

(19)



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS  
ESPAÑA



(11) Número de publicación: **2 913 402**

(51) Int. Cl.:

**C12Q 1/6827** (2008.01)

**C12Q 1/6862** (2008.01)

(12)

## TRADUCCIÓN DE PATENTE EUROPEA

T3

(96) Fecha de presentación y número de la solicitud europea: **08.08.2011 E 18167314 (6)**

(97) Fecha y número de publicación de la concesión europea: **16.03.2022 EP 3418394**

---

(54) Título: **Sistemas de ensayo para análisis genético**

(30) Prioridad:

**06.08.2010 US 37160510 P  
25.01.2011 US 201113013732**

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:  
**02.06.2022**

(73) Titular/es:

**F. HOFFMANN-LA ROCHE AG (100.0%)  
Grenzacherstrasse 124  
4070 Basel, CH**

(72) Inventor/es:

**SPARKS, ANDREW;  
OLIPHANT, ARNOLD;  
ZAHN, JACOB;  
SONG, KEN y  
STUELPNAGEL, JOHN**

(74) Agente/Representante:

**LINAGE GONZÁLEZ, Rafael**

**ES 2 913 402 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**

Sistemas de ensayo para análisis genético

**5 Campo de la invención**

La presente invención se refiere a la detección y cuantificación de la variación genética en muestras mixtas en un ensayo único.

**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 de demostrar, cuando sea apropiado, que los artículos y procedimientos a los que se hace referencia en el presente documento no constituyen una técnica anterior conforme a las disposiciones legales aplicables.

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

Aunque la tecnología convencional proporciona procedimientos de detección para estas diferentes anomalías genéticas, actualmente requiere diferentes técnicas para consultar diferentes clases de mutaciones. Un ensayo que proporciona la detección de una variación en el número de copias con detección simultánea de polimorfismos de genes individuales sería una herramienta poderosa en el tratamiento. La presente invención aborda esta necesidad.

Por tanto existe la necesidad de procedimientos no invasivos de cribado para anomalías genéticas, incluyendo variaciones en el número de copias, en muestras mixtas que comprenden ADN normal y anómalo supuesto. La presente invención aborda esta necesidad.

**35 Sumario de la invención**

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 este 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 escrita, incluyendo los aspectos ilustrados en los dibujos adjuntos y definidos en las reivindicaciones adjuntas.

La presente invención proporciona un procedimiento de ensayo único para determinar la variación genética en muestras mixtas que comprenden material genómico (por ejemplo, ADN libre circulante) de una fuente principal y una o más fuentes secundarias de un sujeto. Más en particular, la presente invención proporciona procedimientos para la detección de la variación en el número de copias de locus en una muestra que comprende ácidos nucleicos de una fuente principal y una fuente secundaria dentro de un individuo. La presente invención utiliza un sistema de ensayo único que permite la identificación de variaciones en el número de copias (VNC) genéticas y la capacidad de distinguir entre estas VNC en las fuentes principal y secundaria.

En un aspecto, la divulgación proporciona sistemas de ensayo único, que no se reivindican, con la capacidad para determinar 1) la presencia o ausencia de VNC en uno o más locus y; 2) la presencia o ausencia de polimorfismos en uno o más locus en una muestra mixta de un individuo. El sistema de ensayo puede detectar específicamente el número de copias de locus seleccionados y polimorfismos presentes en locus seleccionados de una fuente secundaria dentro de la muestra mixta, y distinguir esto del número de copias y polimorfismos presentes en locus seleccionados de una fuente principal en la muestra mixta. Preferentemente, los locus se analizan a través del uso de ácidos nucleicos libres circulantes, y preferentemente los ácidos nucleicos libres circulantes analizados en el sistema de ensayo son ADN libre circulante (ADNlc).

Por tanto, en un aspecto específico, la divulgación proporciona un sistema de ensayo, que no se reivindica, para la detección de la presencia o ausencia de una variación en el número de copias (VNC) de una región genómica y la presencia o ausencia de uno o más polimorfismos en una muestra mixta usando un ensayo único, comprendiendo el ensayo las etapas de introducir un primer conjunto de oligonucleótidos de secuencia fija en una muestra mixta en condiciones que permiten que los oligonucleótidos de secuencia fija se hibriden específicamente a regiones complementarias en uno o más locus en o asociados con una región genómica; introducir un segundo

conjunto de oligonucleótidos de secuencia fija que comprende regiones de cebador universal en la muestra mixta en condiciones que permiten que los oligonucleótidos de secuencia fija se hibriden específicamente a regiones complementarias en uno o más locus con un polimorfismo supuesto; unir los oligonucleótidos hibridados para crear productos de unión contiguos complementarios a los locus correspondientes a una región genómica y/o los locus con un polimorfismo supuesto; 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 correlaciona con el número de copias de una o más regiones genómicas y la presencia o ausencia de un polimorfismo en uno o más locus en la muestra mixta.

10 En determinados aspectos, los oligonucleótidos de secuencia fija se hibridan a regiones adyacentes en los locus, y los oligonucleótidos de secuencia fija se unen directamente durante la etapa de unión del ensayo. En otros aspectos, los oligonucleótidos de secuencia fija se hibridan a regiones no adyacentes en los locus que están separados por uno, unos pocos o varios nucleótidos, y la región entre los oligonucleótidos de secuencia fija se usa como molde para la extensión del cebador para cerrar el hueco entre los oligonucleótidos de secuencia fija antes de la etapa de unión.

Por tanto, en un segundo aspecto específico, la divulgación proporciona un sistema de ensayo, que no se reivindica, para la detección de la presencia o ausencia de una VNC de una región genómica y la presencia o ausencia de uno o más polimorfismos en una muestra mixta usando un ensayo único, comprendiendo el ensayo las etapas de: introducir un primer conjunto de oligonucleótidos de secuencia fija en una muestra mixta en condiciones que permiten que los oligonucleótidos de secuencia fija se hibriden específicamente a regiones complementarias en uno o más locus en o asociados con una región genómica; introducir un segundo conjunto de oligonucleótidos de secuencia fija en la muestra mixta en condiciones que permiten que los oligonucleótidos de secuencia fija se hibriden específicamente a regiones complementarias en uno o más locus con un polimorfismo supuesto; extender la región entre el primer y segundo oligonucleótido hibridado de al menos un conjunto de oligonucleótidos de secuencia fija con una polimerasa y dNTP para crear dos oligonucleótidos de secuencia fija hibridados de forma adyacente de ese conjunto; unir los oligonucleótidos hibridados para crear productos de unión contiguos complementarios a los locus correspondientes a una región genómica y/o los locus con un polimorfismo supuesto; amplificar los productos de unión contiguos usando las regiones de cebador universal para crear productos de amplificación; y detectar los productos de amplificación. La detección de los productos de amplificación se correlaciona con el número de copias de una o más regiones genómicas y la presencia o ausencia de un polimorfismo en uno o más locus en la muestra mixta.

35 En un aspecto preferente, los oligonucleótidos de secuencia fija usados para consultar los locus específicos comprenden regiones de cebador universal que son comunes a sustancialmente todos los oligonucleótidos de secuencia fija usados en un ensayo particular. Esto permite sustancialmente que todos los productos de unión producidos en un ensayo único se amplifiquen, aíslan y/o analicen usando mecanismos similares (por ejemplo, determinación de secuencia o hibridación). El uso de dichas regiones de cebador universal obvia la necesidad de restos detectables individualmente distinguibles asociados con locus o alelos particulares, y proporciona mecanismos más eficaces y rentables para multiplexar la consulta y/o análisis de múltiples locus de una o múltiples muestras. En un aspecto preferente, las regiones de cebador universal se usan en la determinación de secuencia 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.

45 Cada conjunto de ácidos nucleicos de secuencia fija se diseña para hibridarse a al menos dos regiones separadas en un locus seleccionado. En aspectos preferentes, se usan dos o más oligonucleótidos separados en un conjunto para hibridarse a 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.

55 La región genómica puede ser un único locus, o puede ser una región más grande, hasta e incluyendo un cromosoma. Para la determinación de regiones genómicas más grandes, se pueden usar conjuntos de locus para determinar el tamaño y, en determinados aspectos, los límites de dichas regiones genómicas para las que se ha detectado la VNC.

60 En determinados aspectos, los productos de amplificación del ensayo se aíslan antes de la detección. Preferentemente, los productos de amplificación se aíslan como moléculas individuales antes de la detección. Estos productos de amplificación aislados opcionalmente se pueden amplificar además para crear copias idénticas de todos o una porción de los productos de amplificación individuales antes de la detección, o amplificar además para crear copias idénticas de moléculas complementarias a todos o una porción de los productos de amplificación individuales antes de la detección.

65 Los ensayos multiplexados de la invención permiten el análisis de 24 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 simultáneamente. Estos locus seleccionados pueden ser locus diferentes de una única muestra, o pueden ser locus de dos o más individuos. En el último caso, al menos uno de los dos oligonucleótidos de secuencia fija usados para el análisis de un locus seleccionado puede comprender 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 comprende el índice de muestra.
- 5            10            15            20            25            30            35            40            45            50            55            60            65
- Preferentemente, al menos un locus consultado respecto a VNC en una muestra mixta es diferente de todos los locus consultados respecto a polimorfismos en la muestra mixta. En aspectos específicos, varios locus consultados respecto a VNC en una muestra mixta son diferentes de los locus consultados respecto a polimorfismos en la muestra mixta. En aspectos más específicos, la mayoría de los locus consultados respecto a VNC en una muestra mixta son diferentes de los locus consultados respecto a polimorfismos en la muestra mixta.
- En algunos aspectos de la invención, los oligonucleótidos de secuencia fija se hibridan a regiones adyacentes en un locus, y la unión de estos oligonucleótidos da como resultado un producto de unión que une los dos oligonucleótidos de secuencia fija. En otros aspectos, los oligonucleótidos de secuencia fija se hibridan a regiones no adyacentes en un locus, y las regiones entre los oligonucleótidos de secuencia fija se extienden usando una polimerasa y dNTP. En aspectos principales, los oligonucleótidos de secuencia fija se hibridan a regiones no adyacentes en un locus, y uno o más oligonucleótidos puente se hibridan a la región entre y adyacente al conjunto de oligonucleótidos de secuencia fija. En determinados aspectos preferentes, los oligonucleótidos puente usados pueden proporcionar información de contenido sobre polimorfismos en los locus así como información sobre la frecuencia de locus para el análisis de VNC.
- En otros 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 la detección tanto de VNC como de polimorfismos usando el sistema de ensayo de la divulgación, que no se reivindica, se pueden amplificar usando procedimientos de amplificación universal después de la amplificación selectiva inicial de la muestra mixta. El uso de amplificación universal permite amplificar múltiples regiones de ácidos nucleicos de una única o múltiples muestras usando un número único o limitado de cebadores de amplificación, y es especialmente útil en la amplificación de múltiples regiones seleccionadas en una única reacción.
- Por tanto, en un aspecto de la invención, las secuencias complementarias a 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 a partir del procedimiento de amplificación universal.
- 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 en los productos de amplificación y a continuación se pueden combinar productos de amplificación de diferentes muestras. Preferentemente, el índice de muestra se detecta simultáneamente con la detección de VNC o anomalía cromosómica y la detección de polimorfismo, de modo que la VNC y el polimorfismo se pueden asignar apropiadamente a la muestra de origen.
- Las frecuencias de locus seleccionados se pueden determinar para una región genómica de interés y comparar con las frecuencias de locus de una o más de otras regiones genómicas de interés y/o una o más regiones genómicas de referencia para detectar VNC potenciales en base a frecuencias de locus en la muestra mixta.
- En los sistemas de ensayo de la divulgación, que no se reivindica, los productos de amplificación se aíslan opcionalmente antes de la detección. Cuando se aíslan, preferentemente se aíslan como moléculas individuales para ayudar en la detección posterior. Después del aislamiento, los productos de amplificación se pueden amplificar además para crear copias idénticas de todos o una porción de los productos de amplificación individuales antes de la detección. De forma alternativa, los productos de amplificación aislados se pueden amplificar además para crear copias idénticas de moléculas complementarias a todos o una porción de los productos de amplificación individuales antes de la detección.
- Se pueden emplear diversos procedimientos de detección de VNC junto con la detección de polimorfismos en los sistemas de ensayo de la divulgación. Sin embargo, solo se reivindica el procedimiento como se define en las reivindicaciones adjuntas. 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 del primer y segundo locus. Preferentemente, el procedimiento de ensayo amplifica dos o más locus seleccionados de diferentes regiones genómicas, aunque los locus se pueden localizar en la misma región genómica general para la confirmación de VNC que surgen de anomalías cromosómicas en 5 lugar de VNC de un único locus.
- En otros aspectos, los oligonucleótidos puente son oligonucleótidos degenerados. En otros aspectos, los oligonucleótidos puente se proporcionan como grupos de oligonucleótidos con secuencias aleatorias que comprenden sustancialmente cada combinación para el tamaño particular del oligonucleótido puente usado con 10 los oligonucleótidos de secuencia fija.
- Aunque los aspectos de la invención que usan oligonucleótidos puente se describen principalmente usando un único oligonucleótido puente, se contempla que se puedan usar múltiples oligonucleótidos puente que se hibridan 15 a regiones complementarias adyacentes entre oligonucleótidos de secuencia fija en los procedimientos descritos.
- En un aspecto preferente, la invención proporciona un sistema de ensayo, que no se reivindica, para la detección 20 de la presencia o ausencia de variación en el número de copias (VNC) de una región genómica y la presencia o ausencia de uno o más polimorfismos en una muestra mixta usando un ensayo único, comprendiendo el ensayo las etapas de: introducir un primer conjunto de oligonucleótidos de secuencia fija que comprende regiones de cebador universal en una muestra mixta en condiciones que permiten que los oligonucleótidos de secuencia fija se hibriden específicamente a regiones complementarias en uno o más locus en o asociados con una región genómica; introducir un segundo conjunto de oligonucleótidos de secuencia fija que comprende regiones de cebador universal en la muestra mixta en condiciones que permiten que los oligonucleótidos de secuencia fija se hibriden específicamente a regiones no adyacentes complementarias en uno o más locus con un polimorfismo supuesto; introducir uno o más oligonucleótidos puente en condiciones que permiten que los oligonucleótidos puente se hibriden específicamente a regiones en los locus entre las regiones complementarias a los oligonucleótidos de secuencia fija de los conjuntos; unir los oligonucleótidos hibridados para crear productos de unión contiguos complementarios a los locus correspondientes a una región genómica y/o los locus con un polimorfismo supuesto; amplificar los productos de unión contiguos usando las regiones de cebador universal para 25 crear productos de amplificación; y detectar los productos de amplificación. La detección de los productos de amplificación se correlaciona con el número de copias de una o más regiones genómicas y la presencia o ausencia de un polimorfismo en uno o más locus en la muestra mixta.
- En otro aspecto específico, la invención proporciona un sistema de ensayo, que no se reivindica, para la detección 30 de la presencia o ausencia de VNC de una región genómica y la presencia o ausencia de uno o más polimorfismos en una muestra mixta usando un ensayo único, comprendiendo el ensayo las etapas de: introducir un primer conjunto de oligonucleótidos de secuencia fija en una muestra mixta en condiciones que permiten que los oligonucleótidos de secuencia fija se hibriden específicamente a regiones complementarias en uno o más locus en o asociados con una región genómica; introducir un segundo conjunto de oligonucleótidos de secuencia fija en la muestra mixta en condiciones que permiten que los oligonucleótidos de secuencia fija se hibriden específicamente a regiones complementarias en uno o más locus con un polimorfismo supuesto; introducir uno o más oligonucleótidos puente en condiciones que permiten que los oligonucleótidos puente se hibriden específicamente a regiones en los locus entre las regiones complementarias a los oligonucleótidos de secuencia fija del conjunto; extender la región entre al menos un oligonucleótido de secuencia fija y un oligonucleótido puente con una 35 polimerasa y dNTP para crear oligonucleótidos de secuencia fija hibridados de forma adyacente y oligonucleótidos puente; unir los oligonucleótidos hibridados para crear productos de unión contiguos complementarios a los locus correspondientes a una región genómica y/o los locus con un polimorfismo supuesto; amplificar los productos de unión contiguos usando las regiones de cebador universal para crear productos de amplificación; y detectar los 40 productos de amplificación. La detección de los productos de amplificación se correlaciona con el número de copias de una o más regiones genómicas y la presencia o ausencia de un polimorfismo en uno o más locus en la muestra mixta.
- En aspectos preferentes de la invención que usan oligonucleótidos puente, el primer y segundo conjunto de oligonucleótidos de secuencia fija se introducen antes de la introducción de los oligonucleótidos puente. Más 45 preferentemente, los oligonucleótidos de secuencia fija no hibridados se retiran 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 aislan antes de la introducción de los oligonucleótidos puente.
- En un aspecto preferente, el sistema de ensayo proporciona una consulta de locus altamente multiplexada usando 50 uno o más oligonucleótidos puente comunes que son complementarios a regiones en dos o más locus consultados. Por tanto, el número de oligonucleótidos puente usados en el sistema de ensayo multiplexado será 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 se diseñan cada uno para ser compatibles con dos o más locus consultados 55 usando el sistema de ensayo de la divulgación, que no se reivindica.

En estos aspectos, los oligonucleótidos puente usados en el ensayo multiplexado están diseñados preferentemente para tener una  $T_f$  en un intervalo de  $\pm 5$  °C, más preferentemente en un intervalo de  $\pm 2$  °C.

- 5 En determinados aspectos, el sistema de ensayo multiplica la consulta de locus usando uno o más oligonucleótidos puente comunes que son complementarios a regiones en dos o más locus consultados. Por tanto, el número de oligonucleótidos puente usados en el sistema de ensayo multiplexado será 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 se diseñan cada uno para ser compatibles con dos o más locus consultados usando el sistema de ensayo de la divulgación, que no se reivindica.
- 10 10 En determinados aspectos, los oligonucleótidos puente tienen una longitud de entre 2-45 nucleótidos. En un aspecto específico, los oligonucleótidos puente tienen una longitud de entre 3-9 nucleótidos. Aún en otro aspecto específico, los oligonucleótidos tienen una longitud de entre 10-30 nucleótidos.
- 15 15 Los locus consultados respecto a VNC pueden ser indicativos en algunos casos de una duplicación 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.
- 20 20 En dichos aspectos, la divulgación proporciona sistemas de ensayo, que no se reivindican, que comprenden un sistema de ensayo único con la capacidad para detectar dentro de una muestra mixta de un individuo 1) la presencia o ausencia de una anomalía cromosómica asociada con una o más VNC; y 2) la presencia o ausencia de uno o más polimorfismos en uno o más locus seleccionados. La presencia o ausencia de la anomalía cromosómica se detecta preferentemente a través del uso de locus informativos que permiten que el sistema de ensayo distinga entre ácidos nucleicos de una fuente principal y una fuente secundaria.
- 25 30 Por tanto, en un aspecto específico, la divulgación proporciona un sistema de ensayo, que no se reivindica, para la detección de la presencia o ausencia de una anomalía cromosómica y la presencia o ausencia de uno o más polimorfismos en una muestra mixta usando un ensayo único, comprendiendo el ensayo las etapas de: introducir un primer conjunto de oligonucleótidos de secuencia fija que comprenden regiones de cebador universal en una muestra mixta en condiciones que permiten que los oligonucleótidos de secuencia fija se hibriden específicamente a regiones complementarias en uno o más locus con un polimorfismo supuesto; introducir un segundo conjunto de oligonucleótidos de secuencia fija que comprende regiones de cebador universal en la muestra mixta en condiciones que permiten que los oligonucleótidos de secuencia fija se hibriden específicamente a regiones complementarias en dos o más locus de un primer cromosoma; introducir un tercer conjunto de oligonucleótidos de secuencia fija que comprende regiones de cebador universal en la muestra mixta en condiciones que permiten que los oligonucleótidos de secuencia fija se hibriden específicamente a regiones complementarias en dos o más locus de un segundo cromosoma; introducir uno o más oligonucleótidos puente en condiciones que permiten que los oligonucleótidos puente se hibriden específicamente a regiones en los locus entre las regiones complementarias a los oligonucleótidos de secuencia fija del 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 correlaciona con la presencia o ausencia de una anomalía cromosómica y la presencia o ausencia de un polimorfismo en uno o más locus en la muestra mixta.
- 35 45 En otro aspecto específico, la divulgación proporciona un sistema de ensayo, que no se reivindica, para la detección de la presencia o ausencia de una anomalía cromosómica y la presencia o ausencia de uno o más polimorfismos en una muestra mixta usando un ensayo único, comprendiendo el ensayo las etapas de: introducir un primer conjunto de oligonucleótidos de secuencia fija que comprenden regiones de cebador universal en una muestra mixta en condiciones que permiten que los oligonucleótidos de secuencia fija se hibriden específicamente a regiones complementarias en uno o más locus con un polimorfismo supuesto; introducir un segundo conjunto de oligonucleótidos de secuencia fija que comprende regiones de cebador universal en la muestra mixta en condiciones que permiten que los oligonucleótidos de secuencia fija se hibriden específicamente a regiones complementarias en dos o más locus de un primer cromosoma; introducir un tercero conjunto de oligonucleótidos de secuencia fija que comprende regiones de cebador universal en la muestra mixta en condiciones que permiten que los oligonucleótidos de secuencia fija se hibriden específicamente a regiones complementarias en dos o más locus de un segundo cromosoma; introducir uno o más oligonucleótidos puente en condiciones que permiten que los oligonucleótidos puente se hibriden específicamente a regiones en los locus entre las regiones complementarias a los oligonucleótidos de secuencia fija del conjunto; extender la región entre el primer y segundo oligonucleótido hibridado de al menos un conjunto de oligonucleótidos de secuencia fija con una polimerasa y dNTP para crear dos oligonucleótidos de secuencia fija hibridados de forma adyacente de ese 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 correlaciona con la presencia o ausencia de una anomalía cromosómica y la presencia o ausencia de un polimorfismo en uno o más locus en la muestra mixta.
- 50 60
- 55 65

En algunos casos, la anomalía cromosómica se asocia con duplicación génica o 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 porción adicional de un cromosoma en el genoma. Aún en otros casos, la anomalía cromosómica se asocia con aneuploidía de un cromosoma de interés.

