

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 704 701**

51 Int. Cl.:

C12Q 1/68 (2008.01)

C12Q 1/6806 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **01.12.2010 E 17180803 (3)**

97 Fecha y número de publicación de la concesión europea: **24.10.2018 EP 3260555**

54 Título: **Nuevo protocolo de preparación de bibliotecas de secuenciación**

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:

19.03.2019

73 Titular/es:

VERINATA HEALTH, INC. (100.0%)

800 Saginaw Drive

Redwood City CA 94063, US

72 Inventor/es:

RAVA, RICHARD P;

CHINNAPPA, MANJULA;

COMSTOCK, DAVID A.;

HEILEK, GABRIELLE y

RHEES, BRIAN KENT

74 Agente/Representante:

IZQUIERDO BLANCO, María Alicia

ES 2 704 701 T3

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

DESCRIPCIÓN

Nuevo protocolo de preparación de bibliotecas de secuenciación

5 2. ANTECEDENTES DE LA INVENCIÓN.

[0001] La detección y el diagnóstico prenatales son una parte rutinaria de la atención prenatal. Actualmente, el diagnóstico prenatal de afecciones genéticas y cromosómicas involucra pruebas invasivas, como la amniocentesis o el muestreo de vellosidades coriónicas (CVS), realizadas a partir de las 11 semanas de gestación y con un riesgo de aborto $\geq 1\%$. La existencia de ADN libre de células en circulación en la sangre materna (Lo et al., Lancet 350: 485-487 [1997]) se está explotando para desarrollar procesos no invasivos que utilizan ácidos nucleicos fetales de una muestra de sangre periférica materna para determinar anomalías del cromosoma fetal. (Fan HC y Quake SR Anal Chem 79: 7576-7579 [2007]; Fan et al., Proc Natl Acad Sci 105: 16266-16271 [2008]). Estos métodos ofrecen una fuente alternativa y más segura de material genético fetal para el diagnóstico prenatal, y podrían pronunciar de manera efectiva el final de los procedimientos invasivos.

[0002] La secuenciación de ácidos nucleicos está evolucionando rápidamente como una técnica de diagnóstico en el laboratorio clínico. Las aplicaciones que involucran la secuenciación se observan en varias áreas, incluidas las pruebas de cáncer que abarcan pruebas genéticas para la predisposición al cáncer y la evaluación de mutaciones genéticas en el cáncer; la genética abarca las pruebas de portador y el diagnóstico de enfermedades transmitidas genéticamente; y microbiología que abarca genotipos virales y secuencias asociadas con resistencia a fármacos.

[0003] El advenimiento de tecnologías de secuenciación de próxima generación (NGS) que permiten la secuenciación de genomas enteros en tiempo relativamente corto, ha proporcionado la oportunidad de comparar el material genético procedente de un cromosoma a ser comparada con la de otro sin los riesgos asociados con métodos de muestreo invasivo. Sin embargo, las limitaciones de los métodos existentes, que incluyen una sensibilidad insuficiente derivada de los niveles limitados de ADNcf, y el sesgo de secuenciación de la tecnología derivada de la naturaleza inherente de la información genómica, subyacen a la necesidad continua de métodos no invasivos que proporcionen cualquiera o todos de la especificidad, la sensibilidad y la aplicabilidad para diagnosticar de manera confiable las aneuploidías fetales en una variedad de entornos clínicos.

[0004] Ya que la secuenciación de ácido nucleico ha entrado en el área clínica para las pruebas de cáncer, organizaciones tales como la NCCLS (Consejo Nacional de Servicios de Laboratorio clínico) y la Asociación de la Citogenética Clínica han proporcionado pautas para la normalización de las pruebas basadas en secuenciación existentes que utilizan secuenciación de extensión de cebador, terminada en dideoxi, basada en PCR realizada en gel o capilar (NCCLS: Nucleic Acid Sequencing Methods in Diagnostic Laboratory Medicine MM9-A, Vol. 24 Núm. 40), Secuenciación de Sanger y QF-PCR (Asociación para la Sociedad de Citogenética Clínica y Genética Molecular Clínica, Guías de práctica para el análisis e interpretación de secuenciación de Sanger ratificada por el Comité Ejecutivo de CMGS el 7 de agosto de 2009, disponible en la dirección web cmgs.org/BPGs/pdfs%20current%20bpgs/Sequencingv2.pdf QF-PCR para el diagnóstico de las mejores guías de aneuploidía (2007) v2.01). Las directrices se basan en pruebas de consenso de varios protocolos y, entre otras cosas, tienen como objetivo reducir la aparición de eventos adversos en el laboratorio clínico, por ejemplo, mezclas de muestras, al tiempo que se preserva la calidad y la fiabilidad de los ensayos. Ya que los laboratorios clínicos ya están experimentando con NIPD, se desarrollarán procedimientos de calidad para implementar las nuevas tecnologías de secuenciación para proporcionar sistemas de atención de salud apropiados y seguros.

[0005] Chu et al. (Bioinformatics, 25 (10), mayo de 2009, páginas 1244-1250) describe un "Statistical model for whole genome sequencing and its application to minimally invasive diagnosis of fetal genetic disease".

[0006] La presente invención se refiere a métodos de secuenciación de nueva generación fiables que son aplicables al menos para la práctica de diagnóstico prenatal no invasivo, y abarca los procedimientos que aumentan la rapidez y la calidad de los métodos y reducir al mínimo la pérdida de material, y reducir la probabilidad de errores de muestra

55 3. RESUMEN DE LA INVENCIÓN.

[0007] La invención se define por las reivindicaciones adjuntas. La invención proporciona un método para preparar una biblioteca de secuenciación a partir de una muestra de prueba que comprende moléculas de ácido nucleico, en donde el método comprende los pasos consecutivos de la reparación de los extremos, las colas de dA y el adaptador que ligan dichos ácidos nucleicos, y en donde dichos pasos consecutivos excluyen la purificación de los productos reparados de extremo antes de la etapa de colas de dA y excluyen la purificación de los productos de colas de dA antes de la etapa de ligadura adaptadora.

[0008] La invención también proporciona el uso de una biblioteca de secuenciación preparada de acuerdo con el método de la invención en un método de secuenciación de ácido nucleico.

[0009] La presente descripción también se refiere a métodos para determinar la aneuploidía y/o fracción fetal en muestras maternas comprenden ADNcf fetal y materna por secuenciación masiva en paralelo. El método comprende un protocolo novedoso para preparar bibliotecas de secuenciación que inesperadamente mejora la calidad del ADN de la biblioteca mientras que acelera el proceso de análisis de muestras para diagnósticos prenatales.

