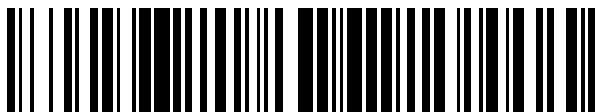




OFICINA ESPAÑOLA DE
PATENTES Y MARCAS
ESPAÑA



⑪ Número de publicación: **2 863 778**

⑮ Int. Cl.:

C12Q 1/6827 (2008.01)

C12Q 1/6862 (2008.01)

⑫

TRADUCCIÓN DE PATENTE EUROPEA

T3

⑯ Fecha de presentación y número de la solicitud europea: **08.08.2011 E 18163473 (4)**

⑯ Fecha y número de publicación de la concesión europea: **27.01.2021 EP 3395955**

⑭ Título: **Procedimientos de ensayo para la determinación de la contribución de una fuente en una muestra**

⑩ Prioridad:

**06.08.2010 US 37160510 P
25.01.2011 US 201113013732**

⑮ Fecha de publicación y mención en BOPI de la traducción de la patente:
11.10.2021

⑬ Titular/es:

**ARIOSA DIAGNOSTICS, INC. (100.0%)
5941 Optical Court
San Jose, CA 95138, US**

⑯ Inventor/es:

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

⑭ Agente/Representante:

LINAGE GONZÁLEZ, Rafael

ES 2 863 778 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

Procedimientos de ensayo para la determinación de la contribución de una fuente en una muestra

5 **CAMPO DE LA INVENCIÓN**

La presente invención se refiere a un procedimiento de ensayo para identificar la variación en el número de copias en muestras mixtas en un único ensayo.

10 **ANTECEDENTES DE LA INVENCIÓN**

En el siguiente análisis se describirán determinados artículos y procedimientos con propósitos preliminares 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 a demostrar, cuando resulte apropiado, que los artículos y procedimientos a los que se hace referencia en el presente documento no constituyen una técnica anterior en virtud de las disposiciones estatutarias aplicables. Los avances recientes en el diagnóstico se han centrado en mecanismos menos invasivos para determinar el riesgo, la presencia y el pronóstico de enfermedades. Los procedimientos de diagnóstico para determinar 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.

20 La identificación de ácidos nucleicos libres circulantes en muestras biológicas, tales como sangre y plasma, permite utilizar técnicas menos invasivas, tales como la extracción de sangre, para la toma de decisiones clínicas. Por ejemplo, se ha encontrado ADN libre circulante procedente de tumores sólidos malignos en la sangre periférica de pacientes con cáncer; el individuo que se ha sometido a un trasplante presenta en su circulación sanguínea ADN libre circulante procedente del órgano transplantado; y se han encontrado ADN y ARN fetales libres circulantes en la sangre y el plasma de mujeres embarazadas. Además, la detección de ácidos nucleicos procedentes de organismos infecciosos, tal como la detección de la carga vírica o la identificación genética de cepas específicas de un agente patógeno vírico o bacteriano, proporciona importantes indicadores de diagnóstico y pronóstico. Los ácidos nucleicos libres circulantes procedentes de una fuente separada de las células normales propias del paciente pueden proporcionar, por tanto, información médica importante, por ejemplo, sobre opciones de tratamiento, diagnóstico, pronóstico y similares.

25 30 35 La sensibilidad de dichas pruebas a menudo depende de la identificación de la cantidad de ácido nucleico procedente de las diferentes fuentes y, en particular, de la identificación de un nivel bajo de ácidos nucleicos de una fuente ante el fondo de un nivel mayor de ácidos nucleicos de una segunda fuente. La detección de la contribución de la especie de ácido nucleico secundaria a los ácidos nucleicos libres circulantes presentes en la muestra biológica puede proporcionar una interpretación estadística exacta de los datos resultantes.

40 45 Por tanto, existe la necesidad de desarrollar procesos para calcular la variación en el número de copias (VNC) en una o más regiones genómicas en una muestra biológica usando información sobre la contribución de los ácidos nucleicos en la muestra. La presente invención aborda esta necesidad.

SUMARIO DE LA INVENCIÓN

50 55 Este sumario se proporciona para introducir una selección de conceptos de forma simplificada que se describen además a continuación en la descripción detallada. No se pretende que el presente sumario identifique rasgos característicos clave o esenciales de la materia objeto reivindicada, ni tampoco se pretende que se use para limitar el alcance de la materia objeto reivindicada. Otros rasgos característicos, detalles, utilidades y ventajas de la materia objeto reivindicada serán evidentes a partir de la siguiente descripción detallada por escrito, incluyendo los aspectos ilustrados en los dibujos adjuntos y definidos en las reivindicaciones adjuntas.

60 65 Los procedimientos de la invención comprenden un sistema de un único ensayo con la capacidad de determinar la contribución de una fuente principal y/o una fuente secundaria dentro de una muestra de un individuo y determinar la información sobre el número de copias para una o más regiones genómicas dentro de una única fuente para determinar una diferencia de valor para la región genómica en comparación con la contribución de la fuente de los locus en la muestra mixta. La presente invención utiliza un sistema de un único ensayo que utiliza tanto la detección no polimórfica como la polimórfica para determinar la contribución de la fuente y las variaciones en el número de copias (VNC) de una única fuente dentro de la muestra mixta. La determinación de la contribución de la fuente principal y/o secundaria dentro de la muestra puede proporcionar información sobre la presencia de suficiente material genético de ambas fuentes para permitir una identificación adecuada de las regiones genómicas para la determinación de la VNC en las muestras mixtas.

En un aspecto, el sistema de ensayo utiliza la amplificación y detección de locus seleccionados en una muestra mixta de un individuo para calcular la contribución de la fuente e identificar el número de copias de una o más regiones genómicas. En un aspecto específico, la invención proporciona sistemas de un único ensayo con la

capacidad de determinar la contribución de ácidos nucleicos procedentes de una fuente principal y/o secundaria en la muestra mixta y la presencia o ausencia de VNC en una o más regiones genómicas de una sola fuente en una muestra mixta. El sistema de ensayo puede detectar específicamente el número de copias de las regiones genómicas presentes en dos o más fuentes dentro de la muestra mixta. Para la determinación de la contribución, los 5 locus seleccionados de una fuente se distinguen de los locus seleccionados de al menos otra fuente en la muestra mixta. Para la determinación de la variación en el número de copias de una región genómica, los locus seleccionados se pueden, pero no se necesitan, distinguir en cuanto a la contribución de la fuente, ya que la 10 variación en el número de copias se puede detectar mediante la comparación de los niveles de dos o más regiones genómicas dentro de una muestra mixta. Preferentemente, los ácidos nucleicos libres circulantes analizados en el sistema de ensayo son ADN libre circulante (ADNlc).

Por tanto, en una primera aplicación, la invención proporciona un sistema de un único ensayo para 1) determinar la contribución de una fuente principal y/o una fuente secundaria en una muestra mixta usando datos de frecuencia 15 derivados de dos o más locus informativos; 2) determinar la frecuencia de una o más regiones genómicas en la fuente principal y secundaria; y 3) identificar la presencia o ausencia de una VNC para una o más regiones genómicas en la fuente principal y/o secundaria. Preferentemente, la identificación de la VNC se basa en la comparación del número de copias de dos o más regiones genómicas de la fuente principal y/o secundaria en la muestra mixta.

20 Preferentemente, los ácidos nucleicos analizados usando los sistemas de la invención son ácidos nucleicos libres circulantes. Más preferentemente, los ácidos nucleicos analizados en el sistema de ensayo comprenden ADN libre circulante (ADNlc).

25 En otro aspecto específico, la invención proporciona un sistema de ensayo para calcular la contribución de una fuente y detectar la presencia o ausencia de VNC en una o más regiones genómicas dentro de una muestra mixta, comprendiendo el ensayo las etapas de introducir un primer conjunto de oligonucleótidos de secuencia fija en una muestra mixta en condiciones que permitan que los oligonucleótidos fijos hibriden específicamente con regiones complementarias en uno o más locus presentes en o asociados con una región genómica; introducir un segundo conjunto de oligonucleótidos de secuencia fija en la muestra mixta en condiciones que permitan que los 30 oligonucleótidos fijos hibriden específicamente con regiones complementarias en dos o más locus informativos; unir los oligonucleótidos hibridados para crear productos de unión contiguos complementarios a los locus seleccionados; amplificar los productos de unión contiguos para crear productos de amplificación; y detectar los productos de amplificación. La detección de los productos de amplificación se usa para calcular la contribución de la fuente y el 35 número de copias de una o más regiones genómicas en la muestra mixta.

35 Los conjuntos de oligonucleótidos fijos comprenden dos o más oligonucleótidos que hibridan con regiones contiguas de las regiones genómicas o los locus informativos. En algunos aspectos preferentes, se consultan conjuntos de locus en cuanto a una VNC que son indicativos de una amplificación de una región genómica más grande, por ejemplo todo o parte de un cromosoma. Preferentemente, los sistemas de ensayo pueden distinguir el número de 40 copias de estos locus entre una fuente principal y una fuente secundaria dentro de una muestra mixta. Los niveles de los locus seleccionados se pueden determinar para una región genómica de interés y comparar con las cantidades de locus de una o más de otras regiones genómicas de interés y/o una o más regiones genómicas de referencia para detectar posibles VNC en base a las frecuencias de los locus en la muestra mixta.

45 La detección de anomalías cromosómicas en una muestra se puede basar en la detección de VNC para múltiples locus seleccionados localizados en o asociados con un único cromosoma de una fuente secundaria y/o principal. Por tanto, en otro aspecto específico, la invención proporciona un procedimiento de un único ensayo para 1) determinar la contribución de una fuente principal y/o una fuente secundaria en una muestra mixta usando datos de frecuencia 50 derivados de dos o más locus informativos; 2) determinar la frecuencia de una o más regiones genómicas en la fuente principal y secundaria; y 3) identificar la presencia o ausencia de una anomalía cromosómica en la fuente principal y/o secundaria en la muestra mixta.

55 Por tanto, la invención proporciona un procedimiento de ensayo para calcular la contribución de una fuente y detectar la presencia o ausencia de una anomalía cromosómica en una muestra mixta usando un único ensayo, comprendiendo el ensayo las etapas de: introducir un primer conjunto de oligonucleótidos de secuencia fija en una muestra mixta en condiciones que permitan que los oligonucleótidos de secuencia fija hibriden específicamente con regiones complementarias en dos o más locus correspondientes a un primer cromosoma; introducir un segundo conjunto de oligonucleótidos de secuencia fija en la muestra mixta en condiciones que permitan que los 60 oligonucleótidos fijos hibriden específicamente con regiones complementarias en dos o más locus correspondientes a un segundo cromosoma; introducir un tercer conjunto de oligonucleótidos de secuencia fija en la muestra mixta en condiciones que permitan que los oligonucleótidos fijos hibriden específicamente con regiones complementarias en dos o más locus informativos; unir los oligonucleótidos hibridados para crear productos de unión contiguos complementarios a los ácidos nucleicos; amplificar los productos de unión contiguos para crear productos de amplificación; y detectar los productos de amplificación. La detección de los productos de amplificación se puede 65 usar para calcular la contribución de una fuente, así como para identificar anomalías cromosómicas en la muestra mixta.

En un aspecto más específico, los procedimientos de ensayo de la invención se usan para identificar el porcentaje de contribución fetal y de anomalías cromosómicas en una muestra materna. La invención proporciona un procedimiento de ensayo que comprende las etapas de: introducir un primer conjunto de oligonucleótidos de secuencia fija en una muestra materna en condiciones que permitan que los oligonucleótidos fijos hibriden específicamente con regiones complementarias en dos o más locus correspondientes a un primer cromosoma;

- 5 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 en dos o más locus correspondientes a un segundo cromosoma; introducir un tercer conjunto de oligonucleótidos de secuencia fija en la muestra materna en condiciones que permitan que los oligonucleótidos fijos hibriden específicamente con regiones complementarias en dos o más locus informativos; unir los oligonucleótidos hibridados para crear productos de unión contiguos complementarios a los ácidos nucleicos; amplificar los productos de unión contiguos para crear productos de amplificación; y detectar los productos de amplificación. La detección de los productos de amplificación se puede usar para calcular la contribución fetal, así como para identificar anomalías cromosómicas en los ácidos nucleicos fetales en la muestra materna.
- 10

En determinados aspectos de la invención, los oligonucleótidos fijos hibridan con regiones inmediatamente adyacentes en la región contigua, de modo que se unen directamente durante la etapa de unión del ensayo. En otros aspectos, sin embargo, puede haber un hueco de uno o más nucleótidos entre los extremos de los oligonucleótidos fijos después de la hibridación en la región contigua. Los oligonucleótidos fijos se unen mediante una combinación de, por ejemplo, extensión del cebador usando una polimerasa y unión.

20 Cada conjunto de ácidos nucleicos de secuencia fija está diseñado para hibridar con al menos dos regiones separadas en un locus seleccionado. En aspectos preferentes, se usan dos o más oligonucleótidos separados en un conjunto para hibridar con estas regiones para proporcionar ácidos nucleicos adyacentes complementarios a los locus seleccionados. En algunos aspectos, sin embargo, un conjunto puede comprender una única sonda con dos o más regiones distintas no adyacentes que son complementarias a los locus seleccionados (por ejemplo, sondas candado), como se describe con más detalle en el presente documento. Los conjuntos de oligonucleótidos de secuencia fija se pueden proporcionar en el ensayo de forma secuencial o simultánea en el ensayo.

25

30 En determinados aspectos preferentes, se usan oligonucleótidos puente para incrementar la especificidad de los conjuntos de oligonucleótidos y/o llenar un hueco entre los oligonucleótidos de secuencia fija. En consecuencia, otro aspecto específico de la invención proporciona un procedimiento de ensayo para calcular la contribución de una fuente y detectar la presencia o ausencia de VNC en una o más regiones genómicas dentro de una muestra mixta,

35 comprendiendo el ensayo las etapas de introducir un primer conjunto de oligonucleótidos de secuencia fija en una muestra mixta en condiciones que permitan que los oligonucleótidos fijos hibriden específicamente con regiones complementarias en uno o más locus presentes en o asociados con una región genómica; introducir un segundo conjunto de oligonucleótidos de secuencia fija en la muestra mixta en condiciones que permitan que los oligonucleótidos fijos hibriden específicamente con regiones complementarias en al menos un locus informativo;

40 introducir uno o más oligonucleótidos puente en condiciones que permitan que los oligonucleótidos puente hibriden específicamente con regiones complementarias en los locus, en los que uno o más oligonucleótidos puente son complementarios a una región de los locus entre e inmediatamente adyacente a las regiones complementarias a los oligonucleótidos de secuencia fija de cada conjunto; unir los oligonucleótidos hibridados para crear productos de unión contiguos complementarios a los ácidos nucleicos; amplificar los productos de unión contiguos para crear

45 productos de amplificación; y detectar los productos de amplificación. La detección de los productos de amplificación se usa para calcular la contribución de la fuente y el número de copias de una o más regiones genómicas en la muestra mixta.

50 Otro aspecto específico de la invención proporciona un procedimiento de ensayo que usa oligonucleótidos puente para calcular la contribución de una fuente e identificar anomalías cromosómicas en una muestra mixta. Este ensayo comprende las etapas de: introducir un primer conjunto de oligonucleótidos de secuencia fija en una muestra mixta en condiciones que permitan que los oligonucleótidos fijos hibriden específicamente con regiones complementarias en dos o más locus correspondientes a un primer cromosoma; introducir un segundo conjunto de oligonucleótidos de secuencia fija en la muestra mixta en condiciones que permitan que los oligonucleótidos fijos hibriden específicamente con regiones complementarias en dos o más locus correspondientes a un segundo cromosoma;

55 introducir un tercer conjunto de oligonucleótidos de secuencia fija en la muestra mixta en condiciones que permitan que los oligonucleótidos fijos hibriden específicamente con regiones complementarias en dos o más locus informativos; introducir uno o más oligonucleótidos puente en condiciones que permitan que los oligonucleótidos puente hibriden específicamente con regiones complementarias en los locus, en los que uno o más oligonucleótidos puente son complementarios a una región de los locus entre e inmediatamente adyacente a las regiones complementarias a los oligonucleótidos de secuencia fija de cada conjunto; introducir uno o más oligonucleótidos puente en condiciones que permitan que los oligonucleótidos puente hibriden específicamente con regiones complementarias en los locus, en los que uno o más oligonucleótidos puente son complementarios a una región de los locus entre e inmediatamente adyacente a las regiones complementarias a los oligonucleótidos de secuencia fija de cada conjunto; unir los oligonucleótidos hibridados para crear productos de unión contiguos complementarios a los ácidos nucleicos; amplificar los productos de unión contiguos para crear productos de amplificación; y detectar los productos de amplificación. La detección de los productos de amplificación se usa para calcular la contribución de la fuente y el número de copias de una o más regiones genómicas en la muestra mixta.

60

65 Otro aspecto específico de la invención proporciona un procedimiento de ensayo que usa oligonucleótidos puente para calcular la contribución de una fuente e identificar anomalías cromosómicas en una muestra mixta. Este ensayo comprende las etapas de: introducir un primer conjunto de oligonucleótidos de secuencia fija en una muestra mixta en condiciones que permitan que los oligonucleótidos fijos hibriden específicamente con regiones complementarias en dos o más locus correspondientes a un primer cromosoma; introducir un segundo conjunto de oligonucleótidos de secuencia fija en la muestra mixta en condiciones que permitan que los oligonucleótidos fijos hibriden específicamente con regiones complementarias en dos o más locus informativos; introducir uno o más oligonucleótidos puente en condiciones que permitan que los oligonucleótidos puente hibriden específicamente con regiones complementarias en los locus, en los que uno o más oligonucleótidos puente son complementarios a una región de los locus entre e inmediatamente adyacente a las regiones complementarias a los oligonucleótidos de secuencia fija de cada conjunto; introducir uno o más oligonucleótidos puente en condiciones que permitan que los oligonucleótidos puente hibriden específicamente con regiones complementarias en los locus, en los que uno o más oligonucleótidos puente son complementarios a una región de los locus entre e inmediatamente adyacente a las regiones complementarias a los oligonucleótidos de secuencia fija de cada conjunto; unir los oligonucleótidos hibridados para crear productos de unión contiguos complementarios a los ácidos nucleicos; amplificar los productos de unión contiguos para crear productos de amplificación; y detectar

los productos de amplificación. La detección de los productos de amplificación se puede usar para calcular la contribución de una fuente (tal como la contribución fetal en una muestra materna), así como para identificar anomalías cromosómicas en la muestra mixta.

5 En determinados aspectos de la invención, los oligonucleótidos fijos hibridan con regiones inmediatamente adyacentes a los oligonucleótidos puente, de modo que se unen directamente durante la etapa de unión del ensayo. En otros aspectos, sin embargo, puede haber un hueco de uno o más nucleótidos entre el extremo de uno o ambos oligonucleótidos fijos y el oligonucleótido puente después de la hibridación del oligonucleótido puente. Los oligonucleótidos fijos y el oligonucleótido puente se unen mediante una combinación de, por ejemplo, extensión del cebador usando una polimerasa y unión.

10 Un rasgo característico importante es que los ensayos multiplexados de la invención permiten el análisis simultáneo de 5 o más, preferentemente 10 o más, preferentemente 16 o más, preferentemente 20 o más, preferentemente 30 o más, preferentemente 32 o más, preferentemente 40 o más, preferentemente 48 o más, preferentemente 50 o más, preferentemente 60 o más, preferentemente 70 o más, preferentemente 80 o más, preferentemente 90 o más, y más preferentemente 96 o más locus seleccionados. Estos locus seleccionados pueden ser locus diferentes de una única muestra, o pueden ser locus de dos o más muestras mixtas. En el último caso, al menos uno de los dos oligonucleótidos de secuencia fija usados para el análisis de un locus seleccionado comprenderá un identificador de muestra (por ejemplo, un "índice de muestra") que permitirá asociar el locus con una muestra particular. De forma alternativa, se puede añadir un índice de muestra durante la amplificación del producto de unión usando un cebador que comprenda el índice de muestra.

15 En aspectos preferentes, la consulta de estos locus utiliza técnicas de amplificación universal que permiten la amplificación de múltiples locus en una única reacción de amplificación. Los ácidos nucleicos seleccionados para el cálculo de la contribución y la detección de VNC y/o anomalías cromosómicas en el sistema de ensayo de la invención se pueden amplificar usando procedimientos de amplificación universal después de la amplificación selectiva inicial de la muestra mixta. El uso de la amplificación universal permite amplificar múltiples regiones de ácidos nucleicos de una única muestra o de múltiples muestras usando un único o un número limitado de cebadores de amplificación, y es especialmente útil para amplificar múltiples regiones seleccionadas en una única reacción. En 20 un aspecto preferente, las regiones de cebador universal se utilizan en la determinación de secuencias de los productos de amplificación. En otro aspecto preferente, se usan las mismas regiones de cebador universal en los oligonucleótidos de secuencia fija usados para la detección de regiones genómicas y los oligonucleótidos de secuencia fija usados para la detección de polimorfismos.

25 35 Por tanto, en un aspecto específico de la invención, las secuencias complementarias a los cebadores para su uso en la amplificación universal se introducen en los locus seleccionados durante o después de la amplificación selectiva. Preferentemente, dichas secuencias se introducen en los extremos de dichos ácidos nucleicos seleccionados, aunque se pueden introducir en cualquier localización que permita la identificación del producto de amplificación obtenido mediante el procedimiento de amplificación universal.

40 45 50 En determinados aspectos preferentes, uno o ambos de los cebadores usados comprenden un índice de muestra u otro identificador. En un aspecto específico, se incluye un índice de muestra en uno o más de los cebadores universales. El índice de muestra se incorpora a los productos de amplificación y, a continuación, se pueden combinar los productos de amplificación de diferentes muestras. El índice de muestra se detecta al mismo tiempo que se detecta la VNC o anomalía cromosómica y se detecta el polimorfismo, de modo que la VNC y el polimorfismo se pueden asignar adecuadamente a la muestra de origen.

55 60 En determinados aspectos, el sistema de ensayo multiplexa la consulta de locus usando uno o más oligonucleótidos puente comunes que son complementarios a regiones de dos o más locus consultados, es decir, se puede usar un único oligonucleótido puente para dos o más conjuntos de oligonucleótidos fijos. Esto permite que el número de oligonucleótidos puente usados en el sistema de ensayo multiplexado sea menor que el número de locus consultados en el ensayo. En determinados aspectos específicos, el sistema de ensayo usa un grupo de oligonucleótidos puente que están diseñados cada uno para ser compatibles con dos o más locus consultados usando el sistema de ensayo de la invención.

65 70 Las frecuencias de los locus seleccionados se pueden determinar para una región genómica de interés y comparar con las frecuencias de los locus de una o más de otras regiones genómicas de interés y/o una o más regiones genómicas de referencia para detectar posibles VNC en base a las frecuencias de los locus en la muestra mixta.

75 80 En algunos casos, la anomalía cromosómica detectada se asocia con la amplificación génica o la expansión de locus en un cromosoma de interés. En otros casos, la anomalía cromosómica se asocia con una translocación que da como resultado la presencia de una parte extra de un cromosoma en el genoma. Aún en otros casos, la anomalía cromosómica es una delección.

85 90 En determinados aspectos preferentes, la anomalía cromosómica se asocia con aneuploidía de un cromosoma de interés. Por ejemplo, las anomalías cromosómicas más comunes en el feto son las trisomías 21, 18, 13, X y/o Y o la

monosomía X. En aspectos específicos preferentes, los sistemas de ensayo de la invención se usan para detectar dichas aneuploidías cromosómicas comunes en el ADN fetal de una muestra materna.

- 5 En los procedimientos de ensayo de la invención, los productos de amplificación se aíslan opcionalmente antes de la detección. Cuando se aíslan, preferentemente se aíslan como moléculas individuales para facilitar la detección posterior. Después del aislamiento, los productos de amplificación se pueden amplificar adicionalmente para crear copias idénticas de todos o una parte de los productos de amplificación individuales antes de la detección. De forma alternativa, los productos de amplificación aislados se pueden amplificar adicionalmente para crear copias idénticas de moléculas complementarias a todos o una parte de los productos de amplificación individuales antes de la detección.
- 10 En los procedimientos de ensayo de la invención se pueden emplear diversos procedimientos de detección de VNC junto con la detección de polimorfismos. En un aspecto general, el sistema de ensayo emplea un procedimiento para la determinación de una VNC en uno o más locus en una muestra mixta, que comprende las etapas de amplificar uno o más ácidos nucleicos seleccionados de una primera región genómica de interés en una muestra mixta; amplificar uno o más ácidos nucleicos seleccionados de un segundo locus de interés en la muestra mixta, determinar la frecuencia relativa de los locus seleccionados, comparar la frecuencia relativa de los locus seleccionados e identificar la presencia o ausencia de una VNC en base a las frecuencias relativas comparadas de los ácidos nucleicos seleccionados en el primer y segundo locus. Preferentemente, el procedimiento de ensayo amplifica dos o más locus seleccionados de diferentes regiones genómicas, aunque los locus pueden estar localizados en la misma región genómica general para confirmar las VNC surgidas de anomalías cromosómicas en lugar de las VNC de un único locus.
- 15 En los procedimientos de ensayo de la invención se pueden emplear diversos procedimientos de detección de VNC junto con la detección de polimorfismos. En un aspecto general, el sistema de ensayo emplea un procedimiento para la determinación de una VNC en uno o más locus en una muestra mixta, que comprende las etapas de amplificar uno o más ácidos nucleicos seleccionados de una primera región genómica de interés en una muestra mixta; amplificar uno o más ácidos nucleicos seleccionados de un segundo locus de interés en la muestra mixta, determinar la frecuencia relativa de los locus seleccionados, comparar la frecuencia relativa de los locus seleccionados e identificar la presencia o ausencia de una VNC en base a las frecuencias relativas comparadas de los ácidos nucleicos seleccionados en el primer y segundo locus. Preferentemente, el procedimiento de ensayo amplifica dos o más locus seleccionados de diferentes regiones genómicas, aunque los locus pueden estar localizados en la misma región genómica general para confirmar las VNC surgidas de anomalías cromosómicas en lugar de las VNC de un único locus.
- 20 En los procedimientos de ensayo de la invención se pueden emplear diversos procedimientos de detección de VNC junto con la detección de polimorfismos. En un aspecto general, el sistema de ensayo emplea un procedimiento para la determinación de una VNC en uno o más locus en una muestra mixta, que comprende las etapas de amplificar uno o más ácidos nucleicos seleccionados de una primera región genómica de interés en una muestra mixta; amplificar uno o más ácidos nucleicos seleccionados de un segundo locus de interés en la muestra mixta, determinar la frecuencia relativa de los locus seleccionados, comparar la frecuencia relativa de los locus seleccionados e identificar la presencia o ausencia de una VNC en base a las frecuencias relativas comparadas de los ácidos nucleicos seleccionados en el primer y segundo locus. Preferentemente, el procedimiento de ensayo amplifica dos o más locus seleccionados de diferentes regiones genómicas, aunque los locus pueden estar localizados en la misma región genómica general para confirmar las VNC surgidas de anomalías cromosómicas en lugar de las VNC de un único locus.
- 25 Más preferentemente, los oligonucleótidos de secuencia fija no hibridados se eliminan antes de la introducción de los oligonucleótidos puente. En algunos aspectos, los oligonucleótidos puente se introducen simultáneamente con la mezcla de unión. En otros aspectos, los productos de hibridación de los oligonucleótidos de secuencia fija y el locus se aíslan antes de la introducción de los oligonucleótidos puente.
- 30 En determinados aspectos específicos, el sistema de ensayo usa un grupo de oligonucleótidos puente que están diseñados cada uno para ser compatibles con dos o más locus consultados usando el sistema de ensayo de la invención. En estos aspectos, los oligonucleótidos puente usados en el ensayo multiplexado están diseñados preferentemente para presentar una T_f en un intervalo de ± 5 °C, más preferentemente en un intervalo de ± 2 °C.
- 35 En determinados aspectos, los oligonucleótidos puente presentan una longitud de entre 2 y 45 nucleótidos. En un aspecto específico, los oligonucleótidos puente presentan una longitud de entre 3 y 9 nucleótidos. Aún en otro aspecto específico, los oligonucleótidos presentan una longitud de entre 10 y 30 nucleótidos.
- 40 Los locus consultados respecto a la VNC pueden en algunos casos ser indicativos de una amplificación de una región genómica más grande, por ejemplo, todo o parte de un cromosoma. Preferentemente, los sistemas de ensayo pueden distinguir el número de copias de estos locus entre una fuente principal y una fuente secundaria dentro de una muestra mixta.
- 45 En otro aspecto, la presente invención utiliza técnicas que permiten la identificación tanto de VNC como de agentes infecciosos en una muestra mixta. Esto puede ser especialmente útil en la monitorización de pacientes en los que el desenlace clínico puede verse comprometido por la presencia de un agente infeccioso. Por ejemplo, un paciente que se ha sometido a un trasplante probablemente estará tomando medicamentos inmunosupresores y, por lo tanto, será más propenso a las infecciones en general. De forma similar, las mujeres embarazadas sufren cambios en su sistema inmunitario y, por tanto, pueden ser más susceptibles a infecciones por agentes patógenos que pueden tener un efecto adverso en la madre y/o el feto. Además, determinados tipos de cáncer se asocian con agentes infecciosos (por ejemplo, cáncer de hígado asociado con infecciones por hepatitis B y C, cáncer de cuello uterino asociado con infecciones por el virus del papiloma humano), y la identificación de los agentes infecciosos puede ser informativa para predecir el desenlace clínico o determinar el curso de tratamiento médico preferente para el paciente.
- 55 Por tanto, en determinados aspectos, la invención proporciona un procedimiento de ensayo para calcular la contribución de una fuente, detectar la presencia o ausencia de una VNC de una región genómica y la presencia o ausencia de un agente infeccioso en una muestra mixta usando un único ensayo, comprendiendo el ensayo las etapas de: introducir un primer conjunto de oligonucleótidos de secuencia fija en una muestra mixta en condiciones que permitan que los oligonucleótidos fijos hibriden específicamente con regiones complementarias en uno o más locus presentes en o asociados con una región genómica; introducir un segundo conjunto de oligonucleótidos de secuencia fija en la muestra mixta en condiciones que permitan que los oligonucleótidos fijos hibriden específicamente con regiones complementarias en al menos un locus informativo; introducir un tercer conjunto de oligonucleótidos de secuencia fija en la muestra mixta en condiciones que permitan que los oligonucleótidos de secuencia fija hibriden específicamente con regiones complementarias en los locus indicativos de un agente infeccioso; unir los oligonucleótidos hibridados para crear un producto de unión contiguo complementario a los locus; amplificar el producto de unión contiguo para crear productos de amplificación; y detectar los productos de
- 60
- 65

amplificación. La detección de los productos de amplificación se correlaciona con el número de copias de la región genómica y la presencia o ausencia de un agente infeccioso en la muestra mixta.

En otro aspecto general, el sistema de ensayo emplea un procedimiento para determinar la presencia o ausencia de una anomalía cromosómica asociada con una VNC en una región genómica, que comprende las etapas de amplificar uno o más locus seleccionados de un primer cromosoma de interés en una muestra mixta; amplificar uno o más locus seleccionados de un segundo cromosoma de interés en la muestra mixta, determinar la frecuencia relativa de las regiones seleccionadas en el primer y segundo cromosomas de interés, comparar la frecuencia relativa de las regiones seleccionadas en el primer y segundo cromosomas de interés e identificar la presencia o ausencia de una anomalía en base a las frecuencias relativas comparadas de las regiones seleccionadas. Preferentemente, se seleccionan dos o más regiones de ácidos nucleicos de cada cromosoma, y más preferentemente se seleccionan cinco o más locus de cada cromosoma.

Aún en otro aspecto general, el sistema de ensayo emplea un procedimiento para determinar la presencia o ausencia de una aneuploidía en una muestra mixta de un individuo, que comprende las etapas de amplificar dos o más locus seleccionados en el ADNlc correspondiente a un primer cromosoma de interés en una muestra mixta; amplificar dos o más locus seleccionados en el ADNlc correspondiente a un segundo cromosoma de interés en la muestra mixta, determinar la frecuencia relativa de las regiones seleccionadas en el primer y segundo cromosomas de interés, comparar la frecuencia relativa de las regiones seleccionadas en el primer y segundo cromosomas de interés e identificar la presencia o ausencia de una aneuploidía en base a las frecuencias relativas comparadas de las regiones seleccionadas. En un aspecto específico, los locus del primer y segundo cromosomas se amplifican en una única reacción, y preferentemente en una única reacción que transcurre dentro de un único recipiente.

