

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 534 986**

51 Int. Cl.:

C12Q 1/68 (2006.01)

C12P 19/34 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **01.12.2010** **E 10830939 (4)**

97 Fecha y número de publicación de la concesión europea: **21.01.2015** **EP 2376661**

54 Título: **Determinación simultánea de aneuploidía y fracción fetal**

30 Prioridad:

19.01.2010 US 296358 P

01.07.2010 US 360837 P

26.10.2010 US 407017 P

26.10.2010 US 455849 P

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

04.05.2015

73 Titular/es:

VERINATA HEALTH, INC (100.0%)

800 Saginaw Drive

Redwood City CA 94063, US

72 Inventor/es:

QUAKE, STEPHEN;

RAVA, RICHARD P.;

CHINNAPPA, MANJULA;

COMSTOCK, DAVID A. y

HEILEK, GABRIELLE

74 Agente/Representante:

IZQUIERDO BLANCO, María Alicia

ES 2 534 986 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).

Determinación simultánea de aneuploidía y fracción fetal**Descripción****5 CAMPO DE LA INVENCION**

La invención se refiere de manera general al campo de diagnósticos, y proporciona un método que es aplicable a la práctica de diagnósticos prenatales no invasivos.

10 ANTECEDENTES DE LA INVENCION

El diagnóstico prenatal para determinar anomalías fetales potenciales proporciona una oportunidad para el cuidado y gestión necesarios durante el embarazo, el periodo neonatal y el parto. Las técnicas de formación de imágenes como ecografía, resonancia magnética y ecocardiografía fetal son útiles para la identificación de anomalías estructurales del feto. La amniocentesis, muestra de vellosidades coriónicas y muestra de sangre fetal proporcionan células y tejidos fetales para el análisis de anomalías cromosómicas, genéticas y bioquímicas, pero son invasivas y plantean riesgo alto al embarazo.

La existencia de ADN libre de células circulante en la sangre materna (Lo et al., Lancet 350:485-487 [1997]) se está aprovechando para desarrollar procesos no invasivos que usan ácidos nucleicos fetales de una muestra de sangre periférica materna para determinar anomalías cromosómicas fetales (Fan HC y Quake SR Anal Chem 79:7576-7579 [2007]; Fan et al., Proc Natl Acad Sci 105:16266-16271 [2008]).

Estos métodos proporcionan un cambio de paradigma en el diagnóstico prenatal, ya que podrían anunciar efectivamente el final de los procedimientos invasivos. Sin embargo, la sensibilidad de la determinación de aneuploidía fetal depende en gran medida de la fracción de ADN fetal, que se ha determinado que es <10% del ADN libre de células (ADNcf) circulante total (Lo et al., Am J Hum Genet 62:768-775 [1998]). Dada la relativamente baja concentración de ácidos nucleicos circulantes fetales, pueden surgir resultados de falsos negativos si hay insuficiente ácido nucleico de partida para el análisis. Por consiguiente, se han desarrollado ensayos para la determinación no invasiva de la fracción de ADN fetal, pro típicamente se basan en comparar el locus específico del feto (como el locus SRY en el cromosoma Y en embarazos masculinos) con los de un locus en cualquiera autosoma que es común tanto para la madre como el feto (Dahllan et al., Lancet 369:474-481 [2007]; Li et al., Clin Chem 1002-1011 [2004]; Fan et al., Proc Natl Acad Sci 105:16266-16271 [2008]). Además, los ensayos usados para cuantificar la fracción fetal se realizan independientemente de los ensayos que se están desarrollando para determinar la presencia o ausencia de aneuploidías en el ADN libre de células circulante.

Los documentos Lo et al., CHEMISTRY, 2008, vol. 54, no. 3, pg. 461-466 y Chu et al., BIOINFORMATICS, 2009, vol. 25, no. 10, pg. 1244-1250 divulgan métodos en los que la secuenciación de ADN en sangre materna se usa para obtener información de la dosis de cromosoma en cachos sospechosos de aneuploidía. Las variaciones anormales en la densidad y/o número de etiquetas de secuencia para cada cromosoma bajo análisis frente a un control para el mismo cromosoma están correlacionados con la presencia de aneuploidía.

Por lo tanto, sería deseable proporcionar un ensayo prenatal que proporcione un control interno para medir la adecuación de los ácidos nucleicos fetales de entrada y evite los diagnósticos incorrectos de anomalías cromosómicas fetales.

La presente invención se refiere a composiciones y métodos que permiten la determinación simultánea de la fracción fetal y la determinación de la presencia o ausencia de aneuploidía a partir de un proceso de secuenciación de diagnóstico individual. El método permite determinar la fracción fetal de una manera independiente del género, que se basa en la cuantificación de alelos en cromosomas múltiples. El método de diagnóstico no invasivo abarca el uso de tecnología de secuenciación de próxima generación (NGS) que puede ser implementada en un proceso racionalizado y rentable para proporcionar diagnósticos prenatales no invasivos de aneuploidías fetales con mayor confianza.

55 RESUMEN DE LA INVENCION

La invención se refiere a composiciones y métodos para determinar simultáneamente la presencia o ausencia de aneuploidía fetal y la cantidad relativa de ácidos nucleicos fetales en una muestra obtenida de una mujer embarazada. El método abarca el uso de tecnologías de secuenciación y explota la aparición de polimorfismos para proporcionar un proceso no invasivo racionalizado aplicable a la práctica del diagnóstico prenatal.

La invención proporciona un método para determinar simultáneamente aneuploidía y fracción fetal en una muestra de sangre, plasma o suero materno que comprende una mezcla de moléculas de ácidos nucleicos fetales y maternos, en donde la mezcla de moléculas de ácidos nucleicos fetales y maternos son moléculas de ADN libre de células (ADNcf) y dicha fracción fetal es la fracción de dichas moléculas de ácidos nucleicos aportadas por un feto a

dicha mezcla de moléculas de ácidos nucleicos fetales y maternos, en donde se sabe que cada uno de dichos ácidos nucleicos comprende al menos un sitio polimórfico, y en donde dicho enriquecimiento comprende: (i) amplificar específicamente dicha pluralidad de ácidos nucleicos objetivo en una porción de dicha mezcla usando pares de cebadores cada uno capaz de amplificar una secuencia de ácidos nucleicos objetivo que comprende un sitio polimórfico en una reacción de PCR multiplex para generar un panel de sitios polimórficos amplificados que contiene un número suficiente de sitios polimórficos de tal manera que al menos dos son sitios polimórficos informativos; y (ii) combinar al menos una porción o todo el producto amplificado con al menos una porción del resto de la muestra no amplificada de la que se retiró la porción; (b) secuenciar al menos una porción de la mezcla enriquecida obtenida en el paso (a), en donde dicha secuenciación comprende proporcionar una pluralidad de lecturas de secuencia que se mapean para un genoma para identificar una pluralidad de etiquetas de secuencia mapeadas; y (c) usar dicha pluralidad de etiquetas de secuencia mapeadas, determinar simultáneamente dicha fracción fetal y dicha aneuploidía, en donde: (i) determinar dicha fracción fetal comprende: (1) usar las etiquetas de secuencia mapeadas para identificar al menos dos sitios polimórficos informativos en dicho panel de sitios polimórficos amplificados, en donde dichos sitios polimórficos informativos se identifican por la diferencia en las secuencias alélicas y el número de etiquetas de secuencia mapeadas para cada uno de los posibles alelos en cada sitio polimórfico, y (2) calcular la fracción fetal en base al número total de etiquetas de secuencia que mapean para un primer alelo y el número total de etiquetas de secuencia que mapean para un segundo alelo en cada uno de dichos sitios polimórficos informativos; y (ii) determinar dicha aneuploidía comprende la cuantificación del número de etiquetas de secuencia que se alinean con un cromosoma de interés, y comparar los resultados obtenidos para el cromosoma de interés con un valor límite, en donde el valor límite es un número que sirve como un límite de diagnóstico de una aneuploidía y en donde la presencia de una aneuploidía para el cromosoma de interés se identifica si el valor límite se supera por los resultados obtenidos para el cromosoma de interés.

En una realización, se divulga un método para determinar simultáneamente aneuploidía y fracción fetal en una muestra materna que comprende una mezcla de moléculas de ácidos nucleicos fetales y maternos, el método comprende: (a) enriquecer dicha mezcla para una pluralidad de ácidos nucleicos objetivo polimórficos; (b) secuenciar al menos una porción de la mezcla enriquecida obtenida en el paso (a), en donde la secuenciación comprende proporcionar una pluralidad de etiquetas de secuencia; y (c) en base a la secuenciación, determinar simultáneamente la fracción fetal y la presencia o ausencia de aneuploidía fetal.

En otra realización, se divulga un método para determinar simultáneamente aneuploidía y fracción fetal en una muestra materna que comprende una mezcla de moléculas de ácidos nucleicos fetales y maternos, el método comprende: (a) enriquecer dicha mezcla para una pluralidad de ácidos nucleicos objetivo polimórficos, en donde enriquecer comprende amplificar una pluralidad de ácidos nucleicos objetivo polimórficos en una porción de dicha mezcla; (b) secuenciar al menos una porción de la mezcla enriquecida obtenida en el paso (a), en donde la secuenciación comprende proporcionar una pluralidad de etiquetas de secuencia; y (c) en base a la secuenciación, determinar simultáneamente la fracción fetal y la presencia o ausencia de aneuploidía fetal.

En otra realización, se divulga un método para determinar simultáneamente aneuploidía y fracción fetal en una muestra materna que comprende una mezcla de moléculas de ácidos nucleicos fetales y maternos, el método comprende: (a) enriquecer dicha mezcla para una pluralidad de ácidos nucleicos objetivo polimórficos, en donde enriquecer comprende amplificar una pluralidad de ácidos nucleicos objetivo polimórficos en una porción de una mezcla purificada de ácidos nucleicos fetales y maternos; (b) secuenciar al menos una porción de la mezcla enriquecida obtenida en el paso (a), en donde la secuenciación comprende proporcionar una pluralidad de etiquetas de secuencia; y (c) en base a la secuenciación, determinar simultáneamente la fracción fetal y la presencia o ausencia de aneuploidía fetal.

En otra realización, se divulga un método para determinar simultáneamente aneuploidía y fracción fetal en una muestra materna que comprende una mezcla de moléculas de ácidos nucleicos fetales y maternos, el método comprende: (a) enriquecer dicha mezcla para una pluralidad de ácidos nucleicos objetivo polimórficos, en donde enriquecer comprende combinar al menos una porción de una primera biblioteca de secuenciación de dicha mezcla de moléculas de ácidos nucleicos fetales y maternos con al menos una porción de una segunda biblioteca de secuenciación de dicha mezcla de moléculas de ácidos nucleicos fetales y maternos; (b) secuenciar al menos una porción de la mezcla enriquecida obtenida en el paso (a), en donde la secuenciación comprende proporcionar una pluralidad de etiquetas de secuencia; y (c) en base a la secuenciación, determinar simultáneamente la fracción fetal y la presencia o ausencia de aneuploidía fetal.

En una realización, se divulga un método para determinar simultáneamente aneuploidía y fracción fetal en una muestra materna que comprende una mezcla de moléculas de ácidos nucleicos fetales y maternos, el método comprende: (a) enriquecer dicha mezcla para una pluralidad de ácidos nucleicos objetivo polimórficos, en donde cada uno de la pluralidad de ácidos nucleicos objetivo polimórficos comprende al menos un polimorfismo de nucleótido simple (SNP); (b) secuenciar al menos una porción de la mezcla enriquecida obtenida en el paso (a), en donde la secuenciación comprende proporcionar una pluralidad de etiquetas de secuencia; y (c) en base a la secuenciación, determinar simultáneamente la fracción fetal y la presencia o ausencia de aneuploidía fetal. En algunas realizaciones, el al menos un SNP, es un SNP individual seleccionado de cada uno de dicha pluralidad de

ácidos nucleicos objetivo polimórficos que comprende un SNP seleccionado de rs560681, rs1109037, rs9866013, rs13182883, rs13218440, rs7041158, rs740598, rs10773760, rs 4530059, rs7205345, rs8078417, rs576261, rs2567608, rs430046, rs9951171, rs338882, rs10776839, rs9905977, rs1277284, rs258684, rs1347696, rs508485, rs9788670, rs8137254, rs3143, rs2182957, rs3739005, y rs530022. Alternativamente, el al menos un SNP es un conjunto de dos SNPs en tándem seleccionado de los conjuntos rs7277033-rs2110153; rs2822654-rs1882882; rs368657-rs376635; rs2822731-rs2822732; rs1475881-rs7275487; rs1735976-rs2827016; rs447340-rs2824097; rs418989-rs13047336; rs987980-rs987981; rs4143392-rs4143391; rs1691324-rs13050434; rs11909758-rs9980111; rs2826842-rs232414; rs1980969-rs1980970; rs9978999-rs9979175; rs1034346-rs12481852; rs7509629-rs2828358; rs4817013-rs7277036; rs9981121-rs2829696; rs455921-rs2898102; rs2898102-rs458848; rs961301-rs2830208; rs2174536-rs458076; rs11088023-rs11088024; rs1011734-rs1011733; rs2831244-rs9789838; rs8132769-rs2831440; rs8134080-rs2831524; rs4817219-rs4817220; rs2250911-rs2250997; rs2831899-rs2831900; rs2831902-rs2831903; rs11088086-rs2251447; rs2832040-rs11088088; rs2832141-rs2246777; rs2832959-rs9980934; rs2833734-rs2833735; rs933121-rs933122; rs2834140-rs12626953; rs2834485-rs3453; rs9974986-rs2834703; rs2776266-rs2835001; rs1984014-rs1984015; rs7281674-rs2835316; rs13047304-rs13047322; rs2835545-rs4816551; rs2835735-rs2835736; rs13047608-rs2835826; rs2836550-rs2212596; rs2836660-rs2836661; rs465612-rs8131220; rs9980072-rs8130031; rs418359-rs2836926; rs7278447-rs7278858; rs385787-rs367001; rs367001-rs386095; rs2837296-rs2837297; y rs2837381-rs4816672.

En una realización, se divulga un método para determinar simultáneamente aneuploidía y fracción fetal en una muestra materna que comprende una mezcla de moléculas de ácidos nucleicos fetales y maternos, el método comprende: (a) enriquecer dicha mezcla para una pluralidad de ácidos nucleicos objetivo polimórficos, en donde cada uno de la pluralidad de ácidos nucleicos objetivo polimórficos comprende al menos un polimorfismo de nucleótido simple (SNP); (b) secuenciar al menos una porción de la mezcla enriquecida obtenida en el paso (a), en donde la secuenciación comprende proporcionar una pluralidad de etiquetas de secuencia; y (c) en base a la secuenciación, determinar simultáneamente la fracción fetal y la presencia o ausencia de aneuploidía fetal. El paso de enriquecer comprende amplificar una pluralidad de ácidos nucleicos objetivo polimórficos en una porción de dicha mezcla. En algunas realizaciones, el al menos un SNP, es un SNP individual seleccionado de cada uno de dicha pluralidad de ácidos nucleicos objetivo polimórficos que comprende un SNP seleccionado de rs560681, rs1109037, rs9866013, rs13182883, rs13218440, rs7041158, rs740598, rs10773760, rs 4530059, rs7205345, rs8078417, rs576261, rs2567608, rs430046, rs9951171, rs338882, rs10776839, rs9905977, rs1277284, rs258684, rs1347696, rs508485, rs9788670, rs8137254, rs3143, rs2182957, rs3739005, y rs530022. Alternativamente, el al menos un SNP es un conjunto de dos SNPs en tándem seleccionado de los conjuntos rs7277033-rs2110153; rs2822654-rs1882882; rs368657-rs376635; rs2822731-rs2822732; rs1475881-rs7275487; rs1735976-rs2827016; rs447340-rs2824097; rs418989-rs13047336; rs987980-rs987981; rs4143392-rs4143391; rs1691324-rs13050434; rs11909758-rs9980111; rs2826842-rs232414; rs1980969-rs1980970; rs9978999-rs9979175; rs1034346-rs12481852; rs7509629-rs2828358; rs4817013-rs7277036; rs9981121-rs2829696; rs455921-rs2898102; rs2898102-rs458848; rs961301-rs2830208; rs2174536-rs458076; rs11088023-rs11088024; rs1011734-rs1011733; rs2831244-rs9789838; rs8132769-rs2831440; rs8134080-rs2831524; rs4817219-rs4817220; rs2250911-rs2250997; rs2831899-rs2831900; rs2831902-rs2831903; rs11088086-rs2251447; rs2832040-rs11088088; rs2832141-rs2246777; rs2832959-rs9980934; rs2833734-rs2833735; rs933121-rs933122; rs2834140-rs12626953; rs2834485-rs3453; rs9974986-rs2834703; rs2776266-rs2835001; rs1984014-rs1984015; rs7281674-rs2835316; rs13047304-rs13047322; rs2835545-rs4816551; rs2835735-rs2835736; rs13047608-rs2835826; rs2836550-rs2212596; rs2836660-rs2836661; rs465612-rs8131220; rs9980072-rs8130031; rs418359-rs2836926; rs7278447-rs7278858; rs385787-rs367001; rs367001-rs386095; rs2837296-rs2837297; y rs2837381-rs4816672.

En una realización, se divulga un método para determinar simultáneamente aneuploidía y fracción fetal en una muestra materna que comprende una mezcla de moléculas de ácidos nucleicos fetales y maternos, el método comprende: (a) enriquecer dicha mezcla para una pluralidad de ácidos nucleicos objetivo polimórficos, en donde cada uno de la pluralidad de ácidos nucleicos objetivo polimórficos comprende al menos un polimorfismo de nucleótido simple (SNP); (b) secuenciar al menos una porción de la mezcla enriquecida obtenida en el paso (a), en donde la secuenciación comprende proporcionar una pluralidad de etiquetas de secuencia; y (c) en base a la secuenciación, determinar simultáneamente la fracción fetal y la presencia o ausencia de aneuploidía fetal. El paso de enriquecer comprende amplificar una pluralidad de ácidos nucleicos objetivo polimórficos en una porción de una mezcla purificada de ácidos nucleicos fetales y maternos. En algunas realizaciones, el al menos un SNP, es un SNP individual seleccionado de rs560681, rs1109037, rs9866013, rs13182883, rs13218440, rs7041158, rs740598, rs10773760, rs4530059, rs7205345, rs8078417, rs576261, rs2567608, rs430046, rs9951171, rs338882, rs10776839, rs9905977, rs1277284, rs258684, rs1347696, rs508485, rs9788670, rs8137254, rs3143, rs2182957, rs3739005, y rs530022. Alternativamente, el al menos un SNP es un conjunto de dos SNPs en tándem seleccionado de los conjuntos rs7277033-rs2110153; rs2822654-rs1882882; rs368657-rs376635; rs2822731-rs2822732; rs1475881-rs7275487; rs1735976-rs2827016; rs447340-rs2824097; rs418989-rs13047336; rs987980-rs987981; rs4143392-rs4143391; rs1691324-rs13050434; rs11909758-rs9980111; rs2826842-rs232414; rs1980969-rs1980970; rs9978999-rs9979175; rs1034346-rs12481852; rs7509629-rs2828358; rs4817013-rs7277036; rs9981121-rs2829696; rs455921-rs2898102; rs2898102-rs458848; rs961301-rs2830208; rs2174536-rs458076; rs11088023-rs11088024; rs1011734-rs1011733; rs2831244-rs9789838; rs8132769-rs2831440; rs8134080-rs2831524; rs4817219-rs4817220; rs2250911-rs2250997; rs2831899-rs2831900; rs2831902-rs2831903;

rs11088086-rs2251447; rs2832040-rs11088088; rs2832141-rs2246777; rs2832959-rs9980934; rs2833734-rs2833735; rs933121-rs933122; rs2834140-rs12626953; rs2834485-rs3453; rs9974986-rs2834703; rs2776266-rs2835001; rs1984014-rs1984015; rs7281674-rs2835316; rs13047304-rs13047322; rs2835545-rs4816551; rs2835735-rs2835736; rs13047608-rs2835826; rs2836550-rs2212596; rs2836660-rs2836661; rs465612-rs8131220; rs9980072-rs8130031; rs418359-rs2836926; rs7278447-rs7278858; rs385787-rs367001; rs367001-rs386095; rs2837296-rs2837297; y rs2837381-rs4816672.

En una realización, se divulga un método para determinar simultáneamente aneuploidía y fracción fetal en una muestra materna que comprende una mezcla de moléculas de ácidos nucleicos fetales y maternos, el método comprende: (a) enriquecer dicha mezcla para una pluralidad de ácidos nucleicos objetivo polimórficos, en donde cada uno de la pluralidad de ácidos nucleicos objetivo polimórficos comprende al menos un polimorfismo de nucleótido simple (SNP); (b) secuenciar al menos una porción de la mezcla enriquecida obtenida en el paso (a), en donde la secuenciación comprende proporcionar una pluralidad de etiquetas de secuencia; y (c) en base a la secuenciación, determinar simultáneamente la fracción fetal y la presencia o ausencia de aneuploidía fetal. El paso de enriquecer comprende combinar al menos una porción de una primera biblioteca de secuenciación de dicha mezcla de moléculas de ácidos nucleicos fetales y maternos con al menos una porción de una segunda biblioteca de secuenciación de ácidos nucleicos objetivo polimórficos amplificados. En algunas realizaciones, el al menos un SNP, es un SNP individual seleccionado de cada uno de dicha pluralidad de ácidos nucleicos objetivo polimórficos que comprende un SNP seleccionado de rs560681, rs1109037, rs9866013, rs13182883, rs13218440, rs7041158, rs740598, rs10773760, rs4530059, rs7205345, rs8078417, rs576261, rs2567608, rs430046, rs9951171, rs338882, rs10776839, rs9905977, rs1277284, rs258684, rs1347696, rs508485, rs9788670, rs8137254, rs3143, rs2182957, rs3739005, y rs530022. Alternativamente, el al menos un SNP es un conjunto de dos SNPs en tándem seleccionados de los conjuntos rs7277033-rs2110153; rs2822654-rs1882882; rs368657-rs376635; rs2822731-rs2822732; rs1475881-rs7275487; rs1735976-rs2827016; rs447340-rs2824097; rs418989-rs13047336; rs987980-rs987981; rs4143392-rs4143391; rs1691324-rs13050434; rs11909758-rs9980111; rs2826842-rs232414; rs1980969-rs1980970; rs9978999-rs9979175; rs1034346-rs12481852; rs7509629-rs2828358; rs4817013-rs7277036; rs9981121-rs2829696; rs455921-rs2898102; rs2898102-rs458848; rs961301-rs2830208; rs2174536-rs458076; rs11088023-rs11088024; rs1011734-rs1011733; rs2831244-rs9789838; rs8132769-rs2831440; rs8134080-rs2831524; rs4817219-rs4817220; rs2250911-rs2250997; rs2831899-rs2831900; rs2831902-rs2831903; rs11088086-rs2251447; rs2832040-rs11088088; rs2832141-rs2246777; rs2832959-rs9980934; rs2833734-rs2833735; rs933121-rs933122; rs2834140-rs12626953; rs2834485-rs3453; rs9974986-rs2834703; rs2776266-rs2835001; rs1984014-rs1984015; rs7281674-rs2835316; rs13047304-rs13047322; rs2835545-rs4816551; rs2835735-rs2835736; rs13047608-rs2835826; rs2836550-rs2212596; rs2836660-rs2836661; rs465612-rs8131220; rs9980072-rs8130031; rs418359-rs2836926; rs7278447-rs7278858; rs385787-rs367001; rs367001-rs386095; rs2837296-rs2837297; y rs2837381-rs4816672.

En otra realización, se divulga un método para determinar simultáneamente aneuploidía y fracción fetal en una muestra materna que comprende una mezcla de moléculas de ácidos nucleicos fetales y maternos, el método comprende: (a) enriquecer dicha mezcla para una pluralidad de ácidos nucleicos objetivo polimórficos, en donde cada uno de la pluralidad de ácidos nucleicos objetivo polimórficos comprende al menos una repetición corta en tándem (STR); (b) secuenciar al menos una porción de la mezcla enriquecida obtenida en el paso (a), en donde la secuenciación comprende proporcionar una pluralidad de etiquetas de secuencia; y (c) en base a la secuenciación, determinar simultáneamente la fracción fetal y la presencia o ausencia de aneuploidía fetal. En algunas realizaciones, la al menos una STR es menor de alrededor de 200 pares de base. En otras realizaciones, cada uno de dicha pluralidad de ácidos nucleicos objetivo polimórficos comprende una STR seleccionada de CSF1PO, FGA, TH01, TPOX, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, D2S1338, Penta D, Penta E, D22S1045, D20S1082, D20S482, D18S853, D17S1301, D17S974, D14S1434, D12ATA63, D11S4463, D10S1435, D10S1248, D9S2157, D9S1122, D8S1115, D6S1017, D6S474, D5S2500, D4S2408, D4S2364, D3S4529, D3S3053, D2S1776, D2S441, D1S1677, D1S1627 y D1GATA113.

En otra realización, se divulga un método para determinar simultáneamente aneuploidía y fracción fetal en una muestra materna que comprende una mezcla de moléculas de ácidos nucleicos fetales y maternos, el método comprende: (a) enriquecer dicha mezcla para una pluralidad de ácidos nucleicos objetivo polimórficos, en donde cada uno de la pluralidad de ácidos nucleicos objetivo polimórficos comprende al menos una repetición corta en tándem (STR); (b) secuenciar al menos una porción de la mezcla enriquecida obtenida en el paso (a), en donde la secuenciación comprende proporcionar una pluralidad de etiquetas de secuencia; y (c) en base a la secuenciación, determinar simultáneamente la fracción fetal y la presencia o ausencia de aneuploidía fetal. El paso de enriquecer comprende amplificar una pluralidad de ácidos nucleicos objetivo polimórficos en una porción de la mezcla. En algunas realizaciones, la al menos una STR es menor de alrededor de 200 pares de base. En otras realizaciones, cada uno de dicha pluralidad de ácidos nucleicos objetivo polimórficos comprende una STR seleccionada de CSF1PO, FGA, TH01, TPOX, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, D2S1338, Penta D, Penta E, D22S1045, D20S1082, D20S482, D18S853, D17S1301, D17S974, D14S1434, D12ATA63, D11S4463, D10S1435, D10S1248, D9S2157, D9S1122, D8S1115, D6S1017, D6S474, D5S2500, D4S2408, D4S2364, D3S4529, D3S3053, D2S1776, D2S441, D1S1677, D1S1627 y D1GATA113.

En otra realización, se divulga un método para determinar simultáneamente aneuploidía y fracción fetal en

una muestra materna que comprende una mezcla de moléculas de ácidos nucleicos fetales y maternos, el método comprende: (a) enriquecer dicha mezcla para una pluralidad de ácidos nucleicos objetivo polimórficos, en donde cada uno de la pluralidad de ácidos nucleicos objetivo polimórficos comprende al menos una repetición corta en tándem (STR); (b) secuenciar al menos una porción de la mezcla enriquecida obtenida en el paso (a), en donde la secuenciación comprende proporcionar una pluralidad de etiquetas de secuencia; y (c) en base a la secuenciación, determinar simultáneamente la fracción fetal y la presencia o ausencia de aneuploidía fetal. El paso de enriquecer comprende amplificar una pluralidad de ácidos nucleicos objetivo polimórficos en una porción de una mezcla purificada de ácidos nucleicos fetales y maternos. En algunas realizaciones, la al menos una STR es menor de alrededor de 200 pares de base. En otras realizaciones, cada uno de dicha pluralidad de ácidos nucleicos objetivo polimórficos comprende una STR seleccionada de CSF1PO, FGA, TH01, TPOX, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, D2S1338, Penta D, Penta E, D22S1045, D20S1082, D20S482, D18S853, D17S1301, D17S974, D14S1434, D12ATA63, D11S4463, D10S1435, D10S1248, D9S2157, D9S1122, D8S1115, D6S1017, D6S474, D5S2500, D4S2408, D4S2364, D3S4529, D3S3053, D2S1776, D2S441, D1S1677, D1S1627 y D1GATA113.

En otra realización, se divulga un método para determinar simultáneamente aneuploidía y fracción fetal en una muestra materna que comprende una mezcla de moléculas de ácidos nucleicos fetales y maternos, el método comprende: (a) enriquecer dicha mezcla para una pluralidad de ácidos nucleicos objetivo polimórficos, en donde cada uno de la pluralidad de ácidos nucleicos objetivo polimórficos comprende al menos una repetición corta en tándem (STR); (b) secuenciar al menos una porción de la mezcla enriquecida obtenida en el paso (a), en donde la secuenciación comprende proporcionar una pluralidad de etiquetas de secuencia; y (c) en base a la secuenciación, determinar simultáneamente la fracción fetal y la presencia o ausencia de aneuploidía fetal. El paso de enriquecer comprende combinar al menos una porción de una primera biblioteca de secuenciación de dicha mezcla de moléculas de ácidos nucleicos fetales y maternos con al menos una porción de una segunda biblioteca de secuenciación de ácidos nucleicos objetivo polimórficos amplificados. En algunas realizaciones, la al menos una STR es menor de alrededor de 200 pares de base. En otras realizaciones, cada uno de dicha pluralidad de ácidos nucleicos objetivo polimórficos comprende una STR seleccionada de CSF1PO, FGA, TH01, TPOX, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, D2S1338, Penta D, Penta E, D22S1045, D20S1082, D20S482, D18S853, D17S1301, D17S974, D14S1434, D12ATA63, D11S4463, D10S1435, D10S1248, D9S2157, D9S1122, D8S1115, D6S1017, D6S474, D5S2500, D4S2408, D4S2364, D3S4529, D3S3053, D2S1776, D2S441, D1S1677, D1S1627 y D1GATA113.

En las realizaciones del método resumido anteriormente y descritas con más detalle a continuación, la muestra materna es una muestra biológica que se puede elegir, pero no está limitada a, sangre, plasma, suero, orina y saliva. Preferiblemente, las moléculas de ácidos nucleicos fetales y maternos en la muestra materna son moléculas de ADN libre de células (ADNcf). Los ácidos nucleicos objetivo polimórficos pueden ser del mismo cromosoma o de cromosomas diferentes.

En las realizaciones del método resumido anteriormente y descrito con más detalle a continuación, la aneuploidía que se determina puede ser una aneuploidía cromosómica o parcial. En algunas realizaciones, la aneuploidía es una aneuploidía cromosómica que se selecciona de trisomía 8, trisomía 13, trisomía 15, trisomía 16, trisomía 18, trisomía 21, trisomía 22, monosomía X y XXX. En algunas realizaciones, determinar la aneuploidía comprende calcular una dosis de cromosoma en base al número de dichas etiquetas de secuencia para el cromosoma de interés y para un cromosoma normalizador, y comparar dicha dosis con un valor límite, mientras que determinar la fracción fetal comprende identificar al menos un sitio polimórfico informativo en dicha mezcla enriquecida, y calcular la fracción fetal de la cantidad de sitios polimórficos fetales y maternos en dicha muestra enriquecida.

En las realizaciones del método resumido anteriormente y descrito con más detalle a continuación, la secuenciación que puede usarse para la determinación simultánea se realiza usando secuenciación de próxima generación (NGS). En algunas realizaciones, la secuenciación es secuenciación masivamente paralela usando secuenciación por síntesis con terminadores de colorante reversibles. En otras realizaciones, la secuenciación es secuenciación por ligadura. En otras realizaciones, la secuenciación es secuenciación de moléculas individuales. La secuenciación de la mezcla enriquecida puede comprender además una amplificación.

Preferiblemente, las moléculas de ácidos nucleicos fetales y maternos son moléculas de ADN libre de células. En algunas realizaciones, la muestra materna es una muestra de plasma. EL kit comprende una composición que comprende al menos un conjunto de cebadores para amplificar al menos un ácido nucleico polimórfico comprendido en las moléculas de ácidos nucleicos fetales y maternos. En algunas realizaciones el ácido nucleico polimórfico comprende un SNP. En otra realización, el ácido nucleico polimórfico comprende una STR.

BREVE DESCRIPCION DE LOS DIBUJOS

Las características nuevas de la invención se exponen con particularidad en las reivindicaciones añadidas. Se obtendrá una mejor comprensión de las características y ventajas de la presente invención con referencia a la

siguiente descripción detallada que expone realizaciones ilustrativas, en las que se utilizan los principios de la invención, y los dibujos acompañantes de los cuales:

La **Figura 1** es un diagrama de flujo de un método **100** para determinar simultáneamente la presencia o ausencia de aneuploidía y la fracción fetal en una muestra de ensayo materna que comprende una mezcla de ácidos nucleicos fetales y maternos.

La **Figura 2** es un diagrama de flujo de un método **200** para determinar simultáneamente la presencia o ausencia de aneuploidía fetal y la fracción fetal en una muestra de ensayo materna enriquecida para ácidos nucleicos polimórficos.

La **Figura 3** es un diagrama de flujo de un método **300** para determinar simultáneamente la presencia o ausencia de aneuploidía fetal y la fracción fetal en una muestra de ensayo de ADN libre de células purificado materno que ha sido enriquecida con ácidos nucleicos polimórficos.

La **Figura 4** es un diagrama de flujo de un método **400** para determinar simultáneamente la presencia o ausencia de aneuploidía fetal y la fracción fetal en una biblioteca de secuenciación construida de ácidos nucleicos fetales y maternos derivados de una muestra de ensayo materna y enriquecida con ácidos nucleicos polimórficos.

La **Figura 5** es un diagrama de flujo de un método **500** para determinar la presencia o ausencia de una variación en el número de copias en una muestra de ensayo que comprende una mezcla de ácidos nucleicos.

La **Figura 6** es un diagrama de barras que muestra la identificación de secuencias polimórficas fetales y maternas (SNPs) usadas para determinar la fracción fetal en una muestra de ensayo. Se muestran el número total de lecturas de secuencias (eje Y) mapeadas para las secuencias SNP identificadas por los números rs (eje X), y el nivel relativo de ácidos nucleicos fetales (*).

La **Figura 7** ilustra la distribución de la dosis de cromosoma para el cromosoma 21 determinada por la secuenciación del ADN libre de células extraído de un conjunto de 48 muestras de sangre obtenidas de sujetos humanos embarazados con fetos masculinos o femeninos. Se muestran las dosis de cromosoma 21 para muestras de ensayo calificadas, es decir, normales para el cromosoma 21 (O), y para trisomía 21 (Δ) para los cromosomas 1-12 y X (Figura 7A), y para los cromosomas 1-22 y X (Figura 7B).

La **Figura 8** ilustra la distribución de la dosis de cromosoma para el cromosoma 18 determinada de la secuenciación de ADN libre de células extraído de un conjunto de 48 muestras de sangre obtenidas de sujetos humanos embarazados con un feto masculino o femenino. Las dosis de cromosoma 18 para las muestras de ensayo calificadas, es decir, normales para el cromosoma 18 (O), y trisomía 18 (Δ) se muestran para los cromosomas 1-12 y X (Figura 8A), y para los cromosomas 1-22 y X (Figura 8B).

La **Figura 9** ilustra la distribución de la dosis de cromosoma para el cromosoma 13 determinada de la secuenciación de ADN libre de células extraído de un conjunto de 48 muestras de sangre obtenidas de sujetos humanos embarazados con un feto masculino o femenino. Las dosis de cromosoma 13 para las muestras de ensayo calificadas, es decir normales para el cromosoma 13 (O), y trisomía 13 (Δ) se muestran para los cromosomas 1-12 y X (Figura 9A), y para los cromosomas 1-22 y X (Figura 9B).

La **Figura 10** ilustra la distribución de las dosis de cromosoma para el cromosoma X determinada de la secuenciación de ADN libre de células extraído de un conjunto de 48 muestras de sangre obtenidas de sujetos humanos embarazados con un feto masculino o femenino. Las dosis de cromosoma X para las muestras masculinas (46,XY; (O)), femeninas (46,XX(Δ)); monosomía X (45,X; (+)), y cariotipos complejos (Cplx (X)) se muestran para los cromosomas 1-12 y X (Figura 10A) y para los cromosomas 1-22 y X (Figura 10B).

La **Figura 11** ilustra la distribución de las dosis de cromosoma para el cromosoma Y determinada de la secuenciación de ADN libre de células extraído de un conjunto de 48 muestras de sangre obtenidas de sujetos humanos embarazados de un feto masculino o femenino. Las dosis de cromosoma Y para las muestras masculinas (46,XY; (Δ)), femeninas (46,XX(O)); monosomía X (45,X; (+)), y cariotipos complejos (Cplx (X)) se muestran para los cromosomas 1-12 (Figura 11A) y para los cromosomas 1-22 (Figura 11B).

La **Figura 12** muestra el coeficiente de variación (CV) para los cromosomas 21 ($\square$), 18 (O) y 13 (Δ) que se determinó de las dosis de cromosoma de muestras calificadas, es decir, no afectadas mostradas en las Figuras 7, 8 y 9, respectivamente.

La **Figura 13** muestra el coeficiente de variación (CV) para los cromosomas X ($\square$) e Y (O) que se determinó de las dosis de cromosoma de muestras calificadas, es decir, no afectadas mostradas en la Figuras 10 y 11, respectivamente.

DESCRIPCION DETALLADA DE LA INVENCION

La invención se refiere a composiciones y métodos para determinar simultáneamente la presencia o ausencia de aneuploidía fetal y la cantidad relativa de ácidos nucleicos fetales en una muestra obtenida de una mujer embarazada. El método abarca el uso de tecnologías de secuenciación, por ejemplo, secuenciación de próxima generación, y explota la aparición de polimorfismos para proporcionar un proceso no invasivo simplificado aplicable a la práctica del diagnóstico prenatal.

A menos que se indique lo contrario, la práctica de la presente invención implica técnicas convencionales

usadas comúnmente en biología molecular, microbiología, purificación de proteínas, diseño de proteínas, secuenciación de proteínas y ADN y campos de ADN recombinante, que están dentro de la técnica. Dichas técnicas son conocidas por los expertos en la técnica y se describen en numerosos textos estándar y trabajos de referencia.

Los intervalos numéricos incluyen los números que definen el intervalo. Se pretende que cada limitación numérica máxima dada a lo largo de esta especificación incluya cada limitación numérica inferior, como si dichas limitaciones numéricas inferiores estuvieran expresamente escritas en la presente. Cada limitación numérica mínima dada a lo largo de esta especificación incluirá cada limitación numérica más alta, como si dichas limitaciones numéricas más altas estuvieran escritas expresamente en la presente. Cada intervalo numérico dado a lo largo de esta especificación incluirá cada intervalo numérico más estrecho que caiga dentro de dicho intervalo numérico más amplio, como si dichos intervalos numéricos más estrechos estuvieran todos expresamente escritos en la presente.

Los encabezamientos proporcionados en la presente no son limitaciones de los varios aspectos o realizaciones de la invención que pueden ser tenidas con referencia a la Especificación como un todo. Por consiguiente, como se ha indicado anteriormente, los términos definidos inmediatamente a continuación se definen más completamente con referencia a la especificación como un todo.

A menos que se defina lo contrario en la presente, todos los términos técnicos y científicos usados en la presente tienen el mismo significado que el que entendería comúnmente alguien experto en la técnica a la que pertenece esta invención. Varios diccionarios científicos que incluyen los términos incluidos en la presente son bien conocidos y están disponibles para los expertos en la técnica. Aunque cualquier método y material similar o equivalente a los descritos en la presente encuentran uso en la práctica o ensayo de la presente invención, se describen algunos métodos y materiales preferidos. Por consiguiente, los términos definidos inmediatamente a continuación se describen más completamente con referencia a la Especificación como un todo. Se debe entender que esta invención no está limitada a la metodología, protocolos y reactivos particulares descritos, ya que estos pueden variar, dependiendo del contexto en el que se usen por los expertos en la técnica.

DEFINICIONES

Como se usa en la presente, los términos singulares "un", "uno" y "el" incluyen la referencia plural a menos que el contexto indique claramente lo contrario. A menos que se indique lo contrario, los ácidos nucleicos están escritos de izquierda a derecha en orientación 5' a 3' y las secuencias de aminoácidos están escritas de izquierda a derecha en orientación amino a carboxi, respectivamente.

El término "evaluar" en la presente se refiere a caracterizar el estado de una aneuploidía cromosómica por uno de los tres tipos de designaciones: "normal", "afectada" y "no designada". Por ejemplo, en presencia de trisomía la designación "normal" se determina por el valor de un parámetro, por ejemplo una dosis de cromosoma de ensayo que está por debajo de un límite definido por el usuario de fiabilidad, la designación "afectada" está determinada por un parámetro, por ejemplo una dosis de cromosoma de ensayo, que está por encima de un límite definido por el usuario de fiabilidad, y el resultado "no designada" se determina por un parámetro, por ejemplo una dosis de cromosoma de ensayo, que se encuentra entre los límites definidos por el usuario de fiabilidad para hacer una designación "normal" o una "afectada".

El término "variación del número de copias" en la presente se refiere a la variación en el número de copias de una secuencia de ácidos nucleicos que es de 1 kb o mayor presente en una muestra de ensayo en comparación con el número de copias de la secuencia de ácidos nucleicos presente en una muestra calificada. Una "variante del número de copias" se refiere a la secuencia de 1kb o mayor del ácido nucleico en el que las diferencias del número de copias se encuentran por comparación de una secuencia de interés en la muestra de ensayo que se presenta en una muestra calificada. Las variantes/variaciones del número de copias incluyen delecciones, incluyendo microdelecciones, inserciones, incluyendo microinserciones, duplicaciones, multiplicaciones, inversiones, translocaciones y variantes multi-sitio complejas. El CNV abarca aneuploidías cromosómicas y aneuploidías parciales.

El término "aneuploidía" en la presente se refiere a un desequilibrio de material genético causado por una pérdida o ganancia de un cromosoma completo, o parte de un cromosoma.

El término "aneuploidía cromosómica" en la presente se refiere a un desequilibrio de material genético causado por una pérdida o ganancia de un cromosoma completo, e incluye aneuploidía de la línea germinal y aneuploidía en mosaico.

El término "aneuploidía parcial" en la presente se refiere a un desequilibrio de material genético causado por una pérdida o ganancia de parte de un cromosoma, por ejemplo, monosomía parcial y trisomía parcial, y abarca desequilibrios resultantes de translocaciones, delecciones e inserciones.

El término "SNPs en tándem" se refiere en la presente a dos o más SNPs que están presentes dentro de

una secuencia de ácidos nucleicos objetivo polimórficos.

El término "ácido nucleico objetivo polimórfico", "secuencia polimórfica", "secuencia de ácidos nucleicos objetivo polimórficos" y "ácido nucleico polimórfico" se usan de manera intercambiable en la presente para referirse a una secuencia de ácidos nucleicos, por ejemplo una secuencia de ADN, que comprende uno o más sitios polimórficos.

El término "sitio polimórfico" se refiere en la presente a un polimorfismo de nucleótido simple (SNP), una delección o inserción multi-base a pequeña escala, un Polimorfismo Multi-Nucleótido (MNP) o una Repetición Corta en Tándem (STR).

El término "pluralidad" se usa en la presente con referencia a un número de moléculas de ácidos nucleicos o etiquetas de secuencia que es suficiente para identificar diferencias significativas en variaciones del número de copias (por ejemplo dosis de cromosoma) en muestras de ensayo y muestras calificadas usando los métodos divulgados en la presente. En algunas realizaciones, se obtienen al menos alrededor de 3×10^6 etiquetas de secuencia, al menos alrededor de 5×10^6 etiquetas de secuencia, al menos alrededor de 8×10^6 etiquetas de secuencia, al menos alrededor de 10×10^6 etiquetas de secuencia, al menos alrededor de 15×10^6 etiquetas de secuencia, al menos alrededor de 20×10^6 etiquetas de secuencia, al menos alrededor de 30×10^6 etiquetas de secuencia, al menos alrededor de 405×10^6 etiquetas de secuencia, al menos alrededor de 50×10^6 etiquetas de secuencia que comprenden lecturas entre de 20 y 40bp para cada muestra de ensayo.

El término "polinucleótido", "ácido nucleico" y "moléculas de ácidos nucleicos" se usan de manera intercambiable y se refieren a secuencias ligadas covalentemente de nucleótidos (es decir, ribonucleótidos para ARN y desoxiribonucleótidos para ADN) en las que la posición 3' de la pentosa de un nucleótido está unida por un grupo fosfodiéster a la posición 5' de la pentosa de la siguiente, incluye secuencias de cualquier forma de ácido nucleico incluyendo, pero no limitado a, moléculas de ARN, ADN y ADN libre de células. El término "polinucleótido" incluye, sin limitación, polinucleótidos de cadena sencilla y doble.

El término "porción" se usa en la presente en referencia a la cantidad de información de secuencia de moléculas de ácidos nucleicos fetales y maternos en una muestra biológica que en suma ascienden a menos de la información de la secuencia de <1 genoma humano.