[0010] En una realización, se da a conocer en el presente documento un método para determinar una aneuploidía cromosómica fetal en una muestra de sangre materna que comprende una mezcla de moléculas de ácidos nucleicos fetales y maternos, en donde el método comprende: (a) preparar una biblioteca de secuenciación de la mezcla de moléculas de ácido nucleico maternas y fetales; en donde la preparación de dicha biblioteca comprende los pasos consecutivos de la reparación del extremo, la cola de dA y la unión del adaptador a dichos ácidos nucleicos; (b) secuenciar al menos una porción de las moléculas de ácido nucleico, obteniendo así información de secuencia para una pluralidad de moléculas de ácido nucleico maternas y fetales de una muestra de sangre materna; (c) usar la información de secuencia para obtener una dosis de cromosoma para un cromosoma aneuploide; y (d) comparar la dosis de cromosoma con al menos un valor umbral, y de ese modo identificar la presencia o ausencia de aneuploidía fetal.

[0011] En otra realización, se da a conocer en el presente documento un método para determinar una aneuploidía cromosómica fetal en una muestra de sangre materna que comprende una mezcla de moléculas de ácidos nucleicos fetales y maternos, en donde el método comprende: (a) preparar una biblioteca de secuenciación de la mezcla de moléculas de ácido nucleico maternas y fetales; en donde la preparación de dicha biblioteca comprende las etapas consecutivas de reparación de extremos, colas de dA y ligadura con adaptador de dichos ácidos nucleicos; (b) secuenciar al menos una porción de las moléculas de ácido nucleico, obteniendo así información de secuencia para una pluralidad de moléculas de ácido nucleico maternas y fetales de una muestra de sangre materna; (c) usar la información de secuencia para obtener una dosis de cromosoma para un cromosoma aneuploide; y (d) comparar la dosis de cromosoma con al menos un valor umbral, y de ese modo identificar la presencia o ausencia de aneuploidía fetal. El método comprende además usar la información de secuencia para identificar un número de etiquetas de secuencia mapeadas para al menos un cromosoma normalizador y para un cromosoma aneuploide; y el uso del número de etiquetas de secuencia mapeadas identificadas para dicho cromosoma aneuploide y el número de etiquetas de secuencia mapeadas identificadas para al menos un cromosoma normalizador para calcular una dosis de cromosoma para dicho cromosoma aneuploide como una proporción del número de etiquetas de secuencia mapeadas identificadas para dicho cromosoma aneuploide y el número de etiquetas de secuencia mapeadas identificadas para al menos un cromosoma normalizador. Opcionalmente, el cálculo de la dosis cromosómica comprende (i) calcular una relación de densidad de etiqueta de secuencia para el cromosoma aneuploide, relacionando el número de etiquetas de secuencia mapeadas identificadas para el cromosoma aneuploide en el paso con la longitud de dicho cromosoma aneuploide; (ii) calcular una relación de densidad de etiqueta de secuencia para al menos un cromosoma normalizador, relacionando el número de etiquetas de secuencia mapeadas identificadas para dicho al menos un cromosoma normalizador con la longitud de al menos un cromosoma normalizador; y (iii) usar las relaciones de densidad de etiquetas de secuencia calculadas en los pasos (i) y (ii) para calcular una dosis de cromosoma para el cromosoma aneuploide, en donde la dosis de cromosoma se calcula como la proporción de la relación de densidad de etiquetas de secuencia para el cromosoma aneuploide y la relación de densidad de la etiqueta de secuencia para al menos un cromosoma normalizador.

[0012] En otra realización, se da a conocer en el presente documento un método para determinar una aneuploidía cromosómica fetal en una muestra de sangre materna que comprende una mezcla de moléculas de ácidos nucleicos fetales y maternos, en donde el método comprende: (a) preparar una biblioteca de secuenciación de la mezcla de moléculas de ácido nucleico maternas y fetales; en donde la preparación de dicha biblioteca comprende las etapas consecutivas de reparación de extremos, colas de dA y ligadura con adaptador de dichos ácidos nucleicos; (b) secuenciar al menos una porción de las moléculas de ácido nucleico, obteniendo así información de secuencia para una pluralidad de moléculas de ácido nucleico maternas y fetales de una muestra de sangre materna; (c) usar la información de secuencia para obtener una dosis de cromosoma para un cromosoma aneuploide; y (d) comparar la dosis de cromosoma con al menos un valor umbral, y de ese modo identificar la presencia o ausencia de aneuploidía fetal. El método comprende además usar la información de secuencia para identificar un número de etiquetas de secuencia mapeadas para al menos un cromosoma normalizador y para un cromosoma aneuploide; y el uso del número de etiquetas de secuencia mapeadas identificadas para dicho cromosoma aneuploide y el número de etiquetas de secuencia mapeadas identificadas para al menos un cromosoma normalizador para calcular una dosis de cromosoma para dicho cromosoma aneuploide como una proporción del número de etiquetas de secuencia mapeadas identificadas para dicho cromosoma aneuploide y el número de etiquetas de secuencia mapeadas identificadas para al menos un cromosoma normalizador. Al menos un cromosoma normalizador es un cromosoma que tiene la variabilidad más pequeña y/o la mayor diferenciabilidad. Opcionalmente, el cálculo de la dosis cromosómica comprende (i) calcular una relación de densidad de etiqueta de secuencia para el cromosoma aneuploide, relacionando el número de etiquetas de secuencia mapeadas identificadas para el cromosoma aneuploide en el paso con la longitud de dicho cromosoma aneuploide; (ii) calcular una relación de densidad de etiqueta de secuencia para al menos un cromosoma normalizador, relacionando el número de etiquetas de secuencia mapeadas identificadas para dicho al menos un cromosoma normalizador con la longitud de al menos un cromosoma normalizador; y (iii) usar las relaciones de densidad de etiquetas de secuencia calculadas en los pasos (i) y (ii) para calcular una dosis de cromosoma para el cromosoma aneuploide, en donde la dosis de cromosoma se

calcula como la proporción de la relación de densidad de etiquetas de secuencia para el cromosoma aneuploide y la relación de densidad de la etiqueta de secuencia para al menos un cromosoma normalizador.