Preferentemente, el sistema de ensayo detecta la presencia o ausencia de locus en muestras que se pueden obtener fácilmente de un sujeto, tales como sangre, plasma, suero y similares. En un aspecto general, el sistema de ensayo utiliza la detección de regiones seleccionadas en el ADNlc en una muestra mixta. En un aspecto más específico, el sistema de ensayo utiliza la detección de regiones seleccionadas en el ADNlc de una muestra mixta de un individuo para identificar la presencia o ausencia de VNC en una región genómica y la presencia o ausencia de un polimorfismo en uno o más locus. El número de copias dentro de una región genómica se puede determinar en base a la detección de las cantidades de locus seleccionados y la comparación con las cantidades de locus seleccionados de otra región genómica y/o las cantidades de locus seleccionados de una región genómica de referencia. En un aspecto particular, la proporción de las frecuencias del ácido nucleico se compara con una proporción media de referencia que se ha determinado para una población estadísticamente significativa de sujetos genéticamente "normales", es decir, sujetos que no tienen una VNC asociada con el locus particular consultado en el sistema de ensayo.

En un aspecto preferente de la invención, los productos de amplificación correspondientes a los ácidos nucleicos seleccionados se aíslan como moléculas individuales para el análisis de los locus seleccionados. Estos productos de amplificación individuales se aíslan unos de otros y, preferentemente, se aíslan físicamente (por ejemplo, sobre un sustrato o en recipientes individuales). Las moléculas individuales se pueden amplificar adicionalmente después del aislamiento para generar múltiples copias idénticas del producto de amplificación, de una parte del mismo o de un ácido nucleico complementario al producto de amplificación o a una parte del mismo.

En un aspecto preferente, los productos de amplificación individuales se analizan a través de la determinación de secuencias. En otros aspectos, los productos de amplificación individuales se analizan usando técnicas de hibridación.

Un rasgo característico de la presente invención es que el número de copias de los locus seleccionados se puede detectar usando procedimientos de detección no polimórfica, es decir, procedimientos de detección que no dependen de la presencia o ausencia de un polimorfismo particular para identificar la región de ácido nucleico seleccionada. En un aspecto preferente, los sistemas de ensayo de detección utilizan procedimientos de detección no polimórfica para "contar" el número relativo de los locus seleccionados presentes en una muestra mixta. Estos valores se pueden utilizar para determinar si, estadísticamente, es probable que una muestra mixta tenga una VNC en una región genómica en una fuente principal y/o secundaria dentro de la muestra mixta. De forma similar, estos valores se pueden utilizar para determinar si, estadísticamente, los ácidos nucleicos de la fuente principal y/o la fuente secundaria presentan uno o más polimorfismos. Dicha información se puede usar para identificar una patología o trastorno genético en particular, para confirmar un diagnóstico o una recidiva de una enfermedad o trastorno, para determinar el pronóstico de una enfermedad o trastorno, para ayudar a determinar posibles opciones de tratamiento, etc.

En algunos aspectos, los procedimientos para la determinación de una aneuploidía usados en el sistema de ensayo miden la variación en el número de copias de múltiples locus seleccionados de dos o más cromosomas en una muestra. Los niveles de los diferentes locus seleccionados correspondientes a cromosomas específicos se pueden cuantificar individualmente y comparar para determinar la presencia o ausencia de una aneuploidía cromosómica en una o más fuentes de células en una muestra mixta. Las regiones cuantificadas individualmente se pueden someter

a un cálculo de normalización, o los datos se pueden someter a una exclusión de valores atípicos antes de la comparación para determinar la presencia o ausencia de una aneuploidía en una muestra mixta.

5 En otros aspectos, las frecuencias relativas de los locus seleccionados se usan para determinar una frecuencia cromosómica del primer y segundo cromosomas de interés, y la presencia o ausencia de una aneuploidía se basa en las frecuencias cromosómicas comparadas del primer y segundo cromosomas de interés.

10 Aún en otros aspectos, las frecuencias relativas de los locus seleccionados se usan para determinar una frecuencia cromosómica de un cromosoma de interés y un cromosoma de referencia, y la presencia o ausencia de una aneuploidía se basa en las frecuencias cromosómicas comparadas del cromosoma de interés y el cromosoma de referencia.

15 Como el procedimiento de ensayo de la invención se configura preferentemente como un sistema altamente multiplexado, se pueden analizar simultáneamente múltiples locus de un único 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 agrupados, dichos datos se identifican preferentemente para las diferentes muestras antes del análisis de aneuploidía. En algunos aspectos, sin embargo, los datos agrupados se pueden analizar para detectar posibles VNC, y posteriormente se pueden analizar muestras individuales del grupo si los resultados iniciales indican que se detecta una posible aneuploidía dentro del grupo agrupado.

20 En determinados aspectos, los sistemas de ensayo utilizan uno o más índices que proporcionan información sobre muestras específicas. Por ejemplo, se puede usar un índice en la amplificación selectiva o universal que sea indicativo de una muestra a partir de la cual se amplificó el ácido nucleico.

25 En un aspecto particular, los locus seleccionados se aislan antes de la detección. Para el análisis, los locus seleccionados se pueden aislar de la muestra mixta usando cualquier medio que aísle selectivamente los ácidos nucleicos particulares presentes en la muestra mixta, por ejemplo, hibridación, amplificación u otra forma de aislamiento basado en secuencia de los ácidos nucleicos de la muestra mixta. Después del aislamiento, los ácidos nucleicos seleccionados se distribuyen individualmente en un formato de detección adecuado, por ejemplo, en una micromatriz o en una cubeta de lectura, para la determinación de secuencias y/o las cantidades relativas de cada ácido nucleico seleccionado en la muestra mixta. Las cantidades relativas de los ácidos nucleicos detectados son indicativas del número de copias de cromosomas que corresponden a los ácidos nucleicos seleccionados presentes en la muestra mixta.

30 35 En un aspecto específico, la invención proporciona un procedimiento de ensayo para detectar la presencia o ausencia de una aneuploidía fetal, que comprende las etapas de proporcionar una muestra mixta que comprende ADNlc materno y fetal, amplificar dos o más locus seleccionados de un primer y segundo cromosomas de interés en la muestra mixta, amplificar dos o más locus seleccionados del primer y segundo cromosomas de interés en la muestra mixta, determinar la frecuencia relativa de las regiones seleccionadas en los cromosomas de interés, comparar la frecuencia relativa de los locus seleccionados en el primer y segundo cromosomas de interés e identificar la presencia o ausencia de una aneuploidía fetal en base a las frecuencias relativas comparadas de los locus seleccionados.

40 45 50 55 En algunos aspectos específicos, las frecuencias relativas de los locus en una región genómica se calculan individualmente, y las frecuencias relativas de los locus individuales se comparan para determinar la presencia o ausencia de una anomalía cromosómica. En otros aspectos específicos, las frecuencias relativas de los locus seleccionados se usan para determinar una frecuencia cromosómica de un primer y un segundo cromosoma de interés y un cromosoma de referencia, y la variación en el número de copias para el cromosoma o una región genómica del cromosoma se basa en las frecuencias cromosómicas comparadas del primer y segundo cromosomas de interés.

60 65 La muestra mixta usada para el análisis se puede obtener o derivar de cualquier muestra que contenga el ácido nucleico de interés que se vaya a analizar usando el procedimiento de ensayo de la invención. Por ejemplo, una muestra mixta puede proceder de cualquier fluido materno que comprenda ácidos nucleicos libres circulantes tanto maternos como fetales, que incluye, pero sin limitarse a, plasma materno, suero materno o sangre materna. Una muestra mixta de un paciente sometido a trasplante sería cualquier fluido o tejido que contenga ácidos nucleicos libres circulantes tanto de las células del donante como de las células del paciente. Una muestra mixta de un paciente con una neoplasia maligna contendría ácidos nucleicos libres circulantes del tejido sano normal del paciente, así como ácidos nucleicos libres circulantes procedentes de las células cancerosas.

Aunque preferentemente el sistema de ensayo se usa para detectar ADNlc en una muestra mixta, en determinados aspectos el ADN de interés que se va a analizar usando el procedimiento de ensayo de la invención comprende ADN obtenido directamente de los diferentes tipos de células en lugar de una muestra mixta que contiene ADN de los tipos de células principal y secundario. Dichas muestras se pueden obtener de diversas fuentes, dependiendo del

5 ADN diana. Por ejemplo, las células fetales para el análisis pueden derivar de muestras tales como líquido amniótico, placenta (por ejemplo, las vellosidades coriónicas) y similares. Se pueden obtener muestras de órganos de donante en un individuo mediante biopsia. Los organismos infecciosos se pueden aislar directamente de un individuo y analizar después del aislamiento. El ADN se puede extraer de células o tejidos cancerosos y usar para el análisis.

10 Otro rasgo característico de la invención es que la gran mayoría de los ácidos nucleicos aislados de la muestra mixta y detectados en el sistema de ensayo proporcionan información relevante sobre la presencia, cantidad y/o naturaleza polimórfica de un locus particular en la muestra mixta. Esto asegura que la mayoría de los ácidos nucleicos analizados en el procedimiento de ensayo de la invención sean informativos.

15 En algunos aspectos, se consulta un conjunto de múltiples locus seleccionados para cada región genómica, y la cantidad del conjunto de regiones seleccionadas presentes en la muestra mixta se suma individualmente para determinar la frecuencia relativa de una región genómica en una muestra mixta. Esto incluye la determinación de la frecuencia del locus para el ADN materno y fetal combinados presentes en la muestra mixta. Preferentemente, la determinación no requiere una distinción entre el ADN de fuentes separadas, aunque en determinados aspectos se puede obtener esta información además de la información sobre las frecuencias relativas en la muestra en su conjunto.

20 En aspectos preferentes, los ácidos nucleicos seleccionados correspondientes a locus informativos se detectan y se suman para determinar la frecuencia relativa de una región genómica en la muestra mixta. Las frecuencias que son mayores de lo esperado para los locus de una primera región genómica en comparación con los locus de un segundo locus en una muestra mixta son indicativas de una VNC de la primera región genómica en la muestra mixta.

25 30 La comparación de regiones genómicas puede ser una comparación de parte o de la totalidad de un cromosoma. Por ejemplo, la región genómica analizada para detectar una VNC puede ser un cromosoma completo del feto (por ejemplo, los cromosomas 18 y 21), donde la probabilidad de que ambos sean aneuploides es mínima. También se puede realizar una comparación de cromosomas donde uno es supuestamente aneuploide (por ejemplo, el cromosoma 21) y el otro actúa como cromosoma de referencia (por ejemplo, un autosoma tal como el cromosoma 2). Aún en otros aspectos, la comparación puede implicar dos o más cromosomas que son supuestamente aneuploides y uno o más cromosomas de referencia.

35 40 En un aspecto, el procedimiento de ensayo de la invención analiza múltiples ácidos nucleicos que representan locus seleccionados en los cromosomas de interés, y la frecuencia relativa de cada locus seleccionado en la muestra se analiza para determinar una frecuencia cromosómica relativa para cada cromosoma de interés particular en la muestra. A continuación, se compara la frecuencia cromosómica de dos o más cromosomas o partes de los mismos para determinar estadísticamente si existe una anomalía cromosómica.

45 50 En otro aspecto, el procedimiento de ensayo de la invención analiza múltiples copias de un conjunto de locus seleccionados en los cromosomas de interés, y la frecuencia relativa de cada uno de los locus seleccionados en la muestra se analiza y se cuantifica independientemente para determinar una frecuencia para cada locus seleccionado en la muestra. La suma de los locus en la muestra se compara para determinar estadísticamente si existe una VNC para uno o más locus en una región genómica de una fuente en una muestra mixta.

55 60 En otro aspecto, se analizan subconjuntos de locus en cada cromosoma para determinar si existe una anomalía cromosómica. La frecuencia de los locus se puede sumar para un cromosoma en particular y las sumas de los locus se pueden usar para determinar la presencia de aneuploidía. En este aspecto de la invención, se suman las frecuencias de los locus individuales en cada región genómica y, a continuación, se compara la suma de los locus en una región genómica de un cromosoma con la de una región genómica de otro cromosoma para determinar si existe una anomalía cromosómica. Los subconjuntos de locus se pueden elegir al azar pero con un número suficiente de locus para proporcionar un resultado estadísticamente significativo en la determinación de si existe una anomalía cromosómica. Se pueden realizar múltiples análisis de diferentes subconjuntos de locus dentro de una muestra mixta para proporcionar mayor potencia estadística. En otro aspecto, se pueden seleccionar locus particulares que se sabe que tienen menor variación entre las muestras, o limitando los datos usados para la determinación de la frecuencia cromosómica, por ejemplo, ignorando los datos de los locus con una frecuencia muy alta o muy baja dentro de una muestra.

65 En un aspecto particular, las cantidades medidas de uno o más locus particulares se normalizan para tener en cuenta las diferencias en la cantidad de locus en la muestra. Esto se puede realizar normalizando la variación conocida debida a fuentes tales como el sistema de ensayo (por ejemplo, temperatura, diferencias entre lotes de

reactivos), la biología subyacente de la muestra (por ejemplo, contenido de ácido nucleico), las diferencias entre operarios o cualquier otra variable.

En determinados aspectos específicos, la determinación del porcentaje relativo de ácidos nucleicos procedentes de la fuente secundaria en una muestra mixta puede ser beneficiosa para realizar el sistema de ensayo, ya que proporcionará información importante sobre la presencia estadística relativa de locus que puede ser indicativa de la variación en el número de copias dentro de la fuente secundaria en esa muestra. La determinación de los locus aportados por la fuente secundaria a la muestra mixta puede proporcionar información que se puede usar para calcular las diferencias estadísticamente significativas en las frecuencias para regiones genómicas de interés. Por tanto, dichos locus podrían proporcionar dos tipos de información en el ensayo: se puede usar la información alélica para determinar el porcentaje de contribución de las células secundarias en una muestra mixta y se puede usar una suma de la información alélica para determinar la frecuencia global relativa de ese locus en una muestra mixta. La información alélica no es necesaria para determinar la frecuencia global relativa de ese locus.

En otro aspecto específico, el procedimiento de ensayo de la invención se puede utilizar para determinar un posible mosaicismo en una población celular y si se deben realizar pruebas de confirmación adicionales para confirmar la identificación del mosaicismo en la fuente principal y/o secundaria. En determinados casos, la determinación del porcentaje de ácidos nucleicos procedentes de la fuente secundaria en una muestra mixta podría ayudar a cuantificar el nivel estimado del mosaicismo. El mosaicismo se podría confirmar posteriormente usando otros procedimientos de prueba que pudieran distinguir una aneuploidía total o parcial en mosaico en células o tejidos específicos.

Aún en otro aspecto específico, el procedimiento de ensayo de la invención se puede utilizar para determinar la contaminación en una muestra, con la especie secundaria representando una especie contaminante.

En otro aspecto, los oligonucleótidos para un ácido nucleico seleccionado dado se pueden conectar en los extremos no específicos de secuencia de modo que una sonda circular o unimolecular se pueda unir a ellos. En este aspecto, el extremo 3' y el extremo 5' de la sonda circular se unen al locus seleccionado, y al menos una región de amplificación universal está presente en la secuencia específica no seleccionada de la sonda circular.

Un rasgo característico importante del ensayo es que los productos de amplificación se pueden analizar directamente sin necesidad de enriquecer las regiones polimórficas de la muestra mixta inicial. Por tanto, la presente invención permite detectar tanto la VNC como los polimorfismos en una muestra materna sin una etapa de enriquecimiento polimórfico intermedia anterior a la determinación de secuencias de los locus seleccionados.

Otro rasgo característico importante del ensayo es que tanto la VNC como la contribución de la fuente se determinan usando un enfoque dirigido de amplificación y detección seleccionadas. Esto permite que la mayoría de la información recopilada en el ensayo sea útil para la determinación de la VNC y/o la contribución de la fuente, y evita la necesidad de generar lecturas de secuencia que deben estar alineadas con una secuencia de referencia.

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

BREVE DESCRIPCIÓN DE LAS FIGURAS

La FIG. 1 es un diagrama de flujo simplificado de las etapas generales utilizadas en los sistemas de ensayo de la invención.

La FIG. 2 ilustra un primer esquema general de un sistema de ensayo de la invención basado en unión.

La FIG. 3 ilustra un segundo esquema general de un sistema de ensayo de la invención basado en unión.

La FIG. 4 es un tercer esquema general de un sistema de ensayo de la invención basado en unión.

La FIG. 5 ilustra el rendimiento del genotipado que se obtuvo usando un sistema de ensayo de la invención.

La FIG. 6 es un gráfico que ilustra los resultados de una determinación del porcentaje fetal usando un ensayo de la invención.

La FIG. 7 ilustra los elementos usados para la detección de aneuploidía y polimorfismo en dos cohortes de muestras maternas.

La FIG. 8 es un sumario de la información y los datos de pacientes y muestras para un subconjunto de una segunda cohorte de mujeres embarazadas.

La FIG. 9 ilustra la detección de aneuploidía del cromosoma 21 lograda usando un aspecto de la invención en una primera cohorte.

5 La FIG. 10 ilustra la detección de aneuploidía del cromosoma 18 lograda usando un aspecto de la invención en una primera cohorte.

La FIG. 11 ilustra la detección de aneuploidía del cromosoma 21 lograda usando un aspecto de la invención en una 10 segunda cohorte.

La FIG. 12 ilustra la detección de aneuploidía del cromosoma 18 lograda usando un aspecto de la invención en una segunda cohorte.

DEFINICIONES

- 15 Se pretende que los términos usados en el presente documento tengan el significado simple y ordinario como entienden los expertos en la técnica. Se pretende que las siguientes definiciones ayuden al lector en el entendimiento de la presente invención, pero no se pretende que varíen o de otro modo limiten el significado de dichos términos a menos que se indique específicamente.
- 20 El término “ácido nucleico amplificado” es cualquier molécula de ácido nucleico cuya cantidad se haya incrementado en al menos dos veces mediante cualquier procedimiento de replicación o amplificación de ácido nucleico realizado *in vitro* en comparación con su cantidad inicial en una muestra mixta.
- 25 El término “producto de amplificación”, como se usa en el presente documento, se refiere al producto resultante de una reacción de amplificación usando el producto de unión contiguo como molde, o al producto resultante de una reacción de amplificación usando una molécula complementaria al producto de unión contiguo como molde.
- 30 El término “anomalía cromosómica” se refiere a cualquier variación genética que afecte a todo o a una parte de un cromosoma más grande que un único locus. Las variantes genéticas pueden incluir, pero sin limitarse a, cualquier VNC tal como amplificaciones o delecciones, translocaciones, inversiones y mutaciones. Los ejemplos de anomalías cromosómicas incluyen, pero no se limitan a, síndrome de Down (trisomía 21), síndrome de Edwards (trisomía 18), síndrome de Patau (trisomía 13), síndrome de Klinefelter (XXY), síndrome de triple X, síndrome de XYY, trisomía 8, trisomía 16, síndrome de Turner, translocación robertsoniana, síndrome de DiGeorge y síndrome de Wolf-Hirschhorn.
- 35 Los términos “complementario” o “complementariedad” se usan en referencia a moléculas de ácido nucleico (es decir, una secuencia de nucleótidos) que están relacionadas por reglas de emparejamiento 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 ADN o ARN monocatenario son sustancialmente complementarias cuando los nucleótidos de una hebra, alineados de manera óptima y con inserciones o delecciones de nucleótidos apropiadas, se emparejan con una complementariedad de al menos aproximadamente un 90 % a aproximadamente un 95 %, y más preferentemente con una complementariedad de aproximadamente un 98 % a aproximadamente un 100 %, e incluso más preferentemente con una complementariedad del 100 %. De forma alternativa, existe complementariedad sustancial cuando una hebra de ARN o ADN híbrido en condiciones de hibridación selectiva con 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 salinas de menos de aproximadamente 1 M, más normalmente de menos de aproximadamente 500 mM y preferentemente de menos de aproximadamente 200 mM. En general, las temperaturas de hibridación son al menos de aproximadamente 2 °C a aproximadamente 6 °C menores que las temperaturas de fusión (T_f).
- 40
- 45
- 50
- 55
- 60
- 65
- El término “variación en el número de copias” o “VNC”, como se usa de manera intercambiable en el presente documento, son alteraciones del ADN de un genoma que dan como resultado una célula que tienen un número anómalo de copias de uno o más locus en el ADN. Las VNC que son clínicamente relevantes pueden estar limitadas a un único gen o incluir un conjunto contiguo de genes. Una VNC también puede corresponder a regiones relativamente grandes del genoma que se han deletreado, invertido o duplicado en determinados cromosomas, hasta incluir una o más copias adicionales de un cromosoma completo. El término VNC, como se usa en el presente documento, no se refiere a ninguna información relacionada con la secuencia, sino más bien a la cantidad o “recuentos” de regiones genéticas presentes en una muestra.
- 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 delección, sustitución o inserción de una o más bases durante la secuenciación, así como cambios de nucleótidos que se pueden producir fuera de la secuenciación, tales como síntesis de oligonucleótidos, amplificación y cualquier otro aspecto del ensayo. Estos índices de corrección pueden ser índices independientes que sean secuencias separadas, o se pueden incluir dentro de otras regiones para ayudar en la confirmación de la exactitud

- de las técnicas experimentales usadas, por ejemplo, un índice de corrección puede ser un subconjunto de secuencias usado para la amplificación universal o un subconjunto de nucleótidos de un locus de muestra.
- 5 El término "herramienta de diagnóstico", 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 para llevar a cabo una prueba de diagnóstico o ensayo en una muestra de paciente.
- 10 El término "rasgo de enfermedad" se refiere a un rasgo monogénico o poligénico asociado con una afección patológica, por ejemplo, una enfermedad, un trastorno, un síndrome o una predisposición.
- 15 El término "región genómica", como se usa en el presente documento, se refiere a cualquier región de uno o más locus que normalmente se encuentran de forma contigua en un genoma. Una región genómica puede variar en tamaño hasta incluir un cromosoma completo.
- 20 20 El término "hibridación" designa en general la reacción por la que se produce el emparejamiento de hebras complementarias de ácido nucleico. Normalmente el ADN es bicatenario, y cuando las cadenas se separan, se rehíbridarán en las condiciones apropiadas. Se pueden formar híbridos ADN-ADN, ADN-ARN o ARN-ARN. Se pueden formar entre una hebra corta y una hebra larga que contenga una región complementaria a la corta. También se pueden formar híbridos imperfectos, pero cuanto más imperfectos sean, menos estables serán (y será menos probable que se formen).
- 25 30 El término "locus informativo", como se usa en el presente documento, se refiere a un locus que es homocigótico para una fuente celular y heterocigótico para una segunda fuente celular en un cromosoma particular o parte de un cromosoma consultado con el propósito de determinar una VNC de todo o parte de ese cromosoma. Los locus informativos para su uso en el sistema de ensayo de la invención incluyen locus usados para consultar un cromosoma de referencia, así como locus usados para consultar un cromosoma que es supuestamente aneuploide en una fuente celular. Los locus informativos también pueden distinguir el número de copias de locus en fuentes celulares de diferentes individuos dentro de un único individuo (por ejemplo, detección de células de donante de trasplante en un receptor de trasplante o detección de un ADN fetal dentro de una muestra mixta materna).
- 35 35 El término "locus", como se usa en el presente documento, se refiere a un locus de localización conocida en un genoma.
- 40 45 El término "fuente principal" se refiere a una fuente de ácidos nucleicos en una muestra de un individuo que es representativa del material genómico predominante en ese individuo.
- 50 55 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 su uso en la invención se obtienen a través de medios relativamente no invasivos, por ejemplo, extracción de sangre u otras técnicas estándar para extraer muestras periféricas de un sujeto.
- 60 65 El término "temperatura de fusión" o T_f se define comúnmente como la temperatura a la que una población de moléculas de ácido nucleico bicatenario se disocia en hebras sencillas. La ecuación para calcular la T_f de los ácidos nucleicos es bien conocida en la técnica. Como se indica en las referencias estándar, se puede realizar una estimación sencilla del valor de T_f mediante la ecuación: $T_f = 81,5 + 16,6(\log_{10}[\text{Na}^+])0,41(\%[\text{G}+\text{C}]) - 675/n - 1,0m$, cuando un ácido nucleico se halla en una solución acuosa que tiene concentraciones de cationes de 0,5 M o menos, el contenido de (G+C) está entre el 30 % y el 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*, 3.^a 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 la T_f .
- 70 "Micromatriz" o "matriz" se refiere a un soporte en fase sólida que tiene una superficie, preferentemente pero no exclusivamente una superficie plana o sustancialmente plana, que tiene una matriz de sitios que contienen ácidos nucleicos, de modo que cada sitio de la matriz comprenda copias idénticas o sustancialmente idénticas de oligonucleótidos o polinucleótidos y esté espacialmente definido y no se superponga con otros sitios miembros de la matriz; es decir, los sitios sean espacialmente discretos. La matriz o micromatriz también puede comprender una estructura que se puede consultar no plana con una superficie tal como una microesfera o un pocillo. Los oligonucleótidos o polinucleótidos de la matriz pueden estar unidos de forma covalente al soporte sólido, o pueden estar unidos de forma no covalente. 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 un análisis, tal como, por ejemplo, análisis de secuencia, de una o más moléculas biológicas usando una micromatriz.
- 75 80 85 El término "fuente secundaria" se refiere a una fuente de ácidos nucleicos dentro de un individuo que está presente en cantidades limitadas y que se puede distinguir de la fuente principal debido a diferencias en su composición

genómica y/o expresión. Ejemplos de fuentes secundarias incluyen, pero no se limitan a, células fetales en una mujer embarazada, células cancerosas en un paciente con una neoplasia maligna, células de un órgano de donante en un paciente trasplantado, ácidos nucleicos de un organismo infeccioso en un huésped infectado y similares.

5 El término “muestra mixta”, como se usa en el presente documento, se refiere a cualquier muestra que comprende material genómico libre circulante (por ejemplo, ADN) de dos o más tipos de células de interés, siendo uno de ellos una fuente principal y el otro una fuente secundaria dentro de un único individuo. Los ejemplos de muestras mixtas incluyen una muestra materna (por ejemplo, sangre, suero o plasma materno que comprende ADN tanto materno como fetal) y una muestra somática derivada periféricamente (por ejemplo, sangre, suero o plasma que comprende diferentes tipos de células, por ejemplo, células hematopoyéticas, células mesenquimales y células circulantes de otros sistemas de órganos). Las muestras mixtas incluyen muestras con material genómico tanto de una fuente principal como de una secundaria en un individuo, que pueden ser, por ejemplo, células somáticas normales y atípicas o células que comprenden genomas de dos individuos diferentes, por ejemplo, una muestra con material genómico tanto materno como fetal o una muestra de un paciente trasplantado que comprende células tanto del donante como del receptor.

10 El término “rasgo monogénico”, como se usa en el presente documento, se refiere a cualquier rasgo, normal o patológico, que se asocia con una mutación o un polimorfismo en un único gen. Dichos rasgos incluyen rasgos asociados con una enfermedad, un trastorno o una predisposición causada por una disfunción en un único gen. Los rasgos también incluyen características no patológicas (por ejemplo, presencia o ausencia de moléculas de la superficie celular en un tipo celular específico).

15 El término alelo “no materno” designa un alelo con un polimorfismo y/o una mutación que se encuentra en un alelo fetal (por ejemplo, un alelo con un SNP o una mutación *de novo*) y/o un alelo paterno, pero que no se encuentra en el alelo materno.

20 Por “no polimórfica”, cuando se usa con respecto a la detección de locus seleccionados, se entiende una detección de dicho locus, que puede contener uno o más polimorfismos, pero en la que la detección no es dependiente de la detección del polimorfismo específico dentro de la región. Por tanto, un locus seleccionado puede contener un polimorfismo, pero la detección de la región usando el sistema de ensayo de la invención se basa en la aparición de la región en lugar de en la presencia o ausencia de un polimorfismo particular en esa región.

25 Como se usa en el presente documento, “nucleótido” se refiere a una combinación base-glúcido-fosfato. Los nucleótidos son unidades monoméricas de una secuencia del ácido nucleico (ADN y ARN). El término nucleótido incluye ribonucleósidos trifosfatos ATP, UTP, CTG, GTP y desoxirribonucleósidos trifosfatos tales como dATP, dCTP, dITP, dUTP, dGTP, dTTP o derivados de los mismos. Dichos derivados incluyen, por ejemplo, [α S]dATP, 7-desaza-dGTP y 7-desaza-dATP, y derivados de nucleótidos que confieren resistencia a la nucleasa en la molécula de ácido nucleico que los contiene. El término nucleótido, como se usa en el presente documento, también se refiere a didesoxirribonucleósidos trifosfatos (ddNTP) y sus derivados. Los ejemplos ilustrados de didesoxirribonucleósidos trifosfatos incluyen, pero no se limitan a, ddATP, ddCTP, ddGTP, ddITP y ddTTP.

30 De acuerdo con la presente invención, un “nucleótido” puede estar no marcado o estar marcado de forma detectable por técnicas bien conocidas. Los marcadores fluorescentes y su unión a oligonucleótidos se describen en muchas revisiones, incluyendo Haugland, *Handbook of Fluorescent Probes and Research Chemicals*, 9.^a ed., Molecular Probes, Inc., Eugene OR (2002); Keller y Manak, *DNA Probes*, 2.^a 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 divulan 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 marcaje también se puede llevar a cabo con puntos cuánticos, como se divulga en las siguientes patentes y publicaciones de patente: 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; 55 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)aminonáftaleno-1-sulfónico (EDANS). Los ejemplos específicos de nucleótidos marcados de forma fluorescente 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, tetrametil-rodamina-

6-dUTP, IR770-9-dATP, fluoresceína-12-ddUTP, fluoresceína-12-UTP y fluoresceína-15-2'-dATP disponibles de Boehringer Mannheim, Indianápolis, IN; y nucleótidos marcados en 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, Rhodamine Green-5-UTP, Rhodamine 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.

El término "oligonucleótidos", como se usa en el presente documento, se refiere a oligómeros lineales de monómeros de ácido nucleico naturales o modificados, incluyendo desoxirribonucleótidos, ribonucleótidos, formas anoméricas de los mismos, monómeros de ácido peptidónucleico (PNA), monómeros de ácido nucleico bloqueado (LNA) y similares, o una combinación de los mismos, que se pueden unir específicamente a un polinucleótido monocatenario por medio de un patrón regular de interacciones monómero con monómero, tal como el tipo de emparejamiento de bases de Watson-Crick, apilamiento de bases, tipos de emparejamiento de bases de Hoogsteen o Hoogsteen inverso o similares. Normalmente, los monómeros se unen por enlaces fosfodiéster o análogos de los mismos para formar oligonucleótidos que varían en tamaño de 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 por el procedimiento de fosforamidita descrito por Beaucage y Carruthers (Tetrahedron Lett., 22:1859-1862 (1981)), o por el procedimiento de triéster de acuerdo con Matteucci, *et al.* (J. Am. Chem. Soc., 103:3185 (1981)), o por otros procedimientos químicos tales como el uso de un sintetizador de oligonucleótidos automatizado comercial.

El término "rasgo poligénico", como se usa en el presente documento, se refiere a cualquier rasgo, normal o patológico, que se asocia con una mutación o un polimorfismo en más de un único gen. Dichos rasgos incluyen rasgos asociados con una enfermedad, un trastorno, un síndrome o una predisposición causada por una disfunción en dos o más genes. Los rasgos también incluyen características no patológicas asociadas con la interacción de dos o más genes.

Como se usa en el presente documento, el término "polimerasa" se refiere a una enzima que une nucleótidos individuales conjuntamente en una hebra larga, usando otra hebra como molde. Existen dos tipos generales de polimerasa: las ADN polimerasas, que sintetizan ADN, y las ARN polimerasas, que sintetizan ARN. Dentro de estas dos clases, existen numerosos subtipos de polimerasas, dependiendo del tipo de ácido nucleico que pueden funcionar como molde y de qué tipo de ácido nucleico se forma.

Como se usa en el presente documento, "reacción en cadena de la polimerasa" o "PCR" se refiere a una técnica para replicar una parte específica de ADN seleccionado *in vitro*, incluso en presencia de un exceso de ADN no específico. Se añaden cebadores al ADN seleccionado, donde los cebadores inician la copia del ADN seleccionado usando nucleótidos y, típicamente, Taq polimerasa o similares. Realizando termociclados, el ADN seleccionado se desnaturiza y copia repetidamente. Se puede amplificar una única copia del ADN seleccionado, incluso si se mezcla con otro ADN aleatorio, para obtener mil 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 partes de ADN adaptadas. En algunos casos, se pueden usar procedimientos de amplificación lineal como alternativa a la PCR.

El término "polimorfismo", como se usa en el presente documento, se refiere a cualquier cambio genético o variante de secuencia en un locus, incluyendo, pero sin limitarse a, polimorfismos mononucleotídicos (SNP), diferencias de metilación, repeticiones cortas en tandem (STR), polimorfismos de un único gen, mutaciones puntuales, repeticiones de trinucleótidos, indels y similares.

En general, un "cebador" es un oligonucleótido usado, por ejemplo, para ceder la extensión, unión 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 del cebador usadas en determinadas reacciones de secuenciación. También se puede usar un cebador en técnicas de hibridación como un medio para proporcionar complementariedad de un locus a un oligonucleótido de captura para la detección de un locus específico.