El término "muestra de ensayo" en la presente se refiere a una muestra que comprende una mezcla de ácidos nucleicos que comprende al menos una secuencia de ácidos nucleicos cuyo número de copias se sospecha que ha sufrido variación. Los ácidos nucleicos presentes en la muestra de ensayo son referidos como "ácidos nucleicos de ensayo".

El término "muestra calificada" en la presente se refiere a una muestra que comprende una mezcla de ácidos nucleicos que están presentes en un número de copias conocido con los que los ácidos nucleicos en la muestra de ensayo se comparan, y es una muestra que es normal, es decir, no aneuploide, para la secuencia de interés, por ejemplo una muestra calificada usada para identificar un cromosoma normalizador para el cromosoma 21 es una muestra que no es una muestra de trisomía 21.

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

El término "ácido nucleico calificado" se usa de manera intercambiable con "secuencia calificado" es una secuencia contra la que se compara la cantidad de una secuencia de ensayo o ácido nucleico de ensayo. Una secuencia calificada es una presente en una muestra biológica preferiblemente en una representación conocida, es decir, la cantidad de una secuencia calificada se conoce. Una "secuencia calificada de interés" es una secuencia calificada para la que la cantidad se conoce en una muestra calificada, y es una secuencia que está asociada con una diferencia en la representación de secuencia en un individuo con una condición médica.

El término "secuencia de interés" en la presente se refiere a una secuencia de ácidos nucleicos que está asociada con una diferencia en la representación de secuencia en individuos sanos frente a enfermos. Una secuencia de interés puede ser una secuencia en un cromosoma que está tergiversada, es decir, sobre- o sub-representado, en una enfermedad o condición genética. Una secuencia de interés puede también ser una porción de un cromosoma, o un cromosoma. Por ejemplo, una secuencia de interés puede ser un cromosoma que está sobre-representado en una condición de aneuploidía, o in gen que codifica un supresor tumoral que está sub-representado en un cáncer. Las secuencias de interés incluyen secuencias que estas sobre- o sub-representadas en la población total, o una subpoblación de células de un sujeto. Una "secuencia calificada de interés" es una secuencia de interés en una muestra calificada. Una "secuencia de ensayo de interés" es una secuencia de interés en una muestra de ensayo.

El término "pluralidad de ácidos nucleicos objetivo polimórficos" en la presente se refiere a un número de secuencias de ácidos nucleicos que comprende cada una al menos un sitio polimórfico de tal forma que al menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40 o más sitios polimórficos diferentes se amplifican de los ácidos nucleicos objetivo polimórficos para identificar y/o cuantificar alelos fetales presentes en las muestras maternas que comprenden ácidos nucleicos fetales y maternos.

El término "secuencia normalizadora" en la presente se refiere a una secuencia que muestra una variabilidad en el número de etiquetas de secuencia que están mapeadas para ella entre las muestras y ejecuciones de secuenciación que mejor se aproxima a la de la secuencia de interés para la que se usa como un parámetro normalizador, y que puede diferenciar mejor una muestra afectada de una o más muestras no afectadas. Un "cromosoma normalizador" es un ejemplo de "secuencia normalizadora".

El término "diferenciabilidad" en la presente se refiere a la característica de un cromosoma normalizador que permite distinguir una o más muestras no afectadas, es decir, normales de una o más muestras afectadas, es decir, aneuploides.

El término "grupo de cromosomas" en la presente se refiere a dos o más cromosomas.

El término "dosis de secuencia" en la presente se refiere a un parámetro que relaciona la densidad de etiqueta de secuencia de una secuencia de interés con la densidad de etiqueta de una secuencia normalizadora. Un a "dosis de secuencia de ensayo" es un parámetro que relaciona la densidad de etiqueta de secuencia de una secuencia de interés, por ejemplo cromosoma 21, con la de una secuencia normalizadora, por ejemplo cromosoma 9, determinada en una muestra de ensayo. De manera similar, una "dosis de secuencia calificada" es un parámetro que relaciona la densidad de etiqueta de secuencia de una secuencia de interés con la de una secuencia normalizadora determinada en una muestra calificada.

El término "densidad de etiqueta de secuencia" en la presente se refiere al número de lecturas de secuencia que se mapean para una secuencia del genoma de referencia, por ejemplo la densidad de etiqueta de secuencia para el cromosoma 21 es el número de lecturas de secuencia generadas por el método de secuenciación que se mapean para el cromosoma 21 del genoma de referencia. El término "proporción de densidad de etiqueta de secuencia" en la presente se refiere a la proporción del número de etiquetas de secuencia que se mapean para un cromosoma del genoma de referencia, por ejemplo cromosoma 21, para la longitud del cromosoma 21 del genoma de referencia.

El término "parámetro" en la presente se refiere a un valor numérico que caracteriza un conjunto de datos cuantitativos y/o una relación numérica entre conjuntos de datos cuantitativos. Por ejemplo, una proporción (o función de una proporción) entre el número de etiquetas de secuencia mapeadas para un cromosoma y la longitud del cromosoma para el que las etiquetas se mapean, es un parámetro.

El término "valor límite" y "valor límite calificado" en la presente se refieren a cualquier número que se calcula usando un conjunto de datos de calificación y sirve como un límite de diagnóstico de una variación del número de copias, por ejemplo una aneuploidía, en un organismo. Si el límite se excede por los resultados obtenidos de la práctica de la invención, un sujeto puede ser diagnosticado con una variación del número de copias, por ejemplo trisomía 21.

El término "lectura" se refiere a una secuencia de ADN de longitud suficiente (por ejemplo, al menos alrededor de 30 bp) que se puede usar para identificar una secuencia o región más grande, por ejemplo, que puede ser alineada y asignada específicamente a un cromosoma o región genómica o gen.

El término "etiqueta de secuencia" se usa en la presente de manera intercambiable con el término "etiqueta de secuencia mapeada" para referirse a una lectura de secuencia que ha sido asignada específicamente, es decir mapeada, a una secuencia más grande, por ejemplo un genoma de referencia, por alineamiento. Las etiquetas de secuencia mapeadas son mapeadas únicamente para un genoma de referencia, es decir son asignadas a una localización única para el genoma de referencia. Las etiquetas que pueden ser mapeadas para más de una localización en un genoma de referencia, es decir etiquetas que no mapean únicamente, no se incluyen en el análisis.

Los términos "alineada", "alineamiento" o "alinear" se refieren a una o más secuencias que se identifican como una equivalencia en términos del orden de sus moléculas de ácidos nucleicos para una secuencia conocida de un genoma de referencia. Dicha alineamiento se puede hacer manualmente o por un algoritmo informático, los ejemplos incluyendo el programa de ordenador Efficient Local Alignment of Nucleotide Data (ELAND) distribuido como parte de la línea Illumina Genomics Analysis. La equivalencia de una lectura de secuencia en la alineamiento puede ser un 100% de equivalencia de secuencia o menor del 100% (equivalencia no perfecta).

El término "genoma de referencia" se refiere a cualquier secuencia del genoma conocida particular, ya sea

parcial o completa, de cualquier organismo o virus que pueda usarse para referenciar secuencias identificadas de un sujeto. Por ejemplo, un genoma de referencia usado para sujetos humanos así como muchos otros organismos se encuentra en el National Center for Biotechnology Information en www.ncbi.nlm.nih.gov. Un "genoma" se refiere a la información genética completa de un organismo o virus, expresada en secuencias de ácidos nucleicos.

El término "genoma de secuencias objetivo artificial" en la presente se refiere a una agrupación de secuencias conocidas que abarca alelos de sitios polimórficos conocidos. Por ejemplo, un "genoma de referencia SNP" es un genoma de secuencias objetivo artificial que comprende una agrupación de secuencias que abarca alelos de SNPs conocidas.

El término "secuencia clínicamente relevante" en la presente se refiere a una secuencia de ácidos nucleicos que se sabe o es sospechosa de estar asociada o implicada con una condición genética o enfermedad. Determinar la ausencia o presencia de una secuencia clínicamente relevante puede ser útil para determinar un diagnóstico o confirmar un diagnóstico de una condición médica, o proporcionar un pronóstico para el desarrollo de una enfermedad.

El término "derivado" cuando se usa en el contexto de un ácido nucleico o mezcla de ácidos nucleicos, en la presente se refiere al medio por el que se obtiene el ácido(s) nucleico de una fuente de la que es originario. Por ejemplo, en una realización, una mezcla de ácidos nucleicos que se deriva de dos genomas diferentes significa que los ácidos nucleicos, por ejemplo ADN libre de células, fueron liberados naturalmente por las células a través de procesos de origen natural como necrosis o apoptosis. En otra realización, una mezcla de ácidos nucleicos que se deriva de dos genomas diferentes significa que los ácidos nucleicos se extrajeron de dos tipos diferentes de células de un sujeto.

El término "muestra materna" se refiere en la presente a una muestra biológica obtenida de un sujeto embarazado, por ejemplo una mujer.

El término "muestra materna original" en la presente se refiere a una muestra biológica obtenida de un sujeto embarazado, por ejemplo una mujer, que sirve como la fuente de la que se extrae una porción para amplificar ácidos nucleicos objetivo polimórficos. La "muestra original" puede ser cualquier muestra obtenida de un sujeto embarazado, y las fracciones procesadas de la misma, por ejemplo una muestra de ADN libre de células purificada extraída de una muestra de plasma materno.

El término "fluido biológico" en la presente se refiere a un líquido tomado de una fuente biológica e incluye, por ejemplo, sangre, suero, plasma, esputo, líquido de lavado, fluido cerebroespinal, orina, semen, sudor, lágrimas, saliva y similares. Como se usa en la presente, los términos "sangre", "plasma" y "suero" abarcan expresamente fracciones o porciones procesadas de los mismos. De manera similar, cuando una muestra se toma de una biopsia, hisopo, frotis, etc., la "muestra" abarca expresamente una fracción procesada o porción derivada de la biopsia, hisopo, frotis, etc.

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

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

El término "sustancialmente libre de células" en la presente se refiere a preparaciones de la muestra deseada de la que los componentes que están normalmente asociados con él se han eliminado. Por ejemplo, una muestra de plasma se vuelve esencialmente libre de células eliminando células sanguíneas, por ejemplo, leucocitos, que están normalmente asociados con ella. En algunas realizaciones, se procesan muestras sustancialmente libres para eliminar células que contribuirían de otra manera al material genético deseado que se va a probar para una aneuploidía.

Como se usa en la presente, el término "fracción fetal" se refiere a la fracción de ácidos nucleicos fetales presentes en una muestra que comprende ácido nucleico fetal y materno.

Como se usa en la presente el término "cromosoma" se refiere al portador del gen que lleva la herencia de una célula viva que se deriva de cromatina y que comprende ADN y componentes de proteína (especialmente histonas). El sistema de numeración de cromosomas del genoma humano individuales internacionalmente reconocido convencional se usa en la presente.

Como se usa en la presente, el término "longitud de polinucleótido" se refiere al número absoluto de

moléculas de ácidos nucleicos (nucleótidos) en una secuencia o en una región de un genoma de referencia. El término "longitud de cromosoma" se refiere a la longitud conocida del cromosoma dada en pares de base, por ejemplo proporcionada en el montaje NCBI36/hg18 del cromosoma humano encontrado en la red mundial en [genome.ucsc.edu/cgi-bin/hgTracks?hgid=167155613 &chromInfoPage=](http://genome.ucsc.edu/cgi-bin/hgTracks?hgid=167155613&chromInfoPage=)

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

El término "condición" en la presente se refiere a "condición médica" como un término amplio que incluye todas las enfermedades y trastornos, pero que también puede incluir lesiones y situaciones de salud normales, como el embarazo, que pueden afectar a la salud de una persona, beneficiarse de la asistencia médica, o tener implicaciones para tratamientos médicos.

DESCRIPCION

El método descrito en la presente es un método de secuenciación que permite la determinación simultánea de la fracción del componente del ácido nucleico fetal menor en una muestra que comprende una mezcla de ácidos nucleicos fetales y maternos. En particular, el método permite la determinación de la fracción de ADN libre de células aportado por un feto a la mezcla de ADN libre de células fetal y materno en una muestra materna, por ejemplo, una muestra de plasma. La diferencia entre la fracción materna y fetal se determina por la aportación relativa de un alelo polimórfico derivado del genoma fetal con la aportación del alelo polimórfico correspondiente derivado del genoma materno. Las secuencias polimórficas se pueden usar en conjunción con ensayos diagnósticos clínicamente relevantes como un control positivo para la presencia de ADN libre de células para resaltar resultados de falsos negativos o falsos positivos derivados de niveles bajos de ADN libre de células por debajo del límite de identificación. El método descrito en la presente es útil a través de un intervalo de estados gestacionales.

Las realizaciones ejemplares del método de la invención se ilustran en las Figuras 1-4 como sigue.

La Figura 1 proporciona un diagrama de flujo de una realización del método de la invención 100 para determinar simultáneamente una aneuploidia fetal y la fracción fetal de ácidos nucleicos fetales en una muestra biológica materna. En el paso 110 se obtiene de un sujeto una muestra de ensayo que comprende una mezcla de ácidos nucleicos fetales y maternos. La muestra es una muestra materna que se obtiene de una mujer embarazada, por ejemplo una mujer embarazada. Se puede usar cualquier muestra biológica materna como fuente de ácidos nucleicos fetales y maternos que estén contenidos en células o sean "libres de células". En algunas realizaciones, es ventajoso obtener una muestra materna que comprenda ácidos nucleicos libres de células, por ejemplo ADN libre de células. Preferiblemente, la muestra biológica materna es una muestra de fluido biológico. Una muestra biológica incluye, como ejemplos no limitativos, muestras de sangre, plasma, suero, sudor, lágrimas, esputo, orina, esputo, flujo del oído, linfa, saliva, fluido cefalorraquídeo, estragos, suspensión de médula ósea, flujo vaginal, lavado transcervical, fluido cerebral, ascitis, leche, secreciones de los tractos respiratorio, intestinal y genitourinario y leucoforesis. En algunas realizaciones, la muestra de fluido biológico es una muestra que se puede obtener fácilmente por procedimientos no invasivos, por ejemplo, sangre, plasma, suero, sudor, lágrimas, esputo, orina, esputos, fluido del oído y saliva. En algunas realizaciones, la muestra biológica es una muestra de sangre periférica, o fracciones del plasma y/o suero de la misma. En otras realizaciones la muestra es una mezcla de dos o más muestras biológicas, por ejemplo, una muestra biológica puede comprender dos o más de una muestra de fluido biológico. Como se usa en la presente, los términos "sangre", "plasma" y "suero" abarcan expresamente fracciones o porciones procesadas de los mismos. En algunas realizaciones, la muestra biológica se procesa para obtener una fracción de la muestra, por ejemplo plasma, que contiene la mezcla de ácidos nucleicos fetales y maternos. En algunas realizaciones, la mezcla de ácidos nucleicos fetales y maternos se procesa adicionalmente a partir de la fracción de la muestra, por ejemplo plasma, para obtener una muestra que comprende una mezcla purificada de ácidos nucleicos fetales y maternos, por ejemplo ADN libre de células. Los ácidos nucleicos libres de células, incluyendo del ADN libre de células, se pueden obtener por varios métodos conocidos en la técnica a partir de muestras biológicas que incluyen, pero no están limitadas a, plasma, suero y orina (Fan et al., Proc Natl Acad Sci 105:16266-16271 [2008]; Koide et al., Prenatal Diagnosis 25:604-607 [2005]; Chen et al., Nature Med. 2: 1033-1035 [1996]; Lo et al., Lancet 350: 485-487 [1997]). Para separar el ADN de células libres de las células, se pueden usar métodos de fraccionamiento, centrifugación (por ejemplo centrifugación en gradiente de densidad), precipitación específica del ADN, o clasificación de células de alto rendimiento y/o separación. Hay disponibles kits comercialmente disponibles para la separación manual o automática de ADN libre de células (Roche Diagnostics, Indianapolis, IN, Qiagen, Valencia, CA, Macherey-Nagel, Duren, DE). En algunas situaciones, puede ser ventajoso fragmentar las moléculas de ácidos nucleicos en la muestra de ácido nucleico. La fragmentación puede ser aleatoria, o puede ser específica, como se consigue, por ejemplo, usando digestión con endonucleasas de restricción. Los métodos para la fragmentación aleatoria son bien conocidos en la técnica, e incluyen, por ejemplo, digestión con ADNasa limitada, tratamiento alcalino y corte físico. En una realización, los ácidos nucleicos de la muestra se

obtienen de ADN libre de células, que no está sometido a fragmentación. En otras realizaciones, los ácidos nucleicos de la muestra se obtienen como ADN genómico, que está sometido a fragmentación en fragmentos de aproximadamente 500 o más pares de bases, y a los que se pueden aplicar fácilmente métodos de NGS.

En el paso **120 (Figura 1)** la mezcla de ácidos nucleicos presente en la muestra se enriquece para ácidos nucleicos objetivo polimórficos que comprenden cada uno un sitio polimórfico. En algunas realizaciones, los ácidos nucleicos que se enriquecen son ADN libre de células. Los ácidos nucleicos objetivo son segmentos de material genético que se sabe que comprenden al menos un sitio polimórfico. En algunas realizaciones, los ácidos nucleicos objetivo comprenden un SNP. En otras realizaciones, los ácidos nucleicos objetivo comprenden una STR. El enriquecimiento de una mezcla de ácidos nucleicos fetales y maternos comprende amplificar secuencias objetivo de una porción de ácidos nucleicos contenidos en la muestra materna original, y combinar parte de o el producto amplificado completo con el resto de la muestra materna original. En el paso **130**, se secuencian al menos una porción de la mezcla enriquecida, se identifican las diferencias de secuencia derivadas de la naturaleza polimórfica de las secuencias objetivo, y la contribución relativa de las de las secuencias polimórficas derivadas del genoma fetal, es decir la fracción fetal, se determina en el paso **140**. En algunas realizaciones, la muestra materna original es una muestra de fluido biológico, por ejemplo plasma. En otras realizaciones, la muestra materna original es una fracción procesada de plasma que comprende ADN libre de células fetal y materno purificado.

Los sitios polimórficos que están contenidos en los ácidos nucleicos objetivo incluyen sin limitación polimorfismos de nucleótido simples (SNPs) SNPs en tándem, delecciones o inserciones multi-base a pequeña escala, denominadas IN-DELS (también denominadas polimorfismos de delección inserción o DIPs), Polimorfismos Multi-Nucleótido (MNP) y Repeticiones Cortas en Tándem (STRs). Los sitios polimórficos que están abarcados por el método de la invención están localizados en cromosomas autosómicos, permitiendo de esta manera la determinación de la fracción fetal independientemente del sexo del feto. Cualquier sitio polimórfico que pueda ser abarcado por las lecturas generadas por los métodos de secuenciación descritos en la presente se pueden usar para determinar simultáneamente la fracción fetal y la presencia o ausencia de una aneuploidía en una muestra materna.

En una realización, la mezcla de ácidos nucleicos fetales y maternos en la muestra está enriquecida para los ácidos nucleicos que comprenden al menos un SNP. En algunas realizaciones, cada ácido nucleico objetivo comprende una SNP individual, es decir una. Las secuencias de ácidos nucleicos objetivo que comprenden SNPs están disponibles en bases de datos públicamente accesibles incluyendo, pero no limitado a la Human SNP Database en la dirección web de la red mundial wi.mit.edu, la NCBI dbSNP Home Page en la dirección web de la red mundial ncbi.nlm.nih.gov, la dirección web de la red mundial lifesciences.perkinelmer.com, la base de datos Celera Human SNP en la dirección web de la red mundial celera.com, la base de datos de SNP del Genome Analysis Group (GAN) en la dirección web de la red mundial gan.iarc.fr. En una realización, los SNPs elegidos para enriquecer el ADN libre de células fetal y materno se seleccionan del grupo de 92SNPs de identificación de individuos (IISNPs) descrito por Pakstis et al. (Pakstis et al. Hum Genet 127:315-324 [2010]), que ha demostrado tener una variación muy pequeña en la frecuencia a través de las poblaciones ($F_{st} \geq 0,06$), y ser altamente informativa en todo el mundo teniendo una heterocigosidad media de $\geq 0,4$. Los SNPs que están abarcados por el método de la invención incluyen SNPs ligados y no ligados. Cada ácido nucleico objetivo comprende al menos un sitio polimórfico, por ejemplo un SNP individual, que difiere del presente en otro ácido nucleico objetivo para generar un panel de sitios polimórficos, por ejemplo SNPs, que contienen un número suficientes de sitios polimórficos de los cuales al menos 1, al menos 2, al menos 3, al menos 4, al menos 5, al menos 6, al menos 7, al menos 8, al menos 9, al menos 10, al menos 11, al menos 12, al menos 13, al menos 14, al menos 15, al menos 16, al menos 17, al menos 18, al menos 19, al menos 20, al menos 25, al menos 30, al menos 35, al menos 40 o más son informativos. Por ejemplo, un panel de SNPs puede ser configurado para comprender al menos un SNP informativo. En una realización, los SNPs que son objetivos para la amplificación se seleccionan de rs560681, rs1109037, rs9866013, rs13182883, rs13218440, rs7041158, rs740598, rs10773760, rs4530059, rs7205345, rs8078417, rs576261, rs2567608, rs430046, rs9951171, rs338882, rs10776839, rs9905977, rs1277284, rs258684, rs1347696, rs508485, rs9788670, rs8137254, rs3143, rs2182957, rs3739005, y rs530022.

En otras realizaciones, cada ácido nucleico objetivo comprende dos o más SNPs, es decir cada ácido nucleico objetivo comprende SNPs en tándem. Preferiblemente, cada ácido nucleico objetivo comprende dos SNPs en tándem. Los SNPs en tándem son analizados como una única unidad como haplotipos cortos, y se proporcionan en la presente como conjuntos de dos SNPs. Para identificar secuencias de SNP en tándem adecuadas, se puede buscar la base de datos International HapMap Consortium (The International HapMap Project, Nature 426:789-796 [2003]). La base de datos está disponible en la red mundial en hapmap.org. En una realización, los SNPs en tándem que son objetivo para la amplificación se seleccionan de los siguientes conjuntos de parejas de tándem de SNPs rs7277033-rs2110153; rs2822654-rs1882882; rs368657-rs376635; rs2822731-rs2822732; rs1475881-rs7275487; rs1735976-rs2827016; rs447340-rs2824097; rs418989- rs13047336; rs987980- rs987981; rs4143392-rs4143391; rs1691324-rs13050434; rs11909758-rs9980111; rs2826842-rs232414; rs1980969-rs1980970; rs9978999-rs9979175; rs1034346-rs12481852; rs7509629-rs2828358; rs4817013-rs7277036; rs9981121-rs2829696; rs455921-rs2898102; rs2898102-rs458848; rs961301-rs2830208; rs2174536-rs458076; rs11088023-rs11088024; rs1011734-rs1011733; rs2831244-rs9789838; rs8132769-rs2831440; rs8134080-rs2831524; rs4817219-rs4817220; rs2250911-rs2250997; rs2831899-rs2831900; rs2831902-rs2831903; rs11088086-rs2251447; rs2832040-rs11088088; rs2832141-

rs2246777; rs2832959-rs9980934; rs2833734-rs2833735; rs933121-rs933122; rs2834140-rs12626953; rs2834485-rs3453; rs9974986-rs2834703; rs2776266-rs2835001; rs1984014-rs1984015; rs7281674-rs2835316; rs13047304-rs13047322; rs2835545-rs4816551; rs2835735-rs2835736; rs13047608-rs2835826; rs2836550-rs2212596; rs2836660-rs2836661; rs465612-rs8131220; rs9980072-rs8130031; rs418359-rs2836926; rs7278447-rs7278858; rs385787-rs367001; rs367001-rs386095; rs2837296-rs2837297; y rs2837381-rs4816672.

En otra realización, la mezcla de ácidos nucleicos fetales y maternos en la muestra está enriquecida para los ácidos nucleicos objetivo que comprenden al menos una STR. Los loci de STR se encuentran en casi cualquier cromosoma en el genoma y pueden ser amplificados usando una variedad de cebadores de reacción en cadena de polimerasa (PCR). Se han preferido las repeticiones de tetranucleótidos entre los científicos forenses debido a su fidelidad en la amplificación por PCR aunque también se usan algunas repeticiones de tri- y pentanucleótidos. Un listado completo de referencias, hechos e información de secuencias sobre STRs, cebadores de PCR publicados, sistemas múltiplex comunes, y datos de poblaciones relacionados están recopilados en la STRBase, a la que se puede acceder a través de la red mundial en ibm4.carb.nist.gov:8800/dna/home.htm. La información de secuencia del GenBank® (<http://www2.ncbi.nlm.nih.gov/cgi-bin/genbank>) para loci de STR usados comúnmente también es accesible a través de la STRBase. La naturaleza polimórfica de las secuencias de ADN repetidas en tándem que están ampliamente extendidas a través del genoma humano las han convertido en marcadores genéticos importantes para estudios de mapeado de genes, análisis de ligamiento, y pruebas de identidad humana. Debido al alto polimorfismo de las STRs, la mayoría de los individuos serán heterocigóticos, es decir la mayoría de la gente poseerá dos alelos (versiones) de cada uno heredado de cada padre con un número diferente de repeticiones. Por lo tanto, la secuencia de STR fetal heredada no maternalmente diferirá en el número de repeticiones de la secuencia materna. La amplificación de estas secuencias de STR resultará en dos productos de la amplificación principales correspondientes a los alelos maternos (y el alelo fetal heredado maternalmente) y un producto secundario correspondiente al alelo fetal no maternalmente heredado. Se informó primero de esta técnica en el 2000 (Pertl et al., Human Genetics 106:45-49 [2002]) y ha sido desarrollada posteriormente usando identificación simultánea de múltiples regiones diferentes de STR usando PCR en tiempo real (Liu et al., Acta Obstet Gyn Scand 86:535-541 [2007]). Por lo tanto, la fracción de ácido nucleico fetal en una muestra materna también puede determinarse secuenciando polimórficamente los ácidos nucleicos objetivo que comprenden STRs, que varían entre individuos en el número de unidades repetidas en tándem entre alelos. En una realización, la determinación simultánea de aneuploidía y fracción fetal comprende secuenciar al menos una porción de ácidos nucleicos fetales y maternos presentes en una muestra materna que se ha enriquecido para secuencias polimórficas que comprende STRs. Dado que el tamaño del ADN libre de células fetal es <300 bp, las secuencias polimórficas comprenden miniSTR, que pueden ser amplificadas para generar amplicones que son de longitudes de alrededor del tamaño de los fragmentos de ADN fetal circulante. El método puede usar uno o una combinación de cualquier número de miniSTRs informativos para determinar la fracción del ácido nucleico fetal. Por ejemplo, cualquiera o una combinación de cualquier número de miniSTRs, por ejemplo se pueden usar los miniSTRs divulgados en la Tabla 22. En una realización, la fracción de ácido nucleico fetal en una muestra materna se realiza usando un método que incluye determinar el número de copias del ácido nucleico materno y fetal presente en la muestra materna amplificando al menos un miniSTR autosómico elegido de CSF1PO, FGA, TH01, TPOX, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, Penta D, Penta E, D2S1338, D1S1677, D2S441, D4S2364, D10S1248, D14S1434, D22S1045, D22S1045, D20S1082, D20S482, D18S853, D17S1301, D17S974, D14S1434, D12ATA63, D11S4463, D10S1435, D10S1248, D9S2157, D9S1122, D8S1115, D6S1017, D6S474, D5S2500, D5S2500, D4S2408, D4S2364, D3S4529, D3S3053, D2S1776, D2S441, D1S1677, D1S1627, y D1GATA113. En otra realización, el al menos un miniSTR autosómico es el grupo de mini STRs CSF1PO, FGA, D13S317, D16S539, D18S51, D2S1338, D21S11 y D7S820.

El enriquecimiento de la muestra para los ácidos nucleicos objetivo se consigue por métodos que comprenden amplificar específicamente secuencias de ácidos nucleicos objetivo que comprenden el sitio polimórfico. La amplificación de las secuencias objetivo puede realizarse por cualquier método que use PCR o variaciones del método incluyendo, pero no limitadas a, PCR asimétrico, amplificación dependiente de la helicasa, PCR hot-start, qPCR, PCR en fase sólida y PCR touchdown. Alternativamente, la replicación de las secuencias de ácidos nucleicos objetivo puede obtenerse por métodos independientes de enzimas, por ejemplo síntesis en fase sólida química usando las fosforamiditas. La amplificación de las secuencias objetivo se consigue usando pares de cebadores cada uno capaz de amplificar una secuencia de ácido nucleico objetivo que comprende el sitio polimórfico, por ejemplo SNP, en una reacción de PCR multiplex. Las reacciones de PCR multiplex incluyen combinar al menos 2, al menos tres, al menos 3, al menos 5, al menos 10, al menos 15, al menos 20, al menos 25, al menos 30 al menos 30, al menos 35, al menos 40 o más conjuntos de cebadores en la misma reacción para cuantificar los ácidos nucleicos objetivo amplificados que comprenden al menos dos, al menos tres, al menos 5, al menos 10, al menos 15, al menos 20, al menos 25, al menos 30, al menos 30, al menos 35, al menos 40 o más sitios polimórficos en la misma reacción de secuenciación. Cualquier panel de conjuntos de cebadores se puede configurar para amplificar al menos una secuencia polimórfica informativa.

Amplificación de SNPs

Un número de cebadores de ácidos nucleicos están ya disponibles para amplificar fragmentos de ADN que

contienen los polimorfismos de SNP y sus secuencias se pueden obtener, por ejemplo, de las bases de datos anteriormente identificadas. También se pueden diseñar cebadores adicionales, por ejemplo, usando un método similar al publicado por Vieux, E. F., Kwok, P-Y and Miller, R. D. en BioTechniques (Junio 2002) Vol. 32. Suplemento: "SNPs: Discovery of Marker Disease, pp. 28-32. En una realización, se elige al menos 1, al menos 2, al menos 3, al menos 4, al menos 5, al menos 6, al menos 7, al menos 8, al menos 9, al menos 10, al menos 11, al menos 12, al menos 13, al menos 14, al menos 15, al menos 16, al menos 17, al menos 18, al menos 19, al menos 20, al menos 25, al menos 30, al menos 35, al menos 40 o más conjuntos de cebadores para amplificar un ácido nucleico objetivo que comprende al menos un SNPs informativo en una porción de una mezcla de ADN libre de células fetal y materno. En una realización, los conjuntos de cebadores comprenden cebadores directos e inversos que abarcan al menos un SNP informativo seleccionado de rs560681, rs1109037, rs9866013, rs13182883, rs13218440, rs7041158, rs740598, rs10773760, rs 4530059, rs7205345, rs8078417, rs576261, rs2567608, rs430046, rs9951171, rs338882, rs10776839, rs9905977, rs1277284, rs258684, rs1347696, rs508485, rs9788670, rs8137254, rs3143, rs2182957, rs3739005, y rs530022. Conjuntos ejemplares de cebadores que se usan para amplificar los SNPs divulgados en la presente se proporcionan en el Ejemplo 7 y las Tablas 10 y 11, y se divulgan como SEQ ID NOs:57-112. En otra realización, el grupo de 13 conjuntos de cebadores SEQ ID NOs:57-82 se usa para amplificar un ácido nucleico objetivo cada uno comprendiendo al menos un SNP, por ejemplo un SNP individual, en una porción de una mezcla de ADN libre de células fetal o materno.

En otra realización, se usa al menos un conjunto de cebadores para amplificar un ácido nucleico objetivo comprendiendo cada uno al menos un SNP en tándem, por ejemplo un conjunto de dos SNPs en tándem, en una porción de una mezcla de ADN libre de células fetal y materno. En una realización, los conjuntos son de cebadores que comprenden cebadores directos e inversos que abarcan al menos un SNP en tándem informativo seleccionado de rs7277033-rs2110153; rs2822654-rs1882882; rs368657-rs376635; rs2822731-rs2822732; rs1475881-rs7275487; rs1735976-rs2827016; rs447340-rs2824097; rs418989-rs13047336; rs987980- rs987981; rs4143392- rs4143391; rs1691324- rs13050434; rs11909758-rs9980111; rs2826842-rs232414; rs1980969-rs1980970; rs9978999-rs9979175; rs1034346-rs12481852; rs7509629-rs2828358; rs4817013-rs7277036; rs9981121-rs2829696; rs455921-rs2898102; rs2898102- rs458848; rs961301-rs2830208; rs2174536-rs458076; rs11088023-rs11088024; rs1011734-rs1011733; rs2831244-rs9789838; rs8132769-rs2831440; rs8134080-rs2831524; rs4817219-rs4817220; rs2250911-rs2250997; rs2831899-rs2831900; rs2831902-rs2831903; rs11088086-rs2251447; rs2832040-rs11088088; rs2832141-rs2246777; rs2832959 -rs9980934; rs2833734-rs2833735; rs933121-rs933122; rs2834140-rs12626953; rs2834485-rs3453; rs9974986-rs2834703; rs2776266-rs2835001; rs1984014-rs1984015; rs7281674-rs2835316; rs13047304-rs13047322; rs2835545-rs4816551; rs2835735-rs2835736; rs13047608-rs2835826; rs2836550-rs2212596; rs2836660-rs2836661; rs465612-rs8131220; rs9980072-rs8130031; rs418359-rs2836926; rs7278447-rs7278858; rs385787-rs367001; rs367001-rs386095; rs2837296-rs2837297; y rs2837381-rs4816672.

Los cebadores usados para amplificar las secuencias objetivos que comprenden los SNPs en tándem están diseñados para abarcar ambos sitios de SNP. Conjuntos ejemplares de cebadores usados para amplificar los SNPs en tándem divulgados en la presente se proporcionan en el Ejemplo 12 y se divulgan como SEQ ID Nos:197-310.

La amplificación de ácidos nucleicos objetivo se realiza usando cebadores específicos de la secuencia que permiten la amplificación específica de la secuencia. Por ejemplo, los cebadores de PCR están diseñados para discriminar contra la amplificación de genes o parálogos similares que están en otros cromosomas tomando ventaja de las diferencias de secuencia entre el ácido nucleico objetivo y cualquier parálogo de otros cromosomas. Los cebadores de PCR directos o inversos están diseñados para hibridar cerca del sitio de SNP y para amplificar una secuencia de ácidos nucleicos de longitud suficiente para ser abarcada en las lecturas generadas por métodos de secuenciación masivamente paralelos. Algunos métodos de secuenciación masivamente paralelos requieren que la secuencia de ácidos nucleicos tenga una longitud mínima (bp) para permitir la amplificación por puente que puede usarse antes de la secuenciación. Por lo tanto, los cebadores de PCR usados para amplificar ácidos nucleicos objetivo están diseñados para amplificar secuencias que son de longitud suficiente para ser amplificadas por puente y para identificar SNPs que están abarcados por las lecturas de secuencia. En algunas realizaciones, el primero de dos cebadores en el conjunto de cebadores que comprende el cebador directo y el inverso para amplificar el ácido nucleico objetivo está diseñado para identificar un SNP individual presente dentro de una lectura de secuencia de alrededor de 20bp, alrededor de 25bp, alrededor de 30bp, alrededor de 35bp, alrededor de 40bp, alrededor de 45bp, alrededor de 50bp, alrededor de 55bp, alrededor de 60bp, alrededor de 65bp, alrededor de 70bp, alrededor de 75bp, alrededor de 80bp, alrededor de 85bp, alrededor de 90bp, alrededor de 95bp, alrededor de 100bp, alrededor de 110bp, alrededor de 120bp, alrededor de 130, alrededor de 140bp, alrededor de 150bp, alrededor de 200bp, alrededor de 250bp, alrededor de 300bp, alrededor de 350bp, alrededor de 400bp, alrededor de 450bp, o alrededor de 500bp. Se espera que los avances tecnológicos en tecnologías de secuenciación masivamente paralelas permitan lecturas de extremos individuales mayores de 500bp. En una realización, uno de los cebadores de PCR está diseñado para amplificar SNPs que están abarcados en lecturas de secuencia de 36 bp. El segundo cebador está diseñado para amplificar el ácido nucleico objetivo como un amplicón de longitud suficiente para permitir la amplificación por puente. En una realización, los cebadores de PCR ejemplares están diseñados para amplificar ácidos nucleicos objetivo que contienen un único SNP seleccionado de los SNPs rs560681, rs1109037, rs9866013, rs13182883, rs13218440, rs7041158, rs740598, rs10773760, rs 4530059, rs7205345, rs8078417, rs576261, rs2567608, rs430046, rs9951171, rs338882, rs10776839, rs9905977, rs1277284, rs258684, rs1347696, rs508485,

rs9788670, rs8137254, rs3143, rs2182957, rs3739005, y rs530022. En otras realizaciones, los cebadores directo e inverso están cada uno diseñados para amplificar ácidos nucleicos objetivo que comprenden cada uno un conjunto de dos SNPs en tándem, cada uno estando presente dentro de una lectura de secuencia de alrededor de 20bp, alrededor de 25bp, alrededor de 30bp, alrededor de 35bp, alrededor de 40bp, alrededor de 45bp, alrededor de 50bp, alrededor de 55bp, alrededor de 60bp, alrededor de 65bp, alrededor de 70bp, alrededor de 75bp, alrededor de 80bp, alrededor de 85bp, alrededor de 90bp, alrededor de 95bp, alrededor de 100bp, alrededor de 110bp, alrededor de 120bp, alrededor de 130, alrededor de 140bp, alrededor de 150bp, alrededor de 200bp, alrededor de 250bp, alrededor de 300bp, alrededor de 350bp, alrededor de 400bp, alrededor de 450bp, o alrededor de 500bp. En una realización, al menos uno de los cebadores está diseñado para amplificar el ácido nucleico objetivo que comprende un conjunto de dos SNPs en tándem como un amplicón de longitud suficiente para permitir la amplificación por puente.

Los SNPs, Los SNPs individuales o en tándem, están contenidos en amplicones de ácidos nucleicos objetivo amplificados de al menos alrededor de 100bp, al menos alrededor de 150bp, al menos alrededor de 200bp, al menos alrededor de 250bp, al menos alrededor de 300bp, al menos alrededor de 350bp, o al menos alrededor de 400bp. En una realización, los ácidos nucleicos objetivo que comprenden un sitio polimórfico, por ejemplo un SNP, son amplificados como amplicones de al menos alrededor de 110 bp, y que comprenden un SNP dentro de 36 bp del extremo 3' o 5' del amplicón. En otra realización, los ácidos nucleicos objetivo que comprenden dos o más sitios polimórficos, por ejemplo dos SNPs en tándem, son amplificados como amplicones de al menos alrededor de 110 bp, y que comprenden el primer SNP dentro de 36 bp del extremo 3' del amplicón, y/o el segundo SNP dentro de 36 bp del extremo 5' del amplicón.

Amplificación de STRs

Hay disponible una variedad de cebadores de ácidos nucleicos para amplificar fragmentos de ADN que contienen las STRs y sus secuencias se pueden obtener, por ejemplo, de las bases de datos identificadas anteriormente. Se han usado varios amplicones de PCR dimensionados para distinguir las distribuciones de tamaño respectivas de las especies de ADN fetal y materno, y han mostrado que las moléculas de ADN fetal en el plasma de mujeres embarazadas son generalmente más cortas que las moléculas de ADN materno (Chan et al., Clin Chem 50:8892 [2004]). El fraccionamiento del tamaño del ADN fetal circulante ha confirmado que la longitud media de los fragmentos de ADN fetal circulante es <300 bp, mientras que se ha estimado que el del ADN materno es de entre alrededor de 0,5 y 1Kb (Li et al., Clin Chem, 50: 1002-1011 [2004]). Estos descubrimientos son consistentes con los de Fan et al., que determinó usando NGS que el ADN libre de células fetal raramente es >340bp (Fan et al., Clin Chem 56:1279-1286 [2010]). El método de la invención abarca determinar la fracción de ácido nucleico fetal en una muestra materna que ha sido enriquecido con ácidos nucleicos objetivo cada uno comprendiendo una miniSTR que comprende cuantificar al menos un alelo fetal y uno materno en una miniSTR polimórfica, que puede ser amplificado para generar amplicones que son de longitudes de alrededor del tamaño de los fragmentos de ADN fetal circulante.

En una realización, el método comprende determinar el número de copias de al menos un alelo fetal y al menos uno materno al menos en una miniSTR polimórfica que se amplifica para generar amplicones que son menores de alrededor 300 bp, menores de alrededor 250 bp, menores de alrededor 200 bp, menores de alrededor 150 bp, menores de alrededor 100 bp o menores de alrededor 50 bp. En otra realización, los amplicones que se generan amplificando las miniSTRs son menores de alrededor de 300 bp. En otra realización, los amplicones que se generan amplificando las miniSTRs son menores de alrededor de 250 bp. En otra realización, los amplicones que se generan amplificando las miniSTRs son menores de alrededor de 200 bp. La amplificación del alelo informativo incluye usar cebadores de miniSTR, que permiten la amplificación de amplicones de tamaño reducido para distinguir alelos de STR que son menores de alrededor de 500 bp, menores de alrededor de 450 bp, menores de alrededor de 400 bp, menores de alrededor de 350 bp, menores de alrededor de 300 pares de bases (bp), menores de alrededor de 250 bp, menores de alrededor de 200 bp, menores de alrededor de 150 bp, menores de alrededor de 100 bp o menores de alrededor de 50 bp. Los amplicones de tamaño reducido generados usando los cebadores de miniSTR son conocidos como miniSTRs que se identifican de acuerdo con el nombre del marcador correspondiente al locus para el que han sido mapeados. En una realización, los cebadores de miniSTR incluyen cebadores de miniSTR que han permitido la máxima reducción de tamaño en el tamaño del amplicón para todos los 13 loci CODIS STR además de D2S1338, Penta D, y pentaE encontrados en los kits de STR comercialmente disponibles (Butler et al., J Forensic Sci 48:1054-1064 [2003]), los loci de miniSTR que no están ligados a los marcadores CODIS se describen por Coble y Butler (Coble and Butler, J Forensic Sci 50:43-53 [2005]), y otros miniSTRs se han caracterizado en el NIST. La información referente a los miniSTRs caracterizados en el NIST es accesible a través de la red mundial en cstl.nist.gov/biotech/strbase/newSTRs.htm. Cualquier par o una combinación de dos o más pares de cebadores de miniSTR se pueden usar para amplificar al menos un miniSTR. Por ejemplo, se selecciona al menos un conjunto de cebadores del conjunto CSF1PO_F (SEQ ID NO:81) y CSF1PO_R (SEQ ID NO:82), conjunto FGA_F (SEQ ID NO:83) y FGA_R (SEQ ID NO:84), conjunto TH01_F (SEQ ID NO:85) y TH01_R (SEQ ID NO:86), conjunto TPOX_F (SEQ ID NO:87) y TPOX_R (SEQ ID NO:88), conjunto vWA_F (SEQ ID NO:89) y vWA_R (SEQ ID NO:90), conjunto D3S1358_F (SEQ ID NO:91) y D3S1358_R (SEQ ID NO:92), conjunto D5S818_F (SEQ ID NO:93) y D5S818_R (SEQ ID NO:94), conjunto D7S820_F (SEQ ID NO:95) y D7S820_R (SEQ ID NO:96), conjunto D7S820_F (SEQ ID NO:97) y D7S820_R (SEQ ID NO:98), conjunto D13S317_F (SEQ ID NO:99) y D13S317_R (SEQ ID NO:100),

conjunto D16S539_F (SEQ ID NO:101) y D16S539_R (SEQ ID NO:102), conjunto D18S51_F (SEQ ID NO:103) y D18S51_R (SEQ ID NO:104), conjunto D21S11_F (SEQ ID NO:105) y D21S11_R (SEQ ID NO:106), conjunto D2S1338_F (SEQ ID NO:107) y D2S1338_R (SEQ ID NO:108), conjunto Penta D_F (SEQ ID NO:109) y Penta D_R (SEQ ID NO:110), conjunto Penta E_F (SEQ ID NO:111) y Penta E_R (SEQ ID NO:112), conjunto (D22S1045_F; SEQ ID NO:113) y D22S1045_F (SEQ ID NO:114), conjunto D20S1082_R (SEQ ID NO:115) y D20S1082_F (SEQ ID NO:116), conjunto D20S482_R (SEQ ID NO:117) y D20S482_F (SEQ ID NO:118), conjunto D18S853_R (SEQ ID NO:119) y D18S853_F (SEQ ID NO:120), conjunto D17S1301_F (SEQ ID NO:121) y D17S1301_R (SEQ ID NO:122), conjunto D17S974_F (SEQ ID NO:123) y D17S974_R (SEQ ID NO:124), conjunto D14S1434_F (SEQ ID NO:125) y D14S1434_R (SEQ ID NO:126), conjunto D12ATA63_F (SEQ ID NO:127) y D12ATA63_R (SEQ ID NO:128), D11S4463_F (SEQ ID NO:129) y D11S4463_R (SEQ ID NO:130), conjunto D10S1435_F (SEQ ID NO:131) y D10S1435_R (SEQ ID NO:132), conjunto D10S1248_F (SEQ ID NO:133) y D10S1248_R (SEQ ID NO:134), conjunto D9S2157_F (SEQ ID NO:135) y D9S2157_R (SEQ ID NO:136), conjunto D9S1122_F (SEQ ID NO:137) y D9S1122_R (SEQ ID NO:138), conjunto D8S1115_F (SEQ ID NO:139) y D8S1115_R (SEQ ID NO:140), conjunto D6S1017_F (SEQ ID NO:141) y D6S1017_R (SEQ ID NO:142), D6S474_F (SEQ ID NO:143) y D6S474_R (SEQ ID NO:144), conjunto D5S2500_F (SEQ ID NO:145) y D5S2500_R (SEQ ID NO:146), conjunto D4S2408_F (SEQ ID NO:147) y D4S2408_R (SEQ ID NO:148), conjunto D4S2364U_F (SEQ ID NO:149) y D4S2364U_R (SEQ ID NO:150), conjunto D3S452_F (SEQ ID NO:151) y D3S452_R (SEQ ID NO:152), conjunto D3S3053_F (SEQ ID NO:153) y D3S3053_R (SEQ ID NO:154), conjunto D2S1776_F (SEQ ID NO:155) y D2S1776_R (SEQ ID NO:156), conjunto D2S441_F (SEQ ID NO:157) y D2S441_R (SEQ ID NO:158), conjunto D1S1677_F (SEQ ID NO:159) y D1S1677_R (SEQ ID NO:160), conjunto D1S1627_F (SEQ ID NO:161) y D1S1627_R (SEQ ID NO:162), y conjunto D1GATA113_F (SEQ ID NO:163) y D1GATA113_R (SEQ ID NO:164).