[0013] En otra realización, se da a conocer en el presente documento un método para determinar una aneuploidía cromosómica fetal en una muestra de sangre materna comprende una mezcla de moléculas de ácidos nucleicos fetales y maternos, en donde el método comprende: (a) preparar una biblioteca de secuenciación de la mezcla de moléculas de ácido nucleico maternas y fetales; en donde la preparación de dicha biblioteca comprende las etapas consecutivas de reparación de extremos, colas de dA y ligadura con adaptador de dichos ácidos nucleicos; (b) secuenciar al menos una porción de las moléculas de ácido nucleico, obteniendo así información de secuencia para una pluralidad de moléculas de ácido nucleico maternas y fetales de una muestra de sangre materna; (c) usar la información de secuencia para obtener una dosis de cromosoma para un cromosoma aneuploide; y (d) comparar la dosis de cromosoma con al menos un valor umbral, y de ese modo identificar la presencia o ausencia de aneuploidía fetal. El método comprende además usar la información de secuencia para identificar un número de etiquetas de secuencia mapeadas para al menos un cromosoma normalizador y para un cromosoma aneuploide; y usando el número de etiquetas de secuencia mapeadas identificadas para dichos cromosomas aneuploides y el número de etiquetas de secuencia mapeadas identificadas para al menos un cromosoma normalizado para calcular una dosis de cromosoma para dicho cromosoma aneuploide como una proporción de la cantidad de etiquetas de secuencia mapeadas identificadas para dicho cromosoma aneuploide y el número de etiquetas de secuencia mapeadas identificadas para al menos un cromosoma normalizante. Opcionalmente, el cálculo de la dosis cromosómica comprende (i) calcular una relación de densidad de etiqueta de secuencia para el cromosoma aneuploide, relacionando el número de etiquetas de secuencia mapeadas identificadas para el cromosoma aneuploide en el paso con la longitud de dicho cromosoma aneuploide; (ii) calcular una relación de densidad de etiqueta de secuencia para al menos un cromosoma normalizador, relacionando el número de etiquetas de secuencia mapeadas identificadas para dicho al menos un cromosoma normalizador con la longitud de al menos un cromosoma normalizador; y (iii) usar las relaciones de densidad de etiquetas de secuencia calculadas en los pasos (i) y (ii) para calcular una dosis de cromosoma para el cromosoma aneuploide, en donde la dosis de cromosoma se calcula como la proporción de la relación de densidad de etiquetas de secuencia para el cromosoma aneuploide y la relación de densidad de la etiqueta de secuencia para al menos un cromosoma normalizador. En realizaciones en las que el cromosoma aneuploide es el cromosoma 21, al menos un cromosoma normalizante se selecciona entre el cromosoma 9, el cromosoma 1, el cromosoma 2, el cromosoma 11, el cromosoma 12 y el cromosoma 14. Alternativamente, el al menos un cromosoma normalizador para el cromosoma 21 es un grupo de cromosomas seleccionados entre el cromosoma 9, el cromosoma 1, el cromosoma 2, el cromosoma 11, el cromosoma 12 y el cromosoma 14. En realizaciones en las que el cromosoma aneuploide es el cromosoma 18, el al menos un cromosoma normalizador se selecciona del cromosoma 8, cromosoma 2, cromosoma 3, cromosoma 5, cromosoma 6, cromosoma 12 y cromosoma 14. Alternativamente, el al menos un cromosoma normalizador para el cromosoma 18 es un grupo de cromosomas seleccionados del cromosoma 8, cromosoma 2, cromosoma 3, cromosoma 5, cromosoma 6, el cromosoma 12 y el cromosoma 14. En realizaciones cuando el cromosoma aneuploide es el cromosoma 13, al menos un cromosoma normalizante se selecciona del cromosoma 2, cromosoma 3, cromosoma 4, cromosoma 5, cromosoma 6 y cromosoma 8. Alternativamente, el al menos un cromosoma normalizador para el cromosoma 13 es un grupo de cromosomas seleccionados del cromosoma 2, cromosoma 3, cromosoma 4, cromosoma 5, cromosoma 6 y cromosoma 8. En realizaciones en las que el cromosoma aneuploide es el cromosoma X, al menos un cromosoma normalizante se selecciona entre el cromosoma 2, el cromosoma 3, el cromosoma 4, el cromosoma 5, el cromosoma 6 y el cromosoma 8. Alternativamente, el al menos un cromosoma normalizador para el cromosoma X es un grupo de cromosomas seleccionados entre el cromosoma 2, el cromosoma 3, el cromosoma 4, el cromosoma 5, el cromosoma 6 y el cromosoma 8.

[0014] La muestra materna utilizada en las realizaciones del método para determinar una aneuploidía cromosómica fetal es un fluido biológico seleccionado a partir de sangre, plasma, suero, orina y saliva. Preferiblemente, la muestra materna es una muestra de plasma. En algunas realizaciones, las moléculas de ácido nucleico comprendidas en la muestra materna son moléculas de ADN libres de células. En algunas realizaciones, los pasos consecutivos comprendidos en la preparación de la biblioteca de secuenciación se realizan en menos de una hora. Preferiblemente, los pasos consecutivos se realizan en ausencia de polietilenglicol. Más preferiblemente, los pasos consecutivos excluyen la purificación. La secuenciación de la biblioteca de secuenciación se realiza mediante los métodos de secuenciación de la próxima generación (NGS). En algunas realizaciones, la secuenciación comprende una amplificación. En otras realizaciones, la secuenciación es una secuenciación masivamente paralela que utiliza la secuenciación por síntesis con terminadores de tinte reversibles. En otras realizaciones, la secuenciación es una secuenciación masivamente paralela utilizando la secuenciación por ligadura. En otras realizaciones más, la secuenciación es la secuenciación de una sola molécula.

[0015] En otra realización, se da a conocer en el presente documento un método para determinar la presencia o ausencia de una aneuploidía en una muestra materna que comprende una mezcla de moléculas de ácido nucleico fetal y materno, en donde el método comprende: (a) preparar una biblioteca de secuenciación a partir de la mezcla; en donde la preparación de dicha biblioteca comprende las etapas consecutivas de reparación de extremos, colas de dA y ligadura con adaptador de dichos ácidos nucleicos maternos y fetales; (b) secuenciar al menos una parte de la biblioteca de secuenciación, en la que la secuencia comprende proporcionar una pluralidad de etiquetas de secuencia; y (c) con base en la secuenciación, determinando la presencia o ausencia de aneuploidía en la muestra.

[0016] En otra realización, se da a conocer en el presente documento un método para determinar la presencia o ausencia de una aneuploidía cromosómica o parcial en una muestra materna que comprende una mezcla de ácidos nucleicos fetales y maternos, en donde el método comprende: (a) preparar una biblioteca de secuenciación de la mezcla; en donde la preparación de dicha biblioteca comprende las etapas consecutivas de reparación de extremos, colas de dA y ligadura con adaptador de dichos ácidos nucleicos maternos y fetales; (b) secuenciar al menos una parte de la biblioteca de secuenciación, en la que la secuencia comprende proporcionar una pluralidad de etiquetas de secuencia; y (c) en función de la secuenciación, la determinación de la presencia o ausencia de la aneuploidía cromosómica o parcial en la muestra.