El término "herramienta de investigación", como se usa en el presente documento, se refiere a cualquier composición o ensayo de la invención usado para investigación científica, de naturaleza académica o comercial, incluyendo el desarrollo de tratamientos farmacéuticos y/o biológicos. No se pretende que las herramientas de investigación de la invención sean terapéuticas o estén sujetas a autorización; en su lugar, se pretende que las herramientas de investigación de la invención faciliten la investigación y ayuden en dichas actividades de desarrollo, incluyendo cualquier actividad realizada con la intención de generar información para respaldar una presentación del expediente de registro.

El término "índice de muestra" se refiere en general a una serie de nucleótidos exclusivos (es decir, cada índice de muestra es exclusivo para una muestra en un sistema de ensayo multiplexado para el análisis de múltiples muestras). Por tanto, se puede usar el índice de muestra para ayudar en la identificación del locus para la multiplexación de diferentes muestras en un único recipiente de reacción, de modo que se pueda identificar cada muestra en base a su índice de muestra. En un aspecto preferente, existe un índice de muestra exclusivo para cada

muestra en un conjunto de muestras, y las muestras se agrupan durante la secuenciación. Por ejemplo, si se agrupan doce muestras en una única reacción de secuenciación, existen al menos doce índices exclusivos de muestra de modo que cada muestra se marque de forma exclusiva.

- 5 El término "locus seleccionado", como se usa en el presente documento, se refiere a un locus correspondiente a un locus consultado, por ejemplo, para determinar el número de copias, la presencia o ausencia de uno o más polimorfismos, la presencia o ausencia de un organismo infeccioso, etc. Dichos locus seleccionados se pueden aislar y amplificar directamente de la muestra para su detección, por ejemplo, basándose en la hibridación y/u otras técnicas basadas en la secuencia, o se pueden amplificar usando la muestra como molde antes de la detección de la secuencia. Las regiones de ácidos nucleicos para su uso en los sistemas de ensayo de la presente invención se pueden seleccionar en base a la variación en el nivel de ADN entre individuos, en base a la especificidad para un cromosoma particular, en base al contenido de CG y/o las condiciones de amplificación requeridas de los locus seleccionados, u otras características que resultarán evidentes para un experto en la técnica al leer la presente divulgación.
- 10
- 15 Los términos "secuenciación", "determinación de secuencias" 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 las bases nucleotídicas en un ácido nucleico.
- 20 El término "contribución de fuente", como se usa en el presente documento, se refiere a la contribución relativa de dos o más fuentes de ácidos nucleicos dentro de un individuo. La contribución de una única fuente se determina en general como porcentaje de los ácidos nucleicos procedentes en una muestra, aunque se puede usar cualquier medición relativa.
- 25 Los términos "se une específicamente", "unión específica" y similares, como se usan en el presente documento, cuando se refieren a un compañero de unión (por ejemplo, una sonda o cebador de ácido nucleico, anticuerpo, etc.), dan como resultado la generación de una señal positiva estadísticamente significativa en las condiciones de ensayo designadas. Típicamente, la interacción dará posteriormente como resultado una señal detectable que es al menos dos veces la desviación estándar de cualquier señal generada como resultado de interacciones no deseadas (fondo).
- 30
- 35 El término "estado", como se usa en el presente documento en relación con un gen, se refiere al estado de la secuencia de los alelos de un gen particular, incluyendo las regiones codificantes y las regiones no codificantes que afectan a la traducción y/o expresión de proteínas de ese gen. El estado de un gen asociado con una enfermedad autosómica dominante tal como la acondroplasia (por ejemplo, el gen que codifica el receptor del factor de crecimiento de fibroblastos) o la enfermedad de Huntington (por ejemplo, el gen de Huntington), o para una enfermedad ligada al cromosoma X en el caso de un feto de sexo masculino, se puede clasificar como afectado, es decir, un alelo posee la mutación o mutaciones causantes de la enfermedad o del trastorno, o no afectado, es decir, ambos alelos carecen de dicha mutación o mutaciones. El estado de un gen asociado con una enfermedad autosómica recesiva o de un gen materno asociado con un trastorno recesivo ligado al cromosoma X se puede clasificar como afectado, es decir, ambos alelos poseen la mutación o mutaciones causantes de la enfermedad o del trastorno; portador, es decir, un alelo posee la mutación o mutaciones causantes de la enfermedad o del trastorno; o no afectado, es decir, ambos alelos carecen de dicha mutación o mutaciones. El estado de un gen también puede indicar la presencia o ausencia de un alelo particular asociado con el riesgo de desarrollar una enfermedad poligénica, por ejemplo, un polimorfismo que protege contra una enfermedad o un trastorno en particular o un polimorfismo asociado con un mayor riesgo de desarrollar una enfermedad o un trastorno en particular.
- 40
- 45

DESCRIPCIÓN DETALLADA DE LA INVENCIÓN

- 50 Los sistemas y procedimientos de ensayo descritos en el presente documento pueden emplear, a menos que se indique de otro modo, técnicas y descripciones convencionales de biología molecular (incluyendo técnicas recombinantes), biología celular, bioquímica, tecnología de micromatrices y de secuenciación, que estén dentro de la habilidad de los expertos en la técnica. Dichas técnicas convencionales incluyen síntesis de matrices poliméricas, hibridación y unión 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 en el presente documento. Sin embargo, por supuesto, también se pueden usar procedimientos convencionales equivalentes. Dichas técnicas y descripciones convencionales se pueden encontrar en manuales de laboratorio estándar, tales 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* (4.^a ed.) W.H. Freeman, Nueva York (1995); Gait, "Oligonucleotide Synthesis: A Practical Approach" IRL Press, Londres (1984); Nelson y Cox, *Lehninger, Principles of Biochemistry*, 3.^a ed., W. H. Freeman Pub., Nueva York (2000); y Berg et al., *Biochemistry*, 5.^a ed., W.H. Freeman Pub., Nueva York (2002). Antes de que se describan las presentes composiciones, herramientas de investigación y
- 55
- 60
- 65

procedimientos, se ha de entender que la presente invención no se limita a los procedimientos, composiciones, dianas y usos específicos descritos, ya que como tales, por supuesto, pueden variar. También se ha de entender que la terminología usada en el presente documento es solo para el propósito de describir aspectos particulares y no se pretende que limite el alcance de la presente invención, que solo se limitará por las reivindicaciones adjuntas.

5 Cabe destacar que, como se usa en el presente documento y en las reivindicaciones adjuntas, las formas en singular "un", "una" y "el/la" incluyen referentes al plural a menos que el contexto lo indique claramente de otro modo. Por tanto, por ejemplo, la referencia a "un locus" se refiere a una, más de una o mezclas de dichas regiones, y la referencia a "un ensayo" incluye la referencia a etapas y procedimientos equivalentes conocidos por los expertos en la técnica, etc.

10 15 Cuando se proporciona un intervalo de valores, se ha de entender que cada valor intermedio entre el límite superior e inferior de ese intervalo y cualquier otro valor establecido o intermedio en ese intervalo establecido se engloba dentro de la invención. Cuando el intervalo establecido incluye los límites superior e inferior, los intervalos que excluyan cualquiera de esos límites incluidos también se incluyen en la invención.

20 25 Salvo que se indique expresamente, se pretende que los términos usados en el presente documento tengan el significado simple y ordinario como entienden los expertos en la técnica. Se pretende que las siguientes definiciones ayuden al lector en el entendimiento de la presente invención, pero no se pretende que varíen o de otro modo limiten el significado de dichos términos a menos que se indique específicamente.

25 En la siguiente descripción se exponen numerosos detalles específicos para proporcionar un entendimiento más completo 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, no se han descrito rasgos característicos y procedimientos bien conocidos para los expertos en la técnica para evitar eclipsar la invención.

La invención en general

30 La presente invención proporciona procedimientos de un único ensayo con la capacidad de detectar variaciones en el número de copias, polimorfismos y ácidos nucleicos asociados con estados patológicos (por ejemplo, ácidos nucleicos de agentes patógenos o asociados con cáncer, diabetes, enfermedad de Alzheimer y similares) en una muestra mixta de un único individuo. El ensayo permite identificar una variación genética en una o más fuentes secundarias presentes en una muestra mixta usando información sobre el número de copias de los locus 35 seleccionados en la muestra e información sobre el porcentaje de contribución de ácidos nucleicos de la fuente principal y una o más fuentes secundarias en la muestra. Estos procedimientos son útiles para cualquier muestra mixta que contenga material genómico (por ejemplo, ADN) de una fuente principal y una secundaria que estén presentes en un único individuo.

40 45 50 El uso de locus seleccionados en los procedimientos de ensayo de la invención proporciona una detección directa de los locus para la determinación de la variación en el número de copias en una o más fuentes dentro de la muestra mixta. Una clara ventaja de la invención es que los locus seleccionados correspondientes a variaciones en el número de copias y/o polimorfismos se pueden analizar además usando una variedad de técnicas de detección y cuantificación, incluyendo, pero sin limitarse a, técnicas de hibridación, PCR digital y técnicas de determinación de secuenciación de alto rendimiento. Las sondas se pueden diseñar para cualquier número de locus seleccionados para cualquier cromosoma. Aunque la amplificación anterior a la identificación y cuantificación de los locus seleccionados no es obligatoria, se puede usar una amplificación limitada de la muestra mixta antes de la detección para expandir el número total de ácidos nucleicos dentro de los materiales de partida.

55 60 65 La figura 1 es un diagrama de flujo simplificado de las etapas generales utilizadas en los sistemas de ensayo de la invención. La figura 1 muestra el procedimiento 100, en el que en una primera etapa 110 se proporciona una muestra mixta de ácidos nucleicos para su análisis. La muestra mixta se puede preparar a partir de prácticamente cualquier muestra, ya que dichas técnicas son conocidas para los expertos en la técnica (véanse, por ejemplo, Tietz Textbook of Clinical Chemistry and Molecular Diagnostics, 4.^a ed., Capítulo 2, Burtis, C. Ashwood E. y Bruns, D. eds. (2006); Chemical Weapons Convention Chemicals Analysis: Sample Collection, Preparation and Analytical Methods, Mesilaakso, M., ed., (2005); Pawliszyn, J., Sampling and Sample Preparation for Field and Laboratory, (2002); Venkatesh Iyengar, G., et al., Element Analysis of Biological Samples: Principles and Practices (1998); Drielak, S., Hot Zone Forensics: Chemical, Biological, and Radiological Evidence Collection (2004); Wells, D., High Throughput Bioanalytical Sample Preparation (Progress in Pharmaceutical and Biomedical Analysis) (2002)). Dependiendo del tipo de muestra mixta elegida, se pueden realizar etapas adicionales de procesamiento y/o purificación para obtener fragmentos de ácido nucleico de pureza o tamaño deseado, usando procedimientos de procesamiento que incluyen, pero no se limitan a, homogeneización ultrasónica, nebulización, purificación en gel, sistemas de purificación por PCR, escisión por nucleasa o una combinación de estos procedimientos. En un aspecto preferente se usan muestras que comprenden ADN libre circulante (ADNlc).

- En la etapa 120 se introduce un primer conjunto de oligonucleótidos de secuencia fija en la muestra mixta de ácidos nucleicos, en condiciones que permitan que el primer conjunto de oligonucleótidos de secuencia fija híbride con la muestra mixta de ácidos nucleicos. El primer conjunto de oligonucleótidos de secuencia fija comprende secuencias de ácido nucleico que son complementarias a uno o más locus seleccionados en la muestra mixta, que, como se describirá en detalle en el presente documento, son útiles para determinar variaciones en el número de copias y/o anomalías cromosómicas. Las secuencias de ácido nucleico que pueden determinar variaciones en el número de copias y/o anomalías cromosómicas incluyen secuencias que permiten identificar anomalías cromosómicas tales como amplificaciones o delecciones, aneuploidías, translocaciones o inversiones.
- 5
- 10 En la etapa 130 se introduce un segundo conjunto de oligonucleótidos de secuencia fija en la muestra mixta de ácidos nucleicos y el primer conjunto de oligonucleótidos de secuencia fija en condiciones que permitan que el segundo conjunto de oligonucleótidos de secuencia fija híbride con la muestra mixta de ácidos nucleicos. El segundo conjunto de oligonucleótidos de secuencia fija comprende secuencias de ácido nucleico que son complementarias a uno o más locus seleccionados en la muestra mixta que pueden detectar polimorfismos. Se pueden incluir opcionalmente etapas de lavado entre las etapas 120 y 130, y 130 y 140.
- 15
- 20 En la etapa 140 se unen el primer y segundo conjuntos de oligonucleótidos de secuencia fija que han hibridado con regiones adyacentes de los locus seleccionados en la muestra mixta, y en la etapa 150 se amplifican los oligonucleótidos unidos. A continuación, los oligonucleótidos unidos y amplificados se detectan y analizan, lo que permite determinar variaciones en el número de copias o anomalías cromosómicas e identificar polimorfismos en la etapa 160.
- 25
- 30 Los conjuntos de ácidos nucleicos de secuencia fija están diseñados para hibridar con al menos dos regiones separadas en un locus seleccionado. En aspectos preferentes, se usan dos o más oligonucleótidos separados para hibridar con estas regiones para proporcionar ácidos nucleicos adyacentes complementarios al locus seleccionado. En algunos aspectos, sin embargo, se puede usar una única sonda que comprende dos o más regiones no adyacentes distintas que son complementarias a los locus seleccionados, incluyendo sondas precirculares tales como las denominadas "sondas candado" o "sondas de inversión molecular (MIP)".
- 35
- 40 La presente invención proporciona un sistema mejorado con respecto a técnicas más aleatorias tales como secuenciación masivamente paralela, secuenciación indiscriminada y el uso de PCR digital aleatoria que han sido usadas por otros para detectar VNC. Los enfoques mencionados anteriormente se basan en la secuenciación de todos o una población estadísticamente significativa de fragmentos de ADN en una muestra, seguida de cartografía de estos fragmentos o asociación de otro modo de los fragmentos a sus cromosomas adecuados. A continuación, los fragmentos identificados se comparan entre sí o con alguna otra referencia (por ejemplo, composición cromosómica normal) para determinar las VNC en cromosomas particulares. Estos procedimientos son inherentemente ineficaces en comparación con la presente invención, ya que los cromosomas principales de interés solo constituyen una minoría de los datos que se generan a partir de la detección de dichos fragmentos de ADN en las muestras mixtas.
- 45
- 50 Los ensayos de la presente invención proporcionan una detección dirigida de locus seleccionados, lo que proporciona tanto información sobre el contenido del locus seleccionado (es decir, la presencia de una región polimórfica) como información sobre la frecuencia del locus seleccionado en una muestra (con o sin detección de cualquier supuesto polimorfismo en esa región). Este rasgo característico clave proporciona la capacidad de detectar tanto el número de copias de los locus seleccionados como la presencia o ausencia de polimorfismos en los locus seleccionados en forma de un único conjunto de datos obtenido mediante la realización de un ensayo multiplexado de la invención.
- 55
- 60 Las técnicas que son dependientes de un muestreo muy amplio de ADN en una muestra proporcionan una cobertura muy amplia del ADN analizado, pero, de hecho, están muestreando el ADN contenido dentro de una muestra en una base de 1X o menos (es decir, submuestreo). Por el contrario, la amplificación de los locus seleccionados usados en los presentes ensayos proporciona una profundidad de cobertura de los locus particulares de interés y proporciona un "supermuestreo" de dichos locus seleccionados con una cobertura de secuencia promedio de preferentemente 2X o más, más preferentemente una cobertura de secuencia de 100X o más, incluso más preferentemente una cobertura de secuencia de 1000X o más de los locus seleccionados (incluyendo los de una o más fuentes secundarias) presentes en la muestra mixta inicial.
- Una clara ventaja de la invención es que los productos de amplificación resultantes de los ensayos se pueden analizar además usando una variedad de técnicas de detección y cuantificación, incluyendo, pero sin limitarse a, técnicas de hibridación, PCR digital y técnicas de determinación de secuenciación de alto rendimiento.
- 65 Los procedimientos de la invención proporcionan un uso de datos más eficaz y económico, y la gran mayoría de las secuencias analizadas después de la amplificación de la muestra dan como resultado información afirmativa sobre la identidad de secuencia y la frecuencia de los locus seleccionados en la muestra mixta. Por tanto, a diferencia de las técnicas que se basan en la secuenciación masivamente paralela o el "recuento" digital aleatorio de regiones cromosómicas y la posterior identificación de datos relevantes de dichos recuentos, el sistema de ensayo de la

invención proporciona un uso mucho más eficaz de los datos recopilados que los enfoques aleatorios enseñados por otros en la técnica.

Procedimientos de ensayo

- 5 Los procedimientos de ensayo de la invención utilizan un esquema general como se describe anteriormente, aunque se pueden emplear muchas configuraciones y variaciones diferentes, algunas de las cuales se describen a continuación y muchas de las cuales se ejemplifican en la patente de EE. UU. con el número de serie 61/371605 presentada el 6 de agosto de 2010.
- 10 La figura 2 ilustra un primer esquema general de un sistema de ensayo de la invención basado en unión. Los oligonucleótidos de secuencia fija 201, 203 comprenden regiones de cebadores universales 209 y 211, respectivamente, y regiones complementarias 205 y 207 a los locus seleccionados, respectivamente. Sin embargo, el sistema de ensayo de la figura 2 emplea además una región de índice de muestra 221 en el primer oligonucleótido de secuencia fija 201. En determinados aspectos, todas o una parte de las secuencias de los locus seleccionados se detectan directamente usando las técnicas descritas, por ejemplo, mediante técnicas de determinación de secuencias o hibridación. En el ejemplo de la figura 2, un índice de muestra se asocia con el primer oligonucleótido de secuencia fija 201. La detección de los índices puede identificar una secuencia de una muestra específica en un sistema de ensayo altamente multiplexado.
- 15 20 En la etapa 202, los oligonucleótidos de secuencia fija 201, 203 se introducen en la etapa 202 en la muestra mixta 200 y se dejan que se unan específicamente al locus 215 seleccionado. Tras la hibridación, los oligonucleótidos de secuencia fija no hibridados se separan preferentemente del resto de la muestra genética (por ejemplo, por lavado, no mostrado). A continuación, se introduce un oligonucleótido puente y se deja que hibride en la etapa 204 con la región del locus 215 situada entre el primer 201 y el segundo 203 oligonucleótido de secuencia fija. En la etapa 206, los oligonucleótidos unidos se unen para crear un ácido nucleico contiguo que abarca y es complementario al locus de interés. En determinados aspectos de la invención, los oligonucleótidos puente presentan una longitud de entre 2 y 45 nucleótidos. En un aspecto específico, los oligonucleótidos puente presentan una longitud de entre 3 y 9 nucleótidos. Aún en otro aspecto específico, los oligonucleótidos presentan una longitud de entre 10 y 30 nucleótidos.
- 25 30 Despues de la unión, el producto de unión se eluye del molde de ADNg. En la etapa 208 se introducen los cebadores universales 217, 219 para amplificar el primer y segundo oligonucleótidos de secuencia fija unidos para crear 210 productos de amplificación 223 que comprenden la secuencia del locus de interés. Estos productos 223 se aislan, detectan, identifican y cuantifican para proporcionar información con respecto a la presencia y cantidad de los locus seleccionados en la muestra mixta. Preferentemente, los productos de amplificación se detectan y cuantifican a través de la determinación de secuencias. En aspectos específicos, es deseable determinar las secuencias tanto del índice de muestra como de los productos de amplificación, por ejemplo, para identificar la muestra así como el locus. Los índices contemplados en la invención, tales como el índice de muestra mostrado aquí, se pueden asociar con los primeros oligonucleótidos de secuencia fija, los segundos oligonucleótidos de secuencia fija o ambos. De forma alternativa o adicional, los índices se pueden asociar con cebadores que se usan para amplificar el primer y segundo oligonucleótidos de secuencia fija unidos, lo que también sirve para incorporar índices en los productos de amplificación.
- 35 40 45 En aspectos preferentes y como se muestra en la figura 2, los índices representativos de la muestra mixta de la que se puede aislar un ácido nucleico se usan para identificar la fuente de los locus seleccionados en un sistema de ensayo multiplexado. En dichos aspectos, los ácidos nucleicos se identifican de forma exclusiva con el índice de muestra. A continuación, los oligonucleótidos identificados de forma exclusiva se pueden combinar, antes de la secuenciación, en un único recipiente de reacción con ácidos nucleicos de otras muestras mixtas. En dicho caso, los datos de secuenciación se segregan por el índice de muestra exclusivo para determinar la frecuencia de cada locus diana para cada muestra mixta y para determinar si existe una anomalía cromosómica en una muestra individual.
- 50 55 60 En los aspectos de la invención en los que se usan índices de muestra, los oligonucleótidos de secuencia fija se diseñan preferentemente de modo que los índices de muestra que comprenden información de identificación estén localizados entre las regiones de cebadores universales 209 y 211 y las regiones complementarias 205 y 207 a los locus seleccionados en la muestra. De forma alternativa, los índices y las secuencias de amplificación universal se pueden añadir al primer y segundo oligonucleótidos de secuencia fija unidos (y al oligonucleótido puente, si está presente) incluyendo estos índices en los cebadores usados para amplificar los productos de unión para muestras separadas. En cualquier caso, los índices preferentemente se codifican corriente arriba de las secuencias específicas del locus, pero corriente abajo de los cebadores universales de modo que se conserven tras la amplificación.
- 65 La figura 3 ejemplifica procedimientos del sistema de ensayo en los que se emplean uno o más oligonucleótidos puente y ejemplifica cómo se pueden detectar e identificar polimorfismos. En la figura 3 se usan dos conjuntos de oligonucleótidos de secuencia fija que comprenden sustancialmente los mismos cebadores universales 309, 311 y regiones específicas de secuencia 305, 307, pero comprenden diferentes índices de muestra 321, 323 en los

oligonucleótidos de secuencia fija del conjunto, donde los diferentes índices corresponden a diferentes secuencias de bases para el polimorfismo mononucleotídico presente en una muestra particular. Las reacciones de unión se llevan a cabo con material de la misma muestra mixta 300, pero en tubos separados con los diferentes conjuntos de oligonucleótidos específicos de alelo. Los oligonucleótidos puente 313, 333 correspondientes a dos posibles nucleótidos para este SNP en los locus seleccionados se usan para detectar el locus seleccionado en cada reacción de unión. Dos índices de alelo 321, 323 que son indicativos de los alelos polimórficos particulares se incorporan a los productos de amplificación, de modo que la determinación de secuencias de la secuencia real del primer y segundo oligonucleótidos y de los oligonucleótidos puente unidos no sea absolutamente necesaria, aunque las secuencias de los productos de unión completos todavía se pueden determinar para identificar y/o proporcionar la confirmación.

Cada uno de los oligonucleótidos de secuencia fija comprende una región complementaria 305, 307 al locus seleccionado y las regiones de cebadores universales 309, 311 usadas para amplificar los diferentes locus seleccionados después de la selección y/o el aislamiento iniciales de los locus seleccionados de la muestra mixta.

15 Las regiones de cebadores universales se localizan en los extremos de los oligonucleótidos de secuencia fija 301, 303 y 323 que flanquean los índices y las regiones complementarias al ácido nucleico de interés, conservando por tanto las secuencias específicas de ácido nucleico y los índices de muestra en los productos de cualquier procedimiento de amplificación universal. Los oligonucleótidos de secuencia fija 301, 303, 323 se introducen en la etapa 302 en una parte alícuota de la muestra genética 300 y se deja que se unan específicamente a los locus seleccionadas 315 o 325. Tras la hibridación, los oligonucleótidos de secuencia fija no hibridados se separan preferentemente del resto de la muestra genética, por ejemplo, por lavado (no mostrado).

20 Se introducen los oligonucleótidos puente correspondientes a un SNP A/T 313 o un SNP G/C 333 y se deja que se unan en la etapa 304 a la región del locus seleccionado 315 o 325 situada entre la primera 305 y segunda 307 regiones complementarias al ácido nucleico de los oligonucleótidos de secuencia fija. De forma alternativa, los oligonucleótidos puente 313, 333 se pueden introducir simultáneamente en la muestra con los oligonucleótidos de secuencia fija. En la etapa 306, los oligonucleótidos unidos se unen en la mezcla de reacción única para crear un ácido nucleico contiguo que abarca y es complementario al locus seleccionado.

25 30 Tras la unión, las reacciones separadas se pueden combinar preferentemente para las etapas de amplificación universal y detección. Los cebadores universales 317, 319 se introducen en las reacciones combinadas en la etapa 308 para amplificar las regiones de molde unidas y crear en la etapa 310 productos unidos del primer y segundo oligonucleótidos de secuencia fija y de los oligonucleótidos puente 327, 329 que comprenden la secuencia del locus seleccionado que representa ambos SNP en el locus seleccionado. Estos productos de unión 327, 329 se detectan y 35 cuantifican a través de la determinación de secuencias del producto de unión, a través del índice de muestra y/o la región del producto que contiene el SNP en el locus seleccionado.

40 En una configuración alternativa de los procedimientos de los sistemas de ensayo de la invención, el oligonucleótido puente puede hibridar con una región que no es directamente adyacente a la región complementaria a uno o ambos de los oligonucleótidos de secuencia fija, y es necesaria una etapa intermedia que requiere la extensión de uno o más de los oligonucleótidos antes de la unión. Por ejemplo, como se ilustra en la figura 4, cada conjunto de oligonucleótidos contiene preferentemente dos oligonucleótidos de secuencia fija 401, 403 y uno o más oligonucleótidos puente 413. Cada uno de los oligonucleótidos de secuencia fija comprende una región complementaria al locus seleccionado 405, 407 y secuencias de cebador, preferentemente secuencias de cebador universal, 409, 411, es decir, regiones de oligonucleótido complementarias a los cebadores universales. Las secuencias de cebador 409, 411 se localizan en o cerca de los extremos de los oligonucleótidos de secuencia fija 401, 403 y, por tanto, conservan las secuencias específicas de ácido nucleico en los productos de cualquier procedimiento de amplificación universal. Los oligonucleótidos de secuencia fija 401, 403 se introducen en la etapa 402 en la muestra mixta 400 y se deja que se unan específicamente a partes complementarias del locus de interés 415. Tras la hibridación, los oligonucleótidos de secuencia fija no hibridados se separan preferentemente del resto de la muestra genética (no mostrado).

45 50 55 A continuación, se introduce el oligonucleótido puente en la etapa 404 y se deja que se una a la región del locus seleccionado 415 situada entre el primer 401 y segundo 403 oligonucleótidos de secuencia fija. De forma alternativa, el oligonucleótido puente se puede introducir simultáneamente con los oligonucleótidos de secuencia fija. En este aspecto ejemplar, el oligonucleótido puente hibrida con una región directamente adyacente a la región del primer oligonucleótido de secuencia fija 405, pero está separado en uno o más nucleótidos de la región complementaria del segundo oligonucleótido de secuencia fija 407. Tras la hibridación de los oligonucleótidos de secuencia fija y puente, el oligonucleótido puente 413 se extiende en la etapa 406, por ejemplo, usando una polimerasa y dNTP, para llenar el hueco existente entre el oligonucleótido puente 413 y el segundo oligonucleótido de secuencia fija 403. Tras la extensión, los oligonucleótidos unidos se unen en la etapa 408 para crear un ácido nucleico contiguo que abarca y es complementario al locus de interés 415. Despues de la unión, se introducen los cebadores universales 417, 419 en la etapa 410 para amplificar el primer y segundo oligonucleótidos y los oligonucleótidos puente unidos para crear en la etapa 412 los productos de amplificación 423 que comprenden la secuencia del locus de interés seleccionado. 60 65 Estos productos de amplificación 423 se aíslan, detectan y cuantifican opcionalmente para proporcionar información sobre la presencia y cantidad del/de los locus seleccionado(s) en la muestra mixta.

Detección de variaciones en el número de copias

- 5 La presente invención proporciona procedimientos para identificar una variación en el número de copias en uno o más locus y la presencia o ausencia de uno o más polimorfismos. Esto se puede realizar usando procedimientos de amplificación para la identificación de locus correspondientes a locus en cromosomas específicos de interés correspondientes a secuencias de un único gen.
- 10 Los sistemas de ensayo utilizan sondas de ácido nucleico diseñadas para identificar, y preferentemente para aislar, locus seleccionados en una muestra mixta. Algunas de las sondas identifican secuencias de interés en locus seleccionados consultados para determinar el número de copias (es decir, la frecuencia de locus), y otras sondas identifican secuencias que corresponden a polimorfismos de interés (es decir, contenido de locus) en ácidos nucleicos correspondientes a una fuente principal o fuente secundaria en un muestra mixta.
- 15 En aspectos específicos, los procedimientos de ensayo de la invención emplean una o más etapas de amplificación selectiva (por ejemplo, usando uno o más cebadores que hibridan específicamente con un locus seleccionado) para aislar, amplificar o analizar sustancialmente todos los locus seleccionados analizados. Esto contrasta directamente con el enfoque de amplificación aleatoria usado por otros que emplean, por ejemplo, secuenciación masivamente paralela, ya que dichas técnicas de amplificación implican, en general, la amplificación aleatoria de todo o una parte sustancial del genoma. Además, la muestra inicial se puede enriquecer opcionalmente usando procedimientos tales como amplificación general para incrementar el número de copias de ácidos nucleicos en la muestra mixta. Preferentemente, las etapas de hibridación, unión y amplificación usadas para identificar los locus de interés se realizan directamente en la muestra mixta.
- 20
- 25 En un aspecto general, el usuario de la invención analiza múltiples locus seleccionados en diferentes cromosomas. Cuando se analizan múltiples locus para una muestra, un modo de realización preferente es amplificar todos los locus seleccionados para cada muestra en un recipiente de reacción. Las frecuencias de los múltiples locus seleccionados se usan para determinar si existe una anomalía cromosómica y, opcionalmente, para identificar polimorfismos en los locus seleccionados.
- 30
- 35 En aspectos preferentes, se pueden amplificar múltiples locus seleccionados de dos o más muestras en un único recipiente de reacción, y la información se puede analizar simultáneamente en un único conjunto de datos, por ejemplo, a través de la determinación de secuencias. A continuación, los datos resultantes se analizan para separar los resultados de las diferentes muestras y se usan para determinar la presencia o ausencia de VNC y la contribución de las fuentes en muestras individuales.
- 40 En un aspecto, las anomalías cromosómicas se identifican en el procedimiento de ensayo de la invención usando múltiples locus seleccionados en múltiples cromosomas, y la frecuencia de los locus seleccionados en los múltiples cromosomas se compara para identificar un incremento de la probabilidad de aneuploidía en base a las proporciones de las frecuencias de los múltiples locus en los cromosomas. La normalización o estandarización de las frecuencias se puede realizar para uno o más locus seleccionados.
- 45
- 50 En otro aspecto, el sistema de ensayo suma las frecuencias de los locus seleccionados en dos o más cromosomas y, a continuación, compara la suma de los locus seleccionados en un cromosoma con la de otro cromosoma para determinar si existe una aneuploidía cromosómica. En otro aspecto, el sistema de ensayo analiza subconjuntos de frecuencias de locus seleccionados en dos o más cromosomas para determinar si existe una aneuploidía cromosómica para uno de los dos cromosomas. La comparación se puede realizar dentro del mismo cromosoma o en diferentes cromosomas.
- 55
- 60 En determinados aspectos, los datos usados para determinar la frecuencia de los locus seleccionados pueden excluir valores atípicos que parece que se deben a un error experimental, o que tienen niveles elevados o reducidos en base a un sesgo genético idiopático dentro de una muestra particular. En un ejemplo, los datos usados para la suma pueden excluir regiones de ADN con una frecuencia particularmente elevada en una o más muestras. En otro ejemplo, los datos usados para la suma pueden excluir locus seleccionados que se encuentran en una abundancia particularmente baja en una o más muestras.
- 65
- 70 En otro aspecto, se pueden elegir subconjuntos de locus seleccionados para proporcionar un resultado estadísticamente significativo al determinar si existe una anomalía cromosómica. Se pueden realizar múltiples análisis de diferentes subconjuntos de locus seleccionados dentro de una muestra mixta para proporcionar mayor potencia estadística. Por ejemplo, si existen 100 locus seleccionados para el cromosoma 21 y 100 locus seleccionados para el cromosoma 18, se podrían realizar una serie de análisis que evalúen menos de 100 regiones para cada uno de los cromosomas. En este ejemplo, los locus seleccionados no se excluyen selectivamente.
- 75 La cantidad de diferentes ácidos nucleicos detectables en determinados cromosomas puede variar dependiendo de una serie de factores, incluyendo representación general de locus en diferentes fuentes celulares en muestras mixtas, tasas de degradación de los diferentes ácidos nucleicos que representan diferentes locus en muestras

mixtas, procedimientos de preparación de muestras y similares. Por tanto, en otro aspecto, la cantidad de locus particulares en un cromosoma se suma para determinar la cantidad de locus para diferentes cromosomas en la muestra. Las frecuencias de los locus se suman para un cromosoma en particular y las sumas de los locus se usan para determinar la presencia de aneuploidía. En este aspecto de la invención, se suman las frecuencias de los locus individuales en cada cromosoma y, a continuación, se compara la suma de los locus en un cromosoma con la de uno o más de otros cromosomas para determinar si existe una anomalía cromosómica.

