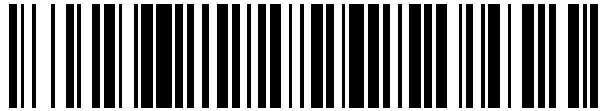


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 721 164**

51 Int. Cl.:

**C12Q 1/6809** (2008.01)

**C12Q 1/6883** (2008.01)

**G06F 19/18** (2011.01)

**G06F 19/22** (2011.01)

**C40B 20/00** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **17.12.2013 PCT/US2013/075683**

87 Fecha y número de publicación internacional: **26.06.2014 WO14099919**

96 Fecha de presentación y número de la solicitud europea: **17.12.2013 E 13864118 (8)**

97 Fecha y número de publicación de la concesión europea: **30.01.2019 EP 2935625**

54 Título: **Detección no invasiva de aneuploidía fetal en embarazos por ovodonación**

30 Prioridad:

**19.12.2012 US 201213720273**

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

**29.07.2019**

73 Titular/es:

**ARIOSIA DIAGNOSTICS, INC. (100.0%)  
5945 Optical Court  
San Jose, CA 95138, US**

72 Inventor/es:

**OLIPHANT, ARNOLD;  
WANG, ERIC y  
STRUBLE, CRAIG**

74 Agente/Representante:

**LINAGE GONZÁLEZ, Rafael**

ES 2 721 164 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).

**DESCRIPCIÓN**

Detección no invasiva de aneuploidía fetal en embarazos por ovodonación

**5 CAMPO DE LA INVENCIÓN**

Esta invención se refiere al diagnóstico de anomalías genéticas en embarazos por ovodonación y sistemas de ensayo para dicho diagnóstico.

**10 ANTECEDENTES DE LA INVENCIÓN**

En la siguiente discusión se describirán determinados artículos y procedimientos con propósitos informativos e introductorios. Nada de lo contenido en el presente documento se debe interpretar como una "admisión" de la técnica anterior. El solicitante se reserva expresamente el derecho de demostrar, cuando corresponda, que los artículos y procedimientos a los que se hace referencia en el presente documento no constituyen la técnica anterior según las disposiciones legales aplicables.

Las anomalías genéticas representan un gran número de patologías, incluyendo patologías causadas por aneuploidía cromosómica (por ejemplo, síndrome de Down), mutaciones hereditarias en genes específicos (por ejemplo, anemia drepanocítica) y patologías causadas por mutaciones somáticas (por ejemplo, cáncer). Los procedimientos de diagnóstico para determinar dichas anomalías genéticas se han convertido en técnicas estándar para identificar enfermedades y trastornos específicos, así como para proporcionar información valiosa sobre la fuente de la enfermedad y las opciones de tratamiento.

Por ejemplo, el cribado y el diagnóstico prenatales se ofrecen de forma rutinaria en la atención prenatal y se considera que son importantes para que las mujeres puedan tomar decisiones informadas acerca de los embarazos afectados por afecciones genéticas. Los procedimientos convencionales de pruebas de diagnóstico prenatal actualmente requieren la extracción de una muestra de células fetales directamente del útero para realizar el análisis genético, usando una biopsia de vellosidades coriónicas (CVS, por sus siglas en inglés) típicamente entre las 11 y 14 semanas de gestación o la amniocentesis típicamente después de las 15 semanas. Sin embargo, estos procedimientos invasivos conllevan un riesgo de aborto espontáneo de alrededor de un 1 %. Mujezinovic y Alfirevic, *Obstet Gynecol* 2007; 110:687-694.

Aunque estos enfoques para la obtención de ADN fetal actualmente proporcionan la prueba de referencia para el diagnóstico prenatal, muchas mujeres deciden no someterse a pruebas invasivas, principalmente porque es desagradable y conlleva un riesgo pequeño pero significativo de aborto espontáneo. Durante mucho tiempo se ha buscado un procedimiento fiable y conveniente para el diagnóstico prenatal no invasivo para reducir este riesgo de aborto espontáneo y permitir pruebas más tempranas. Aunque algunos trabajos han investigado el uso de células fetales obtenidas del moco cervical (Fejgin MD *et al.*, *Prenat Diagn* 2001;21:619-621; Mantzaris *et al.*, *ANZJOG* 2005;45:529-532), la mayoría de las investigaciones se han centrado en estrategias para detectar elementos genéticos del feto presentes en la circulación materna. Se ha demostrado que durante el embarazo existe un tráfico bidireccional entre el feto y la madre (Lo *et al.*, *Blood* 1996; 88:4390-4395), y múltiples estudios han mostrado que tanto las células fetales intactas como los ácidos nucleicos fetales libres circulan por la placenta y circulan por el torrente sanguíneo materno (véase, por ejemplo, Chiu RW y Lo YM, *Semin Fetal Neonatal Med.* 2010, 11 nov). SPARKS *ET AL.*: "Noninvasive prenatal detection and selective analysis of cell -free DNA obtained from maternal blood: evaluation for trisomy 21 and trisomy 18", *AM J OBSTET GYNECOL*, vol. 206, no. 4, abril 2012, páginas 319.E1-319.E9, divulga un procedimiento para determinar el porcentaje de ADN fetal libre circulante en una muestra materna y evaluar la fracción fetal mediante el examen de 192 loci que contienen polimorfismo mononucleotídico (SNP).

En particular, los intentos más recientes de identificar aneuploidías han usado la sangre materna como material de partida. Dichos esfuerzos han incluido el uso de ADN libre circulante para detectar aneuploidía fetal en una muestra de una mujer embarazada, incluyendo el uso de secuenciación paralela masiva de firma (MPSS) para cuantificar con precisión el incremento en los fragmentos de ADNlc (ADN libre circulante) de cromosomas trisómicos. La dosis cromosómica resultante de la aneuploidía fetal, sin embargo, está directamente relacionada con la fracción de ADNlc fetal. La variación de la contribución de los ácidos nucleicos fetales entre las distintas muestras puede complicar el análisis, ya que el nivel de contribución fetal a una muestra materna variará las cantidades que hay que detectar para calcular el riesgo de que un cromosoma fetal sea aneuploide.

Por ejemplo, una muestra ADNlc que contiene un 4 % de ADN de un feto con trisomía 21 debería exhibir un incremento de un 2 % en la proporción de lecturas del cromosoma 21 (chr21) en comparación con un feto normal. Distinguir una trisomía 21 de un feto normal con alta confianza usando una muestra materna con un porcentaje de ácido nucleico fetal de un 4 % requiere un gran número (> 93 K) de observaciones del cromosoma 21, lo cual es complejo y no es rentable usando técnicas no selectivas tales como MPSS.

Existe, por tanto, una necesidad de procedimientos no invasivos de cribado de anomalías genéticas, incluyendo aneuploidías, en muestras mezcladas que comprenden ADN normal y ADN presuntamente anormal. La presente invención aborda esta necesidad.

5 **SUMARIO DE LA INVENCION**

Este Sumario se proporciona para introducir en forma simplificada una selección de conceptos que se describen más en detalle a continuación en la Descripción detallada. Este Sumario no tiene la intención de identificar características clave o esenciales de la materia objeto reivindicada, ni está destinado a ser usado para limitar el alcance de la materia objeto reivindicada. Otras características, detalles, utilidades y ventajas de la materia objeto reivindicada serán evidentes a partir de la siguiente Descripción detallada escrita que incluye los aspectos ilustrados en los dibujos adjuntos y definidos en las reivindicaciones adjuntas.

La presente invención está definida por las reivindicaciones.

La determinación del porcentaje relativo de ADN fetal en una muestra materna de un embarazo por ovodonación proporciona información importante sobre la presencia estadística relativa de regiones de ácido nucleico que pueden ser indicativas de aneuploidía fetal. La determinación de los loci en los que el feto tiene al menos un alelo que es diferente de los alelos de la portadora puede permitir la estimación del ADN fetal en una muestra materna y, por tanto, proporcionar información usada para calcular las diferencias estadísticamente significativas en las frecuencias cromosómicas de los cromosomas de interés. Dichos loci podrían proporcionar dos formas de información en el ensayo: la información alélica se puede usar para determinar el porcentaje de contribución del ADN fetal en la muestra materna y la determinación empírica de la información alélica a partir de las técnicas de detección (por ejemplo, la suma de los niveles de detección alélica) se puede usar para determinar la frecuencia global relativa de ese locus en una muestra materna. Sin embargo, la información alélica no es necesaria para determinar la frecuencia global relativa de ese locus.

Por tanto, en algunos aspectos específicos, la contribución fetal relativa de ADN materno en el alelo de interés se puede comparar con la contribución no materna en ese alelo para determinar la concentración de ADN fetal aproximada en la muestra. La estimación del ADN fetal en una muestra materna se determina en aquellos loci en los que el ADN fetal tiene un alelo que es diferente de los alelos maternos. En un embarazo por ovodonación, esto puede incluir loci en los que el genotipo materno o fetal es homocigótico y el otro es heterocigótico. En esta situación, la cantidad de ADN fetal se puede determinar usando un subconjunto de loci informativos de la donante de óvulos.

En un aspecto específico, la invención proporciona un procedimiento para determinar el porcentaje de ADN fetal libre circulante en una muestra materna de una mujer embarazada con un embarazo por ovodonación que comprende proporcionar una muestra materna que comprende ADN libre circulante materno y fetal; examinar dos o más regiones polimórficas de ácido nucleico seleccionadas; detectar las regiones de ácido nucleico examinadas; determinar una frecuencia relativa de polimorfismos en las regiones de ácido nucleico examinadas para identificar loci informativos de la donante de óvulos; en el que los loci informativos de la donante de óvulos incluyen loci en los que el ADN materno y el ADN fetal difieren en al menos un alelo, y excluyen loci en los que ambos alelos fetales difieren de los alelos maternos; y calcular un porcentaje de ADN fetal libre circulante usando los loci informativos de la donante de óvulos identificados.

En determinados aspectos específicos, la determinación del porcentaje relativo de ADN fetal en una muestra materna puede ser beneficioso para la realización o la optimización de los resultados obtenidos a partir del sistema de ensayo, ya que proporcionará información importante sobre la presencia estadística esperada de los cromosomas fetales y la desviación de esa expectativa puede ser indicativa de aneuploidía fetal. Se pueden usar numerosos enfoques para calcular la contribución relativa del ADN fetal en una muestra materna.

Fuera del alcance de la invención se divulga un sistema de ensayo para proporcionar una probabilidad estadística de la presencia o ausencia de aneuploidía fetal en una mujer embarazada con un embarazo por ovodonación, que comprende proporcionar una muestra materna que comprende ADN libre circulante materno y fetal; examinar dos o más regiones de ácido nucleico de un primer cromosoma; cuantificar una frecuencia relativa de alelos de las regiones de ácido nucleico del primer cromosoma; examinar dos o más regiones de ácido nucleico seleccionadas de un segundo cromosoma; comparar la frecuencia relativa de las regiones de ácido nucleico del primer cromosoma con la frecuencia relativa de las regiones de ácido nucleico del segundo cromosoma; calcular el porcentaje de ADN fetal libre circulante en la muestra materna; y proporcionar una probabilidad estadística de la presencia o ausencia de una aneuploidía fetal usando el porcentaje calculado de ADN fetal libre circulante en la muestra materna.

También se divulga un sistema de ensayo para proporcionar una probabilidad estadística de la presencia o ausencia de aneuploidía fetal en una mujer embarazada con un embarazo por ovodonación, que comprende proporcionar una muestra materna que comprende ADN libre circulante materno y fetal; examinar dos o más regiones de ácido nucleico de un primer cromosoma; detectar las regiones de ácido nucleico examinadas del primer cromosoma; cuantificar una frecuencia relativa de alelos de las regiones de ácido nucleico del primer cromosoma; examinar dos o más regiones de ácido nucleico seleccionadas de un segundo cromosoma; detectar las regiones de ácido nucleico examinadas del

segundo cromosoma; cuantificar una frecuencia relativa de alelos de las regiones de ácido nucleico del segundo cromosoma; comparar la frecuencia relativa de las regiones de ácido nucleico del primer cromosoma con la frecuencia relativa de las regiones de ácido nucleico del segundo cromosoma; calcular el porcentaje de ADN fetal libre circulante en la muestra materna; y proporcionar una probabilidad estadística de la presencia o ausencia de una aneuploidía fetal usando el porcentaje calculado de ADN fetal libre circulante en la muestra materna.

También se divulga además un sistema de ensayo para proporcionar una probabilidad estadística de la presencia o ausencia de aneuploidía fetal en una mujer embarazada con un embarazo por ovodonación, que comprende proporcionar una muestra materna que comprende ADN libre circulante materno y fetal; examinar dos o más regiones de ácido nucleico de un primer cromosoma; cuantificar una frecuencia relativa de alelos de las regiones de ácido nucleico del primer cromosoma; examinar dos o más regiones de ácido nucleico seleccionadas de un segundo cromosoma; cuantificar una frecuencia relativa de alelos de las regiones de ácido nucleico del segundo cromosoma; comparar la frecuencia relativa de las regiones de ácido nucleico del primer cromosoma con la frecuencia relativa de las regiones de ácido nucleico del segundo cromosoma; calcular el porcentaje de ADN fetal libre circulante en la muestra materna; y proporcionar una probabilidad estadística de la presencia o ausencia de una aneuploidía fetal usando el porcentaje calculado de ADN fetal libre circulante en la muestra materna.

En algunos aspectos, el porcentaje de contribución de ADN fetal en una muestra materna se calcula determinando los niveles de uno o más loci sin contribución materna, incluyendo loci de los cromosomas sexuales y loci autosómicos. En aspectos que utilizan loci autosómicos, en general el porcentaje de contribución del ADN fetal se determina comparando una o más variaciones genéticas en los loci no maternos con los loci maternos. En algunos aspectos particulares, estas variaciones genéticas son variaciones en el número de copias. En otros aspectos particulares, estas variaciones genéticas son uno o más polimorfismos de nucleótido único.

El sistema de ensayo analiza preferentemente al menos diez regiones de ácido nucleico polimórficas para cada cromosoma de interés, más preferentemente al menos veinte regiones de ácido nucleico polimórficas para cada cromosoma de interés, más preferentemente al menos cuarenta regiones de ácido nucleico polimórficas para cada cromosoma de interés, e incluso más preferentemente al menos noventa regiones de ácido nucleico polimórficas para cada cromosoma de interés.

El sistema de ensayo se puede configurar como un sistema altamente multiplexado que permite analizar simultáneamente múltiples regiones de ácido nucleico de un único cromosoma o de múltiples cromosomas dentro de una muestra individual y/o de múltiples muestras. En dichos sistemas multiplexados, las muestras se pueden analizar por separado, o se pueden agrupar inicialmente en grupos de dos o más para el análisis de un mayor número de muestras. Cuando se obtienen datos combinados, dichos datos se identifican preferentemente para las diferentes muestras antes del análisis de la aneuploidía. En algunos aspectos, sin embargo, los datos combinados se pueden analizar para detectar potenciales aneuploidías, y las muestras individuales del grupo se analizarán posteriormente si los resultados iniciales indican que se detectó una potencial aneuploidía dentro del grupo combinado.

En determinados aspectos, los sistemas de ensayo utilizarán uno o más índices que proporcionan información para la identificación de un locus o muestra. Por ejemplo, un cebador que se usa en la amplificación selectiva puede tener secuencias adicionales que son específicas de un locus, por ejemplo, una secuencia de ácido nucleico que es indicativa de la región de ácido nucleico seleccionada o un alelo particular de esa región de ácido nucleico. En otro ejemplo se usa un índice en la amplificación selectiva o universal que es indicativa de una muestra a partir de la cual se amplificó el ácido nucleico. En otro ejemplo más se usa un índice de identificación único para distinguir un producto de amplificación particular de otros productos de amplificación obtenidos a partir de los procedimientos de detección. Un solo índice también se puede combinar con cualquier otro índice para crear un índice que proporcione información para dos propiedades (por ejemplo, índice de identificación de muestra, índice del locus de alelo).

En determinados aspectos, las regiones de ácido nucleico se amplifican usando procedimientos de amplificación universal. La amplificación universal se puede realizar, por ejemplo, después de una etapa inicial de amplificación selectiva o enriquecimiento de la muestra materna de un embarazo por ovodonación. El uso de la amplificación universal permite amplificar múltiples regiones de ácido nucleico usando un único o un número limitado de cebadores de amplificación, y es especialmente útil para amplificar múltiples regiones de ácido nucleico seleccionadas en una sola reacción. Esto permite el procesamiento simultáneo de múltiples regiones de ácido nucleico de una o múltiples muestras.

La muestra materna usada para el análisis se puede obtener o derivar de cualquier mujer embarazada con un embarazo por ovodonación. Por ejemplo, una muestra materna puede ser de cualquier fluido materno que comprenda ADN libre circulante tanto materno como fetal, incluyendo, pero sin limitarse a, plasma materno, suero materno o sangre materna.

En aspectos preferentes, se detectan ácidos nucleicos objetivo que corresponden a múltiples regiones de ácido nucleico de un cromosoma y la frecuencia se usa para determinar la frecuencia relativa de un cromosoma en la muestra materna. Las frecuencias que son más altas de lo esperado para una región de ácido nucleico correspondiente a un cromosoma cuando se compara con la cantidad de una región de ácido nucleico correspondiente a otro

5 cromosoma en la muestra materna son indicativas de una duplicación, deleción o aneuploidía fetal. La comparación puede ser una comparación de una región genómica de interés que se inserta o elimina supuestamente en un cromosoma fetal. La comparación también puede ser de cromosomas que cada uno de los cuales pueda ser un supuesto aneuploide en el feto (por ejemplo, los cromosomas 18 y 21), pero en los que la probabilidad de que ambos sean aneuploides es mínima. La comparación también puede ser de cromosomas en los que uno es un supuesto aneuploide (por ejemplo, el cromosoma 21) y el otro actúa como un cromosoma de referencia (por ejemplo, un autosoma tal como el cromosoma 12). En otros aspectos, la comparación puede utilizar dos o más cromosomas que son supuestamente aneuploides y uno o más cromosomas de referencia.

10 Estos y otros aspectos, características y ventajas se proporcionarán con más detalle como se describe en el presente documento.

#### BREVE DESCRIPCIÓN DE LAS FIGURAS

15 La FIG. 1 es un diagrama de flujo simplificado de las etapas generales para la determinación del porcentaje de contribución fetal de ADN libre circulante en una muestra materna.

20 La FIG. 2 es una tabla de combinaciones de genotipos maternos, de donantes de óvulos y fetales en un embarazo por ovodonación en un locus particular.

La FIG. 3 es un gráfico de las comparaciones de fracciones fetales para el conjunto relacionado con la fracción fetal y el conjunto no relacionado de la fracción fetal.

25 Las FIGS. 4A - 4C son gráficos de barras que comparan el porcentaje de contribución estimado usando tanto los ensayos polimórficos como los ensayos de cromosomas sexuales para las muestras mezcladas que comprenden un 75 %, un 80 %, un 85 %, un 90 % y un 95 % de plasma de la mujer.

