQ ID NO:94)
rs258684	7	ATGAAGCCTTCCAC CAACTGCCTGTATG ACTCATCTGGGGAC TTCTGCTCTATACT CAAAGTGGCTTAGT CACTGCCAATGTAT TTCCATATGAGGGA CGATGATTACTAAG GAAATATAGAAAC AACAACTGATC (SEQ ID NO:39)	ATGAAGCCTTCCACC AACTGCCTGTATGAC TCATCTGGGGACTTC TGCTCTATACTCAAA GTGGCTTAGTCACTG CCAATGTATTTCCAT ATGAGGGACG G TGAT TACTAAGGAAATATA GAAACAACAACATGAT C (SEQ ID NO:40)	ATGAAGCC TTCCACCAA CTG (rs258684_C7 _1_F; SEQ ID NO:95)	GATCAGTTGTT GTTTCTATATTT CCTT (rs258684_C7_1_R ; SEQ ID NO:96)
rs1347696	8	ACAACAGAATCAG GTGATTGGAGAAA AGATCACAGGCCTA GGCACCCAAGGCTT GAAGGATGAAAGA ATGAAAGATGGAC GGAA C AAAATTAG GACCTTAATCTTT GTTTCAGTTCAG (SEQ ID NO:41)	ACAACAGAATCAGGT GATTGGAGAAAAGAT CACAGGCCTAGGCAC CCAAGGCTTGAAGGA TGAAAGAATGAAAGA TGGACGGAA G AAAAT TAGGACCTTAATCTTT TGTTTCAGTTCAG (SEQ ID NO:42)	ACAACAGA ATCAGGTG ATTGGA (rs1347696_C 8_4_F; SEQ ID NO:97)	CTGAACTGAAC AAAGAATTAAG GTC (rs1347696_C8_4_ F; SEQ ID NO:98)

(continúa)

SNPs Adicionales para la determinación de la fracción fetal					
SNP ID	Chr	Amplicón: alelo 1	Amplicón: alelo 2	Cebador directo secuencia, nombre y SEQ ID NO.	Cebador inverso secuencia, nombre y SEQ ID NO.
rs508485	11	TTGGGGTAAATTTT CATTGTCATATGTG GAATTTAAATATAC CATCATCTACAAAG AATTCCACAGAGTT AAATATCTTAAGTT AAACACTTAAAATA AGTGTTCGCGTGAT ATTTTGATGACAGAG TAAACAGAGTCTAA TTCCCACCCC (SEQ ID NO:43)	TTGGGGTAAATTTTC ATTGTCATATGTGGA ATTTAAATATACCAT CATCTACAAAGAATT CCACAGAGTTAAATA TCTTAAGTTAAACAC TTAAAATAAGTGTTT GCGTGATATTTTGAT GATAGATAAACAGAG TCTAATCCCACCCC (SEQ ID NO:44)	TTGGGGTA AATTTTCAT TGCA (rs508485_C1 1_1_F; SEQ ID NO:99)	GGGGTGGGAAT TAGACTCTG (rs508485_C11_1_ R; SEQ ID NO:100)
rs9788670	15	TGCAATTCAAATCA GGAAGTATGACCA AAAGACAGAGATC TTTTTTGGATGATC CCTAGCCTAGCAAT	TGCAATTCAAATCAG GAAGTATGACCAAAA GACAGAGATCTTTTT TGGATGATCCCTAGC CTAGCAATGCCTGGC	TGCAATTC AATCAGGA AGTATG (rs9788670_c1 5_2_F; SEQ ID NO:101)	GCAACATCGAG GTTTGTGAG (rs9788670_c15_2 _R; SEQ ID NO:102)
		GCCTGGCAGCCATG CAGGTGCAATGTCA ACCTAAATAATGT ATTGCAAACTCAGA GCTGACAAACCTCG ATGTTGC (SEQ ID NO:45)	AGCCATGCAGGTGCA ATGTCAACCTTAAAT AATGTAATTGCAAAAT CAGAGCTGACAAACC TCGATGTTGC (SEQ ID NO:46)		
rs8137254	22	CTGTGCTCTGCGAA TAGCTGCAGAAGTA ACTTGGGGACCCAA AATAAAGCAGAAT GCTAATGTCAAGTC CTGAGAACCAAGC CCTGGGACTCTGGT GCCATTTGGGATTC TCCATGAGCATGGT (SEQ ID NO:47)	CTGTGCTCTGCGAAT AGCTGCAGAAGTAAC TTGGGGACCCAAAAT AAAGCAGAATGCTAA TGTCAAAGTCTGAGA ACCAAGCCCTGGGAC TCTGGTGCCATTTTG GATTCTCCATGAGCA TGGT (SEQ ID NO:48)	CTGTGCTCT GCGAATAG CTG (rs8137254_c2 2_2_F; SEQ ID NO:103)	ACCATGCTCAT GGAGAATCC (rs8137254_c22_2 _R; SEQ ID NO:104)

5

10

15

20

25

30

35

40

45

50

55

60

65

(continúa)

SNPs Adicionales para la determinación de la fracción fetal					
SNP ID	Chr	Amplicón: alelo 1	Amplicón: alelo 2	Cebador directo secuencia, nombre y SEQ ID NO.	Cebador inverso secuencia, nombre y SEQ ID NO.
rs3143	19	TTTTCCAGCCAAC TCAAGGCCAAAAA AAATTCTTAATAT AGTTATTATGCGAG GGGAGGGGAAGCA AAGGAGCACAGGT AGTCCACAGAATA AGACACAAGAAAC CTCAAGCTGTG (SEQ ID NO:49)	TTTTCCAGCCAAGTC AAGGCCAAAAAAAT TTCTTAATATAGTTAT TATGCGAGGGGAGGG GAAGCAAAGGAGCA CAGGTAGTCCACAGA ATAGGACACAAGAA ACCTCAAGCTGTG (SEQ ID NO:50)	TTTTCCAG CCAAGTCA AGG (rs3143_c19_2 _F: SEQ ID NO:105)	CAAGCTTGAG GTTTCTGTG (rs3143_c19_2_R; SEQ ID NO:106)
rs2182957	13	TCTTCTCGTCCCCT AAGCAAACAACAT CCGCTTGCTTCTGT CTGTGTAACCACAG TGAATGGGTGTGCA CGCTTGATGGGCCT CTGAGCCCCTGTTG CACAAACCAGAAA (SEQ ID NO:51)	TCTTCTCGTCCCCTAA GCAAACAACATCCGC TTGCTTCTGTCTGTGT AACCACAGTGAATGG GTGTGCACGCTTGGT GGGCCTCTGAGCCCC TGTTGCACAAACCAG AAA (SEQ ID NO:52)	TCTTCTCGT CCCCTAAGC AA (rs2182957_c1 3_1_F: SEQ ID NO:107)	TTTCTGGTTTGT GCAACAGG (rs2182957_c13_1 _R; SEQ ID NO:108)
rs3739005	2	CACATGGGGGCATT AAGAATCGCCCAG GGAGGAGGAGGGA GAACGCGTGCCTTTT	CACATGGGGGCATTA AGAATCGCCCAGGGA GGAGGAGGAGGAGAAC GCGTGCCTTTTCAATT	CAATGGG GGCATTAA GAAT (rs3739005_c2 _2_F: SEQ ID NO:109)	ACATCGATGAG CACAAAAACAC (rs3739005_c2_2_ R; SEQ ID NO:110)
		CACATTTGCATTTG AATTTTGGAGTTCC CAGGATGTGTTTTT GTGCTCATCGATGT (SEQ ID NO:53)	TGCATTTGAATTTTGG AGTCCCAGGATGTG TTTTGTGCTCATCGA TGT (SEQ ID NO:54)		
rs530022	1	GGGCTCTGAGGTGT GTGAAATAAAAAAC AAATGTCCATGTCT GTCCTTTTATGGCA TTTTGGGACTTTAC ATTTCAAACATTTT AGACATGTATCACA ACACGAAGGAATA ACAGTTCCAGGGAT ATCT (SEQ ID NO:55)	GGGCTCTGAGGTGTG TGAAATAAAAAACAAA TGTCATGTCTGTCTCT TTTATGGCATTITGGG ACTTTACATTTCAA CATTTAGACATGTA TCACAACACGAGGGA ATAACAGTTCCAGGG ATATCT (SEQ ID NO:56)	GGGCTCTG AGGTGTGT GAAA (rs530022_c1_ 2_F: SEQ ID NO:111)	AGATATCCCTG GAACTGTTATT CC (rs530022_c1_2_R ; SEQ ID NO:112)

Ejemplo 4

Determinación de la fracción fetal en secuenciación masivamente paralela de una biblioteca diana

5 Para determinar la fracción de ADNlc fetal en una muestra materna, se amplificaron secuencias de ácidos nucleicos polimórficas diana que comprendían cada una un SNP y se usaron para preparar una biblioteca diana para la secuenciación en un modo masivamente paralelo.

10 Se extrajo ADNlc como se describe en el Ejemplo 1. Se preparó una biblioteca de secuenciación diana del siguiente modo. Se amplificó el ADNlc contenido en 5 µl de ADNlc purificado en un volumen de reacción de 50 µl que contenía 7,5 µl de una mezcla de cebadores 11 µM (Tabla 1), 10 µl de mezcla maestra NEB 5X y 27 µl de agua. Se realizó ciclo térmico con el gen Amp9700 (Applied Biosystems) usando las siguientes condiciones de ciclo: incubar a 95 °C durante 1 minuto, seguido de 20-30 ciclos a 95 °C durante 20 segundos, 68 °C durante 1 minuto y 68 °C durante 30 s, que se siguió de un incubación final a 68 °C durante 5 minutos. Se añadió un mantenimiento final a 4 °C hasta que las muestras se extrajeron para la combinación con la porción sin amplificar de la muestra de ADNlc purificada. El producto amplificado se purificó usando el sistema de purificación por PCR Agencourt AMPure XP (Parte N° A63881; Beckman Coulter Genomics, Danvers, MA). Se añadió un mantenimiento final a 4 °C hasta que las muestras se extrajeron para preparar la biblioteca diana. El producto amplificado se analizó con un bioanalizador 2100 (Agilent Technologies, Sunnyvale, CA), y se determinó la concentración de producto amplificado. Se preparó una biblioteca de secuenciación de ácidos nucleicos diana amplificados como se describe en el Ejemplo 2, y se secuenció en un modo masivamente paralelo usando secuenciación por síntesis con terminadores de colorante reversible y según el protocolo de Illumina (BioTechniques.RTM. Protocol Guide 2007 publicado en diciembre de 2006: p 29, y en la malla mundial en biotechniques.com/default.asp?page=protocol&subsection=article display&id=112378). El análisis y recuento de marcas mapeadas en un genoma de referencia que consiste en 26 secuencias (13 pares que representan cada una dos alelos) que comprenden un SNP, es decir, SEC ID N°: 1-26, se realizó como se ha descrito.

La Tabla 3 proporciona los recuentos de marcas obtenidos de la secuenciación de la biblioteca diana, y la fracción fetal calculada derivada de datos de secuenciación.

TABLA 3

Determinación de fracción fetal por secuenciación paralela masiva de una librería de ácidos nucleicos polimórficos		
SNP	SNP Conteo Etiquetas	Fracción fetal (%)
rs10773760.1 Chr.12 longitud=128 alelo=A	236590	1.98
rs10773760.2 Chr.12 longitud=128 alelo=A	4680	
rs13182883.1 Chr.5 longitud=111 alelo=A	3607	4.99
rs13182883.2 Chr.5 longitud=111 alelo=G	72347	
rs4530059.1 Chr.14 longitud=110 alelo=A	3698	1.54
rs4530059.1 Chr.14 longitud=110 alelo=G	239801	
rs8078417.1 Chr.17 longitud=110 alelo=C	1E+06	3.66
rs8078417.2 Chr.17 longitud=110 alelo=T	50565	
Fracción fetal (Media $n \pm S.D.$) = 12.4 ± 6.6		

60

65

Los resultados muestran que las secuencias de ácidos nucleicos polimórficas que comprenden cada una al menos un SNP pueden amplificarse a partir del ADNlc derivado de una muestra de plasma materno para construir una biblioteca que puede secuenciarse en un modo masivamente paralelo para determinar la fracción de ácidos nucleicos fetales en la muestra materna.

Ejemplo 5

Determinación de la fracción fetal tras el enriquecimiento de ácidos nucleicos fetales y maternos en una muestra de biblioteca de secuencia de ADNlc.

Para enriquecer el ADNlc fetal y materno contenido en una biblioteca de secuenciación primaria construida usando ADNlc fetal y materno purificado, se usó una porción de una muestra de ADNlc purificada para amplificar secuencias de ácidos nucleicos polimórficas diana, y para preparar una biblioteca de secuenciación de ácidos nucleicos diana polimórficos amplificados, que se usó para enriquecer las secuencias de ácidos nucleicos fetales y maternos comprendidas en la biblioteca primaria.

El método se corresponde con el proceso de trabajo 3 representado en diagrama en la Figura 3. Se preparó una biblioteca de secuenciación diana a partir de una porción del ADNlc purificado como se describe en el Ejemplo 2. Se preparó una biblioteca de secuenciación primaria usando la porción restante del ADNlc purificado como se describe en el Ejemplo 2. El enriquecimiento de la biblioteca primaria para los ácidos nucleicos polimórficos amplificados comprendidos en la biblioteca diana se obtuvo diluyendo las bibliotecas de secuenciación primarias y diana a 10 nM, y combinando la biblioteca diana con la biblioteca primaria a una relación de 1:9 para proporcionar una biblioteca de secuenciación enriquecida. La secuenciación de la biblioteca enriquecida y el análisis de los datos de secuenciación se realizó como se describe en el Ejemplo 2.