5 Los ácidos nucleicos analizados usando los procedimientos de ensayo de la invención preferentemente se amplifican selectivamente y opcionalmente se aislan de la muestra mixta usando cebadores específicos para el locus de interés (por ejemplo, para un locus de interés en una muestra mixta). Los cebadores para la amplificación selectiva se pueden elegir por diversas razones, pero se diseñan preferentemente para 1) amplificar eficazmente una región del cromosoma de interés; 2) tener un intervalo de expresión predecible de fuentes maternas y/o fetales en diferentes muestras mixtas; y 3) ser distintivos del cromosoma particular, es decir, no amplificar regiones homólogas en otros cromosomas. A continuación se indican técnicas ejemplares que se pueden emplear en el sistema de ensayo de la invención.

10 El procedimiento de ensayo de la invención detecta tanto aneuploidías como anomalías cromosómicas específicas a través de la identificación y cuantificación de locus de interés específicos. Dichas anomalías cromosómicas incluyen, pero no se limitan a, mutaciones de delección, mutaciones de inserción, polimorfismos de número de copias, 20 variantes del número de copias, síndrome de delección cromosómica 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 locus seleccionados presentes en el mismo cromosoma o en cromosomas separados, y al menos uno de los locus seleccionados se asocia con la anomalía alélica fetal. A continuación, se comparan las secuencias de los dos locus seleccionados y el número de copias de los dos locus seleccionados para determinar si la anomalía cromosómica 25 está presente y, en caso afirmativo, la naturaleza de la anomalía.

20 Si bien gran parte de la descripción contenida en el presente documento describe la detección de aneuploidías hallando la abundancia de locus seleccionados en uno o más supuestos cromosomas aneuploides y la abundancia de locus seleccionados en uno o más cromosomas normales, se pueden usar las mismas técnicas para detectar variaciones en el número de copias cuando dicha variación en el número de copias se produce solo en una parte de un cromosoma. Para la detección de variaciones en el número de copias, se comparan múltiples locus 30 seleccionados dentro de la supuesta localización de la variación en el número de copias con múltiples locus seleccionados fuera de la supuesta localización de la variación en el número de copias. Por ejemplo, se puede detectar un síndrome de delección cromosómica 22q11 en un feto en una muestra materna seleccionando dos o más locus dentro de la delección 22q11 y dos o más locus fuera de la delección 22q11. Los locus fuera de la delección 22q11 pueden estar en otra región del cromosoma 22 o pueden estar en un cromosoma completamente diferente. 35 La abundancia de cada locus se determina mediante los procedimientos descritos en esta solicitud.

40 En algunos aspectos, se puede usar una amplificación universal para amplificar los locus seleccionados. En algunos aspectos, los locus seleccionados para cada muestra se someten a ensayo en una única reacción en un único recipiente. En otros aspectos, los locus de múltiples muestras se pueden someter a ensayo en una única reacción en un único recipiente.

45 En determinados aspectos de la invención, se puede detectar una delección, incluyendo los límites de dichas delecciones. En algunos aspectos, se pueden usar al menos 24 locus seleccionados dentro de la región de la supuesta delección y se pueden usar al menos 24 locus seleccionados fuera de la región de la supuesta delección. En otro aspecto, se pueden usar al menos 48 locus seleccionados dentro de la región de la supuesta delección y se pueden usar al menos 48 locus seleccionados fuera de la región de la supuesta delección. En otro aspecto, se pueden usar al menos 96 locus seleccionados fuera de la región de la supuesta delección. En otro aspecto, se pueden usar al menos 48 locus seleccionados dentro de la región de la supuesta delección y se pueden usar al menos 192 locus seleccionados fuera de la región de la supuesta delección. En un aspecto preferente, se pueden usar al menos 384 locus seleccionados dentro de la región de la supuesta delección y se pueden usar al menos 384 locus seleccionados fuera de la región de la supuesta delección. Se suman los locus seleccionados dentro de la región de la supuesta 50 delección y los locus seleccionados fuera de la región de la supuesta delección. A continuación, estas sumas se comparan entre sí para determinar la presencia o ausencia de una delección. Opcionalmente, se calcula una proporción de las sumas y esa proporción se puede comparar con una proporción promedio creada a partir de una población normal. Cuando la proporción de uno o más locus seleccionados se encuentra estadísticamente fuera de una proporción esperada, se detecta una delección. El umbral para identificar como positiva una delección puede ser dos o más veces, preferentemente cuatro o más veces la variación calculada en la población normal. Cuando se requiere una pluralidad de locus seleccionados dentro y fuera de la región de la supuesta delección, se pueden identificar los límites de la delección.

Polimorfismos asociados con enfermedades o predisposiciones

- El procedimiento de ensayo de la invención también se puede utilizar para detectar polimorfismos, tales como los asociados con una enfermedad o mutación autosómica dominante o recesiva. Dada la naturaleza multiplexada de los sistemas de ensayo de la invención, la detección tiene lugar en el mismo ensayo que la detección de anomalías cromosómicas. Por tanto, un único sistema de ensayo puede proporcionar información de diagnóstico sobre diferentes clases de mutaciones genéticas. En consecuencia, como los sistemas de ensayo preferentes de la invención están altamente multiplexados y pueden consultar cientos o incluso miles de locus seleccionados dentro de una muestra mixta, en determinados aspectos es deseable consultar la muestra en busca de locus marcadores dentro de la muestra mixta, por ejemplo, locus asociados con un riesgo genético o que identifican la presencia o ausencia de organismos infecciosos. Por tanto, los sistemas de ensayo proporcionan la detección de dichos locus marcadores junto con la detección de los locus seleccionados para la determinación del número de copias en una muestra mixta.
- Por tanto, el procedimiento de ensayo de la invención se puede usar para detectar polimorfismos en una muestra mixta, donde dichos polimorfismos se asocian con genes asociados con trastornos autosómicos recesivos, mutaciones asociadas con trastornos autosómicos dominantes; polimorfismos asociados con el riesgo de desarrollar una enfermedad y/o de progresión de la enfermedad (por ejemplo, metástasis) e indicadores de pronóstico.
- En otros aspectos específicos, el procedimiento de ensayo de la invención se puede usar para detectar mutaciones o polimorfismos fetales en una muestra materna, donde dichas mutaciones o polimorfismos se asocian con trastornos poligénicos tales como cardiopatía coronaria, diabetes, hipertensión, defectos congénitos del corazón y epilepsia. Los ejemplos incluyen mutaciones en genes asociados con predisposiciones, tales como mutaciones en genes de susceptibilidad al cáncer (por ejemplo, mutaciones en BRCA1 o II o en p53); polimorfismos asociados con un riesgo incrementado de desarrollar enfermedades de aparición tardía, tales como el polimorfismo del gen apoE3 asociado con el riesgo de enfermedad de Alzheimer.
- Además de la detección de anomalías cromosómicas y mutaciones o polimorfismos de un único gen asociados con enfermedades, trastornos o predisposiciones monogénicas o poligénicas, los sistemas de ensayo de la invención pueden identificar agentes infecciosos en la muestra mixta.

Amplificación selectiva

- Se pueden usar numerosos procedimientos de amplificación selectiva para proporcionar los ácidos nucleicos amplificados que se analizan en los sistemas de ensayo de la invención, y dichos procedimientos se usan preferentemente para incrementar el número de copias de los locus seleccionados en una muestra mixta de manera que permita la conservación de la información relativa al contenido inicial de los locus seleccionados en la muestra mixta. Aunque no todas las combinaciones de amplificación y análisis se describen en detalle en el presente documento, está dentro de la habilidad 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 consecuentes con la presente memoria descriptiva, y dichas variaciones serán evidentes para un experto en la técnica tras la lectura de la presente divulgación.
- Dichos procedimientos de amplificación incluyen, pero no se limitan a, la 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 el documento WO90/06995), amplificación selectiva de secuencias de polinucleótidos diana (patente de EE. UU. n.º 6.410.276), reacción en cadena de la polimerasa cebada con secuencias 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 basada en secuencias de ácidos nucleicos (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: q-beta replicasa, descrito en la solicitud de patente PCT n.º PCT/US87/00880, procedimientos de amplificación isotérmica, tales como SDA, descritos en Walker *et al.*, Nucleic Acids Res. 20(7):1691-6 (1992) y amplificación en 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 en el documento de EE. UU. con el n.º de serie 09/854.317 y la publicación de EE. UU. n.º 20030143599. En algunos aspectos, el ADN se amplifica por PCR multiplexada específica de locus. En un aspecto preferente, el ADN se amplifica usando unión a adaptador y PCR con cebador único. Otros procedimientos de amplificación disponibles, tales como la PCR equilibrada (Makrigiorgos, *et al.*, Nature Biotechnol, 20:936-9 (2002)) y los procedimientos de amplificación isotérmica, tales como la amplificación de basada en secuencias de ácidos nucleicos (NASBA) y la replicación de secuencias autosostenida (Guatelli *et al.*, PNAS USA 87:1874 (1990)). En base a dichas metodologías, un experto en la técnica puede diseñar fácilmente cebadores en cualquier región 5' y 3' adecuada con respecto a un locus de interés. Dichos cebadores de

pueden usar para amplificar ADN de cualquier longitud siempre que el ADN comprenda los locus de interés seleccionados.

- 5 La longitud de un locus seleccionado amplificado de una región genómica de interés es lo suficientemente larga como para proporcionar suficiente información de secuencia para distinguir el locus amplificado de otros locus que se amplifican y/o seleccionan. En general, un ácido nucleico amplificado correspondiente a un locus seleccionado tiene al menos aproximadamente 16 nucleótidos de longitud, y más típicamente, un ácido nucleico amplificado correspondiente a un locus seleccionado tiene al menos aproximadamente 20 nucleótidos de longitud. En un aspecto preferente de la invención, un ácido nucleico amplificado correspondiente a un locus seleccionado tiene al menos aproximadamente 30 nucleótidos de longitud. En un aspecto más preferente de la invención, un ácido nucleico amplificado correspondiente a un locus seleccionado tiene al menos aproximadamente 32, 40, 45, 50 o 60 nucleótidos de longitud. En otros aspectos de la invención, un ácido nucleico amplificado correspondiente a un locus seleccionado puede tener aproximadamente 100, 150 o hasta 200 de longitud.
- 10 15 En determinados aspectos, la amplificación selectiva comprende una etapa de amplificación lineal inicial, que puede ser particularmente útil si la cantidad inicial de ADN de la muestra mixta 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 de ADN original y ayuda a reducir el error de muestreo cuando se necesita una cuantificación exacta del ADN o de una fracción del ADN (por ejemplo, contribución de ADN fetal en una muestra materna).

20 25 30 Por tanto, en un aspecto, se realiza un número limitado de ciclos de amplificación lineal específica de secuencia en la muestra mixta de partida que comprende ADNlc. El número de ciclos, en general, es menor que el usado para una amplificación por PCR típica, por ejemplo, de 5-30 ciclos o menos. Se pueden diseñar cebadores o sondas para amplificar regiones o segmentos genómicos específicos que comprenden los locus seleccionados. Los cebadores o sondas se pueden modificar con un marcador de extremo en el extremo 5' (por ejemplo, con biotina) o en otra parte a lo largo del cebador o sonda, de modo que los productos de amplificación se puedan purificar o acoplar a un sustrato sólido (por ejemplo, microesfera o matriz) para su aislamiento o análisis adicional. En un aspecto preferente, los cebadores se multiplexan de modo que una única reacción produzca múltiples fragmentos de ADN de diferentes regiones. A continuación, los productos de amplificación de la amplificación lineal se podrían amplificar adicionalmente con procedimientos de PCR estándar o con una amplificación lineal adicional.

35 40 45 Se puede aislar ADNlc, por ejemplo, de sangre, plasma o suero de una mujer embarazada, e incubarlo con cebadores dirigidos a un número establecido de locus seleccionados que correspondan a los cromosomas de interés. Preferentemente, el número de cebadores usados para la amplificación lineal inicial será de 12 o más, más preferentemente de 24 o más, más preferentemente de 36 o más, incluso más preferentemente de 48 o más, e incluso más preferentemente de 96 o más. Cada uno de los cebadores corresponde a un único locus seleccionado y se marca opcionalmente para su identificación y/o aislamiento. Con la amplificación lineal se realiza un número limitado de ciclos, preferentemente de 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 por medio de unión a avidina o estreptavidina en un sustrato sólido. A continuación, los productos de amplificación se someten a procedimientos bioquímicos adicionales, tales como amplificación adicional con otros cebadores y/o técnicas de detección, tales como determinación de secuencias e hibridación.

50 55 60 65 La eficacia de la amplificación lineal puede variar entre sitios y entre ciclos, de modo que en determinados sistemas se puede usar una normalización para asegurar que los productos de la amplificación lineal sean representativos en cuanto a frecuencia y secuencia de los ácidos nucleicos en la muestra mixta. Una persona que aplica el sistema de ensayo de la invención puede utilizar información de múltiples partes alícuotas de una muestra para determinar la variación en la cantidad de diferentes productos de amplificación que representan los locus seleccionados, incluyendo la variación en diferentes locus seleccionados y/o entre locus seleccionados después de la amplificación lineal inicial limitada. Dicha información se puede usar para determinar los niveles iniciales de los locus seleccionados en el contenido de ADN de la muestra, lo que permite la normalización de la frecuencia de los locus seleccionados.

55 **Amplificación universal**

60 65 En aspectos preferentes de la invención, los locus amplificados de forma selectiva preferentemente se amplifican adicionalmente a través de la amplificación universal de todos o sustancialmente todos los diversos locus seleccionados usando los sistemas de ensayo de la invención. Las regiones de cebadores universales se añaden a los oligonucleótidos de secuencia fija de modo que los locus amplificados de forma selectiva se puedan amplificar adicionalmente en una única reacción de amplificación universal. Estas secuencias de cebadores universales se pueden añadir a las regiones de ácidos nucleicos durante el procedimiento de amplificación selectiva, es decir, los cebadores para la amplificación selectiva comprenden secuencias de cebadores universales. De forma alternativa, se pueden añadir adaptadores que comprenden secuencias de amplificación universal a los extremos de los locus seleccionados amplificados de forma selectiva como adaptadores tras la amplificación inicial y después del aislamiento de los locus seleccionados amplificados de forma selectiva de la muestra mixta.

- En un aspecto ejemplar, los locus seleccionados se amplifican inicialmente a partir de una muestra mixta usando cebadores complementarios a los locus de interés seleccionados, seguido de una etapa de amplificación universal para incrementar el número de locus para el análisis. La introducción de regiones de cebadores en los productos de amplificación inicial de una muestra mixta permite la amplificación universal controlada posterior de todos o una parte de los ácidos nucleicos seleccionados antes de o durante el análisis, por ejemplo, la determinación de secuencias.
- El sesgo y la variabilidad se pueden introducir durante la amplificación de ADN, tal como lo observado durante la reacción en cadena de la polimerasa (PCR). En casos donde una reacción de amplificación se multiplexa, existe el potencial de que los diferentes locus seleccionados se amplifiquen a diferentes tasas o eficacias. Esto se puede deber en parte a que algunos cebadores tienen una mejor eficacia en una reacción multiplexada (es decir, cinéticas de hibridación más favorables) debido a condiciones experimentales que favorecen a algunos cebadores sobre otros, tales como el contenido de secuencia del cebador y el ADN molde, las condiciones del tampón y otras condiciones. Una amplificación de ADN universal en un sistema de ensayo multiplexado introduce, en general, menos sesgo y variabilidad entre los locus amplificados.
- En consecuencia, en un aspecto, se realiza un pequeño número (por ejemplo, 1-10, preferentemente 3-5) de ciclos de amplificación selectiva usando secuencias específicas de los locus, seguido de una amplificación universal usando cebadores universales. El número de ciclos en los que se 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 tras un número menor de ciclos de amplificación, se reduce el sesgo de que determinados locus se amplifiquen a tasas mayores que otros.
- Opcionalmente, el sistema de ensayo incluye una etapa entre el procedimiento de amplificación selectiva y el procedimiento de amplificación universal para retirar cualquier ácido nucleico que no se haya amplificado selectivamente en la reacción de amplificación selectiva.
- Para la reacción de amplificación universal se puede usar el producto completo o una parte alícuota del producto de la amplificación selectiva. Se pueden usar las mismas condiciones o condiciones diferentes (por ejemplo, polimerasa, tampones y similares) en las etapas de amplificación, por ejemplo, para garantizar que no se introducen de forma inadvertida sesgo y variabilidad debido a las condiciones experimentales. Además, se pueden usar variaciones en las concentraciones de cebadores para limitar diferencialmente el número de ciclos de amplificación específica de secuencia para algunos locus seleccionados en comparación con otros locus seleccionados.
- 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 multiplexado convencionales que utilizan mecanismos de cebado generales para analizar grandes cantidades de ácidos nucleicos simultáneamente en una reacción en un recipiente. Dichos procedimientos de cebado "universal" permiten un análisis eficaz y de alto volumen de la cantidad de locus presentes en una muestra mixta, y permiten la cuantificación exhaustiva de la presencia de locus seleccionados dentro de dicha muestra mixta para la determinación de una aneuploidía.
- Los ejemplos de dichos procedimientos de ensayo incluyen, pero no se limitan a, procedimientos de multiplexación usados para amplificar y/o genotipar simultáneamente una variedad de muestras, tales como los descritos en Oliphant *et al.*, patente de EE. UU. n.º 7.582.420.
- En algunos aspectos se utilizan reacciones acopladas para la detección multiplexada de secuencias de ácido nucleico cuando los oligonucleótidos de una fase temprana de cada procedimiento contienen secuencias que pueden usar los oligonucleótidos de una fase posterior del procedimiento. Se pueden usar procedimientos ejemplares, solos o en combinación, para amplificar y/o detectar ácidos nucleicos en muestras, que incluyen, pero sin limitarse a, los procedimientos descritos a continuación, cada uno de los cuales se incorpora en su totalidad por referencia.
- En determinados aspectos, el sistema de ensayo de la invención utiliza una de las siguientes técnicas de amplificación selectiva y universal combinadas: (1) reacción de detección de ligasa ("LDR") acoplada a reacción en cadena de la polimerasa ("PCR"); (2) PCR primaria acoplada a PCR secundaria acoplada a LDR; y (3) PCR primaria acoplada a PCR secundaria. Cada uno de estos aspectos de la invención tiene una aplicabilidad particular en la detección de determinadas características de los ácidos nucleicos. Sin embargo, cada uno requiere el uso de reacciones acopladas para la detección multiplexada de diferencias en las secuencias de ácido nucleico cuando los oligonucleótidos de una fase temprana de cada procedimiento contienen secuencias que pueden usar los 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, 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 ácidos nucleicos.

- 5 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 la PCR acoplada a LDR para la detección de ácidos nucleicos.
- 10 Barany *et al.*, patentes de EE. UU. n.º 7.556.924 y 6.858.412, describen el uso de sondas precirculares (también llamadas "sondas candado" o "sondas de inversión múltiple") con LDR y reacción en cadena de la polimerasa ("PCR") acopladas para la detección de ácidos nucleicos.
- 15 Barany *et al.*, patentes de EE. UU. n.º 7.807.431, 7.709.201 y 7.198.814, describen el uso de reacciones combinadas de escisión por endonucleasas y unión para la detección de secuencias del ácidos nucleicos.
- 20 Willis *et al.*, patentes de EE. UU. n.º 7.700.323 y 6.858.412, describen el uso de sondas precirculares en la amplificación, detección y genotipado multiplexados de ácidos nucleicos.
- 25 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 exclusivo y un código de barras.
- 30 20 En un aspecto preferente, los productos de amplificación se multiplexan como se describe previamente. En un aspecto preferente, los productos de amplificación multiplexada se cuantifican mediante análisis de los productos de amplificación. En un aspecto preferente, una muestra representativa de moléculas individuales de los procedimientos de amplificación se aísla del resto de la muestra para su posterior análisis. Para obtener una muestra representativa de moléculas individuales, el número promedio de moléculas por locus debe exceder el ruido de muestreo generado por la reacción multiplexada. En un aspecto, el número promedio por locus es mayor que 100. En otro aspecto, el número promedio por locus es mayor que 500. En otro aspecto, el número promedio por locus es mayor que 1000.
- 35 30 Las moléculas individuales del producto de amplificación preferentemente se aíslan físicamente de las otras moléculas de manera que permita distinguir entre sí los diferentes productos de amplificación en el análisis. En un aspecto preferente, este aislamiento se produce sobre un sustrato sólido. Cada molécula aislada se puede asociar con una dirección física o identificable particular antes del análisis, o la dirección puede llegar a ser conocida para los productos de amplificación particulares en base al resultado del análisis. El sustrato puede ser una superficie plana o una superficie tridimensional tal como una microesfera.
- 40 35 Una vez aislado, el producto de amplificación individual se puede amplificar adicionalmente para hacer múltiples copias idénticas de esa molécula en la misma localización conocida o identificable. La amplificación se puede producir antes o después de que esa localización se convierta en una dirección física o identificable. El producto de amplificación y/o sus copias (que pueden ser idénticas o complementarias al producto de amplificación) se analizan a continuación en base a la secuencia del producto de amplificación o sus copias para identificar el locus y/o alelo particular que representa.
- 45 40 En un aspecto preferente, la longitud completa del producto de amplificación o una parte del producto de amplificación se puede analizar usando determinación de secuencias. El número de bases que es necesario determinar debe ser suficiente para identificar el producto de amplificación de forma exclusiva como perteneciente a un locus y/o alelo específico. En un aspecto preferente, el locus se analiza mediante la determinación de secuencias del producto de amplificación.
- 50 45 Numerosos procedimientos de determinación de secuencias son compatibles con los sistemas de ensayo de la invención. Los procedimientos ejemplares para la determinación de secuencias incluyen, pero no se limitan a, procedimientos basados en hibridación, tales como los que se divultan 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. n.º 2005/0191656, secuenciación por procedimientos de síntesis, por ejemplo, Nyren *et al.*, patentes de EE. UU. n.º 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 por pirofósforo, 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 unión, 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.
- 55 50 La información de secuencia se puede determinar usando procedimientos que determinen muchas (típicamente miles a miles de millones) de secuencias de ácidos nucleicos de una manera intrínsecamente paralela, donde muchas secuencias se leen preferentemente en paralelo usando un procedimiento en serie de alto rendimiento. Dichos procedimientos incluyen, pero no se limitan a, pirosecuenciación (por ejemplo, comercializada por 454 Life Sciences, Inc., Branford, CT); secuenciación por unión (por ejemplo, tal como se comercializa en la tecnología SOLiD™, Life Technology, Inc., Carlsbad, CA); secuenciación por síntesis usando nucleótidos modificados (tal como la comercializada en la tecnología TruSeq™ y HiSeq™ por Illumina, Inc., San Diego, CA; HeliScope™ por Helicos

Biosciences Corporation, Cambridge, MA; y PacBio RS por Pacific Biosciences of California, Inc., Menlo Park, CA), secuenciación mediante tecnologías de detección de iones (Ion Torrent, Inc., South San Francisco, CA); secuenciación de nanobolas de ADN (Complete Genomics, Inc., Mountain View, CA); tecnologías de secuenciación basada en nanoporos (por ejemplo, desarrolladas por Oxford Nanopore Technologies, LTD, Oxford, Reino Unido) y procedimientos de secuenciación altamente paralelizados similares.

De forma alternativa, en otro aspecto, la longitud completa del producto de amplificación o una parte del producto de amplificación se puede analizar usando técnicas de hibridación. Los procedimientos para llevar a cabo ensayos de hibridación de polinucleótidos para la detección se han desarrollado bien 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 a los que se hace referencia en: Maniatis *et al.*, Molecular Cloning: A Laboratory Manual (2.^a 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 repetidas y controladas se han descrito en las patentes de EE. UU. n.^o 5.871.928, 5.874.219, 6.045.996 y 6.386.749, 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, la solicitud de patente de EE. UU. 60/364.731 y la solicitud PCT PCT/US99/06097 (publicada como documento WO99/47964).

Los procedimientos y aparatos para la detección de señales y el procesamiento de datos de intensidad se divultan, 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 documento WO99/47964).

Minimización de la variación dentro de y entre muestras

Un desafío de la detección de anomalías cromosómicas por detección en una muestra mixta reside en que los ácidos nucleicos de la fuente secundaria pueden estar presentes en una abundancia mucho menor que los ácidos nucleicos de la fuente principal normal. En el caso de una muestra materna que contiene ADN_{IC} fetal y materno, el ADN fetal libre circulante como porcentaje del ADN_{IC} total puede variar de menos de un uno a un cuarenta por ciento, y lo más común es que esté presente en o por debajo de un veinte por ciento y con frecuencia 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, el incremento relativo del cromosoma 21 es del 50 % en el ADN fetal y, por tanto, como porcentaje del ADN total en una muestra materna donde, a modo de ejemplo, el ADN fetal representa el 5 % del total, el incremento del cromosoma 21 como porcentaje del total es del 2,5 %. Si esta diferencia se ha de detectar de forma 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 porcentaje de incremento del cromosoma 21 aportado por un tercer cromosoma 21 del ADN fetal.

La variación entre los niveles encontrados entre muestras y/o para locus dentro de una muestra se puede minimizar usando una combinación de procedimientos analíticos, 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 compararla con un cromosoma presente en una abundancia supuestamente anómala, tal como una aneuploidía, en la misma muestra. Aunque el uso de un cromosoma "normal" de este tipo como cromosoma de referencia puede ser suficiente, también es posible usar de dos a varios cromosomas normales como cromosomas de referencia interna para incrementar la potencia estadística de la cuantificación.

Un procedimiento de uso de una referencia interna es calcular una proporción de abundancia de los cromosomas supuestamente anómalos con respecto a la abundancia de los cromosomas normales en una muestra, llamada proporción cromosómica. Al calcular la proporción cromosómica, la abundancia o recuentos de cada uno de los locus seleccionados para cada cromosoma se suman para calcular los recuentos totales para cada cromosoma. A continuación, los recuentos totales de un cromosoma se dividen entre 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 sumando en primer lugar los recuentos de cada uno de los locus seleccionados para cada cromosoma, y, a continuación, dividiendo la suma para un cromosoma entre la suma total para dos o más cromosomas. Una vez calculada, la proporción cromosómica se compara a continuación con la proporción cromosómica promedio de una población normal.

El promedio puede ser la media, la mediana, la moda u otro promedio, con o sin normalización y exclusión 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 una serie de formas, 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 para la muestra se encuentra estadísticamente fuera de la proporción cromosómica promedio para 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 de las tasas aceptables de positivos falsos y negativos falsos para el ensayo. 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 el ejemplo anterior, la proporción cromosómica se determina sumando los recuentos de los locus seleccionados por cromosoma. Típicamente, se usa el mismo número de locus seleccionados para cada cromosoma. Un procedimiento alternativo para generar la proporción cromosómica sería calcular los recuentos promedio para los locus seleccionados para cada cromosoma. El promedio puede ser cualquier estimación de la media, la mediana o la moda, aunque típicamente se usa un promedio. El promedio puede ser la media de todos los recuentos o alguna variación, tal como un promedio truncado o ponderado. Una vez que se hayan calculado los recuentos promedio para cada cromosoma, los recuentos promedio para cada cromosoma se pueden dividir entre los otros para obtener una proporción cromosómica entre dos cromosomas, los recuentos promedio para cada cromosoma se pueden dividir entre la suma de los promedios para todos los cromosomas medidos para obtener una proporción cromosómica para cada cromosoma como se describe anteriormente. Como se resalta anteriormente, la capacidad de detectar una aneuploidía en una muestra materna donde el supuesto ADN se halla en baja abundancia relativa depende en gran medida de la variación en las mediciones de diferentes locus seleccionados en el ensayo. Se pueden usar numerosos procedimientos analíticos que reduzcan esta variación y, por tanto, mejoren la sensibilidad de este procedimiento para detectar una aneuploidía. Un procedimiento para reducir la variabilidad del ensayo es incrementar el número de locus seleccionados usados para calcular la abundancia de los cromosomas. En general, si la variación medida de un único locus seleccionado de un cromosoma es X % y se miden Y locus seleccionados diferentes en el mismo cromosoma, la variación de la medición de la abundancia cromosómica calculada sumando o promediando la abundancia de cada locus seleccionado en ese cromosoma será de aproximadamente X % dividido entre Y^{1/2}. Dicho de otra forma, la variación de la medición de la abundancia de los cromosomas sería aproximadamente la variación promedio de la medición de la abundancia de cada locus seleccionado dividida entre la raíz cuadrada del número de locus.

En un aspecto preferente de la presente invención, el número de locus seleccionados medidos para cada cromosoma es de al menos 24. En otro aspecto preferente de la presente invención, el número de locus seleccionados medidos para cada cromosoma es de al menos 48. En otro aspecto preferente de la presente invención, el número de locus seleccionados medidos para cada cromosoma es de al menos 100. En otro aspecto preferente de la presente invención, el número de locus seleccionados medidos para cada cromosoma es de al menos 200. La medición de cada locus supone un coste incremental y, por tanto, es importante minimizar el número de locus seleccionados. En un aspecto preferente de la presente invención, el número de locus seleccionados medidos para cada cromosoma es menor que 2000. En un aspecto preferente de la presente invención, el número de locus seleccionados medidos para cada cromosoma es menor que 1000. En un aspecto más preferente de la presente invención, el número de locus seleccionados medidos para cada cromosoma es de al menos 48 y menor que 1000. En un aspecto, tras la medición de la abundancia para cada locus seleccionado, se puede usar un subconjunto de los locus seleccionados para determinar la presencia o ausencia de una aneuploidía. Existen muchos procedimientos estándar para elegir el subconjunto de locus seleccionados. Estos procedimientos incluyen la exclusión de valores atípicos, donde los locus seleccionados con niveles detectados inferiores y/o superiores a un determinado percentil se descartan del análisis. En un aspecto, el percentil puede ser el 5 % más bajo y más alto como se mide por la abundancia. En otro aspecto, el percentil puede ser el 10 % más bajo y más alto como se mide por la abundancia. En otro aspecto, el percentil puede ser el 25 % más bajo y más alto como se mide por la abundancia.

Otro procedimiento para elegir el subconjunto de locus seleccionados incluye la eliminación de regiones que se encuentran fuera de cierto límite estadístico. Por ejemplo, los locus seleccionados que se encuentran fuera de una o más desviaciones estándar de la abundancia media se pueden retirar del análisis. Otro procedimiento para elegir el subconjunto de locus seleccionados puede ser comparar la abundancia relativa de un locus seleccionado con la abundancia esperada del mismo locus seleccionado en una población sana y descartar cualquier locus seleccionado que no cumpla la prueba de expectativas. Para minimizar además la variación en el ensayo, se puede incrementar el número de veces que se mide cada locus seleccionado. Como se analiza, a diferencia de los procedimientos aleatorios de detección de aneuploidías, donde se mide el genoma en promedio menos de una vez, los sistemas de ensayo de la presente invención miden múltiples veces de forma intencionada cada locus seleccionado. En general, cuando se cuentan acontecimientos, la variación en el recuento se determina por la estadística de Poisson, y la variación en el recuento es típicamente igual a uno dividido entre la raíz cuadrada del número de recuentos. En un aspecto preferente de la invención, los locus seleccionados se miden cada uno en promedio al menos 100 veces. En un aspecto preferente de la invención, los locus seleccionados se miden cada uno en promedio al menos 500 veces. En un aspecto preferente de la invención, los locus seleccionados se miden cada uno en promedio al menos 1000 veces. En un aspecto preferente de la invención, los locus seleccionados se miden cada uno en promedio al menos 1000 veces.

2000 veces. En un aspecto preferente de la invención, los locus seleccionados se miden cada uno en promedio al menos 5000 veces.

5 En otro aspecto, los subconjuntos de locus se pueden elegir al azar pero con un número suficiente de locus seleccionados para proporcionar un resultado estadísticamente significativo en la determinación de si existe una anomalía cromosómica. Se pueden realizar múltiples análisis de diferentes subconjuntos de locus dentro de una muestra mixta para proporcionar mayor potencia estadística. En este ejemplo, puede ser necesario o no retirar o 10 eliminar cualquier locus seleccionado antes del análisis aleatorio. Por ejemplo, si existen 100 locus seleccionados para el cromosoma 21 y 100 locus seleccionados para el cromosoma 18, se podrían realizar una serie de análisis que evalúen menos de 100 locus para cada uno de los cromosomas.

15 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 la presente solicitud. En general, la variación en el ensayo se puede reducir cuando todos los locus seleccionados para cada muestra se consultan en una única reacción en un único recipiente. De forma similar, se puede reducir la variación en el ensayo cuando se usa un sistema de amplificación universal. Además, la variación del ensayo se puede reducir cuando el número de ciclos de amplificación es limitado.