#### DESCRIPCIÓN DETALLADA DE LA INVENCION

30 Los procedimientos descritos en el presente documento pueden emplear, a menos que se indique lo contrario, técnicas convencionales y descripciones de biología molecular (incluyendo técnicas recombinantes), biología celular, bioquímica y micromatrices y tecnología de secuenciación, que están dentro de los conocimientos de los expertos en la técnica. Dichas técnicas convencionales incluyen síntesis de matriz polimérica, hibridación y ligación de oligonucleótidos, secuenciación de oligonucleótidos y detección de hibridación usando un marcador. Se pueden tener ilustraciones específicas de técnicas adecuadas por referencia a los ejemplos del presente documento. Sin embargo, también se pueden usar procedimientos convencionales equivalentes. Dichas técnicas convencionales y descripciones se pueden encontrar en manuales de laboratorio estándar como Green *et al.*, Eds., *Genome Analysis: A Laboratory Manual Series (Vols. I-IV)* (1999); Weiner, *et al.*, Eds., *Genetic Variation: A Laboratory Manual* (2007); Dieffenbach, Dveksler, Eds., *PCR Primer: A Laboratory Manual* (2003); Bowtell y Sambrook, *DNA Microarrays: A Molecular Cloning Manual* (2003); Mount, *Bioinformatics: Sequence and Genome Analysis* (2004); Sambrook y Russell, *Condensed Protocols from Molecular Cloning: A Laboratory Manual* (2006); y Sambrook y Russell, *Molecular Cloning: A Laboratory Manual* (2002) (todos de Cold Spring Harbor Laboratory Press); Stryer, L., *Biochemistry* (4th Ed.) W.H. Freeman, Nueva York (1995); Gait, "Oligonucleotide Synthesis: A Practical Approach" IRL Press, London (1984); Nelson y Cox, *Lehninger, Principles of Biochemistry*, 3ª ed., W. H. Freeman Pub., Nueva York (2000); y Berg *et al.*, *Biochemistry*, 5ª ed., W.H. Freeman Pub., Nueva York (2002). Antes de que se describan las presentes composiciones, herramientas de investigación y procedimientos, se debe entender que la presente invención no está limitada a los procedimientos, composiciones, dianas y usos específicos descritos, ya que, por supuesto, esto puede variar. También se debe entender que la terminología usada en el presente documento es con la finalidad de describir aspectos particulares solamente, y no se pretende que limite el alcance de la presente invención, que solo estará limitada por las reivindicaciones adjuntas.

50 Cabe señalar que, como se usa en el presente documento y en las reivindicaciones adjuntas, las formas singulares "un", "una", "el" y "la" incluyen los referentes plurales a menos que el contexto indique claramente lo contrario. Así, por ejemplo, la referencia a "una región de ácido nucleico" se refiere a una, más de una o mezclas de dichas regiones, y la referencia a "un ensayo" incluye referencia a etapas y procedimientos equivalentes conocidos por los expertos en la técnica, etc.

60 Cuando se proporciona un intervalo de valores, se debe entender que cada valor que se incluye entre el límite superior e inferior de ese intervalo, y cualquier otro valor indicado o incluido en ese intervalo, se engloba dentro de la divulgación. Cuando el intervalo establecido incluye límites superior e inferior, los intervalos que excluyen cualquiera de esos límites incluidos también están incluidos en la divulgación.

65 En la siguiente descripción se exponen numerosos detalles específicos para proporcionar una comprensión más completa de la presente invención. Sin embargo, será evidente para un experto en la técnica que la presente invención se puede poner en práctica sin uno o más de estos detalles específicos. En otros casos, las características y procedimientos bien conocidos por los expertos en la técnica no se han descrito para evitar complicar la invención.

## Definiciones

A menos que se establezca expresamente, los términos usados en el presente documento pretenden tener el significado simple y corriente tal como lo entienden los expertos en la técnica. Las siguientes definiciones están destinadas a ayudar al lector a comprender la presente invención, pero no pretenden variar o limitar de otra manera el significado de dichos términos a menos que se indique específicamente.

El término "ácido nucleico amplificado" es cualquier molécula de ácido nucleico cuya cantidad se ha incrementado en al menos dos veces mediante cualquier procedimiento de amplificación o replicación de ácido nucleico realizado *in vitro* en comparación con su cantidad de partida.

El término "madre portadora" se refiere a una mujer embarazada que lleva un feto como resultado del uso de un ovocito humano que es de una mujer diferente.

El término "anomalía cromosómica" se refiere a cualquier variante genética para la totalidad o para una parte de un cromosoma. Las variantes genéticas pueden incluir, pero no se limitan a, cualquier variante del número de copias, tal como duplicaciones o deleciones, translocaciones, inversiones y mutaciones.

Los términos "complementario" o "complementariedad" se usan en referencia a moléculas de ácido nucleico (es decir, una secuencia de nucleótidos) que están relacionados por las reglas de apareamiento de bases. Los nucleótidos complementarios son, en general, A y T (o A y U), o C y G. Se dice que dos moléculas de ARN o ADN monocatenarias son sustancialmente complementarias cuando los nucleótidos de una hebra, alineados de manera óptima y con las inserciones o deleciones de nucleótidos apropiadas, se emparejan con al menos aproximadamente un 90 % a aproximadamente un 95 % de complementariedad, y más preferentemente de aproximadamente un 98 % a aproximadamente un 100 % de complementariedad, e incluso más preferentemente con un 100 % de complementariedad. De forma alternativa, existe una complementariedad sustancial cuando una hebra de ARN o ADN hibridará en condiciones de hibridación selectiva a su complemento. Las condiciones de hibridación selectiva incluyen, pero no se limitan a, condiciones de hibridación rigurosas. Las condiciones de hibridación rigurosas incluirán típicamente concentraciones de sal de menos de aproximadamente 1 M, más habitualmente de menos de aproximadamente 500 mM y preferentemente de menos de aproximadamente 200 mM. Las temperaturas de hibridación son, en general, al menos aproximadamente 2 °C a aproximadamente 6 °C inferiores a las temperaturas de fusión ( $T_m$ ).

El término "índice de corrección" se refiere a un índice que puede contener nucleótidos adicionales que permiten la identificación y corrección de errores de amplificación, secuenciación u otros errores experimentales, incluyendo la detección de deleción, sustitución, o inserción de una o más bases durante la secuenciación, así como los cambios de nucleótidos que pueden ocurrir fuera de la secuenciación, tal como síntesis de oligonucleótidos, amplificación y cualquier otro aspecto del ensayo. Estos índices de corrección pueden ser índices independientes que son secuencias separadas, o se pueden incorporar dentro de otros índices para ayudar a confirmar la precisión de las técnicas experimentales usadas, por ejemplo, un índice de corrección puede ser un subconjunto de secuencias de un índice de locus o un índice de identificación.

El término "herramienta de diagnóstico" tal como se usa en el presente documento se refiere a cualquier composición o ensayo de la invención usado en combinación como, por ejemplo, en un sistema con el fin de llevar a cabo una prueba o ensayo de diagnóstico en una muestra de paciente.

El término "embarazo por ovodonación" se refiere a un embarazo resultante del uso de un ovocito humano que no es de la madre portadora.

El término "hibridación" significa, en general, la reacción mediante la cual se produce el emparejamiento de hebras complementarias de ácido nucleico. El ADN suele ser bicatenario y, cuando las hebras se separan, volverán a hibridar en las condiciones apropiadas. Los híbridos se pueden formar entre ADN y ADN, ADN y ARN o ARN y ARN. Se pueden formar entre una hebra corta y una hebra larga que contiene una región complementaria a la corta. Los híbridos imperfectos también se pueden formar, pero cuanto más imperfectos sean, menos estables serán (y menos propensos a formarse).

El término "índice de identificación" se refiere en general a una serie de nucleótidos incorporados en una región de cebador de un procedimiento de amplificación para la identificación única de un producto de amplificación de una región de ácido nucleico. Las secuencias del índice de identificación tienen preferentemente 6 o más nucleótidos de longitud. En un aspecto preferente, el índice de identificación es lo suficientemente largo como para tener una probabilidad estadística de etiquetar cada molécula con una secuencia objetivo de manera única. Por ejemplo, si hay 3000 copias de una secuencia objetivo particular, hay sustancialmente más de 3000 índices de identificación, de modo que es probable que cada copia de una secuencia objetivo particular se etiquete con un índice de identificación único. El índice de identificación puede contener nucleótidos adicionales que permiten la identificación y corrección de errores de secuenciación, incluyendo la detección de deleción, sustitución, o inserción de una o más bases durante la secuenciación, así como los cambios de nucleótidos que pueden ocurrir fuera de la secuenciación, tal como síntesis

de oligonucleótidos, amplificación y cualquier otro aspecto del ensayo. El índice se puede combinar con cualquier otro índice para crear un índice que proporcione información para dos propiedades (por ejemplo, índice de identificación de muestra, índice de identificación de locus).

- 5 El término "probabilidad" se refiere a cualquier valor logrado calculando directamente la probabilidad o cualquier valor que se pueda correlacionar con o de otra manera indicativo de un riesgo.

Los términos "locus" y "loci", tal como se usan en el presente documento, se refieren a una región de ácido nucleico de localización conocida en un genoma.

- 10 El término "índice de locus" se refiere en general a una serie de nucleótidos que corresponden a un locus conocido en un cromosoma. En general, el índice de locus es suficientemente largo como para etiquetar cada región de locus conocida de forma única. Por ejemplo, si el procedimiento usa 192 regiones de locus conocidas correspondientes a 192 secuencias individuales asociadas con los loci conocidos, hay al menos 192 índices de locus únicos, cada uno de los cuales identifica de manera única una región indicativa de un locus particular en un cromosoma. Los índices de locus usados en los procedimientos de la invención pueden ser indicativos de diferentes loci en un solo cromosoma, así como de loci conocidos presentes en diferentes cromosomas dentro de una muestra. El índice de locus puede contener nucleótidos adicionales que permiten la identificación y corrección de errores de secuenciación, incluyendo la detección de deleción, sustitución, o inserción de una o más bases durante la secuenciación, así como los cambios de nucleótidos que pueden ocurrir fuera de la secuenciación, tal como síntesis de oligonucleótidos, amplificación y cualquier otro aspecto del ensayo.

- 25 El término "muestra materna", como se usa en el presente documento, se refiere a cualquier muestra tomada de un mamífero gestante que comprende material genómico libre circulante tanto fetal como materno (por ejemplo, ADN). Preferentemente, las muestras maternas para uso en la invención se obtienen a través de medios relativamente no invasivos, por ejemplo, flebotomía u otras técnicas estándar para extraer muestras periféricas de un sujeto.

- 30 El término "temperatura de fusión" o  $T_m$  se define comúnmente como la temperatura a la que una población de moléculas de ácido nucleico bicatenarias se disocia en hebras simples. La ecuación para calcular la  $T_m$  de los ácidos nucleicos es bien conocida en la técnica. Como lo indican las referencias estándar, se puede calcular una estimación simple del valor de  $T_m$  mediante la ecuación:  $T_m = 81,5 + 16,6 (\log_{10}[\text{Na}^+]) + 0,41 (\% [\text{G}+\text{C}]) - 675 / n - 1,0 m$ , cuando un ácido nucleico está en solución acuosa con concentraciones de cationes de 0,5 M o menos, el contenido de (G+C) es de entre un 30 % y un 70 %, n es el número de bases y m es el porcentaje de emparejamientos erróneos de pares de bases (véase, por ejemplo, Sambrook *J et al.*, *Molecular Cloning, A Laboratory Manual*, 3ª ed., Cold Spring Harbor Laboratory Press (2001)). Otras referencias incluyen cálculos más sofisticados, que tienen en cuenta las características estructurales y de secuencia para el cálculo de  $T_m$ .

- 40 "Micromatriz" o "matriz" se refiere a un soporte de fase sólida que tiene una superficie, preferentemente pero no exclusivamente una superficie plana o sustancialmente plana, que lleva una matriz de sitios que contienen ácidos nucleicos de modo que cada sitio de la matriz comprende copias idénticas o sustancialmente idénticas de oligonucleótidos o polinucleótidos y se define espacialmente y sin superposición con otros sitios miembros de la matriz; es decir, los sitios son espacialmente discretos. La matriz o micromatriz también puede comprender una estructura examinable no plana con una superficie tal como una microesfera o un pocillo. Los oligonucleótidos o polinucleótidos de la matriz se pueden unir covalentemente al soporte sólido, o se pueden unir no covalentemente. La tecnología de micromatrices convencional se revisa, por ejemplo, en Schena, Ed., *Microarrays: A Practical Approach*, IRL Press, Oxford (2000). "Análisis de matriz", "análisis por matriz" o "análisis por micromatriz" se refiere a análisis, tal como por ejemplo, aislamiento de ácidos nucleicos específicos o análisis de secuencia de una o más moléculas biológicas usando una micromatriz.

- 50 Por "no polimórfico", cuando se usa con respecto a la detección de regiones de ácido nucleico seleccionadas, se entiende una detección de dicha región de ácido nucleico, que puede contener uno o más polimorfismos, pero en la que la detección no depende de la detección del polimorfismo específico dentro de la región. Por tanto, una región de ácido nucleico seleccionada puede contener un polimorfismo, pero la detección de la región usando el sistema de ensayo de la invención se basa en la ocurrencia de la región en lugar de en la presencia o ausencia de un polimorfismo particular en esa región.

- 60 Como se usa en el presente documento, "nucleótido" se refiere a una combinación base-azúcar-fosfato. Los nucleótidos son unidades monoméricas de una secuencia de ácido nucleico (ADN y ARN). El término nucleótido incluye trifosfatos de ribonucleósido ATP, UTP, CTG, GTP y trifosfatos de desoxirribonucleósido tales como dATP, dCTP, dITP, dUTP, dGTP, dTTP o derivados de los mismos. Dichos derivados incluyen, por ejemplo, [ $\alpha$ S] dATP, 7-deaza-dGTP y 7-deaza-dATP, y derivados de nucleótidos que confieren resistencia a nucleasa en la molécula de ácido nucleico que los contiene. El término nucleótido, como se usa en el presente documento, también se refiere a trifosfatos de didesoxirribonucleósido (ddNTP) y sus derivados. Ejemplos ilustrados de trifosfatos de didesoxirribonucleósido incluyen, pero no se limitan a, ddATP, ddCTP, ddGTP, ddITP y ddTTP.

65

De acuerdo con la presente invención, un "nucleótido" puede estar no marcado o puede estar marcado de forma detectable mediante técnicas bien conocidas. Los marcadores fluorescentes y su unión a los oligonucleótidos se describen en numerosas revisiones, incluyendo Haugland, Handbook of Fluorescent Probes and Research Chemicals, 9ª ed., Molecular Probes, Inc., Eugene OR (2002); Keller y Manak, DNA Probes, 2ª ed., Stockton Press, Nueva York (1993); Eckstein, Ed., Oligonucleotides and Analogues: A Practical Approach, IRL Press, Oxford (1991); Wetmur, Critical Reviews in Biochemistry and Molecular Biology, 26:227-259 (1991) y similares. Otras metodologías aplicables a la invención se divulgan en la siguiente muestra de referencias: Fung *et al.*, patente de EE. UU. n.º 4.757.141; Hobbs, Jr., *et al.*, patente de EE. UU. n.º 5.151.507; Cruickshank, patente de EE. UU. n.º 5.091.519; Menchen *et al.*, patente de EE. UU. n.º 5.188.934; Begot *et al.*, patente de EE. UU. n.º 5.366.860; Lee *et al.*, patente de EE. UU. n.º 5.847.162; Khanna *et al.*, patente de EE. UU. n.º 4.318.846; Lee *et al.*, patente de EE. UU. n.º 5.800.996; Lee *et al.*, patente de EE. UU. n.º 5.066.580; Mathies *et al.*, patente de EE. UU. n.º 5.688.648; y similares. El marcado también se puede realizar con puntos cuánticos, como se divulga en las siguientes patentes y publicaciones de patentes: patentes de EE. UU. n.º 6.322.901; 6.576.291; 6.423.551; 6.251.303; 6.319.426; 6.426.513; 6.444.143; 5.990.479; 6.207.392; 2002/0045045 y 2003/0017264. Los marcadores detectables incluyen, por ejemplo, isótopos radiactivos, marcadores fluorescentes, marcadores quimioluminiscentes, marcadores bioluminiscentes y marcadores enzimáticos. Los marcadores fluorescentes de nucleótidos pueden incluir, pero no se limitan a, fluoresceína, 5-carboxifluoresceína (FAM), 2',7'-dimetoxi-4',5'-dicloro-6-carboxifluoresceína (JOE), rodamina, 6-carboxirrodamina (R6G), N,N,N',N'-tetrametil-6-carboxirrodamina (TAMRA), 6-carboxi-X-rodamina (ROX), ácido 4-(4'-dimetilaminofenilazo)benzoico (DABCYL), Cascade Blue, Oregon Green, Texas Red, cianina y ácido 5-(2'-aminoetil)aminonaftaleno-1-sulfónico (EDANS). Ejemplos específicos de nucleótidos marcados fluorescentemente incluyen [R6G]dUTP, [TAMRA]dUTP, [R110]dCTP, [R6G]dCTP, [TAMRA]dCTP, [JOE]ddATP, [R6G]ddATP, [FAM]ddCTP, [R110]ddCTP, [TAMRA]ddGTP, [ROX]ddTTP, [dR6G]ddATP, [dR110]ddCTP, [dTAMRA]ddGTP y [dROX]ddTTP disponibles de Perkin Elmer, Foster City, Calif.; FluoroLink desoxinucleótidos, FluoroLink Cy3-dCTP, FluoroLink Cy5-dCTP, FluoroLink FluorX-dCTP, FluoroLink Cy3-dUTP y FluoroLink Cy5-dUTP disponibles de Amersham, Arlington Heights, IL; fluoresceína-15-dATP, fluoresceína-12-dUTP, tetrametilrodamina-6-dUTP, IR770-9-dATP, fluoresceína-12-ddUTP, fluoresceína-12-UTP y fluoresceína-15-2'-dATP disponibles de Boehringer Mannheim, Indianapolis, IN; y nucleótidos marcados con cromosomas, BODIPY-FL-14-UTP, BODIPY-FL-4-UTP, BODIPY-TMR-14-UTP, BODIPY-TMR-14-dUTP, BODIPY-TR-14-UTP, BODIPY-TR-14-dUTP, Cascade Blue-7-UTP, Cascade Blue-7-dUTP, fluoresceína-12-UTP, fluoresceína-12-dUTP, Oregon Green 488-5-dUTP, rodamina Green-5-UTP, rodamina Green-5-dUTP, tetrametilrodamina-6-UTP, tetrametilrodamina-6-dUTP, Texas Red-5-UTP, Texas Red-5-dUTP y Texas Red-12-dUTP disponibles de Molecular Probes, Eugene, OR.

Los términos "oligonucleótidos" u "oligos", como se usan en el presente documento, se refieren a oligómeros lineales de monómeros de ácidos nucleicos naturales o modificados, incluyendo desoxirribonucleótidos, ribonucleótidos, formas anoméricas de los mismos, monómeros de ácidos nucleicos peptídicos (PNA), monómeros de ácidos nucleótidos bloqueados (LNA) y similares, o una combinación de los mismos, capaces de unirse específicamente a un polinucleótido monocatenario por medio de un patrón regular de interacciones monómero-monómero, tales como emparejamiento de bases de tipo Watson-Crick, apilamiento de bases, emparejamiento de bases de tipo Hoogsteen o Hoogsteen inversa, o similares. Normalmente, los monómeros se unen mediante enlaces fosfodiéster o análogos de los mismos para formar oligonucleótidos que varían en tamaño desde unas pocas unidades monoméricas, por ejemplo, 8-12, a varias decenas de unidades monoméricas, por ejemplo, 100-200 o más. Las moléculas de ácido nucleico adecuadas se pueden preparar mediante el procedimiento de fosforamidita descrito por Beaucage y Carruthers (Tetrahedron Lett., 22:1859-1862 (1981)), o mediante el procedimiento del triéster de acuerdo con Matteucci *et al.* (J. Am. Chem. Soc., 103:3185 (1981)) o mediante otros procedimientos químicos, tales como el uso de un sintetizador de oligonucleótidos automatizado comercial.