[0017] En otra realización, se da a conocer en el presente documento un método para determinar la presencia o ausencia de una aneuploidía cromosómica en una muestra materna que comprende una mezcla de ácidos nucleicos fetales y maternos, en donde el método comprende: (a) preparar una biblioteca de secuenciación a partir de la mezcla; en donde la preparación de dicha biblioteca comprende las etapas consecutivas de reparación de extremos, colas de dA y ligadura con adaptador de dichos ácidos nucleicos maternos y fetales; (b) secuenciar al menos una parte de la biblioteca de secuenciación, en la que la secuencia comprende proporcionar una pluralidad de etiquetas de secuencia; y (c) con base en la secuenciación, determinando la presencia o ausencia de la aneuploidía cromosómica en la muestra. Las aneuploidías cromosómicas que pueden determinarse de acuerdo con el método incluyen trisomía 8, trisomía 13, trisomía 15, trisomía 16, trisomía 18, trisomía 21, trisomía 22, monosomía X y XXX.

[0018] En otra realización, se da a conocer en el presente documento un método para determinar la presencia o ausencia de una aneuploidía cromosómica o parcial en una muestra materna que comprende una mezcla de ácidos nucleicos fetales y maternos, en donde el método comprende: (a) preparar una biblioteca de secuenciación de la mezcla; en donde la preparación de dicha biblioteca comprende las etapas consecutivas de reparación de extremos, colas de dA y ligadura con adaptador de dichos ácidos nucleicos maternos y fetales; (b) secuenciar al menos una parte de la biblioteca de secuenciación, en la que la secuencia comprende proporcionar una pluralidad de etiquetas de secuencia; y (c) en función de la secuenciación, la determinación de la presencia o ausencia del cromosoma o una aneuploidía parcial en la muestra que comprende calcular una dosis de cromosoma en función del número de dichas etiquetas de secuencia para un cromosoma de interés y para un cromosoma de normalización, y comparar dicha dosis a un valor umbral.

[0019] En otra realización, se da a conocer en el presente documento un método para determinar la presencia o ausencia de una aneuploidía cromosómica en una muestra materna que comprende una mezcla de ácidos nucleicos fetales y maternos, en donde el método comprende: (a) preparar una biblioteca de secuenciación a partir la mezcla; en donde la preparación de dicha biblioteca comprende las etapas consecutivas de reparación de extremos, colas de dA y ligadura con adaptador de dichos ácidos nucleicos maternos y fetales; (b) secuenciar al menos una parte de la biblioteca de secuenciación, en la que la secuencia comprende proporcionar una pluralidad de etiquetas de secuencia; y (c) basándose en la secuenciación, determinar la presencia o ausencia de la aneuploidía cromosómica en la muestra que comprende calcular una dosis cromosómica basada en el número de dichas etiquetas de secuencia para un cromosoma de interés y para un cromosoma normalizador, y comparar dicha dosis con un valor de umbral. Las aneuploidías cromosómicas que pueden determinarse de acuerdo con el método incluyen trisomía 8, trisomía 13, trisomía 15, trisomía 16, trisomía 18, trisomía 21, trisomía 22, monosomía X y XXX.

[0020] La muestra maternal utilizada en las realizaciones del método para determinar la presencia o ausencia de una aneuploidía es un fluido biológico seleccionado a partir de sangre, plasma, suero, orina y saliva. Preferiblemente, la muestra materna es una muestra de plasma. En algunas realizaciones, las moléculas de ácido nucleico comprendidas en la muestra materna son moléculas de ADN libres de células. En algunas realizaciones, los pasos consecutivos comprendidos en la preparación de la biblioteca de secuenciación se realizan en menos de una hora. Preferiblemente, los pasos consecutivos se realizan en ausencia de polietilenglicol. Más preferiblemente, los pasos consecutivos excluyen la purificación. La secuenciación de la biblioteca de secuenciación se realiza mediante los métodos de secuenciación de la próxima generación (NGS). En algunas realizaciones, la secuenciación comprende una amplificación. En otras realizaciones, la secuenciación es una secuenciación masivamente paralela que usa la secuenciación por síntesis con terminadores de tinte reversibles. En otras realizaciones, la secuenciación es una secuenciación masivamente paralela utilizando la secuenciación por ligadura. En otras realizaciones más, la secuenciación es la secuenciación de una sola molécula.

[0021] En otra realización, se da a conocer en el presente documento un método para determinar la fracción de moléculas de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de moléculas de ácido nucleico fetal y materno, en donde el método comprende: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en una porción de la mezcla; (b) preparar una biblioteca de secuenciación del producto amplificado obtenido en la etapa (a), en la que la preparación de la biblioteca comprende las etapas consecutivas de reparación final, cola de dA y adaptador de ligadura de dichas moléculas de ácido nucleico maternas y fetales; (c) secuenciar al menos una parte de la biblioteca de secuenciación; y (d) basándose en dicha secuenciación, determinando la fracción de las moléculas de ácido nucleico fetal. Opcionalmente, el método puede comprender además determinar la presencia o ausencia de aneuploidía en la muestra materna.

[0022] En otra realización, se da a conocer en el presente documento un método para determinar la fracción de moléculas de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de moléculas de ácido nucleico fetal y materno, en donde el método comprende: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en una porción de la mezcla; (b) preparar una biblioteca de secuenciación del producto amplificado obtenido en la etapa (a), en donde la preparación de la biblioteca comprende las etapas consecutivas de reparación final, cola dA y adaptador de ligadura de dichas moléculas de ácido nucleico maternas y fetales; (c) secuenciar al menos una parte de la biblioteca de secuenciación; y (d) basándose en dicha secuenciación, determinando la fracción de las moléculas de ácido nucleico fetal. La determinación de la fracción comprende determinar el número de etiquetas de secuencia fetal y materna mapeadas a un genoma diana de referencia que comprende al menos un ácido nucleico polimórfico. Opcionalmente, el método puede comprender además determinar la presencia o ausencia de aneuploidía en la muestra materna.