5 Por tanto, en un aspecto, la divulgación proporciona un sistema de ensayo, que no se reivindica, para la detección de la presencia o ausencia de una anomalía cromosómica y la presencia o ausencia de polimorfismos en uno o más locus usando un ensayo único, comprendiendo el ensayo las etapas de introducir un primer conjunto de oligonucleótidos de secuencia fija en una muestra mixta en condiciones que permiten que los oligonucleótidos de secuencia fija se hibriden específicamente a regiones complementarias en dos o más ácidos nucleicos correspondientes a locus informativos en dos o más cromosomas; introducir un segundo conjunto de oligonucleótidos de secuencia fija en la muestra mixta en condiciones que permiten que los oligonucleótidos de secuencia fija se hibriden específicamente a regiones complementarias en ácidos nucleicos indicativas de la presencia o ausencia de polimorfismos en uno o más locus; introducir uno o más oligonucleótidos puente en condiciones que permiten que los oligonucleótidos de secuencia fija se hibriden específicamente a regiones complementarias en los ácidos nucleicos, en el que uno o más oligonucleótidos puente son complementarios a una región de los ácidos nucleicos entre y de inmediato 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 del producto de amplificación se correlaciona con la detección de los locus en la muestra mixta y se puede usar para determinar la cantidad de locus asociados con la presencia o ausencia de la anomalía cromosómica y el estado genético de uno o más locus en la muestra mixta. Los niveles detectados de estos ácidos nucleicos se pueden usar por tanto para determinar la presencia o ausencia de la anomalía cromosómica y la presencia o ausencia de un polimorfismo en uno o más locus de una fuente principal y/o una fuente secundaria dentro de una muestra mixta.

30 Por tanto, en un aspecto, la divulgación proporciona un sistema de ensayo, que no se reivindica, para la detección de la presencia o ausencia de una anomalía cromosómica y la presencia o ausencia de polimorfismos en uno o más locus usando un ensayo único, comprendiendo el ensayo las etapas de introducir un primer conjunto de oligonucleótidos de secuencia fija en una muestra mixta en condiciones que permiten que los oligonucleótidos de secuencia fija se hibriden específicamente a regiones complementarias en dos o más ácidos nucleicos correspondientes a locus informativos en dos o más cromosomas; introducir un segundo conjunto de oligonucleótidos de secuencia fija en la muestra mixta en condiciones que permiten que los oligonucleótidos de secuencia fija se hibriden específicamente a regiones complementarias en ácidos nucleicos indicativas de la presencia o ausencia de polimorfismos en uno o más locus; introducir uno o más oligonucleótidos puente en condiciones que permiten que los oligonucleótidos de secuencia puente se hibriden específicamente a regiones complementarias en los ácidos nucleicos, en el que uno o más oligonucleótidos puente son complementarios a una región de los locus entre las regiones complementarias a los oligonucleótidos de secuencia fija de cada conjunto y de inmediato adyacente a un primer oligonucleótido de secuencia fija; extender la región entre el oligonucleótido puente y el segundo oligonucleótido de secuencia fija usando polimerasa y dNTP; 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 del producto de amplificación se correlaciona con la detección de los locus en la muestra mixta y se puede usar para determinar la cantidad de locus asociados con la presencia o ausencia de la anomalía cromosómica y el estado genético de uno o más locus en la muestra mixta. Los niveles detectados de estos ácidos nucleicos se pueden usar por tanto para determinar la presencia o ausencia de la anomalía cromosómica y la presencia o ausencia de un polimorfismo en uno o más locus de una fuente principal y/o una fuente secundaria dentro de una muestra mixta.

40 50 En otro aspecto, la presente invención utiliza técnicas que permiten la identificación de ambas VNC.

Por tanto, en determinados aspectos, la divulgación proporciona un sistema de ensayo, que no se reivindica, para la detección de la presencia o ausencia de variación en el número de copias (VNC) genéticas de una región genómica y la presencia o ausencia de un agente infeccioso en una muestra mixta usando un ensayo único, comprendiendo el ensayo las etapas de:

introducir un primer conjunto de oligonucleótidos de secuencia fija en una muestra mixta en condiciones que permiten que los oligonucleótidos de secuencia fija se hibriden específicamente a regiones complementarias en uno o más locus en o asociados con una región genómica; introducir un segundo conjunto de oligonucleótidos de secuencia fija en la muestra mixta en condiciones que permiten que los oligonucleótidos de secuencia fija se hibriden específicamente a regiones complementarias en ácidos nucleicos indicativas de un agente infeccioso; unir los oligonucleótidos hibridados para crear un producto de unión contiguo complementario a los ácidos nucleicos; amplificar el producto de unión contiguo para crear productos de amplificación; y detectar los productos de 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 otros aspectos determinados. La divulgación proporciona un sistema de ensayo, que no se reivindica, para la detección de la presencia o ausencia de variación en el número de copias (VNC) de una región genómica, la presencia o ausencia de uno o más polimorfismos y la presencia o ausencia de un agente infeccioso en una muestra mixta de un individuo usando un ensayo único, comprendiendo el ensayo las etapas de: introducir un primer conjunto de oligonucleótidos de secuencia fija en una muestra mixta en condiciones que permiten que los oligonucleótidos de secuencia fija se hibriden específicamente a regiones complementarias en uno o más locus en o asociados con una región genómica; introducir un segundo conjunto de oligonucleótidos de secuencia fija en la muestra mixta en condiciones que permiten que los oligonucleótidos de secuencia fija se hibriden específicamente a regiones complementarias en uno o más locus con un polimorfismo supuesto; introducir un tercer conjunto de oligonucleótidos de secuencia fija en la muestra mixta en condiciones que permiten que los oligonucleótidos de secuencia fija se hibriden específicamente a regiones complementarias en ácidos nucleicos indicativas de un agente infeccioso; unir los oligonucleótidos hibridados para crear un producto de unión contiguo complementario a los ácidos nucleicos; amplificar el producto de unión contiguo para crear productos de amplificación; y detectar los productos de amplificación. La detección de los productos de amplificación se correlaciona con el número de copias de la región genómica, la presencia o ausencia de un polimorfismo en uno o más locus y la presencia o ausencia de un agente infeccioso en la muestra mixta.
- En estos aspectos, los ensayos también pueden comprender introducir uno o más oligonucleótidos puente en condiciones que permiten que los oligonucleótidos puente se hibriden específicamente a regiones complementarias en los ácidos nucleicos, en los que el uno o más oligonucleótidos puente son complementarios a una región de los ácidos nucleicos entre las regiones complementarias a los oligonucleótidos de secuencia fija de cada conjunto.
- 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 del primer y segundo cromosomas de interés, comparar la frecuencia relativa de las regiones seleccionadas del 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 la determinación de 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 del primer y segundo cromosomas de interés, comparar la frecuencia relativa de las regiones seleccionadas del 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 contenida 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 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 con 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 aislan como moléculas individuales para el análisis de los locus seleccionados. Estos productos de amplificación individuales se aislan unos de otros y, preferentemente, se aislan físicamente (por ejemplo, sobre un sustrato o en recipientes individuales). Las moléculas individuales se pueden ampliar además después del aislamiento para preparar múltiples copias idénticas del producto de amplificación, una porción del mismo o un ácido nucleico complementario al producto de amplificación o una porción del mismo. La detección de las moléculas individuales o del producto de amplificación se puede realizar a través de secuenciación.

En algunos aspectos, el producto de unión no se amplifica, sino que se detecta directamente después de la hibridación, por ejemplo, usando técnicas de secuenciación de una única molécula.

- Por tanto, en un aspecto específico, la divulgación proporciona un sistema de ensayo, que no se reivindica, para la detección de la presencia o ausencia de una anomalía cromosómica y polimorfismos en uno o más locus en una muestra mixta usando un ensayo único, comprendiendo el ensayo las etapas de: introducir un primer conjunto de oligonucleótidos de secuencia fija en una muestra mixta en condiciones que permiten que los oligonucleótidos de secuencia fija se hibriden específicamente a regiones complementarias en dos o más ácidos nucleicos correspondientes a un cromosoma particular; introducir un segundo conjunto de oligonucleótidos de secuencia fija en la muestra mixta en condiciones que permiten que los oligonucleótidos de secuencia fija se hibriden específicamente a regiones complementarias en ácidos nucleicos con un polimorfismo supuesto; introducir uno o más oligonucleótidos puente en condiciones que permiten que los oligonucleótidos puente se hibriden específicamente a regiones complementarias en los ácidos nucleicos, en el que uno o más oligonucleótidos puente son complementarios a una región de los ácidos nucleicos entre y de inmediato 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 que tienen la secuencia de los locus; aislar productos de amplificación individuales; y analizar los productos de amplificación individuales para determinar la secuencia de todos o parte de los productos de amplificación individuales. El análisis de los productos de amplificación individuales se correlaciona con la presencia o ausencia de una anomalía cromosómica y el estado genético de uno o más locus en la muestra mixta. Por ejemplo, los niveles de los ácidos nucleicos correspondientes a un cromosoma particular se pueden comparar usando también ácidos nucleicos correspondientes a otro cromosoma, o se pueden comparar con niveles de referencia para el cromosoma que se está consultando.
- En un aspecto preferente, los productos de amplificación individuales se analizan a través de la determinación de secuencia. En otro aspecto, que no se reivindica, los productos de amplificación individuales se analizan usando técnicas de hibridación.

- Es un rasgo característico de la presente invención 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 detección de ensayo utilizan procedimientos de detección no polimórfica para "contar" los números relativos de locus seleccionados presentes en una muestra mixta. Estos números 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 números se pueden utilizar para determinar si, estadísticamente, los ácidos nucleicos de la fuente principal y/o la fuente secundaria tienen uno o más polimorfismos. Dicha información se puede usar para identificar una patología o trastorno genético particular, para confirmar un diagnóstico o recidiva de una enfermedad o trastorno, para determinar el pronóstico de una enfermedad o trastorno, para ayudar a determinar opciones de tratamiento potenciales, etc.

- En algunos aspectos, los procedimientos para la determinación de 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 exclusión de valores atípicos antes de la comparación para determinar la presencia o ausencia de una aneuploidía en una muestra mixta.
- 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.

- 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.

- En un aspecto particular, los locus seleccionados se aislan antes de la detección. 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 para el análisis, 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 la secuencia 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.

Después del aislamiento y la distribución de los ácidos nucleicos seleccionados en un formato adecuado, se identifican las secuencias seleccionadas, por ejemplo, a través de la determinación de secuencia de la secuencia seleccionada.

En un aspecto específico, la invención proporciona un procedimiento de ensayo para la detección de 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 de los cromosomas de interés, comparar la frecuencia relativa de los locus seleccionados del 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.

En algunos aspectos específicos, las frecuencias relativas de los locus de 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 segundo cromosomas 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.

La muestra mixta usada para análisis se puede obtener o derivar de cualquier muestra que contiene el ácido nucleico de interés que se va a analizar usando el procedimiento de ensayo de la invención. La muestra mixta puede ser cualquier fluido materno que comprende ácidos nucleicos libres circulantes tanto maternos como fetales, incluyendo pero sin limitarse a, plasma materno, suero materno o sangre materna.

Preferentemente, el sistema de ensayo se usa para detectar ADNlc en una muestra mixta.

Es un rasgo característico de la invención que los ácidos nucleicos analizados en el procedimiento de ensayo no requieren diferencias polimórficas entre las fuentes celulares para determinar la frecuencia y, por tanto, las VNC potenciales, en locus de una muestra mixta. Otro rasgo característico de la invención es que la gran mayoría de los ácidos nucleicos aislados de la muestra mixta y detectados en el procedimiento de ensayo proporcionan información pertinente sobre la presencia, cantidad y/o naturaleza polimórfica de un locus particular en la muestra mixta. Esto garantiza que la mayoría de los ácidos nucleicos analizados en el procedimiento de ensayo de la invención sean informativos.

En algunos aspectos, se determinan múltiples locus para cada región genómica bajo consulta, y la cantidad de las regiones seleccionadas presentes en la muestra mixta se suman individualmente para determinar la frecuencia relativa de un locus 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 de las frecuencias relativas en la muestra en su conjunto.

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.

La comparación de regiones genómicas puede ser una comparación de parte o de la totalidad de un cromosoma. Por ejemplo, la región genómica detectada para una VNC puede ser un cromosoma completo en el feto (por ejemplo, cromosomas 18 y 21), donde la probabilidad de que ambos sean aneuploides es mínima. También puede ser 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 utilizar dos o más cromosomas que son supuestamente aneuploides y uno o más cromosomas de referencia.

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 de la muestra se analiza para determinar una frecuencia cromosómica relativa para cada cromosoma de interés particular en la muestra. La frecuencia cromosómica de dos o más cromosomas o porciones de los mismos se compara a continuación para determinar estadísticamente si existe una anomalía cromosómica.

En otro aspecto, el procedimiento de ensayo de la invención analiza múltiples ácidos nucleicos que representan

locus seleccionados en cromosomas de interés, y la frecuencia relativa de cada ácido nucleico seleccionado de la muestra se analiza y cuantifica independientemente para determinar una cantidad relativa 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 celular en una muestra mixta.

- 5 En otro aspecto, se analizan subconjuntos de locus en cada cromosoma para determinar si existe una anomalía cromosómica. La frecuencia de locus se puede sumar para un cromosoma particular y las sumas de los locus usar para determinar la aneuploidía. Este aspecto de la invención suma las frecuencias de los locus individuales en cada región genómica y a continuación compara la suma de los locus en una región genómica de un cromosoma con una región genómica de otro cromosoma para determinar si existe una anomalía cromosómica. Los subconjuntos de locus se pueden elegir al azar pero con números suficientes 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 más potencia estadística. En otro aspecto, se pueden seleccionar locus particulares que son conocidos por tener menor variación entre 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.

10 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 de 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.

15 En determinados aspectos específicos, determinar el porcentaje relativo de ácidos nucleicos de la fuente secundaria en una muestra mixta puede ser beneficioso 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. Determinar los locus que contribuyen a la muestra mixta de la fuente secundaria puede proporcionar información usada para calcular las diferencias estadísticamente significativas en frecuencias para regiones genómicas de interés. Por tanto, dichos locus podrían proporcionar dos tipos de información en el ensayo, se puede usar información alélica para determinar el porcentaje de contribución de 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.

20 En otro aspecto específico, el procedimiento de ensayo de la invención se puede utilizar para determinar un mosaicismo potencial en una población celular y si se deben realizar otras pruebas de confirmación para confirmar la identificación del mosaicismo en la fuente principal y/o secundaria. En determinados casos, la determinación del porcentaje de ácidos nucleicos 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 puedan distinguir una aneuploidía total o parcial en mosaico en células o tejidos específicos.

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

30 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.

35 Es un rasgo característico importante del ensayo que los productos de amplificación se analizan directamente sin la necesidad de enriquecer las regiones polimórficas de la muestra mixta inicial. Por tanto, la presente invención permite la detección tanto de VNC como de polimorfismos de una muestra materna sin una etapa de enriquecimiento polimórfico intermedia previa a la determinación de secuencia de los locus seleccionados.

40 Otro rasgo característico importante del ensayo es que tanto la detección de VNC como de polimorfismos 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 polimorfismo consultado en el locus de interés, y obvia la necesidad de generar lecturas de secuencias que se deben alinear con una secuencia de referencia.

45 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

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

divulgación, que no se reivindica.

La FIG. 2 ilustra un primer esquema general para un sistema de ensayo basado en unión de la divulgación, que no se reivindica.

5 La FIG. 3 ilustra un segundo esquema general para un sistema de ensayo basado en unión de la divulgación, que no se reivindica.

10 La FIG. 4 es un tercer esquema general para un sistema de ensayo basado en unión de la divulgación, que no se reivindica.

La FIG. 5 ilustra el rendimiento de genotipado que se obtuvo usando un sistema de ensayo de la divulgación, que no se reivindica.

15 La FIG. 6 ilustra los elementos usados para la detección de aneuploidía y polimorfismo para dos cohortes de muestras maternas.

La FIG. 7 es un sumario de información y datos de pacientes y muestras para un subconjunto de una segunda cohorte de mujeres embarazadas.

20 La FIG. 8 ilustra la detección de aneuploidía del cromosoma 21 lograda usando un aspecto de la invención para una primera cohorte.

25 La FIG. 9 ilustra la detección de aneuploidía del cromosoma 18 lograda usando un aspecto de la invención para una primera cohorte.

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

30 La FIG. 11 ilustra la detección de aneuploidía del cromosoma 18 lograda usando un aspecto de la invención para una segunda cohorte.

### **Definiciones**

35 Los términos usados en el presente documento están destinados a tener el significado simple y ordinario como se entiende por los expertos en la técnica. Las siguientes definiciones están destinadas a ayudar al lector a entender la presente invención, pero no están destinadas a variar o limitar de otro modo el significado de dichos términos a menos que se indique específicamente.

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

45 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.

50 El término "anomalía cromosómica" se refiere a cualquier variación genética que afecta a todo o parte de un cromosoma más grande que un único locus. Las variantes genéticas pueden incluir, pero sin limitarse a, cualquier VNC tal como duplicaciones 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 triple X, síndrome XYY, trisomía 8, trisomía 16, síndrome de Turner, translocación robertsoniana, síndrome de DiGeorge y síndrome de Wolf-Hirschhorn.

55 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 ARN o ADN monocatenarias son sustancialmente complementarias cuando los nucleótidos de una hebra, óptimamente alineados y con inserciones o delecciones de nucleótidos apropiadas, se emparejan con al menos de aproximadamente un 90 % a aproximadamente un 95 % de complementariedad, y más preferentemente de aproximadamente un 98 % a aproximadamente un 100 % de complementariedad, e incluso más preferentemente con un 100 % de complementariedad. De forma alternativa, existe complementariedad sustancial cuando una hebra de ARN o ADN se hibride 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 ( $T_f$ ).

- 5 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 da como resultado una célula que tiene un número anómalo de copias de uno o más locus en el ADN. Las VNC que son clínicamente pertinentes 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 delecionado, 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 secuencia, sino más bien a la cantidad o "recuentos" de regiones genéticas presentes en una muestra.
- 10
- 15 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 experimentales incluyendo la detección de delección, sustitución o inserción de una o más bases durante la secuenciación, así como cambios nucleotídicos 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 son secuencias separadas o se pueden incluir dentro de otras regiones para ayudar a confirmar la exactitud de las técnicas experimentales usadas, por ejemplo, un índice de corrección puede ser un subconjunto de secuencias usadas para la amplificación universal o un subconjunto de nucleótidos de un locus de muestra.
- 20
- 25 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 o ensayo de diagnóstico en una muestra de paciente.
- 30 El término "rasgo de enfermedad" se refiere a un rasgo monogénico o poligénico asociado con una afección patológica, por ejemplo, una enfermedad, trastorno, síndrome o predisposición.
- 35 El término "hibridación" quiere decir 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 hebras se separan, se rehibridarán en las condiciones apropiadas. Se pueden formar híbridos entre ADN-ADN, ADN-ARN o ARN-ARN. Se pueden formar entre una hebra corta y una hebra larga que contiene una región complementaria a la corta. También se pueden formar híbridos imperfectos, pero cuanto más imperfectos sean, menos estables serán (y menos probable que se formen).
- 40 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 porción de un cromosoma consultado para los propósitos de determinar una VNC de todo o parte de ese cromosoma. Los locus informativos para su uso en el procedimiento de ensayo de la invención incluyen locus usados en la consulta de un cromosoma de referencia así como locus usados para la consulta de 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).
- 45
- 50 El término "locus" como se usa en el presente documento se refiere a un locus de localización conocida en un genoma.
- 55 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.
- 60 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 u otras técnicas estándar para extraer muestras periféricas de un sujeto.
- 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 por referencias estándar, una estimación simple del valor de  $T_f$  se puede calcular por 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 está en solución acuosa que tiene concentraciones catiónicas de 0,5 M o menos, el contenido de (G+C) está entre un 30 % y un 70 %, n es el número de bases y m es el porcentaje de emparejamientos erróneos de pares de bases (véase, por ejemplo, Sambrook J et al., *Molecular Cloning, A Laboratory Manual*, 3.<sup>a</sup> ed., Cold Spring Harbor Laboratory Press (2001)). Otras referencias incluyen cálculos más sofisticados que tienen en cuenta características estructurales así como de secuencia para el cálculo de  $T_f$ .