El enriquecimiento de la muestra se obtiene amplificando ácidos nucleicos objetivo contenidos en una porción de la mezcla de ácidos nucleicos fetales y maternos en la muestra original, y combinando al menos una porción o todo el producto amplificado con el resto de la muestra no amplificada original. El enriquecimiento comprende amplificar los ácidos nucleicos objetivo que están contenidos en una porción de la muestra de fluido biológico. En una realización, la muestra que es enriquecida es la fracción de plasma de una muestra de sangre (ver **Figura 2**). Por ejemplo, una porción de una muestra de plasma materna original se usa para amplificar secuencias de ácidos nucleicos objetivo. Posteriormente, algo o todo el producto amplificado se combina con el resto de la muestra de plasma original no amplificada enriqueciéndola de esta manera (ver Ejemplo 10). En otra realización, la muestra que se enriquece en la muestra de ADN libre de células purificada que se extrae del plasma (ver **Figura 3**). Por ejemplo, el enriquecimiento comprende amplificar los ácidos nucleicos objetivo que están contenidos en una porción de una muestra original de mezcla purificada de ácidos nucleicos fetales y maternos, por ejemplo ADN libre de células que ha sido purificada de una muestra de plasma materno, y combinar posteriormente algo o todo el producto amplificado con el resto de la muestra purificada original no amplificada (ver Ejemplo 9). En otra realización, la muestra que es enriquecida es una muestra de biblioteca de secuenciación preparada de una mezcla purificada de ácidos nucleicos fetales y maternos (ver **Figura 4**). Por ejemplo, el enriquecimiento comprende amplificar los ácidos nucleicos objetivo que están contenidos en una porción de una muestra original de mezcla purificada de ácidos nucleicos fetales y maternos, por ejemplo ADN libre de células, que ha sido purificada de una muestra de plasma materno, preparar una primera biblioteca de secuenciación de secuencias de ácidos nucleicos no amplificados, preparar una segunda biblioteca de secuenciación de ácidos nucleicos polimórficos amplificados, y posteriormente combinar algo o toda la segunda biblioteca de secuenciación con algo o todo de la primera biblioteca de secuenciación (ver Ejemplo 8). La cantidad de producto amplificado que se usa para enriquecer la muestra original se selecciona para obtener suficiente información de secuenciación para determinar tanto la presencia como la ausencia de aneuploidía y la fracción fetal de la misma ejecución de secuenciación. Al menos alrededor del 3%, al menos alrededor del 5%, al menos alrededor del 7%, al menos alrededor del 10%, al menos alrededor del 15%, al menos alrededor del 20%, al menos alrededor del 25%, al menos alrededor del 30% o más del número total de etiquetas de secuencia obtenidas de la secuenciación se mapean para determinar la fracción fetal.

En el paso **130 (Figura 1)**, se secuencian la mezcla enriquecida de ácidos nucleicos fetales y maternos. La información de secuencia que se necesita para la determinación simultánea de aneuploidía y fracción fetal se puede obtener usando cualquiera de los métodos de secuenciación de ADN conocidos. En una realización, el método descrito en la presente emplea tecnología de secuenciación de próxima generación (NGS) en la que se secuencian plantillas de ADN amplificadas clonalmente o moléculas de ADN individuales de una forma masivamente paralela dentro de una célula de flujo (por ejemplo como se describe en Volkerding et al. Clin Chem 55:641-658 [2009]; Metzker M Nature Rev 11:31-46[2010]). Además de la información de la secuencia de alto rendimiento, la NGS proporciona información cuantitativa digital, en la que cada lectura de secuencia es una "etiqueta de secuencia" que se puede contar que representa una plantilla de ADN clonal individual o una molécula de ADN individual. Esta cuantificación permite al NGS expandir el concepto digital de contar moléculas de ADN libre de células (Fan et al., Proc Natl Acad Sci U S A 105:16266-16271 [2008]; Chiu et al., Proc Natl Acad Sci U S A 2008;105:20458-20463 [2008]). Las tecnologías de secuenciación de NGS incluyen pirosecuenciación, secuenciación por síntesis con terminadores de colorante reversibles, secuenciación por ligadura de sonda de oligonucleótidos y secuenciación en tiempo real.

Algunas de las tecnologías de secuenciación están disponibles comercialmente, como la plataforma de

secuenciación por hibridación de Affymetrix Inc. (Sunnyvale, CA) y las plataformas de secuenciación por síntesis de Life Sciences (Bradford, CT), Illumina/Solexa (Hayward, CA) y Helicos Biosciences (Cambridge, MA), y la plataforma de ligadura por secuenciación de Applied Biosystems (Foster City, CA), como se describe a continuación. Además de la secuenciación de moléculas individuales realizada usando la secuenciación por síntesis de Helicos Biosciences, se abarcan otras tecnologías de secuenciación de moléculas individuales por el método de la invención e incluyen la tecnología SMRT™ de Pacific Biosciences, la tecnología Ion Torrent™, y la secuenciación por nanoporos que está siendo desarrollada por, por ejemplo Oxford Nanopore Technologies.

Aunque el método Sanger automatizado se considera como una tecnología de "primera generación", la secuenciación Sanger que incluye secuenciación Sanger automatizada, también se puede emplear por el método de la invención. Los métodos de secuenciación adicionales que comprenden el uso de desarrollar tecnologías de formación de imágenes de ácidos nucleicos, por ejemplo microscopía de fuerza atómica (AFM) o microscopía electrónica de transmisión (TEM), también se abarcan por el método de la invención. Las tecnologías de secuenciación ejemplares se describen a continuación.

En una realización, la tecnología de secuenciación de ADN que se usa en el método de la invención es la Helicos True Single Molecule Sequencing (tSMS) (por ejemplo, como se describe en Harris T.D. et al., Science 320:106-109 [2008]). En la técnica tSMS, una muestra de ADN se escinde en cadenas de aproximadamente 100 a 200 nucleótidos, y se añade una secuencia polyA al extremo 3' de cada cadena de ADN. Cada cadena es etiquetada por la adición de un nucleótido de adenosina etiquetado fluorescentemente. Las cadenas de ADN son después hibridadas a una célula de flujo, que contiene millones de sitios de captura de oligo-T que se inmovilizan en la superficie de la célula de flujo. Las plantillas pueden estar a una densidad de alrededor de 100 millones de plantillas/cm². La célula de flujo se carga entonces en un instrumento, por ejemplo, secuenciador HeliScope™, y un laser ilumina la superficie de la célula de flujo, revelando la posición de cada plantilla. Una cámara CCD puede mapear la posición de las plantillas en la superficie de la célula de flujo. La etiqueta fluorescente de la plantilla es después escindida y se hace desaparecer. La reacción de secuenciación empieza introduciendo una polimerasa de ADN y un nucleótido etiquetado por fluorescencia. El ácido nucleico oligo-T sirve como un cebador. La polimerasa incorpora los nucleótidos etiquetados al cebador de una manera dirigida a la plantilla. La polimerasa y los nucleótidos no incorporados se eliminan. Las plantillas que tienen incorporación directa del nucleótido etiquetado por fluorescencia se distinguen por formación de imágenes de la superficie de la célula de flujo. Después de la formación de imágenes, un paso de escindido elimina la etiqueta fluorescente, y el proceso se repite con otros nucleótidos etiquetados por fluorescencia hasta que se consigue la longitud de lectura deseada. La información de la secuencia se recoge con cada paso de adición de nucleótidos.

En una realización, la tecnología de secuenciación que se usa en el método de la invención es la secuenciación 454 (Roche) por ejemplo como se describe en Margulies, M. et al. Nature 437:376-380 [2005]). La secuenciación 454 supone dos pasos. En el primer paso, el ADN es cortado en fragmentos de aproximadamente 300-800 pares de base, y los fragmentos son terminados romos. Los adaptadores de oligonucleótidos son entonces ligados a los extremos de los fragmentos. Los adaptadores sirven como cebadores para la amplificación y secuenciación de los fragmentos. Los fragmentos pueden ser unidos a microesferas de captura de ADN, por ejemplo microesferas recubiertas de estreptavidina usando, por ejemplo Adaptador B, que contiene etiqueta de biotina 5'. Los fragmentos unidos a las microesferas se amplifican por PCR con gotitas de una emulsión de aceite-agua. El resultado son copias múltiples de fragmentos de ADN clonalmente amplificados en cada microesfera. En el segundo paso, las microesferas se capturan en pocillos (de tamaño de pico-litros). La pirosecuenciación se realiza en cada fragmento de ADN en paralelo. La adición de uno o más nucleótidos genera una señal de luz que se registra por una cámara CCD en un instrumento de secuenciación. La fuerza de la señal es proporcional al número de nucleótidos incorporado. La polisequenciación hace uso de pirofosfato (PPi) que es liberado en el momento de la adición de nucleótidos. El PPi se convierte a ATP por ATP sulfurilasa en presencia de adenosina 5' fosfosulfato. La luciferasa usa ATP para convertir luciferina en oxiluciferina, y esta reacción genera luz que es distinguida y analizada.

En una realización, la tecnología de secuenciación de ADN que se usa en el método de la invención es la tecnología SOLiD™ (Applied Biosystems). En la secuenciación por ligadura SOLiD™, el ADN genómico es cortado en fragmentos, y los adaptadores se unen a los extremos 5' y 3' de los fragmentos para generar una biblioteca de fragmentos. Alternativamente, se pueden introducir adaptadores internos ligando adaptadores a los extremos 5' y 3' de los fragmentos, circularizando los fragmentos, digiriendo el fragmento circularizado para generar un adaptador interno, y uniendo los adaptadores a los extremos 5' y 3' de los fragmentos resultantes para generar una biblioteca de compañeros emparejadas. Después, se preparan las poblaciones de microesferas clonales en microreactores que contienen microesferas, cebadores, plantilla y componentes de PCR. Después del PCR, las plantillas se desnaturalizan y las microesferas se enriquecen para separar las microesferas con plantillas extendidas. Las plantillas en las microesferas seleccionadas se someten a una modificación 3' que permite el enlace con un portaobjetos de vidrio. La secuencia se puede determinar por hibridación secuencial y ligadura de oligonucleótidos parcialmente aleatorios con una base central determinada (o pares de bases) que se identifica por un fluoróforo específico. Después de que se registra un color, el oligonucleótido ligado es escindido y eliminado y el proceso es después repetido.

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

En una realización, la tecnología de secuenciación de ADN que se usa en el método de la invención es secuenciación por nanoporos (por ejemplo, como se describe en Soni GV y Meller A. Clin Chem 53: 1996-2001 [2007]). Las técnicas de análisis de ADN de secuenciación por nanoporos se están desarrollando industrialmente por una variedad de compañías, incluyendo Oxford Nanopore Technologies (Oxford, Reino Unido). La secuenciación por nanoporos es una tecnología de secuenciación de una sola molécula por la que una única molécula de ADN se secuencia directamente a medida que pasa a través de un nanoporo. Un nanoporo es un agujero pequeño, del orden de 1 nanómetro de diámetro. La inmersión de un nanoporo en un fluido conductor y la aplicación de un potencial (voltaje) a través de él resulta en una ligera corriente eléctrica debido a la conducción de iones a través del nanoporo. La cantidad de corriente que fluye es sensible al tamaño y forma del nanoporo. A medida que una molécula de ADN pasa a través de un nanoporo, cada nucleótido en la molécula de ADN obstruye el nanoporo a un grado diferente, cambiando la magnitud de la corriente a través del nanoporo en grados diferentes. Así, este cambio en la corriente a medida que la molécula de ADN pasa a través del nanoporo representa una lectura de la secuencia de ADN.

En una realización, la tecnología de secuenciación de ADN que se usa en el método de la invención es la matriz de transistor de efecto de campo sensible a sustancias químicas (chemFET) (por ejemplo, como se describe en la Publicación de Solicitud de Patente U.S. Nº 20090026082). En un ejemplo de la técnica, las moléculas de ADN pueden ser colocadas en cámaras de reacción, y las moléculas de la plantilla pueden ser hibridadas a un cebador de secuenciación enlazado a una polimerasa. La incorporación de uno o más trifosfatos en una nueva cadena de ácidos nucleicos en el extremo 3' del cebador de secuenciación puede distinguirse por un cambio en la corriente por un chemFET. Una matriz puede tener múltiples sensores de chemFET. En otro ejemplo, los ácidos nucleicos individuales pueden estar unidos a microesferas, y los ácidos nucleicos pueden amplificarse en la microesfera, y las microesferas individuales pueden transferirse a cámaras de reacción individuales en una matriz de chemFET, con cada cámara teniendo un sensor de chemFET, y los ácidos nucleicos se pueden secuenciar.

En una realización, la tecnología de secuenciación de ADN que se usa en el método de la invención es el método de Halcyon Molecular que usa microscopía electrónica de transmisión (TEM). El método denominado Nano Transferencia Rápida de Colocación de Moléculas Individuales (IMPRNT), comprende utilizar formación de imágenes por microscopía electrónica de transmisión de resolución de átomos individuales de ADN de alto peso molecular (150kb o mayor) etiquetado selectivamente con marcadores de átomos pesados y organizar estas moléculas en películas ultra finas en matrices paralelas ultra-densas (3nm de cadena a cadena) con separación base a base consistente. El microscopio electrónico se usa para formar imágenes de las moléculas en las películas para determinar la posición de marcadores de átomos pesados y para extraer información de la secuencias de bases del ADN. El método se describe adicionalmente en la publicación de patente de PCT WO2009/046445. El método permite secuenciar completamente genomas humanos en menos de diez minutos.

En una realización, la tecnología de secuenciación de ADN es la secuenciación de moléculas individuales Ion Torrent, que empareja tecnología de semiconductores con una química de secuenciación simple para traducir directamente información codificada químicamente (A, C, G, T) en información digital (0, 1) en un chip semiconductor. En la naturaleza, cuando se incorpora un nucleótido en una cadena de ADN por una polimerasa, se libera un ion de hidrógeno como un subproducto. La Ion Torrent usa una matriz de alta densidad de pocillos de micro-mecanizado para realizar este proceso bioquímico de una manera masivamente paralela. Cada pocillo mantiene una molécula de ADN diferente. Por debajo de los pocillos hay una capa sensible a los iones y por debajo de esa un sensor de iones. Cuando se añade un nucleótido, por ejemplo uno C, a una plantilla de ADN y es después incorporado en una cadena de AN, se liberará un ion de hidrógeno. La carga de ese ion cambiará el pH de la solución, que puede ser identificado por el sensor de iones del Ion Torrent. El secuenciador - esencialmente el medidor de pH en estado sólido más pequeño del mundo, designa la base, yendo directamente de la información química a la información digital. El secuenciador Ion personal Genome Machine (PGM™) inunda después secuencialmente el chip con un nucleótido después de otro. Si el siguiente nucleótido que inunda el chip no es una equivalencia, no se registrará cambio de voltaje y no se designará una base. Si hay dos bases idénticas en la cadena de ADN, el voltaje será doble, y el chip registrará dos bases idénticas designadas. La identificación directa permite el registro de la incorporación de nucleótidos en segundos.

Otros métodos de secuenciación incluyen PCR digital y secuenciación por hibridación. La reacción en cadena de polimerasa digital (PCR digital o dPCR) se puede usar para identificar directamente y cuantificar ácidos nucleicos en una muestra. La PCR digital puede realizarse en una emulsión. Los ácidos nucleicos individuales se separan, por ejemplo, en un dispositivo de cámara de microfluidos, y cada ácido nucleico es amplificado individualmente por PCR. Los ácidos nucleicos pueden ser separados de tal manera que haya una media de aproximadamente 0,5 ácidos nucleicos/pocillo, o no más de un ácido nucleico/pocillo. Se pueden usar diferentes sondas para distinguir alelos fetales y alelos maternos. Los alelos se pueden enumerar para determinar el número de copias. En la secuenciación por hibridación, la hibridación comprende poner en contacto la pluralidad de secuencias de polinucleótidos con una pluralidad de sondas de polinucleótidos, en donde cada una de la pluralidad de sondas de polinucleótidos puede estar atada opcionalmente a un sustrato. El sustrato puede ser una superficie plana que comprende una matriz de secuencias de nucleótidos conocidas. El patrón de hibridación a la matriz puede usarse para determinar las secuencias de polinucleótidos presentes en la muestra. En otras realizaciones, cada sonda está atada a una microesfera, por ejemplo, una microesfera magnética o similar. La hibridación a las microesferas puede identificarse y usarse para identificar la pluralidad de secuencias de polinucleótidos dentro de la muestra.

En una realización, el método emplea secuenciación masivamente paralela de millones de fragmentos de ADN usando la secuenciación por síntesis de Illumina y química de secuenciación basada en terminadores reversible (por ejemplo como se describe en Bentley et al., *Nature* 6:53-59 [2009]). El ADN de la plantilla puede ser ADN genómico, por ejemplo ADN libre de células. En algunas realizaciones, se usa el ADN genómico de células aisladas como la plantilla, y se fragmenta en longitudes de varios cientos de pares de bases. En otras realizaciones, se usa ADN libre de células como la plantilla, y no se requiere fragmentación ya que el ADN libre de células existe como fragmentos cortos. Por ejemplo el ADN libre de células fetal circula en el torrente sanguíneo como fragmentos de <300 bp, y se ha estimado que el ADN libre de células materno circula como fragmentos de entre alrededor de 0,5 y 1 Kb (Li et al., *Clin Chem*, 50: 1002-1011 [2004]). La tecnología de secuenciación Illumina se basa en la unión de ADN genómico fragmentado con una superficie plana, ópticamente transparente en la que los anclajes de oligonucleótidos se enlazan. El ADN de la plantilla se repara en los extremos para generar extremos romos 5'-fosforilados, y la actividad de polimerasa del fragmento Klenow se usa para añadir una única base A al extremo 3' de los fragmentos de ADN fosforilados romos. Esta adición prepara los fragmentos de ADN para la ligadura con adaptadores de oligonucleótidos, que tienen un saliente de una única base T en sus extremos 3' para aumentar la eficiencia de la ligadura. Los oligonucleótidos adaptadores son complementarios a los anclajes de las células de flujo. Bajo condiciones de dilución limitativas se añade ADN de plantilla modificado por adaptadores, de una sola cadena a la célula de flujo y se inmovilizan por hibridación a los anclajes. Los fragmentos de ADN unidos se extienden y se amplifican por puente para crear una célula de flujo de secuenciación de ultra-alta densidad con cientos de millones de clústeres, cada uno conteniendo ~1000 copias de la misma plantilla. En una realización, el ADN genómico fragmentado aleatoriamente, por ejemplo ADN libre de células, se amplifica usando PCR antes de que se someta a amplificación de clústeres. Alternativamente, se usa una preparación de biblioteca genómica libre de amplificación, y el ADN genómico fragmentado aleatoriamente, por ejemplo ADN libre de células se enriquece usando sólo la amplificación de clústeres (Kozarewa et al., *Nature Methods* 6:291-295 [2009]). Las plantillas se secuencian usando una tecnología de secuenciación por síntesis de ADN de cuatro colores robusta que emplea terminadores reversibles con colorantes fluorescentes extraíbles. La identificación por fluorescencia de alta sensibilidad se consigue usando excitación por láser y ópticas de reflexión internas totales. Las lecturas de secuencias cortas de alrededor de 20-40 bp, por ejemplo 36 bp, se alinean contra un genoma de referencia enmascarado de repetición y las diferencias genéticas se designan usando software de línea de análisis de datos especialmente desarrollado. Después de la terminación de la primera lectura, las plantillas pueden generarse in situ para permitir una segunda lectura del extremo opuesto de los fragmentos. Así, se usa cualquier secuenciación de extremos emparejados o de único extremo de los fragmentos de ADN de acuerdo con el método. Se realiza la secuenciación parcial de fragmentos de ADN presentes en la muestra, y se cuentan las etiquetas de secuencia que comprenden lecturas de longitud predeterminada, por ejemplo 36 bp, que son mapeadas para un genoma de referencia conocido. En una realización, la secuencia del genoma de referencia es la secuencia NCB136/hg18 que está disponible en la red mundial en genome.ucsc.edu/cgi-bin/hgGateway?org=Human&db=hg18&hgid=166260105). Otras fuentes de información de secuencias públicas incluyen el GenBank, dbEST, dbSTS, EMBL (el Laboratorio de Biología Molecular Europeo), y el DDBJ (el Banco de datos de ADN de Japón). Hay disponibles un número de algoritmos informáticos para alinear secuencias, incluyendo sin limitación (Altschul et al., 1990), BLITZ (MPsrch) (Sturrock & Collins, 1993), FASTA (Person & Lipman, 1988), BOWTIE (Langmead et al., *Genome Biology* 10:R25.1-R25.10 [2009]), o ELAND (Illumina, Inc., San Diego, CA, USA). En una realización, un extremo de las copias expandidas clonalmente de las moléculas de ADN libre de células de plasma se secuencian y procesa por análisis de alineamiento bioinformático para el Illumina Genome Analyzer que usa el software Efficient Large-Scale Alignment of Nucleotide Databases (ELAND).

La longitud de la lectura de secuencia está asociada con la tecnología de secuenciación particular. Los métodos NGS proporcionan lecturas de secuencias que varían de tamaño de decenas a cientos de pares de bases. En algunas realizaciones del método descrito en la presente las lecturas de secuencia son de alrededor de 20bp, de alrededor de 25bp, de alrededor de 30bp, de alrededor de 35bp, de alrededor de 40bp, de alrededor de 45bp, de

alrededor de 50bp, de alrededor de 55bp, de alrededor de 60bp, de alrededor de 65bp, de alrededor de 70bp, de alrededor de 75bp, de alrededor de 80bp, de alrededor de 85bp, de alrededor de 90bp, de alrededor de 95bp, de alrededor de 100bp, de alrededor de 110bp, de alrededor de 120bp, de alrededor de 130, de alrededor de 140bp, de alrededor de 150bp, de alrededor de 200bp, de alrededor de 250bp, de alrededor de 300bp, de alrededor de 350bp, de alrededor de 400bp, de alrededor de 450bp, o de alrededor de 500bp. Se espera que los avances tecnológicos permitan lecturas de extremos individuales de más de 500 bp permitiendo lecturas de más de alrededor de 1000 bp cuando se generan lecturas de extremos emparejados. En una realización, las lecturas de secuencia son 36 bp. Otros métodos de secuenciación que se pueden emplear por el método de la invención incluyen métodos de secuenciación de moléculas individuales que pueden secuenciar moléculas de ácidos nucleicos de >5000 bp. La cantidad masiva de salida de secuencia se transfiere por una línea de análisis que transforma salida de formación de imagen primaria del secuenciador en cuerdas de bases. Un paquete de algoritmos integrados realiza los pasos de transformación de datos primarios del núcleo: análisis de imágenes, puntuación de intensidad, designación de bases y alineamiento.

En una realización, se realiza la secuenciación parcial de fragmentos de ADN presentes en la muestra y se cuentan las etiquetas de secuencia que comprenden lecturas de longitud predeterminada, por ejemplo 36 bp, que mapean para un genoma de referencia conocido. Sólo se cuentan como etiquetas de secuencia las lecturas de secuencia que alinean únicamente con el genoma de referencia. En una realización, el genoma de referencia es la secuencia del genoma de referencia humano NCBI36/ hg 18, que está disponible en la red mundial en genome.ucsc.edu/cgi-bin/hgGateway?org=Human&db=hg18&hgsid=166260105). Otras fuentes de información de secuencias públicas incluyen GenBank, dbEST, dbSTS, EMBL (el Laboratorio de Biología Molecular Europeo), y el DDBJ (el Banco de datos de ADN de Japón). En otra realización el genoma de referencia comprende la secuencia del genoma de referencia humano NCBI36/hg18 y un genoma de secuencias objetivo artificiales, que incluye secuencias objetivo polimórficas, por ejemplo un genoma SNP. El mapeado de las etiquetas de secuencia se consigue comparando la secuencia de la etiqueta con la secuencia del genoma de referencia para determinar el origen cromosómico del molécula de ácidos nucleicos secuenciada (por ejemplo ADN libre de células), y la información de la secuencia genética específica no se necesita. Hay disponibles una variedad de algoritmos informáticos para alinear secuencias, incluyendo sin limitación BLAST (Altschul et al., 1990), BLITZ (MPsrch) (Sturrock & Collins, 1993), FASTA (Person & Lipman, 1988), BOWTIE (Langmead et al., Genome Biology 10:R25.1-R25.10 [2009]), o ELAND (Illumina, Inc., San Diego, CA, USA). En una realización, un extremo de las copias expandidas clonalmente de las moléculas de ADN libre de células del plasma se secuencian y procesa por análisis de alineamiento bioinformático para el Illumina Genome Analyzer que usa el software Efficient Large-Scale Alignment of Nucleotide Databases (ELAND). El análisis de la información de secuenciación para la determinación de aneuploidía puede permitir un pequeño grado de discrepancia (0-2 discrepancias por etiqueta de secuencia) para dar cuenta de los polimorfismos menores que pueden existir entre el genoma de referencia y los genomas en la muestra mezclada. El análisis de la información de secuenciación para la determinación de la fracción fetal puede permitir un pequeño grado de discrepancia dependiendo de la secuencia polimórfica. Por ejemplo, se puede permitir un pequeño grado de discrepancia si la secuencia polimórfica es una STR. En casos donde la secuencia polimórfica es un SNP, todas las secuencias que coinciden exactamente con cualquiera de los dos alelos en el sitio del SNP se cuentan primero y se filtran de las lecturas restantes, para las que se puede permitir un pequeño grado de discrepancia.

En el paso 140, la información de secuencia obtenida en el paso 130 se analiza y se hace la determinación simultánea de la fracción fetal y la determinación de la presencia o ausencia de aneuploidía.

Se obtienen una pluralidad de etiquetas de secuencia por muestra. En algunas realizaciones, se obtienen al menos alrededor de 3×10^6 etiquetas de secuencia al menos alrededor de 5×10^6 etiquetas de secuencia, al menos alrededor de 8×10^6 etiquetas de secuencia, al menos alrededor de 10×10^6 etiquetas de secuencia, al menos alrededor de 15×10^6 etiquetas de secuencia, al menos alrededor de 20×10^6 etiquetas de secuencia, al menos alrededor de 30×10^6 etiquetas de secuencia, al menos alrededor de 40×10^6 etiquetas de secuencia, o al menos alrededor de 50×10^6 etiquetas de secuencia que comprenden entre 20 y 40bp de lecturas de mapear las lecturas para el genoma de referencia por muestra. En una realización, todas las lecturas de secuencia se mapean para todas las regiones del genoma de referencia. En una realización, se cuentan las etiquetas que comprenden lecturas que han sido mapeadas para todas las regiones, por ejemplo todos los cromosomas del genoma de referencia humano, y les determina a aneuploidía fetal, es decir la sobre- o sub- representación de una secuencia de interés, por ejemplo un cromosoma o porción del mismo, en la muestra de ADN mezclado, y se cuentan las etiquetas que comprenden lecturas que se mapean para el genoma de secuencias objetivo artificiales para determinar la fracción fetal. El método no requiere diferenciación entre los genomas maternos y fetales.

Determinación de aneuploidía

La precisión requerida para determinar correctamente si hay presencia o ausencia de una aneuploidía en una muestra se determina en parte por la variación del número de etiquetas de secuencia que mapean para el genoma de referencia entre las muestras de una ejecución de secuenciación (variabilidad inter-cromosómica), y la variación del número de etiquetas de secuencia que mapean para el genoma de referencia en diferentes ejecuciones

de secuenciación (variabilidad inter-secuenciación). Por ejemplo, las variaciones pueden ser particularmente pronunciadas para etiquetas que mapean para secuencias de referencia ricas en GC o pobres en GC. En una realización, el método usa información de secuenciación para calcular la dosis de cromosoma, que cuenta intrínsecamente para la variabilidad acumulada derivada de la variabilidad intercromosómica, inter-secuenciación y dependiente de la plataforma. Las dosis de cromosomas se determinan de la información de la secuenciación, es decir, el número de etiquetas de secuencia, para la secuencia de interés, por ejemplo cromosoma 21, y el número de etiquetas de secuencia para una secuencia normalizadora. La identificación de una secuencia normalizadora se realiza en un conjunto de muestras calificadas que se sabe no contienen una aneuploidía de la secuencia de interés. El diagrama de flujo proporcionado en la **Figura 5** muestra el proceso **500** por el que se identifican las secuencias normalizadoras, por ejemplo cromosomas normalizadores, y se determina la presencia o ausencia de una aneuploidía. En el paso **510**, se obtiene un conjunto de muestras maternas calificadas para identificar secuencias normalizadoras calificadas, por ejemplo cromosomas normalizadores, y para proporcionar valores de varianza para su uso en determinar la identificación estadísticamente significativa de una aneuploidía en las muestras de ensayo. En el paso **510**, se obtiene una pluralidad de muestras biológicas calificadas de una pluralidad de sujetos que se sabe comprenden células que tienen un número de copias normales para cualquiera de las secuencias de interés, por ejemplo un cromosoma de interés, como el cromosoma asociado con una aneuploidía. En una realización, las muestras calificadas se obtienen de madres embarazadas con un feto que se ha confirmado usando medios citogenéticos que tienen un número de copias normal de cromosomas en relación con el cromosoma de interés. Las muestras maternas calificadas biológicas pueden ser muestras de fluidos biológicos, por ejemplo muestras de plasma, o cualquier muestra adecuada como se ha descrito anteriormente que contenga una mezcla de moléculas de ADN libre de células fetal o materno.

En el paso **520**, se secuencia al menos una porción de cada uno de todos los ácidos nucleicos calificados contenidos en las muestras maternas calificadas. Antes de la secuenciación, la mezcla de ácidos nucleicos fetales y maternos, por ejemplo ADN libre de células purificado, se modifica para preparar una biblioteca de secuenciación para generar lecturas de secuencia de entre 20 y 40 bp, por ejemplo 36 bp, que están alineadas a un genoma de referencia, por ejemplo hg18. En algunas realizaciones, las lecturas de secuencia comprenden alrededor de 20bp, alrededor de 25bp, alrededor de 30bp, alrededor de 35bp, alrededor de 40bp, alrededor de 45bp, alrededor de 50bp, alrededor de 55bp, alrededor de 60bp, alrededor de 65bp, alrededor de 70bp, alrededor de 75bp, alrededor de 80bp, alrededor de 85bp, alrededor de 90bp, alrededor de 95bp, alrededor de 100bp, alrededor de 110bp, alrededor de 120bp, alrededor de 130, alrededor de 140bp, alrededor de 150bp, alrededor de 200bp, alrededor de 250bp, alrededor de 300bp, alrededor de 350bp, alrededor de 400bp, alrededor de 450bp, o alrededor de 500bp. Se espera que los avances tecnológicos permitan lecturas de extremos individuales de más de 500bp permitiendo lecturas de más de alrededor de 1000bp cuando se generan lecturas de extremos emparejados. En una realización, las lecturas de secuencia comprenden 36bp. Las lecturas de secuencia están alineadas a un genoma de referencia humano, y las lecturas que se mapean únicamente para el genoma de referencia humana se cuentan como etiquetas de secuencia. En una realización, se obtienen al menos alrededor de 3×10^6 etiquetas de secuencia calificadas, al menos alrededor de 5×10^6 etiquetas de secuencia calificadas, al menos alrededor de 8×10^6 etiquetas de secuencia calificadas, al menos alrededor de 10×10^6 etiquetas de secuencia calificadas, al menos alrededor de 15×10^6 etiquetas de secuencia calificadas, al menos alrededor de 20×10^6 etiquetas de secuencia calificadas, al menos alrededor de 30×10^6 etiquetas de secuencia calificadas, al menos alrededor de 40×10^6 etiquetas de secuencia calificadas, o al menos alrededor de 50×10^6 etiquetas de secuencia calificadas, que comprende lecturas de entre 20 y 40bp se obtienen de lecturas que mapean únicamente para un genoma de referencia.

En el paso **530**, todas las etiquetas obtenidas de la secuenciación de ácidos nucleicos en las muestras maternas calificadas se cuentan para determinar una densidad de etiqueta de secuencia calificada. En una realización la densidad de etiqueta de secuencia se determina como el número de etiquetas de secuencia calificadas mapeadas para la secuencia de interés en el genoma de referencia. En otra realización, la densidad de etiqueta de secuencia calificada se determina como el número de etiquetas de secuencia calificadas mapeadas para una secuencia de interés normalizada a la longitud de la secuencia calificada de interés para la que se mapean. Las densidades de etiqueta de secuencia que se determinan como una proporción de la densidad de etiqueta en relación a la longitud de la secuencia de interés se refieren en el presente como proporciones de densidad de etiqueta. La normalización a la longitud de la secuencia de interés no se requiere, y puede ser incluida como un paso para reducir el número de dígitos en un número para simplificarla para la interpretación humana. Como todas las etiquetas de secuencia calificadas se mapean y cuentan en cada una de las muestras calificadas, la densidad de etiqueta de secuencia para una secuencia de interés, por ejemplo cromosoma de interés, se determina en las muestras calificadas, como lo son las densidades de etiquetas de secuencia para las secuencias adicionales de las que se identifican posteriormente las secuencias normalizadoras, por ejemplo cromosomas. En una realización, la secuencia de interés es un cromosoma que está asociado con una aneuploidía cromosómica, por ejemplo cromosoma 21, y la secuencia normalizadora calificada es un cromosoma que no está asociado con una aneuploidía cromosómica y cuya variación en la densidad de etiqueta de secuencia se aproxima mejor a la del cromosoma 21. Por ejemplo, una secuencia normalizadora calificada es una secuencia que tiene la variabilidad más pequeña. En algunas realizaciones, la secuencia normalizadora es una secuencia que distingue mejor una o más muestras calificadas de una o más muestras afectadas, es decir, la secuencia normalizadora es una secuencia que tiene la diferenciabilidad más alta. El nivel de diferenciabilidad puede determinarse como una diferencia estadística entre las

dosis de cromosomas en una población de muestras calificadas y las dosis de cromosomas en una o más muestras de ensayo. En otra realización, la secuencia de interés es un segmento de un cromosoma asociado con una aneuploidía parcial, por ejemplo una delección o inserción cromosómica, o translocación cromosómica desequilibrada, y la secuencia normalizadora es un segmento cromosómico que no está asociado con la aneuploidía parcial y cuya variación en la densidad de etiqueta de secuencia mejor se aproxima a la del segmento de cromosoma asociado con la aneuploidía parcial

En el paso **540**, en base a las densidades de etiquetas calificadas calculadas, se determina una dosis de secuencia calificada para una secuencia de interés como la proporción de la densidad de etiqueta de secuencia para la secuencia de interés y la densidad de etiqueta de secuencia calificada para secuencias adicionales de las que se identifican posteriormente secuencias normalizadoras. En una realización, las dosis para el cromosoma de interés, por ejemplo, cromosoma 21, se determinan como una proporción de la densidad de etiqueta de secuencia del cromosoma 21 y la densidad de etiqueta de secuencia para cada uno de los cromosomas restantes, es decir cromosomas 1-20, cromosoma 22, cromosoma X y cromosoma Y.

En el paso **545**, una secuencia normalizadora, por ejemplo un cromosoma normalizador, se identifica para una secuencia de interés, por ejemplo cromosoma 21, en una muestra calificada basada en las dosis de secuencias calculadas. El método identifica secuencias que tienen inherentemente características similares y que son propensas a variaciones similares entre muestras y ejecuciones de secuenciación, y que son útiles para determinar dosis de secuencias en muestras de ensayo. En algunas realizaciones, la secuencia normalizadora es una que diferencia mejor una muestra afectada, por ejemplo una muestra aneuploide, de una o más de las muestras calificadas. En otras realizaciones, una secuencia normalizadora es una secuencia que muestra una variabilidad en el número de etiquetas de secuencia que se mapean para ella entra muestras y ejecuciones de secuenciación que mejor se aproximan a la de la secuencia de interés para la que se usa como un parámetro normalizador y/o que puede diferenciar mejor una muestra afectada de una o más muestras no afectadas.

En algunas realizaciones, se identifica más de una muestra normalizadora. Por ejemplo, la variación, por ejemplo coeficiente de variación, en la dosis de cromosoma para el cromosoma de interés 21 es menor cuando se usa la densidad de etiqueta de secuencia del cromosoma 14. En otras realizaciones, se identifican dos, tres, cuatro, cinco, seis, siete, ocho o más secuencias normalizadoras para su uso al determinar una dosis de secuencia para una secuencia de interés en una muestra de ensayo.

En una realización, la secuencia normalizadora para el cromosoma 21 es seleccionada del cromosoma 9, cromosoma 1, cromosoma 2, cromosoma 3, cromosoma 4, cromosoma 5, cromosoma 6, cromosoma 7, cromosoma 8, cromosoma 10, cromosoma 11, cromosoma 12, cromosoma 13, cromosoma 14, cromosoma 15, cromosoma 16, y cromosoma 17. Preferiblemente, la secuencia normalizadora para el cromosoma 21 se selecciona del cromosoma 9, cromosoma 1, cromosoma 2, cromosoma 11, cromosoma 12, y cromosoma 14. Alternativamente, la secuencia normalizadora para el cromosoma 21 es un grupo de cromosomas seleccionados del cromosoma 9, cromosoma 1, cromosoma 2, cromosoma 3, cromosoma 4, cromosoma 5, cromosoma 6, cromosoma 7, cromosoma 8, cromosoma 10, cromosoma 11, cromosoma 12, cromosoma 13, cromosoma 14, cromosoma 15, cromosoma 16, y cromosoma 17. En otras realizaciones, la secuencia normalizadora para el cromosoma 21 es un grupo de cromosomas seleccionados del cromosoma 9, cromosoma 1, cromosoma 2, cromosoma 11, cromosoma 12, y cromosoma 14.

En una realización, la secuencia normalizadora para el cromosoma 18 se selecciona del cromosoma 8, cromosoma 2, cromosoma 3, cromosoma 4, cromosoma 5, cromosoma 6, cromosoma 7, cromosoma 9, cromosoma 10, cromosoma 11, cromosoma 12, cromosoma 13, y cromosoma 14. Preferiblemente, la secuencia normalizadora para el cromosoma 18 se selecciona del cromosoma 8, cromosoma 2, cromosoma 3, cromosoma 5, cromosoma 6, cromosoma 12, y cromosoma 14. Alternativamente, la secuencia normalizadora para el cromosoma 18 es un grupo de cromosomas seleccionados del cromosoma 8, cromosoma 2, cromosoma 3, cromosoma 4, cromosoma 5, cromosoma 6, cromosoma 7, cromosoma 9, cromosoma 10, cromosoma 11, cromosoma 12, cromosoma 13, y cromosoma 14. En otras realizaciones, la secuencia normalizadora para el cromosoma 18 es un grupo de cromosomas seleccionados del cromosoma 8, cromosoma 2, cromosoma 3, cromosoma 5, cromosoma 6, cromosoma 12, y cromosoma 14.

En una realización, la secuencia normalizadora para el cromosoma X se selecciona del cromosoma 1, cromosoma 2, cromosoma 3, cromosoma 4, cromosoma 5, cromosoma 6, cromosoma 7, cromosoma 8, cromosoma 9, cromosoma 10, cromosoma 11, cromosoma 12, cromosoma 13, cromosoma 14, cromosoma 15, y cromosoma 16. Preferiblemente, la secuencia normalizadora para el cromosoma X se selecciona del cromosoma 2, cromosoma 3, cromosoma 4, cromosoma 5, cromosoma 6, y cromosoma 8. Alternativamente, la secuencia normalizadora para el cromosoma X es un grupo de cromosomas seleccionados del cromosoma 1, cromosoma 2, cromosoma 3, cromosoma 4, cromosoma 5, cromosoma 6, cromosoma 7, cromosoma 8, cromosoma 9, cromosoma 10, cromosoma 11, cromosoma 12, cromosoma 13, cromosoma 14, cromosoma 15, y cromosoma 16. En otras realizaciones, la secuencia normalizadora para el cromosoma X es un grupo de cromosomas seleccionados del cromosoma 2, cromosoma 3, cromosoma 4, cromosoma 5, cromosoma 6, y cromosoma 8.

En una realización, la secuencia normalizadora para el cromosoma 13 es un cromosoma seleccionado del cromosoma 2, cromosoma 3, cromosoma 4, cromosoma 5, cromosoma 6, cromosoma 7, cromosoma 8, cromosoma 9, cromosoma 10, cromosoma 11, cromosoma 12, cromosoma 14, cromosoma 18, y cromosoma 21. Preferiblemente, la secuencia normalizadora para el cromosoma 13 se selecciona del cromosoma 2, cromosoma 3, cromosoma 4, cromosoma 5, cromosoma 6, y cromosoma 8. En otra realización, la secuencia normalizadora para el cromosoma 13 es un grupo de cromosomas seleccionados del cromosoma 2, cromosoma 3, cromosoma 4, cromosoma 5, cromosoma 6, cromosoma 7, cromosoma 8, cromosoma 9, cromosoma 10, cromosoma 11, cromosoma 12, cromosoma 14, cromosoma 18, y cromosoma 21. En otras realizaciones, la secuencia normalizadora para el cromosoma 13 es un grupo de cromosomas seleccionado del cromosoma 2, cromosoma 3, cromosoma 4, cromosoma 5, cromosoma 6, y cromosoma 8.

La variación en la dosis de cromosoma para el cromosoma Y es mayor de 30 independientemente de que cromosoma normalizador se use para determinar la dosis de cromosoma Y. Por lo tanto, cualquier cromosoma, o un grupo de dos o más cromosomas seleccionados de los cromosomas 1-22 y el cromosoma X se puede usar como la secuencia normalizadora para el cromosoma Y. En una realización, el al menos un cromosoma normalizadora es un grupo de cromosomas que consiste de los cromosomas 1-22, y el cromosoma X. En otra realización, el al menos un cromosoma normalizador es un grupo de cromosomas seleccionados del cromosoma 2, cromosoma 3, cromosoma 4, cromosoma 5 y cromosoma 6.

En base a la identificación de la secuencia(s) normalizadora en muestras calificadas, se determina una dosis de cromosoma para una secuencia de interés en una muestra de ensayo que comprende una mezcla de ácidos nucleicos derivados de genomas que difieren en una o más secuencias de interés.

En el paso **515**, se obtiene una muestra de ensayo, por ejemplo muestra de plasma, que comprende ácidos nucleicos fetales y maternos, por ejemplo ADN libre de células, de un sujeto embarazado, por ejemplo una mujer embarazada, para la que se necesita determinar la presencia o ausencia de una aneuploidía fetal.