La Tabla 4 proporciona el número de marcas de secuencia que se mapearon en el genoma de SNP para los SNP informativos identificados a partir de la secuenciación de una biblioteca enriquecida derivada de muestras de plasma de mujeres embarazadas que llevan cada una un feto T21, T13, T18 y de monosomía X, respectivamente. La fracción fetal se calculó del siguiente modo:

$$\% \text{ de alelo}_x \text{ de la fracción fetal} = ((\sum \text{Marcas de secuencia fetal para el alelo}_x) / (\sum \text{Marcas de secuencia maternas para el alelo}_x)) \times 100$$

La Tabla 4 también proporciona el número de marcas de secuencia mapeadas en el genoma de referencia humano. Las marcas mapeadas en el genoma de referencia humano se usaron para determinar la presencia o ausencia de aneuploidía usando la misma muestra de plasma que se utilizó para determinar la fracción fetal correspondiente. El método para usar recuentos de marcas de secuencia para determinar aneuploidía se describe en las solicitudes provisionales de EE.UU. 61/407.017 y 61/455.849778.

TABLA 4

Determinación de fracción feal por secuenciación paralela masiva de una librería enriquecida de acidos nucleicos polimórficos			
Muestra ID (cariotipo)	SNP	SNP Conteo Etiquetas	Fracción fetal (%)
11409 (47, XY+21)	rs13182883.1 Chr.5 longitud=111 alelo=A	261	4.41
	rs13182883.2 Chr.5 longitud=111 alelo=G	5918	
	rs740598.1 Chr.10 longitud=114 alelo=A	5545	7.30
	rs740598.2 Chr.10 longitud=114 alelo=G	405	
	rs8078417.1 Chr.17 longitud=110 alelo=C	8189	6.74
	rs8078417.2 Chr.17 longitud=110 alelo=T	121470	
	rs576261.1 Chr.19 longitud=114 alelo=A	58342	7.62
	rs576261.2 Chr.19 longitud=114 alelo=C	4443	
Fracción fetal (Media \pm S.D.) = 6.5 \pm 1.5			
Muestra ID			
95133 (47,XX+18)	rs1109037.1 Chr.2 longitud=126 alelo=A	12229	2.15
	rs1109037.2 Chr.2 longitud=126 alelo=A	263	
	rs13218440.1 Chr.6 longitud=139 alelo=A	55949	3.09
	rs13218440.2 Chr.6 longitud=139 alelo=G	1729	
	rs7041158.1 Chr.9 longitud=117 alelo=C	7281	4.12
	rs7041158.2 Chr.9 longitud=117 alelo=T	300	
	rs7205345.1 Chr.16 longitud=116 alelo=C	53999	2.14
	rs7205345.2 Chr.16 longitud=116 alelo=G	1154	

5

10

15

20

25

30

35

40

45

50

55

60

65

(continúa)

Fracción fetal (Media \pm S.D.) = 2.9 ± 0.9				
5	Muestra ID			
	51236 (46,XY+13)	rs13218440.1 Chr.6 longitud=139 alelo=A	1119	1.65
		rs13218440.2 Chr.6 longitud=139 alelo=G	67756	
10		rs560681.1 Chr.1 longitud=111 alelo=A	14123	5.18
		rs560681.2 Chr.1 longitud=111 alelo=G	732	
15		rs7205345.1 Chr.16 longitud=116 alelo=C	18176	1.63
		rs7205345.2 Chr.16 longitud=116 alelo=G	296	
		rs9866013.1 Chr.3 longitud=121 alelo=C	117	2.33
20		rs9866013.2 Chr.3 longitud=121 alelo=T	5024	
Fracción fetal (Media \pm S.D.) = 2.7 ± 1.7				
	Muestra ID			
25	54430 (45,XO)	rs1109037.1 Chr.2 longitud=126 alelo=A	19841	1.80
		rs1109037.2 Chr.21 longitud=126 alelo=G	357	
30		rs9866013.1 Chr.3 longitud=121 alelo=C	12931	3.81
		rs9866013.2 Chr.3 longitud=121 alelo=T	493	
		rs7041158.1 Chr.9 longitud=117 alelo=C	2800	4.25
35		rs7041158.2 Chr.9 longitud=117 alelo=T	119	
		rs740598.1 Chr.10 longitud=114 alelo=A	12903	4.87
		rs740598.2 Chr.1 longitud=114 alelo=G	628	
40		rs10773760.1 Chr.12 longitud=128 alelo=A	46324	4.65
		rs10773760.2 Chr.12 longitud=128 alelo=G	2154	
Fracción fetal (Media \pm S.D.) = 3.9 ± 1.2				

45 **Ejemplo 6**

50 **Determinación de la fracción fetal en secuenciación masivamente paralela: Enriquecimiento de ácidos nucleicos fetales y maternos para ácidos nucleicos polimórficos en una muestra de ADNc purificada.**

50 Para enriquecer el ADNc fetal y materno contenido en una muestra de ADNc purificada extraída de una muestra de plasma materno, se usó una porción del ADNc purificado para amplificar secuencias de ácidos nucleicos polimórficos diana que comprendía cada una un SNP elegido del panel de SNP dado en la Tabla 5.

55 El método se corresponde con el proceso de trabajo 2 representado en diagrama en la Figura 3. Se obtuvo plasma libre de células de una muestra materna de sangre, y se purificó el ADNc de la muestra de plasma como se describe en el Ejemplo 1. Se determinó que la concentración final era 92,8 pg/ μ l. Se amplificó el ADNc contenido en 5 μ l de ADNc purificado en un volumen de reacción de 50 μ l que contenía 7,5 μ l de una mezcla de cebadores 1 uM (Tabla 1), 10 μ l de mezcla maestra NEB 5X y 27 μ l de agua. Se realizaron ciclos térmicos con el gen Amp9700 (Applied Biosystems). Usando las siguientes condiciones de ciclos: incubar a 95 °C durante 1 minuto, seguido de 30 ciclos a 95 °C durante 20 segundos, 68 °C durante 1 minuto y 68 °C durante 30 s, que fue seguido de una incubación final a 68 °C durante 5 minutos. Se añadió un mantenimiento final a 4 °C hasta que las muestras se extrajeron para combinar con la porción sin amplificar de la muestra de ADNc purificada. El producto amplificado se purificó usando el sistema de purificación por PCR Agencourt AMPure XP (Parte N° A63881; Beckman Coulter Genomics, Danvers, MA), y la concentración se cuantificó usando Nanodrop 2000 (Thermo Scientific, Wilmington, DE). El producto de amplificación purificado se diluyó 1:10 en agua y se añadieron 0,9 μ l (371 pg) a 40 μ l de muestra de ADNc

purificada para obtener un enriquecimiento del 10 %. El ADNlc fetal enriquecido y materno presente en la muestra de ADNlc purificada se usó para preparar una biblioteca de secuenciación, y se secuenció como se describe en el Ejemplo 2.

5 La Tabla 5 proporciona los recuentos de marcas obtenidos para cada uno de los cromosomas 21, 18, 13, X e Y, es
decir, la densidad de marcas de secuencia, y los recuentos de marcas obtenidos para las secuencias polimórficas
informativas contenidas en el genoma de referencia de SNP, es decir, la densidad de marcas de SNP. Los datos
muestran que la información de secuenciación puede obtenerse a partir de la secuenciación una única biblioteca
10 construida de una muestra de ADNlc materna purificada que se ha enriquecido para secuencias que comprenden
SNP para determinar simultáneamente la presencia o ausencia de aneuploidía y la fracción fetal. La presencia o
ausencia de aneuploidía se determinó usando el número de marcas mapeadas en cromosomas como se describe en
las solicitudes provisionales de EE.UU. 61/407.017 y 61/455.849. En el ejemplo dado, los datos muestran que la
fracción de ADN fetal en la muestra de plasma AFR105 fue cuantificable a partir de los resultados de secuenciación
15 de cinco SNP informativos y se determinó que era del 3,84 %. Se proporcionan densidades de marcas de secuencia
para los cromosomas 21, 13, 18, X e Y.

El ejemplo muestra que el protocolo de enriquecimiento proporciona los recuentos de marcas requeridos para
determinar la aneuploidía y la fracción fetal de un único proceso de secuenciación.

20

25

30

35

40

45

50

55

60

65

TABLA 5

Determinación de fracción feal por secuenciación paralela masiva: Enriquecimiento de Acidos Nucleicos fetales y Maternales para Acidos Nucleicos Polimorficos prurificados en una muestra cfADN					
aneuploidía					
	Cromosoma 21	Cromosoma 18	Cromosoma 13	Cromosoma X	Cromosoma Y
Densidad Secuencia Etiquetas	178763	359529	388204	572330	2219
cariotipo	Inafectado	Inafectado	Inafectado	Inafectado	Inafectado
Fracción fetal					
SNP		SNP Densidad Etiquetas		Fracción fetal (%)	
rs10773760.1 Chr.12	lonaitud=128 alelo=A	18903		2.81	
rs10773760.2 Chr.12	lonaitud=128 alelo=G	532			
rs1109037.1 Chr.2	lonaitud=126 alelo=A	347		5.43	
rs1109037.2 Chr.2	lonaitud=126 alelo=G	6394			
rs2567608.1 Chr.20	lonaitud=110 alelo=A	94503		1.74	
rs2567608.2 Chr.20	lonaitud=110 alelo=G	1649			
rs7041158.1 Chr.9	lonaitud=117 alelo=C	107		5.61	
rs7041158.2 Chr.9	lonaitud=117 alelo=T	6			
rs8078417.1 Chr.17	lonaitud=110 alelo=C	162668		3.61	
rs8078417.2 Chr.17	lonaitud=1 alelo=T	5877			
Fracción fetal		(Media +- SD) = 3.8 +- 1.7			

5

10

15

20

25

30

35

40

45

50

55

60

65

Ejemplo 7

Determinación de la fracción fetal en secuenciación masivamente paralela: Enriquecimiento de ácidos nucleicos fetales y maternos para ácidos nucleicos polimórficos en una muestra de plasma.

Para enriquecer el ADNlc fetal y materno contenido en una muestra original de plasma derivada de una mujer embarazada, se usó una porción de la muestra original de plasma para amplificar secuencias de ácidos nucleicos polimórficas diana que comprendían cada una un SNP elegido del panel de SNP dado en la Tabla 1, y una porción del producto amplificado se combinó con la muestra original de plasma restante.

El método se corresponde con el proceso de trabajo 2 representado en diagrama en la Figura 3. Se amplificó el ADNlc contenido en 15 µl de plasma libre de células en un volumen de reacción de 50 µl que contenía 9 ul de una mezcla 1 µM de cebadores (Tabla 115 plex), 1 µl de ADN polimerasa de sangre Phusion, 25 ul de 2X tampón de PCR de sangre Phusion que contenía trifosfatos de desoxinucleótido (dNTPs: dATP, dCTP, dGTP y dTTP). Se realizaron ciclos térmicos con el gen Amp9700 (Applied Biosystems) usando las siguientes condiciones de ciclos: incubar a 95 °C durante 3 minutos, seguido de 35 ciclos a 95 °C durante 20 segundos, 55 °C durante 30 s y 70 °C durante 1 minuto, que fue seguido de una incubación final a 68 °C durante 5 minutos. Se añadió un mantenimiento final a 4 °C hasta que las muestras se extrajeron para combinar con la porción sin amplificar del plasma libre de células. El producto amplificado se diluyó 1:2 con agua y se analizó usando el bionanalizador. Se diluyeron 3 µl adicionales de producto amplificado con 11,85 µl de agua para obtener una concentración final de 2 ng/µl. Se combinaron 2,2 µl del producto amplificado diluido con la muestra de plasma restante. El ADNlc fetal y materno enriquecido presente en la muestra de plasma se purificó como se ha descrito en el Ejemplo 1, y se usó para preparar una biblioteca de secuenciación. La secuenciación y el análisis de los datos de secuenciación se realizó como se describe en el Ejemplo 2.

Los resultados se facilitan en la Tabla 6. En el ejemplo dado, los datos muestran que la fracción de ADN fetal en la muestra de plasma SAC2517 fue cuantificable a partir de los resultados de secuenciación de un SNP informativo y se determinó que era del 9,5%. En el ejemplo dado, la muestra SAC2517 se mostró por cariotipado que no estaba afectada por aneuploidías de los cromosomas 21, 13, 18, X e Y. Se proporcionan las densidades de marcas de secuencia para los cromosomas 21, 13, 18, X e Y. La presencia o ausencia de aneuploidía se determinó usando recuentos de marcas como se describe en las solicitudes provisionales de EE.UU. 61/407.017 y 61/455.849.

El ejemplo demuestra que puede usarse el enriquecer la mezcla de ADNlc fetal y materno presente en una muestra de plasma para secuencias de ácidos nucleicos que comprenden al menos un SNP informativo para proporcionar la secuencia requerida y recuentos de marcas de SNP para determinar aneuploidía y la fracción fetal de un único proceso de secuenciación en la secuenciación masivamente paralela de una biblioteca preparada a partir del ADNlc contenido en una muestra de plasma que está enriquecida en ácidos nucleicos polimórficos.