20 *Uso de los sistemas de ensayo para la detección en muestras mixtas de pacientes con cáncer*

25 El sistema de ensayo permite detectar alteraciones cuantitativas y cualitativas específicas de tumores en el ADNlc, tales como integridad de las hebras de ADN, frecuencia de mutaciones, anomalías de microsatélites y metilación de genes, como marcadores de diagnóstico, pronóstico y seguimiento en pacientes con cáncer. La capacidad de combinar dicha detección de alteraciones de un único gen (incluyendo mutaciones puntuales, indels y variación en el número de copias) con la detección de la VNC proporciona un procedimiento poderoso para respaldar el diagnóstico clínico, los tratamientos, la predicción de resultados y el seguimiento de la progresión en pacientes con una neoplasia maligna o con sospecha de tenerla.

30 En algunos aspectos, el sistema de ensayo de la invención se usa para propósitos de diagnóstico, por ejemplo, para detectar la presencia y/o la naturaleza de una neoplasia maligna en un paciente o para proporcionar una estimación cuantitativa de la carga tumoral en un paciente. El ADN y los microARN tumorales circulantes se han asociado con determinados cánceres, tales como el cáncer de pulmón (Roth C *et al.*, Mol Oncol. 2011 Jun; 5(3):281-91. Epub del 24 de febrero de 2011). También se han detectado variaciones en el número de copias en determinados cánceres, 35 tales como HER2 y receptor de estrógeno amplificados en el ADNlc de pacientes con cáncer de mama. (Page K., Br J Cancer. 12 de abril de 2011;104(8):1342-8. Epub del 22 de marzo de 2011).

40 En otros aspectos de la invención, el sistema de ensayo se usa en pacientes con cáncer para monitorizar una respuesta al tratamiento y/o seguir el progreso de la enfermedad, por ejemplo, para medir alteraciones de un único gen y ADNlc en pacientes que reciben quimiorradioterapia (CRT). Para determinados cánceres, se ha demostrado que el índice de integridad del ADNlc puede estar asociado de manera significativa e independiente con la respuesta del tumor al tratamiento. Agostini M *et al.*, Ann Surg Oncol. 17 de marzo de 2011. Además, la presencia o ausencia de determinadas alteraciones genéticas y/o diferencias en la variación del número de copias se ha asociado con la respuesta a la quimioterapia y/o el pronóstico de una enfermedad. Véase, por ejemplo, Savas S., Acta Oncol. 45 Noviembre de 2010; 49(8):1217-26. Epub del 29 de julio de 2010, que describe variaciones genéticas útiles para determinar la respuesta al tratamiento y la supervivencia en el cáncer. Por ejemplo, la detección de los niveles de ADNlc combinada con la detección de mutaciones en el gen K-RAS y/o el gen p53 proporcionan una herramienta poderosa y relativamente no invasiva para medir el pronóstico de diversos cánceres, incluyendo el cáncer de ovario, el cáncer de endometrio y los linfomas. Dobrzycka B *et al.*, Ann Oncol., mayo de 2011; 22(5):1133-40. Epub del 23 de noviembre de 2010; Dobrzycka B *et al.*, Int J Cancer, 1 de agosto de 2010; 127(3):612-21; Hosny G *et al.*, Cancer 50 Lett., 18 de marzo de 2009; 275(2):234-9. Epub del 28 de noviembre de 2008. Dicho análisis se puede respaldar aún más usando herramientas tales como Varietas, un portal de bases de datos funcionales para la identificación de variaciones genéticas y la asociación con los resultados del tratamiento y el pronóstico. Paananen J *et al.*, Database (Oxford). 29 de julio de 2010; 2010:baq016.

55 *Uso de los sistemas de ensayo para la detección en muestras mixtas de pacientes trasplantados*

60 Los sistemas de ensayo de la invención se pueden usar para monitorizar la salud de los órganos en un paciente trasplantado usando una combinación de detección de ADNlc y detección de SNP o mutaciones en uno o más genes individuales. Los órganos trasplantados tienen genomas que son distintos del genoma de un paciente receptor, y la salud de los órganos se puede detectar usando el sistema de ensayo. Por ejemplo, se ha demostrado que el rechazo celular agudo está asociado con niveles significativamente incrementados de ADN libre circulante del genoma del donante en los receptores de trasplantes de corazón. Snyder TM *et al.*, Proc Natl Acad Sci USA. 12 de abril de 2011; 108(15):6229-34. Epub del 28 de marzo de 2011. Además, las quimiocinas y las moléculas de adhesión median el rechazo del aloinjerto al reclutar leucocitos hacia el aloinjerto, y se ha demostrado que los SNP localizados en la interleucina (IL)-8, CXCR1, CXCR2, se correlacionan con los resultados del aloinjerto. Ro H. *et al.*, Transplantation. 15 de enero de 2011; 91(1):57-64. Por tanto, los sistemas de ensayo de la invención pueden

proporcionar pruebas no invasivas para monitorizar a los receptores de trasplantes de órganos sólidos y pueden ayudar a identificar los primeros signos de rechazo sin necesidad de realizar biopsias de órganos o de aplicar otras técnicas de diagnóstico o pronóstico más costosas.

5 *Uso de los sistemas de ensayo para la detección en muestras maternas*

En determinados aspectos específicos, la determinación del porcentaje relativo de ADN fetal en una muestra materna puede ser beneficiosa para analizar los productos de amplificación, ya que el porcentaje de ADN fetal en la muestra proporciona información importante sobre la presencia estadística esperada de cromosomas, y la variación de esa expectativa puede ser indicativa de una aneuploidía fetal. Esto puede ser especialmente útil en circunstancias en las que el nivel de ADN fetal en una muestra materna es bajo, ya que el porcentaje de contribución fetal se puede usar para determinar la significación estadística cuantitativa de las variaciones de los niveles de los locus seleccionados identificados en una muestra materna. En otros aspectos, la determinación del porcentaje relativo de ADN_{IC} fetal en una muestra materna puede ser beneficiosa para estimar el nivel de certeza o potencia en la detección de una aneuploidía fetal.

En algunos aspectos específicos, la contribución fetal relativa de ADN materno en el alelo de interés se puede comparar con la contribución paterna en ese alelo para determinar la concentración de ADN fetal aproximada en la muestra. En otros aspectos específicos, la cantidad relativa de secuencias derivadas únicamente del padre (por ejemplo, secuencias del cromosoma Y o polimorfismos específicos del padre) se puede usar para determinar la concentración relativa de ADN fetal en una muestra materna.

Otro enfoque ejemplar para determinar el porcentaje de contribución fetal en una muestra materna es a través del análisis de fragmentos de ADN con diferentes patrones de metilación del ADN entre el ADN fetal y el materno. En un aspecto preferente, el ADN amplificado a partir de ADN libre circulante se obtiene por reacción en cadena de la polimerasa (PCR). También se pueden usar otros mecanismos de amplificación, que incluyen los descritos con más detalle en el presente documento, como resultará evidente para un experto en la técnica tras la lectura de la presente divulgación.

30 En circunstancias en las que el feto es de sexo masculino, el porcentaje de ADN fetal en una muestra se puede determinar mediante la detección de locus específicos del cromosoma Y y la comparación con el contenido de ADN materno calculado. Las cantidades de un locus amplificado específico del cromosoma Y, tal como una región del gen región de determinación sexual del cromosoma Y (SRY), que se localiza en el cromosoma Y y, por tanto, es representativo del ADN fetal, se pueden determinar a partir de la muestra y comparar con uno o más locus seleccionados amplificados que están presentes tanto en el ADN materno como en el ADN fetal y que preferentemente no provienen de un cromosoma que se cree que sea potencialmente aneuploide en el feto, por ejemplo, una región autosómica que no se encuentra en el cromosoma 21 o 18. Preferentemente, esta etapa de amplificación se realiza en paralelo con la etapa de amplificación selectiva, aunque se puede realizar antes o después de la amplificación selectiva dependiendo de la naturaleza del ensayo multiplexado.

40 En aspectos particulares, el porcentaje de ADN fetal libre circulante en una muestra materna se puede determinar mediante PCR usando ADN diluido en serie aislado de la muestra materna, que permite cuantificar con exactitud el número de genomas que comprenden los genes amplificados. La PCR que usa ADN diluido en serie aislado de la muestra materna puede ser preferente cuando se determina el porcentaje de ADN fetal en un feto de sexo masculino. Por ejemplo, si la muestra de sangre contiene 100 % de ADN fetal masculino y se realizan diluciones en serie 1:2, entonces, en promedio, la señal ligada al cromosoma Y desaparecerá 1 dilución antes que la señal autosómica, ya que hay 1 copia del gen ligado al cromosoma Y y 2 copias del gen autosómico.

50 En un aspecto específico, el porcentaje de ADN fetal libre en el plasma materno se calcula para un feto de sexo masculino usando la siguiente fórmula: porcentaje de ADN fetal libre = $(n.º \text{ de copias del gen ligado al cromosoma Y} \times 2 \times 100) / (n.º \text{ de copias del gen autosómico})$, donde el número de copias de cada gen se determina observando la dilución en serie más alta a la que se detectó el gen. La fórmula contiene un factor de multiplicación de 2, que se usa para normalizar el hecho de que en cada genoma, fetal o materno, solo existe 1 copia del gen ligado al cromosoma Y en comparación con dos copias del gen autosómico.

55 *Determinación del contenido de ADN de la fuente secundaria en una muestra mixta*

En determinados aspectos de la invención, la determinación de la contribución de ADN de una fuente secundaria puede ser útil para determinar la variación en el número de copias de los locus en esas muestras. Por ejemplo, en cada muestra materna, el ADN de un feto tendrá aproximadamente un 50 % de sus locus genéticos heredados de la madre y un 50 % de los locus genéticos heredados del padre. La determinación de los locus aportados al feto por fuentes no maternas permite la estimación de ADN fetal en una muestra materna y, por tanto, proporciona información usada para calcular las diferencias estadísticamente significativas en las frecuencias cromosómicas para los cromosomas de interés.

En determinados aspectos, la determinación de polimorfismos en la fuente secundaria requiere un análisis dirigido de SNP y/o mutaciones para identificar la presencia del ADN de la fuente secundaria en una muestra mixta. La información necesaria para este cálculo se puede proporcionar usando el ensayo de la invención. En algunos aspectos, el uso de genotipado previo es útil, por ejemplo el genotipado del donante de un trasplante, el genotipado del padre y de la madre en una muestra materna. En general, sin embargo, esta información correspondiente al genotipado previo no es necesaria antes de realizar el ensayo, y el genotipado se realiza simultáneamente con la determinación del número de copias de los locus seleccionados dentro de una muestra mixta.

5 En un aspecto preferente, el porcentaje de ácidos nucleicos de la fuente secundaria en una muestra mixta se puede cuantificar usando detección multiplexada de SNP sin usar el conocimiento genotípico previo. En este aspecto, se usan dos o más locus polimórficos seleccionados con un SNP conocido en cada región. En un aspecto preferente, los locus polimórficos seleccionados son locus amplificados. En un aspecto preferente, la amplificación es universal. En un modo de realización preferente, los locus polimórficos seleccionados se amplifican en una reacción en un recipiente. Se determina y se cuantifica cada alelo de los locus polimórficos seleccionados en la muestra materna.

10 15 En un aspecto preferente, se usa secuenciación de alto rendimiento para dicha determinación y cuantificación. Se identifican los locus donde los genotipos de las fuentes principal y secundaria son diferentes; por ejemplo, el genotipo del donante es homocigótico y el genotipo del receptor es heterocigótico. Esta identificación se realiza observando una alta frecuencia relativa de un alelo (>80 %) y una baja frecuencia relativa (<20 % y >0,15 %) del otro alelo para un locus seleccionado particular. El uso de múltiples locus es particularmente ventajoso, ya que reduce la cantidad de variación en la medición de la abundancia de los alelos. Se usan todos o un subconjunto de los locus que cumplan este requisito para determinar la concentración de ácido nucleico de la fuente secundaria a través de análisis estadístico. En un aspecto, la concentración se determina sumando los alelos de baja frecuencia de dos o más locus entre sí, dividiendo entre la suma de los alelos de alta frecuencia y multiplicando por dos. En otro aspecto, el porcentaje de ácido nucleico de la fuente secundaria se determina promediando los alelos de baja frecuencia de dos o más locus, dividiendo entre el promedio de los alelos de alta frecuencia y multiplicando por dos.

20 25

Para muchos alelos, las secuencias de ácido nucleico de las fuentes principal y secundaria pueden ser homocigóticas e idénticas, y como esta información no es distinguible, no es útil en la determinación del ácido nucleico de la fuente secundaria en una muestra mixta. La presente invención utiliza información alélica donde existe una diferencia distinguible entre las fuentes celulares (por ejemplo, un alelo fetal que contiene al menos un alelo que difiere del alelo materno) en los cálculos de los porcentajes de ácido nucleico de la fuente secundaria. Los datos relativos a regiones alélicas que son iguales para las fuentes principal y secundaria no se seleccionan, por tanto, para el análisis, o se retiran de los datos pertinentes antes de la determinación del porcentaje para no enturbiar los datos útiles.

30 35 En un aspecto, los locus seleccionados se pueden excluir si la cantidad o frecuencia de la región parece ser un valor atípico debido a un error experimental o de sesgo genético idiopático dentro de una muestra particular. En otro aspecto, los ácidos nucleicos seleccionados se pueden someter a un ajuste estadístico o matemático, tal como normalización, estandarización, agrupamiento o transformación antes de la suma o el promediado. En otro aspecto, los ácidos nucleicos seleccionados se pueden someter tanto a normalización como a exclusión de errores experimentales de los datos antes de la suma o el promediado.

40 45 En un aspecto preferente, se usan 12 o más locus para el análisis. En otro aspecto preferente, se usan 24 o más locus para el análisis. En otro aspecto preferente, se usan 48 o más locus para el análisis. En otro aspecto, se usan uno o más índices para identificar la muestra

50 En un aspecto específico, la contribución de la fuente secundaria se puede cuantificar usando la detección de SNP en tándem. Las técnicas para identificar SNP en tándem en ADN extraído de, por ejemplo, una muestra materna se divulan en Mitchell *et al.*, patente de EE.UU. n.º 7.799.531 y solicitudes de patente de EE.UU. n.º 12/581.070, 12/581.083, 12/689.924 y 12/850.588. Estas describen la diferenciación de los locus fetales y maternos a través de la detección de al menos un polimorfismo mononucleotídico (SNP) en tándem en una muestra materna que tiene un haplotipo diferente entre el genoma fetal y el materno. La identificación y cuantificación de estos haplotipos se puede realizar directamente en la muestra materna, como se describe en las divulgaciones de Mitchell *et al.*, y usar para determinar el porcentaje de contribución fetal en la muestra materna.

60 Dado que en una muestra que no ha sido previamente analizada con respecto al sexo fetal, la probabilidad aproximada de que el feto sea de sexo masculino es del 50 %, las secuencias específicas del cromosoma Y solo serán aplicables en la mitad de dichas muestras. En un aspecto específico, los locus polimórficos usados para la determinación de la contribución de la fuente fetal no están en el cromosoma Y.

65 Una vez que se ha calculado el porcentaje de ADNlc para la fuente secundaria, estos datos se pueden combinar con procedimientos para la detección de aneuploidía para determinar la probabilidad de que una muestra mixta pueda contener una aneuploidía. En un aspecto se usa un procedimiento de detección de aneuploidía que utiliza el análisis

de segmentos de ADN aleatorios, tal como el descrito en, por ejemplo, Quake, solicitud de patente de EE. UU. n.º 11/701.686; y Shoemaker *et al.*, solicitud de patente de EE. UU. n.º 12/230.628. En un aspecto preferente, los procedimientos de detección de aneuploidía que utilizan el análisis de locus seleccionados en una muestra mixta incluyen tanto regiones para determinar el contenido en ADN de la fuente secundaria como regiones no polimórficas de dos o más cromosomas para detectar una anomalía cromosómica en una única reacción. La única reacción ayuda a minimizar el riesgo de contaminación o sesgo que se puede introducir durante las diversas etapas que de otro modo se llevarían a cabo en el sistema de ensayo y podrían producir un sesgo en los resultados si se utilizara el contenido de ADN de la fuente secundaria para ayudar a determinar la presencia o ausencia de una anomalía cromosómica. En otros aspectos, se pueden utilizar un locus o regiones seleccionados tanto para la determinación del contenido de ADN de la fuente secundaria como para la detección de anomalías cromosómicas en la fuente secundaria. La utilización de las mismas regiones tanto para el contenido en ADN como para la detección de anomalías cromosómicas puede ayudar además a minimizar cualquier sesgo debido a un error experimental o contaminación.

- 15 **EJEMPLOS**
- Los siguientes ejemplos se presentan para proporcionar a los expertos en la técnica una divulgación y descripción completas de cómo realizar y usar la presente invención, y no pretenden limitar el alcance de lo que los autores de la invención consideran su invención, ni pretenden representar o dar a entender que los experimentos a continuación son todos o los únicos experimentos realizados. Los expertos en la técnica apreciarán que se puedan realizar numerosas variaciones y/o modificaciones de la invención como se muestra en los aspectos específicos sin apartarse del alcance de la invención como se describe y define ampliamente en las reivindicaciones adjuntas. Por lo tanto, los presentes aspectos se deben considerar en todos los aspectos como ilustrativos y no restrictivos.
- 20 25 Se han realizado esfuerzos para garantizar la exactitud con respecto a los números usados (por ejemplo, cantidades, temperatura, etc.), pero se deben tener en cuenta algunos errores y desviaciones experimentales. A menos que se indique de otro modo, las partes son partes en peso, el peso molecular es el peso molecular promedio ponderado, la temperatura se expresa en grados centígrados y la presión es igual o cercana a la atmosférica.

30 **Ejemplo 1: Aspectos generales de los sistemas de ensayo de la invención**

Se sometieron a prueba varios formatos de ensayo para demostrar la capacidad de realizar la amplificación selectiva y la detección de locus independientes para demostrar la detección multiplexada basada en unión de un gran número (por ejemplo, 96 o más) de locus de interés. Estos locus incluían locus que eran indicativos de la presencia de un cromosoma particular o de la presencia o ausencia de una mutación o polimorfismo en un alelo particular.

35 Estos ensayos se diseñaron en base a secuencias genómicas humanas, y cada consulta consistió en dos oligonucleótidos de secuencia fija por locus seleccionado consultado en el ensayo. El primer oligonucleótido, complementario a la región 3' de una región genómica, comprendía los siguientes elementos de secuencia oligonucleotídica (de 5' a 3'): una secuencia de cebado para PCR universal común para todos los ensayos: TACACCGCGTTATGCGTCGAGAC (SEQ ID NO: 1); un índice de identificación de nueve nucleótidos específico del locus seleccionado; una secuencia específica de locus o locus/alelo de 9 bases que actúa como índice de locus en el primer conjunto independiente de SNP y como índice de locus/alelo en el segundo conjunto específico de polimorfismo; un nucleótido que interrumpe la hibridación y que es diferente de la base correspondiente en el locus genómico; y una secuencia de 20-24 pb complementaria al locus genómico seleccionado. En los casos en los que se detectó un SNP o una mutación en esta parte del locus genómico seleccionado, el conjunto de consulta específico de alelo consistió en dos primeros cebadores de unión en tandem de secuencia fija, cada uno con un índice de locus/alelo diferente y una base específica de alelo diferente en la posición del SNP. Estos primeros oligonucleótidos se diseñaron para cada ácido nucleico seleccionado para proporcionar una T_f prevista uniforme con una variación de dos grados en todas las consultas realizadas en el conjunto de ensayo.

40 45 50 55 50 El segundo oligonucleótido de secuencia fija, complementario a la región 5' de los locus genómicos, comprendía los siguientes elementos de secuencia (de 5' a 3'): una secuencia de 20-24 bases complementaria a la región 5' del locus genómico; un nucleótido que interrumpe la hibridación que era diferente de la base correspondiente en el locus genómico; y una secuencia de cebado para PCR universal que era común para todos los terceros oligonucleótidos en el conjunto de ensayo: ATTGCGGGGACCGATGATCGCGTC (SEQ ID NO: 2).

60 En los casos en los que se detectó un SNP o una mutación en el locus genómico seleccionado, el conjunto de consulta específico de alelo consistió en dos cebadores de unión en tandem, cada uno con un índice de locus/alelo diferente y una base específica de alelo diferente en la posición de la mutación/SNP. Este segundo oligonucleótido de secuencia fija se diseñó para cada ácido nucleico seleccionado para proporcionar una T_f prevista uniforme con una variación de dos grados en todas las consultas realizadas en el conjunto de ensayo que se encontraba sustancialmente en el mismo intervalo de T_f que el primer conjunto de oligonucleótidos.

En determinados aspectos sometidos a prueba, se usaron uno o más oligonucleótidos puente que eran complementarios a la secuencia del locus genómico entre la región complementaria al primer y segundo oligonucleótidos de secuencia fija usados para cada locus seleccionado. En aspectos específicos sometidos a prueba, se usó más de un oligonucleótido puente para cubrir el hueco entre los oligonucleótidos de secuencia fija, y el uno o más oligonucleótidos puente se pueden diseñar opcionalmente para identificar una o más mutaciones o SNP en la secuencia. La longitud de los oligonucleótidos puente usados en los sistemas de ensayo varió de 5 a 36 pares de bases.

Todos los oligonucleótidos usados en los formatos de unión en tandem se sintetizaron usando química convencional en fase sólida. Los segundos oligonucleótidos de secuencia fija y los oligonucleótidos puente se sintetizaron con restos fosfato en 5' para permitir la unión a los extremos hidroxilo en 3' de oligonucleótidos adyacentes.

Ejemplo 2: Preparación de ADN para su uso en procedimientos de unión en tandem

Se obtuvo ADN genómico de un varón caucásico (NA12801) o de una mujer caucásica (NA11995) de los bancos de células de Coriell (Camden, Nueva Jersey) y se fragmentó por cizallamiento acústico (Covaris, Woburn, MA) hasta un tamaño de fragmento medio de aproximadamente 200 pb.

El ADN de Coriell se biotiniló usando procedimientos estándar. En resumen, el ADN fragmentado de Covaris se sometió a reparación final generando la siguiente reacción en un microtubo de 1,5 ml: 5 µg de ADN, 12 µl de tampón ligasa de T4 10X (Enzymatics, Beverly, MA), 50 U de polinucleótido cinasa de T4 (Enzymatics, Beverly, MA) y H₂O hasta 120 µl. Esto se incubó a 37 °C durante 30 minutos. El ADN se diluyó usando Tris 10 mM, EDTA1 mM, pH 8,5 hasta la concentración final deseada de ~2 ng/µl.

Se dispusieron 5 µl de ADN en cada pocillo de una placa de 96 pocillos, y la placa se selló con un sellador de placa adhesivo y se centrifugó durante 10 segundos a 250 x g. A continuación, la placa se incubó a 95 °C durante 3 minutos, se enfrió a 25 °C y se centrifugó nuevamente durante 10 segundos a 250 x g. Se preparó una mezcla maestra de biotinilación en un microtubo de 1,5 ml hasta una concentración final de: tampón TdT 1X (Enzymatics, Beverly, MA), 8 U de TdT (Enzymatics, Beverly, MA), CoCl₂ 250 µM, 0,01 nmol/µl de biotina-16-dUTP (Roche, Nutley, NJ) y H₂O hasta 1,5 ml. Se añadieron partes alícuotas de 15 µl de la mezcla maestra a cada pocillo de una placa de 96 pocillos, y la placa se selló con sellador de placa adhesivo. La placa se centrifugó durante 10 segundos a 250 x g y se incubó a 37 °C durante 60 minutos. Tras la incubación, la placa se centrifugó nuevamente durante 10 segundos a 250 x g, y se añadieron 7,5 µl de mezcla de precipitación (1 µg/µl de azul dextrano, NaOAc 3 mM) a cada pocillo.

La placa se selló con un sellador de placa adhesivo y se mezcló usando un vórtex de placas IKA durante 2 minutos a 3000 rpm. Se añadieron 27,5 µl de isopropanol a cada pocillo, la placa se selló con sellador de placa adhesivo y se agitó en vórtex durante 5 minutos a 3000 rpm. La placa se centrifugó durante 20 minutos a 3000 x g, el sobrenadante se decantó y la placa se invirtió y centrifugó a 10 x g durante 1 minuto sobre una toallita absorbente.

La placa se secó al aire durante 5 minutos y el sedimento se resuspendió en 30 µl de Tris 10 mM, pH 8,0, EDTA 1 mM.

Ejemplo 3: Formatos de ensayo ejemplares que usan unión en tandem

Se sometieron a prueba numerosos formatos de ensayo de unión en tandem usando el ADN biotinilado para ilustrar la prueba de concepto de los sistemas de ensayo de la invención, y se demostró la capacidad de realizar una detección dirigida altamente multiplexada de un gran número de locus independientes usando la serie de diferentes formatos de ensayo. Los sistemas de ensayo ejemplares de la invención se diseñaron para comprender 96 o más consultas por locus en una muestra genética y, en los casos en los que se detectaron SNP, los formatos de ensayo utilizaron 192 o más consultas separadas, cada una de ellas utilizando la detección de alelos diferentes por 96 locus en muestras genéticas. En los ejemplos descritos para cada formato de ensayo se utilizaron dos conjuntos diferentes de oligonucleótidos de secuencia fija y/o oligonucleótidos puente (como se describe en el ejemplo 1), que comprendían un total de 96 o 192 reacciones de consulta para los locus seleccionados dependiendo de si se identificaron o no SNP.

En un primer formato de ensayo ejemplar se utilizaron oligonucleótidos de secuencia fija específicos de locus y oligonucleótidos puente, donde existía un hueco de una base entre el primer oligonucleótido de secuencia fija y los oligonucleótidos puente y un segundo hueco de una base entre los oligonucleótidos puente y el segundo oligonucleótido de secuencia fija. Cada uno de los dos huecos englobaba dos SNP diferentes. En este formato, se usó una ADN polimerasa para incorporar cada una de las bases de los SNP, y se usó una ligasa para cerrar las discontinuidades (nicks) formadas de este modo. La discriminación de bases en los SNP derivó de la fidelidad de la incorporación de bases por parte de la polimerasa y, en el caso de incorporación errónea, de la tendencia de la ligasa a no cerrar discontinuidades adyacentes a bases mal emparejadas.

En el segundo formato de ensayo ejemplar se usaron dos oligonucleótidos de secuencia fija específicos de locus sin un oligonucleótido puente, donde había un hueco de ~15-35 bases entre los oligonucleótidos de secuencia fija y

5 donde el hueco abarcaba uno o más SNP. En este formato, se usó una polimerasa para incorporar las bases faltantes del hueco, y se usó una ligasa para cerrar la discontinuidad formada de este modo. La discriminación de bases en los SNP derivó de la fidelidad de la incorporación de bases por parte de la polimerasa y, en el caso de incorporación errónea, de la tendencia de la ligasa a no cerrar discontinuidades adyacentes a bases mal emparejadas.

10 En un tercer formato de ensayo ejemplar se usaron un primer y un segundo oligonucleótido de secuencia fija específicos de locus sin un oligonucleótido puente, donde había un hueco de ~15-35 bases entre el primer y segundo oligonucleótidos de secuencia fija y donde el hueco abarcaba uno o más SNP. Se usaron dos primeros 15 oligonucleótidos de secuencia fija específicos de alelo separados y dos segundos oligonucleótidos de secuencia fija específicos de alelo separados. Se usó una polimerasa para incorporar las bases faltantes y se usó una ligasa para cerrar la discontinuidad formada de este modo. La discriminación de bases en los SNP derivó de la especificidad de la hibridación, la tendencia de la polimerasa sin corrección a no extender cebadores hibridados con emparejamientos erróneos cerca del extremo 3' y la tendencia de la ligasa a no cerrar discontinuidades adyacentes a bases mal emparejadas.

20 En un cuarto formato ejemplar se usaron oligonucleótidos de secuencia fija específicos de alelo y un oligonucleótido puente específico de locus. En este formato, se usaron dos oligonucleótidos de secuencia fija separados complementarios al extremo 3' de los locus de interés, el primero con una base en 3' específica de un alelo del SNP seleccionado y el segundo con una base en 3' específica del otro alelo del SNP seleccionado. De forma similar, se usaron dos segundos oligonucleótidos de secuencia fija separados, el primero con una base en 5' específica de un alelo de un segundo SNP seleccionado y el segundo con una base en 5' específica del otro alelo del segundo SNP seleccionado. Los oligonucleótidos puente eran complementarios a la región directamente adyacente a las regiones 25 del locus complementarias al primer y segundo oligonucleótidos de secuencia fija y, por tanto, no se necesitó ninguna polimerasa antes de la unión. Se usó una ligasa para sellar las discontinuidades entre los oligonucleótidos de secuencia fija y el oligonucleótido puente. La discriminación de bases en los SNP en este formato de ensayo derivó de la especificidad de la hibridación y la tendencia de la ligasa a no cerrar las discontinuidades adyacentes a las bases mal emparejadas. Este formato ejemplar se sometió a prueba usando la ligasa de T4 o la Taq ligasa para la creación del molde contiguo, y ambas resultaron eficaces en la reacción como se describe a continuación.

30 35 En un quinto formato ejemplar se usaron oligonucleótidos de secuencia fija específicos de locus que eran complementarios a regiones adyacentes en el ácido nucleico de interés y, por tanto, no se creó ningún hueco por hibridación de estos oligonucleótidos. En este formato, no se requirió ninguna polimerasa y se usó una ligasa para cerrar la única discontinuidad entre los oligonucleótidos.

40 45 En un sexto formato ejemplar se usaron oligonucleótidos de secuencia fija específicos de alelo y oligonucleótidos puente específicos de locus, donde existía un hueco de bases pequeño de cinco bases entre la región de los locus complementaria a los oligonucleótidos de secuencia fija. El oligonucleótido puente específico de locus de este ejemplo era un pentámero complementario a las regiones directamente adyacentes a las regiones complementarias a al primer y segundo oligonucleótidos de secuencia fija. En este formato, no se requirió ninguna polimerasa y se usó una ligasa para cerrar las dos discontinuidades entre los oligonucleótidos.

50 55 En un séptimo formato ejemplar se usaron oligonucleótidos de secuencia fija específicos de locus y un oligonucleótido puente específico de locus, donde existía un hueco de bases más pequeño de cinco bases que contenía un SNP en la región complementaria al oligonucleótido puente. En la reacción de hibridación y unión se incluyeron oligonucleótidos puente específicos de alelo correspondientes a los posibles SNP. En este formato, no se requirió ninguna polimerasa y se usó una ligasa para cerrar las dos discontinuidades entre los oligonucleótidos. La discriminación de bases en los SNP en este formato de ensayo derivó de la especificidad de la hibridación y la tendencia de la ligasa a no cerrar las discontinuidades adyacentes a las bases mal emparejadas.

60 65 En un octavo formato ejemplar se usaron oligonucleótidos de secuencia fija específicos de locus y dos oligonucleótidos puente específicos de locus adyacentes, donde existía un hueco de 10 bases entre las regiones complementarias al primer y segundo oligonucleótidos de secuencia fija. Se incluyeron oligonucleótidos puente específicos de locus en la reacción de unión, requiriendo el hueco dos pentámeros contiguos para cerrar el hueco. En este formato, no se requirió ninguna polimerasa y se usó una ligasa para cerrar las tres discontinuidades entre los oligonucleótidos.

70 Para cada uno de los formatos de ensayo descritos anteriormente, se creó una agrupación equimolar (40 nM cada uno) de conjuntos de primeros y segundos oligonucleótidos de secuencia fija específicos de locus o de alelo a partir de los oligonucleótidos preparados como se establece en el ejemplo 2. Asimismo, se creó una agrupación equimolar separada (20 μ M cada uno) de oligonucleótidos puente para los procedimientos de ensayo en base a las secuencias de los locus genómicos seleccionados.

75 80 Se transfirieron 100 μ g de microesferas de estreptavidina a los pocillos de una placa de 96 pocillos, y se retiró el sobrenadante. Se añadieron 60 μ l de tampón BB2 (Tris 100 mM, pH 8,0, EDTA 10 mM, NaCl₂ 500 mM, formamida al 58 %, Tween-80 al 0,17 %), 10 μ l de agrupación de oligonucleótidos de secuencia fija 40 nM y 30 μ l del ADN molde

biotinilado preparado en el ejemplo 2 a las microesferas. La placa se selló con un sellador de placa adhesivo y se agitó en vórtex a 3000 rpm hasta que se resuspendieron las microesferas. Los oligonucleótidos se hibridaron al ADN molde por incubación a 70 °C durante 5 minutos, seguido de un enfriamiento lento hasta temperatura ambiente.

5 La placa se colocó en una placa magnética con varilla elevada durante 2 minutos para poner las microesferas magnéticas y el ADN asociado en el lateral de los pocillos. El sobrenadante se retiró por pipeteo y se reemplazó por 50 µl de BB2 al 60 % (v/v en agua). Las microesferas se resuspendieron por agitación en vórtex, se colocaron nuevamente sobre el imán y se retiró el sobrenadante. Este procedimiento de lavado de microesferas se repitió una vez usando 50 µl de BB2 al 60 %, y se repitió dos veces más usando 50 µl de tampón de lavado (Tris 10 mM, pH 8,0, EDTA 1 mM, NaCl₂ 50 mM).