En el paso **525**, al menos una porción de los ácidos nucleicos de ensayo en la muestra de ensayo se secuencia para generar millones de lecturas de secuencia que comprenden entre 20 y 500 bp, por ejemplo 36bp. Como en el paso **520**, las lecturas generadas de la secuenciación de ácidos nucleicos en la muestra de ensayo se mapean únicamente para un genoma de referencia humano y se cuentan. Como se describe en el paso **520**, se obtienen al menos alrededor de 3×10^6 etiquetas de secuencia calificadas, al menos alrededor de 5×10^6 etiquetas de secuencia calificadas, al menos alrededor de 8×10^6 etiquetas de secuencia calificadas, al menos alrededor de 10×10^6 etiquetas de secuencia calificadas, al menos alrededor de 15×10^6 etiquetas de secuencia calificadas, al menos alrededor de 20×10^6 etiquetas de secuencia calificadas, al menos alrededor de 30×10^6 etiquetas de secuencia calificadas, al menos alrededor de 40×10^6 etiquetas de secuencia calificadas, o al menos alrededor de 50×10^6 etiquetas de secuencia calificadas que comprenden entre 20 y 40bp de lecturas que mapean únicamente para el genoma de referencia humano.

En el paso **535**, todas las etiquetas obtenidas de la secuenciación de ácidos nucleicos en las muestras de ensayo se cuentan para determinar la densidad de etiqueta de secuencia. En una realización, el número de etiquetas de secuencia de ensayo mapeadas para una secuencia de interés se normaliza a la longitud conocida de una secuencia de interés a la que están mapeadas para proporcionar una densidad de etiqueta de secuencia de ensayo. Como se describe para las muestras calificadas, no se requiere la normalización a la longitud conocida de una secuencia de interés, y puede incluirse como un paso para reducir al número de dígitos en un número para simplificarla para la interpretación humana. A medida que todas las etiquetas de secuencia de ensayo mapeadas se cuentan en la muestra de ensayo, se determina en las muestras de ensayo la densidad de etiqueta de secuencia para una secuencia de interés, por ejemplo una secuencia clínicamente relevante como el cromosoma 21, como los son las densidades de etiqueta de secuencia para secuencias adicionales que corresponden con la menos una secuencia normalizadora identificada en las muestras calificadas.

En el paso **550**, en base a la identidad de al menos una secuencia normalizadora en las muestras calificadas, se determina una dosis de secuencia de ensayo para una secuencia de interés en la muestra de ensayo. La dosis de secuencia, por ejemplo dosis de cromosoma, para una secuencia de interés en una muestra de ensayo es una proporción de la densidad de etiqueta de secuencia determinada para la secuencia de interés en la muestra de ensayo y la densidad de etiqueta de secuencia de al menos una secuencia normalizadora determinada en la muestra de ensayo, en donde la secuencia normalizadora en la muestra de ensayo se corresponde con la secuencia normalizadora identificada en las muestras calificadas para la secuencia particular de interés. Por ejemplo, si se determina que la secuencia normalizadora para el cromosoma 21 en las muestras calificadas es el cromosoma 14, entonces la dosis de secuencia de ensayo para el cromosoma 21 (secuencia de interés) se determina como la proporción de la densidad de etiqueta de secuencia para el cromosoma 21 en y la densidad de etiqueta de secuencia para el cromosoma 14 cada una determinada en la muestra de ensayo. De manera similar, se determinan las dosis de cromosomas para los cromosomas 13, 18, X, Y y otros cromosomas asociados con las aneuploidías cromosómicas. Como se ha descrito anteriormente, una secuencia de interés puede ser parte de un cromosoma, por ejemplo un segmento de cromosoma. Por consiguiente, la dosis para un segmento de cromosoma puede

determinarse como la proporción de la densidad de etiqueta de secuencia determinada para el segmento en la muestra de ensayo y la densidad de etiqueta de secuencia para el segmento de cromosoma normalizador en la muestra de ensayo, en donde el segmento normalizador en la muestra de ensayo corresponde con el segmento normalizador identificado en las muestras calificadas para el segmento particular de interés.

En el paso 555, los valores límite se derivan de los valores de la desviación estándar establecidos por una pluralidad de dosis de secuencia calificadas. La clasificación precisa depende de las diferencias entre las distribuciones de probabilidad para las diferentes clases, es decir tipo de aneuploidía. Preferiblemente, los límites se eligen de la distribución empírica para cada tipo de aneuploidía, por ejemplo trisomía 21. Los valores límites posibles que se establecieron para clasificar las aneuploidías de trisomía 13, trisomía 18, trisomía 21 y monosomía X como se describe en los Ejemplos, que describen el uso del método para determinar las aneuploidías cromosómicas por secuenciación de ADN libre de células extraído de una muestra materna comprenden una mezcla de ácidos nucleicos fetales y maternos.

En el paso 560, la variación del número de copias de la secuencia de interés, por ejemplo aneuploidía cromosómica o parcial, se determina en la muestra de ensayo comparando al dosis de secuencia de ensayo para la secuencia de interés con al menos un valor límite establecido de las dosis de secuencia calificadas.

En el paso 560, la dosis calculada para una secuencia de ensayo de interés se compara con ese conjunto como los valores límite que se eligen de acuerdo con un límite definido por el usuario de fiabilidad para clasificar al muestra como una "normal", una "afectada" o una "no designada" en el paso 565. Las muestras "no designadas" son muestras para las que no se puede hacer un diagnóstico definitivo con fiabilidad.

Otra realización divulgada en la presente es un método para proporcionar diagnóstico prenatal de una aneuploidía cromosómica fetal en una muestra biológica que comprende moléculas de ácidos nucleicos fetales y maternos. El diagnóstico se hace en base a la recepción de datos de la secuenciación en al menos una porción de la mezcla de moléculas de ácidos nucleicos fetales y maternos derivados de una muestra de ensayo biológica, por ejemplo una muestra de plasma materno, calculando de los datos de secuenciación una dosis de cromosoma normalizador para uno o más cromosomas de interés, determinando una diferencia estadísticamente significativa entre la dosis de cromosoma normalizador para el cromosoma de interés en la muestra de ensayo y un valor límite establecido en una pluralidad de muestras calificadas (normales), y proporcionando el diagnóstico prenatal en base a la diferencia estadística. Como se describe en el paso 165 del método, se hace un diagnóstico de normal o afectada. Se proporciona una "no designada" en el caso de que el diagnóstico para la normal o afectada no se pueda hacer con confianza.

La cuantificación del número de lecturas de secuencia que alinean con cada cromosoma para determinar aneuploidías cromosómicas también se puede conseguir normalizando el número medio de etiquetas de secuencia para un cromosoma de interés con el número medio de etiquetas para cada uno de los otros autosomas (Fan et al., Proc Natl Acad Sci 105:16266-16271 [2008]). Alternativamente, el número de lecturas únicas que alinean para cada cromosoma se compara con el número total de lecturas que alinean para todos los cromosomas para derivar un porcentaje de representación genómica para cada cromosoma. Se genera una "puntuación z" para representar la diferencia entre el porcentaje de representación genómica del cromosoma de interés y el porcentaje medio de representación para el mismo cromosoma entre un grupo de control euploide, dividido por la desviación estándar (Chiu et al., Clin Chem 56:459-463 [2010]).

Determinación de la fracción fetal

La determinación de la fracción fetal se basa en el número total de etiquetas que mapean para el primer alelo y el número total de etiquetas que mapean para el segundo alelo en un sitio polimórfico informativo, por ejemplo un SNP, contenido en un genoma de referencia. Por ejemplo, el genoma de referencia es la secuencia NCBI36/hg18 del genoma de referencia humano, o el genoma de referencia comprende la secuencia NCBI36/hg18 del genoma de referencia humano y un genoma de secuencias objetivo artificiales, lo que incluye las secuencias polimórficas objetivo. Por ejemplo, el genoma objetivo artificial abarca secuencias polimórficas que comprenden SNPs rs560681, rs1109037, rs9866013, rs13182883, rs13218440, rs7041158, rs740598, rs10773760, rs4530059, rs7205345, rs8078417, rs576261, rs2567608, rs430046, rs9951171, rs338882, rs10776839, rs9905977, rs1277284, rs258684, rs1347696, rs508485, rs9788670, rs8137254, rs3143, rs2182957, rs3739005, y rs530022. En otro ejemplo, el genoma artificial incluye las secuencias objetivo polimórficas de las SEQ ID Nos: 1-56 (ver Ejemplo 5). En otro ejemplo, el genoma artificial comprende secuencias polimórficas que comprenden SNPs en tándem rs7277033-rs2110153; rs2822654-rs1882882; rs368657-rs376635; rs2822731-rs2822732; rs1475881-rs7275487; rs1735976-rs2827016; rs447340-rs2824097; rs418989-rs13047336; rs987980-rs987981; rs4143392-rs4143391; rs1691324-rs13050434; rs11909758-rs9980111; rs2826842-rs232414; rs1980969-rs1980970; rs9978999-rs9979175; rs1034346-rs12481852; rs7509629-rs2828358; rs4817013-rs7277036; rs9981121-rs2829696; rs455921-rs2898102; rs2898102-rs458848; rs961301-rs2830208; rs2174536-rs458076; rs11088023-rs11088024; rs1011734-rs1011733; rs2831244-rs9789838; rs8132769-rs2831440; rs8134080-rs2831524; rs4817219-rs4817220; rs2250911-rs2250997; rs2831899-rs2831900; rs2831902-rs2831903; rs11088086-rs2251447; rs2832040-rs11088088; rs2832141-

rs2246777; rs2832959 -rs9980934; rs2833734-rs2833735; rs933121-rs933122; rs2834140-rs12626953; rs2834485-rs3453; rs9974986-rs2834703; rs2776266-rs2835001; rs1984014-rs1984015; rs7281674-rs2835316; rs13047304-rs13047322; rs2835545-rs4816551; rs2835735-rs2835736; rs13047608-rs2835826; rs2836550-rs2212596; rs2836660-rs2836661; rs465612-rs8131220; rs9980072-rs8130031; rs418359-rs2836926; rs7278447-rs7278858; rs385787-rs367001; rs367001-rs386095; rs2837296-rs2837297; y rs2837381-rs4816672. En otro ejemplo el genoma objetivo artificial abarca secuencias polimórficas que comprenden STRs seleccionadas de CSF1PO, FGA, TH01, TPOX, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, Penta D, Penta E, D2S1338, D1S1677, D2S441, D4S2364, D10S1248, D14S1434, D22S1045, D22S1045, D20S1082, D20S482, D18S853, D17S1301, D17S974, D14S1434, D12ATA63, D11S4463, D10S1435, D10S1248, D9S2157, D9S1122, D8S1115, D6S1017, D6S474, D5S2500, D5S2500, D4S2408, D4S2364, D3S4529, D3S3053, D2S1776, D2S441, D1S1677, D1S1627, y D1GATA113. La composición de las secuencias objetivo artificiales variará dependiendo de las secuencias polimórficas que se usan para determinar la fracción fetal. Por consiguiente, un genoma de secuencias objetivo artificiales no está limitado a las secuencias de SNP o STR ejemplificadas en la presente.

El sitio polimórfico informativo, por ejemplo SNP, se identifica por la diferencia en las secuencias alélicas y la cantidad de cada uno de los alelos posibles. El ADN libre de células está presente a una concentración que es <10% del ADN libre de células materno. Por lo tanto, la presencia de una contribución menor de un alelo a la mezcla de ácidos nucleicos fetales y maternos en relación con la contribución principal del alelo materno puede ser asignada al feto. Los alelos que se derivan del genoma materno son referidos en la presente como alelos principales, y los alelos que se derivan del genoma fetal son referidos en la presente como alelos menores. Los alelos que se representan por niveles similares de etiquetas de secuencias mapeadas representan alelos maternos. Los resultados de una amplificación multiplex ejemplar de ácidos nucleicos objetivo que comprenden SNPs y derivados de una muestra de plasma materno se muestran en la **Figura 6**. Los SNPs informativos son distinguidos del cambio del nucleótido individual en un sitio polimórfico predeterminado, y los alelos fetales se distinguen por su contribución menor relativa a la mezcla de ácidos nucleicos fetales y maternos en la muestra cuando se compara con la contribución mayor a la mezcla por los ácidos nucleicos maternos, es decir las secuencias de SNP son informativas cuando la madre es heterocigótica y hay presente un tercer alelo paternal, permitiendo una comparación cuantitativa entre el alelo heredado maternalmente y el alelo heredado paternalmente para calcular la fracción fetal. Por consiguiente, la abundancia relativa de ADN libre de células fetal en la muestra materna se determina como un parámetro del número total de etiquetas de secuencia únicas mapeadas para la secuencia de ácido nucleicos objetivo en un genoma de referencia para cada uno de los dos alelos del sitio polimórfico predeterminado. En una realización, la fracción de ácidos nucleicos fetales en la mezcla de ácidos nucleicos fetales y maternos se calcula para cada uno de los alelos informativos (alelo_x) de la manera siguiente:

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

y la fracción fetal para la muestra se calcula como la media de la fracción fetal de todos los alelos informativos. Opcionalmente, la fracción de los ácidos nucleicos fetales en la mezcla de ácidos nucleicos fetales y maternos se calcula para cada uno de los alelos informativos (alelo_x) de la manera siguiente:

$$\% \text{ del alelo}_x \text{ de la fracción fetal} = ((2 \times \sum \text{etiquetas de secuencia fetal para el alelo}_x) / (\sum \text{etiquetas de secuencia materna para el alelo}_x)) \times 100$$

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

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

En una realización, el paso de enriquecer la mezcla de ácidos nucleicos fetales y maternos para ácidos nucleicos objetivo polimórficos comprende amplificar los ácidos nucleicos objetivo en una porción de una muestra de ensayo, por ejemplo una muestra de ensayo de plasma, y combinar todo o una porción del producto amplificado con el resto de la muestra de ensayo de plasma. La realización del método **200** se representa en el diagrama de flujo proporcionado en la **Figura 2**. En el paso **210**, se obtiene una muestra de ensayo, por ejemplo una muestra de fluido biológico como una muestra de sangre, de una mujer embarazada, y en el paso **220** se usa una porción del ADN libre de células contenido en la fracción de plasma de la muestra de sangre para amplificar ácidos nucleicos objetivo que comprenden sitios polimórficos, por ejemplo SNPs. En una realización, al menos alrededor del 1%, al menos alrededor del 1,5%, al menos alrededor del 2%, al menos alrededor del 10% del plasma materno se usó para amplificar los ácidos nucleicos objetivo. En el paso **230**, una porción o todos los ácidos nucleicos objetivo amplificados se combina con la mezcla de ADN libre de células fetal y materno presente en la muestra materna, y el ADN libre de células combinado y los ácidos nucleicos amplificados se purifican en el paso **240**, y se usan para preparar una biblioteca que se secuenció en el paso **250**. La biblioteca se purificó de ADN libre de células purificado y comprende al menos alrededor del 10%, al menos alrededor del 15%, al menos alrededor del 20%, al menos

alrededor del 25%, al menos alrededor del 30%, al menos alrededor del 35%, al menos alrededor del 40%, al menos alrededor del 45%, o al menos alrededor del 50% del producto amplificado. En el paso **260**, los datos de las ejecuciones de secuenciación se analizan y se hace la determinación simultánea de la fracción fetal y la presencia o ausencia de aneuploidía.

En una realización, el paso de enriquecer la mezcla de ácidos nucleicos fetales y maternos para los ácidos nucleicos objetivo polimórficos comprende una pluralidad de ácidos nucleicos objetivo polimórficos en una porción de una mezcla de ácidos nucleicos fetales y maternos purificados de una muestra de ensayo materna. En una realización, una porción de una mezcla de ácidos nucleicos fetales y maternos, por ejemplo ADN libre de células, purificada de una muestra de plasma materno se usa para amplificar las secuencias de ácidos nucleicos polimórficos, y una porción del producto amplificado se combina con la mezcla no amplificada o ácidos nucleicos fetales y maternos, por ejemplo ADN libre de células (ver **Figura 3**). La realización del método **300** se representa en el diagrama de flujo proporcionado en la **Figura 3**. En el paso **310**, se obtiene de una mujer embarazada una muestra de ensayo, por ejemplo una muestra de fluido biológico como una muestra de sangre, que comprende una mezcla de ácidos nucleicos fetales y maternos, y la mezcla de ácidos nucleicos fetales y maternos se purifica de la fracción de plasma en el paso **320**. Como se ha descrito anteriormente, los métodos para la separación de ADN libre de células del plasma son bien conocidos. En el paso **330**, una porción del ADN libre de células contenido en la muestra purificada se usa para amplificar ácidos nucleicos objetivo que comprenden sitios polimórficos, por ejemplo SNPs. Al menos alrededor del 5%, al menos alrededor del 10%, al menos alrededor del 15%, al menos alrededor del 20%, al menos alrededor del 25%, al menos alrededor del 30%, al menos alrededor del 35%, al menos alrededor del 40%, al menos alrededor del 45%, o al menos alrededor del 50% del ADN libre de células purificado se usa para amplificar los ácidos nucleicos objetivo. Preferiblemente, la amplificación de las secuencias objetivo se puede realizar por cualquier método que use PCR o variaciones del método incluyendo pero no limitado a PCR asimétrico, amplificación dependiente de la helicasa, CR hot-start, qPCR, PCR en fase sólida y PCR touchdown. En el paso **340**, una porción, por ejemplo al menos alrededor del 0,01% del producto amplificado se combina con la muestra de ADN libre de células purificado no amplificado, y la mezcla de los ácidos nucleicos fetales y maternos amplificados y no amplificados se secuencia en el paso **350**. En una realización, la secuenciación se realiza usando cualquiera de las tecnologías NGS. En el paso **360**, los datos de las ejecuciones de secuenciación se analizan y se hace la determinación simultánea de la fracción fetal y la presencia o ausencia de aneuploidía como se describe en el paso **140** de la realización representada en la **Figura 1**.

En otra realización, el paso **120** de enriquecer la mezcla de ácidos nucleicos fetales y maternos para ácidos nucleicos objetivo polimórficos comprende combinar al menos una porción de una primera biblioteca de secuenciación de moléculas de ácidos nucleicos fetales y maternos no amplificados con al menos una porción de una segunda biblioteca de secuenciación de ácidos nucleicos objetivo polimórficos. Por lo tanto, la muestra que es enriquecida es la muestra de la biblioteca que se prepara para la secuenciación (**Figura 4**). El enriquecimiento de la muestra de la biblioteca para los ácidos nucleicos objetivo se consigue por métodos que comprenden amplificar específicamente las secuencias de ácidos nucleicos que comprenden el sitio polimórfico. En el paso **410**, se obtiene de una mujer embarazada una muestra de ensayo, por ejemplo una muestra de fluido biológico como una muestra de sangre, que comprende una mezcla de ácidos nucleicos fetales y maternos, y la mezcla de ácidos nucleicos fetales y maternos se purifica de la fracción de plasma en el paso **420**. En el paso **430**, se usa una porción del ADN libre de células contenido en la muestra purificada para amplificar ácidos nucleicos objetivo que comprenden sitios polimórficos, por ejemplo SNPs. Al menos alrededor del 5%, al menos alrededor del 10%, al menos alrededor del 15%, al menos alrededor del 20%, al menos alrededor del 25%, o al menos alrededor del 30% del ADN libre de células se usa para amplificar secuencias de ácidos nucleicos objetivo. Preferiblemente, la amplificación de las secuencias objetivo puede realizarse por cualquier método que use PCR o variaciones del método incluyendo, pero no limitadas a, PCR asimétrico, amplificación dependiente de la helicasa, PCR hot-start, qPCR, PCR en fase sólida y PCR touchdown. En el paso **440**, los ácidos nucleicos objetivo amplificados que comprenden los sitios polimórficos, por ejemplo SNPs, se usan para preparar una biblioteca de secuenciación de ácidos nucleicos objetivo. De manera similar, la porción de ADN libre de células no amplificado purificado se usa para preparar una biblioteca de secuenciación primaria en el paso **450**. En el paso **460**, una porción de la biblioteca objetivo se combina con la biblioteca primaria generada de la mezcla no amplificada de ácidos nucleicos, y la mezcla de ácidos nucleicos fetales y maternos comprendida en las dos bibliotecas se secuencia en el paso **470**. La biblioteca enriquecida comprende al menos alrededor del 5%, al menos alrededor del 10%, al menos alrededor del 15%, al menos alrededor del 20%, o al menos alrededor del 25% de la biblioteca objetivo. En el paso **480**, los datos de las ejecuciones de secuenciación se analizan y se hace la determinación simultánea de la fracción fetal y la presencia o ausencia de aneuploidía como se describe en el paso **140** de la realización representada en la **Figura 1**.

Determinación de aneuploidías para diagnósticos prenatales

El ARN y ADN fetal libre de células que circula en la sangre materna puede usarse para el diagnóstico prenatal no invasivo temprano (NIPD) de un número creciente de condiciones genéticas, tanto para la gestión del embarazo y para ayudar a la toma de decisiones reproductoras. La presencia de ADN libre de células que circulan en el torrente sanguíneo se ha conocido durante más de 50 años. Más recientemente, la presencia de pequeñas cantidades de ADN fetal circulante se descubrió en el torrente sanguíneo materno durante el embarazo (Lo et al.,

Lancet 350:485-487 [1997]). Creyendo que se origina de células de la placenta que se mueren, el ADN fetal libre de células (ADNcf) ha demostrado que consiste de fragmentos cortos típicamente de menos de 200 bp de longitud Chan et al., Clin Chem 50:88-92 [2004]), que se pueden distinguir tan pronto como desde las 4 semanas de gestación (Illanes et al., Early Human Dev 83:563-566 [2007]), y que se sabe desaparecen de la circulación materna pocas horas después del parto (Lo et al., Am J Hum Genet 64:218-224 [1999]). Además del ADN libre de células, también se pueden distinguir fragmentos de ARN fetal libre de células (ARNcf) en el torrente sanguíneo materno, originado de genes que es transcriben en el feto o la placenta. La extracción y el posterior análisis de estos elementos genéticos fetales de una muestra de sangre materna ofrece nuevas oportunidades para el NIPD.

El presente método es un método independiente de polimorfismos para el uso en NIPD y que no requiere que el ADN libre de células fetal se distinga del ADN libre de células materno para permitir la determinación de una aneuploidía fetal. En algunas realizaciones, la aneuploidía es una trisomía o monosomía cromosómica completa, o una trisomía o monosomía parcial. Las aneuploidías parciales son causadas por la pérdida o ganancia de parte de un cromosoma, y abarcan desequilibrios cromosómicos resultantes de translocaciones desequilibradas, inversiones desequilibradas, deleciones e inserciones. Con mucho, la aneuploidía más común conocida compatible con la vida es la trisomía 21, es decir Síndrome de Down (DS), que es causado por la presencia de parte o todo el cromosoma 21. Raramente, el DS puede ser causa de un defecto heredado o esporádico por el que una copia extra de todo o parte del cromosoma 21 se uno con otro cromosoma (habitualmente el cromosoma 14) para formar un único cromosoma anormal. El DS está asociado con deficiencia intelectual, dificultades de aprendizaje severas y exceso de mortalidad causado por problemas de salud a largo plazo como enfermedades cardíacas. Otras aneuploidías con significancia clínica conocida incluyen el síndrome de Edward (trisomía 18) y Síndrome de Patau (trisomía 13), que son frecuentemente fatales dentro de los primeros meses de vida. También se conocen anomalías asociadas con el número de cromosomas sexuales e incluyen monosomía X, por ejemplo síndrome de Turner (XO), y síndrome de triple X (XXX) en nacimientos de niñas y síndrome de Klinefelter (XXY) y síndrome de XYY en nacimientos de niños, que están todos asociados con varios fenotipos incluyendo esterilidad y reducción en las habilidades intelectuales. El método de la invención puede ser usado para diagnosticar estas y otras anomalías cromosómicas prenatalmente.

De acuerdo con las realizaciones de la presente invención la trisomía determinada por la presente invención se selecciona de trisomía 21 (T21; Síndrome de Down), trisomía 18 (T18; Síndrome de Edward); trisomía 16 (T16), trisomía 22 (T22; Síndrome de Ojo de Gato); trisomía 15 (T15; Síndrome de Prader Willi), trisomía 13 (T13; Síndrome de Patau), trisomía 8 (T8; Síndrome de Warkany) y trisomías XXY (Síndrome de Klinefelter), XYY o XXX. Se apreciará que varias otras trisomías y trisomías parciales se pueden determinar en ADN libre de células fetal de acuerdo con las enseñanzas de la presente invención. Estas incluyen, pero no están limitadas a, trisomía parcial 1q32-44, trisomía 9 p con trisomía, trisomía, mosaicismo 4, trisomía 17p, trisomía parcial 4q26-qter, trisomía 9, trisomía parcial 2p, trisomía parcial 1q, y/o trisomía parcial 6p/monosomía 6q.

El método de la presente invención también puede usarse para determinar la monosomía X cromosómica, y monosomías parciales como la monosomía 13, monosomía 15, monosomía 16, monosomía 21 y monosomía 22, que se sabe están implicadas en el aborto involuntario del embarazo. La monosomía parcial de cromosomas típicamente implicados en la aneuploidía completa también puede ser determinada por el método de la invención. La monosomía 18p es un trastorno cromosómico raro en el que todo o parte del brazo corto (p) del cromosoma 18 se elimina (monosómico). Este trastorno está caracterizado típicamente por estatura corta, grados variables de retraso mental, retrasos en el habla, malformaciones de la región craneal y facial (craneofacial), y/o anomalías físicas adicionales. Los defectos craneofaciales asociados pueden variar enormemente en rango y severidad de caso a caso. Las condiciones provocadas por cambios en la estructura o número de copias del cromosoma 15 incluyen el Síndrome de Angelman y el Síndrome de Prader-Willi, que implican una pérdida de actividad génica en la misma parte del cromosoma 15, la región 15q11-q13. Se apreciará que varias translocaciones y microdeleciones pueden ser asintomáticas en el padre portador, sin embargo pueden provocar una enfermedad genética importante en la descendencia. Por ejemplo, una madre sana que porta la microdelección 15q11-q13 puede dar a luz a un niño con el síndrome de Angelman, un trastorno neurodegenerativo severo. Por lo tanto, la presente invención puede usarse para identificar dicha delección en el feto. La monosomía parcial 13q es un trastorno cromosómico raro que se produce cuando falta una pieza del brazo largo (q) del cromosoma 13 (monosómico). Los bebés nacidos con monosomía 13q parcial pueden mostrar bajo peso de nacimiento, malformaciones de la cabeza y la cara (región craneofacial), anomalías esqueléticas (especialmente de las manos y pies), y otras anomalías físicas. El retraso mental es característico de esta condición. La tasa de mortalidad durante la niñez es alta entre individuos nacidos con este trastorno. Casi todos los casos de monosomía 13q parcial tienen lugar aleatoriamente sin razón aparente (esporádica). El síndrome de delección 22q11.2, también conocido como síndrome de DiGeorge, es un síndrome causado por la delección de una pieza pequeña del cromosoma 22. La delección (22q11.2) tiene lugar cerca del medio del cromosoma en el brazo largo de uno del par de cromosomas. Las características de este síndrome varían ampliamente, incluso entre miembros de la misma familia, y afectan a muchas partes del cuerpo. Los signos característicos y síntomas pueden incluir defectos de nacimiento como enfermedad cardíaca congénita, defectos en el paladar, más comúnmente relacionados con problemas neuromusculares con cleisis (insuficiencia velo-faríngea), problemas de aprendizaje, diferencias leves en los rasgos faciales e infecciones recurrentes. Las microdeleciones en la región cromosómica 22q11.2 están asociadas con un riesgo aumentado de 20 a 30 veces de esquizofrenia. En

una realización, el método de la invención se usa para determinar monosomías parciales que incluyen per no están limitadas a monosomía 18p, monosomía parcial del cromosoma 15 (15q-11q13), monosomía parcial 13q, y la monosomía parcial del cromosoma 22 también pueden determinarse usando el método.

El método de la invención también puede ser usado para determinar cualquier aneuploidía si uno de los padres es un portador conocido de dicha anomalía. Estas incluyen, pero no están limitadas a, cromosoma marcador supernumerario (SMC) pequeño; translocación t(11;14)(p15;p13); translocación desequilibrada t(8;11)(p23.2;p15.5); microdelección 11q23; delección 17p11.2 síndrome de Smith-Magenis; delección 22q13.3; Microdelección Xp22.3; 10p14 eliminación; Microdelección 20p, el síndrome de Di-George [del (22) (q11.2q11.23)], el síndrome de Williams (delecciones 7q11.23 y 7q36); delección 1p36; microdelección 2p; neurofibromatosis tipo 1 (microdelección 17q11.2), delección Yq; Síndrome de Wolf-Hirschhorn (WHS, microdelección 4p16.3); microdelección 1p36.2; delección 11q14; microdelección 19q13.2; Rubinstein-Taybi (microdelección 16 p13.3); microdelección 7p21; síndrome de Miller-Dieker (17p13.3), delección 17p11.2; y microdelección 2q37.

Composiciones y kits

Se divulgan composiciones que comprenden cebadores para amplificar sitios polimórficos para permitir la cuantificación de la fracción fetal y aneuploidía por mezclas de secuenciación de ácidos nucleicos fetales y maternos, por ejemplo ADN libre de células, presentes en la muestra. Preferiblemente, la muestra es una muestra de plasma de sangre materno. En otra realización, la composición incluye cebadores para amplificar ácidos nucleicos objetivo polimórficos que comprenden cada uno al menos un SNP. El al menos un SNP se selecciona de los SNPs rs560681, rs1109037, rs9866013, rs13182883, rs13218440, rs7041158, rs740598, rs10773760, rs4530059, rs7205345, rs8078417, rs576261, rs2567608, rs430046, rs9951171, rs338882, rs10776839, rs9905977, rs1277284, rs258684, rs1347696, rs508485, rs9788670, rs8137254, rs3143, rs2182957, rs3739005 y rs530022. El conjunto de cebadores correspondiente para amplificar los SNPs se proporcionan en el Ejemplo 3 y se divulgan como SEQ ID Nos: 57-112.

En otra realización, la composición incluye cebadores para amplificar ácidos nucleicos objetivo polimórficos que comprenden cada uno al menos un SNP en tándem. En una realización, la composición incluye cebadores para amplificar los SNPs en tándem ejemplares divulgados en la presente, y la composición comprende los cebadores ejemplares correspondientes de las SEQ ID NOS:57-112.

En otra realización, la composición incluye cebadores para amplificar ácidos nucleicos objetivo polimórficos que comprenden cada uno al menos una STR. Las STRs ejemplares incluyen CSF1PO, FGA, TH01, TPOX, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, D2S1338, Penta D, Penta E, D22S1045, D20S1082, D20S482, D18S853, D17S1301, D17S974, D14S1434, D12ATA63, D11S4463, D10S1435, D10S1248, D9S2157, D9S1122, D8S1115, D6S1017, D6S474, D5S2500, D4S2408, D4S2364, D3S4529, D3S3053, D2S1776, D2S441, D1S1677, D1S1627 y D1GATA113, que pueden estar amplificadas por los conjuntos correspondientes de cebadores proporcionados en el Ejemplo 5 (Tablas 5 y 6) y divulgadas como SEQ ID NOS; 113-196.

Las composiciones divulgadas en la presente invención incluirse en kits para mezclas de secuenciación masivamente paralelas de moléculas de ácidos nucleicos fetales y maternos, por ejemplo ADN libre de células, presentes en una muestra materna, por ejemplo una muestra de plasma. Los kits comprenden una composición que comprende al menos un conjunto de cebadores para amplificar al menos un ácido nucleico objetivo polimórfico en dichas moléculas de ácidos nucleicos fetales y maternos. Los ácidos nucleicos polimórficos pueden comprender un SNP o una STR. Los métodos de secuenciación son métodos NGS de moléculas de ácidos nucleicos individuales o moléculas de ácidos nucleicos amplificadas clonalmente. Los métodos NGS son métodos de secuenciación masivamente paralelos que incluyen pirosecuenciación, secuenciación por síntesis con terminadores de colorante reversibles, secuenciación en tiempo real o secuenciación por ligadura de sonda de oligonucleótidos.

La presente invención se describe con más detalle en los siguientes Ejemplos que no se pretende que limiten de ninguna manera el ámbito de la invención como se reivindica. Se pretende que las Figuras adjuntas se consideren como parte integral de la especificación y descripción de la invención. Los siguientes ejemplos se ofrecen para ilustrar, pero no para limitar la invención reivindicada.

EXPERIMENTAL

Ejemplo 1

Procesamiento de Muestras y Extracción de ADN libre de células

Las muestras de sangre periférica se recogieron de mujeres embarazadas en su primer o segundo trimestre de embarazo y que se consideraron con riesgo de aneuploidía fetal. Se obtuvo el consentimiento informado de cada participante antes de la extracción de sangre. La sangre se recogió antes de la amniocentesis o la muestra de

vellosidades coriónicas. El análisis de cariotipos se realizó usando las muestras de vellosidades coriónicas o amniocentesis para confirmar el cariotipo fetal.

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

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

Se extrajo ADN libre de células (ADNcf) purificado de plasma libre de células usando el QIAamp Blood DNA Mini kit (Qiagen) de acuerdo esencialmente con las instrucciones del fabricante. Se añadieron un mililitro de tampón AL y 100 µl de solución de proteasa a 1ml de plasma. La mezcla se incubó durante 15 minutos a 56° C. Se añadió un mililitro de 100% de etanol a la digestión de plasma. La mezcla resultante se transfirió a minicolumnas QIAamp que se montaron con los VacValves y VacConnectors proporcionado en el montaje de columnas QIAvac 24 Plus (Qiagen). Se aplicó vacío a las muestras, y el ADN libre de células retenido en los filtros de la columna se lavó al vacío con 750 µl de tampón AW1, seguido por un segundo lavado con 750 µl de tampón AW24. La columna se centrifugó a 14.000 RPM durante 5 minutos para eliminar cualquier tampón residual del filtro. El ADN libre de células se eluyó con tampón AE por centrifugación a 14.000 RPM, y se determinó la concentración usando la Qubit™ Quantitation Platform (Invitrogen).

Ejemplo 2

Preparación y secuenciación de bibliotecas de secuenciación primarias y enriquecidas

a. Preparación de bibliotecas de secuenciación

Todas las bibliotecas de secuenciación, es decir bibliotecas primarias y enriquecidas, se prepararon de aproximadamente 2 ng de ADN libre de células purificado que se extrajo de plasma materno. La preparación de la biblioteca se realizó usando reactivos del Conjunto 1 de Reactivos de ADN de Preparación de Muestras NEBNext™ (Parte No. E6000L; New England Biolabs, Ipswich, MA), para Illumina® como sigue. Debido a que el ADN de plasma libre de células está fragmentado en la naturaleza, no se hizo fragmentación adicional por nebulización o sonicación en las muestras de ADN de plasma. Los excesos de aproximadamente 2 ng de fragmentos de ADN libre de células purificados contenidos en 40 µl se convirtieron en extremos romos fosforilados de acuerdo con el NEBNext® End Repair Module incubando en un tubo de microcentrifuga de 1,5 ml el ADN libre de células con 5 µl de tampón de fosforilación10X, 2 µl de mezcla de solución de desoxinucleótidos (10 mM cada dNTP), 1 µl de una dilución 1:5 de ADN polimerasa, 1 µl de ADN Polimerasa T4 y 1µl de Polinucleótido Quinasa T4 proporcionado en el Conjunto 1 de Reactivos de ADN de Preparación de Muestras de ADN NEBNext™ durante 15 minutos a 20° C. Los enzimas fueron entonces inactivados por calor incubando la mezcla de la reacción a 75° C durante 5 minutos. La mezcla se enfrió a 4° C, y se llevó a cabo la adición de colas de dA del ADN de extremo romo usando 10 µl de la mezcla maestra de adición de colas de dA que contenía el fragmento de Klenow (3' a 5' menos exo) (Conjunto 1 de Reactivos de ADN de Preparación de Muestras NEBNext™), e incubando durante 15 minutos a 37° C. Posteriormente, el fragmento de Klenow se inactivó por calor incubando la mezcla de la reacción a 75° C durante 5 minutos. Después de la inactivación del fragmento de Klenow, se usó 1 µl de una dilución 1:5 de Illumina Genomic Adaptor Oligo Mix (Parte No. 1000521; Illumina Inc., Hayward, CA) para ligar los adaptadores de Illumina (sin Índice Adaptadores Y) al ADN de la adición de colas de dA usando 4 µl de la ADN ligasa T4 proporcionada en el Conjunto 1 de Reactivos de ADN de Preparación de Muestras de ADN NEBNext™, incubando la mezcla de la reacción durante 15 minutos a 25° C. La mezcla se enfrió a 4° C, y el ADN libre de células ligado por adaptadores se purificó a partir de adaptadores no ligados, dímeros de adaptadores y otros reactivos usando microesferas magnéticas proporcionadas en el sistema de purificación Agencourt AMPure XP PCR (Parte No. A63881; Beckman Coulter Genomics, Danvers, MA). Se realizaron dieciocho ciclos de PCR para enriquecer selectivamente el ADN libre de células ligado por adaptadores (25 µl) usando Phusion® High-Fidelity Master Mix (25ml; Finnzymes, Woburn, MA) y cebadores de PCR de Illumina 0,5 µM cada uno) complementarios a los adaptadores (Partes No. 1000537 y 1000537). El ADN ligado por adaptadores se sometió a PCR (98° C durante 30 segundos; 18 ciclos de 98° C durante 10 segundos, 65° C durante 30 segundos y 72° C durante 30; extensión final a 72° C durante 5 minutos, y mantener a 4° C) usando Cebadores de PCR Genómicos Illumina (Partes Nos. 100537 y 1000538) y la Mezcla Maestra de PCR Phusion HF proporcionado en el Conjunto 1 de Reactivos de ADN de Preparación de Muestras de ADN NEBNext™, de acuerdo con las instrucciones del fabricante. El producto amplificado se purificó usando el sistema de purificación Agencourt AMPure XP PCR (Agencourt Bioscience Corporation, Beverly, MA), de acuerdo con las instrucciones del fabricante disponibles en www.beckmangenomics.com/products/AMPmeXPPProtocol_000387v001.pdf. El producto amplificado purificado se eluyó en 40 µl de tampón Qiagen EB, y la concentración y la distribución del tamaño de las bibliotecas

amplificadas se analizó usando el Kit Agilent DNA 1000 para el 2100 Bioanalizador (Agilent technologies Inc., Santa Clara, CA).

b. Secuenciación

La secuenciación del ADN de la biblioteca se realizó usando el Analizador de Genomas II (Illumina Inc., San Diego, CA, USA) de acuerdo con los protocolos del fabricante estándar. Se pueden encontrar copias del protocolo para la secuenciación del genoma completo usando tecnología Illumina/Solexa en BioTechniques.RTM. Protocol Guide 2007 Publicado en Diciembre del 2006: p 29, y en la red mundial en biotechniques.com/default.asp?page=protocol&subsection=article_display&id=112378. La biblioteca de ADN se diluyó a 1 nM y se desnaturalizó. La biblioteca de ADN (5pM) se sometió a amplificación de clústeres de acuerdo con el procedimiento descrito en la Cluster Station User Guide and Cluster Station Operations Guide de Illumina, disponible en la red mundial en illumina.com/systems/genome_analyzer/cluster_station.ilmn. El ADN amplificado se secuenció usando el Analizador de Genomas II de Illumina para obtener lecturas de extremos individuales de 36bp. Sólo se necesitan alrededor de 30bp de la información de secuencia aleatoria para identificar una secuencia como perteneciente a un cromosoma humano específico. Las secuencias más largas pueden identificar objetivos más particulares. En el presente caso, se obtuvieron un gran número de las lecturas de 36 bp, cubriendo aproximadamente un 10% del genoma.

Ejemplo 3

Análisis de los datos de secuenciación para la determinación de aneuploidía y fracción fetal

a. Análisis de datos de secuenciación para la determinación de aneuploidía

Al completarse la secuenciación de la muestra, el "Software de Control del Secuenciador" de Illumina transfirió los archivos de imagen y designación de bases a un servidor Unix que estaba ejecutando el software de Illumina "Genome Analyzer Pipeline" versión 1.51. Se ejecutó el programa "Gerald" de Illumina para alinear las secuencias, es decir lecturas de 36bp, al genoma humano de referencia hg18 proporcionado por el National Center for Biotechnology Information (NCBI36/hg18, disponible en la red mundial en <http://genome.ucsc.edu/cgi-bin/hgGateway?org=Human&db=hg18&hgsid=166260105>). Los datos de secuencia generados del procedimiento anterior que alinearon únicamente con el genoma se leyeron de las salidas de Gerald (archivos export.txt) por un programa (c2c.pl) ejecutado en un ordenador con el programa operativo Linux. Se permitieron los alineamientos de secuencia con discrepancias de bases y se incluyeron en los recuentos de alineamientos sólo si se alineaban únicamente con el genoma. Se excluyeron los alineamientos de secuencia con coordenadas de partida y finales idénticas (duplicados).

Se mapearon entre alrededor de 5 y 15 millones de etiquetas de 36 bp con 2 o menos discrepancias únicamente con el genoma humano. Todas las etiquetas mapeadas se contaron y se incluyeron en el cálculo de dosis de cromosomas en ambas muestras de ensayo y calificación. Las regiones que se extienden desde la base 0 a la base 2×10^6 , base 10×10^6 a base 13×10^6 y base 23×10^6 al final del cromosoma Y, fueron excluidas específicamente del análisis debido a que las etiquetas derivas de fetos masculinos o femeninos mapean para estas regiones del cromosoma Y.

b. Análisis de los datos de secuenciación para la determinación de la fracción fetal

Concomitante al análisis para determinar aneuploidía, los datos de secuenciación se analizaron para determinar la fracción fetal. Después de la transferencia de los archivos de imagen y designación de bases al servidor Unix ejecutando el software Illumina "Genome Analyzer Pipeline" versión 1.51 como se describe en el Ejemplo 2c, las lecturas de 36bp se alinearon con un 'genoma de SNP' usando el programa BOWTIE. El genoma de SNP se identificó como la agrupación de 30 secuencias de ADN, es decir SEQ ID NOS: 1-30, que abarcan los alelos de las 15 SNP divulgadas en la Tabla 5 en el Ejemplo 5. Sólo se usaron lecturas que mapeaban únicamente para el genoma de SNP para el análisis de la fracción fetal. Las lecturas que equivalían perfectamente con el genoma de SNP se contaron como etiquetas y se filtraron. De las lecturas restantes, sólo se contaron como etiquetas y se incluyeron en el análisis lecturas que tenían una o dos discrepancias. Se contaron las etiquetas mapeadas para cada uno de los alelos de SNP, y se determinó la fracción fetal como se describe en el Ejemplo 6.

Ejemplo 4

Identificación de Cromosomas Normalizadores para Determinar Aneuploidía

Para identificar cromosomas normalizadores para usarse en la determinación de dosis de cromosomas y la posterior presencia o ausencia de aneuploidía, se obtuvo ADN libre de células de plasma de sangre periférica de 48 voluntarios embarazados como se describe en el Ejemplo 1, y se secuenció como se describe en el Ejemplo 2. Los datos de secuenciación proporcionados en este ejemplo se obtuvieron de la secuenciación de una biblioteca construida de ADN libre de células fetal y materno que había sido enriquecido para ácidos nucleicos objetivo

comprendidos en una segunda biblioteca de secuenciación que ha sido construida de secuencias amplificadas que contienen SNPs como se describe a continuación.

Se determinó el número total de etiquetas de secuencia que se mapearon para cada cromosoma en el genoma de referencia (densidad de etiqueta de secuencia). Alternativamente, el número de etiquetas de secuencia mapeadas puede normalizarse a la longitud del cromosoma para generar una proporción de densidad de etiqueta de secuencia. La normalización a la longitud del cromosoma no es un paso requerido, y puede ser realizada únicamente para reducir el número de dígitos en un número para simplificarlo para la interpretación humana. Las longitudes de cromosomas que se pueden usar para normalizar los recuentos de etiqueta de secuencia pueden ser las longitudes proporcionadas en la red mundial en genome.ucsc.edu/goldenPath/stats.html#hg18.

La Tabla 1 proporciona la proporción calculada para los cromosomas X e Y, y los autosomas 1-22 en una muestra de ADN libre de células ejemplar (11351; 46,XY).