TABLA 6

Determinación de fracción fetal por secuenciación paralela masiva: Enriquecimiento de Acidos Nucleicos fetales y Maternales para Acidos Nucleicos Polimorficos comprimiendo un SNP en una muestra de plasma					
aneuploidía					
	Cromosoma 21	Cromosoma 18	Cromosoma 13	Cromosoma x	Cromosoma Y
SNP Densidad Etiquetas	183851	400582	470526	714055	2449
cariotipo	Inafectado	Inafectado	Inafectado	Inafectado	Inafectado
Fracción fetal					
SNP	Recuento de etiqueta		Fracción fetal (%)		
rs10773760.1 Chr.12 longitud=128 alelo=A	8536		9.5		
rs10773760.2 Chr.12 longitud=128 alelo=G	89924				

Ejemplo 8

Determinación de la fracción fetal en secuenciación masivamente paralela de muestras que comprenden secuencias polimórficas amplificadas: SNP en tándem

5 Para determinar la fracción de ADN fetal en una muestra materna, secuencias de ácidos nucleicos polimórficas diana que comprenden cada una un par de SNP en tándem se amplifican y se usan para preparar una biblioteca diana para secuenciación en un modo masivamente paralelo. Los pares de SNP en tándem pueden seleccionarse de rs7277033-rs2110153; rs2822654-rs1882882; rs368657-rs376635; rs2822731-rs2822732; rs1475881-rs7275487; 10 rs1735976-rs2827016; rs447340-rs2824097; rs418989- rs13047336; rs987980- rs987981; rs4143392-rs4143391; rs1691324-rs13050434; rs11909758-rs9980111; rs2826842-rs232414; rs1980969-rs1980970; rs9978999-rs9979175; rs1034346-rs12481852; rs7509629-rs2828358; rs4817013-rs7277036; rs9981121-rs2829696; rs455921-rs2898102; rs2898102-rs458848; rs961301-rs2830208; rs2174536-rs458076; rs11088023-rs11088024; rs1011734-rs1011733; rs2831244-rs9789838; rs8132769-rs2831440; rs8134080-rs2831524; rs4817219-rs4817220; rs2250911-rs2250997; 15 rs2831899-rs2831900; rs2831902-rs2831903; rs11088086-rs2251447; rs2832040-rs11088088; rs2832141-rs2246777; rs2832959-rs9980934; rs2833734-rs2833735; rs933121-rs933122; rs2834140-rs12626953; rs2834485-rs3453; rs9974986-rs2834703; rs2776266-rs2835001; rs1984014-rs1984015; rs7281674-rs2835316; rs13047304-rs13047322; rs2835545-rs4816551; rs2835735-rs2835736; rs13047608-rs2835826; rs2836550-rs2212596; rs2836660-rs2836661; rs465612-rs8131220; rs9980072-rs8130031; rs418359-rs2836926; rs7278447-rs7278858; 20 rs385787-rs367001; rs367001-rs386095; rs2837296-rs2837297; y rs2837381-rs4816672. Los cebadores usados para amplificar las secuencias diana que comprenden los SNP en tándem se diseñan para englobar ambos sitios de SNP. Por ejemplo, el cebador directo se diseña para englobar el primer SNP, y el cebador inverso se diseña para englobar el segundo del par de SNP en tándem, es decir, cada uno de los sitios de SNP en el par del tándem está englobado dentro de los 36 pb generados por el método de secuenciación. Puede usarse secuenciación de extremos apareados para identificar todas las secuencias que engloban los sitios del SNP en tándem. Conjuntos de cebadores a modo de ejemplo que se usan para amplificar los SNP en tándem desvelados en el presente documento son rs7277033-rs2110153_F:

30 TCCTGGAACAAAAGTATT (SEC ID N°: 197) y rs7277033-rs2110153_R:
AACCTTACAACAAAGCTAGAA (SEC ID N°: 198), set rs2822654-rs1882882_F:
ACTAAGCCTTGGGGATCCAG (SEC ID N°: 199) y rs2822654-rs1882882_R:
TGCTGTGGAATACTAAAAGG (SEC ID N°: 200), set rs368657-rs376635_F:CTCCAGAGGTAATCCTGTGA
(SEC ID N°: 201) y rs368657-rs376635_R:TGGTGTGAGATGGTATCTAGG (SEC ID N°: 202), rs2822731-
s2822732-F:GTATAATCCATGAATCG (SEC ID N°: 203) y rs2822731-
35 rs2822732_R:TTCAAATTGTATATAAGAGAGAT (SEC ID N°: 204), rs1475881-
rs7275487_F:GCAGGAAAGTTATTTTAAAT (SEC ID N°: 205) y rs1475881-
rs7275487_R:TGCTTGAGAAAAGCTAACACTT (SEC ID N°: 206), rs1735976-
rs2827016F:CAGTGTGGAAATTGTCTG (SEC ID N°: 207) y rs1735976-rs2827016
R:GGCACTGGGAGATTATTGTA (SEC ID N°: 208), rs447349-rs2824097_F:TCCTGTTGTTAAGTACACAT
40 (SEC ID N°: 209) y rs447349-rs2824097_R:GGGCCGTAATTACTTTTG (SEC ID N°: 210), rs418989-
rs13047336-F:ACTCAGTAGGCACTTTGTGTC (SEC ID N°: 211) y rs418989-
rs13047336_R:TCTTCCACCACCAATC (SEC ID N°: 212), rs987980-
rs987981_F:TGGCTTTTCAAAGTAAAA (SEC ID N°: 213) y rs987980- rs987981_R:
GCAACGTTAACATCTGAATTT (SEC ID N°: 214), rs4143392-rs4143391_F: rs4143392-rs4143391 (SEC ID N°:
45 215) y rs4143392-rs4143391_R:ATTTTATATGTCATGATCTAAG (SEC ID N°: 216), rs1691324-
ras13050434_F: AGAGATTACAGGTGTGAGC (SEC ID N°: 217) y rs1691324-rs13050434_R:
ATGATCCTCAACTGCCTCT (SEC ID N°: 218), rs1 1909758-rs9980111_F: TGAAACTCAAAGAGAAAAG
(SEC ID N°: 219) y rs11909758-rs9980111_R: ACAGATTTCTACTTAAAATT (SEC ID N°: 220),
rs2826842-rs232414_F: TGAAACTCAAAGAGAAAAG (SEC ID N°: 221) y rs2826842-rs232414_R:
50 ACAGATTTCTACTTAAAATT (SEC ID N°: 22), rs2826842-rs232414_F: GCAAAGGGTACTCTATGTA (SEC
ID N°: 223) y rs2826842-rs232414_R: TATCGGGTCATCTTGTTAAA (SEC ID N°: 224), rs1980969-
rs1980970_F: TCTAACAAAGCTCTGTCCAAA (SEC ID N°: 225) y rs1980969-rs1980970_R:
CCCACTGAATAACTGGAACA (SEC ID N°: 226), rs9978999-rs9979175_F:
GCAAGCAAGCTCTTACCTTC (SEC ID N°: 227) y rs9978999-rs9979175_R:
55 TGTTCTTCCAAAATTCACATGC (SEC ID N°: 228), rs1034346-rs12481852_F:
ATTTCACTATTCCTTCATTTT (SEC ID N°: 229) y rs1034346-rs12481852_R:
TAATTGTTGCACATAAATTAC (SEC ID N°: 230), rs4817013-rs7277036_F:
AAAAAGCCACAGAAATCAGTC (SEC ID N°: 231) y rs4817013-rs7277036_R:
TTCTTATATCTCACTGGGCATT (SEC ID N°: 232), rs9981121-rs2829696_F:
60 GGATGGTAGAAGAGAAGAAAGG (SEC ID N°: 233) y rs9981121-rs2829696_R:
GGATGGTAGAAGAGAAGAAAGG (SEC ID N°: 234), rs455921-rs2898102_F:
TGCAAAGATGCAGAACCAAC (SEC ID N°: 235) y rs455921-rs2898102_R:
TTTTTCCCTTGCTCCTGGCTGA (SEC ID N°: 236), rs2898102-rs458848]:
TGCAAAGATGCAGAACCAAC (SEC ID N°: 237) y rs2898102-rs458848_R:
65 GCCTCCAGCTCTATCCAAGTT (SEC ID N°: 238), rs961301-rs2830208_F:
CCTTAATATCTTCCCATGTCCA (SEC ID N°: 239) y rs961301-rs2830208_R:

ATTGTTAGTGCCTCTTCTGCTT (SEC ID N°: 240), rs2174536-rs458076_F:
 GAGAAGTGAGGTCAGCAGCT (SEC ID N°: 241) y rs2174536-rs458076_R:
 TTTCTAAATTTCCATTGAACAG (SEC ID N°: 242), rs11088023-rs11088024_F:
 5 GAAATTGGCAATCTGATTCT (SEC ID N°: 243) y rs11088023-rs11088024_R:
 CAACTTGTCCCTTTATTGATGT (SEC ID N°: 244), rs1011734-rs1011733_F:
 CTATGTTGATAAAACATTGAAA (SEC ID N°: 245) y rs1011734-rs1011733_R:
 GCGTGTCTGGAATATAGTTT (SEC ID N°: 246), rs2831244-rs9789838_F:
 CAGGGCATATAATCTAAGCTGT (SEC ID N°: 247) y rs2831244-rs9789838_R:
 CAATGACTCTGAGTTGAGCAC (SEC ID N°: 248), rs8132769-rs2831440_F: ACTCTCTCCCTCCCCTCT (SEC
 10 ID N°: 249) y rs8132769-rs2831440_R: TATGGCCCCAAAACATTCT (SEC ID N°: 250), rs8134080-
 rs2831524_F: ACAAGTACTGGGCAGATTGA (SEC ID N°: 251) y rs8134080-rs2831524_R:
 GCCAGGTTTAGCTTTCAAGT (SEC ID N°: 252), rs4817219-rs4817220_F:
 TTTTATATCAGGAGAAACACTG (SEC ID N°: 253) y rs4817219-rs4817220_R:
 CCAGAATTTGGAGGTTTAAT (SEC ID N°: 254), rs2250911-rs2250997_F:
 15 TGTCATTCTCTCTTTATCTCCA (SEC ID N°: 255) y rs2250911-rs2250997_R:
 TTCTTTTGCCTCTCCCAAAG (SEC ID N°: 256), rs2831899-rs2831900_F: ACCCTGGCACAGTGTGACT
 (SEC ID N°: 257) y rs2831899-rs2831900_R: TGGCCTGAGTTGAGAAGAT (SEC ID N°: 258),
 rs2831902-rs2831903_F: AATTTGTAAGTATGTGCAACG (SEC ID N°: 259) y rs2831902-rs2831903_R:
 TTTTCCCATTTCCAACCTCT (SEC ID N°: 260), rs11088086-rs2251447-F:
 AAAAGATGAGACAGCAGGT (SEC ID N°: 261) y rs11088086-rs2251447_R:
 20 ACCCTGTGAATCTCAAAAT (SEC ID N°: 262), rs2832040-rs11088088_F:
 GCACTTGCTTCTATTGTTGT (SEC ID N°: 263) y rs2832040-rs11088088_R:
 CCCTTCTCTCTTCCATTCT (SEC ID N°: 264), rs2832141-rs2246777_F: AGCACTGCAGGTA (SEC ID N°:
 25 265) y rs2832141-rs2246777_R: ACAGATACCAAGAAGTCAA (SEC ID N°: 266), rs2832959-rs9980934_F:
 TGGACACCTTTCAACTTAGA (SEC ID N°: 267) y rs2832959-rs9980934_R:
 GAACAGTAATGTTGAACTTTT (SEC ID N°: 268), rs2833734-rs2833735_F:
 TCTTGCAAAAAGCTTAGCACA (SEC ID N°: 269) y rs2833734-rs2833735_R:
 AAAAAGATCTCAAAGGGTCCA (SEC ID N°: 270), rs933121-rs933122_F: GCTTTTGTGAACATCAAGT (SEC
 ID N°: 271) y rs933121-rs933122_R: CCTCCAGCAGCATAGTCT (SEC ID N°: 272), rs2834140-
 30 rs12626953_F: AAATCCAGGATGTGCAGT (SEC ID N°: 273) y rs2834140-rs12626953_R:
 ATGATGAGGTCAGTGGTGT (SEC ID N°: 274), rs2834485-rs3453_F: CATCACAGATCATAGTAAATGG (SEC
 ID N°: 275) y rs2834485-rs3453_R: AATTATTATTTGCAGGCAAT (SEC ID N°: 276), rs9974986-rs2834703_F:
 CATGAGGCAAACACCTTTCC (SEC ID N°: 277) y rs9974986-rs2834703_R:
 GCTGGACTCAGGATAAAGAACA (SEC ID N°: 278), rs2776266-rs2835001_F:
 35 TGGAAAGCTGAGCTGACTAA (SEC ID N°: 279) y rs2776266-rs2835001_R:CCTTCTTTTCCCCCAGAATC
 (SEC ID N°: 280), rs1984014-rs1984015-F:TAGGAGAACAGAAGATCAGAG (SEC ID N°: 281) y rs1984014-
 rs1984015_R:AAAGACTATTGCTAAATGCTTG (SEC ID N°: 282), rs7281674-rs2835316_F:
 TAAGCGTAGGGCTGTGTGTG (SEC ID N°: 283) y rs7281674-rs2835316_R:
 GGACGGATAGACTCCAGAAGG (SEC ID N°: 284), rs13047304-rs13047322_F:
 40 GAATGACCTTGGCACTTTTATCA (SEC ID N°: 285) y rs13047304-rs13047322_R:
 AAGGATAGAGATATACAGATGAATGGA (SEC ID N°: 286), rs2835735-rs2835736_F:
 CATGCACCGCGCAAATAC (SEC ID N°: 287) y rs2835735-rs2835736_R: ATGCCTACCCACAAACAC (SEC
 ID N°: 288), rs13047608-rs2835826_F: TCCAAGCCCTTCTACTCAC (SEC ID N°: 289) y
 rs13047608-rs2835826_R: CTGGGACGGTGACATTTTCT (SEC ID N°: 290), rs2836550-rs2212596_F:
 45 CCCAGGAAGAGTGGAAGATT (SEC ID N°: 291) y rs2836550-rs2212596_R:
 TTAGCTTGCATGTACCTGTGT (SEC ID N°: 292), rs2836660-rs2836661_F:
 AGCTAGATGGGGTGAATTTT (SEC ID N°: 293) y _R: TGGGCTGAGGGGAGATTC (SEC ID N°: 294),
 rs465612-rs8131220_F: ATCAAGCTAATTAATGTTATCT (SEC ID N°: 295) y rs465612-rs8131220_R:
 AATGAATAAGGTCCTCAGAG (SEC ID N°: 296), rs9980072-rs8130031_F:TTTAATCTGATCATTGCCCTA
 50 (SEC ID N°: 297) y rs9980072-rs8130031_R:
 AGCTGTGGGTGACCTTGA (SEC ID N°: 298), rs418359-rs2836926_F: TGTCCCACCATTGTGTATTA (SEC ID
 N°: 299) y rs418359-rs2836926_R: TCAGACTTGAAGTCCAGGAT (SEC ID N°: 300), rs7278447-rs7278858_F:
 GCTTCAGGGGTGTTAGTTTT (SEC ID N°: 301) y rs7278447-rs7278858_R:
 CTTTGTGAAAAGTCGTCCAG (SEC ID N°: 302), rs385787-rs367001_F:CCATCATGGAAAGCATGG (SEC ID
 55 N°: 303) y rs385787-rs367001_R: TCATCTCCATGACTGCACTA (SEC ID N°: 304), rs367001-rs386095_F:
 GAGATGACGGAGTAGCTCAT (SEC ID N°: 305) y rs367001-rs386095_R:
 CCCAGTGCATGTCTAC (SEC ID N°: 306), rs2837296-rs2837297_F: TCTTGTTC AATCACAGGAC (SEC
 ID N°: 307) y rs2837296-rs2837297_R: ATGCTGTTAGCTGAAGCTCT (SEC ID N°: 308), y
 rs2837381-rs4816672_F: TGAAAGCTCCTAAAGCAGAG (SEC ID N°: 309) andrs2837381-
 60 rs4816672_R:TTGAAGAGATGTGCTATCAT (SEC ID N°: 310). Las secuencias de polinucleótidos, por ejemplo,
 las secuencias de GC clamp, pueden incluirse para garantizar la hibridación específica de cebadores ricos en
 AT (Ghanta et al., PLOS ONE 5(10):
 doi10.1371/journal.pone.0013184 [2010], disponible en la malla mundial en plosone.org). Un ejemplo de una
 secuencia de GC clamp que puede incluirse tanto 5' del cebador directo como 3' del cebador inverso es
 65 GCCGCCTGCAGCCCCGCGCCCCCGTGCCTCCGCCCCGCGCCGGCCGGGGCC (SEC ID N°: 311).