[0023] En otra realización, se da a conocer en el presente documento un método para determinar la fracción de moléculas de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de moléculas de ácido nucleico fetal y materno, en donde el método comprende: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en una porción de la mezcla, en donde cada uno de dicha pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de nucleótido único (SNP); (b) preparar una biblioteca de secuenciación del producto amplificado obtenido en la etapa (a), en la que la preparación de la biblioteca comprende las etapas consecutivas de reparación de extremos, colas de dA y ligadura con adaptador de dichas moléculas de ácido nucleico maternas y fetales; (c) secuenciar al menos una parte de la biblioteca de secuenciación; y (d) basándose en dicha secuenciación, determinando la fracción de las moléculas de ácido nucleico fetal. Opcionalmente, el método puede comprender además determinar la presencia o ausencia de aneuploidía en la muestra materna.

[0024] En otra realización, se da a conocer en el presente documento un método para determinar la fracción de moléculas de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de moléculas de ácido nucleico fetal y materno, en donde el método comprende: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en una porción de la mezcla, en donde cada uno de dicha pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de nucleótido único (SNP); (b) preparar una biblioteca de secuenciación del producto amplificado obtenido en la etapa (a) en la que la preparación de la biblioteca comprende las etapas consecutivas de reparación final, cola dA y adaptador de ligadura de dichas moléculas de ácido nucleico maternas y fetales; (c) secuenciar al menos una parte de la biblioteca de secuenciación; y (d) basándose en dicha secuenciación, determinando la fracción de las moléculas de ácido nucleico fetal. La determinación de la fracción comprende determinar el número de etiquetas de secuencia fetal y materna mapeadas a un genoma diana de referencia que comprende al menos un ácido nucleico polimórfico. Opcionalmente, el método puede comprender además determinar la presencia o ausencia de aneuploidía en la muestra materna.

[0025] En otra realización, se da a conocer en el presente documento un método para determinar la fracción de moléculas de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de moléculas de ácido nucleico fetal y materno, en donde el método comprende: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en una porción de la mezcla, en donde cada uno de dicha pluralidad de ácido nucleico diana polimórfico comprende al menos una repetición corta en tándem (STR); (b) preparar una biblioteca de secuenciación del producto amplificado obtenido en la etapa (a), en la que la preparación de la biblioteca comprende las etapas consecutivas de reparación final, cola dA y adaptador de ligadura de dichas moléculas de ácido nucleico maternas y fetales; (c) secuenciar al menos una parte de la biblioteca de secuenciación; y (d) basándose en dicha secuenciación, determinando la fracción de las moléculas de ácido nucleico fetal. Opcionalmente, el método puede comprender además determinar la presencia o ausencia de aneuploidía en la muestra materna.

[0026] En otra realización, se da a conocer en el presente documento un método para determinar la fracción de moléculas de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de moléculas de ácido nucleico fetal y materno, en donde el método comprende: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en una porción de la mezcla, en donde cada uno de dicha pluralidad de ácido nucleico diana polimórfico comprende al menos una repetición corta en tándem (STR); (b) preparar una biblioteca de secuenciación del producto amplificado obtenido en la etapa (a), en la que la preparación de la biblioteca comprende las etapas consecutivas de reparación final, cola dA y adaptador de ligadura de dichas moléculas de ácido nucleico maternas y fetales; (c) secuenciar al menos una parte de la biblioteca de secuenciación; y (d) basándose en dicha secuenciación, determinando la fracción de las moléculas de ácido nucleico fetal. La determinación de la fracción comprende determinar el número de etiquetas de secuencia fetal y materna mapeadas a un genoma diana de referencia que comprende al menos un ácido nucleico polimórfico. Opcionalmente, el método puede comprender además determinar la presencia o ausencia de aneuploidía en la muestra materna.

[0027] En otra realización, se da a conocer en el presente documento un método para determinar la fracción de moléculas de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de moléculas de ácido nucleico fetal y materno, en donde el método comprende: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en una porción de la mezcla, en donde cada uno de la pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de nucleótido (SNP); (b) preparar una biblioteca de secuenciación del

producto amplificado obtenido en la etapa (a), en la que la preparación de la biblioteca comprende las etapas consecutivas de reparación final, cola dA y adaptador de ligadura de dichas moléculas de ácido nucleico maternas y fetales; (c) secuenciar al menos una parte de la biblioteca de secuenciación; y (d) basándose en dicha secuenciación, determinando la fracción de las moléculas de ácido nucleico fetal. En realizaciones en las que cada uno de la pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un solo nucleótido (SNP), el SNP se selecciona de entre 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 realizaciones en las que cada uno de la pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de nucleótido (SNP), el al menos un SNP es un SNP en tándem seleccionado de pares de SNP en tándem rs7277033-rs2110153; rs2822654-rs1882882; rs368657-rs376635; rs2822731-rs2822732; rs1475881-rs7275487; rs1735976-rs2827016; rs447340-rs2824097; rs418989-rs13047336; rs987980-rs987981; rs4143392-rs4143391; rs1691324-rs13050434; rs11909758-rs9980111; rs2826842-rs232414; rs1980969-rs1980970; rs9978999-rs9979175; rs1034346-rs12481852; rs7509629-rs2828358; rs4817013-rs7277036; rs9981121-rs2829696; rs455921-rs2898102; rs2898102-rs458848; rs961301-rs2830208; rs2174536-rs458076; rs11088023-rs11088024; rs1011734-rs1011733; rs2831244-rs9789838; rs8132769-rs2831440; rs8134080-rs2831524; rs4817219-rs4817220; rs2250911-rs2250997; rs2831899-rs2831900; rs2831902-rs2831903; rs11088086-rs2251447; rs2832040-rs11088088; rs2832141-rs2246777; rs2832959-rs9980934; rs2833734-rs2833735; rs933121-rs933122; rs2834140-rs12626953; rs2834485-rs3453; rs9974986-rs2834703; rs2776266-rs2835001; rs1984014-rs1984015; rs7281674-rs2835316; rs13047304-rs13047322; rs2835545-rs4816551; rs2835735-rs2835736; rs13047608-rs2835826; rs2836550-rs2212596; rs2836660-rs2836661; rs465612-rs8131220; rs9980072-rs8130031; rs418359-rs2836926; rs7278447-rs7278858; rs385787-rs367001; rs367001-rs386095; rs2837296-rs2837297; y rs2837381-rs4816672. Opcionalmente, el método puede comprender además determinar la presencia o ausencia de aneuploidía en la muestra materna.