TABLA 1

Densidad de Etiqueta de Secuencia para los Cromosomas 1-22, X e Y (n=1; muestra 11351, 46 XY)	
Nombre del Cromosoma	Densidad de Etiqueta de Secuencia
chr1	1,857,858
chr2	1,910,676
chr3	1,562,572
chr4	1,376,498
chr5	1,383,453
chr6	1,317,821
chr7	1,192,136
chr8	1,162,856
chr9	914,624
chr10	1,112,763
chr11	1,093,028
chr12	1,051,209
chr13	717,684
chr14	710,878
chr15	675,596
chr16	683,529
chr17	647,571
chr18	615,140
chr19	432,191
chr20	557,068
chr21	284,701
chr22	305,365
chrX	1,060,456
chrY	5380

La densidad de etiqueta de secuencia resultante para cada cromosoma se relacionó con la densidad de etiqueta de secuencia de cada uno de los cromosomas restantes para derivar una dosis de cromosoma calificada, que se calculó como la proporción de la densidad de etiqueta de secuencia para el cromosoma de interés, por ejemplo cromosoma 21, y la densidad de etiqueta de secuencia para cada uno de los cromosomas restantes, es

decir cromosomas 1-20, 22 y X. Las dosis de cromosomas se determinaron para todos los cromosomas en todas las muestras, y las dosis medias para los cromosomas de interés 13, 18, 21, X e Y en las muestras calificadas se proporcionan en la Tabla 2 y se representan en las Figuras 7-11. Las Figuras 7-11 también representan las dosis de cromosomas para las muestras de ensayo. Las dosis de cromosomas para cada uno de los cromosomas de interés en las muestras calificadas proporcionan una medición de la variación en el número total de etiquetas de secuencia mapeadas para cada cromosoma de interés en relación a la de cada uno de los cromosomas restantes. Por lo tanto, las dosis de cromosomas calificadas pueden identificar el cromosoma o un grupo de cromosomas, es decir cromosoma normalizador, que tiene una variación entre muestras que está más cercana a la variación del cromosoma de interés, y que serviría como secuencias ideales para valores normalizadores para evaluación estadística adicional. Las Figuras 12 y 13 representan las dosis de cromosomas medias calculadas determinadas en una población de muestras calificadas para los cromosomas 13, 18 y 21, y los cromosomas X e Y.

En algunas situaciones, el mejor cromosoma normalizador puede no tener la menor variación, pero puede tener una distribución de dosis calificadas que distingue mejor una muestra o muestras de ensayo de las muestras calificadas, es decir el mejor cromosoma normalizador puede no tener la variación más baja, pero puede tener la diferenciabilidad más alta. Por lo tanto, la diferenciabilidad cuenta para la variación en la dosis de cromosoma y la distribución de las dosis en las muestras calificadas.

Las Tablas 3 y 4 proporcionan el coeficiente de variación como la medición de la variabilidad, y los valores de la prueba t de student como una medición de diferenciabilidad para los cromosomas 18, 21, X e Y, en donde cuanto más pequeño el valor de la prueba T, mayor es la diferenciabilidad. La diferenciabilidad para el cromosoma 13 se determinó como la proporción de diferencia entre la dosis de cromosoma media en las muestras calificadas y la dosis para el cromosoma 13 en la única muestra de prueba de T13, y la desviación estándar de la media de la dosis calificada.

Las dosis de cromosomas calificadas también sirven como la base para determinar los valores límite cuando se identifican aneuploidías en muestras de ensayo como se describe a continuación.

TABLA 2

Dosis de Cromosoma Calificada para los Cromosomas 13, 18, 21, X e Y (n=1; muestra 11351, 46 XY)					
Cromosoma	chr 21	chr 18	chr 13	chr X	chrY
chr1	0.153242	0.331102	0.386296	0.570795	0.002896
chr2	0.149005	0.321949	0.375618	0.555016	0.002816
chr3	0.1822	0.393671	0.459297	0.678661	0.003443
chr4	0.20683	0.446888	0.521384	0.770401	0.003908
chr5	0.20579	0.444641	0.518763	0.766528	0.003889
chr6	0.216039	0.466786	0.544599	0.804704	0.004082
chr7	0.238816	0.515998	0.602015	0.889543	0.004513
chr8	0.244829	0.528991	0.617174	0.911941	0.004627
chr9	0.311277	0.672561	0.784677	1.159445	0.005882
chr10	0.255851	0.552804	0.644957	0.952994	0.004835
chr11	0.26047	0.562785	0.656602	0.9702	0.004922
chr12	0.270832	0.585174	0.682722	1.008797	0.005118
chr13	0.396694	0.857118	1	1.477609	0.007496
chr14	0.400492	0.865324	1.009574	1.491755	0.007568
chr15	0.421407	0.910515	1.062298	1.56966	0.007963
chr16	0.416516	0.899947	1.049969	1.551443	0.007871
chr17	0.439644	0.949919	1.108271	1.63759	0.008308
chr18	0.462823	1	1.1667	1.723926	0.008746
chr19	0.658739	1.423306	1.660571	2.453674	0.012448
chr20	0.51107	1.104246	1.288324	1.903638	0.009658
chr21	1	2.160653	2.520834	3.724806	0.018897
chr22	0.93233	2.014442	2.35025	3.472749	0.017618
chrX	0.26847	0.580071	0.676769	1	0.005073
chrY	52.9184	114.3383	133.3985	197.1108	1

TABLA 3

Dosis de Cromosoma Calificada, Varianza y Diferenciabilidad para cromosomas 21 y 18								
	21 (n=35)				18 (n=40)			
	Media	Des Est	CV	Prueba T	Media	Des Est	CV	Prueba T
chr1	0.15332	0.002129	1.39	1.06E-10	0.32451	0.008954	2.76	2.74E-03
chr2	0.15106	0.002053	1.36	8.52E-08	0.31984	0.001783	0.56	5.32E-05
chr3	0.18654	0.004402	2.36	8.07E-07	0.39511	0.002364	0.60	1.93E-05

(continuada)

Dosis de Cromosoma Calificada, Varianza y Diferenciabilidad para cromosomas 21 y 18								
	21 (n=35)				18 (n=40)			
	Media	Des Est	CV	Prueba T	Media	Des Est	CV	Prueba T
chr4	0.21578	0.011174	5.18	1.47E-04	0.45714	0.014794	3.24	1.37E-03
chr5	0.21068	0.005332	2.53	1.08E-06	0.44626	0.003250	0.73	3.18E-05
chr6	0.22112	0.005453	2.47	1.74E-06	0.46818	0.003434	0.73	2.24E-05
chr7	0.24233	0.002314	0.96	2.39E-08	0.51341	0.005289	1.03	1.24E-04
chr8	0.24975	0.003772	1.51	1.06E-07	0.52898	0.002161	0.41	6.32E-05
chr9	0.31217	0.003050	0.98	1.60E-09	0.66100	0.014413	2.18	8.17E-04
chr10	0.25550	0.003164	1.24	2.42E-11	0.54091	0.013953	2.58	2.26E-03
chr11	0.26053	0.002596	1.00	1.32E-10	0.55158	0.013283	2.41	1.29E-03
chr12	0.27401	0.002061	0.75	1.40E-08	0.58032	0.007198	1.24	1.57E-04
chr13	0.41039	0.017637	4.30	3.09E-05	0.86961	0.021614	2.49	2.36E-04
chr14	0.40482	0.002908	0.72	1.10E-08	0.85732	0.011748	1.37	2.16E-04
chr15	0.41821	0.008238	1.97	1.24E-10	0.88503	0.029199	3.30	5.72E-03
chr16	0.40668	0.021232	5.22	2.91E-05	0.86145	0.056245	6.53	1.04E-01
chr17	0.42591	0.027001	6.34	5.85E-04	0.90135	0.068151	7.56	1.24E-01
chr18	0.46529	0.016239	3.49	8.02E-09				
chr19	0.63003	0.063272	10.04	3.30E-02	1.33522	0.150794	11.29	3.04E-01
chr20	0.49925	0.023907	4.79	1.65E-05	1.05648	0.064440	6.10	7.98E-02
chr21					2.06768	0.087175	4.22	5.10E-05
chr22	0.88726	0.083330	9.39	3.43E-02	1.87509	0.198316	10.58	2.43E-01
chrX	0.27398	0.016109	5.88	1.16E-04	0.58665	0.027280	4.65	7.50E-02

TABLA 4

Dosis de Cromosoma Calificada, Varianza y Diferenciabilidad para cromosomas 13, X e Y								
	13 (n=47)				X (n=20)			
	Media	Des Est	CV	Dif	Media	Des Est	CV	Prueba T
chr1	0.37213	0.018589	5.00	2.41	0.58035	0.02706	4.66	5.68E-05
chr2	0.36707	0.010067	2.74	3.03	0.57260	0.01432	2.50	1.53E-09
chr3	0.45354	0.008121	1.79	3.67	0.70741	0.01126	1.59	9.04E-13
chr4	0.52543	0.005306	1.01	2.39	0.82144	0.01192	1.45	5.86E-16
chr5	0.51228	0.008273	1.61	3.95	0.79921	0.01100	1.38	2.32E-13
chr6	0.53756	0.008901	1.66	3.91	0.83880	0.01261	1.50	3.64E-13
chr7	0.58908	0.018508	3.14	2.83	0.91927	0.02700	2.94	1.86E-08
chr8	0.60695	0.015797	2.60	3.05	0.94675	0.02173	2.30	3.40E-10
chr9	0.75816	0.033107	4.37	2.59	1.18180	0.04827	4.08	9.63E-06
chr10	0.62018	0.029891	4.82	2.56	0.96642	0.04257	4.40	4.55E-05

(continuada)

Dosis de Cromosoma Calificada, Varianza y Diferenciabilidad para cromosomas 13, X e Y								
	13 (n=47)				X (n=20)			
	Media	Des Est	CV	Dif	Media	Des Est	CV	Prueba T
chr11	0.63248	0.029204	4.62	2.55	0.98643	0.04222	4.28	1.82E-05
chr12	0.66574	0.023047	3.46	2.76	1.03840	0.03301	3.18	1.26E-07
chr13					1.56355	0.01370	0.88	6.33E-17
chr14	0.98358	0.035331	3.59	2.67	1.58114	0.08076	5.11	2.29E-04
chr15	1.01432	0.055806	5.50	2.39	1.53464	0.12719	8.29	2.01E-02
chr16	0.98577	0.085933	8.72	2.17	1.61094	0.14829	9.21	2.68E-02
chr17	1.03217	0.100389	9.73	2.13	1.74904	0.07290	4.17	1.62E-04
chr18	1.13489	0.040058	3.53	2.62	2.38397	0.30515	12.80	1.07E-01
chr19	1.52678	0.203732	13.34	1.98	1.88186	0.14674	7.80	1.56E-02
chr20	1.20919	0.100371	8.30	2.27	3.71853	0.22406	6.03	4.21E-04
chr21	2.38087	0.132418	5.56	2.29	3.35158	0.40246	12.01	8.66E-02
chr22	2.14557	0.271281	12.64	2.13	0.58035	0.02706	4.66	5.68E-05
chrX	0.66883	0.029157	4.36	1.04				
chr2-6	0.46965	0.006987	1.49	4.17				
chr3-6	0.50496	0.005373	1.06	5.16				
	Y (n=25)							
	Media	Des Est	CV	Prueba T				
Chr 1-22,X	0.00728	0.00227	31.19	1.30E-13				

Los ejemplos de diagnósticos de T21, T13, T18 y síndrome de Turner obtenidos usando los cromosomas normalizadores, dosis de cromosomas y diferenciabilidad para cada uno de los cromosomas de interés se describen en el Ejemplo 6.

Ejemplo 5

Selección de SNPs autosómicos para la determinación de la fracción fetal

Se seleccionó un conjunto de 28 SNPs autosómicos de una lista de 92 SNPs (Pakstis et al., Hum Genet 127:315-324 [2010]), y de las secuencias de SNP disponibles en Applied Biosystems en la dirección de la red mundial appliedbiosystems.com, y validada para su uso en amplificación por PCR multiplexado y para secuenciación masivamente paralela. Los cebadores se diseñaron para hibridar con una secuencia cercana con el sitio de SNPs en el ADN libre de células para asegurar que se incluya en la lectura de 36 bp generada de la secuenciación masivamente paralela en el Analizador GII Illumina, y para generar amplicones de longitud suficiente para someterse a amplificación por puente durante la formación de clústeres. Por lo tanto, los cebadores se diseñaron para generar amplicones que fueron de al menos 110 bp, que cuando se combinan con los adaptadores universales (Illumina Inc., San Diego, CA) usados para amplificación de clústeres, resultó en moléculas de ADN de al menos 200 bp. Se identificaron secuencias de cebadores, y los conjuntos de cebadores, es decir cebadores directos e inversos, se sintetizaron por Integrated DNA Technologies (San Diego, CA), y se almacenaron como una solución de 1µM para ser usada para amplificar secuencias objetivo polimórficas como se describe en los Ejemplos 5-8. La Tabla 5 proporciona los números de ID de entrada de los RefSNP (rs), los cebadores usados para amplificar la secuencia de ADN libre de células objetivo, y las secuencias de los amplicones que comprenden los posibles alelos de SNP que serían generados usando los cebadores. Los SNPs dados en la Tabla 5 se usaron para la amplificación simultánea de 13 secuencias objetivo en un ensayo multiplexado. El panel proporcionado en la Tabla 5 es un panel de SNP ejemplar. Se pueden emplear menos o más SNPs para enriquecer el ADN fetal y materno para ácidos nucleicos objetivo polimórficos. SNPs adicionales que pueden usarse incluyen los SNPs dados en la Tabla 6. Los alelos de SNP se muestran en negrita y están subrayados. Otros SNPs que se pueden usar para determinar la fracción fetal

de acuerdo con el presente método incluyen rs315791, rs3780962, rs1410059, rs279844, rs38882, rs9951171, rs214955, rs6444724, rs2503107, rs1019029, rs1413212, rs1031825, rs891700, rs1005533, rs2831700, rs354439, rs1979255, rs1454361, rs8037429, y rs1490413, que han sido analizados para determinar la fracción fetal por PCR TaqMan y se divulgan en la solicitudes Provisionales US 61/296,358 y 61/360,837.

5

10

15

20

25

30

35

40

45

50

55

60

65

TABLA 5

Panel SNP Panel para la Determinación de la Fracción Fetal					
SNP ID	Chr	Amplificación: Alelo 1	Amplificación: Alelo 2	Secuencia del Cebador Directo, nombre y SEQ ID NO:	Secuencia del Cebador Inverso, nombre y SEQ ID NO:
rs560681	1	CACATGCACAGCCA	CACATGCACAGCCA	CACATGCA	CCCCAAGGTC
		GCAACCCCTGTCAGC	GCAACCCCTGTCAGC	CAGCCAGC	CTGTGACCTGA
		AGGAGTTCCACCA	AGGAGTTCCACCA	AACCC	GT
		GTTTCTTTCTGAGAA	GTTTCTTTCTGAGAA	(rs560681_C	(rs560681_C1_1_
		CATCTGTCAGGTTT	CATCTGTCAGGTTT	1_1_F; SEQ	R; SEQ ID
rs1109037	2	CTCTCCATCTCTATT	CTCTCCATCTCTGTT	ID NO:57)	NO:58)
		TACTCAGGTCACAG	TACTCAGGTCACAG		
		GACCTTGGGG (SEQ	GACCTTGGGG (SEQ		
		ID NO:1)	ID NO:2)		
		TGAGGAAAGTGAGGC	TGAGGAAAGTGAGGC	TGAGGAAG	TGCCAGTGCG
rs9866013	3	TCAGAGGGTAAGAA	TCAGAGGGTAAGAA	TGAGGCTC	AGATGAAAGT
		ACTTTGTCACAGAGC	ACTTTGTCACAGAGC	AGAGGGT	CTTT
		TGGTGGTGAGGGTG	TGGTGGTGAGGGTG	(rs110937_C	(rs110937_C2_1_
		GAGATTTTACACTCC	GAGATTTTACACTCC	2_1_F; SEQ	R; SEQ ID
		CTGCCTCCACACCA	CTGCCTCCACACCA	ID NO:59)	NO:60)
rs9866013	3	GTTTCTCCAGAGTGG	GTTTCTCCGGAGTGG		
		AAAGACTTTTCATCTC	AAAGACTTTTCATCTC		
		GCACTGGCA (SEQ ID	GCACTGGCA (SEQ ID		
		NO:3)	NO:4)		
		GTGCCTTCAGAACCT	GTGCCTTCAGAACCT	GTGCCTTC	TCCCATCCAC
rs9866013	3	TTGAGATCTGATTCT	TTGAGATCTGATTCT	AGAACCTT	CAGCCACCC

Panel SNP Panel para la Determinación de la Fracción Fetal					
SNP ID	Chr	Amplificación: Alelo 1	Amplificación: Alelo 2	Secuencia del Cebador Directo, nombre y SEQ ID NO:	Secuencia del Cebador Inverso, nombre y SEQ ID NO:
rs13182883	5	ATTTTAAAGCTTCT TAGAAGAGAGATTG CAAAGTGGGTGTGTTT CTCTAGCCAGACAG GGCAGGCAAAATAGG GGTGGCTGGTGGGA TGGGA (SEQ ID NO:5)	ATTTTAAAGCTTCT TAGAAGAGAGATTG CAAAGTGGGTGTGTTT CTCTAGCCAGACAG GGCAGGTAAATAGG GGTGGCTGGTGGGA TGGGA (SEQ ID NO:6)	TGAGATCT GAT (rs9866013_C3_1_F; SEQ ID NO:61)	(rs9866013_C3_1_R; SEQ ID NO:62)
		AGGTGTGTCTCTCTT TTGTGAGGGGAGGG GTCCCTTCTGGCCTA GTAGAGGGCCTGGC CTGCAGTGAGCATT AAATCCTCAAGGAA CAGGTGGGAGGT GGGACAAAGG (SEQ ID NO:7)	AGGTGTGTCTCTCTT TTGTGAGGGGAGGG GTCCCTTCTGGCCTA GTAGAGGGCCTGGC CTGCAGTGAGCATT AAATCCTCAGGAA CAGGTGGGAGGT GGGACAAAGG (SEQ ID NO:8)	AGGTGTGT CTCTCTTTT GTGAGGGG (rs13182883_C5_1_F; SEQ ID NO:63)	CCTTTGTCCCA CCTCCCCACC (rs13182883_C5_1_R; SEQ ID NO:64)

Panel SNP Panel para la Determinación de la Fracción Fetal					
SNP ID	Chr	Amplificación: Alelo 1	Amplificación: Alelo 2	Secuencia del Cebador Directo, nombre y SEQ ID NO:	Secuencia del Cebador Inverso, nombre y SEQ ID NO:
rs13218440	6	CCTCGCCTACTGTGC TGTTTCTAACCATCA TGCTTTTCCCTGAAT CTCTTGAGTCTTTT CTGCTGTGGACTGA AACTTGATCCTGAG ATTCACCTCTAGTCC CTCTGAGCAGCCTCC TGGAATACTCAGCT GGGATGG (SEQ ID NO:9)	CCTCGCCTACTGTGC TGTTTCTAACCATCA TGCTTTTCCCTGAAT CTCTTGAGTCTTTT CTGCTGTGGACTGA AACTTGATCCTGAG ATTCACCTCTAGTCC CTCTGAGCAGCCTCC TGGAATACTCAGCT GGGATGG (SEQ ID NO:10)	CCTCGCCT ACTGTGCT GTTTCTAA CC (rs13218440_ C6_1_F; SEQ ID NO:65)	CCATCCCAGCT GAGTATTCCA GGAG (rs13218440_ C6_1_R; SEQ ID NO:66)
		AATTGCAATGGTGA GAGGTTGATGGTAA AATCAAACGGAACT TGTTAATTTGTCAAT CTGATGGACTGGAA CTGAGGATTTTCAAT TTCCCTCTCCCAACCCA AGACACTTCTCACTG G (SEQ ID NO:11)	AATTGCAATGGTGA GAGGTTGATGGTAA AATCAAACGGAACT TGTTAATTTGTCAAT CTGATGGACTGGAA CTGAGGATTTTCAAT TTCCCTCTCCCAACCCA AGACACTTCTCACTG G (SEQ ID NO:12)	AATTGCAA TGGTGAGA GGTTGATG GT (SEQ ID NO:67)	CCAGTGAGAA GTGTCTTGGGT TGG (SEQ ID NO:68)
rs7041158	9	AATTGCAATGGTGA GAGGTTGATGGTAA AATCAAACGGAACT TGTTAATTTGTCAAT CTGATGGACTGGAA CTGAGGATTTTCAAT TTCCCTCTCCCAACCCA AGACACTTCTCACTG G (SEQ ID NO:11)	AATTGCAATGGTGA GAGGTTGATGGTAA AATCAAACGGAACT TGTTAATTTGTCAAT CTGATGGACTGGAA CTGAGGATTTTCAAT TTCCCTCTCCCAACCCA AGACACTTCTCACTG G (SEQ ID NO:12)	AATTGCAA TGGTGAGA GGTTGATG GT (SEQ ID NO:67)	CCAGTGAGAA GTGTCTTGGGT TGG (SEQ ID NO:68)

Panel SNP Panel para la Determinación de la Fracción Fetal					
SNP ID	Chr	Amplición: Alelo 1	Amplición: Alelo 2	Secuencia del Cebador Directo, nombre y SEQ ID NO:	Secuencia del Cebador Inverso, nombre y SEQ ID NO:
rs740598	10	GAAATGCCTTCTCAG	GAAATGCCTTCTCAG	GAAATGCC	GGTTTGAGCA
		GTAATGGAAGGTTA	GTAATGGAAGGTTA	TTCTCAGG	GTTCTGAGAAT
		TCCAAATATTTTCG	TCCAAATATTTTCG	TAATGGAA	GTGGCT (SEQ
		TAAGTATTTCAAATA	TAAGTATTTCAAATA	GGT (SEQ	ID NO:70)
		GCAATGGCTCGTCTA	GCAATGGCTCGTCTA	ID NO:69)	
rs10773760	12	TGGTAGTCTCAG	TGGTAGTCTCGCAG		
		CCACATTCTCAGAAC	CCACATTCTCAGAAC		
		TGCTCAAACC (SEQ	TGCTCAAACC (SEQ		
		ID NO:13)	ID NO:14)		
		ACCCAAAACACTGG	ACCCAAAACACTGG	ACCCAAAA	CCCTTATCTGC
		AGGGCCTCTTCTCA	AGGGCCTCTTCTCA	CACTGGAG	TATGTGGCATA
		TTTTCGGTAGACTGC	TTTTCGGTAGACTGC	GGGCCT	CTTGG (SEQ ID
		AAGTGTTAGCCGTC	AAGTGTTAGCCGTC	(SEQ ID	NO:72)
		GGGACCAGCTTCTGT	GGGACCAGCTTCTGT	NO:71)	
		CTGGAAGTTCGTCA	CTGGAAGTTCGTCA		
		AATTGCAGTTAAGTC	AATTGCAGTTAGGT		
		CAAGTATGCCACAT	CCAAGTATGCCACA		
		AGCAGATAAGGG	TAGCAGATAAGGG		
		(SEQ ID NO:15)	(SEQ ID NO:16)		

Panel SNP Panel para la Determinación de la Fracción Fetal					
SNP ID	Chr	Amplificón: Alelo 1	Amplificón: Alelo 2	Secuencia del Cebador Directo, nombre y SEQ ID NO:	Secuencia del Cebador Inverso, nombre y SEQ ID NO:
rs4530059	14	GCACCAGAAATTTAA	GCACCAGAATTTAA	GCACCAGA	GCACCTGACA
		ACAACGGCTGACAAT	ACAACGGCTGACAAT	ATTTAAC	GGCACATCAG
		AAATATGCAGTCGA	AAATATGCAGTCGA	AACGCTGA	CG (SEQ ID
		TGATGACTTCCCAGA	TGATGACTTCCCAGA	CAA (SEQ	NO:74)
		GCTCCAGAAAGCAAC	GCTCCAGAAAGCAAC	ID NO:73)	
rs7205345	16	TCCAGCACACAGAG	TCCAGCACACGGAG		
		AGGCGCTGATGTGC	AGGCGCTGATGTGC		
		CTGTCAGGTGC (SEQ	CTGTCAGGTGC (SEQ		
		ID NO:17)	ID NO:18)		
		TGACTGTATACCCCA	TGACTGTATACCCCA	TGACTGTA	GCACTAAAGGA
		GGTGACCCCTTGGGT	GGTGACCCCTTGGGT	TACCCAG	TGTGGAAAGTCT
		CATCTCTATCATAGA	CATCTCTATCATAGA	GTGCACCC	AGTGTG (SEQ
		ACTTATCTCACAGAG	ACTTATCTCACAGAG	(SEQ ID	ID NO:76)
		TATAAGAGCTGATTT	TATAAGAGCTGATTT	NO:75)	
		CTGTGCTGCCTCTC	CTGTGCTGCCTCTC		
		ACACTAGACTTCCAC	ACACTAGACTTCCAC		
		ATCCTTAGTGC (SEQ	ATCCTTAGTGC (SEQ		
		ID NO:19)	ID NO:20)		

Panel SNP Panel para la Determinación de la Fracción Fetal					
SNP ID	Chr	Amplicón: Alelo 1	Amplicón: Alelo 2	Secuencia del Cebador Directo, nombre y SEQ ID NO:	Secuencia del Cebador Inverso, nombre y SEQ ID NO:
rs8078417	17	TGTACGTGGTCACCA	TGTACGTGGTCACCA	TGTACGTG	AGTGTGAGAA
		GGGACGCCTGGCG	GGGACGCCTGGCG	GTCACCAG	GAGCCTCAAG
	17	CTGCGAGGGAGGCC	CTGCGAGGGAGGCC	GGGACG	GACAGC (SEQ
		CCGAGCCTCGTGCCC	CCGAGCCTCGTGCCC	(SEQ ID	ID NO:78)
		CCGTGAAGCTTCAG	CCGTGAAGCTTCAG	NO:77)	
		CTCCCCCTCCC <u>CG</u> GGCT	CTCCCCCTCCC <u>TGG</u> CT		
		GTCTTGAGGCTCTT	GTCTTGAGGCTCTT		
		CTCACACT (SEQ ID	CTCACACT (SEQ ID		
		NO:21)	NO:22)		
rs576261	19	CAGTGGACCCTGCT	CAGTGGACCCTGCT	CAGTGGAC	GTGGCAAAGG
		GCACCTTTCTCTCCCC	GCACCTTTCTCTCCCC	CCTGCTGC	AGAGAGTTGT
		TCCCATCAACCTCTT	TCCCATCAACCTCTT	ACCTT	GAGG (SEQ ID
		TTGTGCCTTCCCCCTC	TTGTGCCTTCCCCCTC	(SEQ ID	NO:80)
		CGTGACCACTTCT	CGTGACCACTTCT	NO:79)	
		CTGTCACCA <u>A</u> CCCTG	CTGTCACCA <u>C</u> CCCTG		
		GCCTCACAACTCTCT	GCCTCACAACTCTCT		
		CCTTTGCCAC (SEQ	CCTTTGCCAC (SEQ		
		ID NO:23)	ID NO:24)		

Panel SNP Panel para la Determinación de la Fracción Fetal					
SNP ID	Chr	Amplicón: Alelo 1	Amplicón: Alelo 2	Secuencia del Cebador Directo, nombre y SEQ ID NO:	Secuencia del Cebador Inverso, nombre y SEQ ID NO:
rs2567608	20	CAGTGGCATAAGTAG	CAGTGGCATAAGTAG	CAGTGGCA	CCTCTCCGACA
		TCCAGGGGCTCCTCC	TCCAGGGGCTCCTCC	TAGTAGTC	ACTTCCGCCG
		TCAGCACCTCCAGC	TCAGCACCTCCAGC	CAGGGGCT	(SEQ ID NO:82)
		ACCTTCCAGGAGGC	ACCTTCCAGGAGGC	(SEQ ID	
		AGCAGCGCAGGCAG	AGCAGCGCAGGCAG	NO:81)	
		AGAACCCGCTGGAA	AGAACCCGCTGGAA		
		GAAATCGGCGGAAGT	GGAATCGGCGGAAGT		
		TGTCGGAGAGG (SEQ ID NO:25)	TGTCGGAGAGG (SEQ ID NO:26)		

TABLA6

Panel SNP Panel para la Determinación de la Fracción Fetal				
SNP ID	Chr	Amplicón: Alelo 1	Amplicón: Alelo 2	Secuencia del Cebador Directo, nombre y SEQ ID NO:
rs430046	16	AGGTCTGGGGGCC GCTGAATGCCAAGC TGCGAATCTTAAAT GTTAAGGAACAAG GTCATACAAATGAAT GGTGTGATGTAAAA GCTTGGGAGGTGAT TTCTGAGGGTAGGT GCTGGGTTAATGG GAGGA (SEQ ID NO:27)	AGGTCTGGGGGGCC TGAATGCCAAGCTGG GAATCTTAAATGTTA AGGAACAAGGTCATA CAATGAATGTTGTGA TGTAAGAAGCTTGGGA GGTGATTTTGTAGGG TAGGTGCTGGGTTTA ATGGGAGGA (SEQ ID NO:28)	AGGTCTGG GGGCGCT GAAT (rs430046_C 1_1_F; SEQ ID NO:83)
				TCCTCCCATTA AACCAGCAC CT (rs430046_C1_1_ R; SEQ ID NO:84)
rs9951171	18	ACGGTTCTGTCCCTG TAGGGGAGAAAAAG TCCTCGTTGTTCT CTGGGATGCAACAT GAGAGAGCAGCAC ACTGAGGCTTTATG GATTGCCCTGCCAC AAGTGAACAGG (SEQ ID NO:29)	ACGGTTCTGTCCCTGT AGGGGAGAAAAAGTCC TCGTTGTTCCCTCTGGG ATGCAACATGAGAGA GCAGCACACTGAGGC TTTATGGGTTGCCCT GCCACAAGTGAACAG G (SEQ ID NO:30)	ACGGTTCT GTCCTGTA GGGAGAG (rs9951171_ C1_1_F; SEQ ID NO:85)
				CCTGTTCACCT GTGGCAGGGC A (rs9951171_C1_1_ _R; SEQ ID NO:86)

Panel SNP Panel para la Determinación de la Fracción Fetal					
SNP ID	Chr	Amplicón: Alelo 1	Amplicón: Alelo 2	Secuencia del Cebador Directo, nombre y SEQ ID NO:	Secuencia del Cebador Inverso, nombre y SEQ ID NO:
rs338882	5	GCGCAGTCAGATG	GCGCAGTCAGATGGG	GCGCAGTC	TCCAGCCCCTTG
		GGCGTGCTGGCGTC	CGTGCTGGCGTCTGT	AGATGGGC	TCCCAAACGT
		TGTCTTCTCTCTCTC	CTTCTCTCTCTCTCTC	GTGC	GT
		CTGCTCTCTGGCTT	TCCTGGCTTCATTTT	(rs338882_C	(rs338882_C1_1_
		CATTTTCTCTCTCTT	TCCTCTCTCTCTCTCTC	1_1_F; SEQ	R; SEQ ID
		CTGTCTCACCTTCT	ACCTTCTTTCGTGTGC	ID NO:87)	NO:88)
		TTCGTGTGCTGTG	CTGTGCA T ACACACG		
		CA C ACACACAGTTTG	TTTGGGACAAAGG		
		GGACAAGGG	CTGGA (SEQ ID NO:32)		
		CTGGA (SEQ ID NO:31)			
rs10776839	9	GCCGGACCTGCGA	GCCGGACCTGCGAAA	GCCGGACC	CGGGCAACTG
		AATCCCAAAATGCC	TCCCAAAATGCCAAA	TGCGAAAT	GGGCTCTGATC
		AAACATTCCCGCT	CATTCCCGCCTCACA	CCCAA	(rs10776839_C1_
		CACATGATCCACGA	TGATCCCAGAGAGAG	(rs10776839	1_R; SEQ ID
		GAGAGGGGACCCA	GGGACCCAGTGTCC	C1_1_F; SEQ	NO:90)
		GTGTTCCACAGCTTG	CAGCTTGACGTGAG	ID NO:89)	
		CAGCTGAGGAGCC	GAGCCCGAGTTTGCC		
		CGAGGTTGCCGTCA	GTCAGATCAGAGCCC		
		GATCAGAGCCCCA	CAGTTGCCCG (SEQ ID NO:34)		
		GTTGCCCG (SEQ ID NO:33)			

Panel SNP Panel para la Determinación de la Fracción Fetal					
SNP ID	Chr	Amplificación: Alelo 1	Amplificación: Alelo 2	Secuencia del Cebador Directo, nombre y SEQ ID NO:	Secuencia del Cebador Inverso, nombre y SEQ ID NO:
rs9905977	17	AGCAGCCTCCCTCG ACTAGCTCACACTA CGATAAGGAAAAAT TCATGAGCTGGTGT CCAAAGGAGGGCTG GGTGA C TCGTGGCT CAGTCAGCA T CAAG ATTCCTTTTCGTCTTT CCCCTCTGCC (SEQ ID NO:35)	AGCAGCCTCCCTCGA CTAGCTCACACTACG ATAAGGAAAAATTCAT GAGCTGGTGTCCAAAG GAGGGCTGGGTGACT CGTGGCTCAGTCAGC <u>G</u> TCAAAGATTCCCTTTC GTCTTTCCCTCTGTGCC (SEQ ID NO:36)	AGCAGCCT CCCTCGAC TAGCT (rs9905977_ C1_1_F; SEQ ID NO:91)	GGCAGAGGGG AAAGACGAAA GGA (rs9905977_C1_1_R; SEQ ID NO:92)
		TGGCATTGCCTGTA ATATACATAGCCAT GGTTTTTTATAGGC AATTAAAGATGAAT AGCTTCTAAACTAT AGATAAGTTTCATT ACCCAGGGAAGCT GAACTATAGCTACT TT A CCCCAAAATCAT TAGAATGGTGCTT (SEQ ID NO:37)	TGGCATTGCCTGTAA TATACATAGCCATGG TTTTTTATAGGCAATT TAAGATGAATAGCTT CTAAACTATAGATAA GTTTCATTACCCAG GAAGCTGAACTATAG CTACTTTCCCCAAAA TCATTAGAAATGGTGC TT (SEQ ID NO:38)	TGGCATTG CCTGTAAT ATACATAG (rs1277284_ C4_1_F; SEQ ID NO:93)	AAGCACCATT CTAATGATTTT GG (rs1277284_C4_1_R; SEQ ID NO:94)
rs1277284	4				

Panel SNP Panel para la Determinación de la Fracción Fetal					
SNP ID	Chr	Amplificación: Alelo 1	Amplificación: Alelo 2	Secuencia del Cebador Directo, nombre y SEQ ID NO:	Secuencia del Cebador Inverso, nombre y SEQ ID NO:
rs258684	7	ATGAAGCCTTCCAC	ATGAAGCCTTCCACC	ATGAAGCC	GATCAGTTGTT
		CAACTGCCTGTATG	AACTGCCTGTATGAC	TTCCACCA	GTTTCTATATT
		ACTCATCTGGGGAC	TCATCTGGGGACTTC	ACTG	TCCTT
		TTCTGCTCTATACT	TGCTCTATACTCAAA	(rs258684_C	(rs258684_C7_1_
		CAAAGTGGCTTAGT	GTGGCTTAGTCACTG	7_1_F; SEQ	R; SEQ ID
		CACTGCCAATGTAT	CCAATGTATTCCAT	ID NO:95)	NO:96)
		TTCCATATGAGGGA	ATGAGGGACGGTGAT		
		CGATGATTACTAAG	TACTAAGGAAATATA		
		GAAATATAGAAAC	GAAACAACAACCTGAT		
		AACAACCTGATC	C (SEQ ID NO:40)		
		(SEQ ID NO:39)			
		ACAACAGAAATCAG	ACAACAGAAATCAGGT	ACAACAGA	CTGAACTGAA
		GTGATTGGAGAAA	GATTGGAGAAAAGAT	ATCAGGTG	CAAGAAGATTA
		AGATCACAGGCCTA	CACAGGCCTAGGCAC	ATTGGA	AGGTC
		GGCACCCAAAGGCTT	CCAAAGGCTTGAAAGA	(rs1347696_	(rs1347696_C8_4
		GAAAGGATGAAAGA	TGAAAAGAATGAAAGA	C8_4_F; SEQ	_F; SEQ ID
		ATGAAAGATGGAC	TGGACGGAAAGAAAAT	ID NO:97)	NO:98)
		GGAAACAAAATTAG	TAGGACCTTAATTCTT		
		GACCTTAATTCTTT	TGTTCAAGTTTCAG (SEQ		
		GTTCAAGTTTCAG	ID NO:42)		
		(SEQ ID NO:41)			

Panel SNP Panel para la Determinación de la Fracción Fetal					
SNP ID	Chr	Amplicón: Alelo 1	Amplicón: Alelo 2	Secuencia del Cebador Directo, nombre y SEQ ID NO:	Secuencia del Cebador Inverso, nombre y SEQ ID NO:
rs508485	11	TTGGGGTAAATTTT CATTGTCATATGTG GAAATTTAAATATAC CATCATCTACAAAG AATTCCACAGAGTT AAATATCTTAAAGTT AAACACTTAAATA AGTGTTTGC GTGAT ATTTTGATGACAGAC TAAACAGAGTCTAA TTCCCAACCCC (SEQ ID NO:43)	TTGGGGTAAATTTTC ATTGTCATATGTGGA ATTTAAATATATACCAT CATCTACAAAGAAAT CCACAGAGTTAAATA TCTTAAGTTAAACAC TTAAAAATAAGTGTTT GCGTGATATTTTGAT GATAGATAAACACAGAG TCTAATTCCCAACCCC (SEQ ID NO:44)	TTGGGGTA AATTTTCA TTGTCA (rs508485_C 11_1_F; SEQ ID NO:99)	GGGGTGGGAA TTAGACTCTG (rs508485_C11_1 _R; SEQ ID NO100)

Panel SNP Panel para la Determinación de la Fracción Fetal					
SNP ID	Chr	Amplicón: Alelo 1	Amplicón: Alelo 2	Secuencia del Cebador Directo, nombre y SEQ ID NO:	Secuencia del Cebador Inverso, nombre y SEQ ID NO:
rs9788670	15	TGCAATTCAAATCA GGAAGTATGACCA AAAGACAGAGATC TTTTTTGGATGATC CCTAGCCTAGCAAT GCC TGGCAGCCATG CAGGTGCAATGTCA ACCTTAAATAATGT ATTGCAAACTCAGA GCTGACAAAACCTCG ATGTTGC (SEQ ID NO:45)	TGCAATTCAAATCAG GAAGTATGACCAAAA GACAGAGATCTTTT TGGATGATCCCTAGC CTAGCAATGCCTGGC AGCCATGCAGGTGCA ATGTCAACCTTAAAT AATGTATTGCAAAAT CAGAGCTGACAAACC TCGATGTTGC (SEQ ID NO:46)	TGCAATTC AAATCAGG AAGTATG (rs9788670_c 15_2_F; SEQ ID NO:101)	GCAACATCGA GGTTTGTCTAG (rs9788670_c15_2_R; SEQ ID NO:102)
		CTGTGCTCTGCGAA TAGCTGCAGAAAGTA ACTTGGGGACCCAA AATAAAGCAGAAAT GCTAATGTCAAGTC CTGAGAACCAAGC CCTGGGACTCTGGT GCCATTTTCGGATTTC TCCATGAGCATGGT (SEQ ID NO:47)	CTGTGCTCTGCGAAT AGCTGCAGAAAGTAAC TTGGGGACCCAAAT AAAGCAGAAATGCTAA TGTCAAAGTCCTGAGA ACCAAGCCCTGGGAC TCTGGTGCCATTTTG GATTCTCCATGAGCA TGGT (SEQ ID NO:48)	CTGTGCTC TGCGAATA GCTG (rs8137254_c 22_2_F; SEQ ID NO:103)	ACCATGCTCAT GGAGAAATCC (rs8137254_c22_2_R; SEQ ID NO:104)
rs8137254	22				

Panel SNP Panel para la Determinación de la Fracción Fetal				
SNP ID	Chr	Amplicón: Alelo 1	Amplicón: Alelo 2	Secuencia del Cebador Directo, nombre y SEQ ID NO:
rs3143	19	TTTTTCCAGCCAAC	TTTTTCCAGCCAAC	TTTTTCCA
		TCAAGGCCAAAAA	AAGGCCAAAAAAT	GCCAACTC
rs3143	19	AAATTCTCTTAATAT	TTCTTAATATAGTTAT	AAGG
		AGTTATTATGCGAG	TATGCGAGGGGAGGG	(rs3143_c19_2_F: SEQ ID NO:105)
rs3143	19	GGGAGGGGAAAGCA	GAAGCAAAAGGAGCA	2_F: SEQ ID NO:106)
		AAGGAGCACAGGT	CAGGTAGTCCACAGA	
rs3143	19	AGTCCACAGAATA	ATAGGACACAAAGAA	
		AGACACAAAGAAAC	ACCTCAAGCTGTG	
rs3143	19	CTCAAGCTGTG	(SEQ ID NO:50)	
rs3143	19	(SEQ ID NO:49)		
rs2182957	13	TCTTCTCGTCCCCT	TCTTCTCGTCCCCTAA	TCTTCTCG
		AAGCAAAACAACAT	GCAAAACAACATCCGC	TCCCCTAA
rs2182957	13	CCGCTTGCTTCTGT	TTGCTTCTGTCTGTGT	GCAA
		CTGTGTAACCCACAG	AACCACAGTGAAATGG	(rs2182957_c13_1_F: SEQ ID NO:107)
rs2182957	13	TGAATGGGTGTGCA	GTGTGCACGCTTGGT	
		CGCTTGATGGGCCT	GGGCTCTGAGCCCC	
rs2182957	13	CTGAGCCCCCTGTG	TGTTGCACAAACCAG	
		CACAAACCAGAAA	AAA (SEQ ID NO:52)	
rs2182957	13	(SEQ ID NO:51)		
rs2182957	13			TTTTCTGGTTTG
				TGCAACAGG
rs2182957	13			(rs2182957_c13_1_R: SEQ ID NO:108)

Panel SNP Panel para la Determinación de la Fracción Fetal					
SNP ID	Chr	Amplificación: Alelo 1	Amplificación: Alelo 2	Secuencia del Cebador Directo, nombre y SEQ ID NO:	Secuencia del Cebador Inverso, nombre y SEQ ID NO:
rs3739005	2	CACATGGGGGCATT AAGAATCGCCCAG GGAGGAGGAGGGA GAACGCGTGCTTTT CACATTTCATTTG AATTTTCGAGTTCC CAGGATGTGTTTTT GTGCTCATCGATGT (SEQ ID NO:53)	CACATGGGGGCATTA AGAAATCGCCCAGGA GGAGGAGGAGGAAC GCGTGCCTTTTCACATT TGCATTTGAATTTTIG AGTTCCCAGGATGTG TTTTTGTGCTCATCGA TGT (SEQ ID NO:54)	CACATGGG GGCATTAA GAAT (rs3739005_c 2_2_F; SEQ ID NO:109)	ACATCGATGA GCACAAAAAC AC (rs3739005_c2_2 _R; SEQ ID NO:110)
		GGGCTCTGAGGTGT GTGAAATAAAAAAC AAATGTCCATGTCT GTCCTTTTATGGCA TTTTGGGACTTTAC ATTTCAAACATTTT AGACATGTATCACA ACACGAAGGAATA ACAGTTCAGGGAT ATCT (SEQ ID NO:55)	GGGCTCTGAGGTGTG TGAAATAAAAAACAAA TGTCATGTCTGTCTCT TTTATGGCATTTTGGG ACTTTACATTTCAAA CATTTCAGACATGTA TCACAAACACGAGGGA ATAACAGTTCAGGG ATATCT (SEQ ID NO:56)	GGGCTCTG AGGTGTGT GAAA (rs530022_cl _2_F; SEQ ID NO:111)	AGATATCCCTG GAACTGTTATT CC (rs530022_cl_2_ R; SEQ ID NO:112)
rs530022	1				

Ejemplo 6**Determinación simultánea de Aneuploidía y Fracción fetal: Enriquecimiento de los ácidos nucleicos fetales y maternos en una muestra de biblioteca de secuenciación de ADN libre de células**

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

Se preparó una biblioteca de secuenciación usando ADN libre de células purificado como se describe en el Ejemplo 1.