Las secuencias polimórficas pueden usarse solas o en combinación con ADNlc sin amplificar para determinar tanto la fracción fetal o la presencia o ausencia de aneuploidía como la fracción fetal en una muestra materna como se describe para secuencias de SNP polimórficas. La preparación de muestras y el enriquecimiento de bibliotecas de secuenciación de ADNlc, una muestra de ADNlc purificada y una muestra de plasma se realiza según el método descrito en los Ejemplos 5, 6 y 7, respectivamente. Todas las bibliotecas de secuenciación se preparan como se describe en el Ejemplo 2a, y se realiza secuenciación como se describe en el Ejemplo 2b. El análisis de los datos de secuenciación para la determinación de aneuploidía fetal se realiza como se describe en el Ejemplo 5. Concomitante al análisis para determinar aneuploidía, los datos de secuenciación se analizan para determinar la fracción fetal del siguiente modo. Tras la transferencia de la imagen y los archivos de llamada de bases al servidor de Unix que ejecuta el software "Genoma Analyser Pipeline" de Illumina versión 1.51 como se describe en el Ejemplo 3a, las lecturas de 36 pb se alinean con un 'genoma de SNP en tándem' usando el programa BOWTIE. El genoma de SNP en tándem se identifica como la agrupación de las secuencias de ADN que engloban los alelos de los 58 pares de SNP en tándem desvelados anteriormente. Solo lecturas que se mapearon únicamente con el genoma de SNP en tándem se usan para el análisis de la fracción fetal. Las lecturas que coinciden perfectamente con el genoma de SNP en tándem se cuentan como marcas y se filtran. De las lecturas restantes, solo las lecturas que tienen uno o dos desapareamientos se cuentan como marcas y se incluyen en el análisis. Se cuentan las marcas mapeadas en cada uno de los alelos de SNP en tándem, y se determina la fracción fetal esencialmente como se describe en el Ejemplo 6 anterior, pero representando marcas mapeadas en los dos alelos de SNP en tándem x e y presentes en cada una de las secuencias de ácidos nucleicos polimórficas diana amplificadas que se amplifican para enriquecer las muestras, es decir,

$$\% \text{ fracción fetal alelo}_{x+y} = \left(\frac{\sum \text{secuencias de etiquetas fetales para alelo}_{x+y}}{\sum \text{secuencias de etiquetas maternas para alelo}_{x+y}} \right) \times 100$$

Solo se usan SNP en tándem informativos para determinar la fracción fetal.

Opcionalmente, la fracción de ácidos nucleicos fetales en la mezcla de ácidos nucleicos fetales y maternos se calcula para cada uno del alelo informativo (alelo_{x+y}) del siguiente modo:

$$\% \text{ fracción fetal alelo}_{x+y} = \left(\frac{2 \times \sum \text{secuencias de etiquetas fetales para alelo}_{x+y}}{\sum \text{secuencias de etiquetas maternas para alelo}_{x+y}} \right) \times 100$$

para compensan la presencia de 2 conjuntos de alelos fetales en tándem, estando uno enmascarado por el fondo materno.

El porcentaje la fracción fetal se calcula para al menos 1, al menos 2, al menos 3, al menos 4, al menos 5, al menos 6, al menos 7, al menos 8, al menos 9, al menos 10, al menos 11, al menos 12, al menos 13, al menos 14, al menos 15, al menos 16, al menos 17, al menos 18, al menos 19, al menos 20, o más conjuntos informativos de alelos en tándem. En una realización, la fracción fetal es la fracción fetal promedio determinada para al menos 3 conjuntos informativos de alelos en tándem.

Ejemplo 9

Determinación de la fracción fetal en secuenciación masivamente paralela de muestras que comprenden secuencias polimórficas amplificadas: STR

Para determinar la fracción de ADNlc fetal en una muestra materna, secuencias de ácidos nucleicos polimórficas diana que comprenden cada una una STR se amplifican y se usan para preparar una biblioteca diana para la secuenciación en un modo masivamente paralelo.

Se obtienen muestras de sangre periférica de sujetos embarazos, y el ADNlc se purifica de la fracción de plasma como se describe en Ejemplos 1 y 2. Las STR que se amplifican se eligen de las STR codificantes y no codificantes desveladas en la Tabla 7, y se obtiene la amplificación de las secuencias de STR polimórficas usando los conjuntos correspondientes de cebadores proporcionados. Por ejemplo, las STR enumeradas en la Tabla 7 se amplifican usando los cebadores correspondientes (SEC ID N°: 113-197), y el producto amplificado se usa para generar una biblioteca de secuenciación diana. La biblioteca de secuenciación diana de STR se prepara como se describe para la preparación de la biblioteca diana de SNP como se describe en el Ejemplo 8. Las STR CSF1PO, D13S317, D16S539, D18S51, D21S11, D2S1338D7S820 y FGA se han analizado previamente para determinar la fracción fetal, y se desvelan en las solicitudes provisionales de EE.UU. 61/296.358 y 61/360.837.

TABLA 7

MiniSTRs CODIS y no CODIS				
STR Locus (Nombre marcador)	Localización del cromosoma	Rango de talla (bp)	Acceso a Genbank	Secuencias de los cebadores(anteriores/inversas)
Codis miniSTR loci*				
CSF1PO	5q33.1	89-129	X14720	ACAGTAACTGCCTTCATAGATAG (CSF1PO_F; SEQ ID NO: 113) GTGTCAGACCCTGTTCTAAGTA (CSF1PO_R; SEQ ID NO:114)
FGA	4q31.3	125-281	M64982	AAATAAAATTAGGCATATTTACAAGC(FGA_F; SEQ ID NO: 115) GCTGAGTGATTTGTCTGTAATTG(FGA_R; SEQ ID NO: 116)
TH01	11p15.5	51-98	D00269	CCTGTTCCCTCCCTTATTTCCC(TH01_F; SEQ ID NO: 117) GGGAACACAGACTCCATGGTG(TH01_R; SEQ ID NO: 118)
TPOX	2p25.3	65-101	M68651	CTTAGGGAACCCTCACTGAATG(TPOX_F ; SEQ ID NO:119) GTCCTTGTCAGCGTTTATTTGC(TPOX_R; SEQ ID NO: 120)
vWA	12p13.31	88-148	M25858	AATAATCAGTATGTGACTTGATTGA(v WA_F; SEQ ID NO:121) ATAGGATGGATGGATAGATGGA(vWA_R ; SEQ ID NO: 122)
D3S1358	3p21.31	72-120	NT_005997	CAGAGCAAGACCCTGTCTCAT(D3S1358_F; SEQ ID NO: 123) TCAACAGAGGCTTGCATGTAT(D3 S 1358_R; SEQ ID NO: 124)
D5S818	5q23.2	81-117	AC008512	GGGTGATTTTCTCTTTGGT(D5S818_F; SEQ ID NO: 125) AACATTTGTATCTTTATCTGTATCCTTAT TTAT(D5S818_R; SEQ ID NO: 126)
D7S820	7q21.11	136-176	AC004848	GAACACTTGCATAGTTTGAACGAAC(D7S 820_F; SEQ ID NO: 127) TCATTGACAGAATTGCACCA(D7S820_R; SEQ ID NO: 128)
D8S1179	8q24.13	86-134	AF216671	TTTGATTTTCATGTGTACATTCGTATC(D 7S820_F; SEQ ID NO: 129) ACCTATCCTGTAGATTATTTTCACTGTG(D7S 820_R; SEQ ID NO: 130)
D13S317	13q31.1	88-132	AL353628	TCTGACCCATCTAACGCCTA(D13S317_F; SEQ ID NO:131) CAGACAGAAAGATAGATAGATGATTGA(D13 S317_R; SEQ ID NO:132)
D16S539	16q24.1	81-121	AC024591	ATACAGACAGACAGACAGGTG(D 16S539 _F; SEQ ID NO:133)

(continúa)

MiniSTRs CODIS v no CODIS				
STR Locus (Nombre marcador)	Localización del cromosoma	Rango de talla (bp)	Acceso a Genbank	Secuencias de los cebadores(anteriores/inversas)
Codis miniSTR loci*				
				GCATGTATCTATCATCCATCTCT(D16S53 9_R; SEQ ID NO:134)
D18S51	18q21.33	113-193	AP001534	TGAGTGACAAATTGAGACCTT(D18S51_F ; SEQ ID NO:135) GTCTTACAATAACAGTTGCTACTATT(D18S51_R; SEQ ID NO:136)
D21S11	21q21.1	153-221	AP000433	ATTCCCCAAGTGAATTGC(D21S11_F; SEQ ID NO: 137) GGTAGATAGACTGGATAGATAGACGA(D21S11_R; SEQ ID NO:138)
D2S1338	2q35	90-142	AC01036	TGGAAACAGAAATGGCTTGG(D2S1338_F ; SEQ ID NO:139) GATTGCAGGAGGGAAGGAAG(D2S1338_R; SEQ ID NO: 140)
Penta D	21q22.3	94-167	AP001752	GAGCAAGACACCATCTCAAGAA(Penta D_F; SEQ ID NO:141) GAAATTTTACATTTATGTTTATGATTCTC T(Penta D_R; SEQ ID NO:142)
Penta E	15q26.2	80-175	AC027004	GGCGACTGAGCAAGACTC(Penta E _F; SEQ ID NO: 143) GGTTATTAATTGAGAAAACCTCCTTACA(Penta E _R; SEQ ID NO: 144)
MiniSTRs CODIS y no CODIS loci *				
D22S1045	22q12.3	82-115	AL022314 (17)	ATTTTCCCGATGATAGTAGTCT(D22S1045_F; SEQ ID NO:145) GCGAATGTATGATTGGCAATATTTTT(D22S1045_R; SEQ ID NO:146)
D20S1082	20q13.2	73 - 101	AL158015	ACATGTATCCCAGAACTTAAAGTAAAC(D20S1082_F; SEQ ID NO:147) GCAGAAGGGAAAATTGAAGCTG(D20S1082_R; SEQ ID NO:148)
D20S482	20p13	85 - 126	AL121781 (14)	CAGAGACACCGAACCAATAAGA(D20S482_F; SEQ ID NO:149) GCCACATGAATCAATTCCTATAATAAA(D20S482_R; SEQ ID NO:150)
D18S853	18p11.31	82 - 104	AP005130 (11)	GCACATGTACCCTAAAACCTAAAAT(D18S853_F; SEQ ID NO:151) GTCAACCAAAACCTCAACAAGTAGTAA(D18S853_R; SEQ ID NO:152)
D17S1301	17q25.1	114 - 139	AC016888 (12)	AAGATGAAATTGCCATGTAAAAATA(D17S1301_F; SEQ ID NO:153)

(continúa)

MiniSTRs CODIS y no CODIS loci *					
5				GTGTGTATAACAAAATTCTATGATGG(D17S1301_R; SEQ ID NO:154)	
10	D17S974	17p13.1	114 - 139	AC034303 (10)	GCACCCAAAACGAATGTCATA(D17S974_F; SEQ ID NO:155) GGTGAGGTGAGACCCTGTC(D17S974_R; SEQ ID NO:156)
15	D14S1434	14q32.13	70 - 98	AL121612 (13)	TGTAATAACTCTACGACTGTCTGTCTG(D14S1434_F; SEQ ID NO:157) GAATAGGAGGTGGATGGATGG(D14S1434_R; SEQ ID NO:158)
20	D12ATA63	12q23.3	76 - 106	AC009771 (13)	GAGCGAGACCCTGTCTCAAG(D12ATA63_F; SEQ ID NO:159) GGAAAAGACATAGGATAGCAATTT(D12ATA63_R; SEQ ID NO:160)
25	D11S4463	11q25	88 - 116	AP002806 (14)	TCTGGATTGATCTGTCTGTCC(D11S4463_F; SEQ ID NO:161) GAATTAATACCATCTGAGCACTGAA(D11S4463_R; SEQ ID NO:162)
30	D10S1435	10p15.3	82 - 139	AL354747 (11)	TGTTATAATGCATTGAGTTTATTCTG(D10S1435_F; SEQ ID NO:163) GCCTGTCTCAAAAATAAGAGATAGACA(D10S1435_R; SEQ ID NO:164)
35	D10S1248	10q26.3	79 - 123	AL391869 (13)	TTAATGAATTGAACAAATGAGTGAG(D10S1248_F; SEQ ID NO:165) GCAACTCTGGTTGTATTGTCTTCAT(D10S1248_R; SEQ ID NO:166)
40	D9S2157	9q34.2	71 - 107	AL162417 (10)	CAAAGCGAGACTCTGTCTCAA(D9S2157_F; SEQ ID NO:167) GAAAATGCTATCCTCTTTGGTATAAAT(D9S2157_R; SEQ ID NO:168)
45	D9S1122	9q21.2	93 - 125	AL161789 (12)	GGTATTTCAAGATAACTGTAGATAGG(D9S1122_F; SEQ ID NO:168) GCTTCTGAAAGCTTCTAGTTTACC(D9S1122_R; SEQ ID NO:170)
50	D8S1115	8p11.21	63 - 96	AC090739 (9)	TCCACATCCTCACCAACAC(D8S1115_F; SEQ ID NO:171) GCCTAGGAAGGCTACTGTCAA(D8S1115_R; SEQ ID NO:172)
55	D6S1017	6p21.1	81 - 110	AL035588 (10)	CCACCCGTCCATTTAGGC(D6S1017_F; SEQ ID NO:173) GTGAAAAGTAGATATAATGGTTGGTG(D6S1017_R; SEQ ID NO:174)
60	D6S474	6q21	107 - 136	AL357514 (17)	GGTTTTCCAAGAGATAGACCAATTA(D6S474_F; SEQ ID NO:175) GTCTCTCATAAATCCCTACTCATATC(D6S474_R; SEQ ID NO:176)
65					