10 Las microesferas se resuspendieron en 37 µl de una mezcla de reacción de unión que consistía en tampón Taq ligasa 1X (Enzymatics, Beverly, MA), 1 U de Taq ligasa y una agrupación de oligonucleótidos puente 2 µM (dependiendo del formato de ensayo), y se incubaron a 37 °C durante una hora. Cuando resultó apropiado, y dependiendo del formato de ensayo, en esta mezcla se incluyó una polimerasa termoestable sin corrección más dNTP 200 nM cada uno. La placa se colocó en una placa magnética con varilla elevada durante 2 minutos para poner las microesferas magnéticas y el ADN asociado en el lateral de los pocillos. El sobrenadante se retiró por pipeteo y se reemplazó por 50 µl de tampón de lavado. Las microesferas se resuspendieron por agitación en vórtex, se colocaron nuevamente sobre el imán y se retiró el sobrenadante. Se repitió una vez el procedimiento de lavado.

15 20 Para eluir los productos de las microesferas de estreptavidina, se añadieron 30 µl de Tris 10 mM, EDTA 1 mM, pH 8,0 a cada pocillo de la placa de 96 pocillos. La placa se selló y mezcló usando un vórtex IKA durante 2 minutos a 3000 rpm para resuspender las microesferas. La placa se incubó a 95 °C durante 1 minuto, y el sobrenadante se aspiró usando una pipeta de 8 canales. Se transfirieron 25 µl de sobrenadante de cada pocillo a una nueva placa de 96 pocillos para la amplificación universal.

Ejemplo 4: Amplificación universal de productos unidos en tandem

30 Los ácidos nucleicos polimerizados y/o unidos se amplificaron usando cebadores de PCR universal complementarios a las secuencias universales presentes en el primer y segundo oligonucleótidos de secuencia fija hibridados a los locus de interés. Se usaron 25 µl de cada una de las mezclas de reacción del ejemplo 3 en cada reacción de amplificación. Una reacción de PCR universal de 50 µl que consiste en 25 µl de producto de unión eluido más tampón Pfusion 1X (Finnzymes, Finlandia), betaina 1 M, dNTP 400 nM cada uno, 1 U de ADN polimerasa termoestable con capacidad de corrección de errores Pfusion y los siguientes pares de cebadores:

35 TAATGATACGGCGACCACCGAGATCTACACCGCGCGTATGCGTCGAGA (SEQ ID NO: 3) y TCAAGCAGAAGACGGCATACGAGATXAAACGACGCGATCATCGGTCC CCGCAA (SEQ ID NO: 4), donde X representa uno de los 96 índices de muestra diferentes usados para identificar de forma exclusiva muestras individuales antes del agrupamiento y la secuenciación. La PCR se llevó a cabo en condiciones rigurosas usando un termociclador BioRad Tétrad™.

40 45 Se agruparon 10 µl del producto de la PCR universal de cada una de las muestras y el producto de PCR agrupado se purificó usando microesferas AMPureXP™ SPRI (Beckman-Coulter, Danvers, MA) y se cuantificó usando Quant-iT™ PicoGreen, (Invitrogen, Carlsbad, CA).

Ejemplo 5: Detección y análisis de locus seleccionados

50 Los productos de PCR purificados de cada formato de ensayo se secuenciaron en un único carril de un portaobjetos en un Illumina HiSeq™ 2000 (Illumina, San Diego, CA). Las tandas de secuenciación típicamente dan lugar a ~100 M lecturas brutas, de las que ~85 M (85 %) se cartografián con respecto a las estructuras de ensayo esperadas. Esto se tradujo en un promedio de ~885 K lecturas/muestra en todo el experimento, y (en el caso de un experimento usando 96 locus) 9,2 K lecturas/réplica/locus en 96 locus. Las lecturas cartografiadas se analizaron en recuentos de réplicas/locus/alelos, y se calcularon diversas métricas para cada condición, que incluyen:

55 Rendimiento: una métrica de la proporción de ADN introducido que se consultó en la secuenciación, calculada como el número promedio de lecturas exclusivas por locus (solo contando las lecturas de índices de identificación exclusivas por réplica/locus) dividido entre el número total de equivalentes genómicos contenidos en el ADN introducido.

60 Intervalo de frecuencia del locus en el percentil 80: una métrica de la variabilidad de la frecuencia del locus en los datos de secuenciación, interpretada como el número de intervalos que abarca el 80 % de los locus. Se calculó en base a la distribución de las lecturas totales por locus en todos los locus como el percentil 90 de las lecturas totales por locus dividido entre el percentil 10 de las lecturas totales por locus.

65 Tasa de error en el SNP: una métrica de la tasa de error en la posición del SNP, calculada como la proporción de lecturas que contienen una base discordante en la posición del SNP.

Estos resultados se resumen en la tabla 1:

Tabla 1: Sumario de los resultados de los formatos de ensayo de unión en tándem

5

FORMATO DE ENSAYO	OLIGONUCLEÓTIDO DE SECUENCIA FIJA (1º y/o 2º)	OLIGONUCLEÓTIDO PUENTE USADO	ENZIMA USADA	RENDIMIENTO	INTERVALO DE FREC LOC 80 %	TASA DE ERROR EN SNP
1	ESPECÍFICO DE LOCUS	Específico de locus	pol+lig	9,5 %	5,3	0,18 %
2	ESPECÍFICO DE LOCUS	No	pol+lig	1,4 %	58,3	0,19 %
3	ESPECÍFICO DE ALELO	No	pol+lig	0,4 %	61,7	1,00 %
4	ESPECÍFICO DE ALELO	Específico de locus	Taq lig	5,0 %	5,9	0,92 %
4	ESPECÍFICO DE ALELO	Específico de locus	lig de T4	5,3 %	4,4	0,95 %
5	ESPECÍFICO DE LOCUS	No	Taq lig	22,5%	1,7	N/A
6	ESPECÍFICO DE LOCUS	Específico de locus	Taq lig	12,5	2,9	N/A
7	ESPECÍFICO DE LOCUS	Específico de alelo	Taq lig	14,3	2,8	0,20 %
8	ESPECÍFICO DE LOCUS	Específico de 2 locus	Taq lig	18,5 %	2,8	N/A

La tabla 1 indica que el ensayo de unión en tándem específico de locus en el que se usó un oligonucleótido puente convirtió el ADN molde en un producto seleccionado con alto rendimiento (~10 %), con una alta proporción de producto derivado de los locus seleccionados (el 15 % de las lecturas no contenía las estructuras de ensayo esperadas), con un sesgo de locus limitado (el 80 % de los locus se encontraban dentro de un intervalo de concentración de ~5 veces) y con una alta exactitud para el SNP (tasa de error para SNP del 0,2 %). El ensayo de unión en tándem específico de locus sin el uso de un oligonucleótido puente produjo rendimientos reducidos y un sesgo de locus sustancial, pero todavía arrojó datos de genotipado de SNP de alta exactitud. El ensayo de unión en tándem específico de alelo con un oligonucleótido puente produjo rendimientos intermedios en comparación con el ensayo específico de locus que usaba la Taq ligasa y la ligasa de T4, pero todavía produjo un sesgo de locus limitado y datos de genotipado de SNP de alta exactitud. El ensayo de unión en tándem específico de alelo sin oligonucleótido puente produjo rendimientos reducidos y un sesgo de locus sustancial, pero todavía arrojó datos de genotipado de SNP de alta exactitud.

20 Los formatos de ensayo seis al ocho mostraron que el ADN molde se puede convertir en un producto seleccionado con un alto rendimiento (12-18 %), con una alta proporción de producto derivado de los locus seleccionados (~76 % de las lecturas contenían las estructuras de ensayo esperadas) y con un sesgo de locus limitado (el 80 % de los locus se encontraban dentro de un intervalo de concentración de 2-3 veces). La figura 5 ilustra el resultado del genotipado que se obtuvo usando el formato de ensayo siete, en el que se compararon los recuentos de secuencia para los dos alelos de todos los ensayos polimórficos observados en una única muestra. Nótese la clara separación de los grupos homocigóticos y heterocigóticos, así como los bajos recuentos de fondo observados entre los grupos homocigóticos.

Ejemplo 6: Determinación del porcentaje de ADN fetal usando la unión en tándem

30

Se diseñó un sistema de ensayo ejemplar de la invención para determinar el porcentaje de concentración de ADN fetal en una muestra genética, así como para proporcionar recuentos de locus seleccionados dentro de la muestra. Este ensayo ejemplar comprendió 480 consultas separadas, cada una de las cuales utilizó la detección de diferentes locus en una muestra materna. En el ejemplo inicial se utilizó una determinación del porcentaje de ADN fetal en sujetos portadores de un feto de sexo masculino, por lo que se utilizaron locus en el cromosoma Y así como locus que contenían un SNP fetal heredado del padre que era diferente de la secuencia materna.

40 Específicamente, se consultaron 480 ácidos nucleicos seleccionados usando el sistema de ensayo. Los 480 ácidos nucleicos seleccionados comprendían 48 consultas específicas de secuencia de ácidos nucleicos correspondientes a locus en el cromosoma Y, 192 consultas específicas de secuencia de ácidos nucleicos correspondientes a locus en el cromosoma 21, 192 consultas específicas de secuencia de ácidos nucleicos seleccionados correspondientes a locus en el cromosoma 18 y 144 consultas específicas de secuencia de ácidos nucleicos seleccionados correspondientes a locus polimórficos en los cromosomas 1-16. Estos ensayos se diseñaron en base a secuencias

genómicas humanas, y en cada consulta se usaron tres oligonucleótidos por cada ácido nucleico seleccionado consultado en el ensayo.

5 El primer oligonucleótido usado para cada consulta era complementario a la región 3' de la región genómica seleccionada y comprendía los siguientes elementos de secuencia oligonucleotídica (de 5' a 3'): una secuencia de cebado para PCR universal común para todos los ensayos: **TACACCGCGTATCGTCGAGAC** (SEQ ID NO: 1); un índice de identificación específico de los locus seleccionados que comprendía nueve nucleótidos; y una secuencia de 20-24 pb complementaria al locus genómico seleccionado. Este primer oligonucleótido se diseñó para cada ácido nucleico seleccionado para proporcionar una T_f prevista uniforme con una variación de dos grados en 10 todas las consultas realizadas en el conjunto de 480 ensayos.

15 El segundo oligonucleótido usado para cada consulta fue un oligonucleótido puente complementario a la secuencia del locus genómico directamente adyacente a la región genómica complementaria al primer oligonucleótido. En base a los ácidos nucleicos de interés seleccionados, los oligonucleótidos puente se diseñaron para permitir la utilización de un total de 12 secuencias de oligonucleótidos que pudieran servir como oligonucleótidos puente para las 480 consultas en el conjunto de ensayo.

20 El tercer oligonucleótido usado para cada consulta era complementario a la región 5' del locus genómico seleccionado y comprendía los siguientes elementos de secuencia (de 5' a 3'): una secuencia de 20-24 bases complementaria a la región 5' del locus genómico; un nucleótido que interrumpe la hibridación que era diferente de la base correspondiente en el locus genómico; y una secuencia de cebado para PCR universal que era común para todos los terceros oligonucleótidos en el conjunto de ensayo: **ATTGCGGGGACCGATGATCGCGTC** (SEQ ID NO: 2). Este tercer oligonucleótido se diseñó para cada ácido nucleico seleccionado para proporcionar una T_f prevista uniforme con una variación de dos grados en todas las consultas realizadas en el conjunto de 480 ensayos, 25 y el intervalo de T_f se encontraba sustancialmente en el mismo intervalo de T_f que el primer conjunto de oligonucleótidos.

30 Todos los oligonucleótidos se sintetizaron usando química de fase sólida convencional. El primer oligonucleótido y los oligonucleótidos puente se sintetizaron con restos fosfato en 5' para permitir la unión a los extremos hidroxilo en 3' de oligonucleótidos adyacentes. Se creó una agrupación equimolar de conjuntos del primer y tercer oligonucleótidos usados para todas las consultas en el ensayo multiplexado, y se creó una agrupación equimolar separada de todos los oligonucleótidos puente para permitir reacciones de hibridación separadas.

35 Se aisló ADN genómico a partir de 5 ml de plasma usando el kit Dynal Silane viral NA (Invitrogen, Carlsbad, CA). Se procesaron aproximadamente 12 ng de ADN de cada una de las 37 mujeres, que incluían 7 mujeres no embarazadas, 10 mujeres embarazadas con fetos de sexo masculino y 22 mujeres embarazadas con fetos de sexo femenino. El ADN se biotiniló usando procedimientos estándar y el ADN biotinilado se inmovilizó sobre una superficie sólida recubierta con estreptavidina para permitir la retención del ADN genómico en las etapas posteriores del ensayo.

40 El ADN inmovilizado se hibridó en condiciones de hibridación rigurosas con la primera agrupación que comprendía el primer y tercer oligonucleótidos para cada secuencia consultada. A continuación, los oligonucleótidos no hibridados de la agrupación se lavaron de la superficie del soporte sólido y el ADN inmovilizado se hibridó en condiciones de hibridación rigurosas con la agrupación que comprendía los oligonucleótidos puente. Una vez que se permitió que 45 los oligonucleótidos puente hibridaran con el ADN inmovilizado, los oligonucleótidos restantes no unidos se lavaron de la superficie y los tres oligonucleótidos hibridados unidos a los locus seleccionados se unieron usando ligasa de T4 para proporcionar un molde de ADN contiguo para la amplificación.

50 El ADN unido se amplificó a partir del sustrato sólido usando una ADN polimerasa termoestable con capacidad de corrección de errores, un primer cebador de PCR universal **TAATGATACGGCACCACCGAGATCTACACCGCGTATCGTCGA** GA (SEQ ID NO: 3) y un segundo cebador de PCR universal **TCAAGCAGAACGGCATACGAGATXAAACGACGCGATCGGTC** CCCGCAA (SEQ ID NO: 4), donde **X** representa uno de los 96 índices de muestra diferentes usados para identificar de forma exclusiva muestras individuales antes del agrupamiento y la secuenciación. Se agruparon 10 μ l del producto de la PCR universal de cada una de las 37 muestras descritas anteriormente, y el producto de PCR agrupado se purificó usando microesferas AMPure™ SPRI (Beckman-Coulter, Danvers, MA) y se cuantificó usando Quant-iT™ PicoGreen, (Invitrogen, Carlsbad, CA).

60 El producto de PCR purificado se secuenció en 6 carriles de un único portaobjetos en un Illumina HiSeq™ 2000. La tanda de secuenciación dio lugar a 384 M de lecturas brutas, de las cuales 343 M (89 %) se cartografiaron en los locus genómicos esperados, lo que dio como resultado un promedio de 3,8 M de lecturas por muestra en las 37 muestras y 8 K de lecturas por muestra por locus en los 480 locus. Las lecturas cartografiadas se analizaron en recuentos de muestras y locus, y se calcularon dos métricas separadas del porcentaje de ADN fetal como sigue.

5 El porcentaje de ADN de varón detectado por los locus del cromosoma Y corresponde a la proporción relativa de lecturas derivadas de las consultas de locus del cromosoma Y frente a la proporción relativa de lecturas derivadas de las consultas de locus autosómicos, y se calculó como (número de lecturas del cromosoma Y en un sujeto de prueba/número de lecturas autosómicas en el sujeto de prueba)/(número de lecturas en el sujeto control varón/número de lecturas autosómicas en el sujeto control varón). Esta métrica se usó como medida del porcentaje de ADN fetal en el caso de un feto de sexo masculino usando las lecturas relativas del cromosoma Y.

10 10 El porcentaje de ADN fetal detectado por locus polimórficos corresponde a la proporción de lecturas derivadas de alelos no maternos frente a alelos maternos en los locus donde se puede hacer dicha distinción. En primer lugar, se dividió, para cada locus identificado, el número de lecturas del alelo con el menor número de recuentos (el alelo de baja frecuencia) entre el número total de lecturas para proporcionar una frecuencia de alelo menor (MAF) para cada locus. A continuación, los locus con una MAF entre 0,075 % y 15 % se identificaron como locus informativos. El 15 porcentaje de ADN fetal estimado para la muestra se calculó como la media de la frecuencia de alelo menor de los locus informativos multiplicada por dos, es decir, se calculó como 2X la ocurrencia (de MAF) promedio, donde 0,075 % < MAF < 15 %.

20 20 La figura 6 muestra los resultados de estos cálculos. Como se muestra en la figura 6, el porcentaje de locus masculinos determinado usando las métricas del cromosoma Y descritas anteriormente (círculos grises) permite separar los embarazos que involucran a fetos de sexo masculino de los embarazos que involucran a fetos de sexo femenino (rombos grises) y las muestras de mujeres no embarazadas (círculos negros). Además, el cálculo de la cantidad porcentual fetal en una muestra mediante la métrica de locus polimórficos permite distinguir las muestras de mujeres embarazadas de las de mujeres no embarazadas. Finalmente, existió una correlación entre las estimaciones del porcentaje de ADN fetal para una muestra obtenida locus del cromosoma Y y locus polimórficos en 25 embarazos que involucraban a fetos de sexo masculino. Esta correlación persiste hasta valores fetales porcentuales bastante bajos.

Ejemplo 7: Detección de aneuploidía en una muestra materna

30 30 Los sistemas de ensayo de la invención se usaron para la detección de polimorfismos y anomalías cromosómicas en dos cohortes separadas de mujeres embarazadas. Se sometieron a prueba una primera cohorte de 190 embarazos normales, 36 con T21 y 8 con T18 y una segunda cohorte de 126 embarazos normales, 36 con T21 y 8 con T18 para detectar aneuploidía fetal. Las aneuploidías cromosómicas se detectaron usando 576 ensayos del cromosoma 21 y 576 del cromosoma 18, agrupados y sometidos a ensayo en una única reacción, como se expone a continuación.

35 35 Los elementos usados en los ensayos de detección de aneuploidía se ilustran en la FIG. 7. El ADNlc 701 aislado de muestras maternas se usó como molde para la hibridación, unión y amplificación de múltiples locus seleccionados tanto del cromosoma 21 como del cromosoma 18 en cada muestra materna. Se hibridaron tres oligonucleótidos con cada locus seleccionado para crear productos de unión para amplificación y detección. El oligonucleótido de secuencia fija izquierdo (o primero) comprendía una región complementaria a un locus seleccionado 709 y una primera región de cebador universal 711. El oligonucleótido de secuencia fija derecho 705 (o segundo) comprendía una segunda región complementaria al locus seleccionado 713 y una segunda región de cebador universal 715. Los oligonucleótidos puente 707 usados se diseñaron de modo que cada uno hibridara con las regiones puente de dos o más locus seleccionados usados en el ensayo de detección de aneuploidía. Cuando los oligonucleótidos de secuencia fija 703, 705 y el oligonucleótido puente 707 hibridaron con la región complementaria en el ADNlc 701, sus extremos formaron dos discontinuidades. Tras la unión de los oligonucleótidos hibridados con el ADNlc, se creó 40 40 un producto de unión para cada locus seleccionado que comprendía 703, 705 y 707 y que se usó como molde para los cebadores de amplificación 719, 721.

45 45 A continuación, se utilizaron dos cebadores de amplificación 719, 721 que comprendían regiones complementarias a la primera y segunda regiones de cebador universal, respectivamente, para amplificar el producto de unión. Este producto de amplificación comprendía la secuencia del locus seleccionado. El cebador de amplificación derecho también comprendía un índice de muestra 717 para identificar la muestra particular de la que se obtuvo el locus en el ensayo multiplexado. La amplificación con 96 cebadores de amplificación derechos 729 distintos permitió el agrupamiento y la secuenciación simultánea de 96 productos de amplificación diferentes en un único carril.

50 50 55 Los cebadores de amplificación 719, 721 también contenían una secuencia clúster izquierda 723 (TAATGATACGGCGACCACCGA) (SEQ ID NO: 7) y una secuencia clúster derecha 725 (ATCTCGTATGCCGTCTCTGCTTGA) (SEQ ID NO: 8) que admitían la amplificación de clústeres para secuenciación utilizando el sistema Illumina HiSeq™ 2000 (Illumina, San Diego, CA). Se usó un cebador de secuenciación 727 que comprendía la primera secuencia de cebador universal para determinar la secuencia del producto de amplificación, y se usó un segundo cebador de secuenciación 729 para determinar el índice de muestra 717 del producto de amplificación.

60 60 65 En resumen, se recogieron aproximadamente 10 ml de sangre periférica de cada paciente en un tubo BCT (Streck, Omaha, NE), que se envió por mensajería nocturna a Tandem Diagnostics. El plasma se aisló de los tubos BCT dentro de las 72 h siguientes a la extracción de sangre mediante centrifugación a 1600 x g durante 10 min. El

plasma se transfirió a un segundo tubo y se centrifugó a 16000 x g durante 10 min para eliminar las células restantes. Se aisló el ADNlc de 4-5 ml de plasma por paciente. Se aislaron aproximadamente 15 ng de ADNlc de cada muestra de paciente y se dispusieron en pocillos individuales de una placa de 96 pocillos. Todo el procesamiento posterior se realizó en lotes multiplexados de hasta 96 muestras de ADNlc de pacientes por procedimiento efectuado con el sistema de matriz.

5 El ADNlc aislado de las muestras maternas en cada pocillo se biotiniló, se precipitó y se resuspendió en 30 μ l de TE como en el ejemplo 3 anterior. El ADN molde biotinilado se mezcló con 100 μ g de microesferas magnéticas recubiertas con estreptavidina MyOneCl (Life Technologies, Carlsbad, CA), 60 μ l de tampón BB2 (Tris 100 mM pH

10 8,0, EDTA 10 mM, NaCl₂ 500 mM, formamida al 58 %, Tween-80 al 0,17 %) y 10 μ l de oligonucleótidos de secuencia fija izquierdo 703 y derecho 705 agrupados 40 nM. La mezcla se calentó a 70 °C y se enfrió durante 2 horas. A continuación, las microesferas se inmovilizaron magnéticamente en el lateral del pocillo, se lavaron dos veces con 50 μ l de BB2 al 60 % (v/v con H₂O), se lavaron dos veces más con 50 μ l de tampón de lavado (Tris 10 mM pH 8,0, EDTA 1 mM, NaCl₂ 500 mM) y, a continuación, se resuspendieron en 50 μ l de una mezcla de reacción que contenía

15 1 U de Taq ligasa (Enzymatics, Beverly, MA), tampón de Taq ligasa 1X (Enzymatics) y un oligonucleótido puente pentamérico fosforilado en 5' 707 10 μ M. La mezcla se incubó a 37 °C durante 1 hora. Las microesferas se inmovilizaron de nuevo magnéticamente en un lateral del pocillo, se lavaron dos veces con 50 μ l de tampón de lavado y, a continuación, se resuspendieron en 30 μ l de TE.

20 Los productos de unión se eluyeron de las microesferas inmovilizadas mediante incubación a 95 °C durante 3 minutos. Los productos de unión eluidos se amplificaron mediante 26 ciclos de PCR en 50 μ l de una mezcla de reacción que contenía 1 U de polimerasa Pfusion (Thermo Fisher, Waltham, MA), betaina 1 M, tampón Pfusion 1X y cebadores de amplificación izquierdo y derecho 400 nM (719 y 721, respectivamente). El cebador derecho contenía un índice de muestra de 7 bases (717) que permitía la secuenciación multiplexada de 96 muestras en HiSeq2000

25 (Illumina, San Diego, CA). La secuencia del oligonucleótido de secuencia fija izquierdo era:
TAATGATACGGCGACCGAGATCTACACCGCGTTATGCGTCGAGA C (SEQ ID NO: 5).

Y la secuencia del oligonucleótido de secuencia fija derecho era:

30 TCAAGCAGAAGACGGCATACGAGATNNNNNNAAACGACGCGATCATCG GTCCCCGCAAT (SEQ ID NO: 6).

Los productos de amplificación de una única placa de 96 pocillos se agruparon en un volumen igual y los productos de amplificación agrupados se purificaron con microesferas AMPureXP™ SPRI (Beckman-Coulter, Danvers, MA) de acuerdo con las instrucciones del fabricante. Cada colección agrupada purificada se usó como molde para la amplificación de clústeres en una cubeta de lectura de clústeres del kit Illumina TruSeq v2 SR (Illumina, San Diego, CA) de acuerdo con los protocolos del fabricante. El portaobjetos se procesó en un Illumina HiSeq™ 2000 (Illumina, San Diego, CA) para producir 56 bases de la secuencia específica de locus a partir de un cebador de secuencia izquierdo 723 y se obtuvo una lectura separada de 8 bases de la secuencia específica de muestra a partir del segundo cebador de secuencia 725. Se recogió un promedio de 903 K lecturas brutas por muestra. Se asignó un promedio de 876 K (97 %) lecturas a las estructuras de ensayo esperadas.

35 40 Los productos de amplificación de una única placa de 96 pocillos se agruparon en un volumen igual y los productos de amplificación agrupados se purificaron con microesferas AMPureXP™ SPRI (Beckman-Coulter, Danvers, MA) de acuerdo con las instrucciones del fabricante. Cada colección agrupada purificada se usó como molde para la amplificación de clústeres en una cubeta de lectura de clústeres del kit Illumina TruSeq v2 SR (Illumina, San Diego, CA) de acuerdo con los protocolos del fabricante. El portaobjetos se procesó en un Illumina HiSeq™ 2000 (Illumina, San Diego, CA) para producir 56 bases de la secuencia específica de locus a partir de un cebador de secuencia izquierdo 723 y se obtuvo una lectura separada de 8 bases de la secuencia específica de muestra a partir del segundo cebador de secuencia 725. Se recogió un promedio de 903 K lecturas brutas por muestra. Se asignó un promedio de 876 K (97 %) lecturas a las estructuras de ensayo esperadas.

45 50 La FIG. 8 muestra datos ejemplares para un subconjunto de muestras de pacientes de la segunda cohorte, que se analizaron todas en un ensayo multiplexado en un único carril de una tanda de secuenciación. Inicialmente, se corrieron 96 muestras diferentes en esta tanda en particular, pero luego se excluyeron seis muestras de este conjunto analítico por no cumplir con los umbrales de control de calidad de la muestra. En base a las lecturas obtenidas en el ensayo se calculó una media truncada para cada cromosoma 18 y cromosoma 21 en las muestras. La media truncada se calculó eliminando el 10 % de los recuentos altos y bajos para cada cromosoma por muestra. Los productos de amplificación detectados correspondientes a los diversos locus seleccionados se usaron para calcular una métrica de proporción para el cromosoma 21 y una métrica de proporción para el cromosoma 18 para cada muestra. Para la proporción del cromosoma 21, esta se calculó como la media truncada de los recuentos en los 384 locus seleccionados del cromosoma 21 dividida entre la suma de las medias truncadas de los recuentos de los 576 locus del cromosoma 21 y los 576 locus del cromosoma 18 para cada muestra.

55 60 En promedio, se observaron 834 recuentos de lectura por locus seleccionado en las muestras maternas de la primera cohorte, y se observaron 664 recuentos de lectura por locus seleccionado de la segunda cohorte. Estos recuentos se usaron para calcular las puntuaciones z de la proporción de cromosomas para el cromosoma 21 y el cromosoma 18.

65 70 En resumen, las puntuaciones z se calcularon ajustando a escala la mediana por recuento de locus a un valor común (por ejemplo, 1000) para cada muestra, y los recuentos ajustados a escala se transformaron a base logarítmica 2. Se realizó un modelado logarítmico lineal del RMA y un recorte de mediana (median polish) (Bolstad, B. M *et al.* (2003) *Bioinformatics* 19(2): 185-193; Rafael, A. (2003) *Nucleic Acids Research* 31(4):e15; Irizarry, RA *et al.* (2003) *Biostatistics* 4(2):249-64) para estimar los efectos cromosómicos, los efectos de los loci, los efectos de la muestra y los residuales. Los efectos cromosómicos estimados se establecieron en un valor común, por ejemplo 0, y se calculó 2^(efecto cromosómico + efecto de la muestra + residual) para cada locus para crear recuentos

normalizados. Las puntuaciones z se ajustaron a escala mediante censura iterativa de modo que tuvieran una media de 0 y una desviación estándar de 1.

5 Los datos obtenidos de la primera cohorte de muestras se usaron para determinar las puntuaciones z de la primera cohorte para el cromosoma 21 y el cromosoma 18, que se ilustran en las FIGs. 9 y 10, respectivamente. Las muestras normales se muestran como rombos de color gris oscuro y las muestras con una trisomía se muestran como rombos de color gris claro. 179/180 (99,4 %) muestras normales (rombos de color gris oscuro) obtuvieron puntuaciones z <3; una muestra normal obtuvo una puntuación z para el cromosoma 21 de 3,4 y una puntuación z para el cromosoma 18 de 3,0. 35/35 (100 %) muestras de T21 y 7/7 (100 %) muestras de T18 obtuvieron puntuaciones z para la proporción cromosómica >3. La puntuación z media para T18 fue de 8,5 y el intervalo fue de 5,8 a 10,9. La puntuación z media para T21 fue de 11,5 y el intervalo fue de 6,1 a 19,8.

10 15 Los datos proporcionados en la FIG. 8 se combinaron con los datos de las muestras restantes de la segunda cohorte para determinar las puntuaciones z para el cromosoma 21 y el cromosoma 18, que se ilustran en las FIGs. 11 y 12, respectivamente. Las muestras normales se muestran como rombos de color gris oscuro y las muestras con una trisomía se muestran como rombos de color gris claro. 125/125 muestras normales obtuvieron puntuaciones z <3, 36/36 (100 %) muestras de T21 y 8/8 (100 %) muestras de T18 obtuvieron puntuaciones z >3. La puntuación z media para T18 fue de 9,5 y el intervalo fue de 5,1 a 19,8. La puntuación z media para T21 fue de 11,4 y el intervalo fue de 3,4 a 21,8.

20 25 Además de la detección de aneuploidía en estas cohortes, también se usaron polimorfismos específicos para determinar el porcentaje de contribución fetal a las muestras maternas. La metodología general usada para la determinación de estos porcentajes de contribución fetal se describe en el documento de EE. UU. con el n.º de serie 61/509.188 presentado el 19 de julio de 2011, que se incorpora en su totalidad por referencia.

30 35 En resumen, la secuenciación de determinados locus que tienen polimorfismos detectables identificó estos locus como informativos. Los recuentos de los locus identificados que tienen regiones polimórficas fetales diferentes de las regiones polimórficas maternas se usaron para calcular la contribución fetal aproximada en la muestra materna. Cada uno de los locus usados en el cálculo del porcentaje de contribución fetal a la muestra materna tenía un mínimo de 256 recuentos. Los conjuntos de datos de SNP ejemplares para este cálculo se ilustran a continuación en las tablas 2 y 3. Los datos correspondientes a los locus informativos identificados de estos conjuntos se usaron para el cálculo del porcentaje de contribución. Los locus informativos se muestran en cada tabla en texto en negrita.