Se preparó una biblioteca de secuenciación como sigue: El ADN libre de células contenido en 5 µl de ADN libre de células purificado se amplificó en un volumen de reacción de 50 µl que contenía 7,5µl de una mezcla de cebadores de 1 uM (Tabla 5), 10µl de Mezcla maestra NEB 5X y 27 µl de agua. El ciclado térmico se realizó con el Gene Amp9700 (Applied Biosystems). Usando las siguientes condiciones de ciclado: incubar a 95° C durante 1 minuto, seguido por 30 ciclos a 95° C durante 20 segundos, 68° C durante 1 minuto y 68° C durante 30 segundos, que fue seguido por una incubación final a 68° C durante 5 minutos. Se añadió un mantenimiento final a 4° C hasta que las muestras fueron retiradas para combinarlas con la porción no amplificada de la muestra de ADN libre de células purificada. El producto amplificado se purificó usando el sistema de purificación por PCR Agencourt AMPure XP (Parte No. A63881; Beckman Coulter Genomics, Danvers, MA). Se añadió un mantenimiento final a 4° C hasta que se retiraron las muestras para preparar la biblioteca objetivo. El producto se analizó con un Bioanalizador 2100 (Agilent Technologies, Sunnyvale, CA), y se determinó la concentración del producto amplificado. Se usó un quinto del producto amplificado purificado para preparar una biblioteca de secuenciación objetivo de ácidos nucleicos polimórficos amplificados como se describe en el Ejemplo 2. Las bibliotecas de secuenciación primaria y la objetivo se diluyeron cada una a 10nM, y la biblioteca objetivo se combinó a una proporción de 1:9 con la biblioteca de secuenciación para proporcionar una biblioteca de secuenciación enriquecida. La secuenciación de la biblioteca enriquecida y el análisis de los datos de secuenciación se realizaron como se describe en el Ejemplo 3.

a. Determinación de la fracción fetal

La determinación de la fracción fetal se realizó como se describe en el Ejemplo 5, y la fracción fetal se calculó como se ha descrito anteriormente, es decir:

$$\% \text{ del alelo}_x \text{ de la fracción fetal} = ((\sum \text{etiquetas de secuencia fetal para el alelo}_x) / (\sum \text{etiquetas de secuencia materna para el alelo}_x)) \times 100$$
TABLA 7

Determinación Simultánea de Aneuploidía y Fracción Fetal: Determinación de la Fracción Fetal			
ID de la Muestra (cariotipo)	SNP	RECuentos de ETIQUETAS DE SNP	FRACCION FETAL (%)
11409 (47, XY+21)	rs13182883.1 Chr.5 longitud=111 alelo=A	261	4.41
	rs13182883.2 Chr.5 longitud=111 alelo=G	5918	
	rs740598.1 Chr.10 longitud=114 alelo=A	5545	7.30
	rs740598.2 Chr.10 longitud=114 alelo=G	405	
	rs8078417.1 Chr.17 longitud=110 alelo=C	8189	6.74
	rs8078417.2 Chr.17 longitud=110 alelo=T	121470	
	rs576261.1 Chr.19 longitud=114 alelo=A	58342	7.62
	rs576261.2 Chr.19 longitud=114 alelo=C	4443	
Fracción Fetal (Media±D.S.) = 6.53±1.45			
ID de la Muestra			
95133 (47, XX+18)	rs1109037.1 Chr.2 longitud=126 alelo=A	12229	2.15
	rs1109037.2 Chr.2 longitud=126 alelo=G	263	

(continuada)

ID de la Muestra			
	rs13218440.1 Chr.6 longitud=139 alelo=A	55949	3.09
	rs13218440.2 Chr.6 longitud=139 alelo=G	1729	
	rs7041158.1 Chr.9 longitud=117 alelo=C	7281	4.12
	rs7041158.2 Chr.9 longitud=117 alelo=T	300	
	rs7205345.1 Chr.16 longitud=116 alelo=C	53999	2.14
	rs7205345.2 Chr.16 longitud=116 alelo=G	1154	
Fracción Fetal (Media±D.S.) = 2.9±0.9			
ID de la Muestra			
51236 (46,XY+13)	rs13218440.1 Chr.6 longitud=139 alelo=A	1119	1.65
	rs13218440.2 Chr.6 longitud=139 alelo=G	67756	
	rs560681.1 Chr.1 longitud=111 alelo=A	14123	5.18
	rs560681.2 Chr.1 longitud=111 alelo=G	732	
	rs7205345.1 Chr.16 longitud=116 alelo=C	18176	1.63
	rs7205345.2 Chr.16 longitud=116 alelo=G	296	
	rs9866013.1 Chr.3 longitud=121 alelo=C	117	2.33
	rs9866013.2 Chr.3 longitud=121 alelo=T	5024	
Fracción Fetal (Media±D.S.) = 2.7±1.7			
ID de la Muestra			
54430 (45,XO)	rs1109037.1 Chr.2 longitud=126 alelo=A	19841	1.80
	rs1109037.2 Chr.2 longitud=126 alelo=G	357	
	rs9866013.1 Chr.3 longitud=121 alelo=C	12931	3.81
	rs9866013.2 Chr.3 longitud=121 alelo=T	493	
	rs7041158.1 Chr.9 longitud=117 alelo=C	2800	4.25
	rs7041158.2 Chr.9 longitud=117 alelo=T	119	
	rs740598.1 Chr.10 longitud=114 alelo=A	12903	4.87
	rs740598.2 Chr.10 longitud=114 alelo=G	628	
	rs10773760.1 Chr.12 longitud=128 alelo=A	46324	4.65
	rs10773760.2 Chr.12 longitud=128 alelo=G	2154	
Fracción Fetal (Media±D.S.) = 3.9±1.2			

b. Determinación de aneuploidía

La determinación de aneuploidía de los cromosomas 21, 13, 18 y X se realizó usando dosis de cromosomas como se describe en el Ejemplo 4. La dosis de cromosoma 21 se determinó usando el cromosoma 14 como el cromosoma normalizador; la dosis de cromosoma 13 se determinó usando el grupo de cromosomas 3, 4, 5 y 6 como el cromosoma normalizador; la dosis de cromosoma 18 se determinó usando el cromosoma 8 como el cromosoma normalizador; y la dosis de cromosoma X se determinó usando el cromosoma 4 como el cromosoma normalizador. Los límites se calcularon para ser 2 desviaciones estándar por encima y por debajo de la media determinada en las muestras calificadas.

La Tabla 7 muestra los datos para la determinación de la fracción fetal en muestras ejemplares. Los valores de dosis de cromosoma calculados para los cromosomas 21, 18, 13, X e Y en las muestras de ensayo ejemplares

correspondientes se dan en las Tablas 8, 9, 10, 11 y 12, respectivamente.

Trisomía 18

La Tabla 9 proporciona la dosis calculada para el cromosoma 18 en una muestra de ensayo (95133). El cromosoma 8 se usó como el cromosoma normalizador. En esta situación, el cromosoma 8 tenía la variabilidad más baja y la diferenciabilidad más alta. El límite calculado para el diagnóstico positivo de la aneuploidía T18 se estableció en más de 2 desviaciones estándar de la media de las muestras calificadas (no T18). Se dio un diagnóstico para la T18 en base a que la dosis de cromosoma en la muestra de ensayo era mayor que el límite establecido. Se identificaron ocho muestras de T18 usando dosis de cromosomas, y se confirmaron que eran T18 por cariotipado.

TABLA 9

Dosis de Cromosoma para una aneuploidía T18			
Cromosoma	Densidad de Etiqueta de Secuencia	Dosis de Cromosoma para Chr 18	Límite
Chr18	604,291	0.550731	0.533297
Chr8	1,097,253		

Trisomía 13

Las Tablas 10 y 11 proporcionan la dosis calculada para el cromosoma 13 en una muestra de ensayo (51236). El límite calculado para el diagnóstico positivo de la aneuploidía T13 se estableció en 2 desviaciones estándar de la media de las muestras calificadas (no T13). La dosis de cromosoma para el cromosoma 13 proporcionada en la Tabla 10 se calculó usando la densidad de etiqueta de secuencia para el cromosoma 4 como el cromosoma normalizador, mientras que la dosis dada en la Tabla 11 se determinó usando la media de las proporciones de densidades de etiquetas de secuencia para el grupo de cromosomas 3, 4, 5 y 6 como el cromosoma normalizador. Se dio un diagnóstico para T13 en base a que la dosis de cromosoma en la muestra de ensayo era mayor que el límite establecido. Una muestra de T13 se identificó usando dosis de cromosomas, y se confirmaron que eran T13 por cariotipado.

Los datos muestran que la combinación de cromosomas 3, 4, 5 y 6 proporcionan una variabilidad (1,06) que es similar a la del cromosoma 4 (1,01), demostrando que un grupo de cromosomas se puede usar como el cromosoma normalizador para determinar las dosis de cromosomas e identificar aneuploidías.

TABLA 10

Dosis de Cromosoma para una Aneuploidía T13			
Cromosoma	Densidad de Etiqueta de Secuencia	Dosis de Cromosoma para Chr 13	Límite
Chr13	669,872	0.538140	0.536044
Chr4	1,244,791		

TABLA 11

Dosis de Cromosoma para una Aneuploidía T13			
Cromosoma	Densidad de Etiqueta de Secuencia	Dosis de Cromosoma para Chr 13	Límite
Chr13	669,872	0.532674	0.515706
Chr3	1,385,881		
Chr4	1,244,791		
Chr5	1,229,257		
Chr6	1,170,331		

Síndrome de Turner (monosomía X)

Se identificaron tres muestras que tenían una dosis de cromosoma menor que la del límite establecido que tenían menos de un cromosoma X. Se determinó que las mismas muestras tenían una dosis de cromosoma Y que era menor que el límite establecido, indicando que las muestras no tenían un cromosoma Y.

Las dosis calculadas para los cromosomas X e Y en la muestra de ensayo de monosomía Y ejemplar (54430) se dan en la Tabla 12. El Cromosoma 4 se seleccionó como el cromosoma normalizador para calcular la dosis para el cromosoma X; y todos los cromosomas, es decir 1-22 e Y, se usaron como los cromosomas normalizadores. El límite calculado para el diagnóstico positivo del Síndrome de Turner (monosomía X) se estableció para el cromosoma X en <-2 desviaciones estándar de la media y para la ausencia del cromosoma Y en <-2 desviaciones estándar de la media para las muestras calificadas (no monosomía Y).

TABLA 12

Dosis de Cromosoma para un Síndrome de Turner (monosomía X)			
Cromosoma	Densidad de Etiqueta de Secuencia	Dosis de Cromosoma para Chr X	Límite
ChrX	904,049	0.777990	0.797603
Chr4	1,162,031		
ChrY	390	0.0004462	0.002737754
Chr(1-22, X) (Media)	874,108.1		

Por lo tanto, el método permite la determinación simultánea de aneuploidías cromosómicas y la fracción fetal por secuenciación masivamente paralela de una muestra materna que comprende una mezcla de ADN libre de células fetal y materno que ha sido enriquecido para una pluralidad de secuencias polimórficas comprendiendo cada una un SNP. En este ejemplo, la mezcla de ácidos nucleicos fetales y maternos se enriqueció combinando una porción de una biblioteca de secuenciación que se construyó de secuencias polimórficas fetales y maternas con una biblioteca de secuenciación que se construyó de la mezcla de ADN libre de células fetal y materno original no amplificado.

Ejemplo 7

Determinación Simultánea de Aneuploidía y Fracción fetal: Enriquecimiento de ácidos nucleicos fetales y maternos en una muestra de ADN libre de células purificada

Para enriquecer el ADN libre de células fetal y materno contenido en una muestra purificada de ADN libre de células extraído de una muestra de plasma materno, se usó una porción del ADN libre de células purificado para amplificar secuencias de ácidos nucleicos objetivo polimórficos que comprenden cada una un SNP elegido del panel de SNPs dado en la tabla 6.

El plasma libre de células se obtuvo de una muestra de sangre materna, y el ADN libre de células se purificó de la muestra de plasma como se describe en el Ejemplo 1. La concentración final se determinó que era de 92,8 pg/ μ l.

El ADN libre de células contenido en 5 μ l de ADN libre de células purificado se amplificó en un volumen de reacción de 50 μ l que contenía 7,5 μ l de una mezcla de cebadores de 1 uM (Tabla 5), 10 μ l de Mezcla maestra NEB 5X y 27 μ l de agua. El ciclado térmico se realizó con el Gene Amp9700 (Applied Biosystems). Usando las siguientes condiciones de ciclado: incubar a 95° C durante 1 minuto, seguido por 30 ciclos a 95° C durante 20 segundos, 68° C durante 1 minuto y 68° C durante 30 segundos, que fue seguido por una incubación final a 68° C durante 5 minutos. Se añadió un mantenimiento final a 4° C hasta que las muestras fueron retiradas para combinarlas con la porción no amplificada de la muestra de ADN libre de células purificada. El producto amplificado se purificó usando el sistema de purificación por PCR Agencourt AMPure XP (Parte No. A63881; Beckman Coulter Genomics, Danvers, MA), y se cuantificó la concentración usando el Nanodrop 2000 (Thermo Scientific, Wilmington, DE). El producto de la amplificación purificado se diluyó al 1:10 en agua y se añadieron 0,9 μ l (371 pg) a 40 μ l de muestra de ADN libre de células purificado para obtener una espiga del 10%. El ADN libre de células fetal y materno enriquecido presente en la muestra de ADN libre de células purificado se usó para preparar una biblioteca de secuenciación, y se secuenció como se describe en el Ejemplo 2.

La Tabla 13 proporciona los recuentos de etiquetas obtenidos para cada uno de los cromosomas 21, 18, 13, X e Y, es decir la densidad de etiqueta de secuencia, y los recuentos de etiquetas obtenidos para las secuencias polimórficas informativas contenidas en el genoma de referencia de SNP, es decir densidad de etiqueta de SNP. Los datos muestran que la información de la secuenciación se puede obtener secuenciando una única biblioteca construida de una muestra de ADN libre de células materna purificada que ha sido enriquecida para secuencias que

comprenden SNPs para determinar simultáneamente la presencia o ausencia de aneuploidía y la fracción fetal. En el ejemplo dado los datos muestran que la fracción fetal de ADN en la muestra de plasma AFR105 era cuantificable que los resultados de secuenciación de cinco SNPs informativos y se determinó que era del 3,84%. Las densidades de etiquetas de secuencia se proporcionan para los cromosomas 21, 13, 18, X e Y. La muestra AFR105 fue la única muestra que se sometió al protocolo de enriquecer ADN libre de células purificado para secuencias polimórficas amplificadas. Por lo tanto, no se proporcionaron los coeficientes de variación y ensayos para la diferenciabilidad. Sin embargo, el ejemplo muestra que el protocolo de enriquecimiento proporciona los recuentos de etiquetas requeridos para determinar la aneuploidía y la fracción fetal a partir de un único proceso de secuenciación.

5

10

15

20

25

30

35

40

45

50

55

60

65

TABLA13

Determinación Simultánea de Aneuploidía y Fracción Fetal: Enriquecimiento de ácidos nucleicos fetales y maternos en una muestra de ADN libre de células purificada						
Aneuploidia						
	Cromosoma 21	Cromosoma 18	Cromosoma 13	Cromosoma X	Cromosoma Y	
Densidad de Etiqueta de Secuencia	178763	359529	388204	572330	2219	
Cariotipo	No afectada	No afectada	No afectada	No afectada	No afectada	
Fracción Fetal						
SNP	DENSIDAD DE ETIQUETA DE SNP			FRACCION FETAL (%)		
rs10773760.1 Chr.12 longitud=128 alelo=A	18903			2.81		
rs10773760.2 Chr.12 longitud=128 alelo=G	532					
rs1109037.1 Chr.2 longitud=126 alelo=A	347			5.43		
rs1109037.2 Chr.2 longitud=126 alelo=G	6394					
rs2567608.1 Chr.20 longitud=10 alelo=A	94503			1.74		
rs2567608.2 Chr.20 longitud=110 alelo=G	1649					
rs7041158.1 Chr.9 longitud=17 alelo=C	107			5.61		
rs7041158.2 Chr.9 longitud=117 alelo=T	6					
rs8078417.1 Chr.17 longitud=110 alelo=C	162668			3.61		
rs8078417.2 Chr.17 longitud=110 alelo=T	5877					
Fracción Fetal (Media±D.S.) = 3.8±1.7						

Ejemplo 8**Determinación Simultánea de Aneuploidía y Fracción Fetal: Enriquecimiento de ácidos nucleicos fetales y maternos en una muestra de sangre**

Para enriquecer el ADN libre de células feral y materno contenido en una muestra de plasma original derivada de una mujer embarazada, se usó una porción de la muestra de plasma original para amplificar secuencias de ácidos nucleicos objetivo polimórficas que comprendía cada una un SNP del panel de SNPs proporcionado en la Tabla 14, y una porción del producto amplificado se combinó con el resto de la muestra de plasma original.

El ADN libre de células contenido en 15µl de plasma libre de células se amplificó en un volumen de reacción de 50µl que contenía 9µl de una mezcla de cebadores de 1µm (15 plex Tabla 5), 1µl de ADN polimerasa de sangre Phusion, 25µl del tampón de PCR de sangre Phusion que contenía desoxinucleótidos trifosfatos (dNTPs: dATP, dCTP, dGTP y dTTP). El ciclado térmico se realizó con el Gene Amp9700 (Applied Biosystems) usando las siguientes condiciones de ciclado: incubar a 95° C durante 3 minutos, seguido por 35 ciclos a 95° C durante 20 segundos, 55° C durante 30 segundos, y 70° C durante 1 minuto, a lo que siguió una incubación final a 68° C durante 5 minutos. Se añadió un mantenimiento final a 4° C hasta que las muestras se retiraron para la combinación con la porción no amplificada del plasma libre de células. El producto amplificado se diluyó a 1:2 con agua y se analizó usando el Bioanalizador. Se diluyeron 3 µl adicionales del producto amplificado con 11,85µl de agua para obtener una concentración final de 2 ng/µl. Se combinaron 2,2µl del producto amplificado diluido con el resto de la muestra de plasma. El ADN libre de células fetal y materno enriquecido presente en la muestra de plasma se purificó como se describe en el Ejemplo 1, y se usó para preparar una biblioteca de secuenciación. La secuenciación y el análisis de los datos de secuenciación se realizaron como se describe en los Ejemplos 2 y 3.

Los resultados se dan en la Tabla 14. En el ejemplo dado, los datos muestran que la fracción de ADN fetal en la muestra de plasma SAC2517 era cuantificable de los resultados de secuenciación de un SNP informativo y se determinó que era del 9,5%. En el ejemplo dado, se mostró por cariotipado que la muestra SAC2517 no se vio afectada por aneuploidías de los cromosomas 21, 13, 18, X e Y. Las densidades de etiquetas de secuencia se proporcionan para los cromosomas 21, 13, 18, X e Y. La muestra SAC2517 fue la única muestra que se sometió al protocolo de enriquecimiento de ADN libre de células de plasma para secuencias polimórficas amplificadas. Por lo tanto, los coeficientes de variación y ensayos de diferenciabilidad no se pudieron determinar. El ejemplo demuestra que enriquecer la muestra de ADN libre de células fetal y materno presente en una muestra de plasma para secuencias de ácidos nucleicos que comprenden al menos un SNP informativo se puede usar para proporcionar la secuencia requisito y los recuentos de etiquetas de SNP para determinar la aneuploidía y la fracción fetal a partir de un único proceso de secuenciación.

TABLA14

Determinación Simultánea de Aneuploidía y Fracción Fetal: Enriquecimiento de ácidos nucleicos fetales y maternos en una muestra de plasma						
Aneuploidía						
	Cromosoma 21	Cromosoma 18	Cromosoma 13	Cromosoma X	Cromosoma Y	
Densidad de Etiqueta de Secuencia	183851	400582	470526	714055	2449	
Cariotipo	No afectada	No afectada	No afectada	No afectada	No afectada	
Fracción Fetal						
SNP	RECUENTOS DE ETIQUETAS		FRACCION FETAL (%)			
rs10773760.1 Chr. 12 longitud=128 alelo=A	8536		9.49			
rs10773760.2 Chr. 128 longitud=128 alelo=G	89924					

Ejemplo 9**Determinación Simultánea de Aneuploidía y Fracción Fetal en muestras maternas enriquecidas para secuencias polimórficas que comprenden STRs**

Para determinar simultáneamente la presencia o ausencia de una aneuploidía y la fracción fetal en una mezcla de ADN libre de células fetal y materno obtenido de una muestra materna, la mezcla se enriquece para secuencias polimórficas que comprenden STRs, se secuencian y se analizan los datos. El enriquecimiento puede ser de una biblioteca de secuenciación como se describe en el Ejemplo 6, de una muestra de ADN libre de células purificado como se describe en el ejemplo 7 o de una muestra de plasma como se describe en el ejemplo 8. En cada caso, la información de secuenciación se obtiene de secuenciar una biblioteca individual, lo que permite determinar simultáneamente la presencia o ausencia de una aneuploidía y la fracción fetal. Las STRs que se amplifican se eligen de las STRs codis y no codis divulgadas en la Tabla 9, y la amplificación de las secuencias de STRs polimórficas se obtiene usando los conjuntos correspondientes de cebadores proporcionados. Las STRs de la Tabla 9 han sido divulgadas anteriormente, y se han usado las STRs CSF1PO, FGA, D7S820, D13S317, D16S539, D18S51, D21S11, D2S1338 (ver tabla 5) para determinar la fracción fetal en muestras de ADN libre de células de plasma obtenidas de mujeres embarazadas con fetos masculinos o femeninos (ver Solicitudes provisionales US 61/296,358 y 61/360,837). La cuantificación de las STRs se realizó usando electroforesis capilar (ver Ejemplo 11). El Ejemplo 11 muestra que las STRs se pueden usar para determinar la fracción fetal.

TABLA 15

miniSTRS CODIS y NON-CODIS				
Locus de STR (Nombre del Marcador)	Localización del Cromosoma	Intervalo de Tamaño (bp)	Acceso del GenBank	Secuencias de Cebadores (Directo/Inverso)
Loci de miniSTR Codis*				
CSF1PO	5q33.1	89-129	X14720	ACAGTAACTGCCTTCATAGATAG (CSF1PO_F; SEQ ID NO:113) GTGTCAGACCCTGTTCTAAGTA (CSF1PO_R; SEQ ID NO:114)
FGA	4q31.3	125-281	M64982	AAATAAAATTAGGCATATTTACAAGC (FGA_F; SEQ ID NO:115) GCTGAGTGATTGTCTGTAATTG(FGA_R; SEQ ID NO:116)
TH01	11p15.5	51-98	D00269	CCTGTTCCCTCCCTATTTCCC(TH01_F; SEQ ID NO:117) GGGAACACAGACTCCATGGTG(TH01_R; SEQ ID NO:118)
TPOX	2p25.3	65-101	M68651	CTTAGGGAACCCTCACTGAATG(TPOX_F; SEQ ID NO:119) GTCCTTGTGACGTTTATTTGC(TPOX_R; SEQ ID NO:120)
vWA	12p13.31	88-148	M25858	AATAATCAGTATGTGACTTGATTGA(v WA_F; SEQ ID NO:121) ATAGGATGGATGGATAGATGGA(vWA_R ; SEQ ID NO:122)
D3S1358	3p21.31	72-120	NT_005997	CAGAGCAAGACCCTGTCTCAT(D3S1358_ F; SEQ ID NO:123) TCAACAGAGGCTTGCATGTAT(D3S1358_R; SEQ ID NO:124)
D5S818	5q23.2	81-117	AC008512	GGGTGATTTTCTCTTTGGT(D5S818_F; SEQ ID NO:125) AACATTTGTATCTTTATCTGTATCCTTAT TTAT(D5S818_R; SEQ ID NO:126)

(continuada)

miniSTRS CODIS y NON-CODIS				
Locus de STR (Nombre del Marcador)	Localización del Cromosoma	Intervalo de Tamaño (bp)	Acceso del GenBank	Secuencias de Cebadores (Directo/Inverso)
Loci de miniSTR Codis*				
D7S820	7q21.11	136-176	AC004848	GAACACTTGTTCATAGTTTAGAACGAAC(D7S 820_F; SEQ ID NO:127) TCATTGACAGAATTGCACCA(D7S820_R; SEQ ID NO:128)
D8S1179	8q24.13	86-134	AF216671	TTTGTATTTTCATGTGTACATTCGTATC(D 7S820_F; SEQ ID NO:129) ACCTATCCTGTAGATTATTTTCACTGTG(D7S 820_R; SEQ ID NO:130)
D13S317	13q31.1	88-132	AL353628	TCTGACCCATCTAACGCCTA(D13S317_F; SEQ ID NO:131) CAGACAGAAAGATAGATAGATGATTGA(D13 S317_R; SEQ ID NO:132)
D16S539	16q24.1	81-121	AC024591	ATACAGACAGACAGACAGGTG(D16S539 _F; SEQ ID NO:133) GCATGTATCTATCATCCATCTCT(D16S53 9_R; SEQ ID NO:134)
D18S51	18q21.33	113-193	AP001534	TGAGTGACAAATTGAGACCTT(D18S51_F; SEQ ID NO:135) GTCTTACAATAACAGTTGCTACTATT(D1 8S51_R; SEQ ID NO:136)
D21S11	21q21.1	153-221	AP000433	ATCCCCCAAGTGAATTGC(D21S11_F; SEQ ID NO:137) GGTAGATAGACTGGATAGATAGACGA(D 21S11_R; SEQ ID NO:138)
D2S1338	2q35	90-142	AC01036	TGGAAACAGAAATGGCTTGG(D2S1338_F ; SEQ ID NO:139) GATTGCAGGAGGGAAGGAAG(D2S 13 38_R; SEQ ID NO:140)
Penta D	21q22.3	94-167	AP001752	GAGCAAGACACCATCTCAAGAA(Penta D_F; SEQ ID NO:141) GAAATTTTACATTTATGTTTATGATTCTC T(Penta D_R; SEQ ID NO:142)
Penta E	15q26.2	80-175	AC027004	GGCGACTGAGCAAGACTC(PentaE_F; SEQ ID NO:143) GGTTATTAATTGAGAAAACCTTACA(Penta E _R; SEQ ID NO:144)
Loci de miniSTR No Codis*				
D22S1045	22q12.3	82 - 115	AL022314 (17)	ATTTTCCCGATGATAGTAGTCT(D22S10 45_F; SEQ ID NO:145) GCGAATGTATGATTGGCAATATTTTT(D2 2S1045_R; SEQ ID NO:146)
D20S1082	20q13.2	73 - 101	AL158015	ACATGTATCCCAGAACTTAAAGTAAAC(D20S 1082_F; SEQ ID NO:147)

(continuada)

Loci de miniSTR No Codis*				
				GCAGAAGGGAAAATTGAAGCTG(D20S10 82_R; SEQ ID NO:148)
D20S482	20p13	85 - 126	AL121781 (14)	CAGAGACACCGAACCAATAAGA(D20S48 2_F; SEQ ID NO:149) GCCACATGAATCAATTCCTATAATAAA(D20S 482_R; SEQ ID NO:150)
D18S853	18p11.31	82 - 104	AP005130 (11)	GCACATGTACCCTAAAACCTAAAAT(D18 S853_F; SEQ ID NO:151) GTCAACCAAACTCAACAAGTAGTAA(D 18S853_R; SEQ ID NO:152)
D17S1301	17q25.1	114 - 139	AC016888 (12)	AAGATGAAATTGCCATGTAAAAATA(D1 7S1301_F; SEQ ID NO:153) GTGTGTATAACAAAATTCCTATGATGG(D17S 1301_R; SEQ ID NO:154)
D17S974	17p13.1	114 - 139	AC034303 (10)	GCACCCAAAACCTGAATGTCATA(D17S97 4_F; SEQ ID NO:155) GGTGAGAGTGAGACCCTGTC(D17S974_R ; SEQ ID NO:156)
D14S1434	14q32.13	70 - 98	AL121612 (13)	TGTAATAACTCTACGACTGTCTGTCTG(D 14S1434_F; SEQ ID NO:157) GAATAGGAGGTGGATGGATGG(D 14S 143 4_R; SEQ ID NO:158)
D12ATA63	12q23.3	76 - 106	AC009771 (13)	GAGCGAGACCCTGTCTCAAG(D 12ATA63_F; SEQ ID NO:159) GGAAAAGACATAGGATAGCAATTT(D12 ATA63_R; SEQ ID NO:160)
D11S4463	11q25	88 - 116	AP002806 (14)	TCTGGATTGATCTGTCTGTCC(D11S4463_ F; SEQ ID NO:161) GAATTAAATACCATCTGAGCACTGAA(D 11S4463_R; SEQ ID NO:162)
D10S1435	10p15.3	82 - 139	AL354747 (11)	TGTTATAATGCATTGAGTTTTATTCTG(D 10S1435_F; SEQ ID NO:163) GCCTGTCTCAAAAATAAGAGATAGAC A(D10S1435_R; SEQ ID NO:164)
D10S1248	10q26.3	79 - 123	AL391869 (13)	TTAATGAATTGAACAAATGAGTGAG(D1 0S1248_F; SEQ ID NO:165) GCAACTCTGGTTGTATTGTCTTCAT(D10S 1248_R; SEQ ID NO:166)
D9S2157	9q34.2	71 - 107	AL162417 (10)	CAAAGCGAGACTCTGTCTCAA(D9S2157_ F; SEQ ID NO:167) GAAAATGCTATCCTCTTTGGTATAAAT(D9S2 157_R; SEQ ID NO:168)
D9S1122	9q21.2	93 - 125	AL161789 (12)	GGGTATTTCAAGATAACTGTAGATAGG(D9S 1122_F; SEQ ID NO:168) GCTTCTGAAAGCTTCTAGTTTACC(D9S11 22_R; SEQ ID NO:170)
D8S1115	8p11.21	63 - 96	AC090739 (9)	TCCACATCCTCACCAACAC(D8S1115_F; SEQ ID NO:171)

(continuada)

Loci de miniSTR No Codis*				
				GCCTAGGAAGGCTACTGTCAA(D8S 1115_R; SEQ ID NO:172)
D6S1017	6p21.1	81 - 110	AL035588 (10)	CCACCCGTCCATTTAGGC(D6S1017_F; SEQ ID NO:173) GTGAAAAAGTAGATATAATGGTTGGTG(D6S 1017_R; SEQ ID NO:174)
D6S474	6q21	107 - 136	AL357514 (17)	GGTTTTCCAAGAGATAGACCAATTA(D6S 474_F; SEQ ID NO:175) GTCCTCTCATAAATCCCTACTCATATC(D 6S474_R; SEQ ID NO:176)
D5S2500	5q11.2	85 - 126	AC008791 (17)	CTGTTGGTACATAATAGGTAGGTAGGT(D5S 2500_F; SEQ ID NO:177) GTCGTGGGCCCCATAAATC(D5S2500_R; SEQ ID NO:178)
D4S2408	4p15.1	85 - 109	AC110763 (9)	AAGGTACATAACAGTTCAATAGAAAGC(D4S 2408_F; SEQ ID NO:179) GTGAAATGACTGAAAAATAGTAACCA(D 4S2408_R; SEQ ID NO:180)
D4S2364	4q22.3	67 - 83	AC022317 (9)	CTAGGAGATCATGTGGGTATGATT(D4S2 364U_F; SEQ ID NO:181) GCAGTGAATAAATGAACGAATGGA(D4S 2364_R; SEQ ID NO:182)
D3S4529	3p12.1	111 - 139	AC117452 (13)	CCCCAAATTACTTGAGCCAAT(D3S452_F ; SEQ ID NO:183) GAGACAAAATGAAGAAACAGACAG(D3S 452_R; SEQ ID NO:184)
D3S3053	3q26.31	84 - 108	AC069259 (9)	TCTTTGCTCTCATGAATAGATCAGT(D3S 3053_F; SEQ ID NO:185) GTTTGTGATAATGAACCCACTCAG(D3S3 053_R; SEQ ID NO:186)
D2S1776	2q24.3	127 - 161	AC009475 (11)	TGAACACAGATGTTAAGTGTGTATATG(D2S1 776_F; SEQ ID NO:187) GTCTGAGGTGGACAGTTATGAAA(D2S17 76_R; SEQ ID NO:188)
D2S441	2p14	78 - 110	AC079112 (12)	CTGTGGCTCATCTATGAAACTT(D2S441 _F; SEQ ID NO:189) GAAGTGGCTGTGGTGTATGAT(D2S441_ R; SEQ ID NO:190)
D1S1677	1q23.3	81 - 117	AL513307 (15)	TTCTGTTGGTATAGAGCAGTGTTT(D 1 S 16 77_F; SEQ ID NO:191) GTGACAGGAAGGACGGAATG(D1S1677_R; SEQ ID NO:192)
D1S1627	1p21.1	81 - 100	AC093119 (13)	CATGAGGTTTGCAATACTATCTTAAC(D1S1 627_F; SEQ ID NO:193) GTTTAAATTTCTCCAAATCTCCA(D 1 S 16 27_R; SEQ ID NO:194)
D1GATA113	1p36.23	81 - 105	Z97987 (11)	TCTTAGCCTAGATAGATACTTGCTTCC(D 1GATA113_F; SEQ ID NO:195)

(continuada)

Non-Codis miniSTR loci*				
				GTCAACCTTTGAGGCTATAGGAA(D1GA TA113_R; SEQ ID NO:196)
*(Butler et al., J Forensic Sci 5:1054-1064; Hill et al., Poster #44- 17th International Symposium on Human Identification - 2006)				

La secuenciación de la biblioteca enriquecida para secuencias de STR polimórficas se realiza usando una tecnología NGS, por ejemplo secuenciación por síntesis. Las lecturas de secuencia de longitudes de al menos 100 bp se alinean con un genoma de referencia, por ejemplo la secuencia del genoma de referencia humano NCBI36/hg18, y con un genoma STR, y el número de etiquetas de secuencia mapeadas y etiquetas de STR obtenidas se usa para determinar la presencia o ausencia de aneuploidía y la fracción fetal, respectivamente. El genoma de referencia STR incluye las secuencias de amplicones amplificados de los cebadores dados.

Ejemplo 10

Determinación Simultánea de Aneuploidía y Fracción Fetal en Muestras Maternas Enriquecidas para Secuencias Polimórficas que Comprenden SNPs en tándem

Para determinar simultáneamente aneuploidía y fracción fetal en muestras maternas que comprenden ácidos nucleicos fetales y maternos, se enriquecen muestras de ADN libre de células purificado y muestras de bibliotecas de secuenciación para secuencias de ácidos nucleicos objetivo polimórficos que comprenden cada una un par de SNPs en tándem seleccionados de rs7277033-rs2110153; rs2822654-rs1882882; rs368657-rs376635; rs2822731-rs2822732; rs1475881-rs7275487; rs1735976-rs2827016; rs447340-rs2824097; rs418989-rs13047336; rs987980-rs987981; rs4143392-rs4143391; rs1691324-rs13050434; rs11909758-rs9980111; rs2826842-rs232414; rs1980969-rs1980970; rs9978999-rs9979175; rs1034346-rs12481852; rs7509629-rs2828358; rs4817013-rs7277036; rs9981121-rs2829696; rs455921-rs2898102; rs2898102-rs458848; rs961301-rs2830208; rs2174536-rs458076; rs11088023-rs11088024; rs1011734-rs1011733; rs2831244-rs9789838; rs8132769-rs2831440; rs8134080-rs2831524; rs4817219-rs4817220; rs2250911-rs2250997; rs2831899-rs2831900; rs2831902-rs2831903; rs11088086-rs2251447; rs2832040-rs11088088; rs2832141-rs2246777; rs2832959-rs9980934; rs2833734-rs2833735; rs933121-rs933122; rs2834140-rs12626953; rs2834485-rs3453; rs9974986-rs2834703; rs2776266-rs2835001; rs1984014-rs1984015; rs7281674-rs2835316; rs13047304-rs13047322; rs2835545-rs4816551; rs2835735-rs2835736; rs13047608-rs2835826; rs2836550-rs2212596; rs2836660-rs2836661; rs465612-rs8131220; rs9980072-rs8130031; rs418359-rs2836926; rs7278447-rs7278858; rs385787-rs367001; rs367001-rs386095; rs2837296-rs2837297; y rs2837381-rs4816672. Los cebadores usados para amplificar las secuencias objetivo que comprenden los SNPs en tándem están diseñados para abarcar ambos sitios de SNP. Por ejemplo, el cebador directo está diseñado para abarcar el primer SNP, y el cebador inverso está diseñado para abarcar el segundo del par de SNP en tándem, es decir cada uno de los sitios SNP en el par en tándem está abarcado dentro de los 36 bp generados por el método de secuenciación. La secuenciación de extremos emparejados se puede usar para identificar todas las secuencias que abarcan los sitios de SNP en tándem. Los conjuntos ejemplares de cebadores que se usan para amplificar los SNPs en tándem divulgados en la presente son rs7277033-rs2110153_F: TCCTGGAAACAAAAGTATT (SEQ ID NO:197) y rs7277033-rs2110153_R: AACCTTACAACAAAGCTAGAA (SEQ ID NO:198), set rs2822654-rs1882882_F: ACTAAGCCTTGGGGATCCAG (SEQ ID NO:199) y rs2822654-rs1882882_R: TGCTGTGGAAATACTAAAAGG (SEQ ID NO:200), set rs368657-rs376635_F: CTCCAGAGGTAATCCTGTGA (SEQ ID NO:201) y rs368657-rs376635_R: TGGTGTGAGATGGTATCTAGG (SEQ ID NO:202), rs2822731-rs2822732_F: GTATAATCCATGAATCTTGTTT (SEQ ID NO:203) y rs2822731-rs2822732_R: TTCAAATTGTATATAAGAGAGT (SEQ ID NO:204), rs1475881-rs7275487_F: GCAGGAAAGTTATTTTAAAT (SEQ ID NO:205) y rs1475881-rs7275487_R: TGCTTGAGAAAGCTAACACTT (SEQ ID NO:206), rs1735976-rs2827016_F: CAGTGTGGAAATTGTCTG (SEQ ID NO:207) y rs1735976-rs2827016_R: GGCAGTGGGAGATTATTGTA (SEQ ID NO:208), rs447340-rs2824097_F: TCCTGTTGTTAAGTACACAT (SEQ ID NO:209) y rs447340-rs2824097_R: GGGCCGTAATTACTTTTG (SEQ ID NO:210), rs418989-rs13047336_F: ACTCAGTAGGCACTTTGTGTC (SEQ ID NO:211) y rs418989-rs13047336_R: TCTTCCACCACACCAATC (SEQ ID NO:212), rs987980-rs987981_F: TGGCTTTTCAAAGGTAAAA (SEQ ID NO:213) y rs987980-rs987981_R: GCAACGTTAACATCTGAATTT (SEQ ID NO:214), rs4143392-rs4143391_F: rs4143392-rs4143391 (SEQ ID NO:215) y rs4143392-rs4143391_R: ATTTTATATGTCATGATCTAAG (SEQ ID NO:216), rs1691324-rs13050434_F: AGAGATTACAGGTGTGAGC (SEQ ID NO:217) y rs1691324-rs13050434_R: ATGATCCTCAACTGCCTCT (SEQ ID NO:218), rs11909758-rs9980111_F: TGAAACTCAAAAGAGAAAAAG (SEQ ID NO:219) y rs11909758-rs9980111_R: ACAGATTTCTACTTAAATTT (SEQ ID NO:220), rs2826842-rs232414_F: TGAAACTCAAAAGAGAAAAAG (SEQ ID NO:221) y rs2826842-rs232414_R: ACAGATTTCTACTTAAATTT (SEQ ID NO:222), rs2826842-rs232414_F: GCAAAGGGTACTCTATGTA (SEQ ID NO:223) y rs2826842-rs232414_R: TATCGGGTCATCTTGTTAAA (SEQ ID NO:224), rs1980969-rs1980970_F:

TCTAACAAAGCTCTGTCCAAAA (SEQ ID NO:225) y rs1980969-rs1980970-R: CCACACTGAATAACTGGAACA
 (SEQ ID NO:226), rs9978999-rs9979175_F: GCAAGCAAGCTCTCTACCTTC (SEQ ID NO:227) y rs9978999-
 rs9979175_R: TGTTCTTCCAAAATTCACATGC (SEQ ID NO:228), rs1034346-rs12481852_F:
 5 ATTTCACTATTCCTTCATTTT (SEQ ID NO:229) y rs1034346-rs12481852_R: TAATTGTTGCACACTAAATTAC
 (SEQ ID NO:230), rs4817013-rs7277036_F: AAAAAGCCACAGAAATCAGTC (SEQ ID NO:231) y rs4817013-
 rs7277036_R: TTCTTATATCTCACTGGGCATT (SEQ ID NO:232), rs9981121-rs2829696_F:
 GGATGGTAGAAGAGAAGAAAGG (SEQ ID NO:233) y rs9981121-rs2829696_R:
 GGATGGTAGAAGAGAAGAAAGG (SEQ ID NO:234), rs455921-rs2898102_F: TGCAAAGATGCAGAACCAAC (SEQ
 10 ID NO:235) y rs455921-rs2898102_R: TTTTGTTCCTTGTCTGGCTGA (SEQ ID NO:236), rs2898102- rs458848_F:
 TGCAAAGATGCAGAACCAAC (SEQ ID NO:237) y rs2898102- rs458848_R: GCCTCCAGCTCTATCCAAGTT (SEQ
 ID NO:238), rs961301-rs2830208_F: CCTTAATATCTTCCCATGTCCA (SEQ ID NO:239) y rs961301-rs2830208_R:
 ATTGTTAGTGCCTCTTCTGCTT (SEQ ID NO:240), rs2174536-rs458076_F: GAGAAGTGAGGTCAGCAGCT (SEQ
 15 ID NO:241) y rs2174536-rs458076_R: TTTCTAAATTTCCATTGAACAG (SEQ ID NO:242), rs11088023-
 rs11088024_F: GAAATTGGCAATCTGATTCT (SEQ ID NO:243) y rs11088023-rs11088024_R:
 CAACTTGTCTTTTATTGATGT (SEQ ID NO:244), rs1011734-rs1011733_F: CTATGTTGATAAAACATTGAAA (SEQ
 ID NO:245) y rs1011734-rs1011733_R: GCCTGTCTGGAATATAGTTT (SEQ ID NO:246), rs2831244-rs9789838_F:
 CAGGGCATATAATCTAAGCTGT (SEQ ID NO:247) y rs2831244-rs9789838_R: CAATGACTCTGAGTTGAGCAC
 (SEQ ID NO:248), rs8132769-rs2831440_F: ACTCTCTCCCTCCCCTCT (SEQ ID NO:249) y rs8132769-
 rs2831440_R: TATGGCCCCAAAATATTCT (SEQ ID NO:250), rs8134080-rs2831524_F:
 20 ACAAGTACTGGGCAGATTGA (SEQ ID NO:251) y rs8134080-rs2831524_R: GCCAGGTTTAGCTTTCAAGT (SEQ
 ID NO:252), rs4817219-rs4817220_F: TTTTATATCAGGAGAAACACTG (SEQ ID NO:253) y rs4817219-
 rs4817220_R: CCAGAATTTTGGAGTTTAAT (SEQ ID NO:254), rs2250911-rs2250997_F:
 TGTCATTCTCTTTTATCTCCA (SEQ ID NO:255) y rs2250911-rs2250997_R: TTCTTTTGCCTCTCCCAAAG (SEQ
 25 ID NO:256), rs2831899-rs2831900_F: ACCCTGGCACAGTGTGACT (SEQ ID NO:257) y rs2831899-rs2831900_R:
 TGGGCCCTGAGTTGAGAAGAT (SEQ ID NO:258), rs2831902-rs2831903_F: AATTGTAAAGTATGTGCAACG (SEQ
 ID NO:259) y rs2831902-rs2831903_R: TTTTCCCATTTCCAACCTCT (SEQ ID NO:260), rs11088086-rs2251447_F:
 AAAAGATGAGACAGGCAGGT (SEQ ID NO:261) y rs11088086-rs2251447_R: ACCCCTGTGAATCTCAAAAT
 (SEQ ID NO:262), rs2832040-rs11088088_F: GCATTGCTTCTATTGTTTGT (SEQ ID NO:263) y rs2832040-
 rs11088088_R: CCCTTCCTCTCTTCCATTCT (SEQ ID NO:264), rs2832141-rs2246777_F: AGCACTGCAGGTA
 30 (SEQ ID NO:265) y rs2832141-rs2246777_R: ACAGATACCAAAGAACTGCAA (SEQ ID NO:266), rs2832959-
 rs9980934_F: TGGACACCTTCAACTTAGA (SEQ ID NO:267) y rs2832959- rs9980934_R:
 GAACAGTAATGTTGAACTTTTT (SEQ ID NO:268), rs2833734-rs2833735_F: TCTTGCAAAAAGCTTAGCACA (SEQ
 ID NO:269) y rs2833734-rs2833735_R: AAAAAGATCTCAAAGGGTCCA (SEQ ID NO:270), rs933121-rs933122_F:
 GCTTTTGTGTAACATCAAGT (SEQ ID NO:271) y rs933121-rs933122_R: CCTTCCAGCAGCATAGTCT (SEQ ID
 35 NO:272), rs2834140-rs12626953_F: AAATCCAGGATGTGCAGT (SEQ ID NO:273) y rs2834140-rs12626953_R:
 ATGATGAGGTCAGTGGTGT (SEQ ID NO:274), rs2834485-rs3453_F: CATCACAGATCATAGTAAATGG (SEQ ID
 NO:275) y rs2834485-rs3453_R: AATTATTATTTTGCAGGCAAT (SEQ ID NO:276), rs9974986-rs2834703_F:
 CATGAGGCAAACACCTTTCC (SEQ ID NO:277) y rs9974986-rs2834703_R: GCTGGACTCAGGATAAAGAACA
 (SEQ ID NO:278), rs2776266-rs2835001_F: TGGAAGCCTGAGCTGACTAA (SEQ ID NO:279) y rs2776266-
 rs2835001_R: CCTTCTTTTCCCCCAGAATC (SEQ ID NO:280), rs1984014-
 40 rs1984015_F: TAGGAGAACAGAAGATCAGAG (SEQ ID NO:281) y rs1984014-
 rs1984015_R: AAAGACTATTGCTAAATGCTTG (SEQ ID NO:282), rs7281674-rs2835316_F:
 TAAGCGTAGGGCTGTGTG (SEQ ID NO:283) y rs7281674-rs2835316_R: GGACGGATAGACTCCAGAAGG
 (SEQ ID NO:284), rs13047304-rs13047322_F: GAATGACCTTGGCACTTTTATCA (SEQ ID NO:285) y rs13047304-
 45 rs13047322_R: AAGGATAGAGATATACAGATGAATGGA (SEQ ID NO:286), rs2835735-rs2835736_F:
 CATGCACCGCGCAAATAC (SEQ ID NO:287) y rs2835735-rs2835736_R: ATGCCTCACCCACAAACAC (SEQ ID
 NO:288), rs13047608-rs2835826_F: TCCAAGCCCTTCTCACTCAC (SEQ ID NO:289) y rs13047608-rs2835826_R:
 CTGGGACGGTGACATTTTCT (SEQ ID NO:290), rs2836550-rs2212596_F: CCCAGGAAGAGTGGAAGATT (SEQ
 ID NO:291) y rs2836550-rs2212596_R: TTAGCTTGATGATGCTGCTGT (SEQ ID NO:292), rs2836660-rs2836661_F:
 50 AGCTAGATGGGGTGAATTTT (SEQ ID NO:293) y rs2836660-rs2836661_R: TGGGCTGAGGGGAGATT (SEQ ID NO:294), rs465612-
 rs8131220_F: ATCAAGCTAATTAATGTTATCT (SEQ ID NO:295) y rs465612-rs8131220_R:
 AATGAATAAGGTCTCAGAG (SEQ ID NO:296), rs9980072-rs130031_F: TTTAATCTGATCATTGCCCTA (SEQ ID
 NO:297) y rs9980072-rs130031_R: AGCTGTGGGTGACCTTGA (SEQ ID NO:298), rs418359-rs2836926_F:
 TGTCCCACCATTTGTGTATTA (SEQ ID NO:299) y rs418359-rs2836926_R: TCAGACTTGAAGTCCAGGAT (SEQ ID
 55 NO:300), rs7278447-rs7278858_F: GCTTCAGGGGTGTTAGTTT (SEQ ID NO:301) y rs7278447-rs7278858_R:
 CTTTGTGAAAAGTCGTCCAG (SEQ ID NO:302), rs385787-rs367001_F: CCATCATGGAAAGCATGG (SEQ ID
 NO:303) y rs385787-rs367001_R: TCATCTCCATGACTGCACTA (SEQ ID NO:304), rs367001-rs386095_F:
 GAGATGACGGAGTAGCTCAT (SEQ ID NO:305) y rs367001-rs386095_R: CCCAGCTGCACTGTCTAC (SEQ ID
 NO:306), rs2837296-rs2837297_F: TCTTGTTCGAATCACAGGAC (SEQ ID NO:307) y rs2837296-rs2837297_R:
 60 ATGCTGTTAGCTGAAGCTCT (SEQ ID NO:308), y rs2837381-rs4816672_F: TGAAAGCTCCTAAAGCAGAG (SEQ
 ID NO:309) y rs2837381-rs4816672_R: TTGAAGAGATGTGCTATCAT (SEQ ID NO:310). Se pueden incluir
 secuencias de polinucleótidos, por ejemplo secuencias de abrazadera GC, para asegurar la hibridación específica
 de los cebadores ricos en AT (Ghanta et al., PLOS ONE 5(10): doi10.1371/journal.pone.0013184 [2010], disponible
 en la red mundial en plosone.org). Un ejemplo de una secuencia de abrazadera GC que se puede incluir o en el 5'
 65 del cebador directo o en el 3' del cebador inverso es la

GCCGCCTGCAGCCCGCGCCCCCGTGCCCCCGCCCCGCCGCCGCCGGCCGGCGCC (SEQ ID NO:311). La preparación de la muestra y el enriquecimiento de la biblioteca de secuenciación de ADN libre de células, una muestra de ADN libre de células purificado y una muestra de plasma se realizan de acuerdo con el método descrito en los Ejemplos 6, 7 y 8, respectivamente. Todas las bibliotecas de secuenciación se preparan como se describe en el Ejemplo 2a, y la secuenciación se realiza como se describe en el Ejemplo 2b incluyendo secuenciación de extremos emparejados. El análisis de los datos de secuenciación para la determinación de aneuploidía fetal se realiza como se describe en los Ejemplos 3 y 4. Concomitante al análisis para determinar la aneuploidía, los datos de secuenciación se analizan para determinar la fracción fetal de la manera siguiente. Después de la transferencia de los archivos de imagen y designación de bases al servidor Unix que ejecuta el software "Genome Analyzer Pipeline" versión 1.51 de Illumina como se ha descrito en el Ejemplo 3a, las lecturas de 36bp se alinean con un 'genoma de SNP en tándem' usando el programa BOWTIE. El genoma de SNP en tándem se identifica como el agrupamiento de secuencias de ADN que abarcan los alelos de los 58 pares de SNP en tándem divulgados anteriormente. Sólo se usan lecturas que mapean únicamente para el genoma de SNP en tándem para el análisis de la fracción fetal. Las lecturas que equivalen perfectamente con el genoma de SNP en tándem se cuentan como etiquetas y se filtran. De las lecturas restantes, sólo se cuentan como etiquetas y se incluyen en el análisis las lecturas que tienen una o dos discrepancias. Se cuentan las etiquetas mapeadas para cada uno de los alelos de SNP en tándem, y se determina la fracción fetal esencialmente como se ha descrito en el Ejemplo 6 anterior pero contando para las etiquetas mapeadas para los dos alelos x e y de SNP en tándem presentes en cada una de las secuencias de ácidos nucleicos objetivo polimórficos amplificadas que se amplifican para enriquecer las muestras, ese decir

$\% \text{ de alelo}_{x+y} \text{ de la fracción fetal} = ((\Sigma \text{Etiquetas de secuencia fetal para el alelo}_{x+y}) / (\Sigma \text{Etiquetas de secuencia materna para el alelo}_{x+y})) \times 100.$

Sólo se usan SNPs en tándem informativos para determinar la fracción fetal.