(continúa)

MiniSTRs CODIS y no CODIS loci *					
5	D5S2500	5q11.2	85 - 126	AC008791 (17)	CTGTTGGTACATAATAGGTAGGTAGGT(D5S 2500_F; SEQ ID NO:177) GTCGTGGGCCCCATAAATC(D5S2500_R; SEQ ID NO:178)
10	D4S2408	4p15.1	85 - 109	AC 110763 (9)	AAGGTACATAACAGTTCAATAGAAAGC(D4S 2408_F; SEQ ID NO:179) GTGAAATGACTGAAAAATAGTAACCA(D 4S2408_R; SEQ ID NO:180)
15	D4S2364	4q22.3	67 - 83	AC022317 (9)	CTAGGAGATCATGTGGGTATGATT(D4S2 364U_F; SEQ ID NO:181) GCAGTGAATAAATGAACGAATGGA(D4S 2364_R; SEQ ID NO:182)
20	D3S4529	3p12.1	111 - 139	AC117452 (13)	CCCCAAAATTACTTGAGCCAAT(D3S452_F ; SEQ ID NO:183) GAGACAAAATGAAGAAACAGACAG(D3 S 452_R; SEQ ID NO:184)
25	D3S3053	3q26.31	84 - 108	AC069259 (9)	TCTTTGCTCTCATGAATAGATCAGT(D3S 3053_F; SEQ ID NO:185) GTTTGTGATAATGAACCCACTCAG(D3 S3 053_R; SEQ ID NO:186)
30	D2S1776	2q24.3	127 - 161	AC009475 (11)	TGAACACAGATGTTAAGTGTGTATATG(D2S1 776_F; SEQ ID NO:187) GTCTGAGGTGGACAGTTATGAAA(D2S 17 76_R; SEQ ID NO:188)
35	D2S441	2p14	78 - 110	AC079112 (12)	CTGTGGCTCATCTATGAAAACCTT(D2S441 F; SEQ ID NO:189) GAAGTGGCTGTGGTGTATGAT(D2S441 R; SEQ ID NO:190)
40	D1S1677	1q23.3	81 - 117	AL513307 (15)	TTCTGTTGGTATAGAGCAGTGTTT(D 1 S 16 77_F; SEQ ID NO:191) GTGACAGGAAGGACGGAATG(D 1S1677_ R; SEQ ID NO: 192)
45	D1S1627	1p21.1	81 - 100	AC093119 (13)	CATGAGGTTTGCAAATACTATCTTAAC(D1S1 627_F; SEQ ID NO:193) GTTTTAAT'ITTCTCCAAATCTCCA(D 1 S 16 27_R; SEQ ID NO:194)
50	D1GATA113	1p36.23	81 - 105	Z97987 (11)	TCTTAGCCTAGATAGATACTTGCTTCC(D 1GATA113_F; SEQ ID NO:195) GTCAACCTTTGAGGCTATAGGAA(D1GA TA113_R; SEQ ID NO:196)
55	. * (Butler et al, J Forensic Sci. 5: 1054-1064; Hill et al, cartel # 44- 17o Simposio Internacional sobre Identificación Humana -2006)				1

60

65

Se realiza la secuenciación de la biblioteca enriquecida en secuencias de STR polimórficas usando una tecnología de NGS, por ejemplo, secuenciación por síntesis. La lectura de secuencias de longitudes que engloban las STR, por ejemplo, miniSTR de al menos 100 pb, para un genoma de STR de referencia que consiste en las secuencias polimórficas que se amplificaron en la muestra. Se identifican alelos de STR de informativos por diferencias en la longitud de las repeticiones, y se cuentan el número de marcas de secuencia de STR, y se usan para determinar la fracción fetal. Opcionalmente, la amplificación de las secuencias de STR polimórficas se realiza para enriquecer una muestra de plasma, una muestra de ADNlc purificada o una muestra de biblioteca de secuenciación de ADNlc, como se describe en los Ejemplos 5, 6 y 7, respectivamente.

10 Ejemplo 10

Determinación de la fracción fetal por electroforesis capilar de secuencias polimórficas que comprenden STR

15 Para determinar la fracción fetal en muestras maternas que comprenden ADNlc fetal y materno, se recogieron muestras de sangre periférica de mujeres embarazadas voluntarias que portaban fetos tanto masculinos como femeninos. Se obtuvieron las muestras de sangre periférica y se procesaron para proporcionar ADNlc purificado como se describe en el Ejemplo 1

20 Se analizaron diez microlitros de muestras de ADNlc usando el kit de amplificación por PCR AmpFISTR® MiniFiler™ (Applied Biosystems, Foster City, CA) según las instrucciones del fabricante. Brevemente, se amplificó el ADNlc contenido en 10 µl en un volumen de reacción de 25 µl que contenía 5 µl de cebadores fluorescentemente marcados (conjunto de cebadores AmpF/STR® MiniFiler™), y la mezcla maestra AmpF/STR® MiniFiler™, que incluye ADN polimerasa AmpliTaq Gold® y tampón asociado, sal (MgCl₂ 1,5 mM) y trifosfatos de desoxinucleótido 200 µM (dNTPs: dATP, dCTP, dGTP y dTTP). Los cebadores fluorescentemente marcados son cebadores directos que se marcan con colorantes 6FAM™, VIC™, NED™ y PET™. Se realizaron ciclos térmicos con el gen Amp9700 (Applied Biosystems) usando las siguientes condiciones de ciclos: incubar a 95 °C durante 10 minutos, seguido de 30 ciclos a 94 °C durante 20 segundos, 59 °C durante 2 minutos y 72 °C durante 1 minuto, que fue seguido de una incubación final a 60 °C durante 45 minutos. Se añadió un mantenimiento final a 4 °C hasta que las muestras se extrajeron para el análisis. El producto amplificado se preparó diluyendo 1 µl de producto amplificado en 8,7 µl de formamida Hi-Di™ (Applied Biosystems) y 0,3 µl de patrón de tamaño interno GeneScan™-500 LIZ (Applied Biosystems), y se analizó con un analizador genético ABI PRISM3130x1 (Applied Biosystems) usando recogida de datos HID_G5_POP4 (Applied Biosystems) y una matriz capilar de 36 cm. Todo el genotipado se realizó con el software GeneMapper_ID v3.2 (Applied Biosystems) usando escaleras alélicas y cajas y paneles proporcionados por el fabricante.

Toda la medición del genotipado se realizó en el analizador genético Applied Biosystems 3130x1, usando una "ventana" de ± 0,5 nt alrededor del tamaño obtenido para cada alelo para permitir la detección y la correcta asignación de alelos. Cualquier alelo de muestra cuyo tamaño estuviera fuera de la ventana de ± 0,5 nt se determinó que era OL, es decir, "fuera de la escalera". Los alelos OL son alelos de un tamaño que no está representado en la escalera alélica de AmpF/STR® MiniFiler™ o un alelo que no se corresponde con una escalera alélica, pero cuyo tamaño está justamente fuera de una ventana debido a error de medición. El umbral de altura del pico mínimo de >50 URF se estableció basándose en experimentos de validación realizados para evitar el tipado cuando es probable que efectos estocásticos interfieran con la precisa interpretación de mezclas. El cálculo de la fracción fetal se basa en promediar todos los marcadores informativos. Los marcadores informativos se identifican por la presencia de picos en el electroferograma que entran dentro de los parámetros de predeterminar cajas para las STR que se analizan.

Los cálculos de la fracción fetal se realizaron usando la altura promedio del pico para los alelos principal y secundarios en cada locus de STR determinado de inyecciones por triplicado. Las reglas aplicadas al cálculo son:

1. Datos de alelos fuera de la escalera (OL) para alelos que no están incluidos en el cálculo; y
2. solo alturas de pico derivada de >50 URF (unidades relativas de fluorescencia) están incluidas en el cálculo
3. si solo una caja está presente, el marcador se considera no informativo; y
4. si se llama una segunda caja, pero los picos de la primera y segunda cajas están dentro del 50-70 % de sus unidades relativas de fluorescencia (RFU) en la altura del pico, la fracción secundaria no se mide y el marcado se considera no informativo.

La fracción del alelo secundario para cualquier marcador informativo se calcula dividiendo la altura del pico del componente secundario por la suma de la altura del pico para el componente principal, y expresado como un porcentaje se calculó primero para cada locus informativo como

$$\text{fracción fetal} = \left(\frac{\sum \text{picos de menor alelo}}{\sum \text{picos de mayor alelo(s)}} \right) \times 100$$

65

La fracción fetal para una muestra que comprende dos o más STR informativas se calcularía como el promedio de las fracciones fetales calculadas para los dos o más marcadores informativos.

La Tabla 8 proporciona los datos obtenidos de analizar ADNlc de un sujeto embarazado con un feto masculino.

TABLA 8

Fracción fetal determinada en en cfADN de un Suieto Embarazado oor análisis de STRs								
STR	Alelo 1	Alelo 2	Alelo 3	Alelo 1 Altura	Alelo 2 Altura	Alelo 3 Altura	Fracción Fetal	Fracción Fetal (Media/STR)
AMEL	X	Y		3599	106		2.9	
AMEL	X	Y		3602	110		3.1	
AMEL	X	Y		3652	109		3.0	3.0
CSF1PO	11	12		2870	2730			
CSF1PO	11	12		2924	2762			
CSF1PO	11	12		2953	2786			
D13S317	11	12		2621	2588			
D13S317	11	12		2680	2619			
D13S317	11	12		2717	2659			
D16S539	9	11		1056	1416			
D16S539	9	11		1038	1394			
D16S539	9	11		1072	1437			
D18S51	13	15		2026	1555			
D18S51	13	15		2006	1557			
D18S51	13	15		2050	1578			
D21S11	28	31.2		2450	61		2.5	
D21S11	28	31.2		2472	62		2.5	
D21S11	28	31.2		2508	67		2.7	2.6
D2S1338	20	23		3417	3017			
D2S1338	20	23		3407	3020			
D2S1338	20	23		3493	3055			
D7S820	9	12	13	2373	178	1123	5.1	
D7S820	9	12	13	2411	181	1140	5.1	
D7S820	9	12	13	2441	182	1156	5.1	5.1
FGA	17.2	22	25	68	1140	896	3.3	
FGA	17.2	22	25	68	1144	909	3.1	
FGA	17.2	22	25	68	1151	925	3.3	3.2
Fracción Fetal =3.5								

Los resultados muestran que el ADNlc puede usarse para determinar la presencia o ausencia de ADN fetal como se indica por la detección de un componente secundario en uno o más alelos de STR, para determinar el porcentaje de fracción fetal y para determinar el sexo fetal como se indica por la presencia o ausencia del alelo amelogenina.

Ejemplo 11**Preamplificación de ADNlc para determinar la fracción fetal por electroforesis capilar de secuencias polimórficas que comprenden STR**

Para mejorar la sensibilidad del ensayo de STR en detectar y cuantificar los alelos de STR en el contribuyente secundario de la muestra de ADNlc, se aumentó el número de genomas de partida en las muestras artificiales por una estrategia de amplificación del genoma completo modificada.

Se recogieron muestras de sangre periférica y se procesaron como se ha descrito en el Ejemplo 2. Se extrajo ADN libre de células de 1 ml de plasma libre de células usando el kit de aislamiento de ácidos nucleicos I Roche MagNA Pure Compact - Large Volumen (Roche Applied Science, IN) usando el instrumento MagNA Pure Compact, y se eluyó en 50 μ l de tampón de elución. Se usaron diez microlitros del ADNlc extraído para cuantificar el ADNlc, y el resto se almacenó (véanse las instrucciones de almacenamiento WI0035 Almacenamiento de muestras clínicas). La concentración del ADNlc extraído del plasma se determinó por cuantificación basada en fluorescencia con mediciones de absorbancia de UV usando la plataforma Qubit™ Quantitation (Invitrogen).

Se determinó que la concentración de ADNlc cuantificado en muestras de plasma preparadas usando el kit de aislamiento de ácidos nucleicos I MagnaPure de 16 sujetos embarazados oscilaba entre 20 y 100 pg/ μ l. Como se sabe que el componente fetal del ADNlc de plasma contribuye al 3-10% del ADNlc de plasma total, se crearon muestras de plasma artificiales enriqueciendo alícuotas de ADNlc derivado de plasma de sujetos voluntarios femeninos con ADNlc extraído de plasma de sujetos voluntarios masculinos para imitar las relaciones de ADNlc fetal con respecto a materno encontrado en los sujetos embarazados. Se crearon muestras artificiales para contener 200-1000 pg de ADNlc femenino extraído que se enriquecieron con 45-150 pg de ADNlc masculino extraído en un volumen total de 10 μ l. Cada muestra artificial se enriqueció para contener 3 %, 5 % y 10 % de ADNlc masculino.