- "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 lleva una matriz de sitios que contienen ácidos nucleicos de modo que cada sitio de la matriz comprende copias idénticas o sustancialmente idénticas de oligonucleótidos o polinucleótidos y está espacialmente definida y no se superpone con otros sitios miembros de la matriz; es decir, los sitios son espacialmente distintos. 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 se pueden unir de forma covalente al soporte sólido, o se pueden unir de forma no covalente. La tecnología de micromatrices convencional se revisa, por ejemplo, en Schena, Ed., 5 *Microarrays: A Practical Approach*, IRL Press, Oxford (2000). "Análisis de matriz", "análisis por matriz" o "análisis por micromatriz" se refiere al análisis, tal como, por ejemplo, análisis de secuencia, de una o más moléculas biológicas usando una micromatriz.
- El término "fuente secundaria" se refiere a una fuente de ácidos nucleicos dentro de un individuo que está presente 10 en cantidades limitadas y que se puede distinguir de la fuente principal debido a diferencias en su composición genómica y/o expresión. Los 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 con trasplante, ácidos nucleicos de un organismo infeccioso en un huésped infectado y similares.
- El término "muestra mixta" como se usa en el presente documento se refiere a cualquier muestra que comprende 15 material genómico libre circulante (por ejemplo, ADN) de dos o más tipos celulares de interés, siendo uno una fuente principal y siendo el otro una fuente secundaria dentro de un único individuo. Las muestras mixtas incluyen muestras con material genómico de una fuente tanto principal como 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 con trasplante que comprende células tanto del donante como del receptor. Las muestras mixtas se derivan 20 preferentemente de forma periférica, por ejemplo, de sangre, plasma, suero, etc.
- 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 25 asociados con una enfermedad, trastorno o predisposición provocada 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).
- El término alelo "no materno" quiere decir un alelo con un polimorfismo y/o una mutación que se encuentra en un 30 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.
- Por "no polimórfica", cuando se usa con respecto a la detección de locus seleccionados, se quiere decir una 35 detección de dicho locus, que puede contener uno o más polimorfismos, pero en la que la detección no depende de la detección del polimorfismo específico dentro de la región. Por tanto, un locus seleccionado puede contener un polimorfismo, pero la detección de la región usando el sistema de ensayo de la divulgación, que no se reivindica, se basa en la aparición de la región en lugar de la presencia o ausencia de un polimorfismo particular en esa 40 región.
- Como se usa en el presente documento, "nucleótido" se refiere a una combinación base-glúcido-fosfato. Los nucleótidos son unidades monoméricas de una secuencia de ácido nucleico (ADN y ARN). El término nucleótido 45 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, [ $\alpha$ 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 50 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.
- De acuerdo con la presente invención, un "nucleótido" puede estar no marcado o estar marcado de forma 55 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.<sup>a</sup> ed., Molecular Probes, Inc., Eugene OR (2002); Keller y Manak, *DNA Probes*, 2.<sup>a</sup> ed., Stockton Press, New York (1993); Eckstein, Ed., *Oligonucleotides and Analogues: A Practical Approach*, IRL Press, Oxford (1991); Wetmur, *Critical 60 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. 65 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; 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'-aminoethyl)aminonaftaleno-1-sulfónico (EDANS). Los ejemplos específicos de nucleótidos marcados de forma fluorescente incluyen [R6G]dUTP, [TAMRA]dUTP, [R100]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-dGTP, 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-dCTP, 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 anómeras 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 a 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 fosforamida 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, trastorno, síndrome o predisposición provocada 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 puede funcionar como molde y del tipo de ácido nucleico que se forma.

Como se usa en el presente documento, "reacción en cadena de la polimerasa" o "PCR" se refiere a una técnica para replicar un trozo específico de ADN seleccionado *in vitro*, incluso en presencia de un exceso de ADN inespecífico. 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 desnaturaliza 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 trozos de ADN adaptados. 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 cebar 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 de muestra exclusivos de modo que cada muestra se marque de forma exclusiva.

El término "locus seleccionado" como se usa en el presente documento se refiere a un locus correspondiente a locus consultados, por ejemplo, respecto a número de copias, la presencia o ausencia de uno o más polimorfismos, 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, en base a 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 el procedimiento 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.

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

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).

El término "estado" como se usa en el presente documento en relación a un gen se refiere al estado de 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(s) mutación/mutaciones causantes de la enfermedad o trastorno, o no afectado, es decir, ambos alelos carecen de dicha(s) mutación/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 mutación/mutaciones causante(s) de la enfermedad o trastorno; portador, es decir, un alelo posee mutación/mutaciones causante(s) de la enfermedad o trastorno; o no afectado, es decir, ambos alelos carecen de dicha(s) mutación/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 trastorno particular o un polimorfismo asociado con un mayor riesgo de desarrollar una enfermedad o trastorno particular.

#### **Descripción detallada de la invención**

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 micromatrizes y 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.<sup>a</sup> ed.) W.H. Freeman, New York (1995); Gait, "Oligonucleotide Synthesis: A Practical Approach" IRL Press, London (1984); Nelson y Cox, *Lehninger, Principles of Biochemistry*, 3.<sup>a</sup> ed., W. H. Freeman Pub., New York (2000); y Berg *et al.*, *Biochemistry*, 5.<sup>a</sup> ed., W.H. Freeman Pub., New York (2002). Antes de que se describan las presentes composiciones, herramientas de investigación y procedimientos, se debe entender que la presente invención no se limita a los procedimientos, composiciones, dianas y usos específicos descritos, ya que, por supuesto, como tales pueden variar. También se debe 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.
- 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.
- Cuando se proporciona un intervalo de valores, se debe 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 excluyen cualquiera de esos límites incluidos también se incluyen en la invención.
- A menos que se establezca expresamente, los términos usados en el presente documento están destinados a tener el significado simple y ordinario como se entiende por los expertos en la técnica. Las siguientes definiciones están destinadas a ayudar al lector a entender la presente invención, pero no están destinadas a variar o limitar de otro modo el significado de dichos términos a menos que se indique específicamente. Todas las publicaciones mencionadas en el presente documento se incorporan por referencia para el propósito de describir y divulgar las formulaciones y metodologías que se describen en la publicación y que se podrían usar en relación con la presente invención descrita.
- En la siguiente descripción se exponen numerosos detalles específicos para proporcionar una comprensión más completa de la presente invención. Sin embargo, será evidente para un experto en la técnica que la presente invención se puede practicar 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 confundir la invención.
- #### 40 La invención en general
- La presente divulgación proporciona sistemas de ensayo único, que no se reivindican, con la capacidad para detectar variaciones en el número de copias, polimorfismos y enfermedad infecciosa en una muestra mixta de un único individuo. Los procedimientos de ensayo comprenden identificar el número de copias de locus seleccionados (incluyendo múltiples locus que se asocian con anomalías cromosómicas) y la detección de polimorfismos que se encuentran en una fuente principal y/o una fuente secundaria. Estos procedimientos son útiles para cualquier muestra mixta que contiene material genómico (por ejemplo, ADN) de una fuente principal y una secundaria que estén presentes en un único individuo.
- El uso de locus seleccionados en los procedimientos de ensayo de la invención proporciona la amplificación de locus para la detección de la variación en el número de copias. 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 de selección se pueden diseñar para cualquier número de locus para cualquier cromosoma. Aunque la amplificación de la muestra mixta antes de la identificación y cuantificación de las regiones de ácidos nucleicos de selección no es obligatoria, se puede realizar una amplificación limitada antes de la detección, en particular si se limitan las cantidades iniciales de ácido nucleico.
- La figura 1 es un diagrama de flujo simplificado de las etapas generales utilizadas en los sistemas de ensayo de la divulgación, que no se reivindica. La figura 1 muestra el procedimiento 100, en el que en una primera etapa 110 se proporciona una muestra de ácido nucleico mixta 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.<sup>a</sup> 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 de procesamiento y/o

5 purificación adicionales 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).

10 En la etapa 120, se introduce un primer conjunto de oligonucleótidos de secuencia fija en la muestra de ácido nucleico mixta, en condiciones que permiten que el primer conjunto de oligonucleótidos de secuencia fija se hibride a la muestra de ácido nucleico mixta. El primer conjunto de oligonucleótidos de secuencia fija comprende secuencias de ácido nucleico que son complementarias a un locus seleccionado(s) en la muestra mixta, que como 15 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 la identificación de anomalías cromosómicas tales como duplicaciones o delecciones, aneuploidías, translocaciones o inversiones.

20 En la etapa 130 se introduce un segundo conjunto de oligonucleótidos de secuencia fija en la muestra de ácido nucleico mixta y el primer conjunto de oligonucleótidos de secuencia fija en condiciones que permiten que el segundo conjunto de oligonucleótidos de secuencia fija se hibride a la muestra de ácido nucleico mixta. El segundo conjunto de oligonucleótidos de secuencia fija comprende secuencias de ácido nucleico que son complementarias a un locus seleccionado(s) en la muestra mixta, que pueden detectar polimorfismos. Se pueden incluir 25 opcionalmente etapas de lavado entre las etapas 120 y 130, y 130 y 140.

25 Aunque la invención se describe como los dos conjuntos de oligonucleótidos introducidos en la muestra mixta secuencialmente, el orden de los conjuntos se puede invertir del descrito en las figuras o, en aspectos preferentes, se pueden introducir simultáneamente.

30 En la etapa 140 se unen el primer y segundo conjuntos de oligonucleótidos de secuencia fija que se han hibridado a 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 la determinación de variaciones en el número de copias o anomalías cromosómicas y la identificación de 35 polimorfismos en la etapa 160.

35 Los conjuntos de ácidos nucleicos de secuencia fija se diseñan para hibridarse al menos a dos regiones separadas en un locus seleccionado. En aspectos preferentes, se usan dos o más oligonucleótidos separados para hibridarse a 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 40 distintas que son complementarias a los locus seleccionados incluyendo sondas precirculares tales como las denominadas "sondas candado" o "sondas de inversión molecular (MIP)".

45 La presente invención proporciona un procedimiento mejorado con respecto a técnicas más aleatorias tales como secuenciación masivamente paralela, secuenciación indiscriminada y el uso de PCR digital aleatoria que se han usado 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 apropiados. 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 50 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.

55 Los ensayos de la presente invención proporcionan una detección dirigida de locus seleccionados, lo que proporciona tanto información sobre el contenido de la región seleccionada (es decir, presencia de una región polimórfica) como información sobre la frecuencia de la región detectada 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 regiones seleccionadas como la presencia o ausencia de polimorfismos en una región seleccionada como un único conjunto de datos del rendimiento de un ensayo multiplexado de la 60 invención.

65 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 selectiva usada en los presentes ensayos se diseña específicamente para proporcionar profundidad de cobertura de locus particulares de interés y proporcionar 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 de más, incluso más preferentemente una cobertura de secuencia de 1000X o más de los locus seleccionados (incluyendo de la una o más fuentes secundarias) presentes en la muestra mixta inicial.

- 5 Una clara ventaja de la invención es que los productos de unión resultantes de los ensayos correspondientes a anomalías cromosómicas y/o anomalías cromosómicas y polimorfismos se pueden analizar 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.
- 10 Los procedimientos de ensayo de la invención proporcionan un uso más eficaz y económico de los datos, y la gran mayoría de las secuencias analizadas después de la amplificación de la muestra dan como resultado información afirmativa sobre la presencia de una VNC particular 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 pertinentes de dicho procedimiento de ensayo de la invención proporciona un uso mucho más eficaz de la recopilación de datos que los enfoques aleatorios enseñados por otros en la técnica.
- 15

#### **Procedimientos de ensayo**

- 20 Los sistemas de ensayo de la invención utilizan un esquema general como se describe anteriormente, aunque se pueden emplear muchas configuraciones y variaciones diferentes, de las que unas pocas se describen a continuación y de las que más se exemplifican en el documento de EE. UU. con n.º de serie 61/371605 presentado el 6 de agosto de 2010 y el documento de EE. UU. con n.º de serie 13/013.732, de los que ambos se incorporan en el presente documento por referencia en su totalidad.
- 25 La figura 2 ilustra un primer esquema general para un sistema de ensayo basado en unión de la divulgación, que no se reivindica. Los oligonucleótidos de secuencia fija 201, 203 comprenden regiones de cebador universal 209 y 211, respectivamente, y regiones complementarias 205 y 207 a los locus seleccionados, respectivamente. Sin embargo, además, el sistema de ensayo en la figura 2 emplea una región de índice de muestra 221 en el primer oligonucleótido de secuencia fija 201. En determinados aspectos, toda o una porción de las secuencias de los locus seleccionados se detectan directamente usando las técnicas descritas, por ejemplo, por técnicas de determinación de secuencia 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.
- 30
- 35 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. Después de 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 se hibride en la etapa 204 a la región del locus 215 entre el primer 201 y segundo 203 oligonucleótidos de secuencia fija. Los oligonucleótidos unidos se unen en la etapa 206 para crear un ácido nucleico contiguo que abarca y es complementario al locus de interés. En determinados aspectos de la invención, los oligonucleótidos puente tienen una longitud de entre 2-45 nucleótidos. En un aspecto específico, los oligonucleótidos puente tienen una longitud de entre 3-9 nucleótidos. Aún en otro aspecto específico, los oligonucleótidos puente tienen una longitud de entre 10-30 nucleótidos.
- 40
- 45 Despues de la unión, el producto de unión se eluye del molde de ADNg. Los cebadores universales 217, 219 se introducen en la etapa 208 para amplificar el primer y segundo oligonucleótidos de secuencia fija unidos para crear 210 los 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 secuencia. En aspectos específicos, es deseable determinar las secuencias tanto del índice como de los productos de amplificación, por ejemplo, para proporcionar identificación de la muestra así como del locus. Los índices contemplados en la invención 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.
- 50
- 55
- 60 En aspectos preferentes, se usan índices representativos de la muestra mixta de la que se puede aislar un ácido nucleico 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. Los oligonucleótidos identificados de forma exclusiva se pueden combinar a continuación en un único recipiente de reacción con ácidos nucleicos de otras muestras mixtas antes de la secuenciación. 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.
- 65

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 se localicen entre las regiones de cebador universal 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.

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 fijos de oligonucleótidos de secuencia 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 secuencia 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 al locus seleccionado 305, 307 y las regiones de cebador universal 309, 311 usadas para amplificar los diferentes locus seleccionados después de la selección y/o aislamiento iniciales de los locus seleccionados de la muestra mixta. Las regiones de cebador universal 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 seleccionados 315 o 325. Después de la hibridación, los oligonucleótidos de secuencia fija no hibridados se separan preferentemente del resto de la muestra genética, por ejemplo, por lavado (no mostrado).

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.

Después de 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 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 cuantifican a través de la determinación de secuencia del producto de unión, a través del índice de muestra y/o la región del producto que contiene el SNP en el locus seleccionado.

En una configuración alternativa de los procedimientos de la invención, el oligonucleótido puente se puede hibridar a 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 oligonucleotídicas complementarias a 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 porciones complementarias del locus de interés 415. Después de la hibridación, los oligonucleótidos de secuencia fija no hibridados se separan preferentemente del resto de la muestra genética (no mostrado).

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 se hibrida a una región directamente adyacente a la región del primer oligonucleótido de secuencia fija 405, pero está separado por uno o más nucleótidos de la región complementaria del segundo oligonucleótido de secuencia fija 407. Después de 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 entre el oligonucleótido puente 413 y el segundo oligonucleótido de secuencia fija 403. Después de la extensión, los oligonucleótidos unidos se unen en la etapa 408 para crear un ácido nucleico contigo que abarca y es complementario al locus de interés 415. Después 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. Estos productos de amplificación 423 se aislan, 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**

Los sistemas de ensayo utilizan sondas de ácido nucleico diseñadas para identificar, y preferentemente para aislar, 20 regiones de ácidos nucleicos seleccionadas en una muestra mixta. Algunas de las sondas identifican secuencias de interés en locus seleccionados consultados respecto al número de copias (es decir, 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.

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 se hibridan específicamente a 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 porción sustancial del genoma. Además, aunque la muestra inicial se puede enriquecer usando procedimientos tales como amplificación general para incrementar el número de copias de ácidos nucleicos en la muestra mixta, preferentemente no se usan etapas de enriquecimiento antes de las etapas de hibridación, unión y amplificación usadas para identificar los locus de interés.

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. Se analiza la frecuencia o cantidad de los múltiples locus seleccionados para determinar si existe una anomalía cromosómica, y se analiza la presencia o ausencia de un polimorfismo para determinar la presencia o ausencia o cálculo de probabilidad de una anomalía cromosómica en una fuente en la muestra mixta.

En aspectos preferentes, se pueden amplificar múltiples locus seleccionados de dos o más muestras en un único recipiente de reacción, y analizar simultáneamente la información en un único conjunto de datos, por ejemplo, a través de la determinación de secuencia. 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/o la presencia o ausencia de polimorfismos para muestras individuales.

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 en la probabilidad de aneuploidía en base a las proporciones de los cromosomas. La normalización o estandarización de las frecuencias se puede realizar para uno o más locus seleccionados.

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 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 o bien en diferentes cromosomas.

En determinados aspectos, los datos usados para determinar la frecuencia de los locus seleccionados pueden excluir datos atípicos que parece que se deben a 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 en particular 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 en particular baja en una o más muestras.

- En otro aspecto, 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 más potencia estadística. Por ejemplo, si existen 100 regiones seleccionadas para el cromosoma 21 y 100 regiones seleccionadas para el cromosoma 18, se podría 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.
- La cantidad de diferentes ácidos nucleicos detectables en determinados cromosomas puede variar dependiendo de varios factores, incluyendo representación general de locus en diferentes fuentes celulares en muestras maternas, 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 locus se suman para un cromosoma particular y la suma de los locus se usa para determinar la presencia de aneuploidía. Este aspecto de la invención suma las frecuencias de los locus individuales en cada cromosoma y a continuación compara la suma de los locus en un cromosoma con la de otro cromosoma para determinar si existe una anomalía cromosómica.
- 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 dicha amplificación selectiva diseñados para aislar regiones se pueden elegir por diversos motivos, pero se diseñan preferentemente para 1) amplificar eficazmente una región del cromosoma de interés; 2) tener un intervalo predecible de expresión de fuentes maternas y/o fetales en diferentes muestras mixtas; 3) ser distintivos del cromosoma particular, es decir, no amplificar regiones homólogas en otros cromosomas.
- El procedimiento de ensayo de la invención detecta tanto aneuploidías fetales como anomalías cromosómicas específicas a través de la identificación y cuantificación de locus específicos de interés. Dichas anomalías adicionales incluyen, pero no se limitan a, mutaciones de delección, mutaciones de inserción, polimorfismos de número de copias, variantes del número de copias, síndrome de delección 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 secuencias de ácido nucleico seleccionadas presentes en los mismos cromosomas o en cromosomas separados, y al menos uno de los locus seleccionados se asocia con la anomalía alélica fetal. Las secuencias de los dos locus seleccionados y el número de copias de los dos locus seleccionados se comparan a continuación para determinar si la anomalía cromosómica está presente y si está, la naturaleza de la anomalía.
- Si bien gran parte de la descripción contenida en el presente documento describe la detección de aneuploidías contando la abundancia de locus en uno o más supuestos cromosomas aneuploides y la abundancia de locus en uno o más cromosomas normales, se pueden usar las mismas técnicas para detectar variaciones en el número de copias donde dicha variación en el número de copias se produce solo en una porción de un cromosoma. En esta detección de las variaciones en el número de copias, se comparan múltiples locus dentro de la supuesta localización de la variación en el número de copias con múltiples locus 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 regiones nucleicas 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. La abundancia de cada locus se determina por los procedimientos descritos en esta solicitud.
- En algunos aspectos, se puede usar una amplificación universal para amplificar los locus. En algunos aspectos, los locus 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.
- 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. Los locus dentro de la delección se suman a continuación como también los locus fuera de la delección. Estas sumas se comparan a continuación entre sí para determinar la presencia o ausencia de una delección. Opcionalmente, las sumas se ponen en una proporción y esa proporción se puede comparar con una proporción promedio creada a partir de una población normal. Cuando la proporción

para una muestra se encuentra estadísticamente fuera de una proporción esperada, se detecta la delección. El umbral para la detección de una delección puede ser de dos o más veces, preferentemente cuatro o más veces la variación calculada en la población normal.