Como se usa en el presente documento, el término "polimerasa" se refiere a una enzima que une los nucleótidos individuales juntos en una cadena larga, usando otra cadena como molde. Hay dos tipos generales de polimerasas; ADN polimerasas que sintetizan ADN y ARN polimerasas que sintetizan ARN. Dentro de estas dos clases hay numerosos subtipos de polimerasas, dependiendo de qué tipo de ácido nucleico pueda funcionar como molde y qué tipo de ácido nucleico se forme.

Como se usa en el presente documento, "reacción en cadena de la polimerasa" o "PCR" se refiere a una técnica para la replicación de una pieza específica de ADN objetivo *in vitro*, incluso en presencia de un exceso de ADN no específico. Los cebadores se agregan al ADN objetivo, donde los cebadores inician la copia del ADN objetivo usando nucleótidos y, típicamente, Taq polimerasa o similares. Al ciclar la temperatura, el ADN objetivo se desnaturaliza y se copia repetidamente. Una sola copia del ADN objetivo, incluso si se mezcla con otro ADN aleatorio, se puede amplificar para obtener miles de millones de réplicas. La reacción en cadena de la polimerasa se puede usar para detectar y medir cantidades muy pequeñas de ADN y para crear piezas de ADN personalizadas. En algunos casos, los procedimientos de amplificación lineal se pueden usar como una alternativa a la PCR.

El término "polimorfismo", como se usa en el presente documento, se refiere a cualquier cambio genético en un locus que puede ser indicativo de ese locus particular, incluyendo, pero no limitado a polimorfismos de un solo nucleótido (SNP), diferencias de metilación, repeticiones cortas en tándem (STR) y similares.

65



En general, un "cebador" es un oligonucleótido usado para, por ejemplo, cebar la extensión, ligación y/o síntesis de ADN, tal como en la etapa de síntesis de la reacción en cadena de la polimerasa o en las técnicas de extensión de cebadores usadas en determinadas reacciones de secuenciación. Un cebador también se puede usar en técnicas de hibridación como un medio para proporcionar complementariedad de una región de ácido nucleico a un oligonucleótido de captura para la detección de una región de ácido nucleico específica.

El término "herramienta de investigación", como se usa en el presente documento, se refiere a cualquier composición o ensayo de la invención usado para la investigación científica, académica o comercial en la naturaleza, incluyendo el desarrollo de terapias farmacéuticas y/o biológicas. Las herramientas de investigación de la invención no están destinadas a ser terapéuticas o estar sujetas a aprobación reglamentaria; en su lugar, las herramientas de investigación de la invención están destinadas a facilitar la investigación y ayudar en dichas actividades de desarrollo, incluidas las actividades realizadas con la intención de producir información para respaldar una presentación reglamentaria.

El término "índice de muestra" se refiere en general a una serie de nucleótidos únicos (es decir, cada índice de muestra es único para una muestra en un sistema de ensayo multiplexado para el análisis de múltiples muestras). Por tanto, el índice de muestra se puede usar para ayudar en la identificación de la región de ácido nucleico para multiplexar diferentes muestras en un solo recipiente de reacción, de modo que cada muestra se pueda identificar en función de su índice de muestra. En un aspecto preferente, hay un índice de muestra único para cada muestra de un conjunto de muestras, y las muestras se combinan durante la secuenciación. Por ejemplo, si se combinan doce muestras en una única reacción de secuenciación, hay al menos doce índices de muestra únicos, de modo que cada muestra se etiqueta de forma única. El índice se puede combinar con cualquier otro índice para crear un índice que proporcione información para dos propiedades (por ejemplo, índice de identificación de muestra, índice de locus de muestra).

El término "región de ácido nucleico seleccionada", como se usa en el presente documento, se refiere a una región de ácido nucleico correspondiente a un cromosoma individual. Dichas regiones de ácido nucleico seleccionadas se pueden aislar directamente y se pueden enriquecer a partir de la muestra para detección, por ejemplo, en base a técnicas de hibridación y/u otras técnicas basadas en secuencias, o se pueden amplificar usando la muestra como un molde antes de la detección de la secuencia. Las regiones de ácidos nucleicos para uso en los sistemas de ensayo de la presente invención se pueden seleccionar en base a la variación del nivel de ADN entre individuos, en base a la especificidad para un cromosoma particular, en base al contenido de CG y/o a las condiciones de amplificación requeridas de las regiones de ácido nucleico seleccionadas, u otras características que serán evidentes para un experto en la técnica al leer la presente divulgación.

El término "amplificación selectiva", "amplificar selectivamente" y similares se refiere a un procedimiento de amplificación que depende en su totalidad o en parte de la hibridación de un oligo a una secuencia en una región genómica seleccionada. En determinadas amplificaciones selectivas, los cebadores usados para la amplificación son complementarios a una región genómica seleccionada. En otras amplificaciones selectivas, los cebadores usados para la amplificación son cebadores universales, pero solo dan como resultado un producto si una región del ácido nucleico usado para la amplificación es complementaria a una región genómica de interés.

Los términos "secuenciación", "determinación de la secuencia" y similares, como se usan en el presente documento, se refieren en general a cualquiera y a todos los procedimientos bioquímicos que se pueden usar para determinar el orden de bases de nucleótidos en un ácido nucleico.

El término "se une específicamente", "unión específica" y similares, como se usa en el presente documento, cuando se refiere a un sitio de unión (por ejemplo, una sonda de ácido nucleico o cebador, anticuerpo, etc.) que da como resultado la generación de una señal positiva estadísticamente significativa en las condiciones de ensayo designadas. Típicamente, la interacción dará como resultado una señal detectable que es al menos el doble de la desviación estándar de cualquier señal generada como resultado de interacciones no deseadas (fondo).

El término "universal", cuando se usa para describir un procedimiento de amplificación, se refiere al uso de un único cebador o un conjunto de cebadores para una pluralidad de reacciones de amplificación. Por ejemplo, en la detección de 96 secuencias objetivo diferentes, todos los moldes pueden compartir las secuencias de cebado universales idénticas, lo que permite la amplificación múltiple de las 96 secuencias diferentes usando un único conjunto de cebadores. El uso de dichos cebadores simplifica enormemente la multiplexación, ya que solo se necesitan dos cebadores para amplificar una pluralidad de secuencias de ácidos nucleicos seleccionadas. El término "universal" cuando se usa para describir un sitio de cebado es un sitio con el que hibridará un cebador universal.

Cabría destacar también que se pueden usar "conjuntos" de secuencias de cebado / cebadores universales. Por ejemplo, en reacciones altamente multiplexadas, puede ser útil usar varios conjuntos de secuencias universales, en lugar de un solo conjunto; por ejemplo, 96 ácidos nucleicos diferentes pueden tener un primer conjunto de secuencias de cebado universales, y los segundos 96 pueden tener un conjunto diferente de secuencias de cebado universales, etc.

## La invención en general

La presente invención se define en las reivindicaciones y permite procedimientos mejorados para la identificación de variantes de número de copias de regiones genómicas particulares, incluyendo cromosomas completos (por ejemplo, aneuploidías), en muestras mezcladas que comprenden ADN libre circulante de una mujer embarazada con un embarazo por ovodonación. Estos procedimientos son útiles para cualquier muestra mezclada de un individuo que comprende material genómico libre circulante (por ejemplo, ADN) de dos o más individuos distintos, por ejemplo, muestras mezcladas que comprenden ADN libre circulante materno y fetal, muestras mezcladas que comprenden ADN libre circulante de un donante y un receptor de trasplante, y similares.

En determinados aspectos, el porcentaje de contribución fetal en una muestra materna de un embarazo por ovodonación se estima usando los procedimientos de ensayo de la presente invención. En determinados aspectos específicos y como se describe en el presente documento, la determinación del porcentaje de contribución fetal del ADN libre circulante en una muestra materna se puede usar para incrementar la precisión de los cálculos de frecuencia para las regiones de ácido nucleico seleccionadas usadas para determinar la probabilidad de la presencia o ausencia de aneuploidía fetal.

En determinados aspectos, los procedimientos de ensayo de la invención proporcionan un enriquecimiento seleccionado de regiones de ácido nucleico de cromosomas de interés y/o cromosomas de referencia para la detección de variantes de número de copias. Una clara ventaja de la invención es que las regiones de ácido nucleico seleccionadas se pueden analizar adicionalmente usando una variedad de técnicas de detección y cuantificación, que incluyen, pero no se limitan a, técnicas de hibridación, PCR digital y técnicas de determinación de secuenciación de alto rendimiento. Las sondas de selección se pueden diseñar para un número cualquiera de regiones de ácido nucleico para cualquier cromosoma. Aunque la amplificación anterior a la identificación y cuantificación de las regiones de ácidos nucleicos seleccionadas no es obligatoria, es preferente una amplificación limitada antes de la detección.

#### **Determinación del contenido de ADN fetal en una muestra materna**

Se divulgan procedimientos para identificar aneuploidías cromosómicas fetales en muestras maternas que comprenden tanto ADN libre circulante materno de una madre portadora como ADN libre circulante fetal en un embarazo por ovodonación y sistemas de ensayo para dicha identificación. En determinados aspectos, determinar el porcentaje relativo de ADN fetal en una muestra materna puede ser beneficioso al realizar el sistema de ensayo, ya que proporcionará información importante sobre la presencia estadística esperada de regiones genómicas y la variación con respecto a esa expectativa puede ser indicativa de una variación en el número de copias asociada a inserciones, deleciones o aneuploidía. Esto puede ser especialmente útil en circunstancias en las que el nivel de ADN fetal en una muestra materna sea bajo, ya que el porcentaje de contribución fetal se puede usar para determinar la significación estadística cuantitativa en las variaciones de los niveles de regiones de ácido nucleico identificadas en una muestra materna. En otros aspectos, la determinación del porcentaje relativo de ADN libre circulante fetal en una muestra materna puede ser beneficiosa para estimar el nivel de certeza o capacidad en la detección de una aneuploidía fetal. La estimación inexacta del porcentaje de contribución de ADN libre circulante fetal puede causar una determinación de riesgo inexacta de la presencia o ausencia de aneuploidía fetal, lo que lleva a un resultado falso positivo o falso negativo.

Los procedimientos de la invención incluyen el examen de loci informativos de la donante de óvulos en las regiones de ácidos nucleicos seleccionados que se detectan a continuación para determinar una frecuencia relativa de alelos para identificar loci informativos de la donante de óvulos tales como alelos de baja frecuencia y alelos de alta frecuencia. Los alelos de alta frecuencia son alelos maternos, estén o no presentes en el genoma fetal, ya que los alelos maternos son naturalmente más abundantes en la muestra materna. Los alelos de baja frecuencia son alelos fetales en los que al menos una copia del alelo está ausente en el genoma materno. La frecuencia de los alelos de baja frecuencia y la frecuencia de los alelos de alta frecuencia se usan para calcular el porcentaje de ADN libre circulante fetal en la muestra, por ejemplo, las sumas de los alelos de baja frecuencia se pueden dividir por las sumas de los alelos de alta y baja frecuencia.

La FIG. 1 es un diagrama de flujo simplificado de las etapas generales utilizadas en la determinación del porcentaje de contribución fetal de ADN libre circulante en una muestra materna de un embarazo por ovodonación. La FIG. 1 muestra el procedimiento 100, en el que en una primera etapa 101 se proporciona una muestra materna de una mujer embarazada con un embarazo por ovodonación que comprende ADN libre circulante materno y fetal. La muestra materna puede estar en forma de sangre total, plasma o suero. Opcionalmente, el ADN libre circulante se aísla de la muestra antes de un análisis adicional. En la etapa 103, dos o más regiones de ácido nucleico polimórficas se examinan en el ADN libre circulante de la muestra. El examen puede comprender la secuenciación de ácidos nucleicos polimórficos seleccionados seguida de una amplificación selectiva de esas regiones de ácido nucleico, incrementando el número de copias de las regiones de ácido nucleico seleccionadas de modo que permita la conservación de la cantidad relativa de las regiones de ácido nucleico seleccionadas en la muestra inicial. Opcionalmente, las regiones de ácido nucleico se pueden amplificar universalmente a continuación, por ejemplo, a través del uso de secuencias de cebadores universales. La amplificación universal se puede realizar antes o durante una etapa de determinación. En la etapa 105, se detectan las regiones de ácido nucleico examinadas. La detección puede comprender procedimientos de detección tales como determinación de secuencia o hibridación. La detección de secuencias de

ácido nucleico seleccionadas se puede llevar a cabo mediante a través de la detección de uno o más índices asociados con la región polimórfica seleccionada, por ejemplo, índices de locus o índices de alelos. En la etapa 107 se determina una frecuencia relativa de polimorfismos en las regiones de ácido nucleico examinadas para identificar los loci informativos de la donante de óvulos. En el paso 109, el porcentaje de ADN libre circulante fetal se calcula usando los loci informativos de la donante de óvulos identificados. El cálculo del porcentaje fetal puede comprender dividir la suma de los alelos de baja frecuencia por la suma de los alelos combinados de baja y alta frecuencia. Estos conceptos serán discutidos con más detalle a continuación.

Los embarazos por ovodonación se diferencian de los embarazos tradicionales en la determinación del porcentaje de contribución fetal en una muestra materna, dado que la madre portadora y el feto no comparten aproximadamente la mitad de la información genética de la madre. En una muestra materna en la que el feto es la descendencia genética de la madre, el ADN materno y el ADN fetal tienen en general al menos un alelo en común en cualquier locus en particular, aunque determinados alelos pueden diferir debido, por ejemplo, a mutaciones. Sin embargo, en un embarazo por ovodonación, el ADN materno y el ADN fetal pueden no tener ningún alelo en común en un locus en particular. Por ejemplo, tanto el genoma materno como el fetal pueden comprender alelos homocigóticos en el mismo locus, pero ambos alelos maternos pueden ser diferentes de los alelos fetales.

Los loci informativos de la donante de óvulos para uso en la presente invención son loci en los que un alelo fetal es diferente de los alelos maternos. Por ejemplo, el genoma materno puede comprender alelos homocigóticos en un locus particular, mientras que el genoma fetal comprende alelos heterocigóticos, en los que los genomas materno y fetal tienen un alelo en común. De forma alternativa, el genoma materno puede comprender alelos heterocigóticos en un locus particular y el genoma fetal puede comprender alelos homocigóticos, en los que el ADN materno y fetal tienen un alelo en común.

La FIG. 2 es una tabla que ilustra las 14 combinaciones posibles de genotipo materno, genotipo de la donante de óvulos y genotipo fetal en un embarazo por ovodonación en un locus particular, y además muestra sus clasificaciones como loci informativos versus loci no informativos. En la FIG. 2, A se usa para representar un primer tipo de alelo, mientras que B se usa para representar un tipo diferente de alelo. En la Combinación 201, por ejemplo, el genotipo materno es AA, el genotipo de la donante de óvulos es AA y el genotipo fetal es AA. En esta situación, el locus no es informativo para los propósitos de la evaluación del porcentaje de contribución fetal, ya que el ADN fetal no comprende al menos un alelo que es diferente del ADN materno. En otro ejemplo, la Combinación 204 comprende un genotipo materno AA, un genotipo de la donante de óvulos AB y un genotipo fetal AB. Este es un locus informativo de la donante de óvulos porque el ADN fetal tiene al menos un alelo que es diferente de los alelos maternos en ese locus. En otro ejemplo más, la Combinación 105 comprende un genotipo materno AA, un genotipo de la donante de óvulos AB y un genotipo fetal BB. Este es un locus informativo de la donante de óvulos, que queda fuera del alcance de la invención, ya que ambos alelos del ADN fetal son diferentes del ADN materno en ese locus.

En algunos aspectos específicos, la contribución fetal relativa de ADN materno en un alelo de interés se puede comparar con la contribución no materna en ese alelo para determinar la concentración de ADN fetal aproximada en la muestra.

***Determinación del contenido de ADN fetal en una muestra materna mediante el uso de polimorfismos autosómicos fetales y variaciones genéticas***

Como se describe anteriormente, en una muestra materna de un embarazo por ovodonación, el ADN de un feto puede no tener ningún alelo en común con la madre portadora en un locus particular. La determinación de los loci en los que el feto tiene al menos un alelo que es diferente de los alelos de la madre portadora puede permitir la estimación del porcentaje de ADN fetal en una muestra materna y, por tanto, proporcionar información usada para calcular las diferencias estadísticamente significativas en las frecuencias cromosómicas de los cromosomas de interés.

En determinados aspectos, la determinación de polimorfismos fetales requiere SNP seleccionados y/o análisis de mutaciones para identificar la presencia de ADN fetal en una muestra materna. En algunos aspectos se puede realizar el uso de genotipos anteriores del padre y la madre. Por ejemplo, los progenitores pueden haber sido sometidos a dicha determinación de genotipo para la identificación de marcadores de enfermedad, por ejemplo, la determinación del genotipo para trastornos tales como la fibrosis quística, la distrofia muscular, la atrofia muscular espinal o incluso el estado del gen RhD. Dicha diferencia en polimorfismos, variantes de número de copias o mutaciones se puede usar para determinar el porcentaje de contribución fetal en una muestra materna.

Fuera del alcance de la invención se divulga que el porcentaje de ADN libre circulante fetal en una muestra materna se puede cuantificar usando la detección de SNP multiplexada sin usar el conocimiento previo del genotipo materno o paterno. En este aspecto, se usan dos o más regiones de ácido nucleico polimórficas seleccionadas con un SNP conocido en cada región. En un aspecto preferente, las regiones de ácido nucleico polimórficas seleccionadas están localizadas en un cromosoma autosómico que es poco probable que sea aneuploide, por ejemplo, los cromosomas 1-12. Las regiones de ácido nucleico polimórficas seleccionadas de la madre se amplifican. En un aspecto preferente, la amplificación es universal.

Las regiones de ácido nucleico polimórficas seleccionadas se amplifican en una reacción en un solo recipiente. A continuación, se determina y cuantifica cada alelo de las regiones de ácido nucleico polimórficas seleccionadas en la muestra materna. En un aspecto preferente se usa la secuenciación de alto rendimiento para dicha determinación y cuantificación. Después de la determinación de la secuencia, los loci se identifican cuando los genotipos materno y fetal son diferentes, por ejemplo, el genotipo materno es homocigótico y el genotipo fetal es heterocigótico, o tanto el genotipo materno como el fetal son homocigóticos y ambos alelos fetales son diferentes de los alelos maternos. Esta identificación se puede hacer observando una alta frecuencia relativa de un alelo (> 60 %) y una baja frecuencia relativa (< 20 % y > 0,15 %) del otro alelo para una región de ácido nucleico seleccionada en particular. El uso de múltiples loci es particularmente ventajoso, ya que reduce la cantidad de variación en la medición de la abundancia de los alelos. Todos o un subconjunto de los loci que cumplen este requisito se usan para determinar la concentración fetal a través del análisis estadístico.

Fuera del alcance de la invención se divulga que la concentración fetal en un embarazo por ovodonación se determina teniendo en cuenta los loci informativos de la donante de óvulos en los que ambos alelos fetales son diferentes de los alelos maternos. La detección de ácidos nucleicos polimórficos adicionales puede dar como resultado una sobreestimación del porcentaje de contribución fetal, ya que dichos loci no se tienen en cuenta antes o después de una etapa de determinación. En el procedimiento de la invención, los alelos fetales que difieren ambos de los alelos maternos se excluyen antes o después de la determinación. Fuera del alcance de la invención se divulga que la concentración fetal se estima determinando una frecuencia relativa de alelos de baja frecuencia de dos o más loci, determinando una frecuencia relativa de alelos de alta frecuencia y comparando la frecuencia relativa de alelos de baja frecuencia con la frecuencia de alelos tanto de alta como de baja frecuencia.