[0028] En otra realización, se da a conocer en el presente documento un método para determinar la fracción de moléculas de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de moléculas de ácido nucleico fetal y materno, en donde el método comprende: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en una porción de la mezcla, en donde cada uno de dicha pluralidad de ácido nucleico diana polimórfico comprende al menos una repetición corta en tándem (STR); (b) preparar una biblioteca de secuenciación del producto amplificado obtenido en la etapa (a), en la que la preparación de la biblioteca comprende las etapas consecutivas de reparación de extremos, colas de dA y ligadura con adaptador de dichas moléculas de ácido nucleico maternas y fetales; (c) secuenciar al menos una parte de la biblioteca de secuenciación; y (d) basado en dicha secuenciación, determinando la fracción del ácido nucleico fetal y moléculas. El al menos un STR se selecciona entre CSF1PO, FGA, TH01, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, D2S1338, Penta D, Penta E, D22S1045, D20S1082, D20S482, D18S853, D17S1301, D17S974, D14S1434, D12ATA63, D11S4463, D10S1435, D10S1248, D9S2157, D9S1122, D8S1115, D6S1017, D6S474, D5S2500, D4S2408, D4S2364, D3S4529, D3S3053, D2S1776, D2S441, D1S1677, D1S1627, y D1GATA113. Opcionalmente, el método puede comprender además determinar la presencia o ausencia de aneuploidía en la muestra materna.

[0029] En otra realización, se da a conocer en el presente documento un método para determinar la fracción de moléculas de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de moléculas de ácido nucleico fetal y materno, en donde el método comprende: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en una porción de la mezcla, en donde cada uno de la pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de nucleótido (SNP); (b) preparar una biblioteca de secuenciación del producto amplificado obtenido en la etapa (a), en la que la preparación de la biblioteca comprende las etapas consecutivas de reparación de extremos, colas de dA y ligadura con adaptador de dichas moléculas de ácido nucleico maternas y fetales; (c) secuenciar al menos una parte de la biblioteca de secuenciación; y (d) basándose en dicha secuenciación, determinando la fracción de las moléculas de ácido nucleico fetal. La determinación de la fracción comprende determinar el número de etiquetas de secuencia fetal y materna mapeadas a un genoma diana de referencia que comprende al menos un ácido nucleico polimórfico. En realizaciones en las que cada uno de la pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un solo nucleótido (SNP), el SNP se selecciona de entre 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 realizaciones en las que cada uno de la pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de nucleótido (SNP), el al menos un SNP es un SNP en tándem seleccionado de pares de SNP en tándem rs7277033-rs2110153; rs2822654-rs1882882; rs368657-rs376635; rs2822731-rs2822732; rs1475881-rs7275487; rs1735976-rs2827016; rs447340-rs2824097; rs418989-rs13047336; rs987980-rs987981; rs4143392-rs4143391; rs1691324-rs13050434; rs11909758-rs9980111; rs2826842-rs232414; rs1980969-rs1980970; rs9978999-rs9979175; rs1034346-rs12481852; rs7509629-rs2828358; rs4817013-rs7277036; rs9981121-rs2829696; rs455921-rs2898102; rs2898102-rs458848; rs961301-rs2830208; rs2174536-rs458076; rs11088023-rs11088024; rs1011734-rs1011733; rs2831244-rs9789838; rs8132769-rs2831440; rs8134080-rs2831524; rs4817219-rs4817220; rs2250911-rs2250997; rs2831899-rs2831900; rs2831902-rs2831903; rs11088086-rs2251447; rs2832040-rs11088088; rs2832141-rs2246777; rs2832959-rs9980934; rs2833734-

rs2833735; rs933121-rs933122; rs2834140-rs12626953; rs2834485-rs3453; rs9974986- rs2834703; rs2776266-rs2835001; rs1984014-rs1984015; rs7281674-rs2835316; rs13047304-rs13047322; rs2835545-rs4816551; rs2835735-rs2835736; rs13047608-rs2835826; rs2836550-rs2212596; rs2836660-rs2836661; rs465612-rs8131220; rs9980072-rs8130031; rs418359-rs2836926; rs7278447-rs7278858; rs385787-rs367001; rs367001-rs386095; rs2837296-rs2837297; y rs2837381-rs4816672. Opcionalmente, el método puede comprender además determinar la presencia o ausencia de aneuploidía en la muestra materna.

[0030] En otra realización, se da a conocer en el presente documento un método para determinar la fracción de moléculas de ácidos nucleicos fetales en una muestra materna que comprende una mezcla de moléculas de ácido nucleico fetal y materno, en donde el método comprende: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en una porción de la mezcla, en donde cada uno de dicha pluralidad de ácido nucleico diana polimórfico comprende al menos una repetición corta en tándem (STR); (b) preparar una biblioteca de secuenciación del producto amplificado obtenido en la etapa (a), en la que la preparación de la biblioteca comprende las etapas consecutivas de reparación de extremos, colas de dA y ligadura con adaptador de dichas moléculas de ácido nucleico maternas y fetales; (c) secuenciar al menos una parte de la biblioteca de secuenciación; y (d) basándose en dicha secuenciación, determinando la fracción de las moléculas de ácido nucleico fetal. La determinación de la fracción comprende determinar el número de etiquetas de secuencia fetal y materna mapeadas a un genoma diana de referencia que comprende al menos un ácido nucleico polimórfico. El al menos un STR se selecciona entre CSF1PO, FGA, TH01, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, D2S1338, Penta D, Penta E, D22S1045, D20S1082, D20S482, D18S853, D17S1301, D17S974, D14S1434, D12ATA63, D11S4463, D10S1435, D10S1248, D9S2157, D9S1122, D8S1115, D6S1017, D6S474, D5S2500, D4S2408, D4S2364, D3S4529, D3S3053, D2S1776, D2S441, D1S1677, D1S1627, and D1GATA113. Opcionalmente, el método puede comprender además determinar la presencia o ausencia de aneuploidía en la muestra materna.

[0031] La muestra materna utilizada en las realizaciones del método para determinar la fracción de moléculas de ácido nucleico fetal, es un fluido biológico seleccionado de sangre, plasma, suero, orina y saliva. Preferiblemente, la muestra materna es una muestra de plasma. En algunas realizaciones, las moléculas de ácido nucleico comprendidas en la muestra materna son moléculas de ADN libres de células. En algunas realizaciones, los pasos consecutivos comprendidos en la preparación de la biblioteca de secuenciación se realizan en menos de una hora. Preferiblemente, los pasos consecutivos se realizan en ausencia de polietilenglicol. Más preferiblemente, los pasos consecutivos excluyen la purificación. La secuenciación de la biblioteca de secuenciación se realiza mediante los métodos de secuenciación de la próxima generación (NGS). En algunas realizaciones, la secuenciación comprende una amplificación. En otras realizaciones, la secuenciación es una secuenciación masivamente paralela que usa la secuenciación por síntesis con terminadores de tinte reversibles. En otras realizaciones, la secuenciación es una secuenciación masivamente paralela utilizando la secuenciación por ligadura. En otras realizaciones más, la secuenciación es la secuenciación de una sola molécula.