##### 5   **Polimorfismos asociados con enfermedades o predisposiciones**

Los sistemas de ensayo de la divulgación, que no se reivindican, se utilizan para detectar polimorfismos, tales como los asociados con una enfermedad autosómica dominante o recesiva o un trastorno de predisposición. Dada la naturaleza multiplexada del sistema de ensayo de la divulgación, esta detección tiene lugar en el mismo ensayo que la detección de anomalías cromosómicas. Por tanto, un sistema de ensayo único puede proporcionar información de diagnóstico sobre diferentes clases de mutaciones genéticas. En consecuencia, como los sistemas de ensayo preferentes de la divulgación están altamente multiplexados y pueden consultar cientos o incluso miles de ácidos nucleicos dentro de una muestra mixta, en determinados aspectos es deseable consultar la muestra respecto a marcadores de ácido nucleico dentro de la muestra mixta, por ejemplo, ácidos nucleicos asociados con riesgo genético o que identifican la presencia o ausencia de organismos infecciosos. Por tanto, los sistemas de ensayo proporcionan la detección de dichos ácidos nucleicos junto con la detección de ácidos nucleicos para la determinación del número de copias dentro de una muestra mixta.

Por tanto, el sistema de ensayo de la divulgación, que no se reivindica, 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 desarrollo de una enfermedad y/o de progresión de la enfermedad (por ejemplo, metástasis) e indicadores de pronóstico.

En otro aspecto específico, el sistema de ensayo de la divulgación, que no se reivindica, 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, anomalías congénitas 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 BRCAI 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 divulgación, que no se reivindica, 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 el procedimiento de ensayo de la invención, y dichos procedimientos se usan preferentemente para incrementar el número de copias de un locus de interés en una muestra mixta de manera que permita la conservación de información sobre el contenido inicial del locus en la muestra mixta. Aunque no todas las combinaciones de amplificación y análisis se describen en el presente documento en detalle, 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 ácido nucleico (NASBA). (Véanse las patentes de EE. UU. n.º 5.409.818, 5.554.517 y 6.063.603, de las que cada una se incorpora en el presente documento por referencia). 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, de los que cada uno se incorpora en el presente

documento por referencia. 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 basada en secuencias de ácido nucleico (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 se pueden usar para amplificar ADN de cualquier longitud siempre que contenga el locus de interés en su secuencia.

- 5 10 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 distinguirlo de otros ácidos nucleicos que se amplifican y/o seleccionan. En general, un ácido nucleico amplificado tiene una longitud de al menos aproximadamente 16 nucleótidos, y más típicamente, un ácido nucleico amplificado tiene una longitud de al menos aproximadamente 20 nucleótidos. En un aspecto preferente de la invención, un ácido nucleico amplificado tiene una longitud de al menos aproximadamente 30 nucleótidos. En un aspecto más preferente de la invención, un ácido nucleico amplificado tiene una longitud de al menos aproximadamente 32, 40, 45, 50 o 60 nucleótidos. En otros aspectos de la invención, un ácido nucleico amplificado puede tener una longitud de aproximadamente 100, 150 o hasta 200.
- 15 20 25 En determinados aspectos, la amplificación seleccionada comprende una etapa de amplificación lineal inicial. Esto puede ser en particular útil si la cantidad de partida de ADN de la muestra mixta es bastante limitada, por ejemplo, donde 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 donde se necesita una cuantificación exacta del ADN o de una fracción del ADN (por ejemplo, contribución de ADN fetal en una muestra materna).

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. 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 proporciona múltiples fragmentos de ADN de diferentes regiones. A continuación, los productos de amplificación de la amplificación lineal se podrían amplificar además con procedimientos de PCR estándar o con una amplificación lineal adicional.

- 30 35 40 45 50 55 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 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 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 se someten a procedimientos bioquímicos adicionales, tales como amplificación adicional con otros cebadores y/o técnicas de detección, tales como determinación de secuencia e hibridación.
- 55 Las eficacias de la amplificación lineal pueden variar entre sitios y entre ciclos, de modo que en determinados sistemas se puede usar una normalización para garantizar que los productos de la amplificación lineal sean representativos del contenido de ácido nucleico en el material de partida. Quien practica el procedimiento de ensayo de la invención puede utilizar información de diversas muestras para determinar la variación en los niveles de ácido nucleico, incluyendo la variación en diferentes locus en muestras individuales y/o entre los mismos locus en diferentes muestras después de la amplificación lineal inicial limitada. Dicha información se puede usar en la normalización para evitar la desviación de los niveles iniciales de contenido de ADN.

### **Amplificación universal**

- 60 65 En los aspectos preferentes de la invención, los locus amplificados selectivamente se amplifican además preferentemente a través de la amplificación universal de todos o sustancialmente todos los diversos locus que se van a analizar usando el procedimiento de ensayo de la invención. Las regiones de cebador universal se añaden a los oligonucleótidos de secuencia fija de modo que los locus amplificados selectivamente se puedan amplificar además en una única reacción de amplificación universal. Estas secuencias de cebador universal se pueden añadir a las regiones de ácidos nucleicos durante el procedimiento de amplificación selectiva, es decir, los cebadores para la amplificación selectiva tienen secuencias de cebador universal que flanquean un locus. De forma

alternativa, los adaptadores que comprenden secuencias de amplificación universal se pueden añadir a los extremos de los ácidos nucleicos seleccionados como adaptadores después de la amplificación y aislamiento de los ácidos nucleicos seleccionados de la muestra mixta.

- 5 En un aspecto ejemplar, los ácidos nucleicos se amplifican inicialmente a partir de una muestra mixta usando cebadores complementarios a regiones seleccionadas de los cromosomas de interés, seguido de una etapa de amplificación universal para incrementar el número de locus para el análisis. Esta introducción de regiones de cebador en los productos de amplificación inicial de una muestra mixta permite una amplificación universal controlada posterior de todos o una porción de los ácidos nucleicos seleccionados antes de o durante el análisis, 10 por ejemplo, la determinación de secuencia.

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 locus se amplifiquen a diferentes tasas o eficacia. Parte de esto se puede deber a la variedad 15 de cebadores en una reacción múltiple teniendo algunos mejor eficacia (es decir, hibridación) que otros, o funcionando algunos mejor en condiciones experimentales específicas debido a la composición de bases. Cada conjunto de cebadores para un locus dado se puede comportar de forma diferente en base al contexto de secuencia del cebador y ADN molde, condiciones del tampón y otras condiciones. Una amplificación de ADN universal para 20 un sistema de ensayo multiplexado introducirá en general menos sesgo y variabilidad.

20 En consecuencia, en un aspecto, se realiza un pequeño número (por ejemplo, 1-10, preferentemente 3-5) de ciclos de amplificación seleccionada o enriquecimiento de ácido nucleico de la muestra inicial en una reacción de mezcla multiplexada, seguido de amplificación universal usando cebadores universales introducidos. El número de ciclos que usan cebadores universales variará, pero preferentemente será de al menos 10 ciclos, más preferentemente 25 de al menos 5 ciclos, incluso más preferentemente de 20 ciclos o más. Al pasar a la amplificación universal después de un número menor de ciclos de amplificación, se reduce el sesgo de que determinados locus se amplifiquen a tasas mayores que otros.

30 Opcionalmente, el sistema de ensayo incluirá una etapa entre la amplificación seleccionada y la amplificación universal para retirar cualquier exceso de ácidos nucleicos que no se amplifique específicamente en la amplificación seleccionada.

35 El producto completo o una alícuota del producto de la amplificación seleccionada se puede usar para la amplificación universal. Se pueden usar las mismas o diferentes condiciones (por ejemplo, polimerasa, tampones, y similares) en las etapas de amplificación, por ejemplo, para garantizar que no se introducen de forma inadvertida sesgo y variabilidad debido a condiciones experimentales. Además, se pueden usar variaciones en las concentraciones de cebador para limitar eficazmente el número de ciclos de amplificación específicos de secuencia.

40 En determinados aspectos, las regiones de cebador universal 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 45 la cuantificación exhaustiva de la presencia de locus dentro de una muestra mixta de este tipo para la determinación de aneuploidía.

50 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.

55 Algunos aspectos utilizan reacciones acopladas para la detección multiplexada de secuencias de ácido nucleico donde los oligonucleótidos de una fase temprana de cada procedimiento contienen secuencias que se pueden usar por 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, de los que cada uno se incorpora por referencia en su totalidad.

60 En determinados aspectos, el procedimiento 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 con 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 aplicabilidad particular en 65 la detección de determinadas características de ácidos nucleicos. Sin embargo, cada uno requiere el uso de reacciones acopladas para la detección multiplexada de diferencias en las secuencias de ácido nucleico donde los oligonucleótidos de una fase temprana de cada procedimiento contienen secuencias que se pueden usar por 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 ácido nucleico.

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 con LDR para la detección de ácido nucleico.

10 Barany *et al.*, patentes de EE. UU. n.º 7.556.924 y 6.858.412, describen el uso de sondas precírculo (también llamadas "sondas candado" o "sondas de inversión múltiple") con LDR acoplada y reacción en cadena de la polimerasa ("PCR") para la detección de ácido nucleico.

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 y unión de endonucleasas para la detección de secuencias de ácido nucleico.

Willis *et al.*, patentes de EE. UU. n.º 7.700.323 y 6.858.412, describen el uso de sondas precirculares en la amplificación, detección y genotipado de ácido nucleico multiplexado.

20 Ronaghi *et al.*, patente de EE. UU. n.º 7.622.281 describe técnicas de amplificación para marcar y amplificar un ácido nucleico usando un adaptador que comprende un cebador exclusivo y un código de barras.

25 En algunos casos, un ensayo único puede emplear una combinación de los procedimientos descritos anteriormente. Por ejemplo, algunos de los locus se pueden detectar usando oligonucleótidos de secuencia fija que se unen a regiones complementarias adyacentes en un locus, mientras que otros locus se pueden detectar usando locus puente en el mismo ensayo. En otro ejemplo, algunos de los locus se pueden detectar usando oligonucleótidos de secuencia fija que se unen a regiones complementarias adyacentes en un locus, mientras que otros locus pueden requerir una etapa de extensión del cebador para unirse a los oligonucleótidos de secuencia fija.

30 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 por análisis de los productos de amplificación. En un aspecto preferente, una muestra representativa de moléculas individuales de los procesos de amplificación se aísla de las otras moléculas para su análisis posterior. 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 40 En un aspecto preferente, los productos de amplificación se aislan 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. La molécula aislada se puede asociar con una dirección física o identificable particular antes del análisis, o bien 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.

45 Una vez aislado, el producto de amplificación individual se puede amplificar además 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.

50 55 En un aspecto preferente, la longitud completa del producto de amplificación o una porción del producto de amplificación se puede analizar usando determinación de secuencia. 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 producto de amplificación se analiza a través de la determinación de secuencia del producto de amplificación seleccionado.

60 65 Numerosos procedimientos de determinación de secuencia son compatibles con los procedimientos de ensayo de las invenciones. Los procedimientos ejemplares para la determinación de secuencia incluyen, pero no se limitan a, incluyendo, pero sin limitarse a, procedimientos basados en hibridación, tales como los divulgados en Drmanac, patentes de EE. UU. n.º 6.864.052; 6.309.824; y 6.401.267; y Drmanac *et al.*, publicación de patente EE. UU. n.º 2005/0191656, que se incorporan por referencia, procedimientos de secuenciación por 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 con pirofosfato como se describe en Ronaghi *et al.*, patentes de EE. UU. n.º 7.648.824, 7.459.311, 6.828.100 y 6.210.891; y procedimientos de determinación de secuenciación basados 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.

La información de secuencia se puede determinar usando procedimientos que determinen muchas (típicamente miles a miles de millones) de secuencias de ácido nucleico 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, como se comercializa por 454 Life Sciences, Inc., Branford, CT); secuenciación por unión (por ejemplo, como se comercializa en la tecnología SOLiD™, Life Technology, Inc., Carlsbad, CA); secuenciación por síntesis usando nucleótidos modificados (tal como se comercializa 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), tecnologías de secuenciación por detección de iones (Ion Torrent, Inc., South San Francisco, CA); secuenciación de nanobolas de ADN (Complete Genomics, Inc., Mountain View, CA); tecnologías de secuenciación basadas en nanoporos (por ejemplo, como se desarrolla 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 porción del producto de amplificación se puede analizar usando técnicas de hibridación. Los procedimientos para realizar ensayos de hibridación de polinucleótidos para la detección se han desarrollado bien en la técnica. Los procedimientos y condiciones de ensayo de hibridación variarán dependiendo de la aplicación y se seleccionan de acuerdo con los procedimientos de unión generales conocidos incluyendo los mencionados en: Maniatis *et al.* *Molecular Cloning: A Laboratory Manual* (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). Se han descrito procedimientos y aparatos para llevar a cabo reacciones de hibridación repetidas y controladas en las patentes de EE. UU. n.º 5.871.928, 5.874.219, 6.045.996 y 6.386.749, 6.391.623, de las que cada una se incorpora en el presente documento por referencia.

La presente divulgación también contempla la detección de señales de hibridación entre ligandos en determinados aspectos preferentes. Véanse las patentes de EE. UU. n.º 5.143.854, 5.578.832; 5.631.734; 5.834.758; 5.936.324; 5.981.956; 6.025.601; 6.141.096; 6.185.030; 6.201.639; 6.218.803; y 6.225.625, en la solicitud de patente de EE. UU. 60/364.731 y en la solicitud PCT PCT/US99/06097 (publicada como documento WO99/47964) de las que cada una también se incorpora en el presente documento por referencia en su totalidad para todos los propósitos.

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) de las que cada una también se incorpora en el presente documento por referencia en su totalidad para todos los propósitos.

#### **Minimización de la variación dentro de y entre muestras**

Un desafío con la detección de anomalías cromosómicas en un feto por detección en una muestra mixta es que los ácidos nucleicos de la fuente secundaria están presentes en una abundancia mucho menor que los ácidos nucleicos del tipo celular normal.

La variación entre los niveles encontrados entre muestras y/o para locus dentro de una muestra se puede minimizar en una combinación de procedimientos analíticos, de los que muchos 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. Si bien el uso de un cromosoma "normal" de este tipo como cromosoma de referencia puede ser suficiente, también es posible el uso de dos o más 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 para cada cromosoma se suman para calcular los recuentos totales para cada cromosoma. Los recuentos totales de un cromosoma se dividen a continuación 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 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.

- 5 El promedio puede ser la media, mediana, moda u otro promedio, con o sin normalización y exclusión de datos atípicos. En un aspecto preferente, se usa la media. Al desarrollar el conjunto de datos para la proporción cromosómica de la población normal, se calcula la variación normal de los cromosomas medidos. Esta variación se puede expresar de varias 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 deseado. En general, este umbral puede ser un múltiplo de la variación observada en la proporción cromosómica. En un ejemplo, este umbral es tres o más veces la variación de la proporción cromosómica. En otro ejemplo, es cuatro o más veces la variación de la proporción cromosómica. En otro ejemplo, es cinco o más veces la variación de la proporción cromosómica. En otro ejemplo, es seis o más veces la variación de la proporción cromosómica. En el ejemplo anterior, la proporción cromosómica se determina sumando los recuentos de locus 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 para cada cromosoma. El promedio puede ser cualquier estimación de la media, mediana o 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 se han 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 para 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<sup>1/2</sup>. 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.
- 40 En un aspecto preferente de la presente invención, el número de locus 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. Hay un coste incremental para medir cada locus y por tanto es importante minimizar el número de cada locus seleccionado. 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, después de 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 abundancia. En otro aspecto, el percentil puede ser el 10 % más bajo y más alto como se mide por abundancia. En otro aspecto, el percentil puede ser el 25 % más bajo y más alto como se mide por abundancia.
- 60 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 2000 veces. En un aspecto preferente de la invención, los locus seleccionados se miden cada uno en promedio al menos 5000 veces.
- 5           10          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 más potencia estadística. En este ejemplo, puede ser necesario o no retirar o eliminar cualquier locus 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, de las que muchas se describen anteriormente en esta solicitud. En general, la variación en el ensayo se puede reducir cuando todos los locus 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 se limita el número de ciclos de amplificación.

## 25          **Uso de sistemas de ensayo para detección en muestras mixtas de pacientes con cáncer**

- El sistema de ensayo permite la detección de 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           35          En algunos aspectos, el procedimiento 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., junio de 2011; 5(3):281 - 91. Epub de 24 de febrero de 2011). También se han detectado variaciones en el número de copias en determinados cánceres, tales como HER2 y receptor de estrógeno amplificados en el ADNlc de pacientes con cáncer de mama. (Page K., Br J Cancer. 12 abril de 2011; 104(8):1342-8. Epub de 22 de marzo de 2011).

40           45          En otros aspectos de la divulgación, el sistema de ensayo, que no se reivindica, se usa en pacientes con cáncer para seguir una respuesta al tratamiento y/o controlar el progreso de la enfermedad, por ejemplo, para medir alteraciones de un único gen y ADNlc en pacientes que reciben quimiorradioterapia (QRT). Para determinados cánceres, se ha demostrado que el índice de integridad del ADNlc puede estar asociado significativa e independientemente 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 en el número de copias se ha asociado con la respuesta a la quimioterapia y/o el pronóstico de una enfermedad. 50           Véase, por ejemplo, Savas S., Acta Oncol. Noviembre de 2010; 49(8):1217-26. Epub de 29 de julio de 2010, que describe variaciones genéticas útiles para la determinación de la respuesta al tratamiento y la supervivencia en el cáncer. Por ejemplo, la detección de niveles de ADNlc combinada con la detección de mutaciones en el gen KRAS y/o el gen p53 proporcionan una herramienta poderosa y relativamente no invasiva para medir el pronóstico de diversos cánceres, incluyendo cáncer de ovario, cáncer de endometrio y linfomas. Dobrzycka B *et al.*, Ann Oncol., mayo de 2011; 22(5):1133-40. Epub de 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 Lett. 18 de marzo de 2009; 275(2):234-9. Epub de 28 de noviembre de 2008. Dicho análisis se puede respaldar además 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.

## 60          **Uso de sistemas de ensayo para detección en muestras mixtas de pacientes con trasplante**

El sistema de ensayo de la divulgación, que no se reivindica, se puede usar para seguir la salud de los órganos en un paciente con trasplante usando una combinación de detección de ADNlc y detección de SNP o mutaciones en uno o más genes únicos. Los órganos transplantados 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 de donante en receptores de trasplante de corazón. Snyder TM *et al.*, Proc Natl Acad Sci USA. 12 de abril de 2011; 108(15):6229-34. Epub de 28 de marzo de 2011. Además, las quimiocinas y moléculas de adhesión median en 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, el sistema de ensayo de la divulgación puede proporcionar pruebas no invasivas para seguir a receptores de trasplantes de órganos sólidos y puede ayudar en la identificación de los primeros signos de rechazo sin la necesidad de biopsias de órganos u otras técnicas de diagnóstico o pronóstico más costosas.

#### **10 Uso de sistemas de ensayo para 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 donde 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 en las variaciones de niveles de locus seleccionados identificados en una muestra materna. En otros aspectos, la determinación del porcentaje de ADN<sub>f</sub> fetal relativo 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 de ADN entre ADN fetal y materno. En un aspecto preferente, el ADN amplificado a partir de ADN libre de plasma es 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.

En aspectos particulares, el porcentaje de ADN fetal libre en la muestra materna se puede determinar por PCR usando ADN diluido en serie aislado de la muestra materna, lo que puede cuantificar con exactitud el número de genomas que comprenden los genes amplificados.

En circunstancias donde el feto es de sexo masculino, el porcentaje de ADN fetal en una muestra se puede determinar a través de la detección de locus específicos del cromosoma Y y la comparación con el contenido de ADN materno calculado. Las cantidades de un locus específico del cromosoma Y amplificado, 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 está 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 bien después de la amplificación selectiva dependiendo de la naturaleza del ensayo multiplexado.

En aspectos particulares, el porcentaje de ADN fetal libre circulante en una muestra materna se puede determinar por PCR usando ADN diluido en serie aislado de la muestra materna, lo que puede 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 un 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.

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.<sup>o</sup> de copias del gen ligado al cromosoma Y x 2 x 100)/(n.<sup>o</sup> 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.

### 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 heredados de la madre

5 y un 50 % de los locus heredados del padre. Determinar los locus contribuidos al feto a partir de fuentes paternas puede permitir la estimación de ADN fetal en una muestra materna y por tanto proporcionar información usada para calcular las diferencias estadísticamente significativas en las frecuencias cromosómicas para los cromosomas de interés.