En un aspecto, el porcentaje de contribución fetal se determina por análisis de un subconjunto de loci, tales como un subconjunto en el que el genotipo materno o el genotipo fetal son homocigóticos, y el otro genotipo es heterocigótico. En un aspecto específico, la concentración fetal se determina usando la frecuencia de alelos de baja frecuencia de dos o más loci seleccionados y la frecuencia de alelos de alta y baja frecuencia. En un aspecto, la concentración fetal se determina sumando los alelos de baja frecuencia de dos o más loci y dividiendo por la suma de los alelos de alta y baja frecuencia. En otro aspecto, el porcentaje de ADN libre circulante fetal se determina promediando los alelos de baja frecuencia de dos o más loci, dividiendo por el promedio de los alelos de alta y baja frecuencia y multiplicando por dos. El aislamiento del subconjunto de loci se puede hacer antes de la determinación de la frecuencia relativa de los alelos o después de la determinación de la frecuencia relativa de los alelos.

Fuera del alcance de la invención se divulga que el porcentaje de contribución fetal se determina analizando un subconjunto de loci en el que tanto el genotipo materno como el fetal son homocigóticos y ambos alelos fetales son diferentes de los alelos maternos. En un aspecto específico, la concentración fetal se determina usando la frecuencia de alelos de baja frecuencia de dos o más loci seleccionados y la frecuencia de alelos de alta y baja frecuencia. En determinados aspectos específicos, el porcentaje de ADN libre circulante fetal se determina sumando los alelos de baja frecuencia de dos o más loci y dividiendo por la suma de los alelos de alta y baja frecuencia. En otro aspecto, el porcentaje de contribución fetal se determina promediando los alelos de baja frecuencia de dos o más loci y dividiendo por el promedio de los alelos de alta y baja frecuencia. El aislamiento de un subconjunto de loci se puede hacer antes de la determinación de la frecuencia relativa de los alelos o después de la determinación de la frecuencia relativa de los alelos.

Fuera del alcance de la invención se divulga además que el porcentaje de contribución fetal se determina por análisis de todos los loci informativos de la donante de óvulos que comprenden ácidos nucleicos polimórficos. Por ejemplo, se usan todos los loci en los que el genotipo materno o fetal es homocigótico y el otro es heterocigótico y todos los loci en los que tanto el genotipo materno como el fetal son homocigóticos y ambos alelos fetales son diferentes de los alelos maternos. En un aspecto, el porcentaje de contribución fetal se determina usando una combinación de cálculos: para los loci en los que tanto el genotipo materno como el fetal son homocigóticos y ambos alelos fetales son diferentes de los alelos maternos, usando la frecuencia de alelos de baja frecuencia de dos o más loci seleccionados y la frecuencia de alelos de alta y baja frecuencia y, para los loci en los que el genotipo materno o fetal es homocigótico y el otro es heterocigótico, usando la frecuencia de alelos de baja frecuencia de dos o más loci y la frecuencia de alelos de alta y baja frecuencia.

Para muchos alelos, las secuencias materna y fetal pueden ser homocigóticas e idénticas y, dado que esta información no distingue entre el ADN materno y el fetal, no es útil en la determinación del porcentaje de ADN fetal en una muestra materna. La presente invención utiliza información de los alelos cuando existe una diferencia distinguible entre el ADN fetal y el materno (es decir, un alelo fetal que contiene al menos un alelo que difiere del alelo materno) en los cálculos del porcentaje fetal. Los datos pertenecientes a regiones alélicas que son iguales para el ADN materno y fetal, por tanto, no se seleccionan para el análisis, o se eliminan de los datos pertinentes antes de la determinación del porcentaje de ADN fetal para no saturar los datos útiles.

Procedimientos ejemplares para cuantificar el ADN fetal en plasma materno se pueden encontrar, por ejemplo, en Chu *et al.*, Prenat Diagn 2010; 30:1226-1229.

En un aspecto, las regiones de ácido nucleico seleccionadas pueden ser excluidas si la cantidad o la frecuencia de la región parece ser un valor atípico debido a un error experimental, o por sesgo genético idiopático dentro de una muestra particular. En otro aspecto, los ácidos nucleicos seleccionados se pueden someter a ajustes estadísticos o matemáticos, tales como normalización, estandarización, agrupación o transformación antes de la determinación de la frecuencia de alelos de baja y alta frecuencia. En otro aspecto, los ácidos nucleicos seleccionados se pueden someter tanto a normalización como a exclusión de datos de errores experimentales antes de la determinación de la frecuencia de alelos de baja y alta frecuencia.

En un aspecto preferente se usan 12 o más loci para el análisis. En otro aspecto preferente se usan 24 o más loci para el análisis. En otro aspecto preferente se usan 48 o más loci para el análisis. En otro aspecto se usan uno o más índices para identificar la muestra, el locus, el alelo o el ácido nucleico.

En un aspecto preferente, el porcentaje de contribución fetal en una muestra materna se puede cuantificar usando la detección de SNP en tándem en los alelos maternos y fetales. Las técnicas para identificar SNP en tándem en ADN extraído de una muestra materna se divulgan en Mitchell *et al.*, patente de EE. UU. n.º 7.799.531 y en las solicitudes de patente de EE. UU. n.º 2010/184043, 2010/184044, 2011/117548 y 2011/059451. Estas describen la diferenciación de los loci fetales y maternos a través de la detección de al menos un polimorfismo de un solo nucleótido (SNP) en tándem en una muestra materna que tiene un haplotipo diferente entre el genoma fetal y el genoma materno. La identificación y cuantificación de estos haplotipos se puede realizar directamente en la muestra materna, como se describe en la divulgación de Mitchell *et al.*, y se usa para determinar el porcentaje de contribución fetal en la muestra materna.

#### **Detección de aneuploidías cromosómicas**

Fuera del alcance de la invención se divulgan procedimientos que proporcionan mecanismos para identificar aneuploidías cromosómicas fetales en muestras maternas de un embarazo por ovodonación que comprenden ADN materno y fetal. En determinados aspectos, la identificación de aneuploidía fetal se puede llevar a cabo mediante cualquier procedimiento adecuado, como será evidente para un experto en la técnica al leer la presente memoria descriptiva. Los procedimientos incluyen la detección polimórfica, tal como la detección de SNP de ácidos nucleicos específicos o, preferentemente, la detección no polimórfica basada en secuencias fetales y maternas y, preferentemente, secuencias no polimórficas conservadas entre la madre y el feto. Procedimientos ejemplares son los siguientes:

En algunos aspectos, los ácidos nucleicos se pueden seleccionar a partir de una muestra materna antes de la detección, es decir, aislar selectivamente de una muestra materna antes de la detección usando técnicas de amplificación o de captura, tales como la hibridación. En otro aspecto específico, los ácidos nucleicos usados en la detección de aneuploidías fetales en un embarazo por ovodonación se pueden seleccionar después de la detección, por ejemplo, filtrando los datos de frecuencia obtenidos por técnicas tales como la secuenciación paralela masiva de firma de ácidos nucleicos dentro de la muestra materna.

En algunos aspectos específicos, la detección de aneuploidías fetales en un embarazo por ovodonación emplea procedimientos de secuenciación selectivos que examinan loci específicos de cromosomas, lo que permite la secuenciación altamente multiplexada de loci seleccionados de cromosomas específicos de interés. La secuenciación selectiva de cromosomas se puede usar para analizar simultáneamente loci polimórficos y no polimórficos en una sola reacción, permitiendo la determinación tanto de aneuploidías fetales como del porcentaje de contribución de los ácidos nucleicos fetales en la muestra materna. Posteriormente, se puede emplear un novedoso cálculo de riesgos de la invención que aprovecha las estimaciones de aneuploidías fetales y contribución fetal para calcular la probabilidad de aneuploidías fetales (por ejemplo, trisomía fetal) en cada sujeto.

En un aspecto, el procedimiento utiliza el análisis de segmentos aleatorios de ADN, tales como se describe, por ejemplo, en Quake *et al.*, patentes de EE. UU. n.º 8.008.018 y 7.888.017, y Shoemaker *et al.*, para determinar la aneuploidía fetal. Brevemente, la cantidad de ácidos nucleicos dentro de una muestra mezclada, tal como una muestra materna, se puede detectar diferencialmente usando secuencias de ácido nucleico seleccionadas. Los ácidos nucleicos pueden ser ADN o ARN genómicos y preferentemente son ARNm. En el caso del ARNm, uno puede elegir secuencias objetivo correspondientes a genes que están altamente expresados en el feto. Los ácidos nucleicos en cada muestra se detectan con una o más sondas específicas de secuencia dirigidas a al menos una de las dos secuencias objetivo en los ácidos nucleicos para obtener un producto de reacción detectable. Una sonda específica para un cromosoma examinado se combina con la muestra de reacción, junto con una sonda de control específica para otro cromosoma (por ejemplo, no examinado). En la mayoría de los casos, los productos de reacción procederán de ácidos nucleicos maternos, pero una pequeña cantidad de productos de reacción procederán de ácidos nucleicos fetales. Para distinguir la variación aleatoria de los resultados fetales se llevan a cabo un gran número de reacciones y se aplican procedimientos estadísticos a los resultados. El etiquetado y la detección en el presente procedimiento se usan para distinguir la presencia o ausencia de una única secuencia objetivo, denominado "análisis digital", aunque se puede realizar con procedimientos de detección de ácidos nucleicos sensibles que distinguen entre una y más de una secuencia objetivo en una muestra discreta.

En otro ejemplo, la secuenciación masiva en paralelo de ácidos nucleicos (por ejemplo, fragmentos de ADN seleccionados al azar de la muestra) se usa para determinar la secuencia de los ácidos nucleicos en la muestra materna para determinar la frecuencia de los ácidos nucleicos seleccionados dentro de la muestra materna. Técnicas ejemplares para la realización de secuenciación masiva en paralelo de secuencias de ácido nucleico incluyen las divulgadas en las patentes de EE. UU. n.º 8.318.430, 8.296.076 y 8.195.415 en las solicitudes de patente de EE. UU. US20090029377, 20090087847, 20100112590, 20110312503, 20110319272, 20110246083, 20110318734, 20120003635, 20120003636, 20120003637, 20120190559 y 20120208708. Para la detección de una anomalía en la frecuencia cromosómica (por ejemplo, una trisomía), los ácidos nucleicos secuenciados se identifican como provenientes de un primer cromosoma, y las cantidades totales de ácidos nucleicos de al menos un primer cromosoma en la muestra materna se comparan con las cantidades totales de ácidos nucleicos de al menos un segundo cromosoma en la muestra materna. Las cantidades totales de ácido nucleico incluyen los ácidos nucleicos tanto del feto como de la madre en la muestra materna, y los ácidos nucleicos del feto no se diferencian de los maternos al determinar la frecuencia de los ácidos nucleicos correspondientes a la frecuencia del cromosoma. Cuando se presupone que un primer cromosoma es euploide, y se sospecha que el segundo cromosoma es aneuploide, se comparan los números totales de ácidos nucleicos para el primer y el segundo cromosoma para determinar la presencia o ausencia de dicha aneuploidía.

En aspectos más específicos, las muestras usadas para la secuenciación masiva en paralelo de ácidos nucleicos se enriquecen de regiones polimórficas. Técnicas ejemplares para realizar el enriquecimiento incluyen las divulgadas e, por ejemplo, los documentos WO2011091063, WO2011091046 y la solicitud de patente de EE. UU. n.º 20110230358. Brevemente, una porción de una muestra materna que comprende ADN libre circulante se amplifica para aumentar el número de copias de la una o más secuencias polimórficas en la muestra, y las porciones amplificadas de ácidos nucleicos se vuelven a agregar a la muestra original para la secuenciación. De forma alternativa, la muestra se somete a una secuenciación del genoma completo para obtener una pluralidad de etiquetas de secuencia, y las secuencias de las etiquetas se comparan con la secuencia de múltiples polimorfismos de referencia.

En algunos aspectos, los ácidos nucleicos se secuencian usando procedimientos de hibridación basados en matrices, tales como los descritos en la publicación de patente de EE. UU. n.º 2011/0172111. En otros aspectos, las biomoléculas se detectan usando la tecnología de detección de nanoporos, tales como las descritas en la publicación de patente de EE. UU. n.º 2011/0124518.

En otro aspecto, los ácidos nucleicos se secuencian y se comparan usando polimorfismos que diferencian entre alelos maternos y fetales en una muestra, usando procedimientos tales como los descritos en las patentes de EE. UU. n.º 7.727.720, 7.718.370, 7.598.060, 7.442.506, 7.332.277, 7.208.274 y 6.977.162. Brevemente, los procedimientos utilizan detección polimórfica para identificar anomalías cromosómicas. Las secuencias se determinan en alelos que son homocigóticos en la madre y heterocigóticos en el feto, y se determina una proporción para los alelos heterocigóticos. La proporción para los alelos heterocigóticos se usa para indicar la presencia o ausencia de una anomalía cromosómica.

En otro aspecto más, la determinación de aneuploidías fetales utiliza la identificación de polimorfismos en tándem, tales como la descrita, por ejemplo, en la patente de EE. UU. n.º 7.799.531 y las publicaciones de EE. UU. n.º 2011/0117548, 2011/0059451, 2010/0184044, 2010/184043 y 2008/0020390. En resumen, los SNP en tándem se detectan y usan para diferenciar los alelos maternos y fetales en una muestra materna para detectar anomalías cromosómicas fetales a través de la comparación del ADN materno con el ADN fetal.

En un aspecto preferente, la determinación de aneuploidía fetal utiliza la amplificación seleccionada de loci representativos. Dichas técnicas se divulgan, por ejemplo, en las solicitudes de patente de EE. UU. n.º 2012/034603, 2012/034685, 2012/040859 y 2013/040375. Estas técnicas utilizan la detección de regiones genómicas usando oligonucleótidos de secuencia fija y uniendo los oligonucleótidos de secuencia fija mediante ligación y/o extensión. Esto se puede lograr usando una combinación de ligación y amplificación, por ejemplo, la ligación de dos o más oligonucleótidos de secuencia fija y opcionalmente un oligonucleótido puente que es complementario a una región entre los oligonucleótidos de secuencia fija. En otro ejemplo, esto se puede lograr usando una combinación de extensión, ligación y amplificación.

En algunos aspectos se determinan variaciones para la población normal a partir de muestras normales que tienen una proporción similar de ADN fetal. Por ejemplo, una dosis cromosómica esperada para la trisomía en una muestra de ADN con un porcentaje específico de ADN libre circulante fetal se puede calcular agregando el porcentaje de contribución del cromosoma aneuploide. La dosis cromosómica para la muestra se puede comparar con la dosis cromosómica para un feto normal y con la dosis cromosómica esperada si es triploide para determinar estadísticamente, usando la variación de la dosis cromosómica, si es más probable que la muestra sea normal o triploide, y la probabilidad estadística de que sea una o la otra.

En un aspecto preferente, las regiones de ácido nucleico seleccionadas para el análisis en la muestra materna incluyen en una única reacción ambas regiones de ácido nucleico para la determinación del porcentaje de contribución fetal, así como regiones de ácido nucleico que corresponden a dos o más cromosomas usados para detectar una anomalía en la dosis cromosómica. El uso de una sola reacción ayuda a minimizar el riesgo de contaminación o sesgo que se

pueden introducir si se usan separadas, que de otro modo pueden distorsionar los resultados. De hecho, los procedimientos se realizan preferentemente como reacciones multiplexadas o incluso altamente multiplexadas, en las que los loci polimórficos y no polimórficos (para determinar el porcentaje de contribución fetal y la aneuploidía fetal, respectivamente) se examinan en una sola reacción para cada muestra. En modos de realización preferentes se usan los ensayos de multiplexación descritos en las solicitudes de patente de EE. UU. 2012/034603, 2012/034685, 2012/040859 y 2013/040375, ya que estos ensayos evalúan loci polimórficos y no polimórficos en una muestra materna en una sola reacción multiplexada.

En otros aspectos, una o más regiones de ácido nucleico seleccionadas pueden examinar tanto para la determinación de la proporción de ácido nucleico fetal como la detección de anomalías cromosómicas fetales. El uso de las mismas regiones tanto para el porcentaje de contribución fetal como para la detección de aneuploidías fetales ayuda aún más a minimizar el sesgo debido a un error experimental o a contaminación.

### Procedimientos de amplificación

Numerosos procedimientos de amplificación se pueden usar para proporcionar los ácidos nucleicos amplificados que se analizan en los sistemas de ensayo y dichos procedimientos se usan preferentemente para incrementar el número de copias de una región de ácido nucleico de interés en una muestra materna de un embarazo por ovodonación de una manera que permite la conservación de la información relativa al contenido inicial de la región de ácido nucleico en la muestra materna. Aunque en el presente documento no se describen en detalle todas las combinaciones de amplificación y análisis, está dentro de la experiencia de los expertos en la técnica utilizar diferentes procedimientos de amplificación y/o herramientas analíticas para aislar y/o analizar los ácidos nucleicos de la región compatibles con esta memoria descriptiva, y dichas variaciones serán evidentes para un experto en la materia al leer la presente divulgación.

Dichos procedimientos de amplificación incluyen, pero no se limitan a, reacción en cadena de la polimerasa (PCR) (patentes de EE. UU. n.º 4.683.195 y 4.683.202; PCR Technology: Principles and Applications for DNA Amplification, ed. H. A. Erlich, Freeman Press, NY, N.Y., 1992), reacción en cadena de la ligasa (LCR) (Wu y Wallace, Genomics 4:560, 1989; Landegren *et al.*, Science 241:1077, 1988), amplificación por desplazamiento de hebra (SDA) (patentes de EE. UU. n.º 5.270.184 y 5.422.252), amplificación mediada por transcripción (TMA) (patente de EE. UU. n.º 5.399.491), amplificación lineal enlazada (LLA) (patente de EE. UU. n.º 6.027.923) y similares, replicación de secuencia autosostenida (Guatelli *et al.*, Proc. Nat. Acad. Sci. USA, 87, 1874 (1990) y documento WO90/06995), amplificación selectiva de secuencias de polinucleótidos objetivo (patente de EE. UU. n.º 6.410.276), reacción en cadena de la polimerasa cebada con secuencia consenso (CP-PCR) (patente de EE. UU. n.º 4.437.975), reacción en cadena de la polimerasa cebada arbitrariamente (AP-PCR) (patentes de EE. UU. n.º 5.413.909 y 5.861.245) y amplificación de secuencia basada en ácido nucleico (NASBA). (Véanse las patentes de EE. UU. n.º 5.409.818, 5.554.517 y 6.063.603). Otros procedimientos de amplificación que se pueden usar incluyen: replicasa Qbeta, descrita en la solicitud de patente PCT n.º PCT/US87/00880, procedimientos de amplificación isotérmica tales como SDA, descritos en Walker *et al.*, 1992, Nucleic Acids Res. 20 (7):1691-6, 1992, y amplificación por círculo rodante, descrita en la patente de EE. UU. n.º 5.648.245. Otros procedimientos de amplificación que se pueden usar se describen en las patentes de EE. UU. n.º 5.242.794, 5.494.810, 4.988.617 y 6.582.938 y la publicación de EE. UU. n.º 20030143599. En algunos aspectos, el ADN se amplifica por PCR múltiple específica de locus. En un aspecto preferente, el ADN se amplifica usando la ligación con adaptador y la PCR de cebador único. Otros procedimientos disponibles de amplificación incluyen PCR equilibrada (Makrigiorgos *et al.*, (2002), Nat Biotechnol, Vol. 20, pp.936-9) y replicación de secuencia autosostenida (Guatelli *et al.*, Proc. Natl. Acad. Sci USA 87: 1874, 1990). Basándose en dichas metodologías, un experto en la técnica puede diseñar fácilmente cebadores en cualquier región adecuada 5' y 3' para una región de ácido nucleico de interés. Dichos cebadores se pueden usar para amplificar ADN de cualquier longitud, siempre que contenga la región de ácido nucleico de interés en su secuencia.