Las muestras artificiales que tienen concentraciones de ADNlc total inferiores a aproximadamente 50 pg/ μ l se preamplificaron usando amplificación por PCR de amplificación por extensión de cebadores mejorada modificada (mIPEP) según el método de Hanson y Ballantyne, (Hanson y Ballantyne, Analytical Biochem 346:246-257 [2005]) del siguiente modo. Se amplificaron diez microlitros de ADNlc de plasma enriquecido en un volumen de reacción de 25 μ l que contenía dNTPs 1 mM, MgCl₂ 2,5 mM (Applied Biosystems), 1X tampón Expand High Fidelity (Nº 3), 10,5 U de mezcla de enzimas Expand High Fidelity (Roche Diagnostics) y cebador PEP 40 μ M (5' - NNNNNNNNNNNNNNNN-3', Qiagen). La amplificación se realizó en un ciclador térmico GeneAmp PCR System 9700 bajo las siguientes condiciones: (1) 20 y 30 ciclos de 94 °C durante 1 minuto, 37 °C durante 2 minutos, y 0,1 °C/s de rampa hasta 55 °C durante 4 minutos. El producto de amplificación se purificó usando una columna Qiagen. La concentración del producto de amplificación se determinó usando la plataforma de cuantificación Qubit™ como se ha descrito anteriormente. Se realizó análisis de STR como se ha descrito en el Ejemplo 9 anterior, excepto que solo se incluyeron alturas de pico >100 URF en los cálculos.

Los resultados se muestran en las Tablas 9, 10 y 11. Los resultados proporcionados en la Tabla 9 muestran que el ADNlc contenido en 10 μ l de ADNlc de muestras artificiales ART23 y ART24 que tienen una concentración inicial de ADNlc de 46,2 y 50,2 pg/ μ l, respectivamente, se amplificó aproximadamente 5 y 10 veces tras 20 y 30 ciclos de amplificación por PCR, respectivamente.

Estos datos indican que una pre-amplificación de ADNlc usando el método de mIPEP proporcionado potenció los niveles de ADNlc total, haciendo el nivel del componente secundario más susceptible al análisis de STR.

TABLA 9

Preamplificación con mIPEP			
Muestra	cfADN sin mIPEP (pg/ μ l)	cfADN con mIPEP: 20 ciclos PCR (pg/50 μ l)	cfADN con mIPEP: 30 ciclos PCR (pg/50 μ l)
ART23	46.2	2265	4125
ART24	50.2	2085	3875

La Tabla 10 muestra mediciones por triplicado que perfilan 9 loci del ADNlc de muestras enriquecidas ART23 y ART24 siguiendo el procedimiento de mIPEP con 20 y 30 ciclos de amplificación, como se ha descrito anteriormente.

5 Los datos en la Tabla 11 indican que la pre-amplificación de ADNlc permite la detección y cuantificación del componente secundario en la mayoría de los loci probados en muestras artificialmente mezcladas que tienen una concentración de ADNlc inicial que, por lo demás, no permitiría un análisis preciso de los alelos de STR secundarios.

TABLA 10

10

15

20

25

30

35

40

45

50

55

60

Preamplificación de mIPEP y Detección de Componente Menor								
		ART23 (453pg)	ART23 (82 5pg)	ART23 (462pg)		ART24 (417pg)	ART24 (775pg)	ART24 (502pg)
		mIPEP amplificado 20 ciclos	mIPEP amplificado 20 ciclos	Extracto inamplificad cfADN		mIPEP amplificado 30 ciclos	mIPEP amplificado 30 ciclos	Extracto inamplificad cfADN
STR Locus	Alelo	Alelo altura	Alelo altura	Alelo altura	Alelo	Alelo altura	Alelo altura	Alelo altura
AMEL	X/Y	291/95	397/170	535/832	X/Y	695/359	1878/1148	1564/1959
AMEL	X/Y	425/147	428/188	675/1048	X/Y	1216/619	1551/954	1573/1943
AMEL	X/Y	267/94	455/203	664/1043	X/Y	718/363	1479/924	1621/2024
CSF1PO	10/11	800/979	725/1009	1429/1325	11/12	2029/1317	4159/2317	2990/3083
CSF1PO	10/11	1147/1432	789/1102	1779/1650	11/12	3449/2223	3460/113/1890	2996/3118
CSF1PO	10/11	729/906	831/1162	1783/1657	11/12	2006/1309	3362/1840	3072/3183
D13S317	12	743	515	1229	11	955	1490	3634
D13S317	12	1079	563	1534	11	1631	1198	3631
D13S317	12	668	583	1520	11	968	1170	3795
D16S539	9/10	239/140	370/466	835/676	10/11	513/512	1173/1472	1678/973
D16S539	9/10	347/203	64*(OL)/3 91/489	1046/864	10/11	859/870	973/1212	1730/999
D16S539	9/10	227/134	441/515	1055/860	10/11	530/513	960/1183	1784/1044
D18S51	14/15	359/464	363/220	785/541	12/18	1044/576	1840/786	2559/1507
		512/645	391/226	999/672	12/18	1769/994	1511/643	2565/1469
		313/402	409/245	994/685	12/18	1033/567	1496/631	2643/1523
D21S11	29/32	103/104	114/173	605/413	31.2	381	661	3276
		149/153	130/182	759/523	31.2	650	536	3028
		85/86	131/196	760/525	31.2	380	520	3282
D2S1338	18/20	572/383	428/363	1116/1013	19/20	1066/433	2315/1243	2962/2968
		827/553	454/386	1428/1279	19/20	1821/757	1901/101	2942/2942

65

(Continúa)

Preamplificación de mIPEP y Detección de Componente Menor									
		ART23 (453pg) mIPEP amplificado 20 ciclos	ART23 (82 5pg) mIPEP amplificado 20 ciclos	ART23 (462pg) Extracto inamplificad cfADN		ART24 (417pg) mIPEP amplificado 30 ciclos	ART24 (775pg) mIPEP amplificado 30 ciclos	ART24 (502pg) Extracto inamplificad cfADN	
5									
10		530/351	482/408	1431/1275	19/20	1063/444	1859/1012	3072/3067	
	D7S820	11/12	262/167	149/270	557/627	11/12	256/138	520/322	1550/1548
15		62/366/23.1	162/292	699/775	11/12	448/236	419/258	1484/1466	
		224/146	169/307	689/779	11/12	253/141	406/250	1579/1573	
20	FGA	21/23	263/146	181/88	596/365	22/24	228/244	375/429	1272/1064
		384/215	191/92	762/450	22/24	409/425	303/345	1221/1023	
		230/136	202/102	749/456	22/24	232/250	297/348	1298/1087	
25	**OL" significa "Medida fuera de Escala"								

30

35

40

45

50

55

60

65

TABLA 11

Fracción Fetal Determinada en una muestra siguiendo Preamplificación usando mPEP							
5	Marcador STR	Alelo 1/ altura	Alelo 2/ altura	Alelo 3/ altura	Alelo 4/ altura	Fracción de porcentaje	Fracción de porcentaje
						menor/STR-menor >100 RFU	menor/STR-menor <100 RFU
	Amelogenin	X/2799	Y/207				
10	Amelogenin	X/2751	Y/198				
	Amelogenin	X/3109	Y/232				
		X/2886	Y/212			7	
15	CSF1PO	10/2377	11/1869	12/508			
	CSF1PO	10/2299	11/1814	12/498			
	CSF1PO	10/2616	11/206	12/562			
20		10/2431	11/1917	12/523		12	
	D13S317	10/1232	11/1600	13/186			
	D13S317	10/1208	11/1548	13/182			
25	D13S317	10/1386	11/1758	13/212			
		10/1275	11/1635	13/193		12	
	D16S539	11/757	12/933				
30	D16S539	11/729	12/885				
	D16S539	11/836	12/1031				
		11/774	12/950			12	
35	D18S51	<u>QL/80</u>	14/3137	15/371			
	D18S51	<u>11/73</u>	14/3082	15/362			
	D18S51	<u>QL/83</u>	14/3488	15/413			
40		<u>QL</u>	14/3236	15/382			
	D21S11	29/953	30/941				
	D21S11	29/921	30/908				
45	D21S11	29/1046	30/1045				
		29/973	30/965				
	D2S1338	17/461	18/366	20/2280	24/1760		
50	D2S1338	17/460	18/360	20/2240	24/1712		
	D2S1338	17/508	18/409	20/2563	24/1971		
		17/476	18/378	20/2361	24/1814	20	
55	D7S820	8/1409	9/60	12/1059			
	D7S820	8/1380	9/60	12/1036			
	D7S820	8/1561	9/69	12/1166			
60		8/1450	9/63	12/1087			2
	FGA	19/825	21/850	25/279			
	FGA	19/807	21/841	25/265			
65	FGA	19/913	21/958	25/306			

Fracción Fetal Determinada en una muestra siguiendo Preamplificación usando mPEP						
Marcador STR	Alelo 1/ altura	Alelo 2/ altura	Alelo 3/ altura	Alelo 4/ altura	Fracción de porcentaje menor/STR-menor	Fracción de porcentaje menor/STR-menor
					≥100 RFU	<100 RFU
	19/848	21/883	25/283		16	
%Fracción Fetal > 100RFU Para alelo menor +12					12	
%Fracción Fetal incluye < 100RFU Para alelo menor +11						11

20 **Ejemplo 12**

Correlación de la fracción fetal determinada por análisis de SNP y STR fetales y maternos

25 Para verificar que la fracción fetal calculada, es decir, la fracción de componente de ácido nucleico secundario, determinada usando el ensayo de SNP y STR como se describe en los ejemplos precedentes proporcionó una medición precisa de la fracción fetal, se comparó el porcentaje de la fracción fetal de ADNc en el plasma de los mismos sujetos embarazados.

30 Se obtuvieron muestras de sangre periférica de 48 sujetos voluntarios, 24 de los sujetos estuvieron embarazadas con fetos masculinos y 24 estuvieron embarazadas con fetos femeninos. Se preparó ADNc como se describe en el Ejemplo 1. Se determinó la fracción fetal usando SNP en secuenciación masivamente paralela por síntesis como se describe en el Ejemplo 5, y la fracción fetal usando STR se determinó usando electroforesis capilar como se describe en el Ejemplo 10

35 Los resultados mostrados en la Figura 6 indican que existe una correlación positiva entre la fracción determinada usando el ensayo de STR y la fracción determinada usando la secuenciación de SNP. Estos datos validan adicionalmente el uso de secuencias polimórficas que comprenden STR o SNP para determinar la fracción de ADNc fetal en una muestra de plasma.

40 **Ejemplo 13**

Uso de la fracción fetal para establecer umbrales y estimar el tamaño de muestra mínimo en la detección de aneuploidía

45 Se manipulan los recuentos de coincidencias de secuencias con cromosomas diferentes para generar una puntuación que variará con el número de copias cromosómicas que pueden interpretarse para identificar amplificación o delección cromosómica. Por ejemplo, una puntuación tal podría generarse comparando la cantidad relativa de una marca de secuencia en un cromosoma que experimenta cambios del número de copias con un cromosoma que se sabe que es euploide. Ejemplos de puntuaciones que pueden usarse para identificar la amplificación o delección incluyen, pero no se limitan a: recuentos para el cromosoma de interés divididos por los recuentos de otro cromosoma de la misma serie experimental, los recuentos para el cromosoma de interés divididos por el número total de recuentos de la serie experimental, comparación de recuentos de la muestra de interés frente a una muestra de control separada. Sin pérdida de generalidad, puede asumirse que las puntuaciones aumentarán a medida que aumenta el número de copias. El conocimiento de la fracción fetal puede usarse para establecer umbrales de "corte" para llamar estados de "aneuploidía", "normales", o "marginales" (inciertos). Entonces, se realizan cálculos para estimar el número mínimo de secuencias requeridas para lograr sensibilidad adecuada (es decir, probabilidad de identificar correctamente un estado de aneuploidía).

60 La Figura 7 es un gráfico de dos poblaciones de puntuaciones diferentes. El eje x es la puntuación y el eje y es la frecuencia. La puntuaciones en muestras de cromosomas sin aneuploidía pueden tener una distribución mostrada en la Figura 7A. La Figura 7B ilustra una distribución hipotética de una población de puntuaciones en muestras con un cromosoma amplificado. Sin pérdida de generalidad, los gráficos y ecuaciones muestran el caso de una puntuación monofactorial en la que la condición de aneuploidía representa una amplificación del número de copias. Los casos multifactoriales y/o las anomalías de reducción/delección son simples extensiones o transposiciones de las descripciones dadas y está previsto que entren dentro del alcance de esta materia.

La cantidad de “solapamiento” entre las poblaciones puede determinar cómo de bien pueden discriminarse los casos normales y de aneuploidía. En general, al aumentar la fracción fetal, ff, aumenta la potencia de discriminación separando los dos centros de poblaciones (moviendo “C2”, el “Centro de puntuaciones de aneuploidía”, y aumentando “d”, haciendo que las poblaciones se solapen menos. Además, un aumento en el valor absoluto de la magnitud, m, (por ejemplo, que tiene cuatro copias del cromosoma en lugar de una trisomía) de la amplificación también aumentará la separación de centros de la población, conduciendo a mayor potencia (es decir, mayor probabilidad de identificar correctamente estados de aneuploidía).

El aumento en el número de secuencias generadas, N, reduce las desviaciones estándar “sdevA” y/o “sdevB”, la diseminación de las dos poblaciones de puntuaciones, que también hace que las poblaciones se solapen menos.

Establecimiento de umbrales y estimación del tamaño de muestra

El siguiente procedimiento puede usarse para establecer “c”, el valor crítico para llamar estados de “aneuploidía”, “normales” o “marginales” (incierto). Sin pérdida de generalidad, a continuación se usan pruebas estadísticas unilaterales.

Primero, se decide una tasa de positivos falsos aceptable, FP (algunas veces también llamada “error de tipo I” o “especificidad”), que es la probabilidad de un positivo falso o llamar falsamente una neuploidía. Por ejemplo, FP puede ser al menos, o aproximadamente 0,001, 0,002, 0,003, 0,004, 0,005, 0,006, 0,007, 0,008, 0,009, 0,01, 0,02, 0,03, 0,04, 0,05, 0,06, 0,07, 0,08, 0,09 o 0,1.