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

15 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. En un aspecto preferente, se usa secuenciación de alto rendimiento para dicha determinación y cuantificación. Se identifican los locus donde los genotipos de la fuente 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 en particular 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 Para muchos alelos, las secuencias de ácido nucleico de la fuente 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 pertenecientes a regiones alélicas que son iguales para la fuente 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.

25 40 45 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.

50 55 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

60 65 En un aspecto específico, la contribución de la fuente secundaria se puede cuantificar usando la detección de SNP en tandem. Las técnicas para identificar SNP en tandem 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 tandem 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.

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; 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 análisis de locus seleccionados en una muestra mixta incluyen tanto regiones para la determinación del contenido en ADN de la fuente secundaria así 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 diversas etapas en el sistema de ensayo que de otro modo podría sesgar los resultados cuando se utiliza contenido en 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 de ADN como para la detección de anomalías cromosómicas puede ayudar además a minimizar cualquier sesgo debido a un error experimental o contaminación.

### 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. Se apreciará por los expertos en la técnica que se pueden 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.

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.

#### Ejemplo 1: aspectos generales del sistema de ensayo de la divulgación

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 fueron indicativos de la presencia de un cromosoma particular o de la presencia o ausencia de una mutación o polimorfismo en un alelo particular.

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: **TACACCGGCGITATGCGTCGAGAC** (SEQ ID NO: 1); un índice de identificación de nueve nucleótidos específico para el 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 un índice de locus/alelo en el segundo conjunto específico de polimorfismo; un nucleótido que interrumpe la hibridación 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 donde se detectó un SNP o una mutación en esta porción 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 en el conjunto de ensayo.

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' en el 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: **ATTGCCGGGACCGATGATCCCGTC** (SEQ ID NO:2).

En los casos donde 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 que cada ácido nucleico seleccionado proporcione una  $T_f$  prevista uniforme con una variación de dos grados en todas las consultas en el conjunto de ensayo que fue sustancialmente 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 fueron 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 hombre de raza blanca (NA12801) o una mujer de raza blanca (NA11995) de los depósitos celulares 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 de extremo 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 T4 polinucleótido cinasa (Enzymatics, Beverly, MA) y H<sub>2</sub>O hasta 120 µl. Esto se incubó a 37 °C durante 30 minutos. El ADN se diluyó usando Tris 10 mM, EDTA 1 mM, pH 8,5 hasta la concentración final deseada de ~2 ng/µl.

Se colocaron 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, se incubó la placa a 95 °C durante 3 minutos, se enfrió a 25 °C y se centrifugó de nuevo 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), CoCl<sub>2</sub> 250 µM, 0,01 nmol/µl de biotina-16-dUTP (Roche, Nutley, NJ) y H<sub>2</sub>O hasta 1,5 ml. Se alicuotaron 15 µl de la mezcla maestra en cada pocillo de una placa de 96 pocillos, y se selló la placa con sellador de placa adhesivo. La placa se centrifugó durante 10 segundos a 250 x g y se incubó durante 37 °C durante 60 minutos. Despues de la incubación, la placa se centrifugó de nuevo 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 en 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 resuspendería 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 donde se detectaron SNP, los formatos de ensayo utilizaron 192 o más consultas separadas, utilizando cada una la detección de alelos diferentes por 96 locus en muestras genéticas. Los ejemplos descritos para cada formato de ensayo utilizaron dos conjuntos diferentes de oligonucleótidos de secuencia fija y/u 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.

Un primer formato de ensayo ejemplar usó 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 SNP, y se usó una ligasa para sellar las mellas

(nicks) formadas de este modo. La discriminación de bases de SNP derivó de la fidelidad de la incorporación de bases por la polimerasa y, en el caso de incorporación errónea, de la tendencia de la ligasa a no sellar mellas adyacentes a bases mal emparejadas.

5 El segundo formato de ensayo ejemplar usó dos oligonucleótidos de secuencia fija específicos de locus sin un oligonucleótido puente, donde existía un hueco de ~15-35 bases entre los oligonucleótidos de secuencia fija y 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 sellar la mella formada de este modo. La discriminación de bases de SNP derivó de la fidelidad de la incorporación de bases por la polimerasa y, en el caso de incorporación errónea, de la tendencia de la ligasa a no sellar mellas adyacentes a bases mal emparejadas.

10 Un tercer formato de ensayo ejemplar usó un primer y segundo oligonucleótidos de secuencia fija específicos de alelo sin un oligonucleótido puente, donde existía 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 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 sellar la mella formada de este modo. La discriminación de bases de SNP derivó de la especificidad de 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 sellar mellas adyacentes a bases mal emparejadas.

15 20 Un cuarto formato ejemplar usó 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 para un alelo del SNP dirigido y el segundo con una base en 3' específica para el otro alelo del SNP dirigido. De forma similar, se usaron dos segundos oligonucleótidos de secuencia fija separados, el primero con una base en 5' específica para un alelo de un segundo SNP dirigido y el segundo con una base en 5' específica para el otro alelo del segundo SNP dirigido. Los oligonucleótidos puente eran complementarios a la región directamente adyacente a las regiones de 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 mellas entre los oligonucleótidos de secuencia fija y el oligonucleótido puente. La discriminación de bases de SNP en este formato de ensayo derivó de la especificidad de hibridación y la tendencia de la ligasa a no sellar las mellas adyacentes a las bases mal emparejadas. Este formato ejemplar se sometió a prueba usando T4 ligasa o bien Taq ligasa para la creación del molde contiguo, y ambos demostraron ser eficaces en la reacción como se describe a continuación.

25 30 35 Un quinto formato ejemplar usó oligonucleótidos de secuencia fija específicos de locus que fueron 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 sellar la única mella entre los oligonucleótidos.

40 45 Un sexto formato ejemplar usó 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 locus complementaria a los oligonucleótidos de secuencia fija. El oligonucleótido puente específico de locus de este ejemplo fue un pentámero complementario a las regiones directamente adyacentes a las regiones complementarias al primer y segundo oligonucleótidos de secuencia fija. En este formato, no se requirió ninguna polimerasa y se usó una ligasa para sellar las dos mellas entre los oligonucleótidos.

50 55 Un séptimo formato ejemplar usó 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 sellar las dos mellas entre los oligonucleótidos. La discriminación de bases de SNP en este formato de ensayo derivó de la especificidad de hibridación y la tendencia de la ligasa a no sellar las mellas adyacentes a las bases mal emparejadas.

60 65 Un octavo formato ejemplar usó 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 sellar las tres mellas entre los oligonucleótidos.

Para cada uno de los formatos de ensayo descritos anteriormente, se creó un grupo 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 expone en el ejemplo 2. Asimismo, se creó un grupo equimolar separado (20 µM cada uno) de oligonucleótidos puente para los procedimientos de ensayo en base a las secuencias de los locus genómicos seleccionados.

Se transfirieron 100 µg de microesferas de estreptavidina a los pocillos de una placa de 96 pocillos, y se retiró el sobrenadante. Se añadieron 60 µl de tampón BB2 (Tris 100 mM, pH 8,0, EDTA 10 mM, NaCl<sub>2</sub> 500 mM, formamida al 58 %, Tween-80 al 0,17 %), 10 µl de grupo de oligonucleótidos de secuencia fija 40 nM y 30 µl 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 de nuevo 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, NaCl<sub>2</sub> 50 mM).

10 15 Las microesferas se resuspendieron en 37 µl de mezcla de reacción de unión que consistía en 1X tampón Taq ligasa (Enzymatics, Beverly, MA), 1 U de Taq ligasa y grupo de oligonucleótidos puente 2 µM (dependiendo del formato del ensayo) y se incubó a 37 °C durante una hora. Cuando fue 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 de nuevo sobre el imán y se retiró el sobrenadante. Se repitió una vez el procedimiento de lavado.

20 25 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 35 Los ácidos nucleicos polimerizados y/o unidos se amplificaron usando cebadores de PCR universales 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 consistía en 25 µl de producto de unión eluido más 1X tampón Pfusion, betaína 1 M, dNTP 400 nM cada uno, 1 U de ADN polimerasa termoestable con corrección de errores Pfusion (Thermo Fisher, Waltham MA), y los siguientes pares de cebadores:

40 TAATGATA CGGC GACC ACCGAG ATCTAC ACCGGCGT ATGCGTCGAGA (SEQ ID NO:3) y  
TCAAGCAGA AGACGG CATA CGAGAT XAAACGAC GCGATCAT CGGTCC

45 CCGCAA (SEQ ID NO:4), donde X representa uno de los 96 índices de muestra diferentes usado 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 Tetrad™.

50 Se agruparon 10 µl del producto de 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**

55 Los productos de PCR purificados de cada formato de ensayo se secuencian 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 ~885K lecturas/muestra en todo el experimento, y (en el caso de un experimento usando 96 locus) 9,2K 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:

60 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.

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 un 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.

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

10 Estos resultados se resumen en la tabla 1:

**Tabla 1: Sumario de resultados de formatos de ensayo de unión en tandem**

FORMATO DE ENSAYO	OLIGONUCLEÓTIDO DE SECUENCIA FIJA (1. <sup>º</sup> y/o 2. <sup>º</sup> )	OLIGONUCLEÓTIDO PUENTE USADO	ENZIMA USADA	RENDIMIENTO	INTERVALO FREC. LOC. 80 %	TASA DE ERROR DE 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	T4 lig	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

15 La tabla 1 indica que el ensayo de unión en tandem específico de locus usando un oligonucleótido puente convirtió el ADN molde en un producto dirigido con alto rendimiento (~10 %), con una alta proporción de producto derivado de los locus dirigidos (un 15 % de las lecturas no contenía las estructuras de ensayo esperadas), con un sesgo de locus limitado (un 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 de SNP de un 0,2 %). El ensayo de unión en tandem específico de locus sin el uso de un oligonucleótido puente produjo rendimientos reducidos y un sesgo de locus sustancial, pero todavía produjo datos de genotipado de SNP de alta exactitud. El ensayo de unión en tandem específico de alelo con un oligonucleótido puente produjo rendimientos intermedios en comparación con el ensayo específico de locus usando tanto T4 ligasa como Taq ligasa, pero todavía produjo un sesgo de locus limitado y datos de genotipado de SNP de alta exactitud. El ensayo de unión en tandem específico de alelo sin oligonucleótido puente produjo rendimientos reducidos y un sesgo de locus sustancial, pero todavía produjo 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 dirigido con un alto rendimiento (12-18 %), con una alta proporción de producto derivado de los locus dirigidos (~76 % de las lecturas contenían las estructuras de ensayo esperadas) y con un sesgo de locus limitado (un 80 % de los locus se encontraban dentro de un intervalo de concentración de 2-3 veces). La figura 5 ilustra el rendimiento de 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.

#### 25 **Ejemplo 6: detección de aneuploidía en muestras de pacientes de mujeres embarazadas**

30 Los sistemas de ensayo de la divulgación se usaron en 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.

Los elementos usados en los ensayos de detección de aneuploidía se ilustran en la FIG. 6. El ADNlc 601 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 a cada locus seleccionado para crear productos de unión para amplificación y detección. El

5 oligonucleótido de secuencia fija izquierdo (o primero) comprendía una región complementaria a un locus seleccionado 609 y una primera región de cebador universal 611. El oligonucleótido de secuencia fija derecho (o segundo) 605 comprendía una segunda región complementaria al locus seleccionado 613 y una segunda región de cebador universal 615. Los oligonucleótidos puente 607 usados se diseñaron de modo que cada uno se hibridara a las regiones puente de dos o más locus seleccionados usados en el ensayo de detección de 10 aneuploidía. Cuando los oligonucleótidos de secuencia fija 603, 605 y el oligonucleótido puente 607 se hibridaron a la región complementaria en el ADNlc 601, sus extremos formaron dos mellas. Tras la unión de los oligonucleótidos hibridados al ADNlc, se creó un producto de unión para cada locus seleccionado que comprendía 15 603, 605 y 607 y que se usó como molde para los cebadores de amplificación 619, 621.

15 A continuación, se usaron dos cebadores de amplificación 619, 621 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 617 para identificar la muestra particular de la que se obtuvo el locus 20 en el ensayo multiplexado. La amplificación con 96 cebadores de amplificación derechos 629 distintos permitió el agrupamiento y la secuenciación simultánea de 96 productos de amplificación diferentes en un único carril.

25 Los cebadores de amplificación 619, 621 también contenían una secuencia de grupo izquierda 623 (TAATGATAACGGCGACCACCGA)SEQ ID NO:7) y una secuencia de grupo derecha 625 (ATCTCGTATGCCGTCTCTGCTTGA) SEQ ID NO:8) que soportó la amplificación de grupos para secuenciación usando el sistema Illumina HiSeq™ 2000 (Illumina, San Diego, CA). Se usó un cebador de secuenciación 627 que comprendía la primera secuencia de cebador universal para determinar la secuencia del producto de amplificación, y se usó un segundo cebador de secuenciación 629 para determinar el índice de muestra 617 del producto de amplificación.

30 En resumen, se extrajeron aproximadamente 10 ml de sangre periférica de cada paciente en un tubo BCT (Streck, Omaha, NE), que se envió por medio de mensajería nocturna a Tandem Diagnostics. El plasma se aisló de los tubos BCT dentro de las 72 h siguientes a la extracción de sangre por 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 retirar cualquier célula restante. Se aisló 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 35 posterior se realizó en lotes multiplexados de hasta 96 muestras de ADNlc de pacientes por procedimiento de sistema de matriz.

40 El ADNlc aislado de las muestras maternas en cada pocillo se biotiniló, se precipitó y se resuspendió en 30 ul de TE como en el ejemplo 3 anterior. El ADN molde biotinilado se mezcló con 100 ug de microesferas magnéticas recubiertas con estreptavidina MyOneC1 (Life Technologies, Carlsbad, CA), 60 µl de tampón BB2 (Tris 100 mM, pH 8,0, EDTA 10 mM, NaCl<sub>2</sub> 500 mM, formamida al 58 %, Tween-80 al 0,17 %) y 10 µl de oligonucleótidos de secuencia fija izquierdo 603 y derecho 605 agrupados 40 nM. La mezcla se calentó hasta 70 °C y se enfrió 2 horas. A continuación, las microesferas se inmovilizaron magnéticamente en el lateral del pocillo, se lavaron dos veces 45 con 50 ul de BB2 al 60 % (v/v con H<sub>2</sub>O), se lavaron dos veces más con 50 ul de tampón de lavado (Tris 10 mM, pH 8,0, EDTA 1 mM, NaCl<sub>2</sub> 50 mM), y a continuación se resuspendió en una reacción de 50 µl que contenía 1 U de Taq ligasa (Enzymatics, Beverly MA), 1X tampón Taq ligasa (Enzymatics) y 10 uM de un oligonucleótido puente pentámero fosforilado en 5' 607. 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 ul de tampón de lavado y a 50 continuación se resuspendieron en 30 µl de TE.

55 Los productos de unión se eluyeron de las microesferas inmovilizadas por incubación a 95 °C durante 3 minutos. Los productos de unión eluidos se amplificaron por 26 ciclos de PCR en una reacción de 50 ul que contenía 1 U de Pfusion (Finnzymes), betaína 1 M, 1X tampón Pfusion y cebadores de amplificación izquierdo y derecho 400 nM (619, 621 respectivamente). El cebador derecho contenía un índice de muestra de 7 bases (617) que permitía la secuenciación multiplexada de 96 muestras en HiSeq2000 (Illumina, San Diego, CA). La secuencia del oligonucleótido de secuencia fija izquierdo fue:

**TAATGATAACGGCGACCACCGAGATCTACACCGGGCGITATGCGTCGAGA**

**C**

60  
(SEQ ID NO: 5)

Y la secuencia del oligonucleótido de secuencia fija derecho fue:

TCAAGCAGAAGACGGCATACGAGATNNNNNNAAACGACGCGATCATCG  
GTCCCCGCAAT (SEQ ID NO: 6)

- 5 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 grupos en una cubeta de lectura de grupos 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 623 y se obtuvo una lectura separada de 8 bases de la secuencia específica de muestra a partir del segundo cebador de secuencia 625. Se recogió un promedio de 903K lecturas brutas por muestra. Se asignó un promedio de 876K (97 %) lecturas a las estructuras de ensayo esperadas.
- 10 15 La FIG. 7 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 procesaron 96 muestras diferentes en esta tanda particular, pero después se excluyeron seis muestras de este conjunto analítico por no cumplir los umbrales de control de calidad de muestra.
- 20 25 En base a las lecturas producidas en el ensayo se calculó una media truncada para cada cromosoma 18 y cromosoma 21 en las muestras. La media truncada se calculó retirando un 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.
- 30 35 40 45 50 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. 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 de RMA y un refinado de mediana (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 efectos cromosómicos, efectos de locus, efectos de muestra y residuales. Los efectos cromosómicos estimados se establecieron en un valor común, por ejemplo 0, y se calculó  $2^{\text{efecto cromosómico} + \text{efecto de muestra} + \text{residual}}$  para cada locus para crear recuentos normalizados. Las puntuaciones z se ajustaron a escala por censura iterativa de modo que tuvieran una media de 0 y una desviación estándar de 1. 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. 8 y 9, respectivamente. Las muestras normales se muestran como rombos gris oscuro y las muestras con una trisomía se muestran como rombos gris claro. 179/180 (99,4 %) muestras normales (rombos gris oscuro) tuvieron puntuaciones z <3; una muestra normal tuvo 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 tuvieron puntuaciones z para la proporción cromosómica >3. La puntuación z media para T18 fue de 8,5 y el intervalo fue de 5,8-10,9. La puntuación z media para T21 fue de 11,5 y el intervalo fue de 6,1-19,8. Los datos proporcionados en la FIG. 7 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. 10 y 11, respectivamente. Las muestras normales se muestran como rombos gris oscuro y las muestras con una trisomía se muestran como rombos 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 5,1-19,8. La puntuación z media para T21 fue de 11,4 y el intervalo fue de 3,4-21,8. Además de la detección de aneuploidía en estas cohortes, también se determinaron polimorfismos específicos para estas muestras en un mismo ensayo. Se obtuvo información específica para locus individuales, así como información polimórfica más general, tal como el número de locus en los que el locus fetal presentó un polimorfismo

mononucleotídico en un alelo diferente de los polimorfismos mononucleotídicos en el locus materno (FIG. 7, n.º Locus DifPoli). Esta determinación también identificó la presencia de polimorfismos específicos en el genoma fetal. Por ejemplo, el estado de tres polimorfismos ejemplares se determinó usando una combinación de oligonucleótidos puente que se diseñaron para unirse tanto al residuo **A** como al **T** en las siguientes regiones polimórficas ejemplares:

**Tabla 2: polimorfismos individuales consultados usando el ensayo de la invención**

Cromosoma	Localización	RSID
Ch01	01_010303942	rs11582123
	<b>TTTACATGTCTTGGGCATTAGGT[A/T]GAGTGAAATCTAGGCCTTGCAAATC</b> (SEQ ID NO:7)	
Ch03	03_098690592	rs2470750
	<b>TTGTGTAACGTTAACCTCAGGGACCA[A/T]GAGATGTACTTAGTATTAAATTGCC</b> (SEQ ID NO:8)	
Ch04	04_055495793	rs6815910
	<b>GGAAGAAGTGCAGTGTAGTAGACAAC[A/T]CTGGCATTGTGTTTGTGAAGTGGG</b> (SEQ ID NO:9)	

10 **Tabla 3: Estado materno y fetal previsto para SNP rs11582123.**

Muestra	SNP	Recuentos A	Recuentos T	Estado fetal previsto	Estado materno previsto
1	rs11582123	294	26	A/T	A/A
2		181	134	A/A	A/T
4		34	330	A/T	T/T
5		241	21	A/T	A/A
6		166	134	A/T	A/T
7		137	182	T/T	A/T
8		199	135	A/A	A/T
9		0	267	T/T	T/T
10		0	284	T/T	T/T
11		151	154	A/T	A/T
12		294	1	A/T	A/A
13		131	114	A/A	A/T
14		118	159	T/T	A/T
15		257	10	A/T	A/A
16		309	31	A/T	A/A
17		20	289	A/T	T/A
18		137	166	T/T	A/T
19		138	143	A/T	A/T
20		24	242	A/T	T/T
21		140	161	A/T	A/T
22		159	118	A/A	A/T
23		119	122	A/T	A/T
24		0	250	T/T	T/T
25		0	285	T/T	T/T
26		120	130	A/T	A/T
28		134	113	A/A	A/T
29		109	118	A/T	A/T
30		0	271	T/T	T/T
31		148	139	A/T	A/T
32		29	253	A/T	T/T
33		0	304	T/T	T/T
34		1	278	T/T	T/T
35		103	188	T/T	A/T
36		18	269	A/T	T/T
37		279	34	A/T	A/A
38		0	250	T/T	T/T
39		0	263	T/T	T/T
40		136	142	A/T	A/T
41		147	145	A/T	A/T
42		15	270	A/T	T/T
43		44	222	A/T	T/T