La longitud de un ácido nucleico seleccionado amplificado a partir de una región genómica de interés es en general suficientemente larga para proporcionar suficiente información de la secuencia para distinguirla de otros ácidos nucleicos que se amplifican y/o seleccionadas. En general, un ácido nucleico amplificado tiene una longitud de al menos aproximadamente 16 nucleótidos, y más típicamente, un ácido nucleico amplificado tiene una longitud de al menos aproximadamente 20 nucleótidos. En un aspecto preferente, un ácido nucleico amplificado tiene una longitud de al menos aproximadamente 30 nucleótidos. En un aspecto más preferente, un ácido nucleico amplificado tiene una longitud de al menos aproximadamente 32, 40, 45, 50 o 60 nucleótidos. En otros aspectos, un ácido nucleico amplificado puede tener una longitud de aproximadamente 100, 150 o hasta 200 nucleótidos.

En determinados aspectos, la amplificación selectiva usa una o unas pocas rondas de amplificación con pares de cebadores que comprenden ácidos nucleicos complementarios a los ácidos nucleicos seleccionados. En otros aspectos, la amplificación selectiva comprende una etapa de amplificación lineal inicial. Estos procedimientos pueden ser particularmente útiles si la cantidad inicial de ADN es bastante limitada, por ejemplo, cuando el ADN libre circulante en una muestra está disponible en cantidades limitadas. Este mecanismo incrementa la cantidad de moléculas de ADN que son representativas del contenido original de ADN, y ayuda a reducir el error de muestreo cuando se necesita una cuantificación exacta del ADN o de una fracción del ADN (por ejemplo, la contribución del ADN fetal en una muestra materna).

Por tanto, en un aspecto, se realiza un número limitado de ciclos de amplificación específica de secuencia en la muestra materna inicial que comprende ADN libre circulante. El número de ciclos es en general menor que el usado para una amplificación de PCR típica, por ejemplo, de 5 a 30 ciclos o menos. Los cebadores o sondas se pueden diseñar para amplificar segmentos o regiones genómicas específicas. Los cebadores o sondas se pueden modificar con una etiqueta terminal en el extremo 5' (por ejemplo, con biotina) o en otro lugar a lo largo del cebador o la sonda, de modo que los productos de amplificación se puedan purificar o unir a un sustrato sólido (por ejemplo, microesfera o matriz) para un mayor aislamiento o análisis. En un aspecto preferente, los cebadores se multiplexan de modo que una sola reacción produce múltiples fragmentos de ADN de diferentes regiones. Los productos de amplificación de la amplificación lineal se podrían amplificar aún más con procedimientos de PCR estándar o con amplificación lineal adicional.

Por ejemplo, el ADN libre circulante se puede aislar de sangre, plasma o suero de una mujer embarazada e incubarse con cebadores contra un conjunto determinado de regiones de ácido nucleico que corresponden a los cromosomas de interés. Preferentemente, el número de pares de cebadores usados para la amplificación inicial será 12 o más, más preferentemente 24 o más, más preferentemente 36 o más, incluso más preferentemente 48 o más, y aún más preferentemente 96 o más. Cada uno de los cebadores corresponde a una única región de ácido nucleico y está etiquetado opcionalmente para identificación y/o aislamiento. Se realiza un número limitado de ciclos, preferentemente 10 o menos. Los productos de amplificación se aíslan posteriormente, por ejemplo, cuando los cebadores se unen a una molécula de biotina, los productos de amplificación se pueden aislar mediante la unión a avidina o estreptavidina en un sustrato sólido. A continuación, los productos se someten a procedimientos bioquímicos adicionales, como una amplificación adicional con otros cebadores (por ejemplo, cebadores universales) y/o técnicas de detección, tales como determinación de secuencias e hibridación.

Las eficiencias de amplificación pueden variar entre sitios y entre ciclos, de modo que en determinados sistemas se puede usar la normalización para garantizar que los productos de la amplificación sean representativos del material de partida que contiene ácidos nucleicos. Una persona que practica el sistema de ensayo puede utilizar información de varias muestras para determinar la variación en los niveles de ácidos nucleicos, incluyendo la variación en diferentes regiones de ácido nucleico en muestras individuales y/o entre las mismas regiones de ácido nucleico en diferentes muestras después de la amplificación lineal inicial limitada. Dicha información se puede usar en la normalización para evitar la distorsión de los niveles iniciales de contenido de ADN.

### **Amplificación universal**

En aspectos preferentes de la invención, las regiones de ácido nucleico amplificado selectivamente se amplifican preferentemente después de la amplificación o enriquecimiento selectivo, ya sea antes o durante las técnicas de detección de las regiones de ácido nucleico. En otro aspecto de la invención, las regiones de ácido nucleico se amplifican selectivamente durante la técnica de detección de las regiones de ácido nucleico sin ninguna amplificación previa. En un sistema de ensayo multiplexado, esto se realiza preferentemente a través del uso de la amplificación universal de las diversas regiones de ácido nucleico que se van a analizar usando los sistemas de ensayo de la invención. Las secuencias de cebadores universales se agregan a las regiones de ácido nucleico amplificadas selectivamente, ya sea durante o después de la amplificación selectiva, de modo que se puedan amplificar aún más en una única reacción de amplificación universal. Por ejemplo, estas secuencias de cebadores universales se pueden agregar a las regiones de ácido nucleico durante el procedimiento de amplificación selectiva, es decir, los cebadores para amplificación selectiva tienen secuencias de cebadores universales que flanquean un locus. De forma alternativa, se pueden agregar adaptadores que comprenden secuencias de amplificación universal a los extremos de los ácidos nucleicos seleccionados como adaptadores después de la amplificación y el aislamiento de los ácidos nucleicos seleccionados de la muestra materna.

En un aspecto ejemplar, los ácidos nucleicos se amplifican inicialmente a partir de una muestra materna usando cebadores que comprenden una región complementaria a regiones seleccionadas de los cromosomas de interés y sitios de cebado universales. La amplificación selectiva inicial va seguida de una etapa de amplificación universal para incrementar el número de regiones de ácido nucleico para el análisis. Esta introducción de regiones de cebadores en los productos de amplificación inicial permite una amplificación universal controlada posterior de todos o de una parte de los ácidos nucleicos seleccionados antes o durante el análisis, por ejemplo, la determinación de la secuencia.

El sesgo y la variabilidad se pueden introducir durante la amplificación del ADN, tal como la observada durante la reacción en cadena de la polimerasa (PCR). En los casos en que una reacción de amplificación sea multiplexada, existe el potencial de que los loci se amplifiquen a diferentes velocidades o eficiencias. Parte de esto se puede deber a la variedad de cebadores en una reacción multiplexada, algunos con mayor eficiencia (es decir, hibridación) que otros, o algunos que funcionan mejor en condiciones experimentales específicas debido a la composición de bases. Cada conjunto de cebadores para un locus dado se puede comportar de manera diferente en base al contexto de secuencia del cebador y el molde de ADN, las condiciones del tampón y otras condiciones. Una amplificación universal de ADN para un sistema de ensayo multiplexado introducirá en general menos sesgo y variabilidad.



- En consecuencia, en un aspecto preferente, se realiza un pequeño número (por ejemplo, 1-10, preferentemente 3-5) de ciclos de amplificación selectiva o enriquecimiento de ácidos nucleicos en una reacción de mezcla multiplexada, seguido de amplificación universal usando sitios de cebado universal introducidos. El número de ciclos que usan cebadores universales variará, pero preferentemente será de al menos 10 ciclos, más preferentemente de al menos 5 ciclos, incluso más preferentemente de 20 ciclos o más. Al pasar a la amplificación universal después de uno o unos pocos ciclos de amplificación selectiva, se reduce el sesgo de tener determinados loci que se amplifican a mayor velocidad que otros.
- Opcionalmente, el sistema de ensayo incluirá una etapa entre la amplificación selectiva y la amplificación universal para eliminar cualquier exceso de ácidos nucleicos que no se hayan amplificado específicamente en la amplificación selectiva.
- Para la amplificación universal se puede usar el producto entero o una parte alícuota del producto de la amplificación selectiva. Se pueden usar las mismas o diferentes condiciones (por ejemplo, polimerasa, tampones y similares) en las etapas de amplificación, por ejemplo, para garantizar que no se introduzcan inadvertidamente sesgos y variabilidad debido a las condiciones experimentales. Además, se pueden usar variaciones en las concentraciones de cebadores para limitar eficazmente el número de ciclos de amplificación específicos de secuencia.
- En determinados aspectos, las regiones de cebadores universales de los cebadores o adaptadores usados en el sistema de ensayo se diseñan para ser compatibles con los procedimientos de ensayo multiplexados convencionales que utilizan mecanismos de cebado generales para analizar un gran número de ácidos nucleicos simultáneamente en una reacción en un solo recipiente. Dichos procedimientos de cebado "universales" permiten un análisis eficiente y de gran volumen de la cantidad de regiones de ácido nucleico presentes en una muestra materna de un embarazo por ovodonación, y permiten una cuantificación exhaustiva de la presencia de regiones de ácido nucleico dentro de dicha muestra materna para la determinación de aneuploidía.
- Ejemplos de dichos procedimientos de ensayo incluyen, pero no se limitan a, procedimientos de multiplexación usados para amplificar y/o genotipar una variedad de muestras simultáneamente, tales como las descritas en Oliphant *et al.*, patente de EE. UU. n.º 7.582.420.
- Algunos aspectos utilizan reacciones acopladas para la detección múltiple de secuencias de ácido nucleico en las que los oligonucleótidos de una fase temprana de cada procedimiento contienen secuencias que pueden ser usadas por oligonucleótidos de una fase posterior del procedimiento. Se pueden usar procedimientos ejemplares para amplificar y/o detectar ácidos nucleicos en muestras, solos o en combinación, incluyendo, pero sin limitarse a los procedimientos que se describen a continuación.
- En determinados aspectos, el sistema de ensayo utiliza una de las siguientes técnicas de amplificación selectiva y universal combinadas: (1) LDR acoplada a PCR; (2) PCR primaria acoplada a PCR secundaria acoplada a LDR; y (3) PCR primaria acoplada a PCR secundaria. Cada uno de estos aspectos tiene una aplicabilidad particular en la detección de determinadas características de ácido nucleico. Sin embargo, cada uno requiere el uso de reacciones acopladas para la detección múltiple de las diferencias de secuencia de ácido nucleico, en las que los oligonucleótidos de una fase temprana de cada procedimiento contienen secuencias que pueden ser usadas por oligonucleótidos de una fase posterior del procedimiento.
- Barany *et al.*, patentes de EE. UU. n.º 6.852.487, 6.797.470, 6.576.453, 6.534.293, 6.506.594, 6.312.892, 6.268.148, 6.054.564, 6.027.889, 5.830.711 y 5.494.810, describen el uso del ensayo de reacción en cadena de la ligasa (LCR) para la detección de secuencias específicas de nucleótidos en una variedad de muestras de ácido nucleico.
- Barany *et al.*, patentes de EE. UU. n.º 7.807.431, 7.455.965, 7.429.453, 7.364.858, 7.358.048, 7.332.285, 7.320.865, 7.312.039, 7.244.831, 7.198.894, 7.166.434, 7.097.980, 7.083.917, 7.014.994, 6.949.370, 6.852.487, 6.797.470, 6.576.453, 6.534.293, 6.506.594, 6.312.892 y 6.268.148, describen el uso de la reacción de detección de la ligasa con reacción de detección ("LDR") acoplada a reacción en cadena de la polimerasa ("PCR") para la detección de ácidos nucleicos.
- Barany *et al.*, patentes de EE. UU. n.º 7.556.924 y 6.858.412, describen el uso de sondas candado (también denominadas "sondas de precírculo" o "sondas de inversión múltiple") con reacción de detección de ligasa ("LDR") acoplada y reacción en cadena de la polimerasa ("PCR") para la detección de ácidos nucleicos.
- Barany *et al.*, patentes de EE. UU. n.º 7.807.431, 7.709.201 y 7.198.814, describen el uso de reacciones de escisión y ligación combinadas de endonucleasas para la detección de secuencias de ácido nucleico.
- Willis *et al.*, patentes de EE. UU. n.º 7.700.323 y 6.858.412, describen el uso de sondas de precírculo en la amplificación, detección y genotipado multiplexado de ácidos nucleicos.
- Ronaghi *et al.*, patente de EE. UU. n.º 7.622.281, describen técnicas de amplificación para marcar y amplificar un ácido nucleico usando un adaptador que comprende un cebador único y un código de barras.

Además de las diversas técnicas de amplificación, numerosos procedimientos de determinación de secuencia son compatibles con los sistemas de ensayo. Preferentemente, dichos procedimientos incluyen procedimientos de secuenciación de "próxima generación". Procedimientos ejemplares para la determinación de secuencias incluyen, pero no se limitan a, procedimientos basados en hibridación, tales como los divulgados en Drmanac, patentes de EE. UU. n.º 6.864.052; 6.309.824 y 6.401.267; y Drmanac *et al.*, publicación de patente de EE. UU. 2005/0191656, secuenciación por procedimientos de síntesis, por ejemplo, Nyren *et al.*, patente de EE. UU. 7.648.824, 7.459.311 y 6.210.891; Balasubramanian, patentes de EE. UU. n.º 7.232.656 y 6.833.246; Quake, patente de EE. UU. n.º 6.911.345; Li *et al.*, Proc. Natl Acad Sci., 100: 414-419 (2003); secuenciación con pirofosfato como se describe en Ronaghi *et al.*, patentes de EE. UU. n.º 7.648.824, 7.459.311, 6.828.100 y 6.210.891; y procedimientos de determinación de secuenciación basada en ligación, por ejemplo, Drmanac *et al.*, solicitud de patente de EE. UU. n.º. 20100105052, y Church *et al.*, solicitudes de patente de EE. UU. n.º. 20070207482 y 20090018024.

De forma alternativa, las regiones de ácido nucleico de interés se pueden seleccionar y/o identificar usando técnicas de hibridación. Los procedimientos para llevar a cabo ensayos de hibridación de polinucleótidos para detección han sido bien desarrollados en la técnica. Los procedimientos y condiciones del ensayo de hibridación variarán dependiendo de la aplicación y se seleccionan de acuerdo con los procedimientos de unión generales conocidos, incluyendo los mencionados en: Maniatis *et al.*, Molecular Cloning: A Laboratory Manual (2ª ed. Cold Spring Harbor, N.Y., 1989); Berger y Kimmel, Methods in Enzymology, Vol. 152, Guide to Molecular Cloning Techniques (Academic Press, Inc., San Diego, Calif., 1987); Young y Davis, P.N.A.S., 80: 1194 (1983). Los procedimientos y aparatos para llevar a cabo reacciones de hibridación controladas y repetidas se han descrito en las patentes de EE. UU. n.º 5.871.928, 5.874.219, 6.045.996, 6.386.749 y 6.391.623.

La presente invención también contempla la detección de señales de hibridación entre ligandos en determinados aspectos preferentes. Véanse las patentes de EE. UU. n.º 5.143.854, 5.578.832; 5.631.734; 5.834.758; 5.936.324; 5.981.956; 6.025.601; 6.141.096; 6.185.030; 6.201.639; 6.218.803 y 6.225.625, en la solicitud de patente de EE. UU. 60/364.731 y en la solicitud PCT PCT/US99/06097 (publicada como WO99/47964).

Los procedimientos y aparatos para la detección de señales y el procesamiento de datos de intensidad se divulgan, por ejemplo, en las patentes de EE. UU. n.º 5.143.854, 5.547.839, 5.578.832, 5.631.734, 5.800.992, 5.834.758, 5.856.092, 5.902.723, 5.936.324, 5.981.956, 6.025.601, 6.090.555, 6.141.096, 6.185.030, 6.201.639, 6.218.803 y 6.225.625, en la solicitud de patente de EE. UU. 60/364.731 y en la solicitud PCT PCT/US99/06097(publicada como WO99/47964).

### Uso de índices en los sistemas de ensayo de la invención

En determinados aspectos, la totalidad o una porción de los ácidos nucleicos de interés se detectan directamente usando las técnicas descritas, por ejemplo, determinación de secuencia o hibridación. Sin embargo, en determinados aspectos, los ácidos nucleicos de interés se asocian con uno o más índices que se identifican para una región de ácido nucleico seleccionada o una muestra particular que se está analizando. La detección de uno o más índices puede servir como un mecanismo de detección sustituto de la región de ácido nucleico seleccionada, o como confirmación de la presencia de una región de ácido nucleico seleccionada particular si se determinan tanto la secuencia del índice como la secuencia de la región de ácido nucleico. Estos índices se asocian preferentemente con los ácidos nucleicos seleccionados durante una etapa de amplificación usando cebadores que comprenden tanto el índice como las regiones de secuencia que hibridan específicamente con la región de ácido nucleico.

En un ejemplo, los cebadores usados para la amplificación selectiva de una región de ácido nucleico se diseñan para proporcionar un índice de locus entre la región complementaria a un locus de interés y un sitio de cebado de amplificación universal. El índice de locus es único para cada región de ácido nucleico seleccionada y es representativo de un locus en un cromosoma de interés o cromosoma de referencia, de modo que la cuantificación del índice de locus en una muestra proporcione datos de cuantificación para el locus y el cromosoma particular que contiene el locus.

En otro ejemplo, los cebadores usados para la amplificación de una región de ácido nucleico seleccionada se diseñan para proporcionar un índice de alelo entre la región complementaria a un locus de interés y un sitio de cebado de amplificación universal. El índice de alelo es único para alelos particulares de una región de ácido nucleico seleccionada y es representativo de una variación de locus presente en un cromosoma de interés o cromosoma de referencia, de modo que la cuantificación del índice de alelo en una muestra proporcione datos de cuantificación para el alelo y de modo que las frecuencias relativas de los índices de alelo para un locus particular proporcionen datos de cuantificación tanto para el locus como para el cromosoma particular que contiene el locus.

En otro aspecto, los cebadores usados para la amplificación de las regiones de ácido nucleico seleccionadas que se van a analizar para una muestra materna de un embarazo por ovodonación se diseñan para proporcionar un índice de identificación entre la región complementaria a un locus de interés y un sitio de cebado de amplificación universal. En dicho aspecto, está presente un número de índices de identificación suficiente para identificar de manera única cada región de ácido nucleico seleccionada en la muestra. Cada región de ácido nucleico que se va a analizar está asociada con un índice de identificación único, de modo que el índice de identificación se asocie de manera única a la

región de ácido nucleico seleccionada. La cuantificación del índice de identificación en una muestra proporciona datos de cuantificación para la región de ácido nucleico seleccionada asociada y el cromosoma correspondiente a la región de ácido nucleico seleccionada. El locus de identificación también se puede usar para detectar cualquier sesgo de amplificación que se produzca después del aislamiento inicial de las regiones de ácido nucleico seleccionadas de una muestra.

En determinados aspectos, solo el índice de locus y/o el índice de identificación (si está presente) se detectan y se usan para cuantificar las regiones de ácido nucleico seleccionadas en una muestra. En otro aspecto, se realiza un recuento del número de veces que se produce cada índice de locus con un índice de identificación único para determinar la frecuencia relativa de una región de ácido nucleico seleccionada en una muestra.