[0032] En otra realización, se da a conocer en el presente documento un medio legible por ordenador que tiene almacenado en el mismo instrucciones legibles por computadora para llevar a cabo el método para determinar la presencia o ausencia de una aneuploidía por ejemplo una aneuploidía cromosómica fetal, en una muestra de sangre materna que comprende una mezcla de moléculas de ácidos nucleicos fetales y maternos.

[0033] En una realización, el medio legible por ordenador que ha almacenado en el mismo instrucciones legibles por ordenador para llevar a cabo el método comprende las etapas de (a) utilizar información de la secuencia obtenida a partir de una pluralidad de moléculas de ácido nucleico fetal y materno en una muestra de plasma materno a identificar una serie de etiquetas de secuencia mapeadas para un cromosoma de interés; (b) usar información de secuencia obtenida de una pluralidad de moléculas de ácido nucleico materno y fetal en una muestra de plasma materno para identificar un número de etiquetas de secuencia mapeadas para al menos un cromosoma normalizador; (c) usar el número de etiquetas de secuencia mapeadas identificadas para un cromosoma de interés en el paso (a) y el número de etiquetas de secuencia mapeadas identificadas para el al menos un cromosoma normalizador en el paso (b) para calcular una dosis de cromosoma para dicho cromosoma de interés; y (d) comparar dicha dosis de cromosoma con al menos un valor umbral, y de ese modo identificar la presencia o ausencia de aneuploidía fetal. Los cromosomas de interés pueden ser cualquiera de los cromosomas 21, 13, 18 y X.

[0034] En otra realización, se da a conocer en el presente documento un sistema de procesamiento de ordenador que se adapta o configura para realizar el método para determinar la presencia o ausencia de una aneuploidía por ejemplo una aneuploidía cromosómica fetal, en una muestra de sangre materna comprende una mezcla de fetal y moléculas de ácidos nucleicos maternos.

[0035] En una realización, el sistema de procesamiento de ordenador está adaptado o configurado para realizar los pasos siguientes: (a) el uso de información de la secuencia obtenida a partir de una pluralidad de moléculas de ácido nucleico fetal y materno en una muestra de plasma materno para identificar un número de mapeado etiquetas de secuencia para un cromosoma de interés; (b) usar información de secuencia obtenida de una pluralidad de moléculas de ácido nucleico materno y fetal en una muestra de plasma materno para identificar un número de etiquetas de secuencia mapeadas para al menos un cromosoma normalizador; (c) usar el número de etiquetas de

secuencia mapeadas identificadas para un cromosoma de interés en el paso (a) y el número de etiquetas de secuencia mapeadas identificadas para al menos un cromosoma normalizador en el paso (b) para calcular una dosis de cromosoma para un cromosoma de interés; y (d) comparar dicha dosis de cromosoma con al menos un valor umbral, y de ese modo identificar la presencia o ausencia de aneuploidía fetal. Los cromosomas de interés pueden ser cualquiera de los cromosomas 21, 13, 18 y X.

[0036] En otra realización, se da a conocer en el presente documento un aparato adaptado o configurado para identificar la aneuploidía fetal en una muestra de plasma materno que comprende moléculas de ácido nucleico fetal y materno, y en donde dicho aparato comprende: (a) un dispositivo de secuenciación adaptado o configurado para secuenciar al menos una porción de las moléculas de ácido nucleico en una muestra de plasma materno que comprende moléculas de ácido nucleico materno y fetal, generando así información de secuencia; y (b) un sistema de procesamiento computacional configurado para realizar los siguientes pasos: (i) usar la información de secuencia generada por el dispositivo de secuenciación para identificar un número de etiquetas de secuencia mapeadas para un cromosoma de interés; (ii) usar la información de secuencia generada por el dispositivo de secuenciación para identificar un número de etiquetas de secuencia mapeadas para al menos un cromosoma normalizador; (iii) usar el número de etiquetas de secuencia mapeadas identificadas para un cromosoma de interés en el paso (i) y el número de etiquetas de secuencia mapeadas identificadas para el al menos un cromosoma normalizador en el paso (ii) para calcular una dosis de cromosoma para un cromosoma de interés; y (iv) comparar dicha dosis de cromosoma con al menos un valor umbral, y de ese modo identificar la presencia o ausencia de aneuploidía fetal. Los cromosomas de interés pueden ser cualquiera de los cromosomas 21, 13, 18 y X.

[0037] Aunque los ejemplos en el presente documento se refieren a seres humanos y la lengua se dirige principalmente a las preocupaciones humanas, el concepto de esta invención es aplicable a genomas de cualquier planta o animal.

4. BREVE DESCRIPCIÓN DE LOS DIBUJOS

[0038]

La Figura 1 es un diagrama de flujo de un método **100** para determinar la presencia o ausencia de una aneuploidía cromosómica en una muestra de prueba que comprende una mezcla de ácidos nucleicos.

La Figura 2 es un diagrama de flujo de un método **200** para determinar simultáneamente la presencia o ausencia de aneuploidía y la fracción fetal en una muestra de prueba materna que comprende una mezcla de ácidos nucleicos fetales y maternos.

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 prueba de plasma materno 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 muestra de prueba de ADNcf purificada materna que ha sido enriquecida con ácidos nucleicos polimórficos.

La Figura 5 es un diagrama de flujo de un método **500** para determinar simultáneamente la presencia o ausencia de aneuploidía fetal y la fracción fetal en una biblioteca de secuenciación construida a partir de ácidos nucleicos maternos y fetales derivados de una muestra de prueba materna y enriquecida con ácidos nucleicos polimórficos.

La Figura 6 es un diagrama de flujo de un método **600** para determinar la fracción fetal mediante la secuenciación de una biblioteca de ácidos nucleicos diana polimórficos amplificados a partir de una porción de una mezcla purificada de ácidos nucleicos maternos y fetales.

La Figura 7 muestra los electroferogramas de una biblioteca de secuenciación de ADNcf preparada de acuerdo con el protocolo abreviado descrito en el Ejemplo 2a (**A**), y el protocolo descrito en el Ejemplo 2b (**B**).