# ES 2 913 402 T3

Muestra	SNP	Recuentos A	Recuentos T	Estado fetal previsto	Estado materno previsto
44		140	159	T/T	A/T
45		0	259	T/T	T/T
46		1	304	T/T	T/T
47		162	127	A/A	A/T
48		0	335	T/T	T/T
49		1	247	T/T	T/T
50		153	154	A/T	A/T
51		118	182	T/T	A/T
52		145	134	A/T	A/T
53		146	132	A/T	A/T
54		7	319	A/T	T/T
55		152	174	T/T	A/T
56		1	319	T/T	T/T
57		147	150	A/T	A/T
58		136	157	T/T	A/T
59		83	162	A/A	T/T
60		14	215	A/T	T/T
61		157	121	A/A	A/T
62		281	0	A/A	A/A
63		0	260	T/T	T/T
64		0	305	T/T	T/T
65		18	252	A/T	T/T
66		0	303	T/T	T/T
67		99	161	T/T	A/T
68		141	127	A/T	A/T
69		0	237	T/T	T/T
70		0	315	T/T	T/T
71		132	139	A/T	A/T
73		112	120	A/T	A/T
75		1	268	T/T	T/T
76		166	123	A/A	A/T
78		0	245	T/T	T/T
79		12	264	A/T	T/T
80		15	281	A/T	T/T
81		21	269	A/T	T/T
82		108	160	T/T	A/T
83		106	144	T/T	A/T
84		137	135	A/T	A/T
85		115	151	T/T	A/T
86		0	262	T/T	T/T
87		0	269	T/T	T/T
89		0	284	T/T	T/T
90		0	261	T/T	T/T
91		143	137	A/T	A/T
92		0	308	T/T	T/T
93		1	256	T/T	T/T
94		158	105	A/A	A/T
95		149	103	A/A	A/T

**Tabla 4: Estado materno y fetal previsto para SNP rs2470750.**

Muestra	SNP	Recuentos A	Recuentos T	Estado fetal previsto	Estado materno previsto
1	rs2470750	243	15	A/T	A/A
2		265	0	A/A	A/A
4		170	107	A/A	A/T
5		196	30	A/T	A/A
6		141	144	A/T	A/T
7		139	137	A/T	A/T
8		272	0	A/A	A/A
9		218	0	A/A	A/A
10		216	0	A/A	A/A
11		228	0	A/A	A/A
12		6	224	A/T	T/T

# ES 2 913 402 T3

Muestra	SNP	Recuentos A	Recuentos T	Estado fetal previsto	Estado materno previsto
13		126	93	A/A	A/T
14		125	123	A/T	A/T
15		234	20	A/T	A/A
16		147	113	A/A	A/T
17		235	2	A/T	A/A
18		129	142	A/T	A/T
19		132	114	A/A	A/T
20		214	0	A/A	A/A
21		1	245	T/T	T/T
22		141	111	A/A	A/T
23		135	128	A/A	A/T
24		121	160	T/T	A/T
25		209	21	A/T	A/A
26		0	239	T/T	T/T
27		203	4	A/T	A/A
28		101	115	T/T	A/T
29		212	10	A/T	A/A
30		86	101	T/T	A/T
31		118	116	A/T	A/T
32		135	121	A/A	A/T
33		111	128	T/T	A/T
34		120	118	A/T	A/T
35		246	0	A/A	A/A
36		113	115	A/T	A/T
37		96	126	T/T	A/T
38		107	88	A/A	A/T
39		241	0	A/A	A/A
40		116	118	A/T	A/T
41		135	89	A/A	A/T
42		129	85	A/A	A/T
43		0	205	T/T	T/T
44		138	88	A/A	A/T
45		129	86	A/A	A/T
46		108	123	T/T	A/T
47		14	246	A/T	T/T
48		129	148	T/T	A/T
49		108	110	A/T	A/T
50		120	124	A/T	A/T
51		212	22	A/T	A/T
52		237	0	A/A	A/A
53		104	147	T/T	A/T
54		134	126	A/T	A/T
55		128	82	A/A	A/T
56		225	5	A/T	A/A
57		213	11	A/T	A/A
58		125	116	A/T	A/T
59		226	1	A/A	A/A
60		103	119	T/T	A/T
61		84	91	T/T	A/T
62		130	104	A/A	A/T
63		251	0	A/A	A/A
64		243	0	A/A	A/A
65		127	115	A/A	A/T
66		113	104	A/A	A/T
67		26	190	A/T	T/T
68		80	83	A/T	A/T
69		122	132	T/T	A/T
70		0	235	T/T	T/T
71		90	123	T/T	A/T
73		174	0	A/A	A/A
75		0	233	T/T	T/T
76		220	0	A/A	A/A
78		115	115	A/T	A/T
79		112	144	T/T	A/T

# ES 2 913 402 T3

Muestra	SNP	Recuentos A	Recuentos T	Estado fetal previsto	Estado materno previsto
80		10	248	A/T	T/T
81		241	0	A/A	A/A
82		228	0	A/A	A/A
83		243	16	A/T	A/A
84		133	104	A/A	A/T
85		101	99	A/T	A/T
86		1	209	A/A	A/A
87		224	7	A/T	A/A
89		122	101	A/A	A/T
90		130	89	A/A	A/T
91		128	151	T/T	A/T
92		231	0	A/A	A/A
93		107	118	T/T	A/T
94		93	100	A/T	A/T
95		132	119	A/A	A/T

**Tabla 5: Estado materno y fetal previsto para SNP rs6815910.**

Muestra	SNP	Recuentos A	Recuentos T	Estado fetal previsto	Estado materno previsto
1	rs6815910	295	32	A/T	A/A
10		133	107	A/A	A/T
11		115	131	T/T	A/T
12		311	10	A/T	A/A
13		18	252	A/T	T/T
14		132	178	T/T	A/T
15		288	0	A/A	A/A
16		325	1	A/A	A/A
17		11	276	A/T	T/T
18		282	0	A/A	A/A
19		131	133	A/T	A/T
2		7	311	A/T	T/T
20		135	116	A/A	A/T
21		121	140	T/T	A/T
22		287	11	A/T	A/A
23		148	146	A/T	A/T
24		185	138	A/A	A/T
25		116	126	T/T	A/T
26		235	0	A/A	A/A
27		288	0	A/A	A/A
28		242	0	A/A	A/A
29		239	12	A/T	A/A
30		235	24	A/T	A/A
31		126	148	T/T	A/T
32		25	256	A/T	T/T
33		286	1	A/A	A/A
34		158	156	A/T	A/T
35		287	0	A/A	A/A
36		118	133	T/T	A/T
37		163	119	A/A	A/T
38		273	10	A/T	A/A
39		132	148	T/T	A/T
4		0	343	T/T	T/T
40		143	177	T/T	A/T
41		0	308	T/T	T/T
42		297	0	A/A	A/A
43		117	130	T/T	A/T
44		296	1	A/A	A/A
45		276	0	A/A	A/A
46		140	134	A/T	A/T
47		158	139	A/A	A/T
48		0	304	T/T	T/T
49		251	13	A/T	A/A

# ES 2 913 402 T3

Muestra	SNP	Recuentos A	Recuentos T	Estado fetal previsto	Estado materno previsto
5		138	115	A/A	A/T
50		142	162	T/T	A/T
51		0	306	T/T	T/T
52		249	21	A/T	A/A
53		111	170	T/T	A/T
54		140	151	A/T	A/T
55		102	217	T/T	A/T
56		315	0	A/A	A/A
57		123	158	T/T	A/T
58		146	168	T/T	A/T
59		226	50	A/T	A/A
6		309	0	A/A	A/A
60		122	133	T/T	A/T
61		240	28	A/T	A/A
62		132	124	A/T	A/T
63		291	9	A/T	A/A
64		0	304	T/T	T/T
65		273	0	A/A	A/A
66		154	139	A/T	A/T
67		145	153	A/T	A/T
68		110	163	T/T	A/T
69		131	134	A/T	A/T
7		186	127	A/A	A/T
70		167	163	A/T	A/T
71		238	26	A/T	A/A
73		18	244	A/T	T/T
75		130	129	A/T	A/T
76		133	113	A/A	A/T
78		237	2	A/T	A/A
79		0	278	T/T	T/T
8		192	159	A/A	T/T
80		153	131	A/A	T/T
81		25	229	A/T	T/T
82		0	256	T/T	T/T
83		152	142	A/T	A/T
84		290	2	A/T	A/A
85		270	0	A/A	A/A
86		0	242	T/T	T/T
87		150	134	A/A	A/T
89		169	117	A/A	A/T
9		271	1	A/A	A/A
90		109	144	T/T	A/T
91		261	12	A/T	A/A
92		258	0	A/A	A/A
93		0	309	T/T	T/T
94		116	146	T/T	A/T
95		123	116	A/T	A/T

La localización de estos SNP se indica usando dbSNP versión 132 y GRCH37/UCSC hg 19. Los datos para estos polimorfismos se obtuvieron en el mismo conjunto de datos que los datos de aneuploidía ilustrados en las FIGS. 10 y 11. Por tanto, un ensayo único demostró la capacidad para identificar aneuploidía fetal, diferencias polimórficas entre locus fetales y maternos y la información de SNP real para locus fetales seleccionados en un ensayo único.

## LISTADO DE SECUENCIAS

<110> Ariosa Diagnostics, Inc.  
 5 <120> SISTEMAS DE ENSAYO PARA ANÁLISIS GENÉTICO  
 <130> KSG/FP7350572  
 10 <140> EP  
 <141> 08/08/2011  
 <150> EP11745881.0  
 <151> 08/08/2011  
 15 <150> PCT/US2011/046963  
 <151> 08/08/2011  
 <150> 13/013.732  
 <151> 25/01/2011  
 20 <150> 61/371605  
 <151> 06/08/2010  
 <160> 11  
 25 <170> PatentIn versión 3.5  
 <210> 1  
 <211> 24  
 30 <212> ADN  
 <213> SECUENCIA ARTIFICIAL  
 <220>  
 <223> Oligonucleótido sintético  
 35 <400> 1  
tacaccggcg ttatgcgtcg agac 24  
 <210> 2  
 40 <211> 24  
 <212> ADN  
 <213> ORGANISMO ARTIFICIAL  
 <220>  
 45 <223> Oligonucleótido sintético  
 <400> 2  
attgcgggga ccgatgtatcg cgtc 24  
 50 <210> 3  
 <211> 48  
 <212> ADN  
 <213> SECUENCIA ARTIFICIAL  
 55 <220>  
 <223> Oligonucleótido sintético  
 <400> 3  
taatgatacg ggcgaccacccg agatctacac cggcgttatcg cgtcgaga 48  
 60 <210> 4  
 <211> 53  
 <212> ADN  
 <213> ORGANISMO ARTIFICIAL  
 65

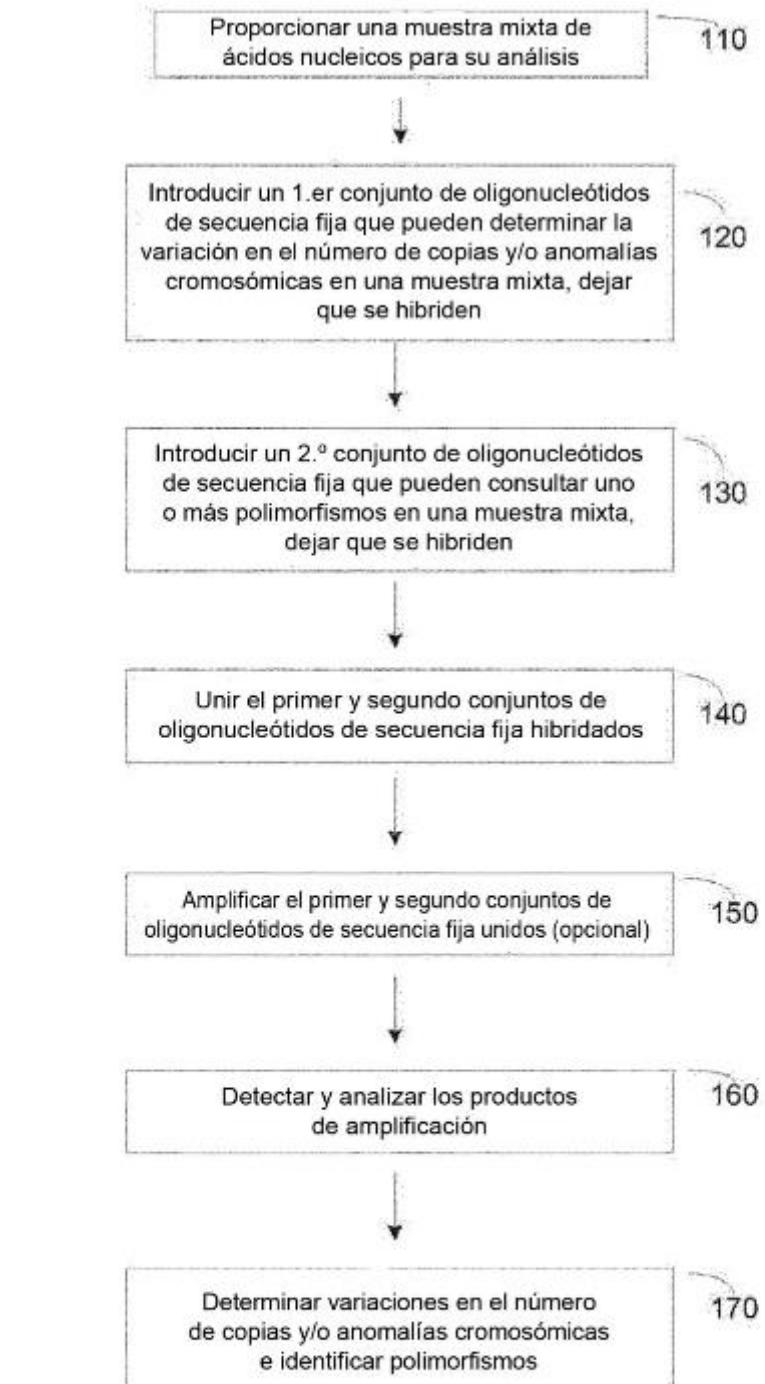
<220>  
 <223> Oligonucleótido sintético  
 5 <220>  
 <221> Índice  
 <222> (26)..(26)  
 <223> n es un índice de secuencia.  
 10 <220>  
 <221> misc\_feature  
 <222> (26)..(26)  
 <223> n es a, c, g o t  
 15 <400> 4  
**tcaaggcagaa gacggcatac gagatnaaac gacgcgatca tgggtccccg caa** 53  
 20 <210> 5  
 <211> 21  
 <212> ADN  
 <213> SECUENCIA ARTIFICIAL  
 <220>  
 <223> Oligonucleótido sintético  
 25 <400> 5  
**taatgatacg gcgaccacccg a** 21  
 <210> 6  
 <211> 25  
 30 <212> ADN  
 <213> SECUENCIA ARTIFICIAL  
 <220>  
 <223> **ATCTCGTATGCCGTCTTCTGCTTGA**  
 35 <400> 6  
**atctcgatcg ccgttttctg cttaga** 25  
 <210> 7  
 40 <211> 48  
 <212> ADN  
 <213> SECUENCIA ARTIFICIAL  
 <220>  
 45 <223> Oligonucleótido sintético  
 <400> 7  
**aatgatacgg cgaccacccga gatctacacc ggccgttatgc gtcgagac** 48  
 50 <210> 8  
 <211> 60  
 <212> ADN  
 <213> SECUENCIA ARTIFICIAL  
 55 <220>  
 <223> Oligonucleótido sintético  
 <220>  
 <221> n  
 60 <222> (26)..(32)  
 <223> n = G, A, T o C  
 <220>  
 <221> misc\_feature  
 65 <222> (26)..(32)

	<223> n es a, c, g o t	
5	<400> 8 tcaagcagaa gacggcatac gagatnnnnn nnaaacgacg cgatcategg tccccgcaat	60
	<210> 9	
	<211> 52	
	<212> ADN	
10	<213> SECUENCIA ARTIFICIAL	
	<220>	
	<223> Oligonucleótido sintético	
15	<220>	
	<221> w	
	<222> (27)..(27)	
	<223> w = A o T	
20	<400> 9 tttacatgtc tttgggcatt ttaggtwgag taaaatctag gccttgcaaa tc	52
	<210> 10	
	<211> 52	
25	<212> ADN	
	<213> SECUENCIA ARTIFICIAL	
	<220>	
	<223> Oligonucleótido sintético	
30	<220>	
	<221> w	
	<222> (27)..(27)	
	<223> w = A o T	
35	<400> 10 tttgtgttaacg ttaacctcaq ggaccawgag atgtacttag tattaatttg cc	52
	<210> 11	
	<211> 52	
40	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Oligonucleótido sintético	
45	<220>	
	<221> w	
	<222> (27)..(27)	
	<223> w = A o T	
50	<400> 11 ggaagaagtg cagtgttagta gacaacwctg gcattgtgtt ttgtgaactg gg	52

## REIVINDICACIONES

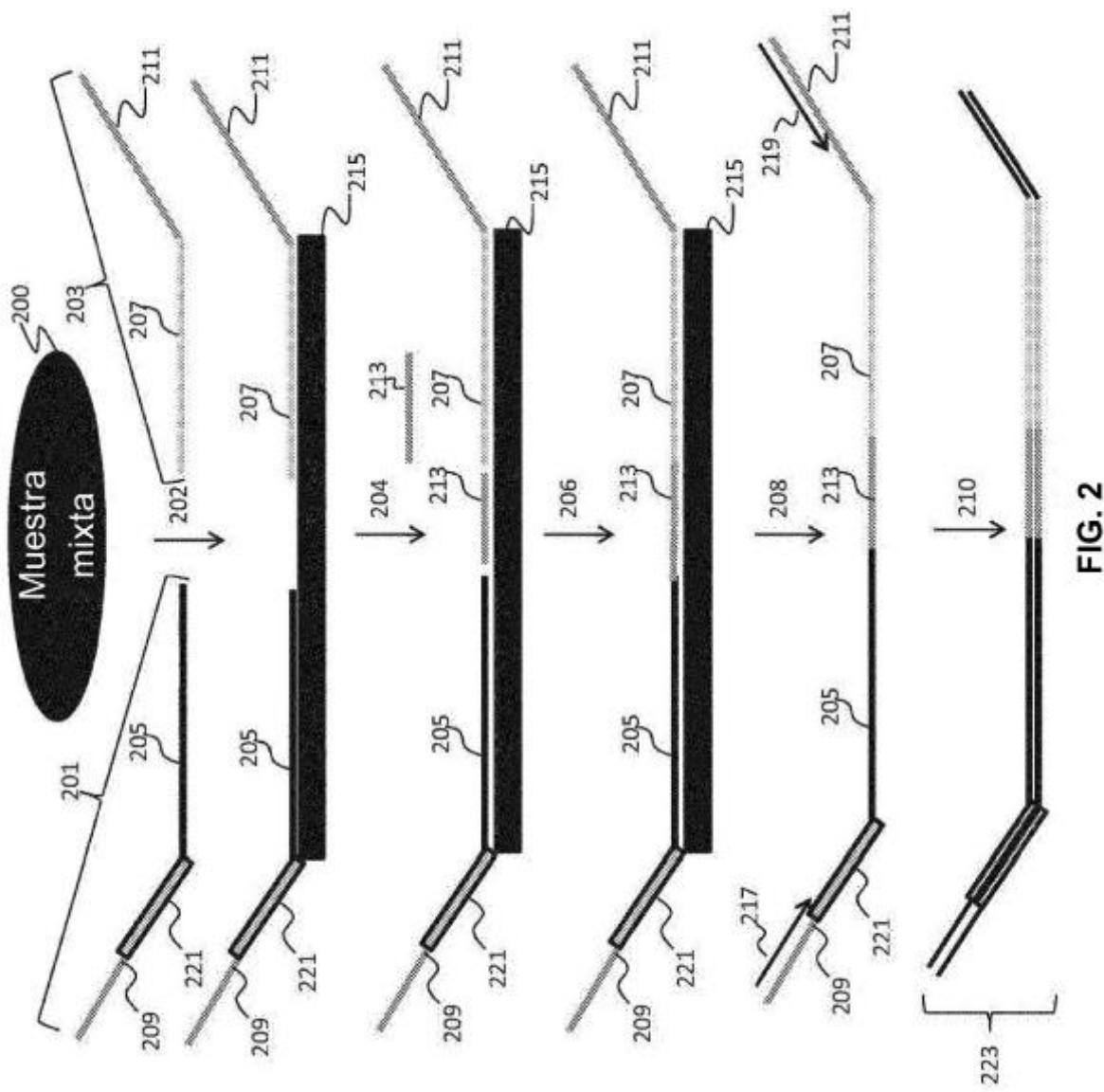
1. Un procedimiento de ensayo único para la detección de la presencia o ausencia de variación en el número de copias (VNC) de una región genómica en una muestra mixta usando un ensayo único, en el que la muestra mixta es de cualquier fluido materno que comprende ácidos nucleicos libres circulantes tanto maternos como fetales,  
 5 comprendiendo el procedimiento de ensayo las etapas de:
- introducir un primer conjunto de oligonucleótidos de secuencia fija en una muestra mixta en condiciones que permiten que los oligonucleótidos de secuencia fija se hibriden específicamente a regiones complementarias en 10 24 o más locus en un primer cromosoma, en el que al menos uno de los oligonucleótidos de secuencia fija comprende regiones de cebador universal;
- 15 introducir un segundo conjunto de oligonucleótidos de secuencia fija en la muestra mixta en condiciones que permiten que los oligonucleótidos de secuencia fija se hibriden específicamente a regiones complementarias en 24 o más locus en un segundo cromosoma, en el que al menos uno de los oligonucleótidos de secuencia fija comprende regiones de cebador universal;
- 20 introducir uno o más oligonucleótidos puente en condiciones que permiten que los oligonucleótidos puente se hibriden específicamente a regiones en los locus entre las regiones complementarias a los oligonucleótidos de secuencia fija de los conjuntos;
- 25 unir los oligonucleótidos hibridados para crear productos de unión contiguos complementarios a los locus del primer cromosoma y los locus del segundo cromosoma;
- 30 amplificar los productos de unión contiguos usando las regiones de cebador universal para crear productos de amplificación; y  
 detectar y cuantificar los productos de amplificación por secuenciación de alto rendimiento, en el que los locus seleccionados se miden cada uno en promedio al menos 100 veces;
- 35 en el que la detección de los productos de amplificación se correlaciona con el número de copias de los locus del primer y segundo cromosomas en la muestra mixta.
3. El procedimiento de ensayo único de la reivindicación 1, en el que la muestra mixta es de plasma materno, suero materno o sangre materna.
- 40 4. El procedimiento de ensayo único de una cualquiera de las reivindicaciones precedentes, en el que uno o ambos conjuntos de oligonucleótidos de secuencia fija comprenden una sonda precírculo.
- 45 5. El procedimiento de ensayo único de una cualquiera de las reivindicaciones precedentes, en el que los productos de amplificación se aislan antes de la detección.
6. El procedimiento de ensayo único de una cualquiera de las reivindicaciones precedentes, en el que los productos de amplificación se aislan como moléculas individuales antes de la detección.
- 50 7. El procedimiento de ensayo único de la reivindicación 6, en el que los productos de amplificación aislados individuales se amplifican además para crear copias idénticas de todos o una porción de los productos de amplificación individuales antes de la detección.
- 55 8. El procedimiento de ensayo único de la reivindicación 6, en el que los productos de amplificación aislados individuales se amplifican además para crear copias idénticas de moléculas complementarias a todos o una porción de los productos de amplificación individuales antes de la detección.
9. El procedimiento de ensayo único de una cualquiera de las reivindicaciones 1 a 8, en el que la VNC se limita a una región subcromosómica.
- 60 10. El procedimiento de ensayo único de una cualquiera de las reivindicaciones 1 a 8, en el que la VNC se limita a un único gen.
- 65 11. El procedimiento de ensayo único de una cualquiera de las reivindicaciones precedentes, en el que la VNC es una aneuploidía cromosómica.