Opcionalmente, la fracción de los ácidos nucleicos fetales en la mezcla de los ácidos nucleicos fetales y maternos se calcula para cada uno de los alelos informativos (alelo_{x+y}) de la manera siguiente:

$\% \text{ de alelo}_{x+y} \text{ de la fracción fetal} = ((2 \times \Sigma \text{Etiquetas de secuencia fetal para el alelo}_{x+y}) / (\Sigma \text{Etiquetas de secuencia materna para el alelo}_{x+y})) \times 100,$

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

El porcentaje de fracción fetal se calcula para al menos 1, al menos 2, al menos 3, al menos 4, al menos 5, al menos 6, al menos 7, al menos 8, al menos 9, al menos 10, al menos 11, al menos 12, al menos 13, al menos 14, al menos 15, al menos 16, al menos 17, al menos 18, al menos 19, al menos 20, al menos 25, al menos 30, al menos 35, al menos 40 o más conjuntos informativos de alelos en tándem. En una realización, la fracción fetal es la fracción fetal media determinada para al menos 3 conjuntos informativos de alelos en tándem.

Ejemplo 11

Determinación de Fracción Fetal por Electroforesis Capilar de Secuencias Polimórficas que Comprenden STRs

Para determinar la fracción fetal en muestras maternas que comprenden ADN libre de células fetal y materno, se recogieron muestras de sangre periférica de mujeres embarazadas voluntarias que llevaban fetos masculinos o femeninos. Las muestras de sangre periférica se obtuvieron y procesaron para proporcionar ADN libre de células purificado como se describe en el Ejemplo 1.

Se analizaron diez microlitros de muestras de ADN libre de células usando el kit de amplificación por PCR AmpF1STR® MiniFiler™ (Applied Biosystems, Foster City, CA) de acuerdo con las instrucciones del fabricante. Brevemente, se amplificó el ADN libre de células contenido en 10 µl en un volumen de reacción de 25 µl que contenía 5µl de cebadores etiquetados por fluorescencia (Conjunto de cebadores AmpFISTR® MiniFiler™) y la Mezcla Maestra AmpFISTR® MiniFiler™, que incluye ADN polimerasa AmpliTaq Gold® y el tampón asociado, sal (1,5 mM MgCl₂) y 200 µM de desoxinucleótido trifosfatos (dNTPs: dATP, dCTP, dGTP y dTTP). Los cebadores etiquetados por fluorescencia son cebadores directos que están etiquetados con colorantes 6FAM™, VIC™, NED™, y PET™. Se realizó el ciclado térmico con el Gene Amp9700 (Applied Biosystems) usando las siguientes condiciones de ciclado: incubación a 95° C durante 10 minutos, seguido por 30 ciclos a 94° C durante 20 segundos, 59° C durante 2 minutos y 72° C durante 1 minuto, a lo que siguió una incubación final a 60° C durante 45 minutos. Se añadió un mantenimiento final a 4° C hasta que las muestras se retiraron para análisis. El producto amplificado se preparó diluyendo 1ul de producto amplificado en 8,7ul de Hi-DiTM formamida (Applied Biosystems) y 0,3 µl de estándar de tamaño interno GeneScanTM-500 LIZ_ (Applied Biosystems), y se analizó con un Analizador Genético ABI PRISM3130x1 (Applied Biosystems) usando el Data Collection HID_G5_POP4 (Applied Biosystems), y una

matriz capilar de 36-cm. Todo el genotipado se realizó con el software GeneMapper_ID v3.2 (Applied Biosystems) usando las escaleras alélicas e intervalos y paneles proporcionados por el fabricante.

Toda medición del genotipado se realizó en el Analizador Genético Applied Biosystems 3130xl, usando una "ventana" de $\pm 0,5$ -nt de alrededor del tamaño obtenido para cada alelo para permitir la detección y asignación correcta de los alelos. Cualquier alelo de la muestra cuyo tamaño estaba fuera de la ventana de $\pm 0,5$ -nt se determinó que era OL, es decir "Fuera de la escalera". Los alelos OL son alelos de un tamaño que no está representado en la Escalera Alélica AmpFISTR® MiniFiler™ e un alelo que no corresponde con una escalera alélica, pero cuyo tamaño está justo fuera de una ventana debido a un error de medición. Se estableció el límite de altura del pico mínimo de >50 RFU en base a los experimentos de validación realizados para evitar el tipificado cuando es probable que los efectos estocásticos interfieran con la interpretación precisa de las mezclas. El cálculo de la fracción fetal se basa en hacer la media de todos los marcadores informativos. Los marcadores informativos se identifican por la presencia de picos en el electroferograma que cae dentro de los parámetros de los intervalos preestablecidos para las STRs que se analizan.

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

1. Los datos de alelos fuera de escalera (OL) para alelos no incluidos en el cálculo; y
2. Sólo se incluyen en el cálculo alturas de pico derivadas de > 50 RFU (unidades de fluorescencia relativas)
3. Si uno de los intervalos está presente el marcador se considera no informativo; y
4. Si se designa un segundo intervalo pero los picos del primer y segundo intervalos están dentro del 50-70% de sus unidades de fluorescencia relativa (RFU) en la altura del pico, la fracción minoritaria no se mide y el marcador se considera no informativo.

La fracción del alelo menor para cualquier marcador informativo dado se calcula dividiendo la altura del pico del componente menor por la suma de la altura del pico para el componente principal, y se expresa como un porcentaje que fue primero calculado para cada locus informativo como

$$\text{fracción fetal} = (\text{altura del pico del alelo menor} / \text{altura del pico del alelo(s) principal}) \times 100$$

La fracción fetal para una muestra que comprende dos o más STRs informativas, se calcularía como la media de las fracciones fetales calculadas para dos o más marcadores informativos.

La Tabla 16 proporciona los datos obtenidos de analizar ADN libre de células de un sujeto embarazado con un feto masculino.

TABLA 16

Fracción Fetal Determinada en ADN libre de células de un Sujeto Embarazado por análisis de STRs								
STR	Alelo 1	Alelo2	Alelo3	Altura del Alelo 1	Altura del Alelo 2	Altura del Alelo 3	Fracción Fetal	Fracción Fetal (Media/STR)
AMEL	X	Y		3599	106		2.9	
AMEL	X	Y		3602	110		3.1	
AMEL	X	Y		3652	109		3.0	3.0
CSF1PO	11	12		2870	2730			
CSF1PO	11	12		2924	2762			
CSF1PO	11	12		2953	2786			
D13S317	11	12		2621	2588			
D13S317	11	12		2680	2619			
D13S317	11	12		2717	2659			
D16S539	9	11		1056	1416			
D16S539	9	11		1038	1394			
D16S539	9	11		1072	1437			
D18S51	13	15		2026	1555			
D18S51	13	15		2006	1557			
D18S51	13	15		2050	1578			
D21S11	28	31.2		2450	61		2.5	
D21S11	28	31.2		2472	62		2.5	
D21S11	28	31.2		2508	67		2.7	2.6
D2S1338	20	23		3417	3017			
D2S1338	20	23		3407	3020			
D2S1338	20	23		3493	3055			
D7S820	9	12	13	2373	178	1123	5.1	
D7S820	9	12	13	2411	181	1140	5.1	
D7S820	9	12	13	2441	182	1156	5.1	5.1
FGA	17.2	22	25	68	1140	896	3.3	
FGA	17.2	22	25	68	1144	909	3.1	
FGA	17.2	22	25	68	1151	925	3.3	3.2
Fracción Fetal = 3.5								

Los resultados muestran que los miniSTRs pueden usarse para distinguir alelos fetales y maternos en ADN libre de células de una muestra de plasma materno. Se espera que los miniSTRs se puedan usar en secuenciación masivamente paralela para la determinación simultánea de aneuploidía y fracción fetal.

Reivindicaciones

1. Un método para determinar simultáneamente aneuploidía y fracción fetal en una muestra materna de sangre, plasma o suero que comprende una mezcla de moléculas de ácidos nucleicos fetales y maternos, en el que la mezcla de moléculas de ácidos nucleicos fetales y maternos son moléculas de ADN libre de células (ADNcf) y dicha fracción fetal es la fracción de dichas moléculas de ácidos nucleicos proporcionadas por un feto a dicha mezcla de moléculas de ácidos nucleicos fetales y maternos, dicho método comprende:

(a) enriquecer dicha mezcla para una pluralidad de ácidos nucleicos objetivo polimórficos, en donde cada uno de dichos ácidos nucleicos objetivo se sabe que comprende al menos un sitio polimórfico, y en donde dicho enriquecimiento comprende:

(i) amplificar específicamente dicha pluralidad de ácidos nucleicos objetivo en una porción de dicha mezcla usando pares de cebadores cada uno capaz de amplificar una secuencia de ácidos nucleicos objetivo que comprende un sitio polimórfico en una reacción de PCR multiplex para generar un panel de sitios polimórficos amplificados que contiene un número suficiente de sitios polimórficos de tal forma que al menos dos son sitios polimórficos informativos; y

(ii) combinar al menos una porción de todo el producto amplificado con al menos una porción del resto de la muestra no amplificada de la que se obtuvo la porción;

(b) secuenciar al menos una porción de la mezcla enriquecida obtenida en el paso (a), en donde dicha secuenciación comprende proporcionar una pluralidad de lecturas de secuencia que se mapean para un genoma para identificar una pluralidad de etiquetas de secuencia mapeadas; y

(c) usar dicha pluralidad de etiquetas de secuencia mapeadas, determinar simultáneamente dicha fracción fetal y dicha aneuploidía, en donde:

(i) determinar dicha fracción fetal comprende:

(1) usar las etiquetas de secuencia mapeadas para identificar al menos dos sitios polimórficos informativos en dicho panel de sitios polimórficos amplificados, en donde dichos sitios polimórficos informativos se identifican por la diferencia en las secuencias alélicas y el número de etiquetas de secuencia mapeadas para cada uno de los alelos posibles en cada sitio polimórfico; y

(2) calcular la fracción fetal en base al número total de etiquetas de secuencia que mapean para el primer alelo y el número total de etiquetas de secuencia que mapean para un segundo alelo en cada uno de dichos sitios polimórficos informativos; y

(ii) determinar dicha aneuploidía comprende la cuantificación del número de etiquetas de secuencia que alinean con un cromosoma de interés, y comparar los resultados obtenidos para el cromosoma de interés con un valor límite, en donde el valor límite es un número que sirve como un límite de diagnóstico de una aneuploidía y en donde la presencia de una aneuploidía para el cromosoma de interés se identifica si el valor límite es excedido por los resultados obtenidos para el cromosoma de interés.

2. El método de la reivindicación 1, en donde dicho paso de enriquecimiento (a) comprende:

(i) amplificar una pluralidad de ácidos nucleicos objetivo polimórficos para generar un panel de sitios polimórficos que contiene un número suficiente de sitios polimórficos de tal manera que al menos tres son sitios polimórficos informativos;

(ii) amplificar una pluralidad de ácidos nucleicos objetivo polimórficos para generar un panel de sitios polimórficos que contiene un número suficiente de sitios polimórficos de tal manera que al menos cuatro son sitios polimórficos informativos; o

(iii) amplificar una pluralidad de ácidos nucleicos objetivo polimórficos para generar un panel de sitios polimórficos que contiene un número suficiente de sitios polimórficos de tal manera que al menos cinco son sitios polimórficos informativos.

3. El método de la Reivindicación 1 o la Reivindicación 2, en donde determinar dicha fracción fetal en el paso (c) comprende calcular la fracción fetal en base a:

(i) el número total de etiquetas de secuencia que mapean par aun primer alelo y el número total de etiquetas de secuencia que mapean para un segundo alelo para cada uno de los al menos tres alelos informativos;

(ii) el número total de etiquetas de secuencia que mapean par aun primer alelo y el número total de etiquetas de secuencia que mapean para un segundo alelo para cada uno de los menos cuatro alelos informativos; o

(iii) el número total de etiquetas de secuencia que mapean par aun primer alelo y el número total de etiquetas de secuencia que mapean para un segundo alelo para cada uno de los al menos cinco alelos informativos.

4. El método de cualquiera de las Reivindicaciones 1-3, en donde dicho paso de enriquecimiento (a) comprende

amplificar una pluralidad de ácidos nucleicos objetivo polimórficos en una porción de una mezcla purificada de ácidos nucleicos fetales y maternos.

5. El método de cualquiera de las Reivindicaciones 1-3, en donde dicho enriquecimiento comprende:

- (i) preparar una primera biblioteca de secuenciación de dicha mezcla de moléculas de ácidos nucleicos fetales y maternos;
- (ii) preparar una segunda biblioteca de secuenciación de dichos ácidos nucleicos objetivo polimórficos; y
- (iii) combinar al menos una porción de dicha primera biblioteca de secuenciación con al menos una porción de dicha segunda biblioteca de secuenciación.

6. El método de cualquiera de las Reivindicaciones anteriores, en donde dichos ácidos nucleicos objetivo polimórficos están localizados en el mismo cromosoma.

7. El método de cualquiera de las Reivindicaciones anteriores, en donde dichos ácidos nucleicos objetivo polimórficos están localizados en cromosomas diferentes.

8. El método de cualquiera de las Reivindicaciones anteriores, en donde cada uno de dicha pluralidad de ácidos nucleicos objetivo polimórficos comprende al menos un polimorfismo de nucleótido simple (SNP).

9. El método de la Reivindicación 8, en donde el SNP es un SNP individual.

10. El método de la Reivindicación 9, en donde dicho al menos un SNP es un SNP individual seleccionado de cada uno de dicha pluralidad de ácidos nucleicos objetivo polimórficos que comprende un SNP seleccionado de rs560681, rs1109037, rs9866013, rs13182883, rs13218440, rs7041158, rs740598, rs10773760, rs4530059, rs7205345, rs8078417, rs576261, rs2567608, rs430046, rs9951171, rs338882, rs10776839, rs9905977, rs1277284, rs258684, rs1347696, rs508485, rs9788670, rs8137254, rs3143, rs2182957, rs3739005 y rs530022.

11. El método de la Reivindicación 8, en donde el SNP es un SNP en tándem.

12. El método de la Reivindicación 11, en donde dicho al menos un SNP es un SNP en tándem seleccionado de los conjuntos de SNPs en tándem rs7277033-rs2110153; rs2822654-rs1882882; rs368657-rs376635; rs2822731-rs2822732; rs1475881-rs7275487; rs1735976-rs2827016; rs447340-rs2824097; rs418989-rs13047336; rs987980-rs987981; rs4143392-rs4143391; rs1691324-rs13050434; rs11909758-rs9980111; rs2826842-rs232414; rs1980969-rs1980970; rs9978999-rs9979175; rs1034346-rs12481852; rs7509629-rs2828358; rs4817013-rs7277036; rs9981121-rs2829696; rs455921-rs2898102; rs2898102-rs458848; rs961301-rs2830208; rs2174536-rs458076; rs11088023-rs11088024; rs1011734-rs1011733; rs2831244-rs9789838; rs8132769-rs2831440; rs8134080-rs2831524; rs4817219-rs4817220; rs2250911-rs2250997; rs2831899-rs2831900; rs2831902-rs2831903; rs11088086-rs2251447; rs2832040-rs11088088; rs2832141-rs2246777; rs2832959-rs9980934; rs2833734-rs2833735; rs933121-rs933122; rs2834140-rs12626953; rs2834485-rs3453; rs9974986-rs2834703; rs2776266-rs2835001; rs1984014-rs1984015; rs7281674-rs2835316; rs13047304-rs13047322; rs2835545-rs4816551; rs2835735-rs2835736; rs13047608-rs2835826; rs2836550-rs2212596; rs2836660-rs2836661; rs465612-rs8131220; rs9980072-rs8130031; rs418359-rs2836926; rs7278447-rs7278858; rs385787-rs367001; rs367001-rs386095; rs2837296-rs2837297; y rs2837381-rs4816672.

13. El método de cualquiera de las Reivindicaciones 1-7, en donde cada uno de dicha pluralidad de ácidos nucleicos objetivo comprende al menos una repetición corta en tándem (STR).

14. El método de la Reivindicación 13, en donde cada uno de dicha pluralidad de ácidos nucleicos objetivo polimórficos es una STR seleccionada de CSF1PO, FGA, TH01, TPOX, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, D2S1338, Penta D, Penta E, D22S1045, D20S1082, D20S482, D18S853, D17S1301, D17S974, D14S1434, D12ATA63, D11S4463, D10S1435, D10S1248, D9S2157, D9S1122, D8S1115, D6S1017, D6S474, D5S2500, D4S2408, D4S2364, D3S4529, D3S3053, D2S1776, D2S441, D1S1677, D1S1627 y D1GATA113.

15. El método de la Reivindicación 13, en donde dicha al menos una STR es menor de alrededor de 300 pares de bases.

16. El método de cualquiera de las Reivindicaciones anteriores, en donde dicha secuenciación:

- (i) es secuenciación de próxima generación (NGS);
- (ii) es secuenciación masivamente paralela usando secuenciación por síntesis con terminadores de colorante reversibles;
- (iii) es secuenciación por ligadura;
- (iv) comprende una amplificación; o

(v) es secuenciación de moléculas individuales.

17. El método de cualquiera de las Reivindicaciones 1-16, en donde dicha aneuploidía es una aneuploidía cromosómica completa; preferiblemente una aneuploidía elegida de trisomía 8, trisomía 13, trisomía 15, trisomía 16, trisomía 18, trisomía 21, trisomía 22, monosomía X, XXX, XXY y XYY.

18. El método de cualquiera de las Reivindicaciones 1-16, en donde dicha aneuploidía es una aneuploidía cromosómica parcial.

5

10

15

20

25

30

35

40

45

50

55

60

65

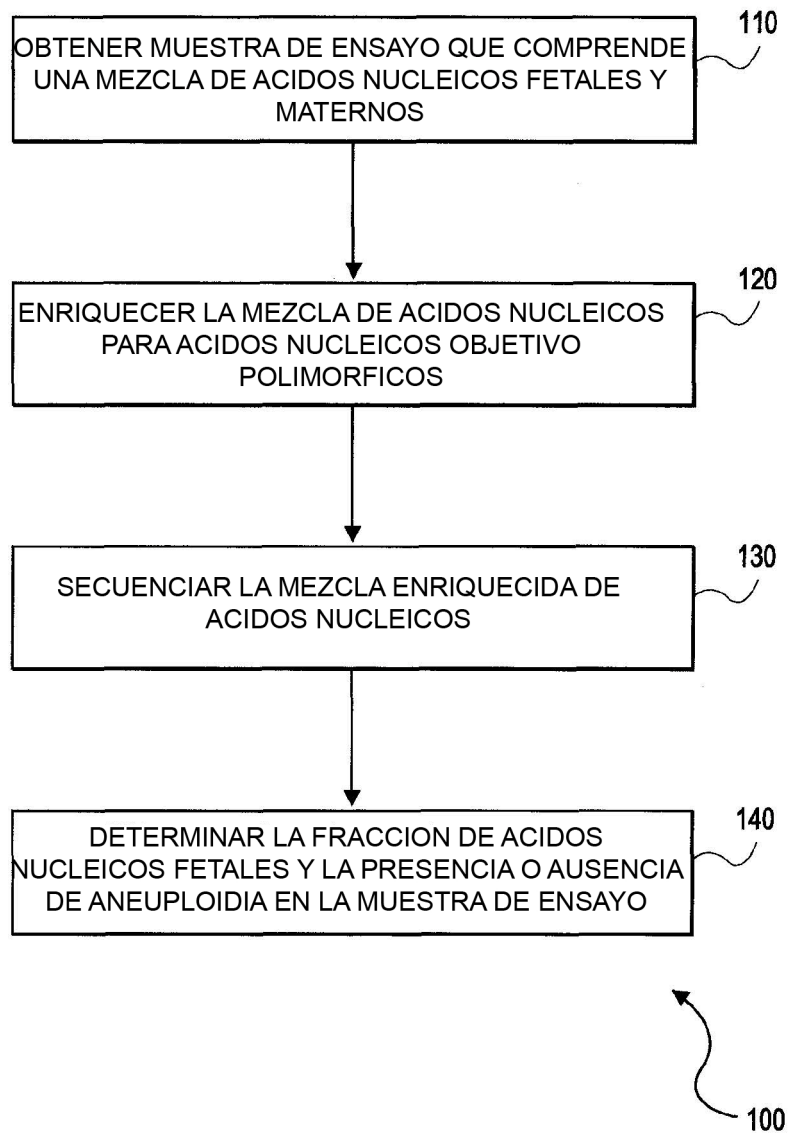
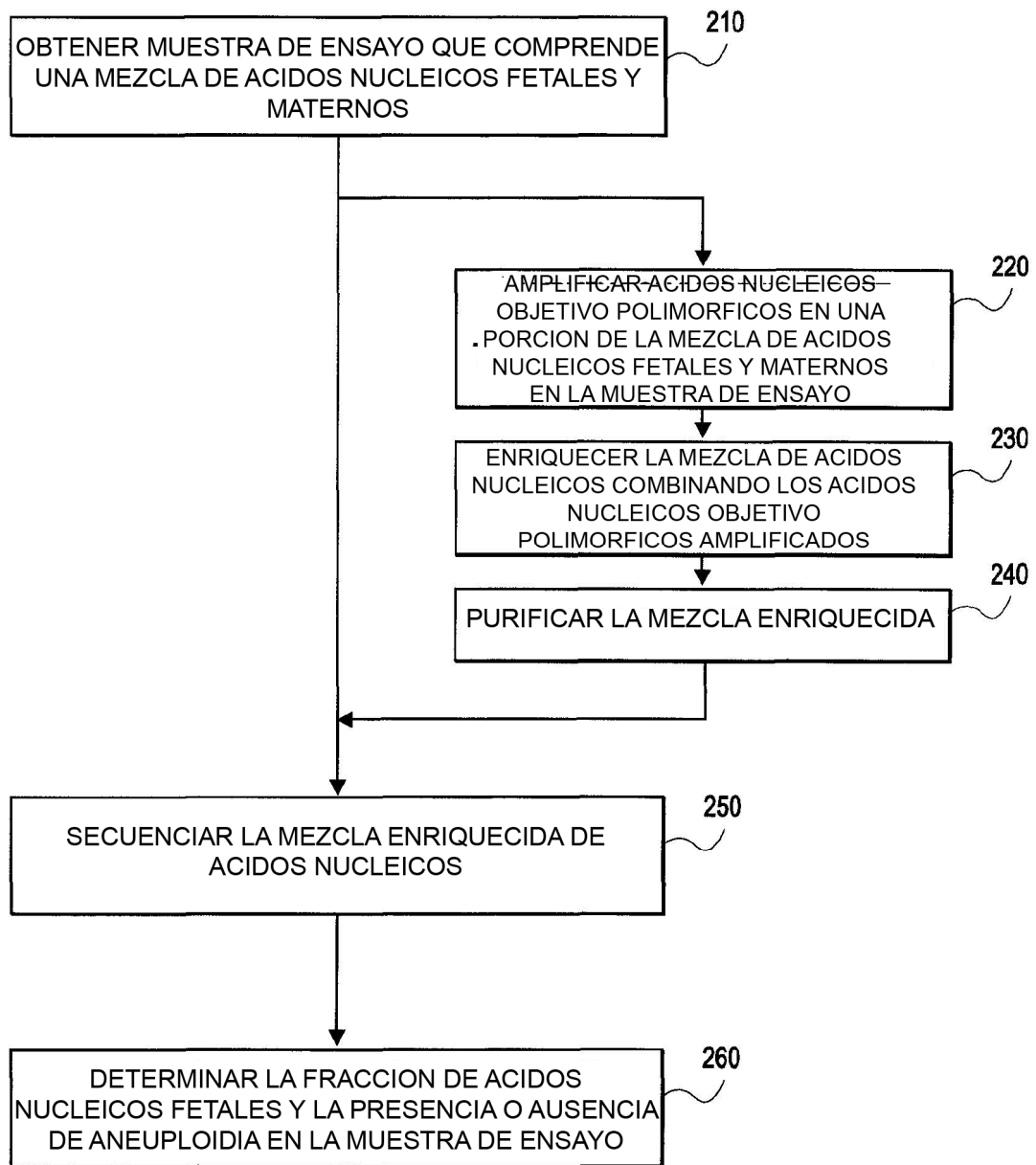
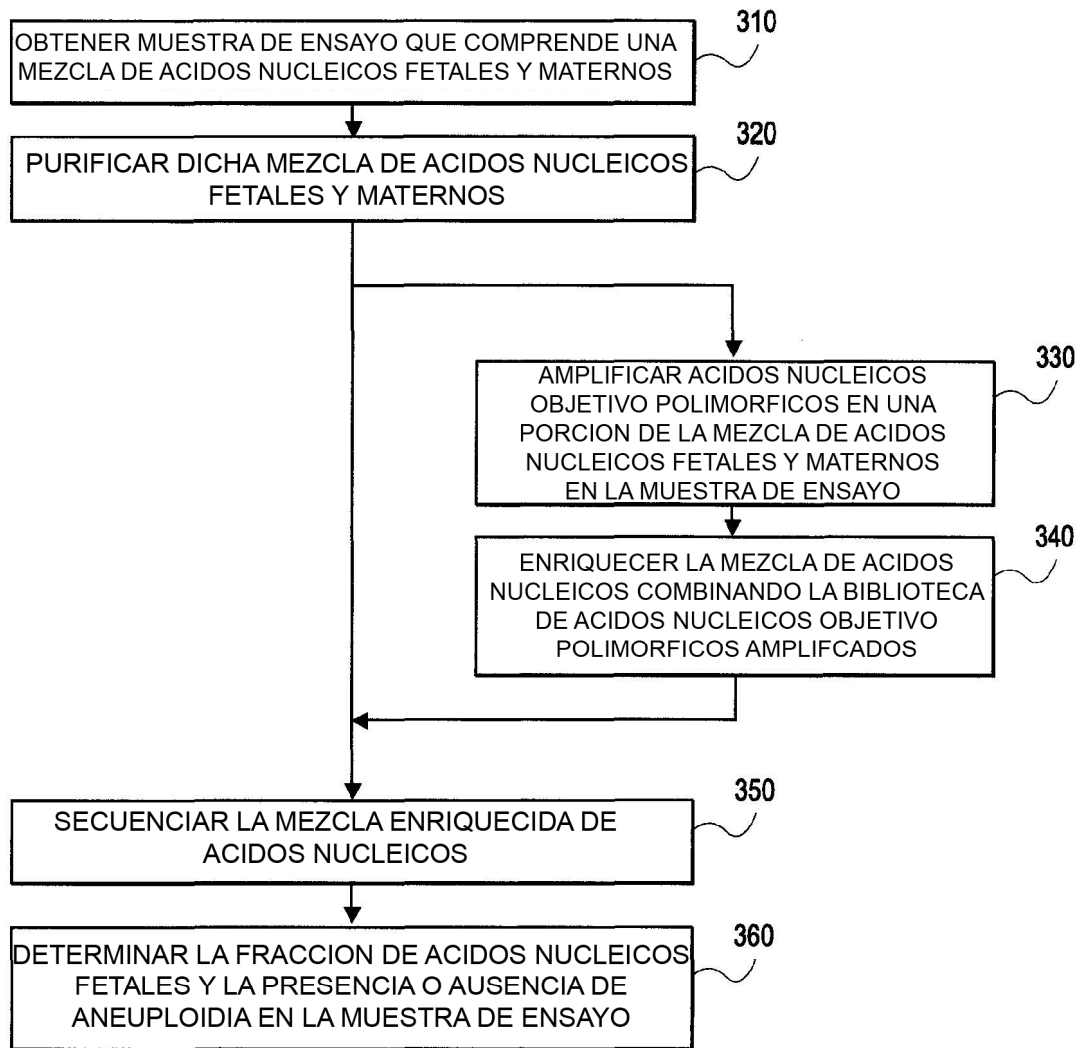