Segundo, el valor de “c” puede determinarse resolviendo la ecuación: FP = integral de c al infinito de (f1 (x)dx).

$$FP = \int_c^{\infty} f1(x)dx \tag{ecuación 1}$$

Una vez se ha determinado un valor crítico, c, puede estimarse el número mínimo de secuencias requeridas para lograr una cierta TP = Tasa de positivos verdaderos. La tasa de positivos verdaderos puede ser, por ejemplo, aproximadamente 0,5, 0,6, 0,7, 0,8 o 0,9. En una realización, la tasa de positivos verdaderos puede ser 0,8. En otras palabras, N es el número mínimo de secuencias requeridas para identificar aneuploidía el 100*TP por ciento de las veces. N = número mínimo de forma que TP = integral desde c hasta el infinito de f2(x,ff)dx > 0,8. N se determina resolviendo

$$\min_N \text{ s.t. } \{TB \geq \int_c^{\infty} f2(x, N)dx\} \tag{ecuación 1}$$

En las pruebas estadísticas clásicas f1 y f2 son frecuentemente F, distribuciones F no centradas (un caso especial de distribuciones t y y no centradas), aunque esto no es una condición necesaria para la presente solicitud.

Establecer “niveles” de umbrales para dar más control de errores

Los umbrales también pueden establecerse en etapas usando los métodos anteriores. Por ejemplo, un umbral puede establecerse para llamada de alta confianza de “aneuploidía”, llamémosle ca, que usa FP 0,001 y un umbral “marginal”, llamémosle cb, que usa FP 0,05. En este caso, si la puntuación, S:

(S > ca)	Llamado “Trisomía”
(cb > S <= ca)	Llamado “Marginal”
(S < cb)	Llamado “Normal”

Algunas generalizaciones triviales que se encuentran dentro del alcance de esta materia

Pueden usarse diferentes valores para los umbrales tales como TP, FP, etc. Los procedimientos pueden realizarse en cualquier orden. Por ejemplo, puede empezarse con N y resolver para c, etc. Las distribuciones pueden depender de ff de manera que f1(x,N,ff), f2(x,N,ff) y/u otras variables. Las ecuaciones integrales anteriores pueden resolverse por referencia a tablas o métodos informáticos iterativos. Puede estimarse un parámetro no de centralidad y la potencia puede leerse de tablas estadísticas estándar. La potencia estadística y los tamaños de muestras pueden

5 derivarse del cálculo o estimación de medias cuadradas estimadas. Pueden usarse distribuciones teóricas de forma
cerrada tales como f, t, t no centrada, normal, etc. o estimaciones (kernel u otras) para modelar las distribuciones f1,
f2. Puede usarse el establecimiento de umbrales empíricos y la selección de parámetros usando curvas de eficacia
diagnóstica (ROC) y cotejarse con la fracción fetal. Pueden usarse diversas estimaciones de la dispersión de la
distribución (varianza, desviación absoluta media, intervalo intercuartílico, etc.). Pueden usarse diversas
estimaciones del centro de distribución (media, mediana, etc.). Pueden usarse pruebas estadísticas bilaterales a
diferencia de unilaterales. La prueba de hipótesis simple puede reformularse como regresión lineal o no lineal.
10 Pueden usarse métodos combinatorios, simulación (por ejemplo, Montecarlo), maximización (por ejemplo,
maximización de la esperanza), métodos iterativos u otros independientemente o conjuntamente con lo anterior para
establecer potencia estadística o umbrales.

Aunque las realizaciones preferidas de la presente invención se han mostrado y descrito en el presente documento,
será obvio para aquellos expertos en la materia que tales realizaciones se proporcionan a modo de ejemplo solo.
15 Numerosas variaciones, cambios y sustituciones se producirán ahora por aquellos expertos en la materia sin
apartarse de la invención. Debe entenderse que pueden emplearse diversas alternativas a las realizaciones de la
invención descritas en el presente documento en la práctica la invención.

20

25

30

35

40

45

50

55

60

65

Reivindicaciones

1. Un método de determinación de la fracción de ácidos nucleicos fetales en una muestra materna de sangre que comprende una mezcla de ADN fetal y genómico, en el que dicho ADN fetal y genómico es ADN libre de células (ADNlc), comprendiendo dicho método:
- (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en dicha mezcla, en la que cada ácido nucleico diana polimórfico amplificado comprende al menos un sitio polimórfico, usando pares de cebadores cada uno capaz de amplificar una secuencia de ácidos nucleicos diana que comprende un sitio polimórfico en una reacción de PCR de múltiplex, para generar un panel de sitios polimórficos amplificados que contiene un número suficiente de sitios polimórficos tal que al menos 25 sean sitios polimórficos informativos;
- (b) realizar secuenciación masivamente paralela de al menos una porción del producto amplificado obtenido en (a), en la que dicha secuenciación comprende proporcionar una pluralidad de marcas de secuencia para dicho panel de sitios polimórficos amplificados; y
- (c) usar dicha pluralidad de marcas de secuencia para:
- (i) identificar al menos un sitio polimórfico informativo en dicho panel de sitios polimórficos amplificados, en el que dicho al menos un sitio polimórfico informativo se identifica por la diferencia en las secuencias alélicas en cada sitio polimórfico y el número de marcas de secuencia mapeadas en cada uno de los posibles alelos en cada sitio polimórfico, y
- (ii) determinar dicha fracción de ácidos nucleicos fetales usando el número total de marcas que se mapean en un primer alelo en un sitio polimórfico identificado informativo y el número total de marcas que se mapean en un segundo alelo en el sitio polimórfico identificado informativo para calcular dicha fracción de ácidos nucleicos fetales.
2. El método de la reivindicación 1, en el que cada uno de dicha pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un único nucleótido (SNP).
3. El método de la reivindicación 1, en el que cada uno de dicha pluralidad de ácido nucleico diana polimórfico comprende al menos una repetición corta en tándem (STR).
4. El método de la reivindicación 1, en el que dicha secuenciación masivamente paralela es:
- (a) secuenciación por síntesis con terminadores de colorante reversible; o
- (b) secuenciación por ligación.
5. El método de la reivindicación 1, en el que dicha secuenciación:
- (a) es secuenciación de una molécula única; o
- (b) comprende una amplificación.
6. El método de la reivindicación 1, en el que dicha muestra materna es plasma o suero.
7. El método de la reivindicación 2, en el que dicho al menos un SNP es un SNP único seleccionado de rs560681, ras1109037, rs9866013, rs13182883, rs13218440, rs7041158, rs740598, rs10773760, rs4530059, rs7205345, rs8078417, rs576261, rs2567608, rs430046, rs9951171, rs338882, rs10776839, rs9905977, rs1277284, rs258684, rs1347696, rs508485, rs9788670, rs8137254, rs3143, rs2182957, rs3739005 y rs530022.
8. El método de la reivindicación 2, en el que dicho al menos un SNP es un SNP en tándem seleccionado de los pares de SNP en tándem rs7277033-rs2110153; rs2822654-rs1882882; rs368657-rs376635; rs2822731-rs2822732; rs1475881-rs7275487; rs1735976-rs2827016; rs447340-rs2824097; rs418989- rs13047336; rs987980- rs987981; rs4143392-rs4143391; rs1691324-rs13050434; rs11909758-rs9980111; rs2826842-rs232414; rs1980969-rs1980970; rs9978999-rs9979175; rs1034346-rs12481852; rs7509629-rs2828358; rs4817013-rs7277036; rs9981121-rs2829696; rs455921-rs2898102; rs2898102-rs458848; rs961301-rs2830208; rs2174536-rs458076; rs11088023-rs11088024; rs1011734-rs1011733; rs2831244-rs9789838; rs8132769-rs2831440; rs8134080-rs2831524; rs4817219-rs4817220; rs2250911-rs2250997; rs2831899-rs2831900; rs2831902-rs2831903; rs1 1088086-rs2251447; rs2832040-rs11088088; rs2832141-rs2246777; rs2832959-rs9980934; rs2833734-rs2833735; rs933121-rs933122; rs2834140-rs12626953; rs2834485-rs3453; rs9974986-rs2834703; rs2776266-rs2835001; rs1984014-rs1984015; rs7281674-rs2835316; rs13047304-rs13047322; rs2835545-rs4816551; rs2835735-rs2835736; rs13047608-rs2835826; rs2836550-rs2212596; rs2836660-rs2836661; rs465612-rs8131220; rs9980072-rs8130031; rs418359-rs2836926; rs7278447-rs7278858; rs385787-rs367001; rs367001-rs386095; rs2837296-rs2837297; y rs2837381-rs4816672.
9. El método de la reivindicación 3, en el que dicha al menos una STR está seleccionada de CSF1PO, FGA, TH01, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, D2S1338, Penta D, Penta E, D22S1045, D20S1082, D20S482, D18S853, D17S1301, D17S974, D14S1434, D12ATA63, D11S4463, D10S1435,

D10S1248, D9S2157, D9S1122, D8S1115, D6S1017, D6S474, D5S2500, D4S2408, D4S2364, D3S4529, D3S3053, D2S1776, D2S441, D1S1677, D1S1627 y D1GATA113.

- 5
10. El método de la reivindicación 1, en el que dicho método es un método independiente del sexo del feto.
11. El método de la reivindicación 1, en el que dicha pluralidad de ácidos nucleicos polimórficos se localiza en una:
- (a) pluralidad de cromosomas diferentes; o
 - (b) cromosoma distintos del cromosoma 13, 18, 21, X o Y.
- 10
12. El método de la reivindicación 1, en el que dicha pluralidad de ácidos nucleicos polimórficos comprende al menos 25, al menos 30 o al menos 40 sitios polimórficos.
- 15
13. El método de la reivindicación 1, en el que dicho panel de sitios polimórficos amplificados contiene un número suficiente de sitios polimórficos tal que al menos 30, o al menos 40, sean sitios polimórficos informativos.
14. El método de cualquier reivindicación precedente, en el que dicho método se combina con:
- (a) un método usado para determinar la presencia o ausencia de aneuploidía fetal en la muestra materna de sangre;
 - o
 - (b) un ensayo para determinar la presencia o ausencia de una afección prenatal asociada al feto.
- 20
15. El método de la reivindicación 14(a), en el que dicha aneuploidía fetal es una monosomía o trisomía.
- 25
- 30
- 35
- 40
- 45
- 50
- 55
- 60
- 65