12. El procedimiento de ensayo único de la reivindicación 11, en el que la aneuploidía cromosómica es trisomía 21, trisomía 18, trisomía 13, síndrome triple X, síndrome XYY, trisomía 8, trisomía 16 o síndrome de Turner.
- 5 13. El procedimiento de ensayo único de una cualquiera de las reivindicaciones 1 a 10, en el que la VNC comprende una delección cromosómica.
14. El procedimiento de ensayo único de la reivindicación 13, en el que la delección cromosómica es una delección 22q11.



100

**FIGURA 1**

**FIG. 2**

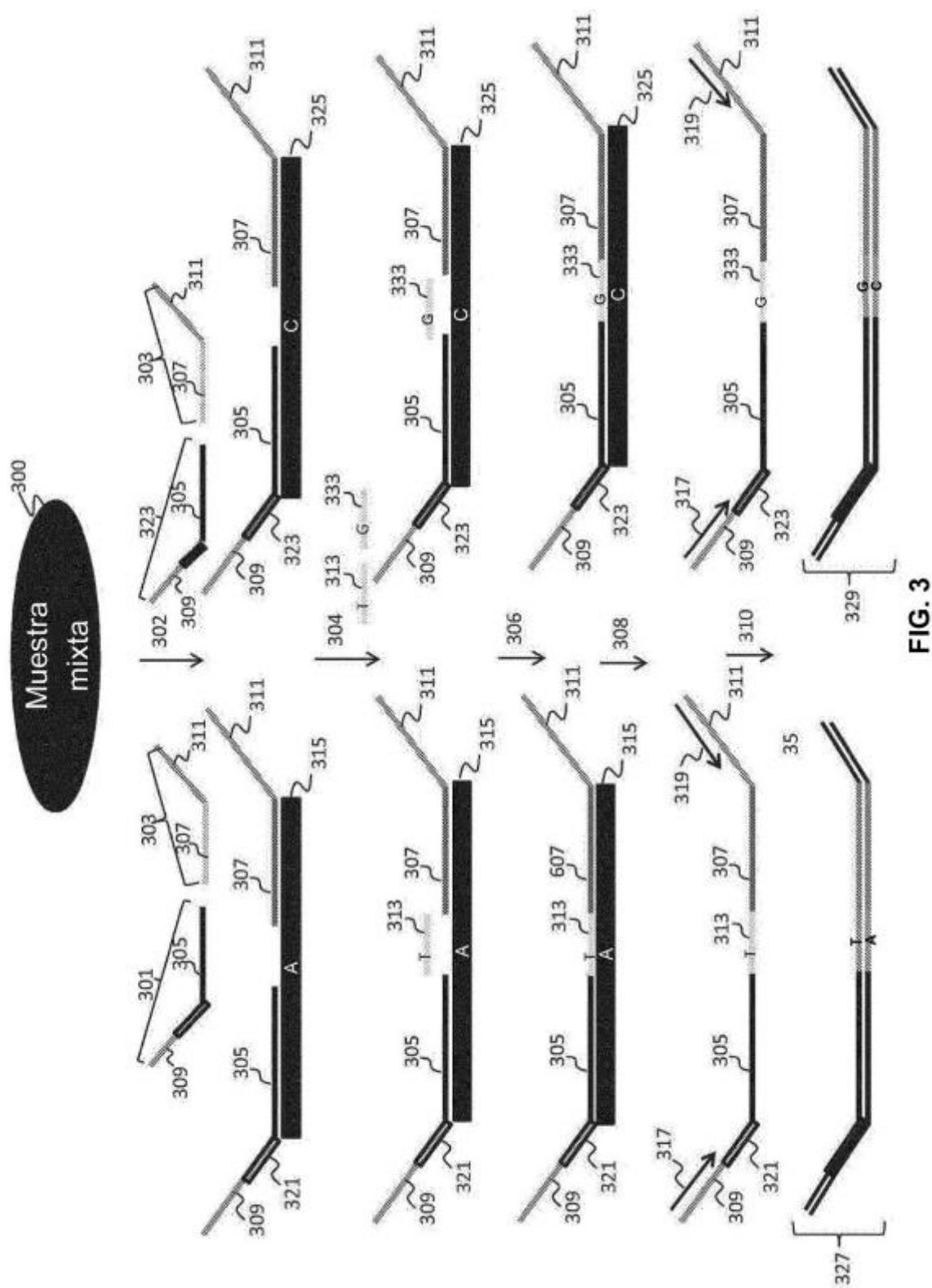
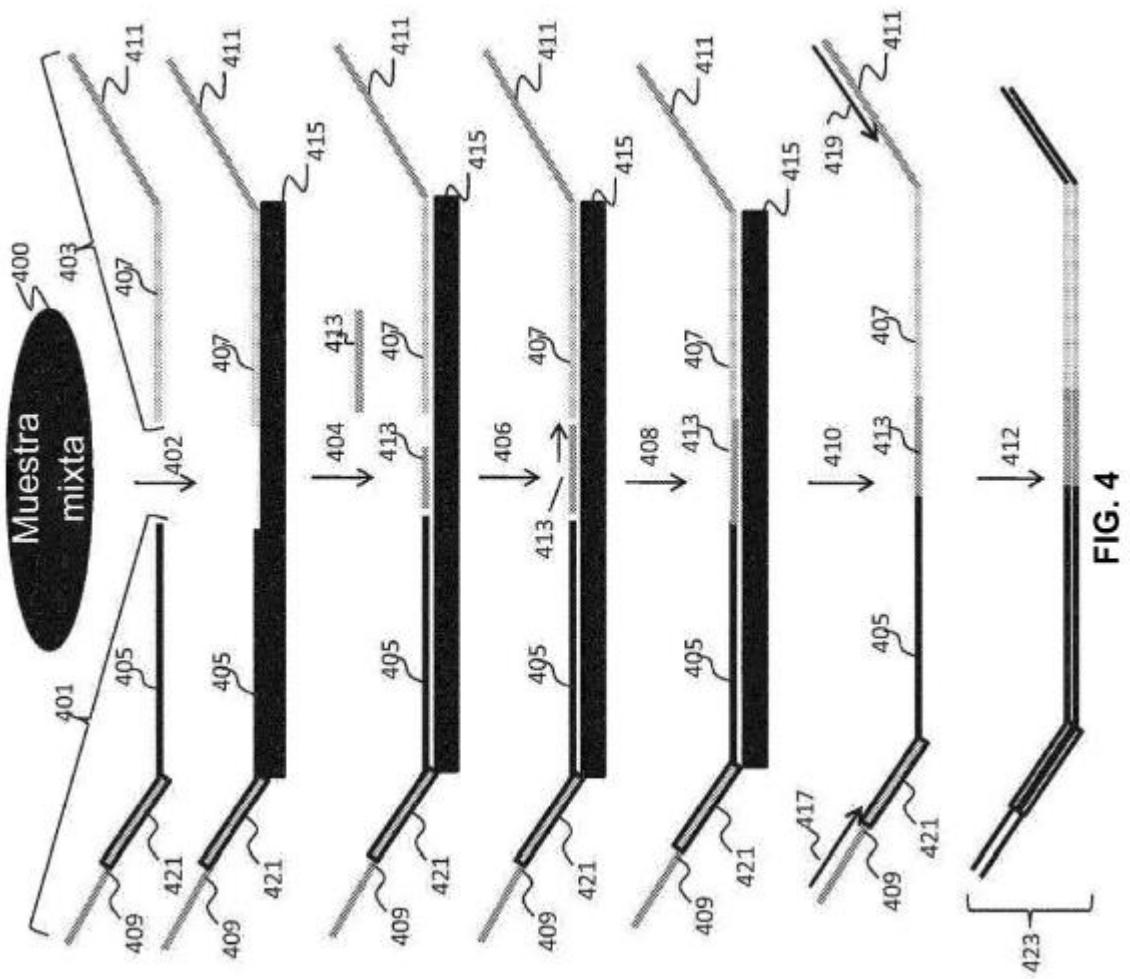


FIG. 3

**FIG. 4**

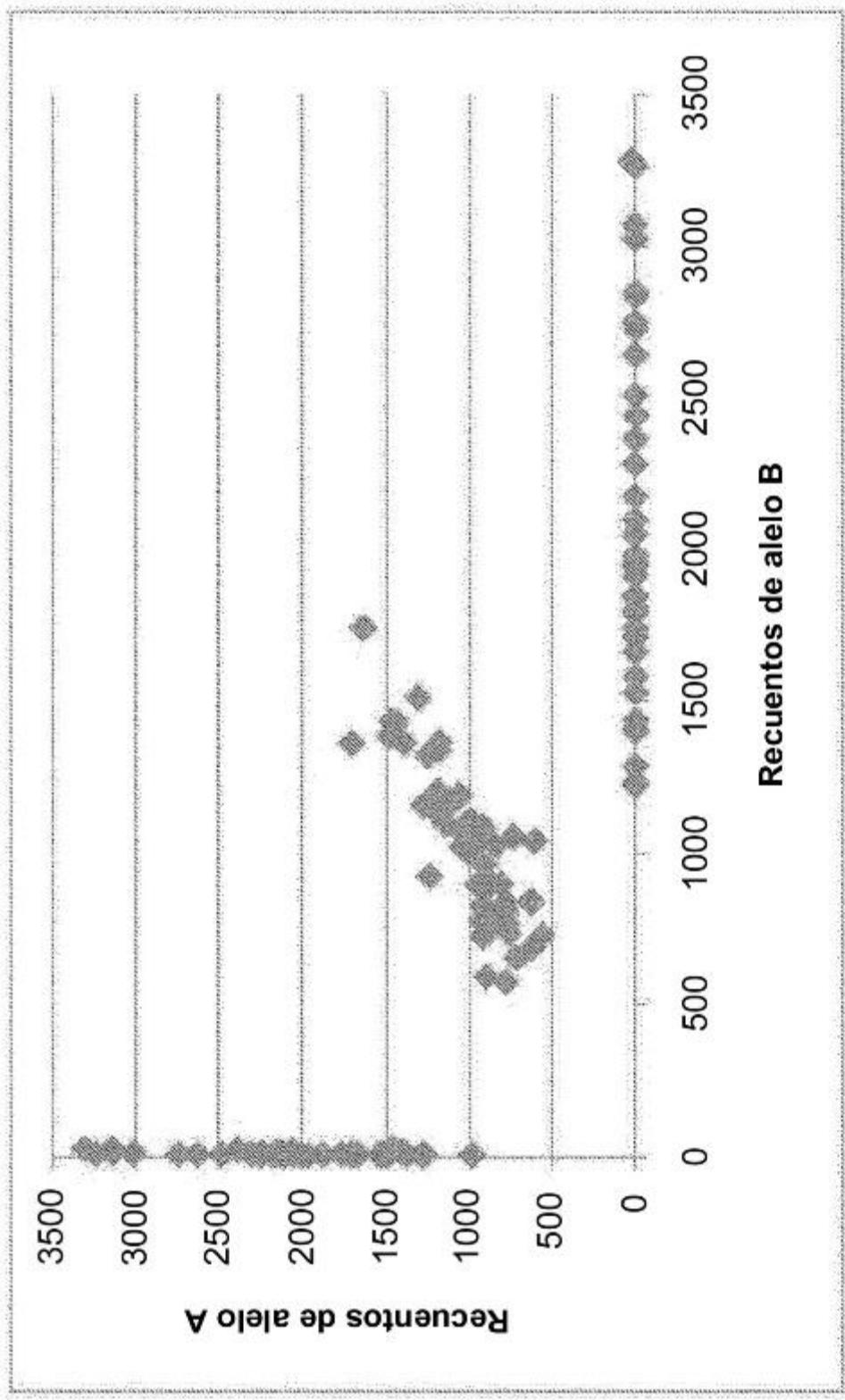


FIG. 5

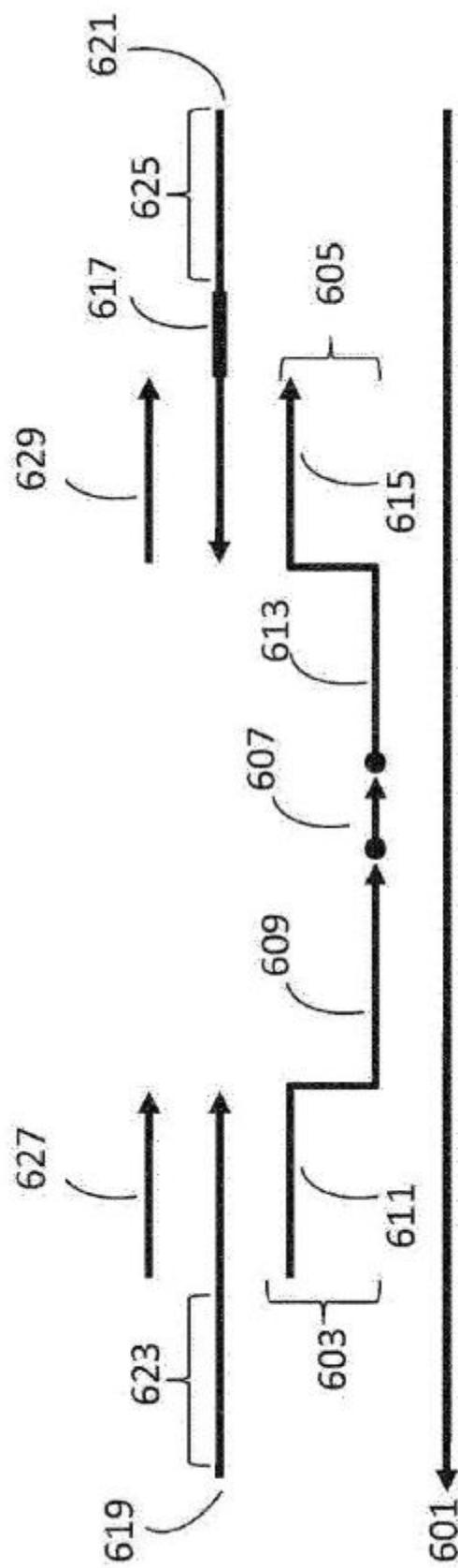


FIG. 6

# ES 2 913 402 T3

Muestra	Marca de muestra	Sexo fetal	Equiv. genoma materno	Cariotipo fetal	Cariotipo madre	Edad Gest madre (día)	Nº Locus DifPoli	% fetal	Media18	Media21	Prop18	Prop21	ZSat18	ZSat21
1	TGTCCCC	M	2293	46XX	46XY	45	110	38	0,210138	1005,426	999,4412	0,501493	0,498507	0,48436
2	ATCCTCC	F	3593	46XX	46XX	37	131	45	0,096075	996,6228	998,7823	0,499459	0,500541	-0,84491
4	GCCAACC	F	1469	46XX	46XX	34	121	26	0,235051	991,0624	997,5193	0,498376	0,501624	-1,55236
5	CCACGAC	F	1296	46XX	46XX	36	129	36	0,140539	1010,048	1004,468	0,501385	0,498615	0,413907
6	TCCTCGT	F	2110	46XX	46XX	33	136	24	0,049392	994,7681	1000,632	0,498531	0,501469	-1,45169
7	CCTTGCT	F	2772	46XX	46XX	43	111	42	0,095395	997,3443	986,4745	0,50274	0,49776	-1,29935
8	TCACTCT	M	2506	46XX	46XY	34	126	32	0,158725	1004,807	996,2886	0,502128	0,497872	0,899842
9	CATCTTC	M	3996	46XX	46XY	36	121	34	0,110817	1005,714	1002,283	0,500854	0,499146	0,067201
10	TTCGGCG	F	4032	46XX	46XX	42	84	29	0,159776	999,7097	1000,399	0,499828	0,500172	-0,60393
11	GTGCCCC	F	4212	46XX	46XX	36	175	39	0,119147	1004,796	995,1465	0,502412	0,497588	-1,085457
12	GCCCTG	F	3816	46XX	46XX	36	111	27	0,085664	1002,909	999,3516	0,500888	0,499112	0,089291
13	CTTACAC	F	2160	46XX	46XX	41	123	40	0,117371	1002,012	995,0877	0,501734	0,498266	0,641893
14	TCGGAGC	F	2790	46XX	46XX	51	125	36	0,119314	1008,515	999,3439	0,502284	0,497716	1,001374
15	CGGCCAT	F	2783	46XX	46XX	51	125	50	0,123598	1005,525	998,8667	0,501661	0,498339	0,594272
16	CCAGGCTG	M	2272	46XX	46XY	34	110	40	0,119257	1010,173	1004,256	0,501468	0,498532	-0,46856
17	AGCACCT	M	2401	46XX	46XY	34	140	43	0,111573	988,7338	986,5967	0,500541	0,499459	-0,137679
18	CGATACC	F	3960	46XX	46XX	41	126	27	0,099872	996,8338	999,2182	0,499403	0,500597	-0,88162
19	CCCGAAT	M	3816	46XX	46XY	37	133	30	0,095735	999,3982	991,0248	0,502103	0,497897	0,883556
20	CCCCAGGG	M	3474	46XX	46XY	38	123	32	0,115184	1006,395	992,8395	0,50339	0,49661	1,724514
21	CAGCACCG	M	1080	46XX	46XY	38	97	28	0,13036	994,7498	1003,364	0,497844	0,502156	-1,90014