FIG. 1



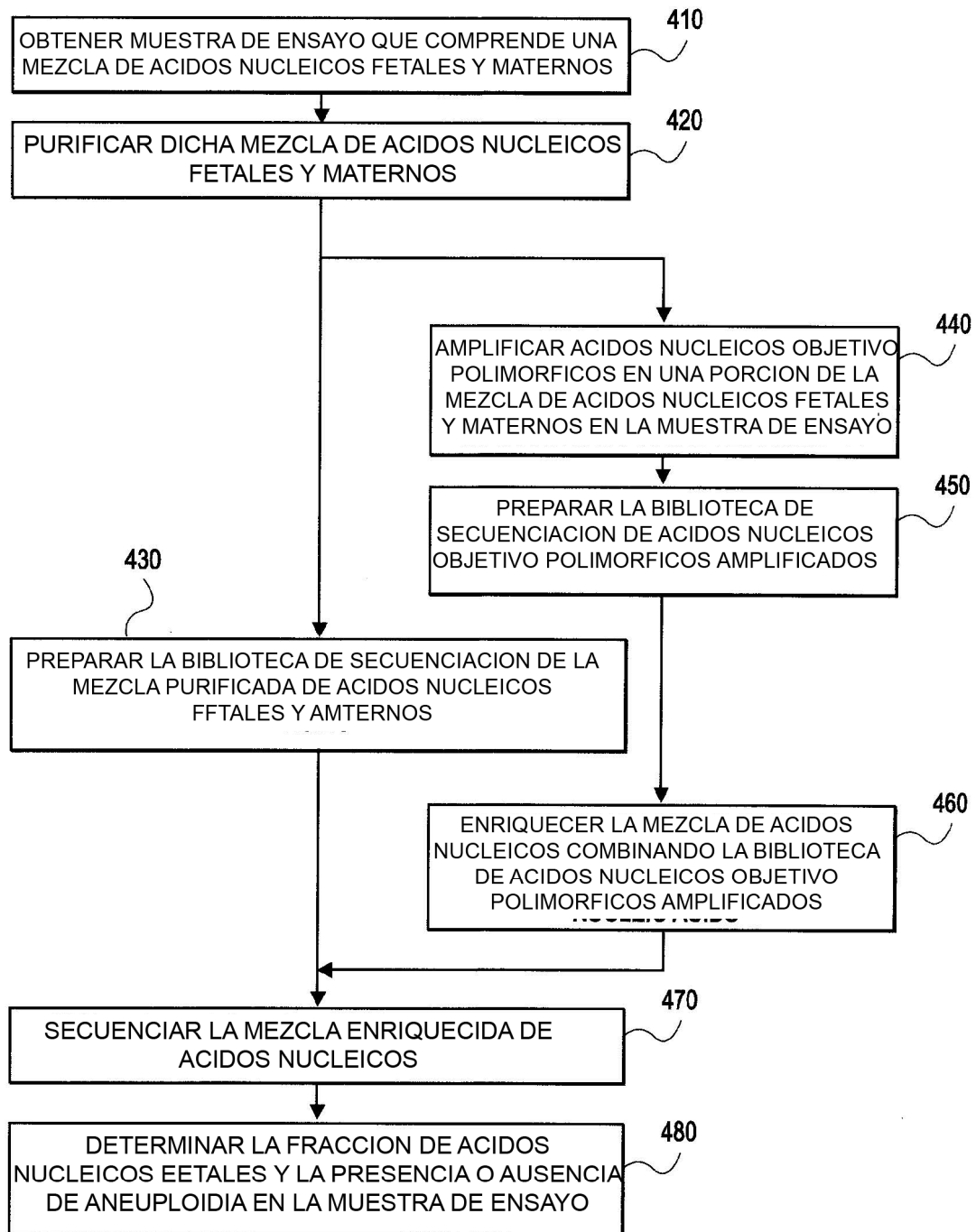
200

FIG. 2



300

FIG. 3



400

FIG. 4

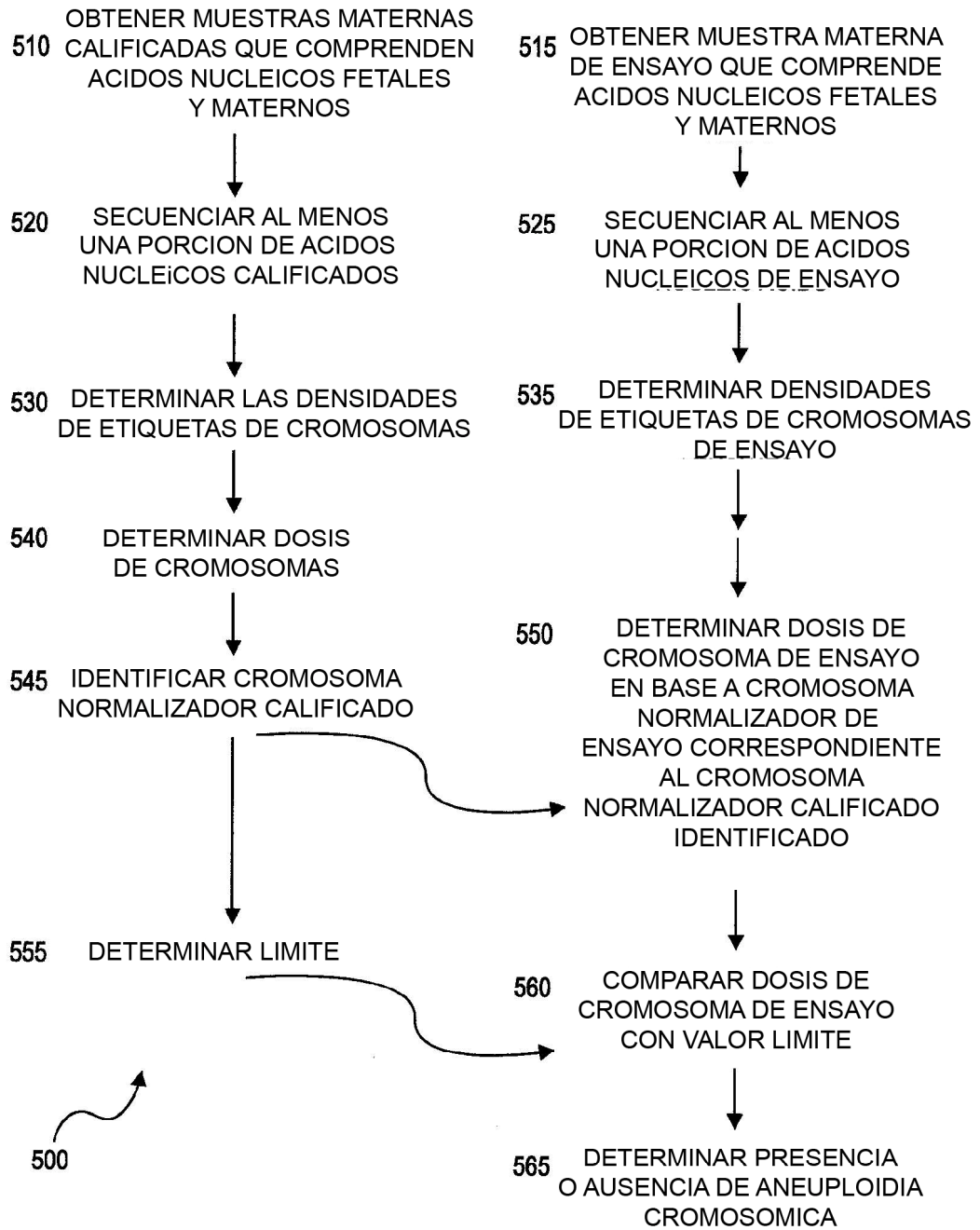


FIG. 5

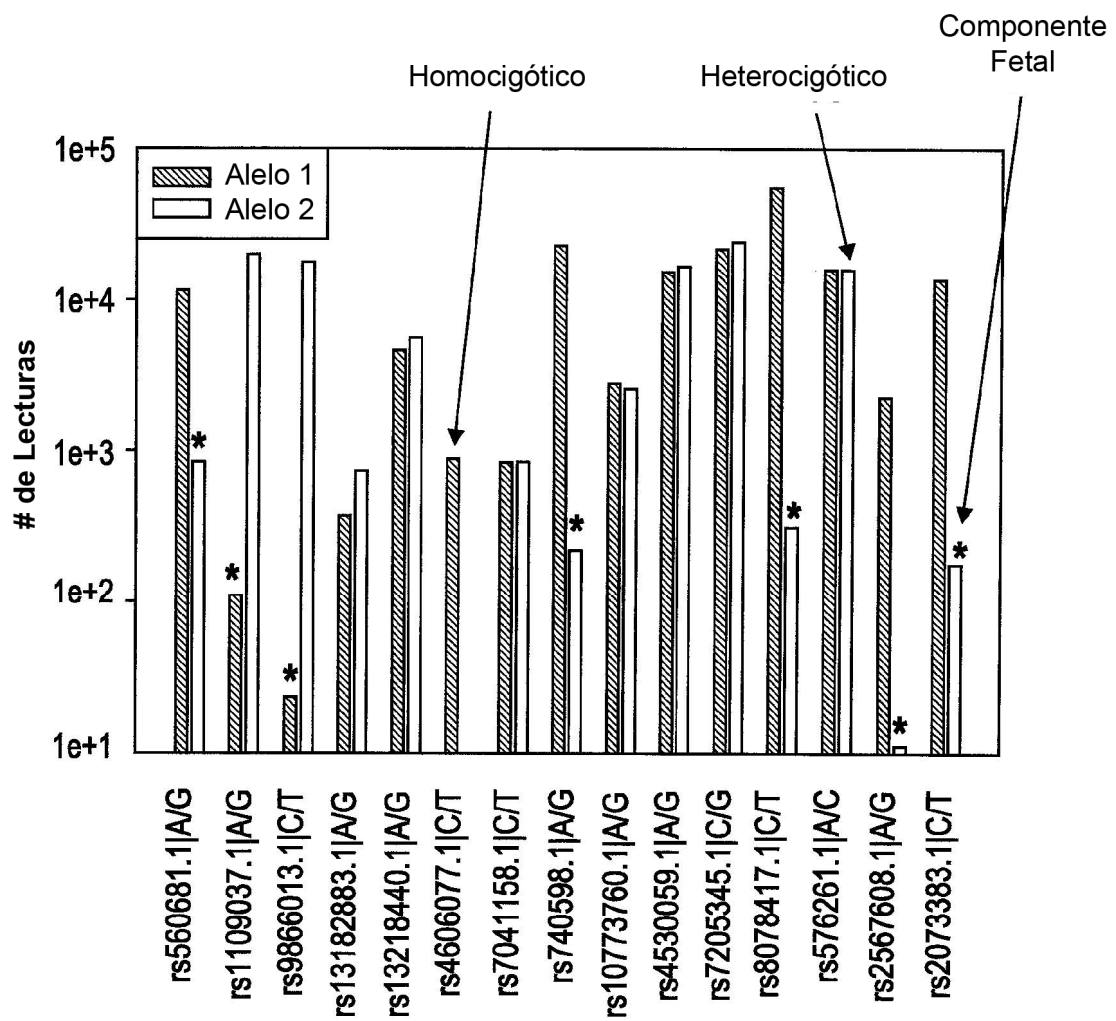


FIG. 6

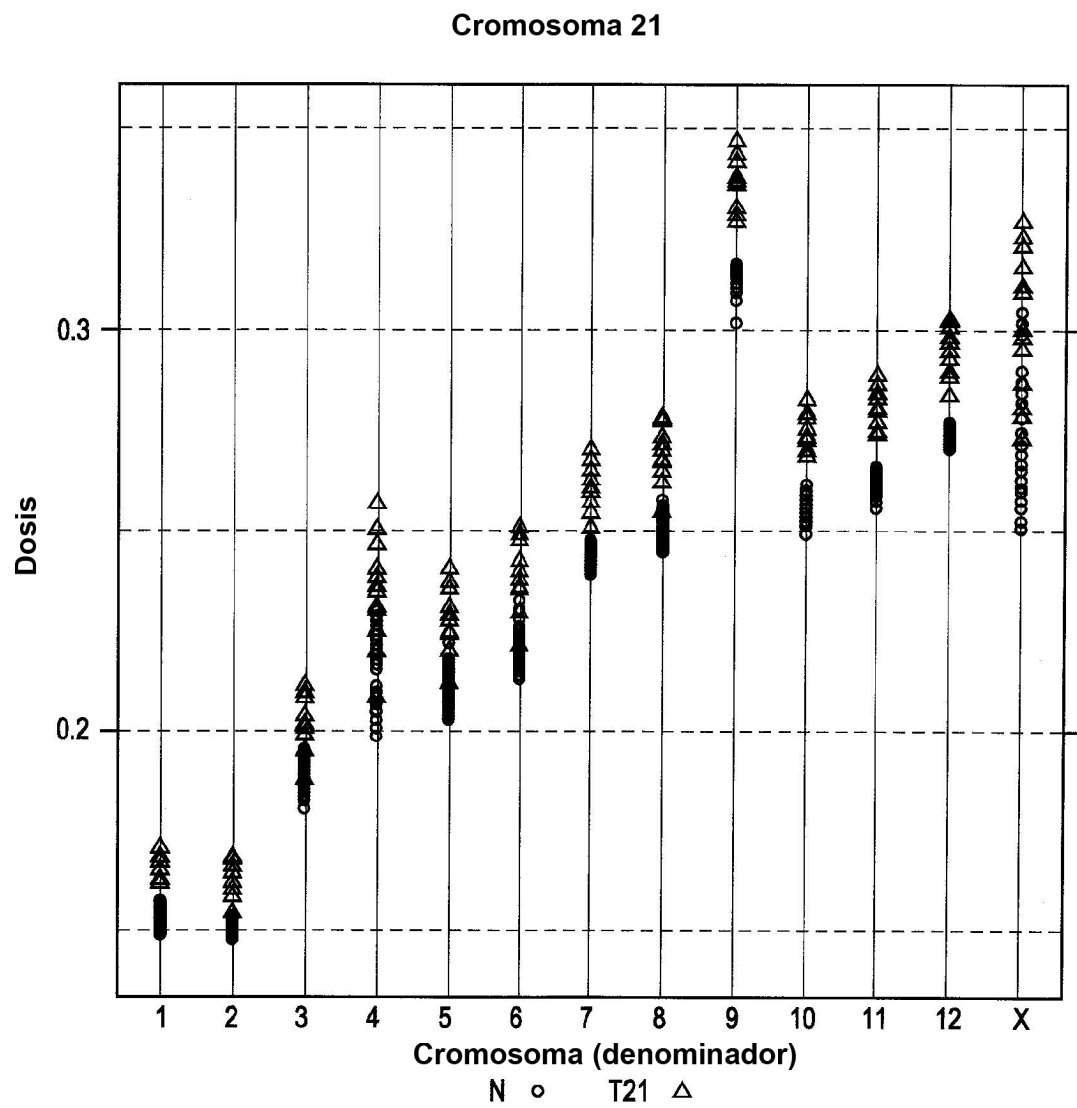


FIG. 7A

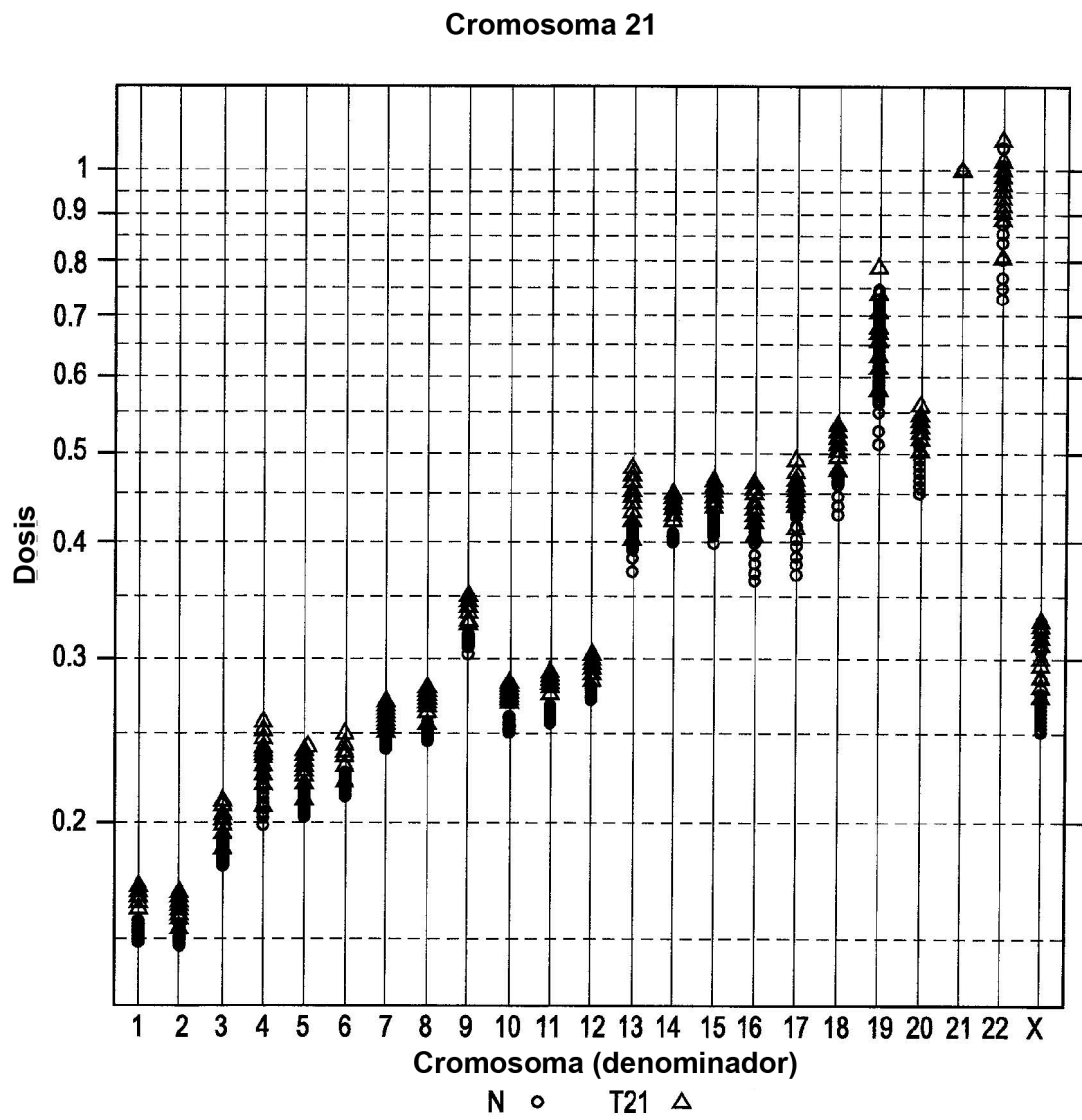


FIG. 7B

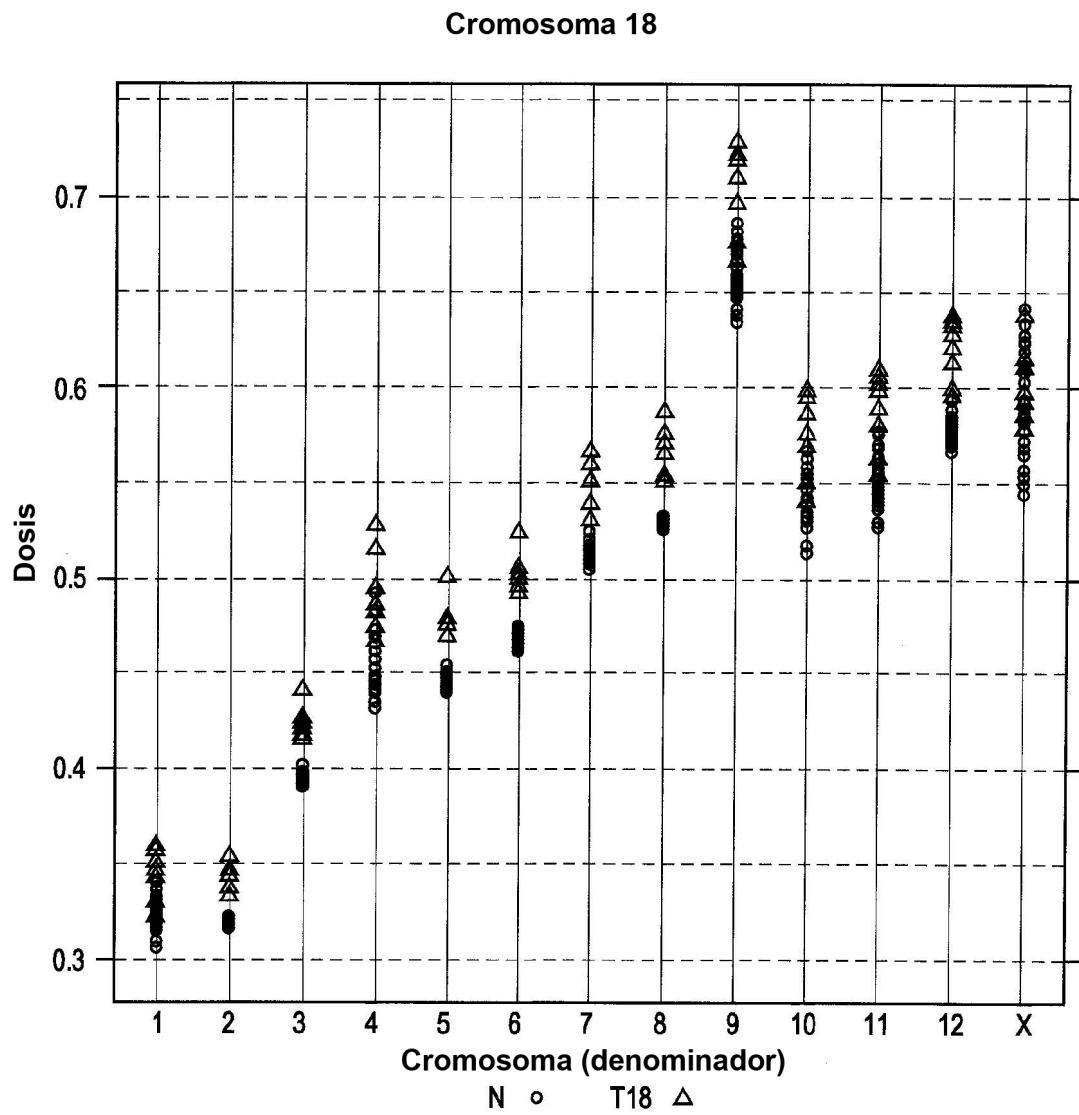


FIG. 8A

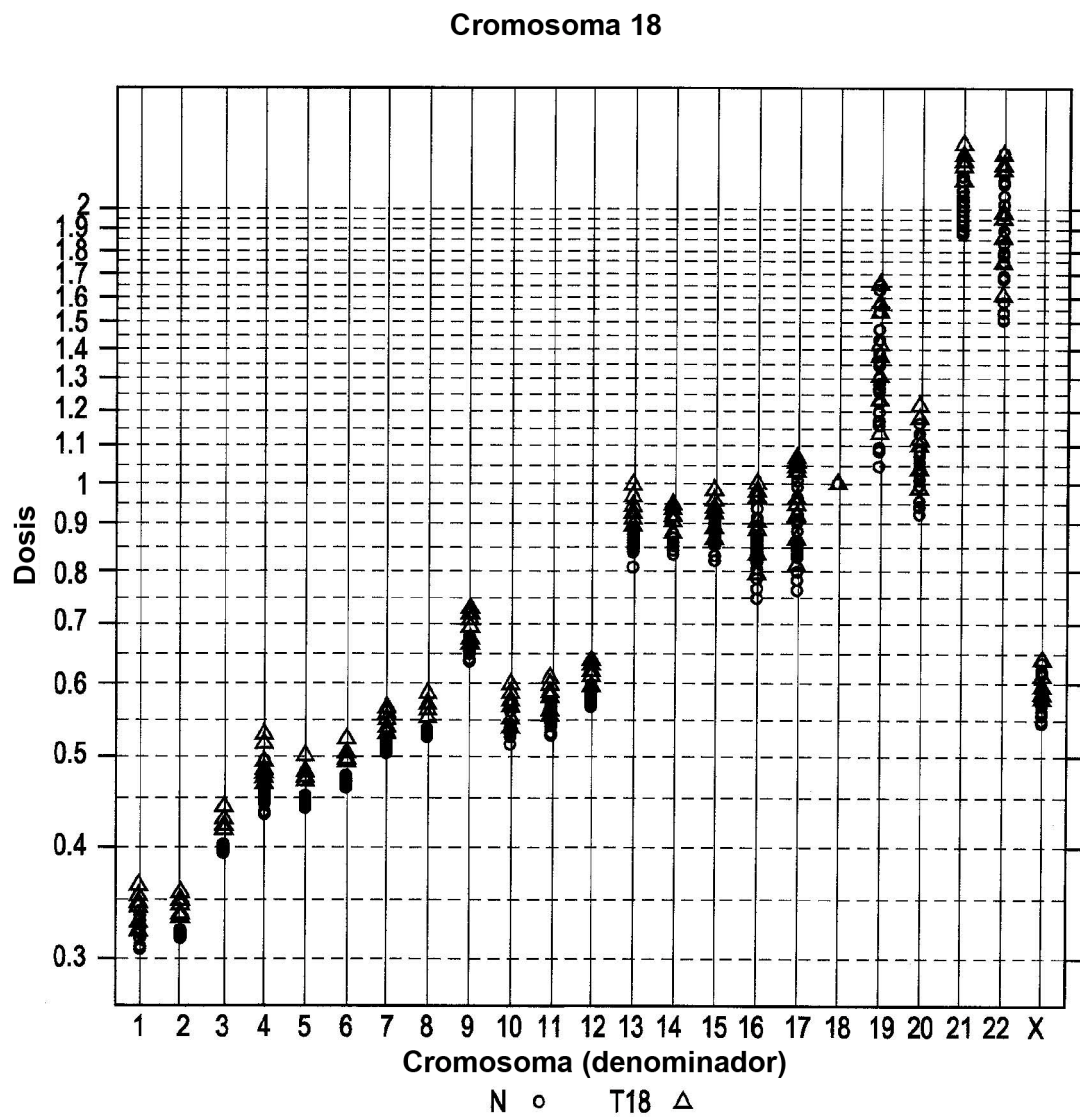


FIG. 8B

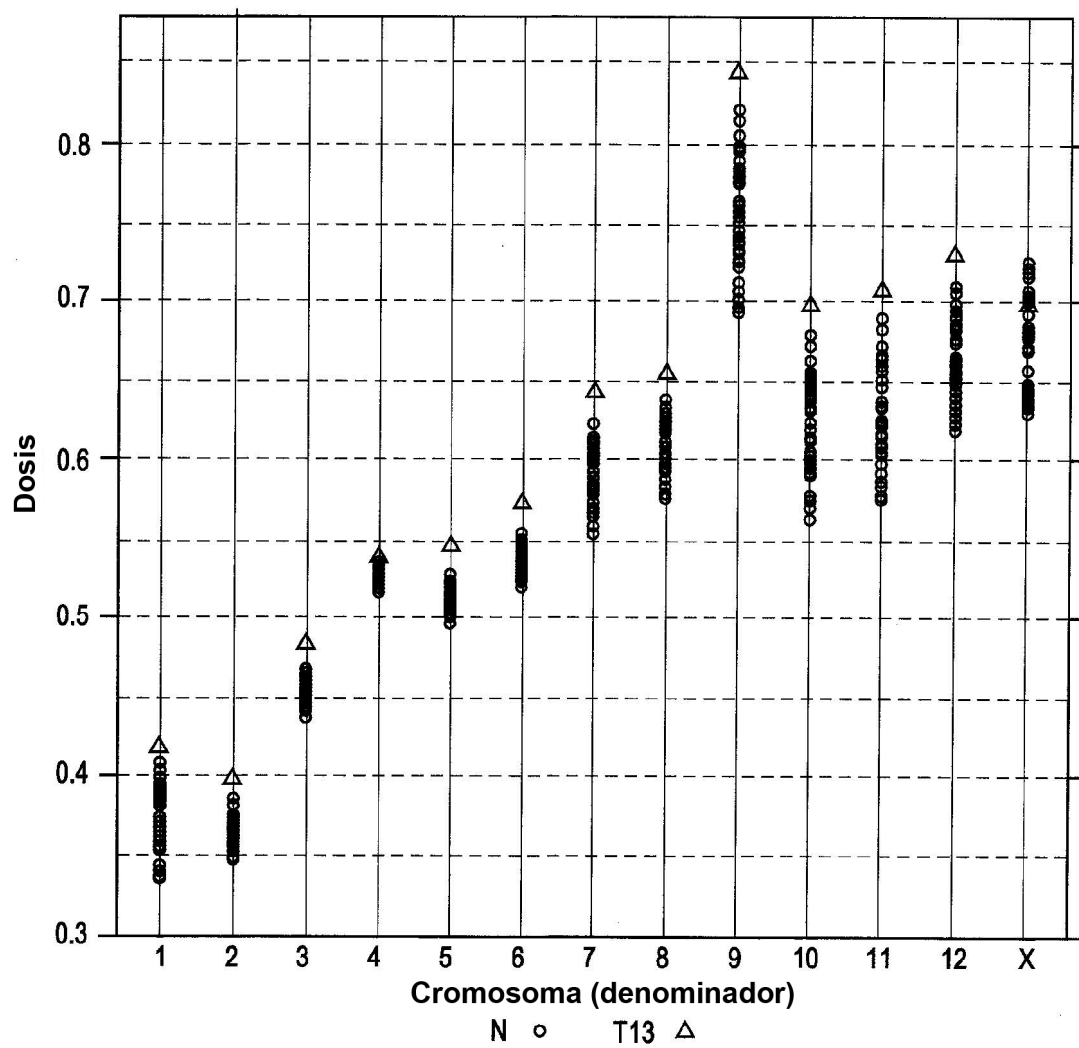


FIG. 9A

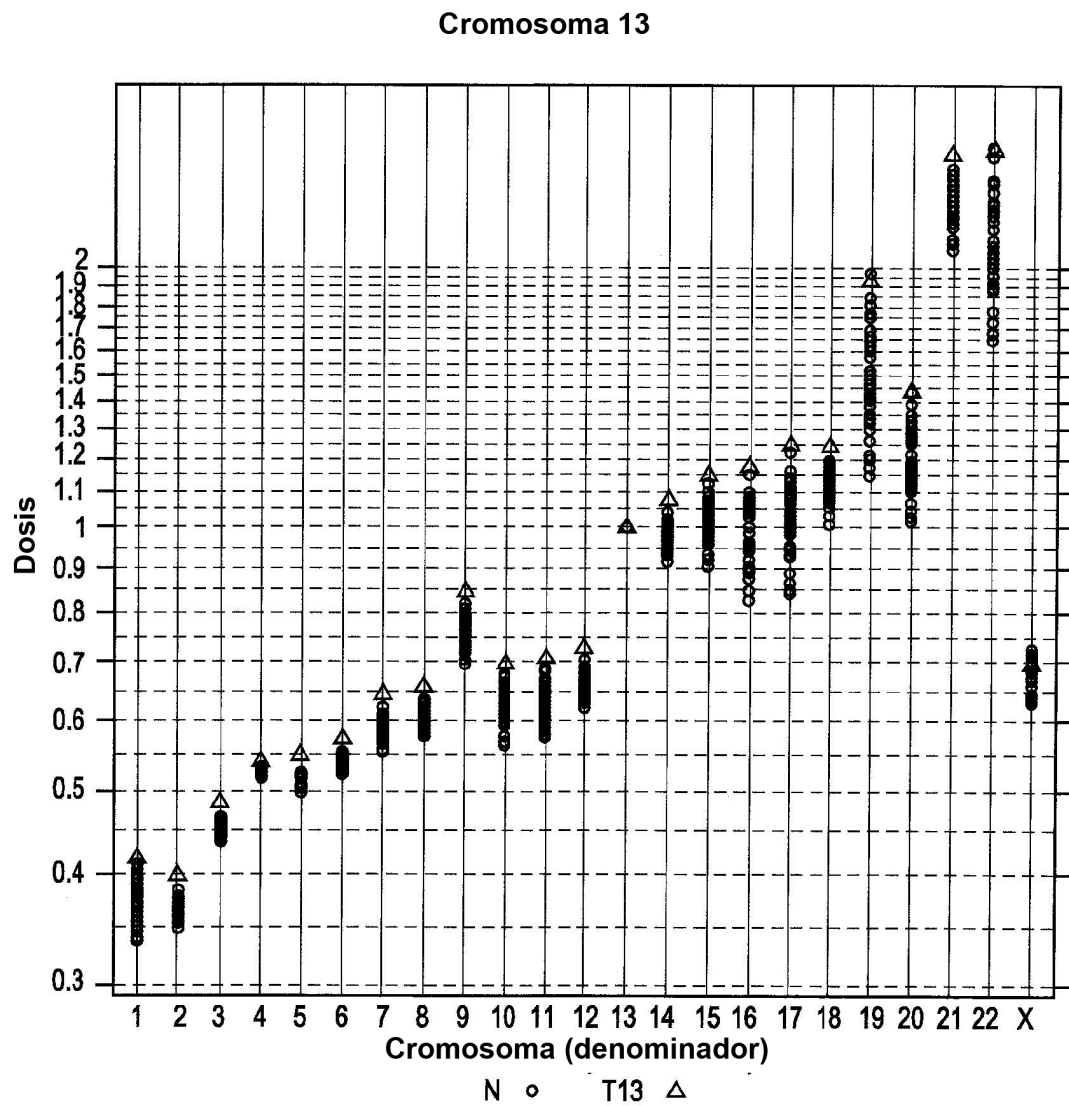


FIG. 9B

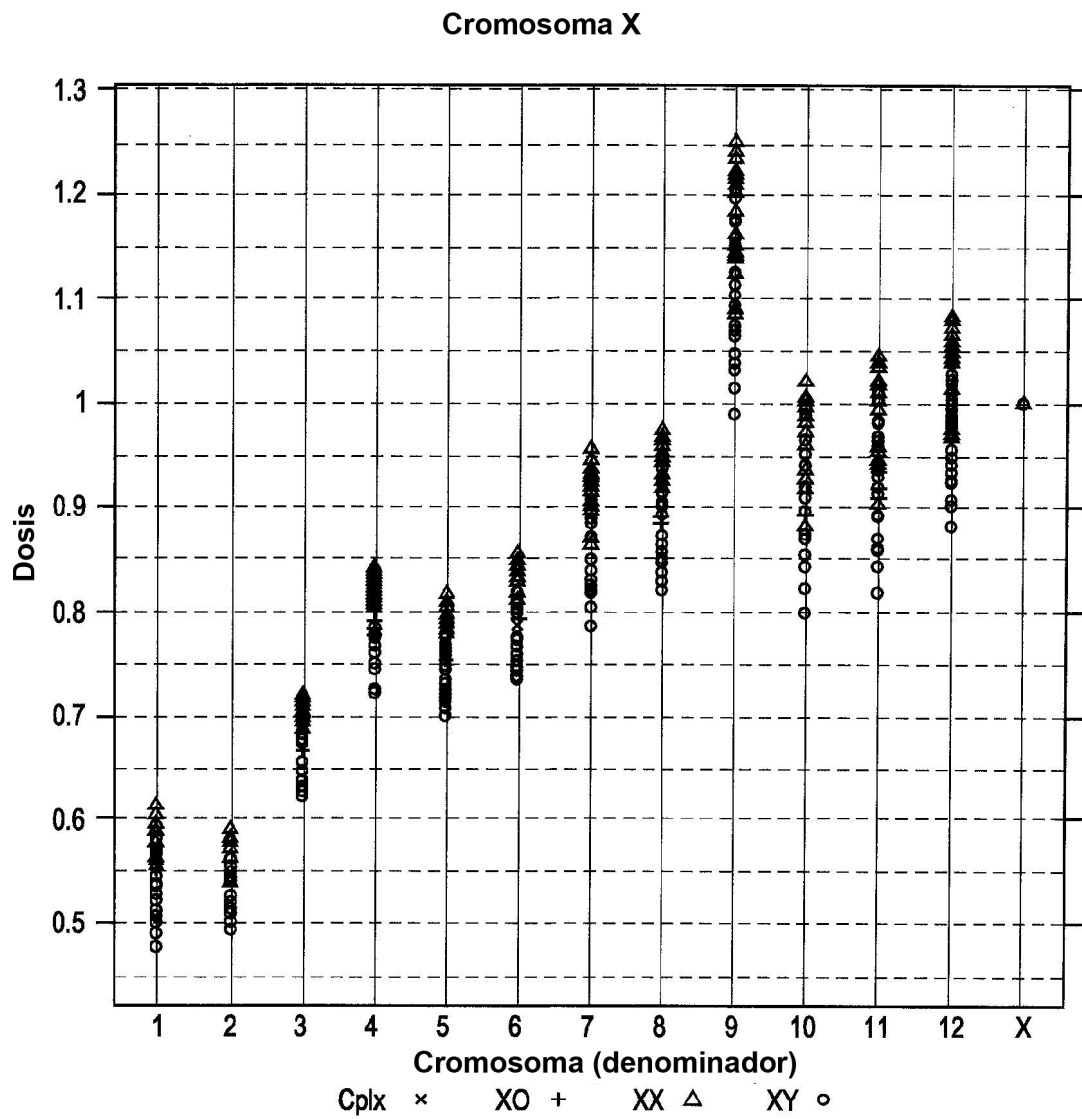


FIG. 10A

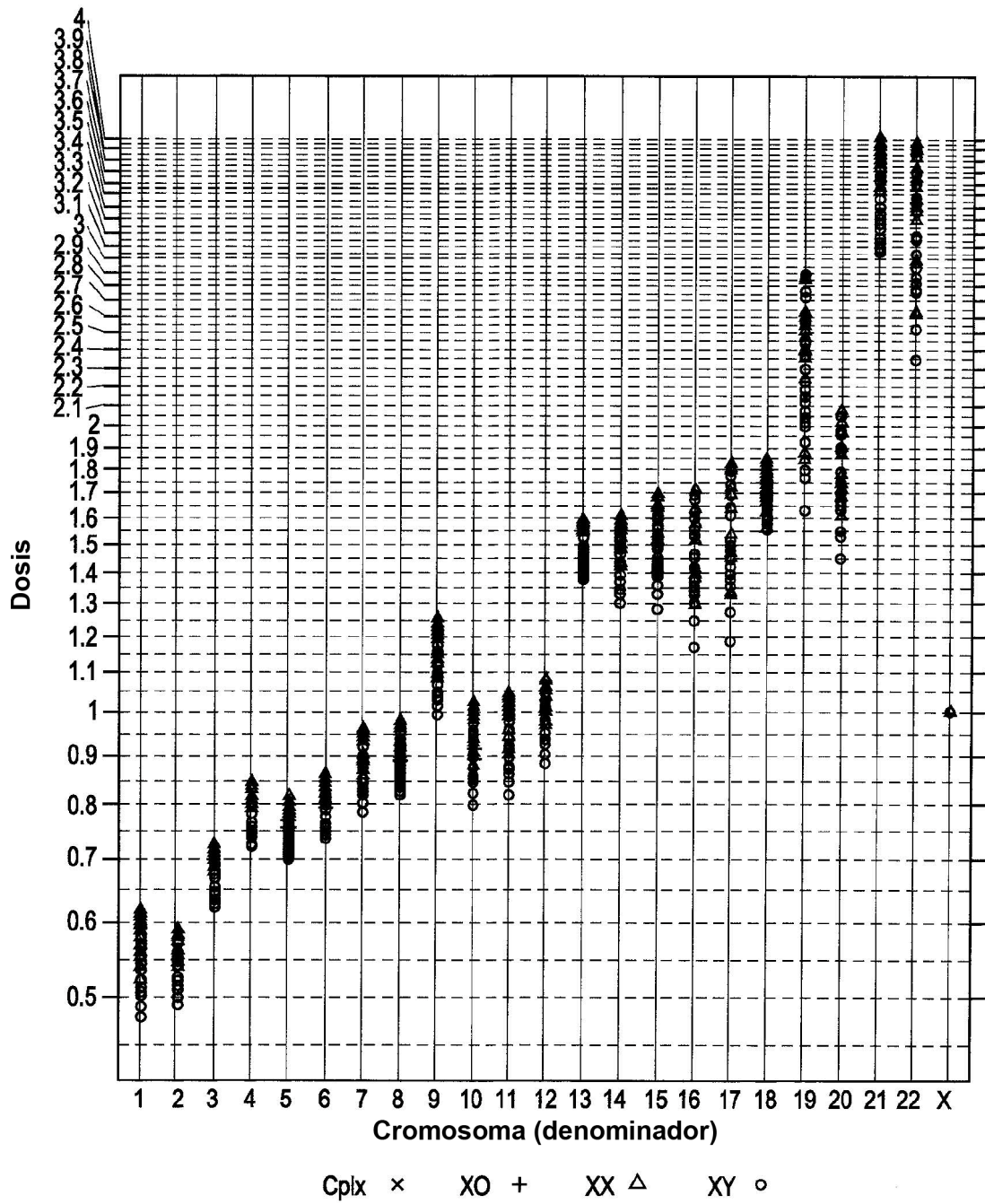


FIG. 10B

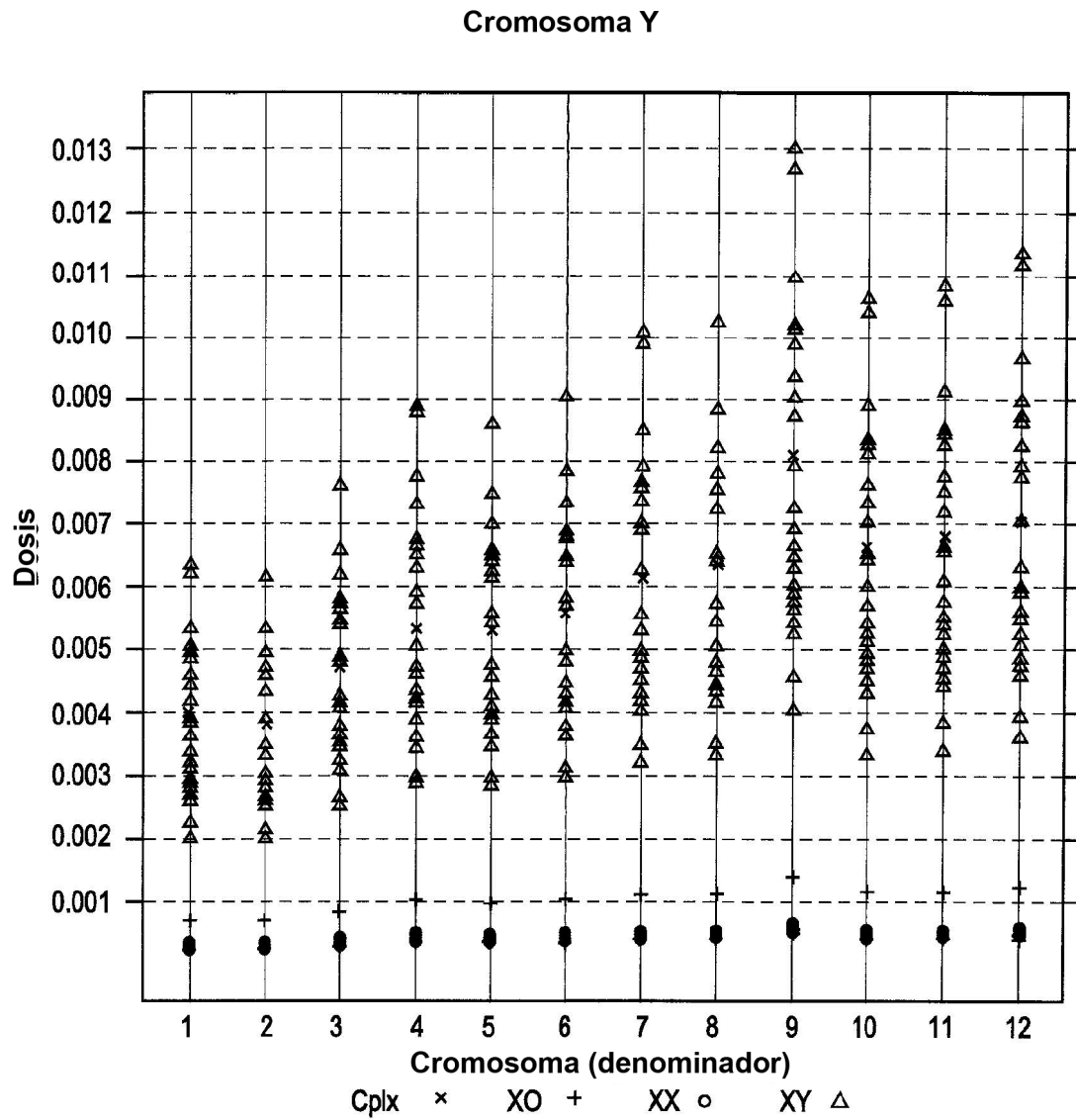


FIG. 11A

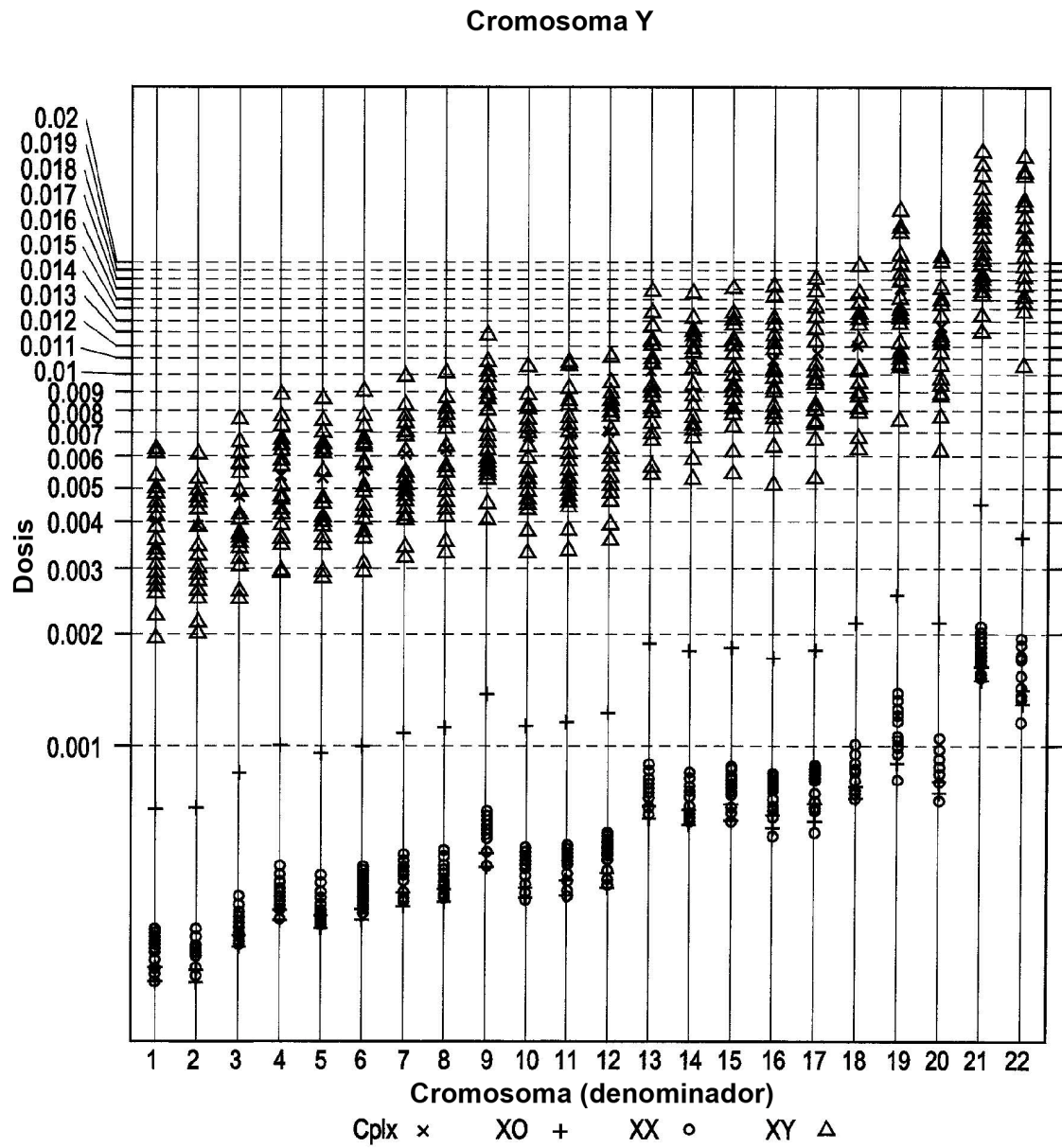


FIG. 11B

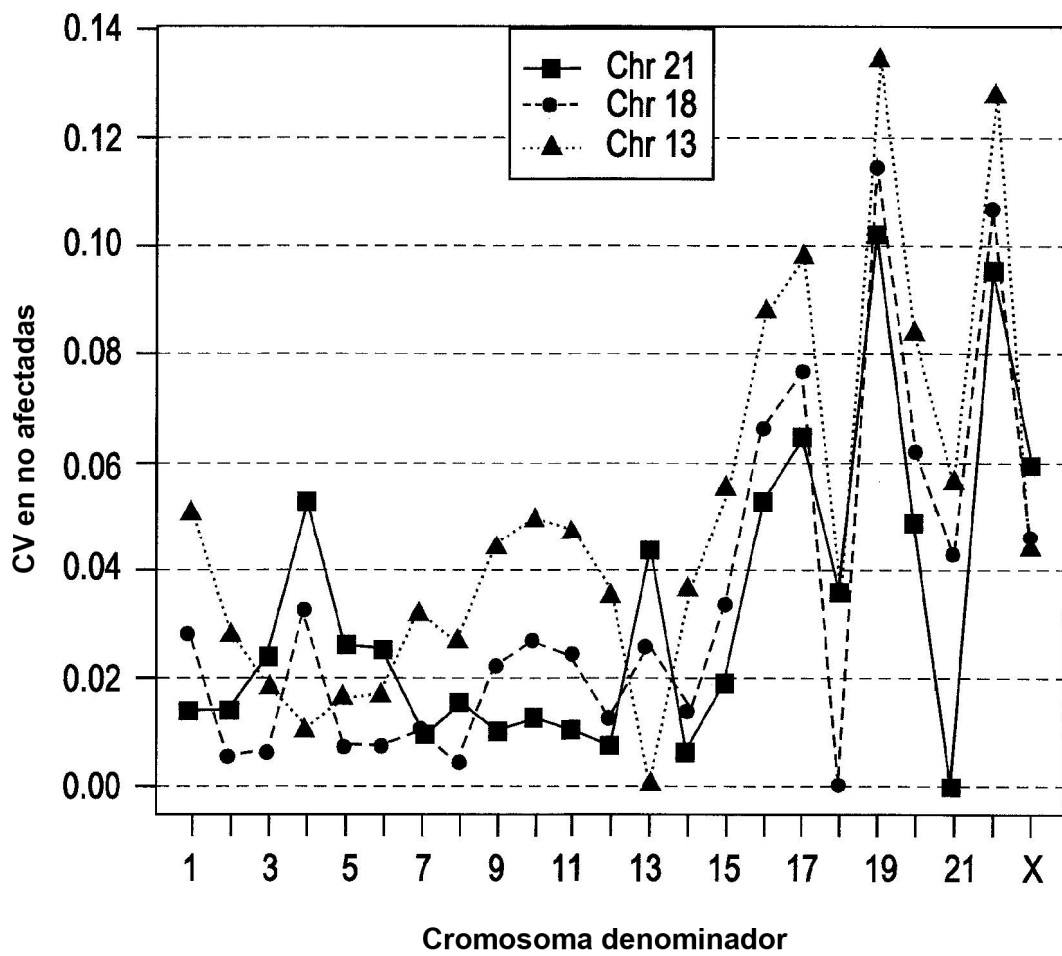


FIG. 12

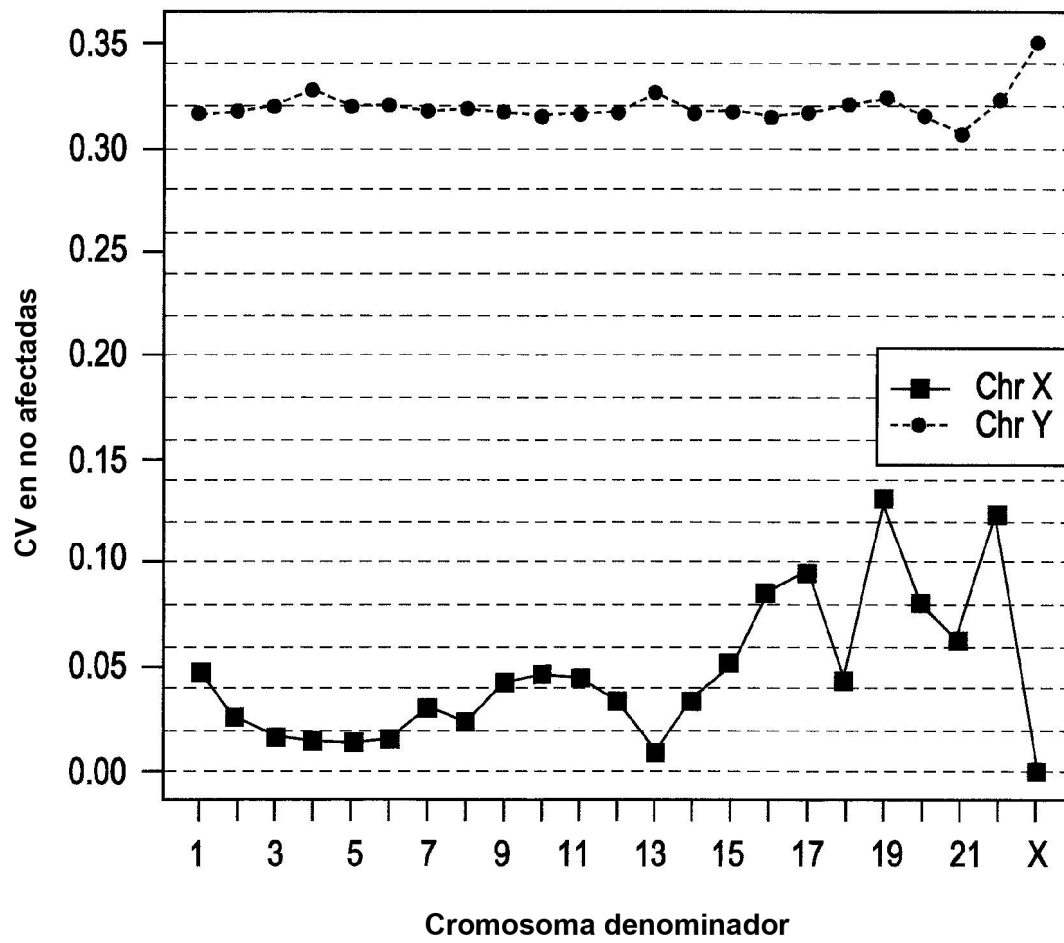


FIG. 13