En algunos aspectos, los índices representativos de la muestra de la que se aísla un ácido nucleico se usan para identificar la fuente del ácido nucleico en un sistema de ensayo de multiplexado. En dichos aspectos, los ácidos nucleicos se identifican de forma única con el índice de muestra. Esos oligonucleótidos identificados de forma única se pueden combinar en un solo recipiente de reacción con ácidos nucleicos de otras muestras antes de la secuenciación. Los datos de secuenciación se separan primero por cada índice de muestra único antes de determinar la frecuencia de cada locus objetivo para cada muestra y antes de determinar si hay una anomalía cromosómica para cada muestra. Para la detección, se secuencian los índices de muestra, los índices de locus y los índices de identificación (si están presentes).

En aspectos que usan índices, los cebadores de amplificación selectiva se diseñan preferentemente de modo que los índices que comprenden información de identificación están codificados en uno o ambos extremos del cebador. De forma alternativa, los índices y las secuencias de amplificación universal se pueden agregar a los ácidos nucleicos amplificados selectivamente después de la amplificación selectiva inicial.

Los índices son secuencias no complementarias pero únicas usadas en el cebador para proporcionar información relevante para la región de ácido nucleico selectiva que se aísla y/o amplifica usando el cebador. La ventaja de esto es que la información sobre la presencia y la cantidad de la región de ácido nucleico seleccionada se pueden obtener sin necesidad de determinar la secuencia real en sí misma, aunque en determinados aspectos puede ser deseable hacerlo. Sin embargo, en general, la capacidad de identificar y cuantificar una región de ácido nucleico seleccionada a través de la identificación de uno o más índices disminuirá la duración de la secuenciación requerida, ya que la información de los loci se captura en el extremo 3' o 5' de la región de ácido nucleico seleccionada aislada. El uso de la identificación de índices como un sustituto para la identificación de regiones de ácido nucleico seleccionadas también puede reducir el error, ya que las lecturas de secuenciación más largas son más propensas a la introducción de errores.

Además de índices de locus, índices de alelos e índices de identificación, se pueden introducir índices adicionales en los cebadores para ayudar en la multiplexación de muestras. Por ejemplo, los índices de corrección que identifican el error experimental (por ejemplo, los errores introducidos durante la amplificación o la determinación de la secuencia) se pueden usar para identificar potenciales discrepancias en los procedimientos experimentales y/o procedimientos de detección en los sistemas de ensayo. El orden y la ubicación de estos índices, así como la longitud de estos índices, pueden variar, y se pueden usar en varias combinaciones.

Los cebadores usados para la identificación y cuantificación de una región de ácido nucleico seleccionada pueden estar asociados con regiones complementarias al extremo 5' de la región de ácido nucleico seleccionada, o en determinados regímenes de amplificación los índices pueden estar presentes en uno o ambos de un conjunto de cebadores de amplificación que comprenden secuencias complementarias a las secuencias de la región de ácido nucleico seleccionada. Los cebadores se pueden usar para multiplexar el análisis de múltiples regiones de ácido nucleico seleccionadas que se van a analizar dentro de una muestra, y se pueden usar en solución o en un sustrato sólido, por ejemplo, en una micromatriz o en una microesfera. Estos cebadores se pueden usar para amplificación o replicación lineal, o pueden crear construcciones circulares para análisis posteriores.

#### **Minimización de variaciones dentro de y entre muestras**

Uno de los retos de la detección de anomalías cromosómicas en una muestra materna de un embarazo por ovodonación es que, a menudo, el ADN del tipo de célula con la supuesta anomalía cromosómica está presente en mucha menor abundancia que el ADN de tipo celular normal. En el caso de una muestra materna mezclada que contiene ADN libre circulante fetal y materno, el ADN libre circulante fetal como porcentaje del ADN libre circulante total puede variar desde menos de un uno hasta un cuarenta por ciento, y más comúnmente está presente en o por debajo de un veinte por ciento y frecuentemente en o por debajo de un diez por ciento. En la detección de una aneuploidía tal como la trisomía 21 (síndrome de Down) en el ADN fetal de dicha muestra materna mezclada, el incremento relativo del cromosoma 21 es de un 50 % en el ADN fetal; por tanto, como porcentaje del ADN total en una muestra materna en la que, como ejemplo, el ADN fetal sea un 5 % del total, el incremento del cromosoma 21 como porcentaje del total es de un 2,5 %. Si se va a detectar esta diferencia de manera sólida a través de los procedimientos descritos en el presente documento, la variación en la medición del cromosoma 21 tiene que ser mucho menor que el incremento porcentual del cromosoma 21.

La variación entre los niveles encontrados entre las muestras y/o para regiones de ácido nucleico dentro de una muestra se puede minimizar en una combinación de procedimientos de análisis, muchos de los cuales se describen en esta solicitud. Por ejemplo, la variación se reduce usando una referencia interna en el ensayo. Un ejemplo de una referencia interna es el uso de un cromosoma presente en una abundancia "normal" (por ejemplo, disomía para un autosoma) para comparar con respecto a un cromosoma presente en una abundancia supuestamente anómala, como aneuploidía, en la misma muestra. Si bien el uso de uno de dichos cromosomas "normales" como cromosoma de referencia puede ser suficiente, también es posible usar muchos cromosomas normales como cromosomas de referencia interna para incrementar el poder estadístico de la cuantificación.

Un procedimiento para usar una referencia interna es calcular una proporción entre la abundancia de los cromosomas supuestamente anómalos y la abundancia de los cromosomas normales en una muestra, llamada proporción cromosómica. Al calcular la proporción cromosómica se cuantifican la abundancia o los recuentos de cada una de las regiones de ácido nucleico para cada cromosoma para calcular los recuentos totales para cada cromosoma. Los recuentos totales para un cromosoma se dividen a continuación por los recuentos totales para un cromosoma diferente para crear una proporción cromosómica para esos dos cromosomas.

De forma alternativa, se puede calcular una proporción cromosómica para cada cromosoma cuantificando primero la frecuencia relativa (por ejemplo, sumando) de cada una de las regiones de ácido nucleico para cada cromosoma, y comparando a continuación las frecuencias relativas de un cromosoma con la frecuencia relativa de dos o más cromosomas (por ejemplo, dividiendo la suma para un cromosoma por la suma total para dos o más cromosomas). Una vez calculada, la proporción cromosómica se compara con la proporción cromosómica promedio de una población normal.

El promedio puede ser la media, mediana, modo o de otro promedio, con o sin normalización y exclusión de los datos de valores atípicos. En un aspecto preferente, se usa la media. Al desarrollar el conjunto de datos para la proporción cromosómica de la población normal, se calcula la variación normal de los cromosomas medidos. Esta variación se puede expresar de varias maneras, lo más típicamente como el coeficiente de variación, o CV. Cuando la proporción cromosómica de la muestra se compara con la proporción cromosómica promedio de una población normal, si la proporción cromosómica de la muestra queda estadísticamente fuera de la proporción cromosómica promedio de la población normal, la muestra contiene una aneuploidía. Los criterios para establecer el umbral estadístico para declarar una aneuploidía dependen de la variación en la medición de la proporción cromosómica y las tasas aceptables de falsos positivos y falsos negativos para el ensayo deseado. En general, este umbral puede ser un múltiplo de la variación observada en la proporción cromosómica. En un ejemplo, este umbral es tres o más veces la variación de la proporción cromosómica. En otro ejemplo, es cuatro o más veces la variación de la proporción cromosómica. En otro ejemplo, es cinco o más veces la variación de la proporción cromosómica. En otro ejemplo, es seis o más veces la variación de la proporción cromosómica. En el ejemplo anterior, la proporción cromosómica se estima determinando las frecuencias relativas (por ejemplo, sumando los recuentos) de regiones de ácido nucleico por cromosoma. Típicamente, se usa el mismo número de regiones de ácido nucleico para cada cromosoma. Un procedimiento alternativo para obtener la proporción cromosómica sería calcular la frecuencia relativa promedio para las regiones de ácido nucleico para cada cromosoma. El promedio puede ser cualquier estimación de la media, la mediana o el modo, aunque típicamente se usa un promedio. El promedio puede ser la media de todas las frecuencias relativas o alguna variación, tal como un promedio truncado o ponderado. Una vez que se han calculado las frecuencias relativas promedio para cada cromosoma, las frecuencias relativas promedio para cada cromosoma se pueden comparar con las de otros cromosomas para obtener una proporción cromosómica entre dos cromosomas, las frecuencias relativas promedio para cada cromosoma se pueden comparar con las frecuencias relativas para todos los cromosomas medidos para obtener una proporción cromosómica para cada cromosoma como se describe anteriormente. Como se destacó anteriormente, la capacidad de detectar una aneuploidía en una muestra materna de un embarazo por ovodonación en la que el ADN supuesto tiene una baja abundancia relativa depende en gran medida de la variación en las mediciones de diferentes regiones de ácido nucleico en el ensayo. Se pueden usar numerosos procedimientos analíticos que reducen esta variación y, por tanto, mejoran la sensibilidad de este procedimiento para detectar aneuploidías.

Un procedimiento para reducir la variabilidad del ensayo es incrementar el número de regiones de ácido nucleico usadas para calcular la abundancia de los cromosomas. En general, si la variación medida de una sola región de ácido nucleico de un cromosoma es  $X\%$  y se miden  $Y$  regiones de ácido nucleico diferentes en el mismo cromosoma, la variación en la medición de la abundancia cromosómica, calculada sumando o promediando la abundancia de cada región de ácido nucleico en ese cromosoma, será aproximadamente de  $X\%$  dividido por  $Y^{1/2}$ . Dicho de otra manera, la variación en la medición de la abundancia del cromosoma sería aproximadamente la variación promedio en la medición de la abundancia de cada región de ácido nucleico dividida por la raíz cuadrada del número de regiones de ácido nucleico.

En un aspecto preferente, el número de regiones de ácido nucleico medidas para cada cromosoma es de al menos 24. En otro aspecto preferente, el número de regiones de ácido nucleico medidas para cada cromosoma es de al menos 48. En otro aspecto preferente, el número de regiones de ácido nucleico medidas para cada cromosoma es de al menos 100. En otro aspecto preferente, el número de regiones de ácido nucleico medidas para cada cromosoma

es de al menos 200. Hay un coste incremental de medir cada región de ácido nucleico y, por tanto, es importante minimizar el número de regiones de ácido nucleico. En un aspecto preferente, el número de regiones de ácido nucleico medidas para cada cromosoma es inferior a 2000. En un aspecto preferente, el número de regiones de ácido nucleico medidas para cada cromosoma es inferior a 1000. En un aspecto más preferente, el número de regiones de ácido nucleico medidas para cada cromosoma es de al menos 48 e inferior a 1000. En un aspecto, después de la medición de la abundancia para cada región de ácido nucleico, se puede usar un subconjunto de las regiones de ácido nucleico para determinar la presencia o ausencia de aneuploidía. Hay muchos procedimientos estándar para elegir el subconjunto de regiones de ácido nucleico. Estos procedimientos incluyen la exclusión de valores atípicos, en la que se descartan del análisis las regiones de ácido nucleico con niveles detectados inferiores y/o superiores a un determinado percentil. En un aspecto, el percentil puede ser el 5 % más bajo y más alto medido por abundancia. En otro aspecto, el percentil puede ser el 10 % más bajo y más alto medido por abundancia. En otro aspecto, el percentil puede ser el 25 % más bajo y más alto medido por abundancia.

Otro procedimiento para elegir el subconjunto de regiones de ácido nucleico incluye la eliminación de las regiones que quedan fuera de un cierto límite estadístico. Por ejemplo, las regiones que quedan fuera de una o más desviaciones estándar de la abundancia media se pueden retirar del análisis. Otro procedimiento para elegir el subconjunto de regiones de ácido nucleico puede ser comparar la abundancia relativa de una región de ácido nucleico con la abundancia esperada de la misma región de ácido nucleico en una población sana y descartar cualquier región de ácido nucleico que no supere la prueba de expectativas. Para minimizar aún más la variación en el ensayo, se puede incrementar el número de veces que se mide cada región de ácido nucleico. Como se discutió, a diferencia de los procedimientos aleatorios de detección de aneuploidía en los que el genoma se mide en promedio menos de una vez, los sistemas de ensayo de la presente invención miden intencionalmente cada región de ácido nucleico múltiples veces. En general, al contabilizar eventos, la variación en el recuento se determina mediante estadísticas de Poisson, y la variación en el recuento es típicamente igual a uno dividido por la raíz cuadrada del número de recuentos. En un aspecto preferente, las regiones de ácido nucleico se miden cada una en promedio al menos 100 veces. En un aspecto preferente, las regiones de ácido nucleico se miden cada una en promedio al menos 500 veces. En un aspecto preferente, las regiones de ácido nucleico se miden cada una en promedio al menos 1000 veces. En un aspecto preferente, las regiones de ácido nucleico se miden cada una en promedio al menos 2000 veces. En un aspecto preferente, las regiones de ácido nucleico se miden cada una en promedio al menos 5000 veces.

En otro aspecto, subconjuntos de loci se pueden elegir al azar usando un número suficiente para proporcionar un resultado estadísticamente significativo en la determinación de si existe una anomalía cromosómica. Se pueden realizar múltiples análisis de diferentes subconjuntos de loci dentro de una muestra materna para obtener más poder estadístico. En este ejemplo, puede o no ser necesario retirar o eliminar cualquier loci antes del análisis aleatorio. Por ejemplo, si hay 100 regiones seleccionadas para el cromosoma 21 y 100 regiones seleccionadas para el cromosoma 18, se podrían realizar una serie de análisis que evalúen menos de 100 regiones para cada uno de los cromosomas.

Además de los procedimientos anteriores para reducir la variación en el ensayo, se pueden usar en combinación otras técnicas analíticas, muchas de las cuales se describen anteriormente en esta solicitud. En general, la variación en el ensayo se puede reducir cuando todas las regiones de ácido nucleico para cada muestra se examinan en una sola reacción en un solo recipiente. De manera similar, la variación en el ensayo se puede reducir cuando se usa un sistema de amplificación universal. Además, la variación en el ensayo se puede reducir cuando el número de ciclos de amplificación es limitado.

#### 45 **Uso del porcentaje de ADN libre circulante fetal para optimizar la detección de aneuploidía**

Una vez que se ha calculado el porcentaje de ADN libre circulante fetal, estos datos se pueden combinar con procedimientos para la detección de aneuploidía para determinar la probabilidad de que una muestra materna pueda contener una aneuploidía. En un aspecto se usa un procedimiento de detección de aneuploidía que utiliza análisis de segmentos de ADN aleatorios, tal como el que se describe, por ejemplo, en Quake, solicitud de patente de EE. UU. n.º 2007/202525; Shoemaker *et al.*, solicitud de patente de EE. UU. n.º 2010/136529. En un aspecto preferente se usan procedimientos de detección de aneuploidía que utilizan análisis de regiones de ácido nucleico seleccionadas. En este aspecto se calcula el porcentaje de ADN libre circulante fetal para una muestra. La proporción cromosómica para esa muestra, una proporción cromosómica para la población normal y una variación de la proporción cromosómica para la población normal se determinan como se describe en el presente documento.

En un aspecto preferente, la relación cromosómica y su variación para la población normal se determinan a partir de muestras normales que tienen un porcentaje similar de ADN fetal. Una proporción cromosómica de aneuploidía esperada para una muestra de ADN con ese porcentaje de ADN libre circulante fetal se calcula agregando el porcentaje de contribución del cromosoma con aneuploidía. La proporción cromosómica para la muestra se puede comparar con la proporción cromosómica para la población normal y con la proporción cromosómica de aneuploidía esperada para determinar estadísticamente, usando la variación de la proporción cromosómica, si es más probable que la muestra sea normal o aneuploide, y la probabilidad estadística de que sea una o la otra.

En un aspecto preferente, las regiones seleccionadas de una muestra materna de un embarazo por ovodonación incluyen ambas regiones para la determinación del contenido de ADN fetal, así como regiones no polimórficas de dos

o más cromosomas para detectar una anomalía cromosómica fetal en una única reacción. La reacción única ayuda a minimizar el riesgo de contaminación o sesgo que se pueden introducir durante varias etapas en el sistema de ensayo que de otro modo pueden distorsionar los resultados cuando se utiliza el contenido de ADN fetal para ayudar a determinar la presencia o ausencia de una anomalía cromosómica.

5 En otros aspectos, se pueden usar una o varias regiones seleccionadas tanto para la determinación del contenido de ADN fetal como para la detección de anomalías cromosómicas fetales. Los alelos para las regiones seleccionadas se pueden usar para determinar el contenido de ADN fetal y estas mismas regiones seleccionadas se pueden usar a continuación para detectar anomalías cromosómicas fetales que ignoran la información alélica. El uso de las mismas regiones tanto para el contenido de ADN fetal como para la detección de anomalías cromosómicas puede ayudar aún más a minimizar cualquier sesgo debido a un error experimental o a contaminación.

10 En determinados aspectos, el porcentaje de ADN libre circulante fetal se puede usar como una herramienta de control de calidad en la determinación de la presencia o ausencia de aneuploidía fetal. La determinación del porcentaje de ADN libre circulante fetal en una muestra materna en un embarazo por ovodonación puede proporcionar un mecanismo para determinar qué análisis adicionales, en su caso, se pueden realizar en la muestra. En algunas muestras maternas, el nivel de ADN libre circulante fetal en la muestra es bajo, y en algunos casos puede ser tan bajo que determinados análisis no se puedan realizar de manera fiable en la muestra, tal como los análisis para determinar la presencia o ausencia de aneuploidía fetal.

15 En la aneuploidía fetal, tal como en el caso de la trisomía 21, un incremento relativo del cromosoma 21 en el ADN fetal es de un 50 % en comparación con la disomía 21. Como porcentaje del ADN total en la muestra materna, en la que, como ejemplo, el ADN fetal es un 5 % del ADN total en la muestra, el incremento en el cromosoma 21 como porcentaje del total es de un 2,5 %. A medida que el porcentaje de ADN libre circulante fetal en una muestra disminuye, el incremento en el cromosoma 21 como porcentaje del ADN total en la muestra disminuye y puede ser difícil de detectar. Para garantizar una determinación exacta de la presencia o ausencia de aneuploidía fetal, se puede usar un nivel umbral de porcentaje de ADN libre circulante fetal como herramienta de control de calidad. Un "nivel umbral" es un nivel mínimo de ADN libre circulante fetal que se considera aceptable para la realización de análisis adicionales en la muestra, por ejemplo, solo las muestras maternas de embarazos por ovodonación con un nivel de ADN libre circulante fetal en o por encima del nivel umbral se usan para determinar la presencia o ausencia de aneuploidía fetal. El nivel umbral puede ser un nivel predeterminado establecido antes de la determinación del porcentaje fetal, o se puede establecer después del análisis del desempeño de la prueba y/o el análisis adicional en la prueba deseada para detectar anomalías fetales.

20 El uso de un nivel umbral en el análisis de muestras maternas de embarazos por ovodonación proporciona un estándar para garantizar la calidad y fiabilidad de los análisis de muestras. En determinados ejemplos específicos, la probabilidad de la presencia o ausencia de aneuploidía fetal se puede determinar en muestras en las que se puede demostrar que comprenden ADN libre circulante fetal en o por encima de un nivel umbral.