FIG. 7

ES 2 913 402 T3

	CITCTATG	M	2002	46XXX	46XX+21	30	123	30	0.08322	994,7936	999,994	0,498697	0,501303	-1,3432	1,343199
	Marca de muestra	Sexo ferial	Equiv. genoma	Cariotipo materno	Cariotipo fetal	Edad madre	Edad fetal (dia)	Locus informativos	% fetal	Media18	Media21	Prop18	Prop21	ZStat18	ZStat21
23	ACTTAC	M	3042	46XXX	46XY	35	130	33	0.138751	1004,97	1003,101	0,500465	0,499535	-0,18713	0,187128
24	CGTATCG	F	1922	46XXX	46XXX	23	159	29	0.254234	1003,218	1006,4	0,499208	0,500792	-1,008668	1,008678
25	AACCATT	M	3708	46XXX	46XY	38	148	31	0,145101	996,0673	995,7268	0,500085	0,499915	-0,435338	0,435383
26	CTGGGC	M	4032	46XXX	46XY	30	126	27	0,077438	1005,191	1003,628	0,500389	0,499611	-0,22694	0,23694
27	TGCCGAG	F	2383	46XXX	46XXX	35	138	32	0,131712	996,445	1002,698	0,498436	0,501564	-1,51345	1,513451
28	ATACCA	F	4104	46XXX	46XXX	23	99	33	0,099894	991,6721	1001,903	0,497434	0,502566	-2,16839	2,168385
29	GAAACCG	F	2437	46XXX	46XXX	33	149	25	0,067606	1004,187	1006,778	0,499356	0,500644	-0,91231	0,912305
30	AAAGGCC	F	2765	46XXX	46XXX	23	126	29	0,144098	1001,308	995,2881	0,501508	0,498492	0,494093	-0,49409
31	TCTAATT	M	2416	46XXX	46XY	23	119	28	0,105371	1006,535	1005,069	0,500364	0,499636	-0,23309	0,25309
32	GTGATC	M	2653	46XXX	46XY	34	148	29	0,164518	1003,738	998,0583	0,501419	0,498581	0,433957	-0,43596
33	GATGCT	M	7488	46XXX	46XY	36	148	21	0,056835	1012,154	1002,257	0,502456	0,497544	1,114241	-1,11424
34	AGTCGT	M	2653	46XXX	46XY	26	140	39	0,21197	1004,959	996,9646	0,501997	0,498003	0,813799	-0,81378
35	AATTCTG	F	3780	46XXX	46XX	35	133	43	0,259642	997,0006	997,6736	0,499831	0,500169	-0,60149	0,601493
36	CTAAGTT	M	8136	46XXX	46XY	26	151	33	0,082773	1007,02	1006,742	0,500069	0,499631	-0,44559	0,445592
37	GGCGGTT	M	2228	46XXX	46XY	37	114	37	0,155551	1004,912	1003,615	0,500323	0,499677	-0,28021	0,280209
38	TATAGGC	F	4248	46XXX	46XXX	27	165	48	0,064569	1014,599	995,8518	0,5044662	0,495338	2,556064	-2,55606
39	TAGGTAC	M	3960	46XXX	46XY	35	124	41	0,167184	997,8251	998,4044	0,499855	0,500145	-0,586068	0,586081
40	CAATTGG	F	4140	46XXX	46XX	24	127	43	0,129398	997,6805	996,6003	0,500271	0,499729	-0,31423	0,314232
41	TCATAAG	M	4320	46XXX	46XY	39	139	29	0,227198	1002,31	991,6654	0,502669	0,497331	1,253298	-1,2533
42	GAACGGT	F	3744	46XXX	46XXX	38	118	39	0,094147	1007,442	994,0998	0,503333	0,496667	1,687191	-1,687191

FIG. 7 (cont.)

ES 2 913 402 T3

FIG. 7 (cont.)

ES 2 913 402 T3

64	TATCAAT	M	8712	46XX	46XY	35	119	33	0.077373	1004.37	996,8203	0.501886	0.498114	0.741595	-0.74159
Muestra	Marca de muestra	Sexo fetal	Equiv. genoma materno	Cariotipo fetal	Cariotipo	Edad madre (años)	Edad Gest	Locus informativos	% fetal	Media18	Media21	Prop18	Prop21	ZStat18	ZStat21
65	TAAGAGT	F	2552	46XX	46XX+13	35	90	41	0.144442	999,7136	988,3962	0.502846	0.497154	1,369064	-1,36906
66	TAAAGAT	F	7668	46XX	46XX+13	30	141	37	0.057703	998,6741	990,8426	0.501968	0.498032	0.795151	-0.79515
67	TCTGCAC	M	3996	46XX	46XY+13	35	161	26	0.21564	1001,317	998,1811	0.500784	0.499216	0.021337	-0.021334
68	GTGGGCT	F	7560	46XX	46XX+21	35	125	28	0.092562	983,1548	1024,755	0.489641	0.510359	-7,2619	7,2619
69	GTGACTT	F	13248	46XX	46XX+21	19	132	30	0.085949	980,027	1020,997	0.489763	0.510237	-7,18225	7,182254
70	GGTCAGC	F	7452	46XX	46XX+21	40	91	31	0.056151	990,5328	1011,076	0.494868	0.505132	-3,845332	3,845324
71	GGCGGAC	M	5508	46XX	46XY+21	18	166	32	0.140102	965,1286	1037,339	0.48197	0.51803	-12,2758	12,27579
73	GGCIAAC	M	5184	46XX	46XY+18	29	147	29	0.157072	1044,147	960,7049	0.52081	0.47919	13,11012	-13,1101
75	GAGTGCG	F	2689	46XX	46XX+18	40	93	24	0.129256	1038,787	958,3378	0.520141	0.479859	12,67306	-12,6731
76	GAATGAC	F	7632	46XX	46XX+18	43	133	24	0.048391	1015,894	982,9916	0.50823	0.49177	4,887983	-4,88798
78	GCTTG	F	7128	46XX	46XX+18	25	128	30	0.06035	1018,978	985,6703	0.508308	0.491692	4,938613	-4,93861
79	GCAGAAC	M	4464	46XX	46XY+18	40	82	31	0.082749	1022,638	974,749	0.511988	0.488012	7,344068	-7,34407
80	ATGGAAT	F	2100	46XX	46XX+18	42	98	39	0.094043	1023,162	977,342	0.511452	0.488548	6,993832	-6,99383
81	ATCAGAT	M	4176	46XX	46XY+21	35	147	33	0.148476	957,3759	1035,729	0.479622	0.520378	-13,8102	13,81016
82	AGTTACT	M	4500	46XX	46XY+21	39	124	19	0.205496	952,9748	1049,25	0.475958	0.524042	-16,205	16,20502
83	AGTAGAC	M	8640	46XX	46XY+21	37	125	30	0.130907	973,3572	1028,835	0.486146	0.513854	-9,54628	9,546281
84	AGGTCTT	M	16092	46XX	46XY+21	43	94	35	0.126545	976,6225	1029,797	0.486749	0.513251	-9,1521	9,152097
85	AGGATAT	M	4212	46XX	46XY+21	40	157	24	0.119821	969,1251	1025,932	0.485763	0.514237	-9,79643	9,796425
86	AGGACGG	M	4644	46XX	46XY+21	33	97	31	0.053737	992,2248	1011,59	0.495168	0.504832	-3,64941	3,649405
87	AGAGATT	M	4320	46XX	46XY+21	31	119	37	0.08022	978,7159	1014,132	0.491114	0.508886	-6,29894	6,298937

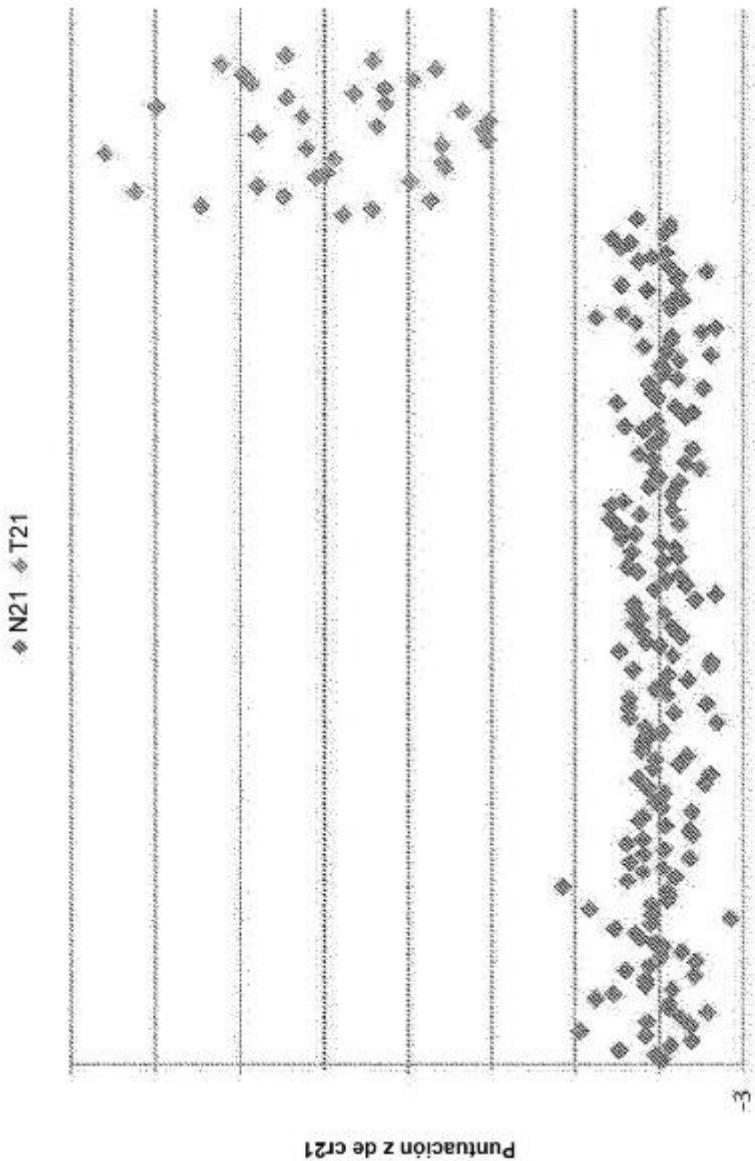
FIG. 7 (cont.)

# ES 2 913 402 T3

Muestra	Sexo	Equiv. genoma	Cariotipo materno	Cariotipo fetal	Edad gest. (dia)	Locus informativos	% feml	Media18	Media21	Prop18	Prop21	ZStat18	ZStat21
89	AGCTGCC F	6768	46XX	46XXX+21	31	98	35	0,126696	963,2256	1024,818	0,484509	0,515491	-10,616 10,61595
90	AGGGCTG M	1375	46XX	46XY+21	36	162	36	0,133202	961,1764	1032,034	0,482225	0,517775	-12,1088 12,10879
91	AGCATGC M	5155	46XX	46XY+21	41	154	37	0,095349	967,2823	1017,613	0,487322	0,512678	-8,77777 8,777767
92	AATGGCTT F	2549	46XX	46XXX+21	41	119	32	0,09505	965,3134	1025,717	0,484831	0,515169	-10,4056 10,4056
93	AAGTAAG F	9216	46XX	46XXX+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

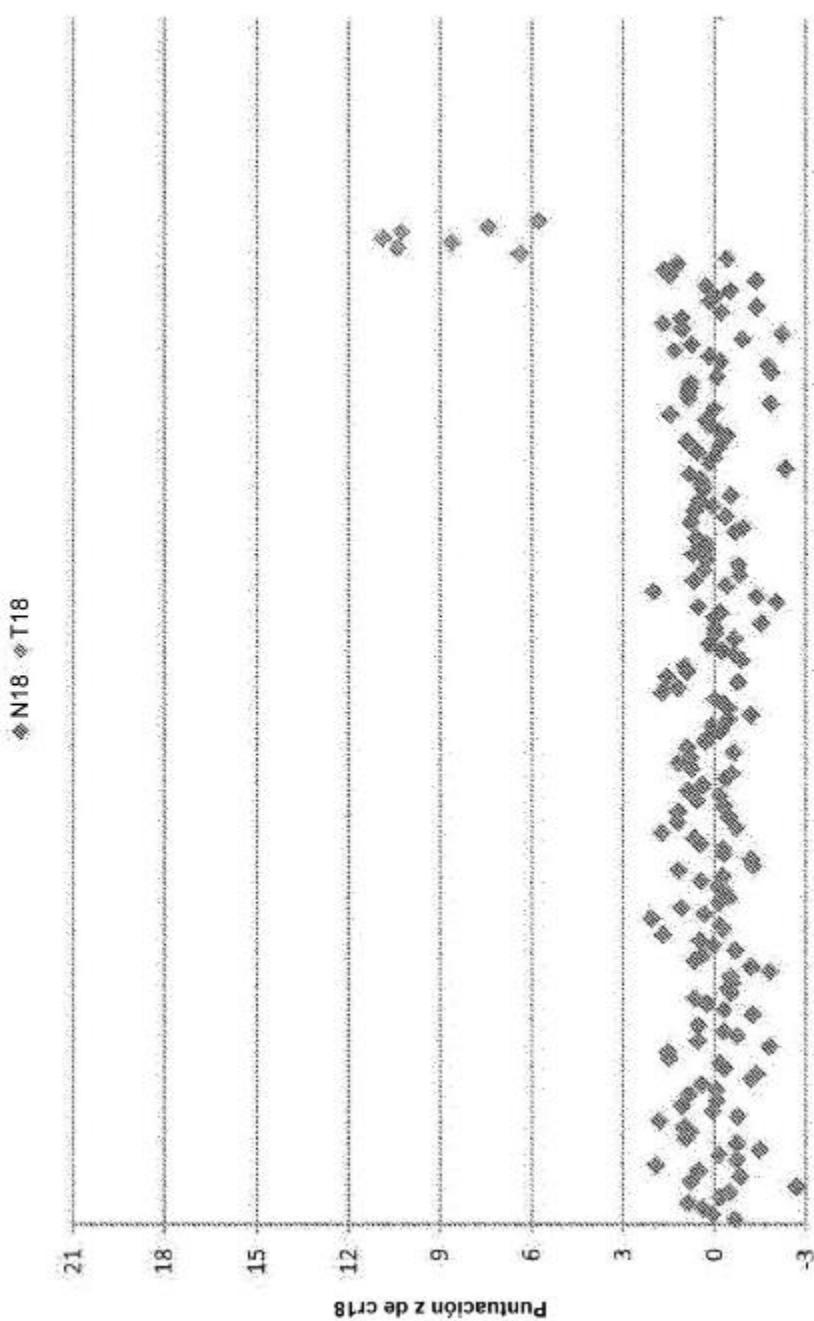
FIG. 7 (cont.)

# ES 2 913 402 T3



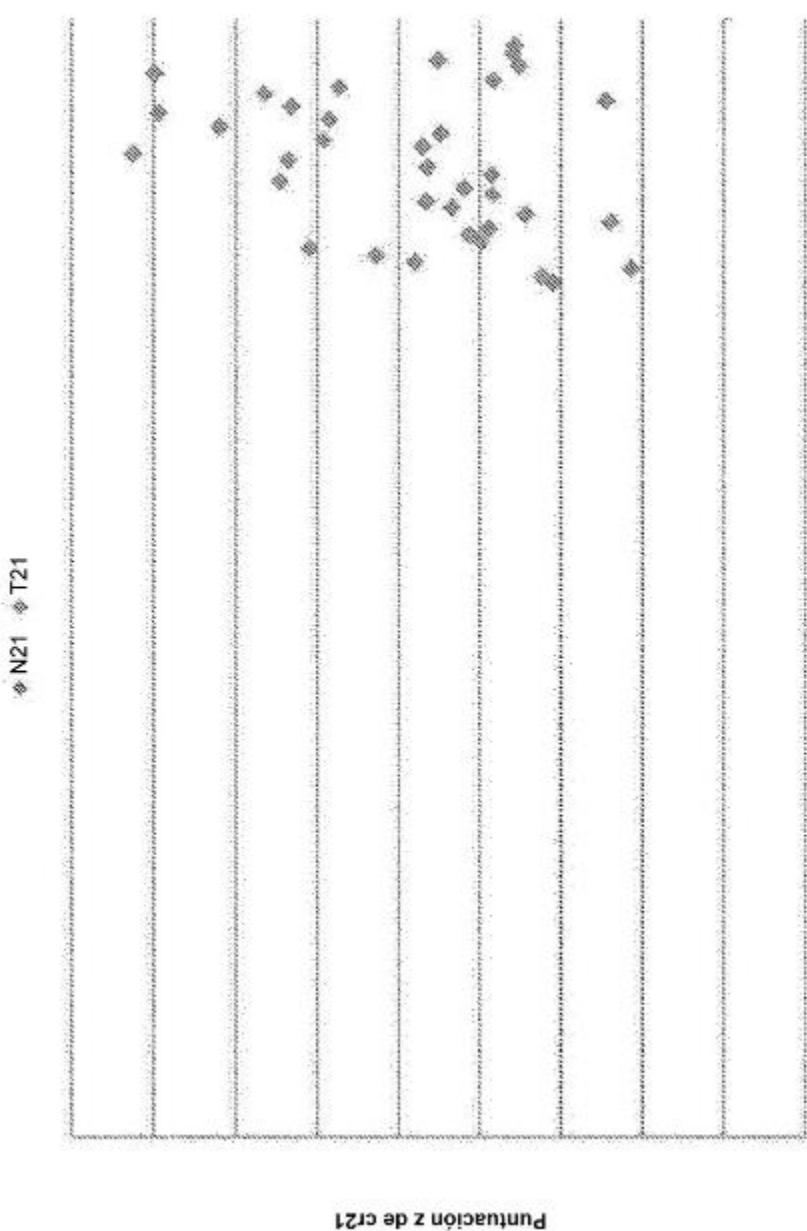
**FIG. 8**

# ES 2 913 402 T3



**FIG. 9**

# ES 2 913 402 T3



**FIG. 10**

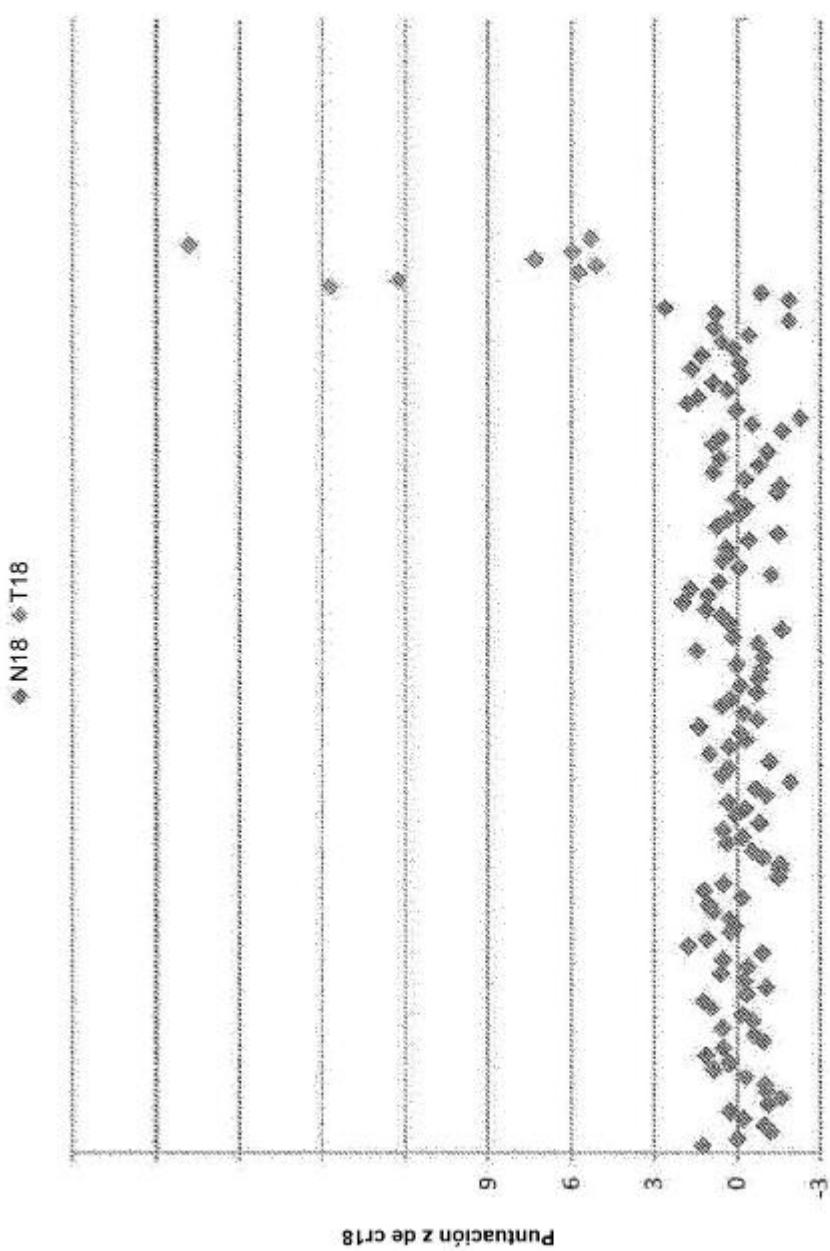


FIG. 11