Tabla 2: Detección de SNP y porcentaje fetal calculado para la muestra materna 1

Cromosoma/Locus	Recuentos SNP_1	Recuentos SNP_2	Recuentos totales	Porcentaje fetal calculado
Ch01_Lc067487	294	26	320	0,210138
Ch01_Lc067489	187	167	354	
Ch01_Lc067490	389	1	390	
Ch01_Lc067491	233	113	346	
Ch01_Lc067492	0	267	267	
Ch01_Lc067493	145	132	277	
Ch01_Lc067495	106	172	278	
Ch01_Lc067496	308	28	336	
Ch01_Lc067497	298	0	298	
Ch01_Lc067498	310	1	311	
Ch01_Lc067499	256	1	257	
Ch01_Lc067501	26	273	299	
Ch01_Lc067503	296	0	296	
Ch01_Lc067504	134	149	283	
Ch02_Lc067508	0	337	337	
Ch02_Lc067510	37	324	361	
Ch02_Lc067511	138	147	285	
Ch02_Lc067512	180	251	431	
Ch02_Lc067514	0	383	383	
Ch02_Lc067515	316	31	347	
Ch02_Lc067516	276	2	278	
Ch02_Lc067519	42	276	318	
Ch02_Lc067521	312	47	359	
Ch02_Lc067522	158	170	328	
Ch02_Lc067523	38	328	366	
Ch02_Lc067524	177	127	304	
Ch02_Lc067525	292	0	292	
Ch02_Lc067526	361	0	361	

ES 2 863 778 T3

Ch02_Lc067527	261	26	287
Ch02_Lc067529	140	146	286
Ch03_Lc067530	0	268	268
Ch03_Lc067531	217	178	395
Ch03_Lc067532	245	153	398
Ch03_Lc067533	1	286	287
Ch03_Lc067534	384	38	422
Ch03_Lc067535	192	114	306
Ch03_Lc067537	32	276	308
Ch03_Lc067538	243	15	258
Ch03_Lc067539	132	247	379
Ch03_Lc067540	162	105	267
Ch03_Lc067541	239	35	274
Ch03_Lc067542	3	406	409
Ch03_Lc067544	2	271	273
Ch03_Lc067545	373	0	373
Ch03_Lc067546	354	0	354
Ch03_Lc067547	1	256	257
Ch03_Lc067548	365	0	365
Ch03_Lc067549	187	111	298
Ch04_Lc067550	33	312	345
Ch04_Lc067552	323	1	324
Ch04_Lc067553	217	119	336
Ch04_Lc067557	35	236	271
Ch04_Lc067558	184	166	350
Ch04_Lc067559	295	32	327
Ch04_Lc067560	140	141	281
Ch04_Lc067561	160	123	283
Ch04_Lc067562	313	2	315
Ch04_Lc067566	142	191	333
Ch04_Lc067569	117	206	323
Ch05_Lc067570	0	403	403
Ch05_Lc067571	229	219	448
Ch05_Lc067572	185	134	319
Ch05_Lc067573	271	22	293
Ch05_Lc067575	261	142	403
Ch05_Lc067578	0	399	399
Ch05_Lc067579	307	46	353
Ch05_Lc067581	189	109	298
Ch05_Lc067582	0	268	268
Ch05_Lc067583	167	203	370
Ch05_Lc067585	209	119	328
Ch05_Lc067586	3	327	330
Ch05_Lc067587	321	0	321
Ch06_Lc067589	286	0	286
Ch06_Lc067590	2	344	346
Ch06_Lc067591	124	179	303
Ch06_Lc067592	0	330	330
Ch06_Lc067593	0	286	286
Ch06_Lc067594	396	2	398
Ch06_Lc067595	349	0	349
Ch06_Lc067597	340	1	341
Ch06_Lc067598	0	412	412
Ch06_Lc067599	182	93	275
Ch06_Lc067600	44	307	351
Ch06_Lc067601	43	324	367
Ch06_Lc067602	358	1	359
Ch07_Lc067603	160	141	301
Ch07_Lc067604	302	0	302
Ch07_Lc067605	37	414	451
Ch07_Lc067606	269	290	559
Ch07_Lc067607	166	159	325
Ch07_Lc067609	1	396	397
Ch07_Lc067610	225	134	359
Ch07_Lc067611	48	391	439
Ch07_Lc067612	2	333	335

ES 2 863 778 T3

Ch07_Lc067614	200	246	446
Ch07_Lc067615	188	184	372
Ch07_Lc067616	167	116	283
Ch07_Lc067617	204	186	390
Ch07_Lc067618	281	28	309
Ch07_Lc067619	44	297	341
Ch07_Lc067620	336	0	336
Ch07_Lc067621	48	342	390
Ch08_Lc067622	313	1	314
Ch08_Lc067623	414	0	414
Ch08_Lc067624	230	142	372
Ch08_Lc067625	0	377	377
Ch08_Lc067626	41	357	398
Ch08_Lc067627	133	258	391
Ch08_Lc067628	388	1	389
Ch08_Lc067629	348	0	348
Ch08_Lc067630	37	314	351
Ch08_Lc067631	185	129	314
Ch08_Lc067632	49	308	357
Ch08_Lc067633	186	195	381
Ch08_Lc067634	174	217	391
Ch08_Lc067635	161	152	313
Ch08_Lc067637	0	284	284
Ch08_Lc067638	343	50	393
Ch09_Lc067639	164	99	263
Ch09_Lc067640	185	186	371
Ch09_Lc067641	344	0	344
Ch09_Lc067642	294	0	294
Ch09_Lc067643	36	336	372
Ch09_Lc067644	221	144	365
Ch09_Lc067645	315	36	351
Ch09_Lc067646	141	143	284
Ch09_Lc067647	33	270	303
Ch09_Lc067648	43	349	392
Ch09_Lc067649	147	152	299
Ch09_Lc067650	201	187	388
Ch09_Lc067651	176	151	327
Ch10_Lc067652	29	277	306
Ch10_Lc067653	134	157	291
Ch10_Lc067654	174	196	370
Ch10_Lc067655	189	181	370
Ch10_Lc067656	125	174	299
Ch10_Lc067657	375	2	377
Ch10_Lc067658	0	345	345
Ch10_Lc067659	204	174	378
Ch10_Lc067661	236	271	507
Ch10_Lc067662	39	325	364
Ch11_Lc067663	2	378	380
Ch11_Lc067664	42	298	340
Ch11_Lc067666	196	200	396
Ch11_Lc067667	220	164	384
Ch11_Lc067668	20	290	310
Ch11_Lc067670	1	356	357
Ch12_Lc067671	212	195	407
Ch12_Lc067673	1	298	299
Ch12_Lc067674	242	30	272
Ch12_Lc067675	2	292	294
Ch12_Lc067676	1	381	382
Ch12_Lc067677	139	179	318

Tabla 3: Detección de SNP y porcentaje fetal calculado para la muestra 2

Cromosoma/Locus	Recuentos SNP_1	Recuentos SNP_2	Recuentos totales	Porcentaje fetal calculado
Ch01_Lc067487	181	134	315	0,096075
Ch01_Lc067489	18	337	355	
Ch01_Lc067490	17	333	350	
Ch01_Lc067491	356	0	356	
Ch01_Lc067492	12	264	276	
Ch01_Lc067493	7	385	392	
Ch01_Lc067494	140	118	258	
Ch01_Lc067495	163	118	281	
Ch01_Lc067496	198	172	370	
Ch01_Lc067498	15	301	316	
Ch01_Lc067499	162	170	332	
Ch01_Lc067501	11	252	263	
Ch01_Lc067502	129	130	259	
Ch01_Lc067503	148	172	320	
Ch01_Lc067504	157	146	303	
Ch02_Lc067508	188	196	384	
Ch02_Lc067510	1	356	357	
Ch02_Lc067511	3	308	311	
Ch02_Lc067512	262	193	455	
Ch02_Lc067513	318	0	318	
Ch02_Lc067514	1	440	441	
Ch02_Lc067515	169	166	335	
Ch02_Lc067516	189	149	338	
Ch02_Lc067519	218	133	351	
Ch02_Lc067521	153	180	333	
Ch02_Lc067522	14	326	340	
Ch02_Lc067523	180	178	358	
Ch02_Lc067524	330	1	331	
Ch02_Lc067526	202	185	387	
Ch02_Lc067527	149	192	341	
Ch02_Lc067529	140	160	300	
Ch03_Lc067530	132	128	260	
Ch03_Lc067531	20	392	412	
Ch03_Lc067532	202	270	472	
Ch03_Lc067533	328	0	328	
Ch03_Lc067534	224	223	447	
Ch03_Lc067535	188	142	330	
Ch03_Lc067537	315	1	316	
Ch03_Lc067538	265	0	265	
Ch03_Lc067539	191	214	405	
Ch03_Lc067540	166	124	290	
Ch03_Lc067542	256	208	464	
Ch03_Lc067543	139	123	262	
Ch03_Lc067544	190	180	370	
Ch03_Lc067545	378	12	390	
Ch03_Lc067546	22	352	374	
Ch03_Lc067547	184	162	346	
Ch03_Lc067548	363	0	363	
Ch03_Lc067549	0	312	312	
Ch04_Lc067550	162	156	318	
Ch04_Lc067552	331	0	331	
Ch04_Lc067553	225	155	380	
Ch04_Lc067555	121	146	267	

ES 2 863 778 T3

Ch04_Lc067557	284	13	297
Ch04_Lc067558	229	208	437
Ch04_Lc067559	7	311	318
Ch04_Lc067560	154	136	290
Ch04_Lc067561	12	258	270
Ch04_Lc067562	410	1	411
Ch04_Lc067566	320	18	338
Ch04_Lc067569	0	289	289
Ch05_Lc067570	19	444	463
Ch05_Lc067571	498	0	498
Ch05_Lc067572	169	182	351
Ch05_Lc067573	294	0	294
Ch05_Lc067575	422	0	422
Ch05_Lc067578	18	388	406
Ch05_Lc067579	17	304	321
Ch05_Lc067580	156	149	305
Ch05_Lc067581	303	19	322
Ch05_Lc067583	23	347	370
Ch05_Lc067585	22	293	315
Ch05_Lc067586	391	0	391
Ch05_Lc067587	434	1	435
Ch05_Lc067588	157	129	286
Ch06_Lc067589	274	0	274
Ch06_Lc067590	23	320	343
Ch06_Lc067591	10	342	352
Ch06_Lc067592	181	177	358
Ch06_Lc067593	0	296	296
Ch06_Lc067594	267	200	467
Ch06_Lc067595	212	201	413
Ch06_Lc067596	329	12	341
Ch06_Lc067597	319	1	320
Ch06_Lc067598	243	186	429
Ch06_Lc067600	0	341	341
Ch06_Lc067601	417	0	417
Ch06_Lc067602	1	340	341
Ch07_Lc067603	168	185	353
Ch07_Lc067604	333	11	344
Ch07_Lc067605	211	287	498
Ch07_Lc067606	2	542	544
Ch07_Lc067607	310	0	310
Ch07_Lc067609	178	189	367
Ch07_Lc067610	425	0	425
Ch07_Lc067611	1	449	450
Ch07_Lc067612	169	149	318
Ch07_Lc067614	0	446	446
Ch07_Lc067615	0	427	427
Ch07_Lc067616	13	278	291
Ch07_Lc067617	239	246	485
Ch07_Lc067618	17	267	284
Ch07_Lc067619	0	335	335
Ch07_Lc067620	319	0	319
Ch07_Lc067621	0	354	354
Ch08_Lc067622	25	232	257
Ch08_Lc067623	470	0	470
Ch08_Lc067624	0	376	376
Ch08_Lc067625	212	202	414
Ch08_Lc067626	377	0	377

Ch08_Lc067627	379	16	395
Ch08_Lc067628	189	210	399
Ch08_Lc067629	16	338	354
Ch08_Lc067630	152	153	305
Ch08_Lc067631	0	379	379
Ch08_Lc067632	25	355	380
Ch08_Lc067633	186	236	422
Ch08_Lc067634	2	375	377
Ch08_Lc067635	169	159	328
Ch08_Lc067636	13	274	287
Ch08_Lc067637	373	11	384
Ch08_Lc067638	28	431	459
Ch09_Lc067640	367	0	367
Ch09_Lc067641	359	0	359
Ch09_Lc067642	307	1	308
Ch09_Lc067643	350	32	382
Ch09_Lc067644	168	241	409
Ch09_Lc067645	174	179	353
Ch09_Lc067646	133	177	310
Ch09_Lc067647	199	188	387
Ch09_Lc067648	24	450	474
Ch09_Lc067649	340	0	340
Ch09_Lc067650	22	348	370
Ch09_Lc067651	20	365	385
Ch10_Lc067652	0	292	292
Ch10_Lc067653	0	279	279
Ch10_Lc067654	0	396	396
Ch10_Lc067655	160	170	330
Ch10_Lc067656	175	121	296
Ch10_Lc067657	212	188	400
Ch10_Lc067658	0	356	356
Ch10_Lc067659	400	13	413
Ch10_Lc067661	263	268	531
Ch10_Lc067662	20	324	344
Ch11_Lc067663	12	357	369
Ch11_Lc067664	142	179	321
Ch11_Lc067665	278	0	278
Ch11_Lc067666	180	232	412
Ch11_Lc067667	13	438	451
Ch11_Lc067668	0	365	365
Ch11_Lc067669	15	263	278
Ch11_Lc067670	0	374	374
Ch12_Lc067671	233	183	416
Ch12_Lc067672	0	269	269
Ch12_Lc067673	412	11	423
Ch12_Lc067676	37	436	473
Ch12_Lc067677	168	197	365

Los datos para estos polimorfismos se obtuvieron en el mismo conjunto de datos que los datos de aneuploidía ilustrados en las FIGs. 11 y 12. Por tanto, un único ensayo demostró la capacidad de identificar una aneuploidía fetal, y las diferencias polimórficas entre los locus fetal y materno permitieron la identificación de locus informativos y el cálculo del porcentaje estimado de ADNlc fetal en la muestra en base a los locus informativos.

LISTA DE SECUENCIAS

<110> Ariosa Diagnostics, Inc.

5 <120> SISTEMAS DE ENSAYO PARA LA DETERMINACIÓN DE LA CONTRIBUCIÓN DE UNA FUENTE EN UNA
MUESTRA

<130> EAH/FP7347180

10 <140> EP
<141> 08/08/2011

<150> EP11745883.6
<151> 08/08/2011

15 <150> PCT/US2011/046981
<151> 08/08/2011

<150> US 13/013.732
<151> 25/01/2011

20 <150> US 61/371605
<151> 06/08/2010

25 <160> 8
<170> PatentIn versión 3.5

<210> 1
30 <211> 24
<212> ADN
<213> SECUENCIA ARTIFICIAL

<220>
35 <223> Oligonucleótido sintético

<400> 1
24
 tacaccggcg ttatgcgtcg agac

40 <210> 2
<211> 24
<212> ADN
<213> ORGANISMO ARTIFICIAL

45 <220>
<223> Oligonucleótido sintético
24
 <400> 2
 attgcgggga ccgatgatcg cgtc

50 <210> 3
<211> 48
<212> ADN
<213> SECUENCIA ARTIFICIAL

55 <220>
<223> Oligonucleótido sintético
48
 <400> 3
 taatgatacg gcgaccacccg agatctacac cggcgttatg cgtcgaga

<210> 4
<211> 53
<212> ADN
65 <213> ORGANISMO ARTIFICIAL

<220>
 <223> Oligonucleótido sintético

5 <220>
 <221> Índice
 <222> (26)..(26)
 <223> n es un índice de secuencia.

10 <220>
 <221> misc_feature
 <222> (26)..(26)
 <223> n es a, c, g o t

15 <400> 4
tcaaggcagaa gacggcatac gagatnnaaac gacgcgatca tcggtccccg caa 53

<210> 5
 <211> 21

20 <212> ADN
 <213> SECUENCIA ARTIFICIAL

<220>
 <223> Oligonucleótido sintético

25 <400> 5
taatgatacgtgcgaccaccg a 21

<210> 6
 <211> 25

30 <212> ADN
 <213> SECUENCIA ARTIFICIAL

<220>
 <223> **ATCTCGTATGCCGTCTTCTGCTTGA**

<400> 6
atctcgatgcgcgtttctg cttga 25

40 <210> 7
 <211> 48
 <212> ADN
 <213> SECUENCIA ARTIFICIAL

45 <220>
 <223> Oligonucleótido sintético

<400> 7
aatgatacgg cgaccaccga gatctacacc ggcttatgc gtcgagac 48

50 <210> 8
 <211> 60
 <212> ADN
 <213> SECUENCIA ARTIFICIAL

55 <220>
 <223> Oligonucleótido sintético

60 <220>
 <221> n
 <222> (26)..(32)
 <223> n = G, A, T o C

<220>

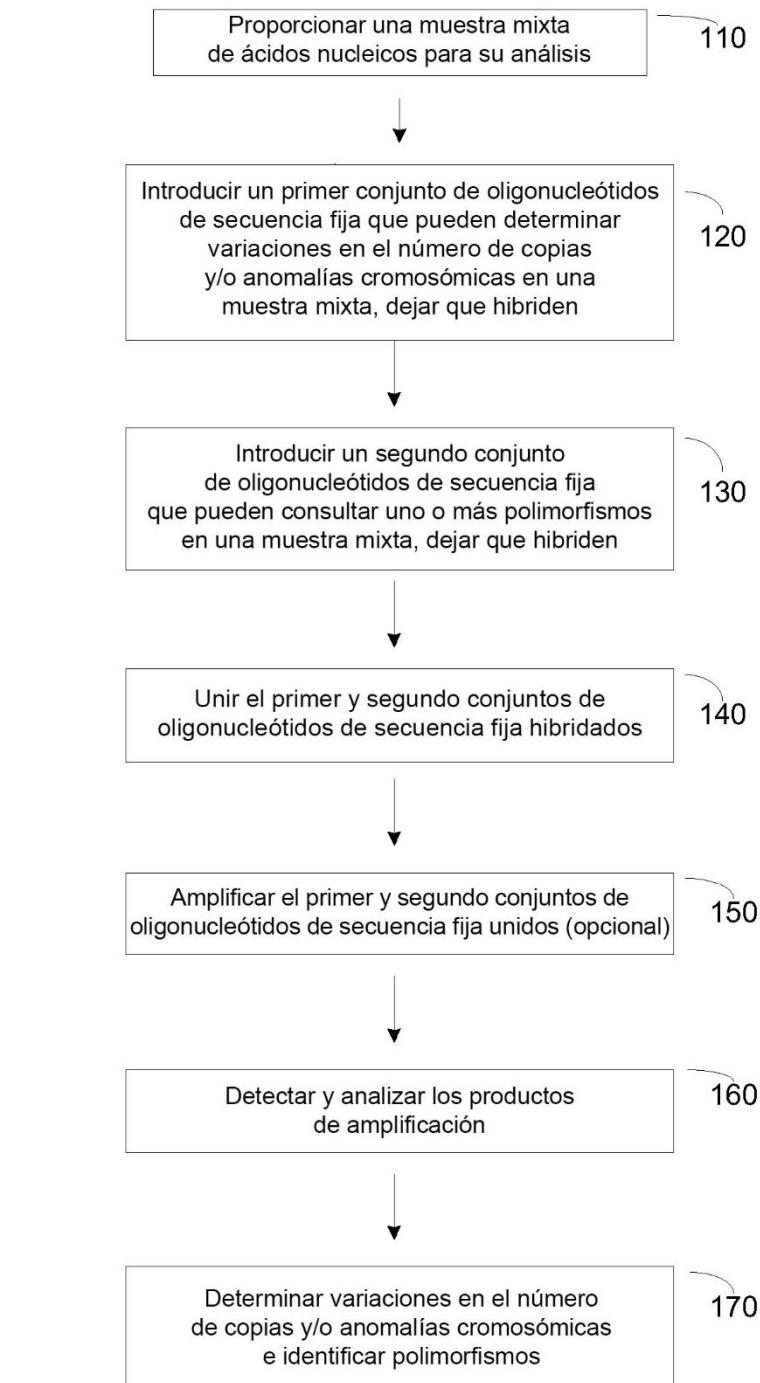
ES 2 863 778 T3

<221> misc_feature
<222> (26)..(32)
<223> n es a, c, g o t

5 <400> 8
tcaaggcagaa gacggcatac gagatnnnnn nnnaacgacg cgatcatcgg tccccgcaat 60

REIVINDICACIONES

10. Procedimiento de ensayo de la reivindicación 8, en el que los productos de amplificación individuales aislados se amplifican adicionalmente para crear copias idénticas de moléculas complementarias a todos o una parte de los productos de amplificación individuales antes de la detección.
5
11. Procedimiento de ensayo de una cualquiera de las reivindicaciones precedentes, en el que al menos un oligonucleótido de secuencia fija de cada uno de los conjuntos primero y segundo de oligonucleótidos de secuencia fija comprende un índice de muestra.
- 10 12. Procedimiento de ensayo de una cualquiera de las reivindicaciones precedentes, en el que la muestra mixta es una muestra materna que comprende ADNlc materno y fetal.
13. Procedimiento de ensayo de una cualquiera de las reivindicaciones precedentes, en el que
15 (a) la región genómica está localizada en un cromosoma seleccionado del grupo que consiste en los cromosomas 13, 18 y 21; o
(b) la región genómica está localizada en un cromosoma Y o X.
- 20 14. Procedimiento de ensayo de una cualquiera de las reivindicaciones precedentes, en el que
25 (a) la muestra mixta es una muestra materna que comprende ADNlc materno y fetal o el ADNlc de la muestra materna y el sistema de ensayo identifica la presencia o ausencia de anomalías cromosómicas en la muestra, en el que las anomalías cromosómicas se seleccionan de síndrome de Down (trisomía 21), síndrome de Edwards (trisomía 18), síndrome de Patau (trisomía 13), síndrome de Klinefelter (XXY), síndrome de triple X, síndrome de XYY, trisomía 8, trisomía 16, síndrome de Turner, síndrome robertsoniano, síndrome de DiGeorge y síndrome de Wolf-Hirschhorn; y/o
30 (b) en el que la muestra mixta es una muestra materna que comprende ADNlc materno y fetal o el ADNlc de la muestra materna y el sistema de ensayo identifica el porcentaje de contribución fetal en la muestra, y en el que el feto es de sexo masculino y donde el porcentaje de ADN fetal en la muestra se determina a través de la detección de locus específicos del cromosoma Y; y/o
35 (c) en el que se determinan opcionalmente cantidades de un locus específico del cromosoma Y amplificado a partir de la muestra y se comparan con uno o más locus seleccionados amplificados que están presentes tanto en el ADN materno como en el ADN fetal y que provienen de una región autosómica que no se encuentra en el cromosoma 13, 18, o 21.
- 40 15. Procedimiento de ensayo de una cualquiera de las reivindicaciones precedentes, en el que el al menos un locus polimórfico está localizado en un cromosoma seleccionado del grupo que consiste en los cromosomas 1-16.