25 En determinados aspectos, el cálculo del porcentaje de ADN libre circulante fetal en la muestra materna se puede realizar antes de la determinación de la presencia o ausencia de aneuploidía fetal. En estos aspectos, si el porcentaje de ADN libre circulante fetal no alcanza el valor umbral, se puede determinar que la muestra no debe someterse a análisis adicionales. De forma alternativa, se pueden realizar análisis adicionales en la muestra, pero los resultados no se pueden usar o se pueden clasificar como poco fiables. En determinados aspectos, el cálculo del porcentaje de ADN libre circulante fetal en la muestra materna se puede realizar simultáneamente a la determinación de la presencia o ausencia de aneuploidía fetal. En estos aspectos, si el ADN libre circulante fetal en la muestra está por debajo de un valor umbral, la determinación de la presencia o ausencia de aneuploidía fetal puede no se puede usar o se puede clasificar como no fiable.

30 En determinados aspectos, el nivel umbral de ADN libre circulante fetal en una muestra materna puede estar entre un 1 y un 5 % del ADN libre circulante fetal, tal como entre un 1,5 y un 4,5 % del ADN libre circulante fetal, o más específicamente entre un 2 y un 4 % del ADN libre circulante fetal. En determinados aspectos específicos, el nivel umbral de porcentaje de ADN libre circulante fetal puede ser de un 1,0, un 1,5, un 2,0, un 2,5, un 3,0, un 3,5, un 4,0, un 4,5 o un 5,0 % del ADN libre circulante fetal, o cualquier número intermedio contenido en el mismo.

### 55 **Detección de mutaciones genéticas**

En determinados aspectos, el sistema de ensayo detecta tanto aneuploidías fetales como otras alteraciones genéticas (incluyendo anomalías cromosómicas) en loci de interés específicos. Dichas alteraciones genéticas adicionales incluyen, pero no se limitan a, mutaciones de delección, mutaciones de inserción, polimorfismos de número de copias, variantes del número de copias, síndrome de delección 22q11, síndrome de delección 11q en el cromosoma 11, síndrome de delección 8p en el cromosoma 8 y similares. En general, se analizan al menos dos secuencias de ácido nucleico objetivo presentes en el mismo cromosoma o en cromosomas separados, y al menos una de las secuencias objetivo se asocia con la anomalía alélica fetal. Las secuencias de las dos secuencias objetivo y el número de copias de las dos secuencias objetivo se comparan a continuación para determinar si la anomalía cromosómica está presente, y si es así, la naturaleza de la anomalía.

Si bien gran parte de la descripción contenida en el presente documento describe la detección de aneuploidía calculando la abundancia de regiones de ácido nucleico en uno o más supuestos cromosomas aneuploides y la abundancia de regiones de ácido nucleico en uno o más cromosomas normales, las mismas técnicas se pueden usar para detectar variaciones en el número de copias en las que dicha variación solo se produce una parte de un cromosoma. En esta detección de las variaciones en el número de copias, múltiples regiones de ácido nucleico dentro de la supuesta localización de variaciones en el número de copias se comparan con múltiples regiones de ácido nucleico fuera de la supuesta localización de variaciones en el número de copias. Otros aspectos de la invención descritos para la aneuploidía se pueden usar para la detección de la variación en el número de copias. Por ejemplo, se puede detectar un síndrome de delección 22q11 en un feto en una muestra materna mezclada seleccionando dos o más regiones de ácido nucleico dentro de la delección 22q11 y dos o más regiones de ácido nucleico fuera de la delección 22q11. Las regiones de ácido nucleico fuera de la delección 22q11 pueden estar en otra región del cromosoma 22 o pueden estar en un cromosoma completamente diferente. La abundancia de cada una de las regiones de ácido nucleico se determina mediante los procedimientos descritos en esta solicitud.

Las regiones de ácido nucleico dentro de la delección se cuantifican a continuación para determinar una frecuencia relativa y se hace lo mismo para las regiones de ácido nucleico fuera de la delección. Estas frecuencias relativas se comparan a continuación entre sí para determinar la presencia o ausencia de una delección. Opcionalmente, las frecuencias relativas se colocan en una proporción y esa proporción se puede comparar con una proporción promedio creada a partir de una población normal. Cuando la proporción para una muestra queda estadísticamente fuera de la proporción esperada, se detecta la delección. El umbral para la detección de una delección puede ser cuatro o más veces la variación calculada en la población normal.

#### Uso de otros procedimientos de detección fetal

Los procedimientos divulgados se pueden usar junto con la detección de otros factores de riesgo conocidos (por ejemplo, edad materna, antecedentes familiares, información genética materna o paterna) y/o medios para detectar anomalías fetales, y preferentemente con otros mecanismos de diagnóstico de anomalías fetales que sean relativamente no invasivos (por ejemplo, mediciones de uno o más marcadores bioquímicos en una muestra materna y/o mediciones o detección estructural a partir de una ecografía). El uso combinado de estos factores de riesgo y los mecanismos de diagnóstico con los procedimientos puede proporcionar una mejorada determinación del riesgo de anomalías fetales y, en particular, la presencia o ausencia de una mutación genética conocida, tal como una trisomía.

Por tanto, en algunos aspectos preferentes, los resultados obtenidos en los sistemas de ensayo se combinan con los resultados de la detección bioquímica de factores de riesgo, la detección por ecografía de factores de riesgo u otros determinantes de riesgo de anomalías fetales.

En algunos aspectos específicos, los resultados obtenidos en los sistemas de ensayo se combinan con la detección de marcadores bioquímicos asociados con un riesgo incrementado de anomalías fetales. Los marcadores bioquímicos se pueden determinar en base a una muestra que comprende sangre materna, suero, plasma u orina. Dichos marcadores bioquímicos incluyen, pero no se limitan a, beta hCG libre, proteína A plasmática asociada al embarazo (PAPP-A), alfa-fetoproteína en sangre materna, hCG en sangre materna, estriol no conjugado en sangre materna, inhibina dimérica A en sangre materna, estriol total en orina materna, fragmento beta-core en orina materna, hCG hiperglicosilada en orina materna, hCG hiperglicosilada en sangre materna e inhibina A (preferentemente inhibina A dimérica). En algunos aspectos, el mecanismo de evaluación adicional es el análisis multimarcador, tal como el descrito en Orlandi *et al.*, patente de EE. UU. n.º 7.315.787 o Wald *et al.*, patente de EE. UU. n.º 6.573.103. La detección de la presencia y/o de los niveles de estos y otros marcadores se puede combinar con los resultados de los sistemas de ensayo de la invención para proporcionar un resultado final a la paciente.

En otros aspectos específicos, los resultados obtenidos en los sistemas de ensayo se combinan con los resultados obtenidos de las imágenes ecográficas, que incluyen pero no se limitan a: grosor de translucidez nuchal (NT) o edema, grosor del pliegue nuchal, anomalía del sistema venoso (incluyendo el conducto venoso, la vena porta y hepática y la vena cava inferior), hueso nasal ausente o hipoplásico, longitud del fémur, longitud del húmero, intestino hiperecogénico, pielectasia renal, focos ecogénicos en el corazón, frecuencia cardíaca fetal y determinadas anomalías cardíacas. En aspectos específicos, la evaluación adicional de la anomalía fetal se realiza a través del análisis de forma, como se describe en las patentes de EE. UU. n.º 7.780.600 y 7.244.233. En un aspecto específico, la evaluación adicional se basa en la determinación de puntos de referencia basados en imágenes, como se describe en la patente de EE. UU. n.º 7.343.190. La detección de estos y otros parámetros físicos se puede combinar con los resultados de los sistemas de ensayo de la invención para proporcionar un resultado final a la paciente.

Se sabe que la mayoría de los marcadores de cribado y las características físicas varían con la edad gestacional. Para tener en cuenta esta variación, cada nivel de marcador se puede expresar como un múltiplo de la mediana de nivel (MoM) para embarazos no afectados de la misma edad gestacional. Especialmente para los marcadores derivados de las ecografías, la medición de la longitud cefalocaudal (CRL) o del diámetro biparietal (DBP) son medidas alternativas de la edad gestacional. Las MoM se pueden ajustar de una manera conocida para tener en cuenta los factores que se

sabe que afectan a los niveles de marcadores, tales como el peso materno, el grupo étnico, el estado diabético y el número de fetos en desarrollo.

El uso de las técnicas anteriores se puede realizar en una sola etapa del embarazo o se puede obtener secuencialmente en dos o más etapas diferentes del embarazo. Estos niveles de marcadores también se pueden interpretar en combinación con variables maternas como la edad materna, el peso, el grupo étnico, etc. para obtener una estimación del riesgo. La estimación del riesgo se realiza usando técnicas estadísticas estándar. Por ejemplo, los procedimientos conocidos se describen en Wald NJ *et al.*, BMJ (1992); 305(6850):391-4; Wald NJ *et al.* (1988) BMJ 297:883-887 y en Royston P, Thompson SG Stat Med. (1992) 11(2):257-68.

**Detección de otros agentes o factores de riesgo en muestras maternas de un embarazo por ovodonación**

Dada la naturaleza multiplexada de los sistemas de ensayo, en determinados aspectos puede ser beneficioso utilizar el ensayo para detectar otros ácidos nucleicos que podrían representar un riesgo para la salud del sujeto o sujetos o afectar de otro modo a decisiones clínicas sobre el tratamiento o pronóstico para un sujeto. Dichos ácidos nucleicos podrían incluir, pero no se limitan a, indicadores de enfermedad o riesgo tales como alelos maternos, polimorfismos o mutaciones somáticas que se sabe que presentan un riesgo para la salud materna o fetal. Dichos indicadores incluyen, pero no se limitan a, genes asociados con el estado Rh; mutaciones o polimorfismos asociados con enfermedades tales como diabetes, hiperlipidemia, hipercolesterolemia, trastornos sanguíneos tales como anemia drepanocítica, hemofilia o talasemia, afecciones cardíacas, etc.; ácidos nucleicos exógenos asociados con infecciones activas o latentes; mutaciones somáticas o variaciones en el número de copias asociadas con trastornos autoinmunitarios o neoplasias malignas (por ejemplo, cáncer de mama) o cualquier otro problema de salud que pueda afectar al sujeto, y en particular a las opciones clínicas que pueden estar disponibles en el tratamiento y/o prevención de riesgos para la salud en un sujeto en base las consecuencias de los resultados del ensayo.

En consecuencia, dado que los sistemas de ensayo preferentes están altamente multiplexados y pueden examinar cientos o incluso miles de ácidos nucleicos dentro de una muestra materna, en determinados aspectos es deseable examinar a la muestra en busca de marcadores de ácido nucleico dentro de la muestra materna, por ejemplo, ácidos nucleicos asociados con riesgo genético o que identifican la presencia o ausencia de organismos infecciosos. Por tanto, en determinados aspectos, los sistemas de ensayo proporcionan la detección de dichos ácidos nucleicos junto con la detección de ácidos nucleicos para la determinación del número de copias dentro de una muestra materna.

Por ejemplo, en determinadas muestras maternas de embarazos por ovodonación, muestras de sujetos con enfermedad autoinmunitaria y muestras de pacientes sometidas a quimioterapia, la supresión inmunitaria del sujeto puede incrementar el riesgo de la enfermedad debido a cambios en el sistema inmunitario del sujeto. La detección de agentes exógenos en una muestra materna puede ser indicativa de exposición a e infección por un agente infeccioso, y este hallazgo tiene un impacto en el cuidado de la paciente o en la gestión de una enfermedad infecciosa para la cual un sujeto ha dado positivo en una prueba para dicho agente infeccioso.

Específicamente, los cambios en la inmunidad y la fisiología durante el embarazo pueden hacer que las mujeres embarazadas sean más susceptibles o resulten más gravemente afectadas por enfermedades infecciosas. De hecho, el embarazo en sí mismo puede ser un factor de riesgo para adquirir determinadas enfermedades infecciosas, tales como toxoplasmosis, enfermedad de Hansen y listeriosis. Además, en las mujeres embarazadas o sujetos con sistemas inmunitarios suprimidos, determinadas enfermedades infecciosas tales como la gripe y la varicela pueden tener un curso clínico más grave, mayor tasa de complicaciones y mayor tasa de letalidad. Por lo tanto, la identificación de agentes de enfermedades infecciosas puede permitir un mejor tratamiento para la enfermedad materna durante el embarazo, lo que lleva a un mejor resultado general tanto para la madre como para el feto.

Además, determinados agentes infecciosos se pueden transmitir al feto por medio de transmisión vertical, es decir, la transmisión de infecciones de la madre al bebé. Estas infecciones pueden ocurrir mientras el feto todavía está en el útero, durante el parto o después del parto (tal como durante la lactancia).

Por tanto, en algunos aspectos preferentes, el sistema de ensayo puede incluir la detección de secuencias exógenas, por ejemplo, secuencias de organismos infecciosos que pueden tener un efecto adverso sobre la salud y/o la viabilidad del feto o el bebé, para proteger a la salud de la madre, del feto y/o del bebé.

Ejemplos de infecciones que se pueden transmitir por medio de transmisión vertical, y que se someten a prueba usando los procedimientos de ensayo, incluyen pero no se limitan a infecciones congénitas, infecciones perinatales e infecciones postnatales.

Las infecciones congénitas se transmiten en el útero atravesando la placenta para infectar el feto. Muchos microbios infecciosos pueden causar infecciones congénitas, lo que ocasiona problemas en el desarrollo fetal o incluso la muerte. TORCH es un acrónimo de varias de las infecciones congénitas más comunes. Estas son: toxoplasmosis, otras infecciones (por ejemplo, sífilis, hepatitis B, virus de Coxsackie, virus de Epstein-Barr, virus de varicela zóster y parvovirus B19 humano (eritema infeccioso)), rubéola, citomegalovirus (CMV) y virus del herpes simple.



Las infecciones perinatales se refieren a las infecciones que ocurren cuando el bebé pasa a través de un canal de parto infectado o a través de contaminación con heces durante el parto. Estas infecciones pueden incluir, pero no se limitan a, enfermedades de transmisión sexual (por ejemplo, gonorrea, clamidia, virus del herpes simple, virus del papiloma humano, etc.), CMV y estreptococos del grupo B (GBS).

Las infecciones que se transmiten de la madre al bebé después del parto se conocen como infecciones postnatales. Estas infecciones se pueden transmitir durante la lactancia a través de microbios infecciosos que se encuentran en la leche materna de la madre. Algunos ejemplos de infecciones postnatales son el CMV, el virus de inmunodeficiencia humana (VIH), el virus de la hepatitis C (VHC) y los GBS.

## EJEMPLOS

Los ejemplos siguientes se proponen con el fin de proporcionar a los expertos en la técnica una divulgación y una descripción completas de cómo hacer y usar la presente invención, y no se pretende que limiten el alcance de lo que los inventores consideran como su invención ni se pretende que representen o impliquen que los experimentos que figuran a continuación son todos o los únicos experimentos realizados. Por lo tanto, los aspectos actuales se deben considerar en todos los sentidos como ilustrativos y no restrictivos.

La eficiencia y la exactitud de la identificación del porcentaje de contribución fetal usando los procedimientos de la presente invención se demuestra en los Ejemplos siguientes, en los que en un grupo de 24 individuos, los procedimientos determinaron correctamente el porcentaje de contribución de cada individuo en las muestras mezcladas.

### Ejemplo 1: Sujetos y muestras

Los sujetos se inscribieron prospectivamente después de proporcionar un consentimiento informado según los protocolos aprobados por las juntas de revisión institucional. Los sujetos debían tener al menos 18 años de edad. Se seleccionó para su inclusión en este estudio a un subconjunto de sujetos inscritos, que consistía en 16 individuos de los que 8 individuos eran de una primera población étnica y 8 individuos de una segunda población étnica. Para imitar una muestra mezclada de un embarazo por ovodonación en la que la muestra materna comprende ADN de dos individuos distintos que no comparten aproximadamente la mitad de la información genética del otro, se prepararon muestras mezcladas que comprendían plasma de un individuo de la primera población étnica y de un individuo de la segunda población étnica. Para cada pareja de individuos, las muestras se mezclaron por volumen en fracciones de 100 %, 95 %, 90 %, 85 %, 15 %, 10 %, 5 % y 0 % del plasma de los individuos de la primera población étnica.

### Ejemplo 2: Análisis de loci polimórficos para evaluar el porcentaje de contribución de individuos no relacionados

Para evaluar el porcentaje de contribución de cada individuo en las muestras mezcladas, los ensayos se diseñaron para un conjunto de 192 loci que contienen SNP en los cromosomas 1 a 12, en los que se usaron dos oligos intermedios que diferían en una base para examinar cada SNP. Los SNP se optimizaron para la frecuencia de alelos menores en el conjunto de datos HapMap 3. Duan, *et al.*, *Bioinformatics*, 3(3):139-41(2008); publicación electrónica el 9 de noviembre de 2008.

Los oligonucleótidos fueron sintetizados por IDT (Integrated DNA Technologies, Coralville, IA) y se agruparon para crear un único conjunto de ensayo multiplexado. El producto del ensayo se generó a partir de cada muestra mezclada, y los productos de las muestras mezcladas se agruparon y usaron como molde para la amplificación de grupos en un solo carril de un portaobjetos de flujo TruSeq v2 (Illumina, San Diego, CA). El portaobjetos se procesó en un Illumina HiSeq 2000 para generar un promedio de 1180 lecturas de secuencia sin procesar por muestra. Un promedio de 1150 lecturas tuvo menos de 3 emparejamientos erróneos con las estructuras de ensayo esperadas, lo que dio como resultado un promedio de 854 lecturas/locus/muestra.

Los loci polimórficos informativos se definieron como loci en los que al menos un alelo de un primer individuo de una muestra mezclada difería de los alelos de un segundo individuo de una muestra mezclada. Debido a que el ensayo muestra especificidades de alelo superiores a un 99 %, los loci informativos de la donante de óvulos se identificaron fácilmente cuando se midió que la proporción de alelos de un locus de un individuo con una proporción menor era de entre un 1 y un 20 %. Para propósitos de comparación, se usaron dos conjuntos de loci informativos de donantes de óvulos para analizar los ensayos para las muestras mezcladas. En el primer conjunto de loci, denominado en el presente documento como el conjunto de "porcentaje de contribución homocigótica no idéntica no inclusiva", se usaron modelos binomiales para no considerar al menos una porción significativa de los ensayos en los cuales los alelos del primer individuo son homocigóticos y los alelos del segundo individuo son homocigóticos, pero los alelos del primer individuo son diferentes de los alelos del segundo individuo. En el segundo conjunto de loci, denominado en el presente documento como el conjunto de "porcentaje de contribución homocigótica no idéntica inclusiva", no se eliminaron loci informativos de la donante de óvulos del conjunto de ensayo. Para ambos conjuntos de loci, se estimó una probabilidad máxima usando una distribución binomial, para determinar la proporción fetal más probable en base a las mediciones

de varios loci informativos de donantes de óvulos. Los resultados correlacionaron bien ( $R^2 > 0,99$ ) con el enfoque del promedio ponderado presentado por Chu y colaboradores (véase Chu *et al.*, Prenat. Diagn., 30:1226-29 (2010)).

### Ejemplo 3: Resultados

La FIG. 3 es un gráfico de comparación de estimaciones del porcentaje de contribución para el conjunto de porcentaje de contribución homocigótica no idéntica inclusiva y el conjunto de porcentaje de contribución homocigótica no idéntica no inclusiva. Se esperaba que el conjunto de porcentaje de contribución homocigótica no idéntica inclusiva diera una estimación de porcentaje de contribución más alta que el conjunto de porcentaje de contribución homocigótica no idéntica no inclusiva debido a la presencia detectada de alelos polimórficos adicionales. Con la excepción de las muestras por debajo de los límites detectables de los procedimientos (aproximadamente un 3,5 %), el análisis del conjunto de porcentaje de contribución homocigótica no idéntica no inclusiva mostró una estimación de fracción fetal más baja que el análisis del conjunto de porcentaje de contribución homocigótica no idéntica inclusiva.