FIGURA 1

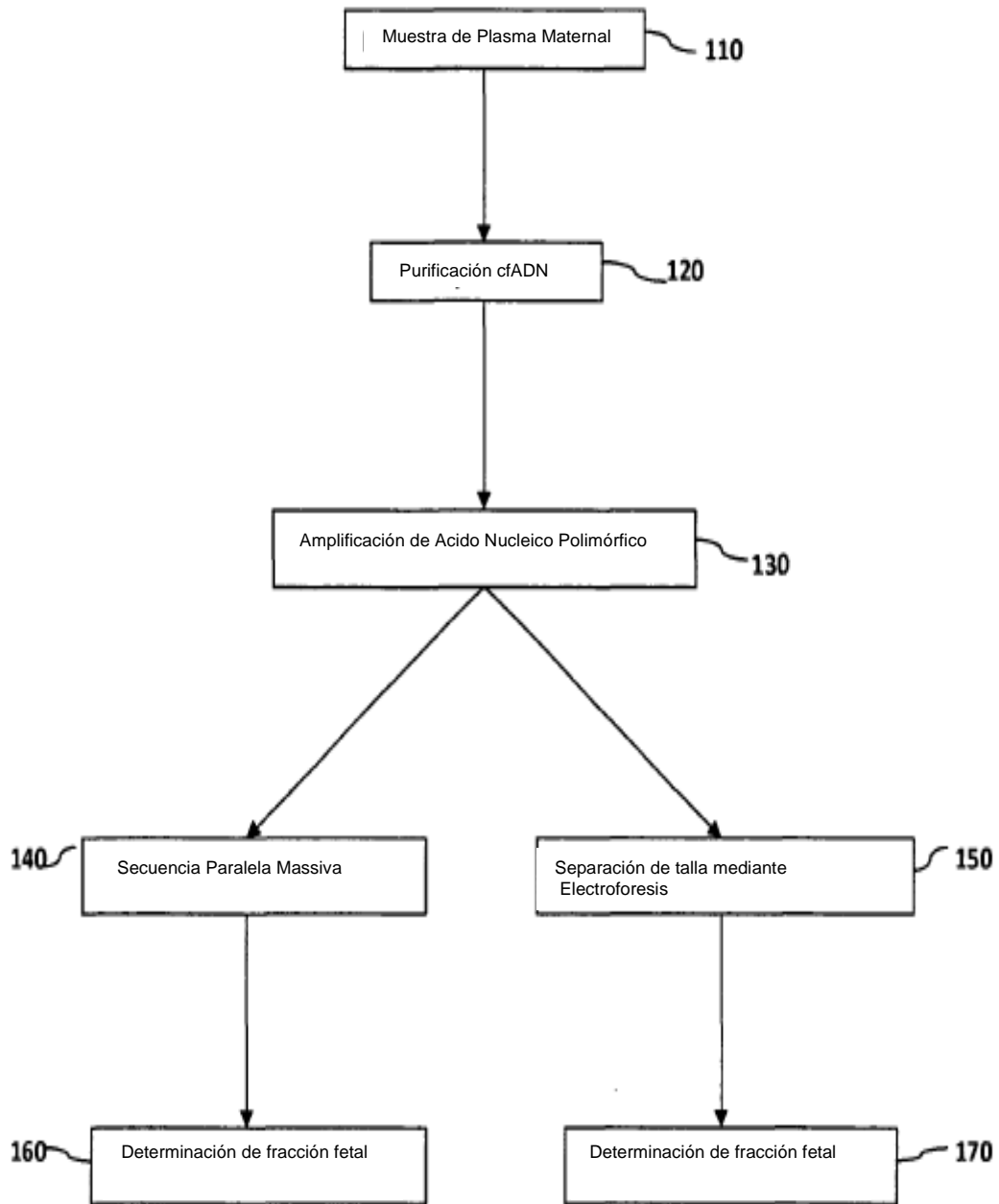


FIGURA 2

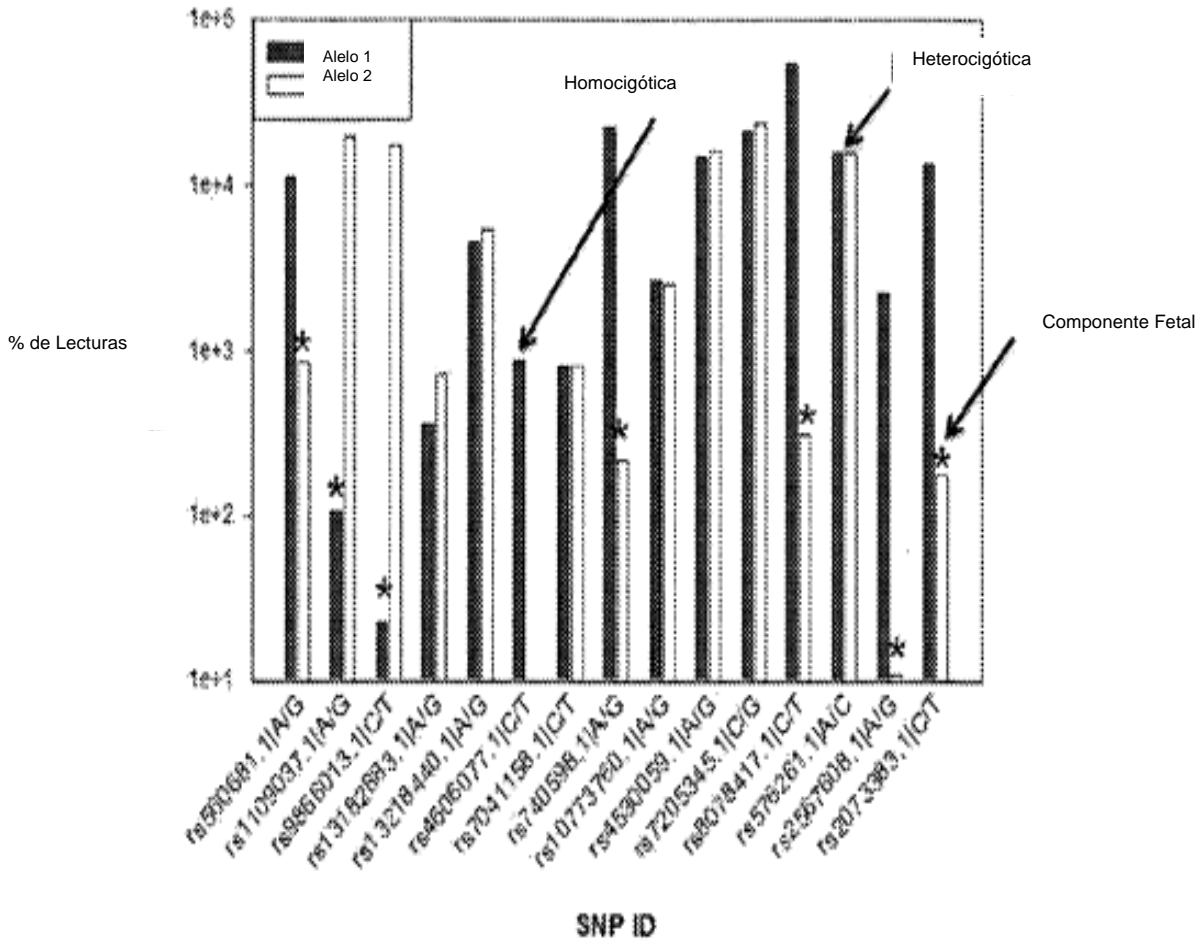


FIGURA 3

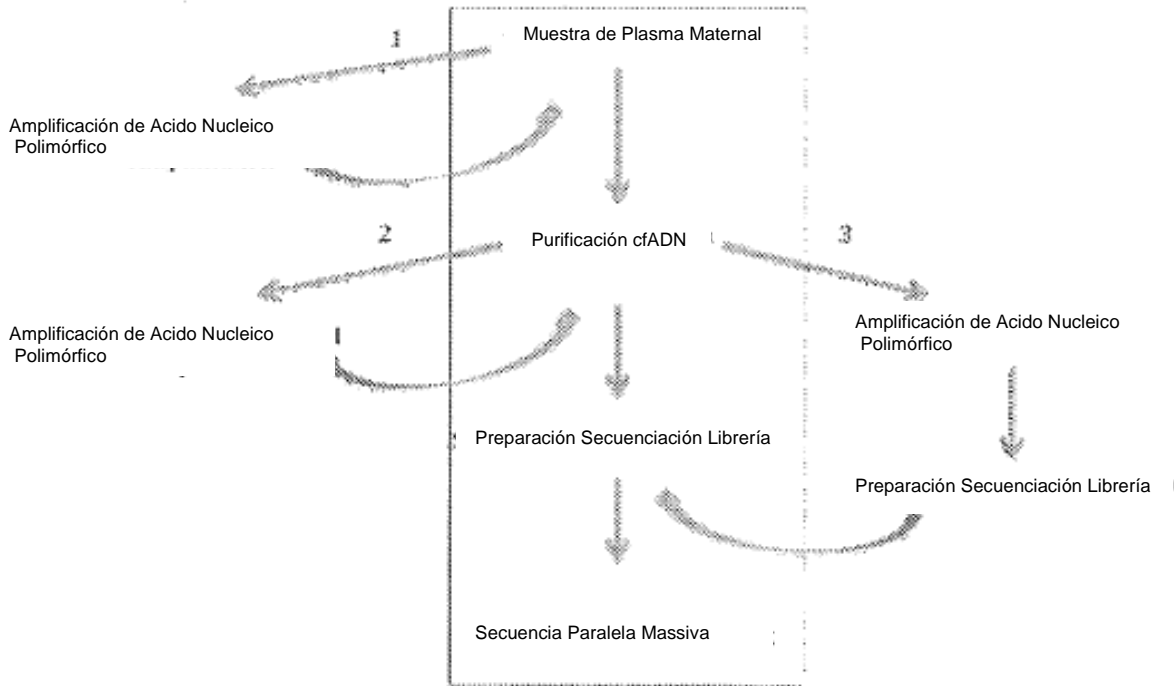


FIGURA 4

Loci Amplificados: Designación Locus	Localización Cromosoma	Alelos incluidos en Identificador de escala Alelica	Etiqueta de tinte	Control ADN 9947A
D8S1179	8	8, 9 10, 11, 12, 13, 14, 15, 16, 17, 18, 19	6-FAM	13a
D21S11	21q11.2-q21	24, 24.2, 25, 26, 27, 28, 28.2, 29, 29.2, 30, 30.2, 31, 31.2, 32, 32.2, 33, 33.2, 34, 34.2, 35, 35.2, 36, 37, 38		30b
D7S820	7q11.21-22	6, 7, 8, 9, 10, 11, 12, 13, 14, 15		10, 11
CSFIPO	5q33.3-34	6, 7, 8, 9, 10, 11, 12, 13, 14, 15		10, 12
D3S1358	3p	12, 13, 14, 15, 16, 17, 18, 19	VIC	14, 15
TH01	11p15.5	4, 5, 6, 7, 8, 9, 9.3, 10, 11, 13.3		8, 9.3
D13S317	13q22-31	8, 9, 10, 11, 12, 13, 14, 15		11c
D16S539	16q24-qter	5, 8, 9, 10, 11, 12, 13, 14, 15		11, 12
D2S1338	2q35-37.1	15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28		19, 23
D19S433	19q12-13.1	9, 10, 11, 12, 12.2, 13, 13.2, 14, 14.2, 15, 15.2, 16, 16.2, 17, 17.2	NED	14, 15
vWA	12p12-pter	11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24		17, 18
TPOX	2p23-2per	6, 7, 8, 9, 10, 11, 12, 13		8d
D18S51	18q21.3	7, 9, 10, 10.2, 11, 12, 13, 13.2, 14, 14.2, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27		15, 19
Amelogenina	X: p22.1-22.3 Y: p11.2	X, Y	PET	X
D5S818	5q21-31	7, 8, 9, 10, 11, 12, 13, 14, 15, 16		11e
FGA	4q28	17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 26.2, 27, 28, 29, 30, 30.2, 31.2, 32.2, 33.2, 42.2, 43.2, 44.2, 45.2, 46.2, 47.2, 48.2, 50.2, 51.2		23, 24

FIGURA 5

Loci Amplificado: Designación Locus	Localización Cromosoma	Alelos incluidos en Identificador de escala Alelica	Etiqueta de tinte	Control ADN 9947A
D8S1179	8	8, 9 10, 11, 12, 13, 14, 15, 16, 17, 18, 19	6-FAM	13a
D21S11	21q11.2-q21	24, 24.2, 25, 26, 27, 28, 28.2, 29, 29.2, 30, 30.2, 31, 31.2, 32, 32.2, 33, 33.2, 34, 34.2, 35, 35.2, 36, 37, 38		30b
D7S820	7q11.21-22	6, 7, 8, 9, 10, 11, 12, 13, 14, 15		10, 11
CSFIPO	5q33.3-34	6, 7, 8, 9, 10, 11, 12, 13, 14, 15		10, 12
D3S1358	3p	12, 13, 14, 15, 16, 17, 18, 19	VIC	14, 15
TH01	11p15.5	4, 5, 6, 7, 8, 9, 9.3, 10, 11, 13.3		8, 9.3
D13S317	13q22-31	8, 9, 10, 11, 12, 13, 14, 15		11c
D16S539	16q24-qter	5, 8, 9, 10, 11, 12, 13, 14, 15		11, 12
D2S1338	2q35-37.1	15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28		19, 23
D19S433	19q12-13.1	9, 10, 11, 12, 12.2, 13, 13.2, 14, 14.2, 15, 15.2, 16, 16.2, 17, 17.2	NED	14, 15
vWA	12p12-pter	11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24		17, 18
TPOX	2p23-2per	6, 7, 8, 9, 10, 11, 12, 13		8d
D18S51	18q21.3	7, 9, 10, 10.2, 11, 12, 13, 13.2, 14, 14.2, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27		15, 19
Amelogenina	X: p22.1-22.3 Y: p11.2	X, Y	PET	X
D5S818	5q21-31	7, 8, 9, 10, 11, 12, 13, 14, 15, 16		11e
FGA	4q28	17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 26.2, 27, 28, 29, 30, 30.2, 31.2, 32.2, 33.2, 42.2, 43.2, 44.2, 45.2, 46.2, 47.2, 48.2, 50.2, 51.2		23, 24

FIGURA 6

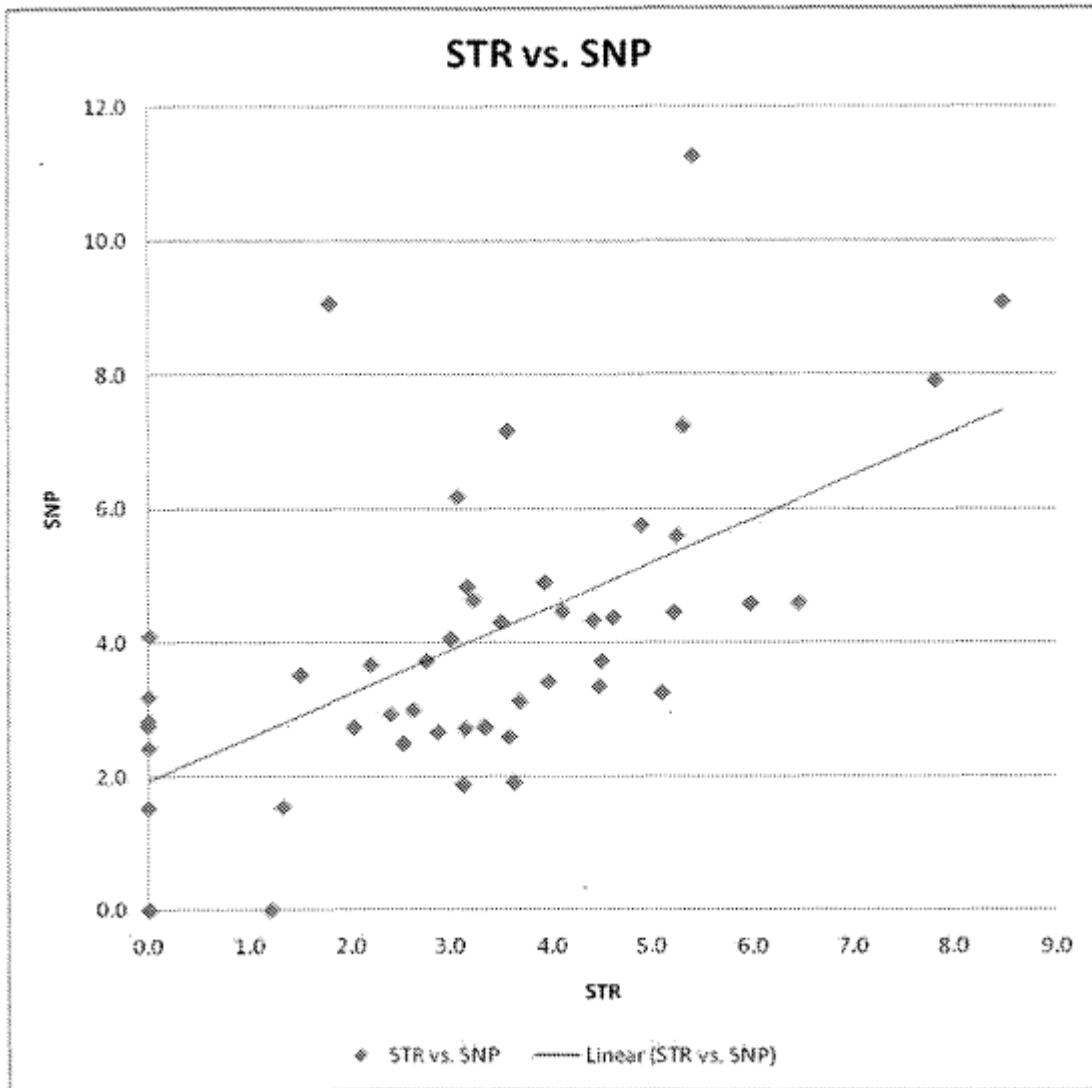


FIGURA 7