100

FIGURA 1

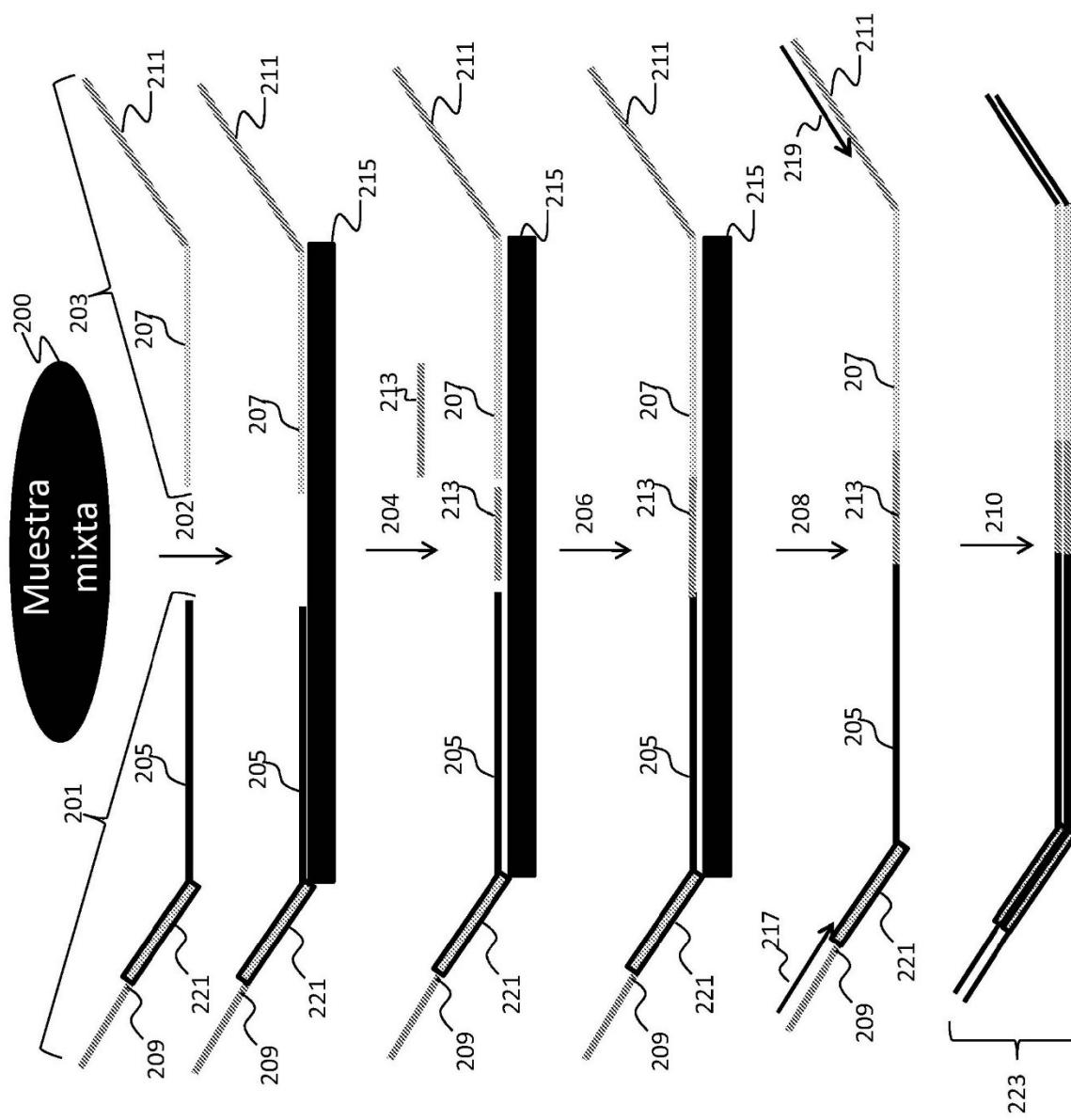


FIG. 2

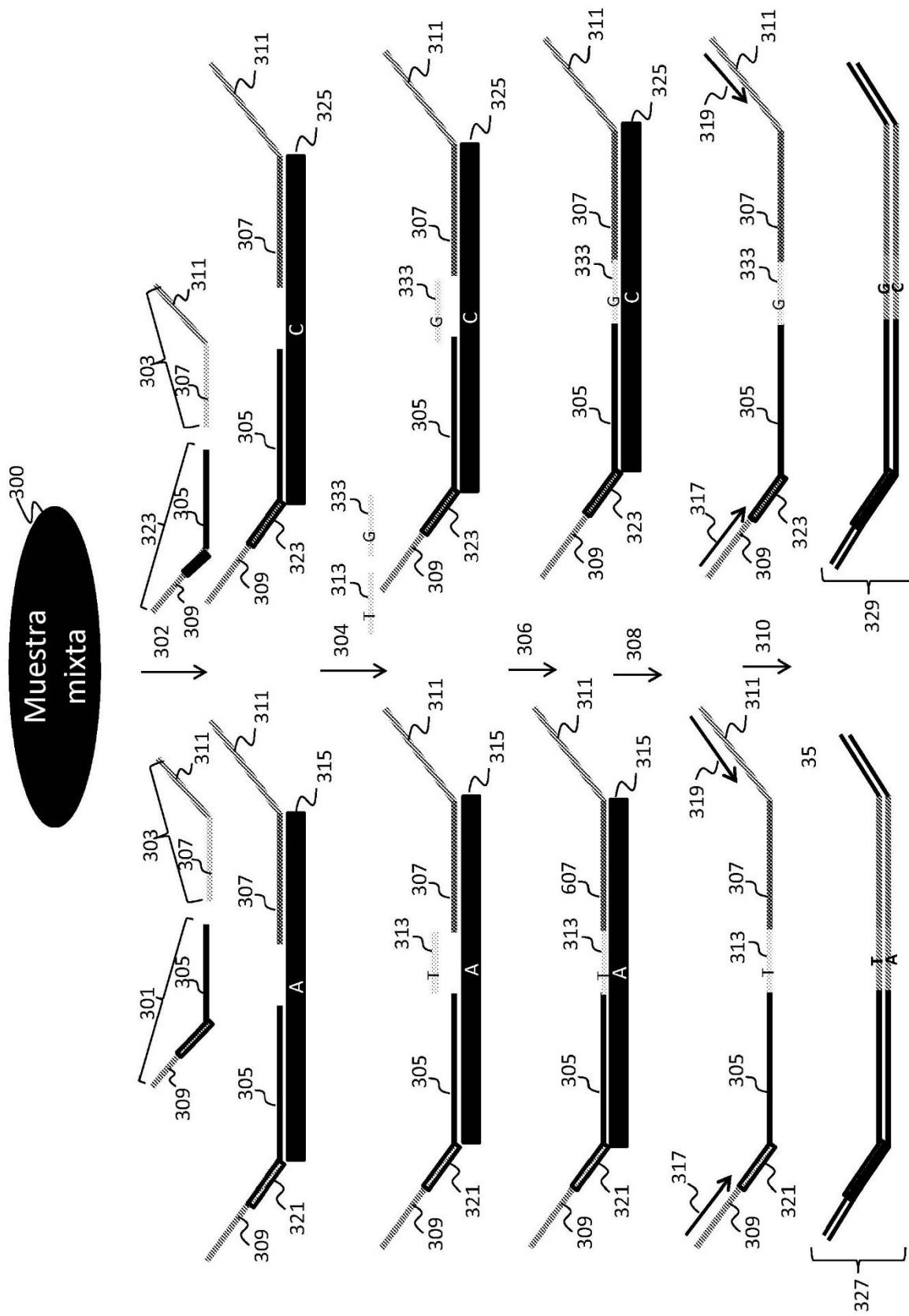


FIG. 3

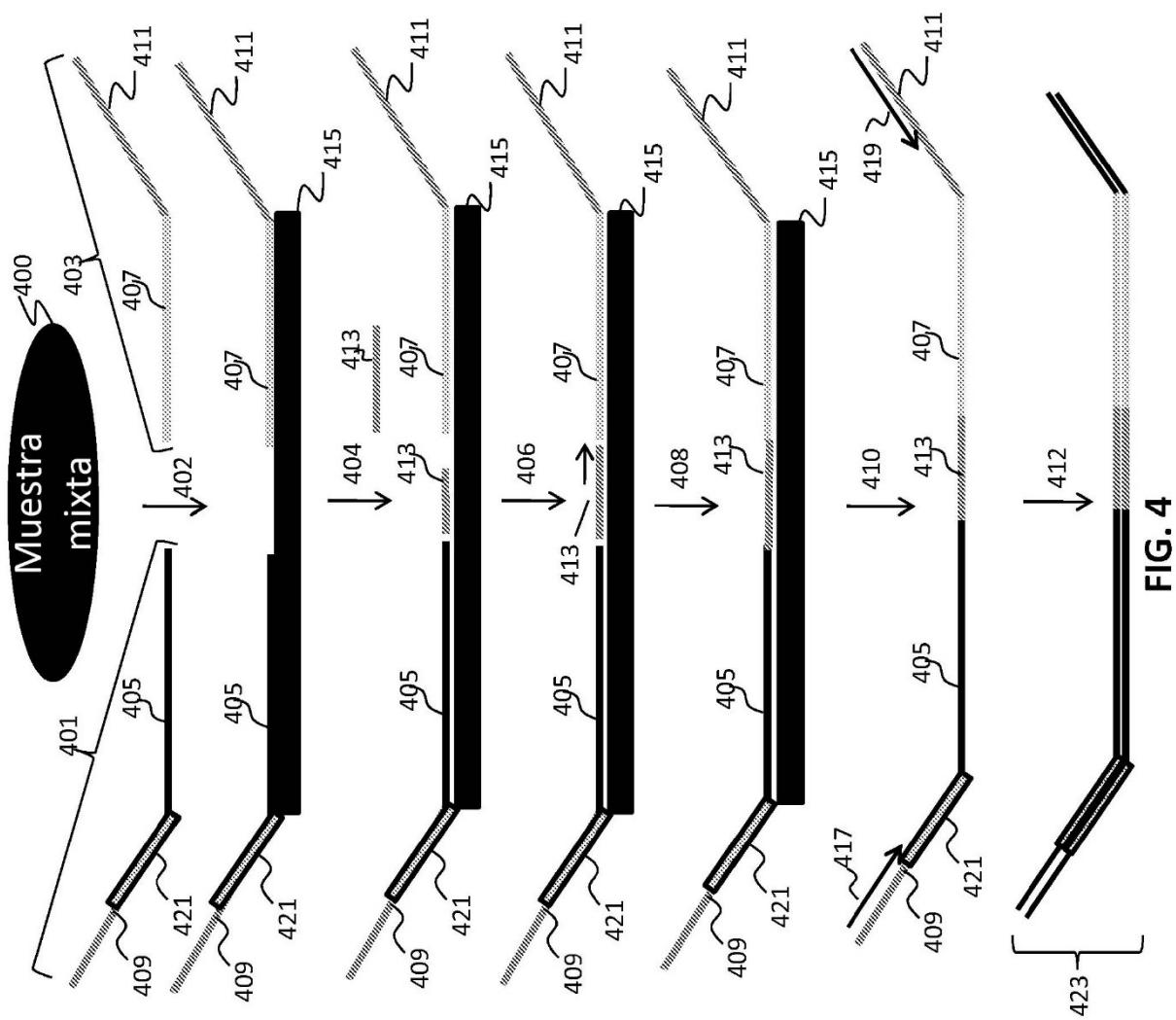


FIG. 4

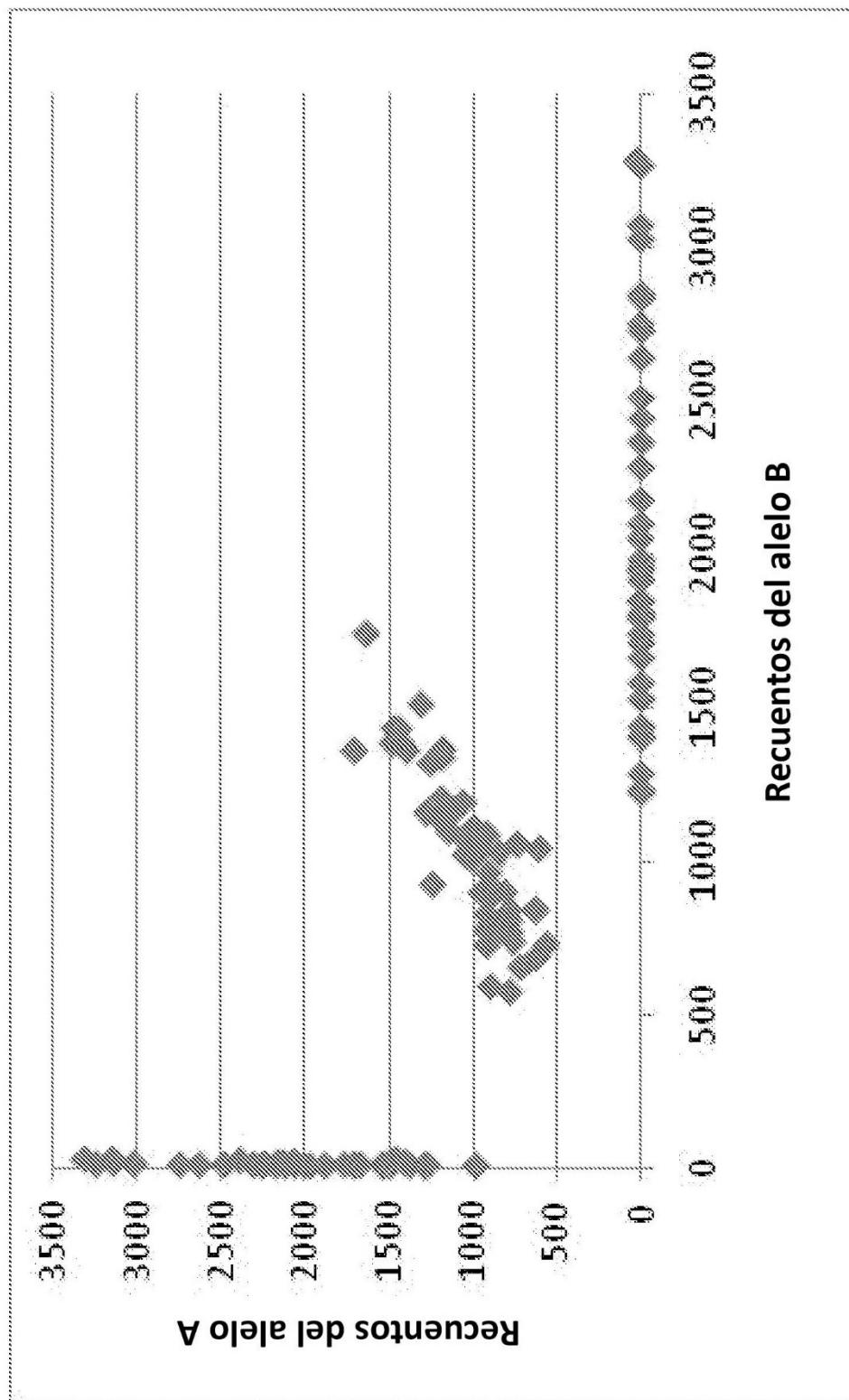
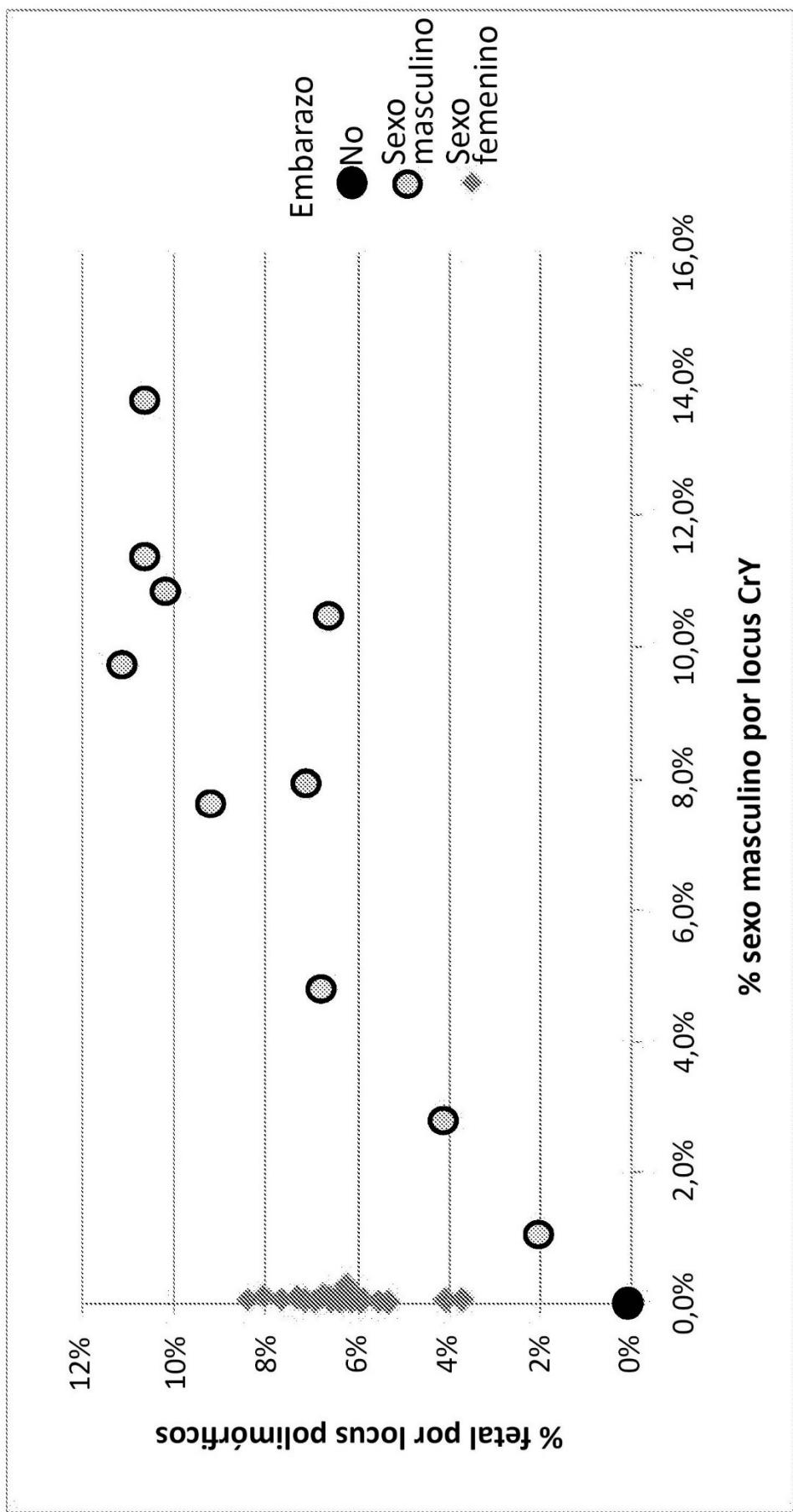


FIG. 5



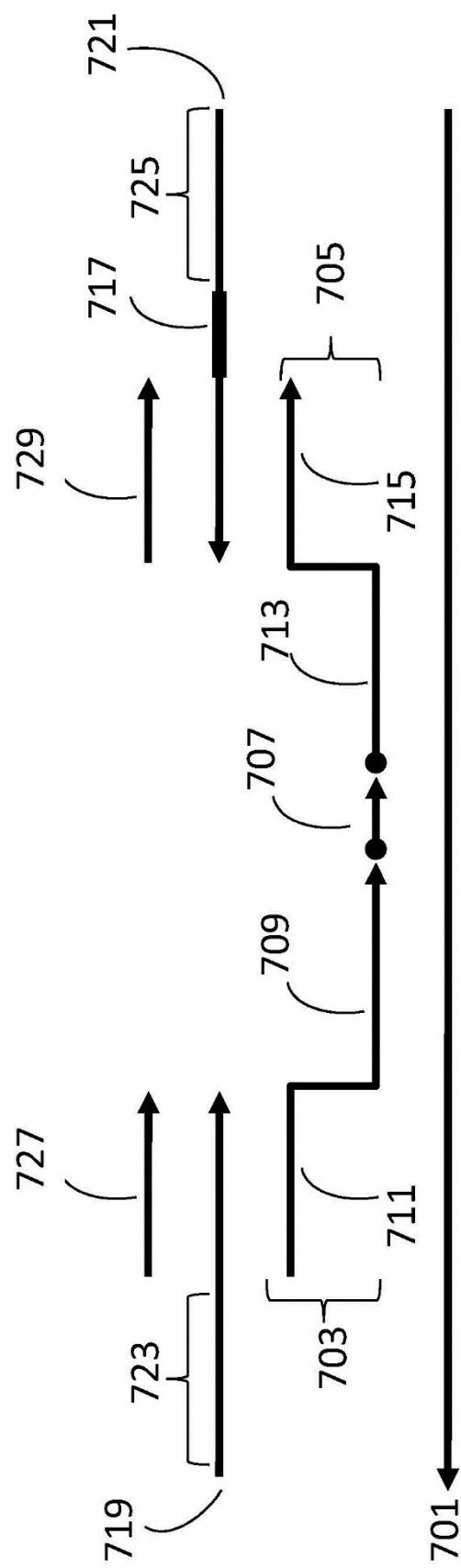


FIG. 7

ES 2 863 778 T3

Muestra	Marca de muestra	Sexo fetal	Equiv. genoma	Cariotipo materno	Cariotipo fetal	Edad madre	Edad Gest (día)	Nº Locus DiffPol	% Fetal	Media18	Media21	Prop18	Prop21	EstZ18	EstZ21
1	TGTCCCC	M	2293	46XX	46XY	45	110	38	0,210138	1005,426	999,4412	0,501493	0,498507	0,48436	-0,48436
2	ATCCTCC	F	3593	46XX	46XX	37	131	45	0,096075	996,6228	998,7823	0,499459	0,500541	-0,84491	0,84491
4	GCCAAACC	F	1469	46XX	46XX	34	121	26	0,235051	991,0624	997,5193	0,498376	0,501624	-1,552358	1,552358
5	CCACGAC	F	1296	46XX	46XX	36	129	36	0,140539	1010,048	1004,468	0,501385	0,498615	0,413907	-0,41391
6	TCCTCGT	F	2110	46XX	46XX	33	136	24	0,049392	994,7681	1000,632	0,498531	0,501469	-1,45169	1,451689
7	CCTTGCT	F	2772	46XX	46XX	43	111	42	0,095395	997,3443	986,4745	0,50274	0,49726	1,299353	-1,299353
8	TCACTCT	M	2506	46XX	46XY	34	126	32	0,158725	1004,807	996,2886	0,502128	0,497872	0,899842	-0,899842
9	CATCTTC	M	3996	46XX	46XY	36	121	34	0,110817	1005,714	1002,283	0,500854	0,499146	0,067201	-0,067201
10	TTCGGCG	F	4032	46XX	46XX	42	84	29	0,159776	999,7097	1000,399	0,499828	0,500172	-0,60393	0,60393
11	GTGTCCC	F	4212	46XX	46XX	36	175	39	0,119147	1004,796	995,1465	0,502412	0,497588	1,083457	-1,08546
12	GCCCCTG	F	3816	46XX	46XX	36	111	27	0,085664	1002,909	999,3516	0,500888	0,499112	0,089291	-0,08929
13	CTTACAC	F	2160	46XX	46XX	41	123	40	0,111731	1002,012	995,0877	0,501734	0,498266	0,641893	-0,641893
14	TCGCAGC	F	2790	46XX	46XX	51	125	36	0,119314	1008,515	999,3439	0,502284	0,497716	1,001374	-1,001374
15	CGGCCAT	F	2783	46XX	46XY	51	125	50	0,123598	1005,525	998,867	0,501661	0,498339	0,594272	-0,594272
16	CCAGCTG	M	2272	46XX	46XY	34	110	40	0,1119257	1010,173	1004,256	0,501468	0,498532	0,468561	-0,46856
17	AGCACCT	M	2401	46XX	46XY	34	140	43	0,111573	988,7338	986,5967	0,500541	0,499459	-0,3768	0,37679
18	CGATACC	F	3960	46XX	46XX	41	126	27	0,099872	996,8338	999,2182	0,499403	0,500597	-0,88162	0,88162
19	CCCGAAT	M	3816	46XX	46XY	37	133	30	0,095735	999,3982	991,0248	0,502103	0,497897	0,883556	-0,883556
20	CCCAAGGG	M	3474	46XX	46XY	38	123	32	0,115184	1006,395	992,8395	0,50339	0,49661	1,724514	-1,724514
21	CAGCACG	M	1080	46XX	46XY	38	97	28	0,13036	994,7498	1003,364	0,497844	0,502156	-1,90014	1,900139

ES 2 863 778 T3

Muestra	Marca de muestra	Sexo fetal	Equiv. genoma	Cariotipo materno	Cariotipo fetal	Edad madre (día)	Edad Gest	Locus informativos	% Fetal	Media18	Media21	Prop18	Prop21	EstZ18	EstZ21
22	CTCTATG	M	2002	46XX	46XY	30	123	30	0,08322	994,7936	999,994	0,498697	0,501303	-1,3432	1,343199
23	ACTTAC	M	3042	46XX	46XY	35	130	33	0,138751	1004,97	1003,101	0,500465	0,499535	-0,18713	0,187128
24	CGTATCG	F	1922	46XX	46XX	23	159	29	0,254234	1003,218	1006,4	0,499208	0,500792	-1,00868	1,008678
25	AACCAIT	M	3708	46XX	46XY	38	148	31	0,145101	996,0673	995,7268	0,500085	0,499915	-0,435388	0,435383
26	CTGGGGC	M	4032	46XX	46XY	30	126	27	0,077438	1005,191	1003,628	0,500389	0,499611	-0,23694	0,23694
27	TGCCGAG	F	2383	46XX	46XX	35	138	32	0,131712	996,445	1002,698	0,498436	0,501564	-1,51345	1,513451
28	ATACCAG	F	4104	46XX	46XX	23	99	33	0,099894	991,6721	1001,903	0,497434	0,502566	-2,16839	2,168385
29	GAACCCG	F	2437	46XX	46XX	33	149	25	0,067606	1004,187	1006,778	0,499356	0,500644	-0,91231	0,912305
30	AAAGGCC	F	2765	46XX	46XX	23	126	29	0,144098	1001,308	995,2881	0,501508	0,498492	0,494093	-0,49409
31	TCTAAATT	M	2416	46XX	46XY	23	119	28	0,105371	1006,535	1005,069	0,500364	0,499636	-0,25309	0,25309
32	GTGATC	M	2653	46XX	46XY	34	148	29	0,164518	1003,738	998,0583	0,501419	0,498581	0,435957	-0,43596
33	GATGTCT	M	7488	46XX	46XY	36	148	21	0,056835	1012,154	1002,257	0,502456	0,497544	1,114241	-1,11424
34	AGTCTGT	M	2653	46XX	46XY	26	140	39	0,21197	1004,959	996,9646	0,501997	0,498003	0,813799	-0,8138
35	AATCTCTG	F	3780	46XX	46XX	35	133	43	0,259642	997,0006	997,6736	0,499831	0,500169	-0,60149	0,601493
36	CTAAGT	M	8136	46XX	46XY	26	151	33	0,082773	1007,02	1006,742	0,500069	0,499931	-0,44599	0,445992
37	GGGGGIT	M	2228	46XX	46XY	37	114	37	0,155551	1004,912	1003,615	0,500323	0,499677	-0,28021	0,280209
38	TATAGGC	F	4248	46XX	46XX	27	165	48	0,064569	1014,599	995,8518	0,504662	0,495338	2,556064	-2,55606
39	TAGGTAC	M	3960	46XX	46XY	35	124	41	0,167184	997,8251	998,4044	0,499855	0,500145	-0,58608	0,586081
40	CAAATGG	F	4140	46XX	46XX	24	127	43	0,129398	997,6805	996,6003	0,500271	0,499729	-0,31423	0,314232
41	TCATAAG	M	4320	46XX	46XY	39	139	29	0,227198	1002,31	991,6654	0,502669	0,497331	1,253298	-1,2533
42	GAACGGT	F	3744	46XX	46XX	38	118	39	0,094147	1007,442	994,0998	0,503333	0,496667	1,687191	-1,68719

ES 2 863 778 T3

Muestra	Marca de muestra	Sexo fetal	Equiv. genoma materno	Cariotipo fetal	Edad madre (día)	Edad Gest	Locus informativos	% Fetal	Media21	Prop18	Prop21	EstZ18	EstZ21	
43	AAGGGCGT	F	6984	46XX	46XX	37	130	36	0,228387	998,8644	995,701	0,500793	0,499207	-0,027059
44	AGGAATC	F	5112	46XX	46XX	18	188	39	0,244261	1001,465	995,804	0,501417	0,498583	-0,435058
45	AACTATG	F	3503	46XX	46XX	41	141	22	0,106969	999,5094	990,8389	0,502178	0,497822	-0,932375
46	TTTGGAT	F	2606	46XX	46XX	39	143	32	0,196098	1001,959	1000,069	0,500472	0,499528	-0,18279
47	TTGATGT	F	5292	46XX	46XX	36	111	30	0,109084	1004,708	1002,777	0,500481	0,499519	-0,17694
48	ATGGTTG	M	5184	46XX	46XY	24	160	40	0,060517	1010,204	1006,291	0,50097	0,49903	-0,143
49	GGGTAGT	M	3996	46XX	46XY	23	153	30	0,151707	1003,952	998,371	0,501394	0,498606	-0,419553
50	AGATGAT	F	5256	46XX	46XX	26	118	26	0,126345	998,1654	996,1826	0,500497	0,499503	-0,16635
51	GTGAAAG	M	4068	46XX	46XY	33	140	29	0,214532	1004,428	996,9586	0,501866	0,498134	-0,728436
52	AGTGAAG	M	11412	46XX	46XY	18	196	24	0,140258	1007,84	998,7992	0,502253	0,497747	-0,9811
53	ACAAACGC	M	3146	46XX	46XY	21	147	32	0,180869	1010,525	999,8358	0,502659	0,497341	-1,24639
54	GACGAGG	F	5508	46XX	46XX	35	114	32	0,053622	994,8149	1005,452	0,497341	0,502659	-2,22916
55	GGATAAC	F	6624	46XX	46XX	28	140	37	0,21238	999,7187	1006,639	0,498275	0,501725	-1,61846
56	GTTCGCG	M	6192	46XX	46XY	44	115	25	0,083131	1009,487	999,2224	0,502555	0,497445	-1,17876
57	GTCTTAT	F	7056	46XX	46XX	25	115	37	0,135201	1005,611	996,4	0,5023	0,4977	-1,01226
58	TAGGTT	F	5148	46XX	46XX	22	99	38	0,124127	996,363	999,6895	0,499167	0,500833	-1,03587
59	TGCTTIC	M	4104	46XX	46XY	30	197	28	0,310846	1004,509	1009,571	0,498743	0,501257	-1,31256
60	TCTGTGG	M	5364	46XX	46XY	27	122	26	0,118328	997,1746	992,0824	0,50128	0,49872	-0,345307
61	TGGGCTC	F	3636	46XX	46XX	28	178	30	0,158146	989,5388	999,8624	0,497405	0,502595	-2,18709
62	TGAGTAT	M	17856	46XX	46XY	NA	NA	45	0,064088	993,7189	996,5081	0,499299	0,500701	-0,94924
63	TGCAAAT	F	5004	46XX	46XX	35	109	39	0,076783	1003,939	992,7595	0,5028	0,4972	-1,33858

ES 2 863 778 T3

ES 2 863 778 T3

89	AGCTGCC	F	6768	46XX	46XX+21	31	98	35	0,126696	963,2256	1024,818	0,484509	0,515491	-10,616	10,61595			
Muestra	Marca de muestra	Sexo fetal	Equiv. genoma	Cariotipo materno	Cariotipo fetal	Edad madre	Edad Gest (día)	Locus informativos	% Fetal	Media18	Media21	Prop18	Prop21	EstZ18	EstZ21			
90	AGCGCTG	M	1375	46XX	46XY+21	36	162	36	0,133202	961,1764	1032,034	0,482225	0,517775	-12,1088	12,10879			
91	AGCATGCC	M	5155	46XX	46XY+21	41	154	37	0,095349	967,2823	1017,613	0,487322	0,512678	-8,77777	8,777767			
92	AATGGGT	F	2549	46XX	46XXX+21	41	119	32	0,09505	965,3,134	1025,717	0,484831	0,515169	-10,4056	10,4056			
93	AAAGTAAG	F	9216	46XX	46XX+21	42	253	32	0,209915	946,0902	1052,948	0,473273	0,526727	-17,96	17,96002			
94	AAATAGC	F	2041	46XX	46XXX+21	36	161	32	0,102554	973,7957	1025,545	0,487058	0,512942	-8,94982	8,949816			
95	AAATCCT	M	1843	46XX	46XY+21	42	91	35	0,137238	965,4467	1031,552	0,483449	0,516551	-11,309	11,30904			

FIG. 8 (cont.)

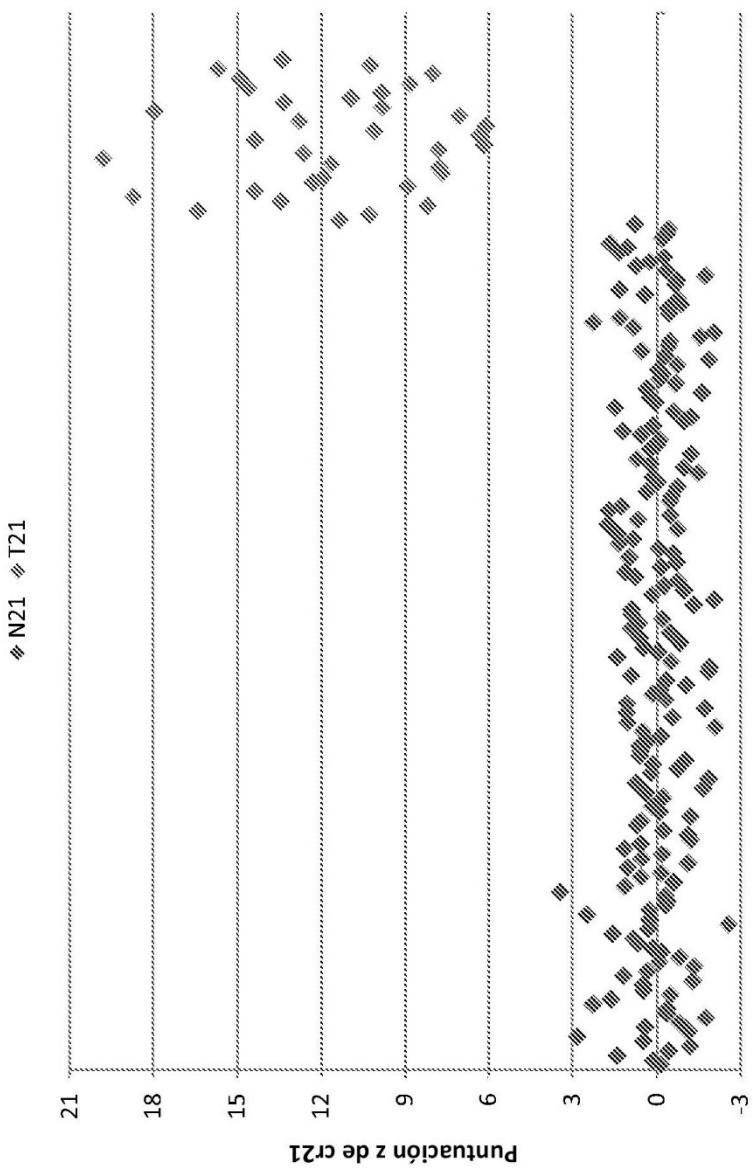


FIG. 9

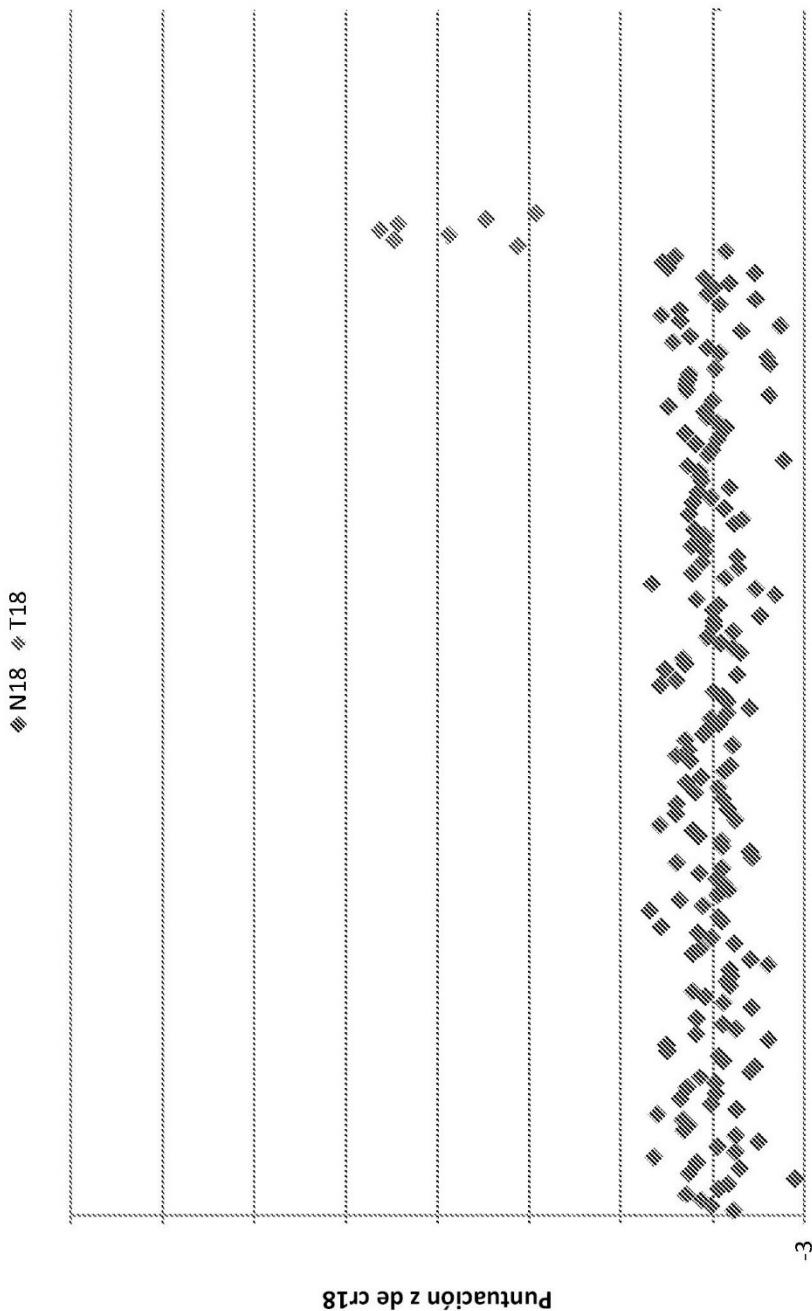


FIG. 10

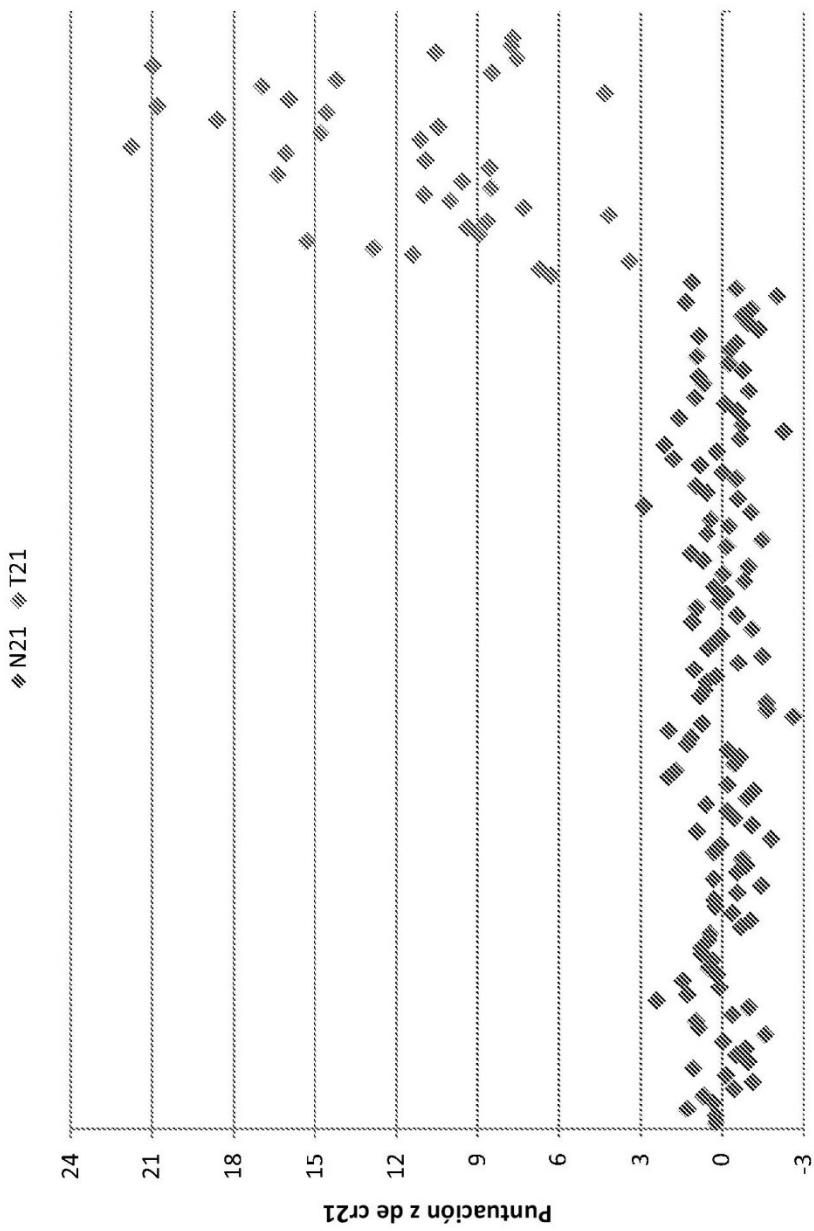


FIG. 11

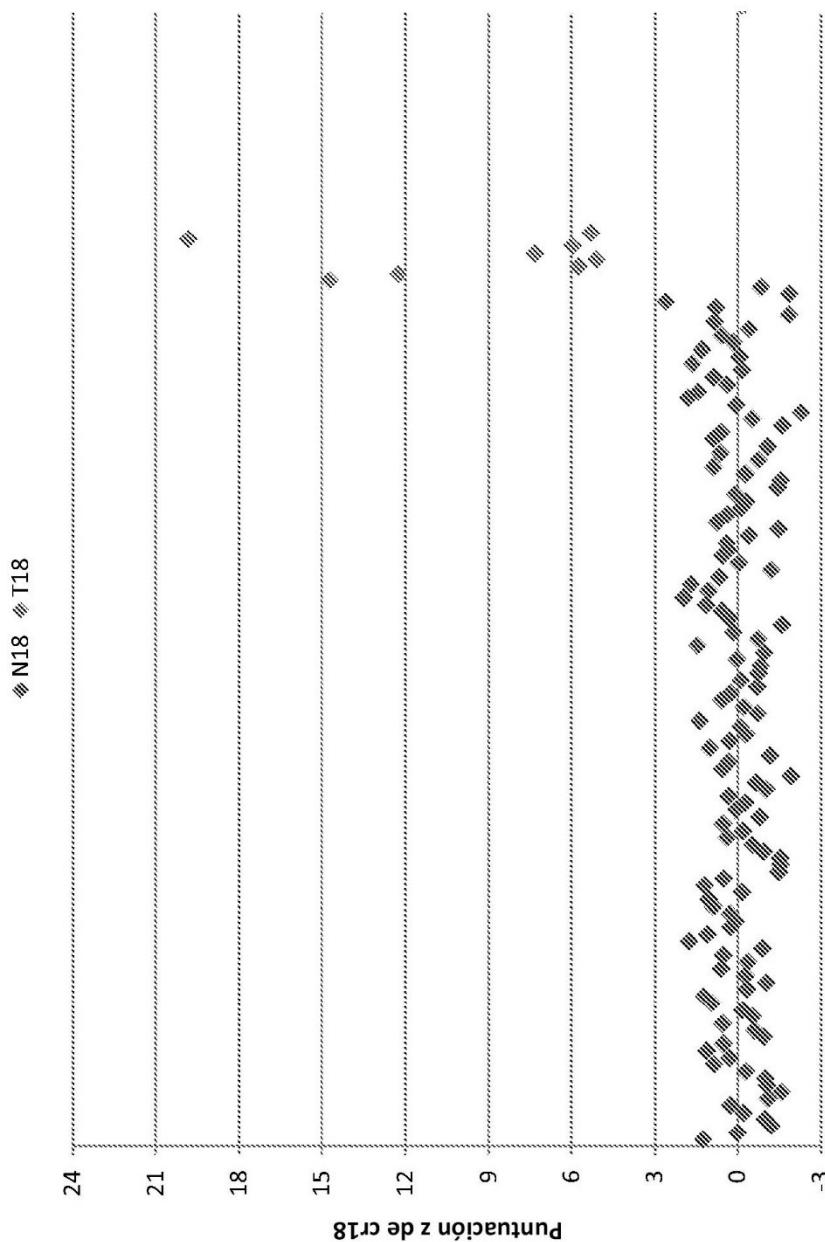


FIG. 12