### Ejemplo 4: Validación de rendimiento: sujetos y muestras

Los sujetos se inscribieron prospectivamente después de proporcionar un consentimiento informado según los protocolos aprobados por las juntas de revisión institucional. Los sujetos debían tener al menos 18 años de edad. Se seleccionó para su inclusión en este estudio a un subconjunto de sujetos inscritos, que consistía en 8 individuos de los que 4 eran mujeres y 4 hombres. Para imitar una muestra mezclada de un embarazo por ovodonación en el que el feto es un macho, se prepararon muestras mezcladas que comprendían plasma de una hembra y un macho. Para cada pareja mujer-hombre, las muestras se mezclaron por volumen en fracciones de 100 %, 95 %, 90 %, 85 %, 15 %, 10 %, 5 % y 0 % del plasma de la mujer.

### Ejemplo 5: Validación de rendimiento: análisis de loci polimórficos para evaluar el porcentaje de contribución de individuos no relacionados

Para evaluar el porcentaje de contribución de cada individuo en las muestras mezcladas, los ensayos se diseñaron para un conjunto de 192 loci que contienen SNP en los cromosomas 1 a 12, en los que se usaron dos oligos intermedios que diferían en una base para examinar cada SNP. Los SNP se optimizaron para la frecuencia de alelos menores en el conjunto de datos HapMap 3. Duan, *et al.*, Bioinformatics, 3(3):139-41(2008); publicación electrónica el 9 de noviembre de 2008.

Los oligonucleótidos fueron sintetizados por IDT (Integrated DNA Technologies, Coralville, IA) y se agruparon para crear un único conjunto de ensayo multiplexado. El producto del ensayo se generó a partir de cada muestra mezclada, y los productos de las 8 muestras mezcladas se agruparon y usaron como molde para la amplificación de grupos en un solo carril de un portaobjetos de flujo TruSeq v2 (Illumina, San Diego, CA). El portaobjetos se procesó en un Illumina HiSeq 2000 para generar un promedio de 1180 lecturas de secuencia sin procesar por muestra. Un promedio de 1150 lecturas tuvo menos de 3 emparejamientos erróneos con las estructuras de ensayo esperadas, lo que dio como resultado un promedio de 854 lecturas/locus/muestra.

Los loci polimórficos informativos se definieron como loci en los que al menos un alelo de un primer individuo de una muestra mezclada difería de los alelos de un segundo individuo de una muestra mezclada. Debido a que el ensayo muestra especificidades de alelo superiores a un 99 %, los loci informativos de la donante de óvulos se identificaron fácilmente cuando se midió que la proporción de alelos de un locus de un individuo con una proporción menor era de entre un 1 y un 20 %. Para propósitos de comparación, se usaron dos conjuntos de loci informativos de donantes de óvulos para analizar los ensayos para las muestras mezcladas. En el primer conjunto de loci, denominado en el presente documento como el conjunto relacionado con la fracción fetal, se usaron funcionalidades binomiales para no considerar los ensayos en los cuales los alelos del primer individuo son homocigóticos y los alelos del segundo individuo son homocigóticos, pero los alelos del primer individuo son diferentes de los alelos del segundo individuo. En el segundo conjunto de loci, denominado en el presente documento como el conjunto no relacionado con la fracción fetal, no se eliminaron loci informativos de la donante de óvulos del conjunto de ensayo. Para ambos conjuntos de loci, se estimó una probabilidad máxima usando una distribución binomial, para determinar la proporción fetal más probable en base a las mediciones de varios loci informativos de donantes de óvulos. Los resultados correlacionaron bien ( $R^2 > 0,99$ ) con el enfoque del promedio ponderado presentado por Chu y colaboradores (véase Chu *et al.*, Prenat. Diagn., 30:1226-29 (2010)).

### Ejemplo 6: Validación de rendimiento: análisis de loci en cromosomas sexuales para evaluar el porcentaje de contribución de individuos no relacionados

Con el fin de validar la estimación del porcentaje de contribución usando loci polimórficos, el porcentaje de contribución se estimó a través del análisis de loci en los cromosomas sexuales de cada individuo en las muestras mezcladas y los resultados de ambas estimaciones se compararon. Los ensayos se diseñaron para un conjunto de 8 ensayos diseñados para los cromosomas sexuales y otros ensayos se diseñaron para los cromosomas 13, 18 y 21.

5 Los oligonucleótidos fueron sintetizados por IDT (Integrated DNA Technologies, Coralville, IA) y se agruparon para crear un único conjunto de ensayo multiplexado. El producto del ensayo se generó a partir de cada muestra mezclada, y los productos de las muestras mezcladas se agruparon y usaron como molde para la amplificación de grupos en un solo carril de un portaobjetos de flujo TruSeq v2 (Illumina, San Diego, CA). El portaobjetos se procesó en un Illumina HiSeq 2000 para generar un promedio de 1180 lecturas de secuencia sin procesar por muestra. Un promedio de 1150 lecturas tuvo menos de 3 emparejamientos erróneos con las estructuras de ensayo esperadas, lo que dio como resultado un promedio de 854 lecturas/locus/muestra.

10 El porcentaje de contribución no normalizado usando Y (FPY) se estimó usando las medianas de los recuentos de los ensayos de Y en comparación con las medianas de los recuentos de los ensayos en los cromosomas 13, 18 y 21.

$$FPY = \frac{\text{mediana } (Y)}{\text{mediana } (13,18,21)}$$

15 Dado que el porcentaje de contribución usando Y no está normalizado, se esperaba que el porcentaje estimado de contribución usando los ensayos polimórficos estaría fuertemente correlacionado, pero no sería idéntico.

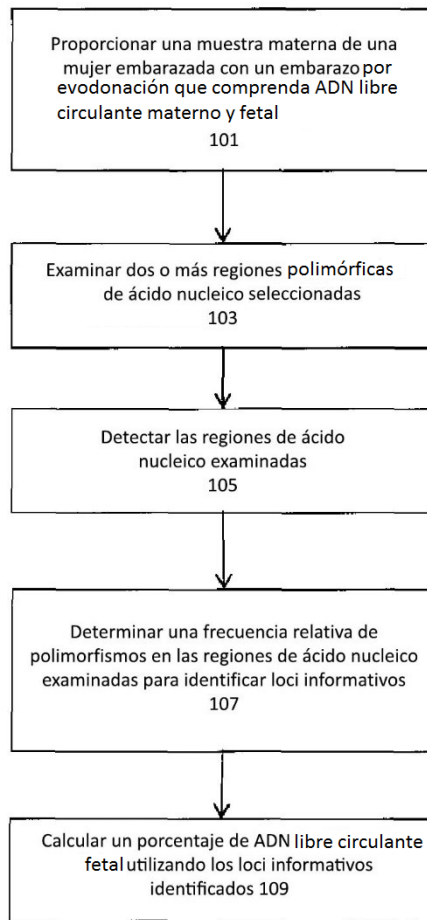
**Ejemplo 7: Validación de rendimiento: resultados**

20 Las FIGS. 4A-C son gráficos de barras que comparan el porcentaje de contribución estimado usando tanto los ensayos polimórficos como los ensayos de cromosomas sexuales para las muestras mezcladas que comprenden un 75 %, un 80 %, un 85 %, un 90 % y un 95 % de plasma de la mujer. En cada uno de estos casos, mientras que la estimación para cada muestra mezclada no es idéntica, existe una fuerte correlación entre la estimación del porcentaje de contribución usando los ensayos polimórficos y la estimación usando la evaluación de cromosomas sexuales.

25 Aunque la presente invención se satisface por aspectos en muchas formas diferentes, como se describe en detalle en relación con aspectos preferentes de la invención, se entiende que la presente divulgación ha de ser considerada como ejemplo de los principios de la invención y no está destinada a limitar la invención a los aspectos específicos ilustrados y descritos en el presente documento. El resumen y el título no se deben interpretar como limitantes del alcance de la presente invención, ya que su propósito es permitir a las autoridades apropiadas, así como al público en  
30 general, determinar rápidamente la naturaleza general de la invención.

**Reivindicaciones**

- 5 **1.** Un procedimiento para determinar el porcentaje de ADN libre circulante fetal en una muestra materna de una mujer embarazada con un embarazo por ovodonación, que comprende:
- proporcionar una muestra materna que comprende ADN libre circulante materno y fetal;  
 examinar dos o más regiones polimórficas de ácido nucleico seleccionadas;  
 detectar las regiones de ácido nucleico examinadas;  
 10 determinar una frecuencia relativa de polimorfismos en las regiones de ácido nucleico examinadas para identificar los loci informativos de la donante de óvulos, en el que los loci informativos de la donante de óvulos incluyen loci en los que el ADN materno y el ADN fetal difieren en al menos un alelo, y excluyen loci en los que ambos alelos fetales difieren de los alelos maternos; y  
 calcular un porcentaje de ADN fetal libre circulante usando los loci informativos de la donante de óvulos identificados.
- 15 **2.** El procedimiento de la reivindicación 1, en el que calcular el porcentaje de ADN libre circulante fetal comprende comparar una frecuencia relativa de alelos de baja frecuencia con una frecuencia relativa de alelos tanto de alta como de baja frecuencia.
- 20 **3.** El procedimiento de la reivindicación 1, en el que la etapa de examinar dos o más regiones polimórficas de ácido nucleico seleccionadas comprende:
- introducir un primer conjunto de oligonucleótidos de secuencia fija en la muestra materna en condiciones que permitan que los oligonucleótidos fijos hibriden específicamente con regiones complementarias de una  
 25 primera región polimórfica de ácido nucleico;  
 introducir un segundo conjunto de oligonucleótidos de secuencia fija en la muestra materna en condiciones que permitan que los oligonucleótidos fijos hibriden específicamente con regiones complementarias de una segunda región polimórfica de ácido nucleico; y  
 30 ligar los oligonucleótidos hibridados para crear productos de ligación contiguos complementarios de los ácidos nucleicos.
- 4.** El procedimiento de la reivindicación 3, en el que los oligonucleótidos de secuencia fija comprenden una región de cebador universal.
- 35 **5.** El procedimiento de la reivindicación 4, en el que el primer conjunto de oligonucleótidos de secuencia fija y el segundo conjunto de oligonucleótidos de secuencia fija comprenden la misma región de cebador universal.
- 6.** El procedimiento de la reivindicación 4, en el que el primer conjunto de oligonucleótidos de secuencia fija comprende una primera región de cebador universal y el segundo conjunto de oligonucleótidos de secuencia fija comprende una segunda región de cebador universal.
- 40 **7.** El procedimiento de la reivindicación 4, y que comprende además amplificar los productos de ligación contiguos usando la región de cebador universal para crear productos de amplificación.
- 45 **8.** El procedimiento de la reivindicación 7, en el que la amplificación comprende realizar la amplificación mediante reacción en cadena de la polimerasa (PCR) en los moldes de productos de ligación contiguos.
- 9.** El procedimiento de la reivindicación 1, en el que la muestra materna es plasma o suero materno.
- 50 **10.** El procedimiento de la reivindicación 1, en el que examinar dos o más regiones polimórficas de ácido nucleico seleccionadas comprende examinar al menos diez regiones de ácido nucleico.
- 11.** El procedimiento de la reivindicación 10, en el que examinar dos o más regiones polimórficas de ácido nucleico seleccionadas comprende examinar al menos veinte regiones de ácido nucleico.
- 55 **12.** El procedimiento de la reivindicación 11, en el que examinar dos o más regiones polimórficas de ácido nucleico seleccionadas comprende examinar al menos cuarenta regiones de ácido nucleico.
- 13.** El procedimiento de la reivindicación 12, en el que examinar dos o más regiones polimórficas de ácido nucleico seleccionadas comprende examinar al menos noventa regiones de ácido nucleico.
- 60 **14.** El procedimiento de la reivindicación 13, en el que examinar dos o más regiones polimórficas de ácido nucleico seleccionadas comprende examinar al menos 300 regiones de ácido nucleico.



**FIG. 1**

N.º de combinación	Genotipo materno	Genotipo de donante de óvulos	Genotipo fetal	¿Informativo?
201	AA	AA	AA	NO
202	AA	AA	AB	SÍ
203	AA	AB	AA	NO
204	AA	AB	AB	SÍ
205	AA	AB	BB	SÍ
206	AA	BB	AB	SÍ
207	AA	BB	BB	SÍ
208	AB	AA	AA	SÍ
209	AB	AA	AB	NO
210	AB	AB	AA	SÍ
211	AB	AB	AB	NO
212	AB	AB	BB	SÍ
213	AB	BB	AB	NO
214	AB	BB	BB	SÍ

FIG. 2

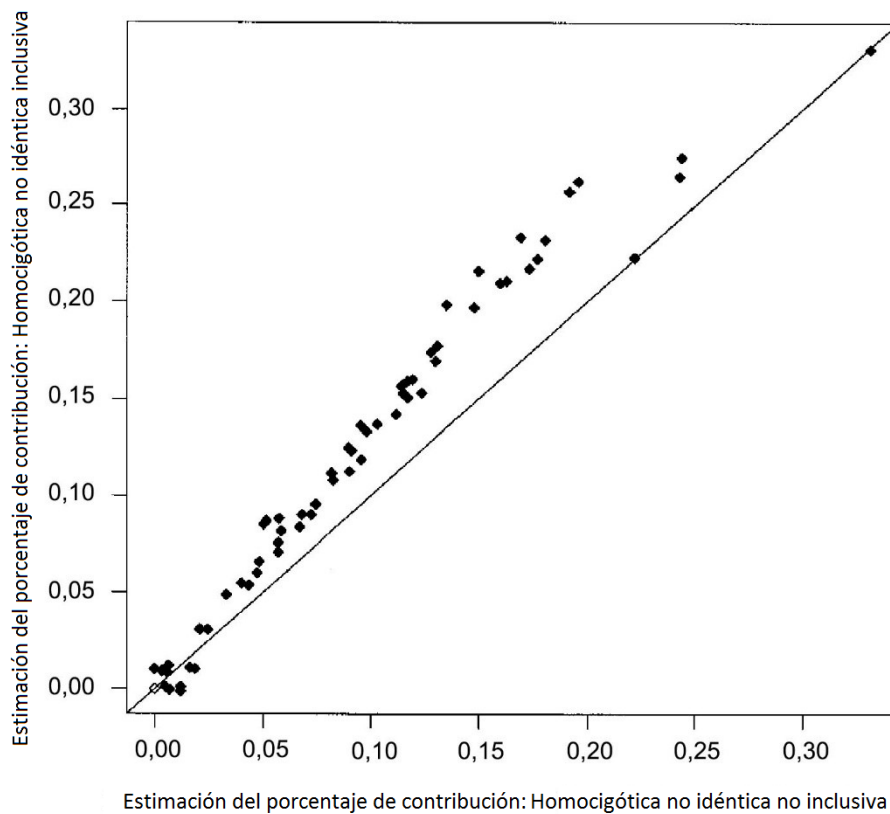


FIG. 3

FIG. 4A

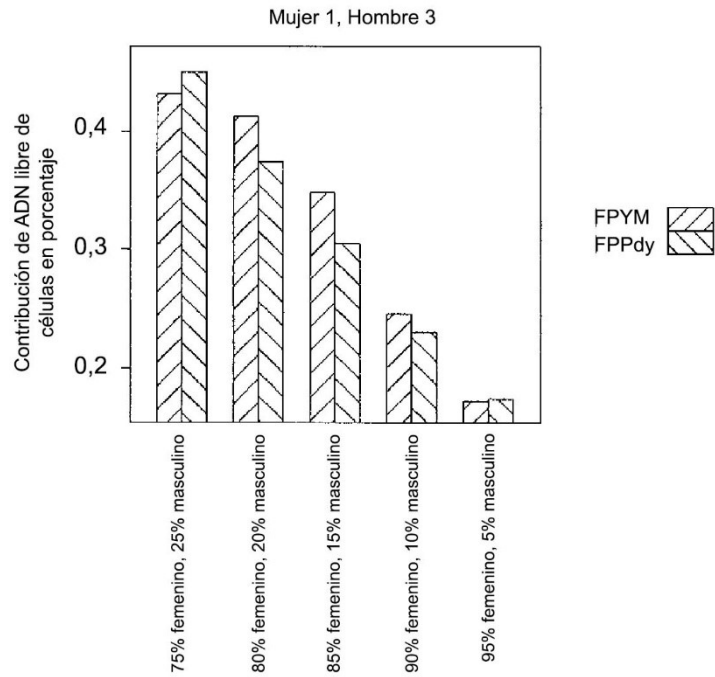
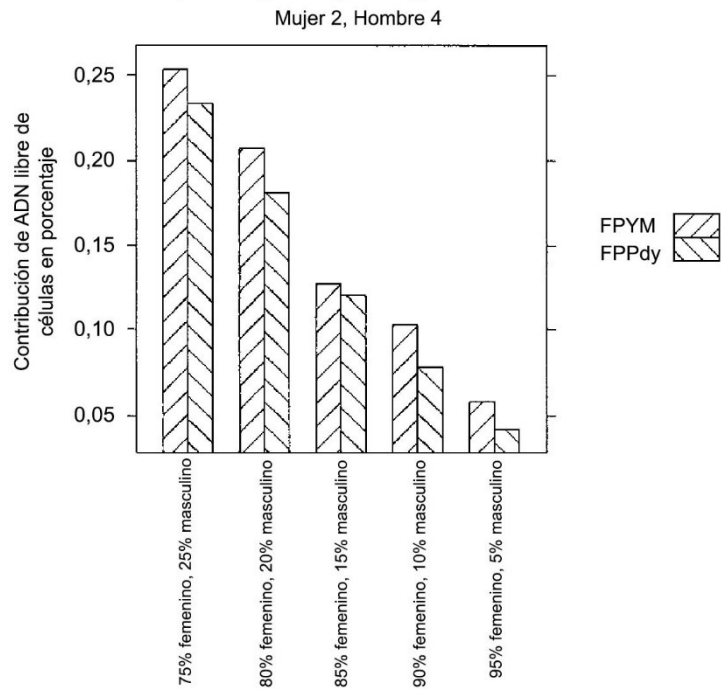


FIG. 4B



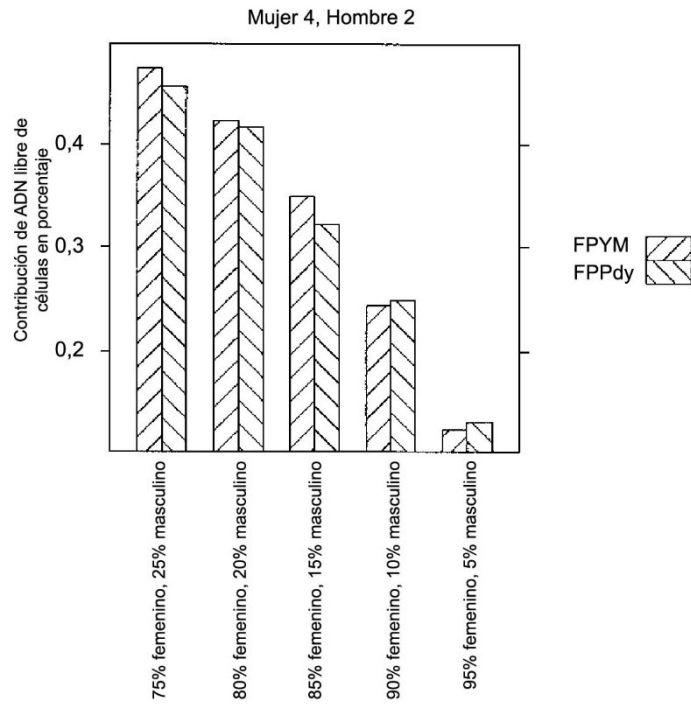


FIG. 4C