La Figura 8 muestra en el eje Y la proporción del número de etiquetas de secuencia asignadas a cada cromosoma (eje X) y el número total de etiquetas asignadas a todos los cromosomas (1-22, X e Y) para la muestra M11281 cuando la biblioteca se preparó utilizando el protocolo abreviado del Ejemplo 2a (**♦**) y cuando se preparó de acuerdo con el protocolo de longitud completa del Ejemplo 2b (**■**). También se muestran las relaciones de las etiquetas para la muestra M11297 obtenidas de la secuenciación de una biblioteca preparada de acuerdo con el protocolo abreviado del Ejemplo 2a (**A**) y de acuerdo con el protocolo de longitud completa del Ejemplo 2b (**X**).

La Figura 9 muestra la distribución de la dosis de cromosoma para el cromosoma 21 determinada a partir de la secuenciación del ADNcf extraído de un conjunto de 48 muestras de sangre obtenidas de sujetos humanos, cada una de ellas con un feto masculino o femenino. Dosis de cromosoma 21 para muestras de ensayo calificadas es decir, normales para el cromosoma 21 (**O**), y la trisomía 21 se muestran (**Δ**) para los cromosomas 1-12 y X (**A**), y para los cromosomas 1-22 y X (**B**).

La Figura 10 muestra la distribución de la dosis de cromosoma para el cromosoma 18 determinada a partir de la secuenciación del ADNcf extraído de un conjunto de 48 muestras de sangre obtenidas de sujetos humanos, cada uno de ellos con un feto masculino o femenino. Dosis de cromosoma 18 para muestras de ensayo calificadas, es decir, normales para el cromosoma 18 (**O**), y la trisomía 18 (**Δ**) se muestran para los cromosomas 1-12 y X (**A**), y para los cromosomas 1-22 y X (**B**).

La Figura 11 muestra la distribución de la dosis de cromosoma para el cromosoma 13 determinada a partir de la secuenciación del ADNfc extraído de un conjunto de 48 muestras de sangre obtenidas de sujetos humanos, cada una de ellas con un feto masculino o femenino. Dosis de cromosoma 13 para muestras de ensayo cualificadas, es decir, normales para el cromosoma 13 (O), y la trisomía 13 (Δ) se muestran para los cromosomas 1-12 y X (A), y para los cromosomas 13- 21 y Y (B).

La Figura 12 muestra la distribución de las dosis de cromosomas para el cromosoma X determinada a partir de la secuenciación del ADNfc extraído de un conjunto de 48 muestras de sangre analizadas de sujetos humanos, cada una de ellas con un feto masculino o femenino. Cromosoma X dosis para hombres (46, XY; (O)), mujeres (46, XX; (Δ)); se muestran muestras de monosomía X (45, X; (+)) y cariotipos complejos (Cplx (X)) para los cromosomas 1-12 y X (A) y para los cromosomas 1-22 y X (B). **La Figura 13** muestra la distribución de las dosis de cromosomas para el cromosoma Y determinada a partir de la secuenciación del ADNfc extraído de un conjunto de 48 muestras de sangre de prueba obtenidas de sujetos humanos, cada una de ellas con un feto masculino o femenino. Dosis de cromosoma Y para hombres (46, XY; (Δ)), mujeres (46, XX; (O)); se muestran muestras de monosomía X (45, X; (+)) y cariotipos complejos (Cplx (X)) para los cromosomas 1-12 (A) y para los cromosomas 1-22 (B).

La Figura 14 muestra el coeficiente de variación (CV) para los cromosomas 21 ($\blacksquare$), 18 ($\bullet$) y 13 ($\blacktriangle$) que se determinó a partir de las dosis mostradas en las Figuras 9, 10 y 11, respectivamente.

La Figura 15 muestra el coeficiente de variación (CV) para los cromosomas X ($\blacksquare$) e Y ($\bullet$) que se determinó a partir de las dosis mostradas en las Figuras 12 y 13, respectivamente.

La Figura 16 muestra la secuencia de dosis (eje Y) para un segmento del cromosoma 11 (81000082-103000103bp) determinada a partir de la secuenciación del ADNfc extraído de un conjunto de 7 muestras calificadas (O) obtenidas y 1 muestra de prueba ($\blacklozenge$) de mujeres embarazadas. Se identificó una muestra de un sujeto que llevaba un feto con una aneuploidía parcial del cromosoma 11 ($\blacklozenge$).

La Figura 17 muestra un gráfico de la proporción del número de etiquetas de secuencia asignadas a cada cromosoma y el número total de etiquetas asignadas a todos los cromosomas (1-22, X e Y) obtenidos de la secuenciación de una biblioteca de ADNfc no enriquecida ($\bullet$), y biblioteca ADNfc enriquecida con 5% ($\blacksquare$) o 10% ($\blacklozenge$) de biblioteca SNP multiplex amplificada.

La Figura 18 muestra un diagrama de barras que muestra la identificación de las secuencias polimórficas fetales y maternas (SNP) utilizadas para determinar la fracción fetal en una muestra de prueba. Se muestra el número total de lecturas de secuencia (eje Y) mapeadas a las secuencias SNP identificadas por números rs (eje X), y el nivel relativo de ácidos nucleicos fetales (*).

La Figura 19 representa una realización del uso de la fracción fetal para determinar los umbrales de corte para la detección de aneuploidía.

La Figura 20 ilustra la distribución de las dosis de cromosomas normalizados para el cromosoma 21 (A), el cromosoma 18 (B), el cromosoma 13 (C), el cromosoma X (D) y el cromosoma Y (E) en relación con la desviación estándar de la media (eje Y) para la dosis de cromosomas correspondiente en muestras no afectadas.

5. DESCRIPCIÓN DETALLADA DE LA INVENCION

[0039] La descripción se refiere a métodos para determinar la presencia o ausencia de una aneuploidía por ejemplo aneuploidía cromosómica o parcial, y/o fracción fetal en muestras maternas que comprenden ácidos nucleicos fetales y maternos por secuenciación masiva en paralelo. El método comprende un nuevo protocolo para preparar bibliotecas de secuenciación que mejora inesperadamente la calidad del ADN de la biblioteca mientras que acelera el proceso de análisis de muestras para diagnósticos prenatales. Los métodos también permiten determinar las variaciones en el número de copias (CNV) de cualquier secuencia de interés en una muestra de prueba que comprende una mezcla de ácidos nucleicos que se conocen o se sospecha que difieren en la cantidad de una o más secuencias de interés, y/o la determinación de la fracción de una de al menos dos poblaciones de ácidos nucleicos contribuyó a la muestra por diferentes genomas. Las secuencias de interés incluyen secuencias genómicas que van desde cientos de bases hasta decenas de megabases hasta cromosomas completos que se sabe o se sospecha que están asociados con una condición genética o de enfermedad. Ejemplos de secuencias de interés incluyen cromosomas asociados con aneuploidías bien conocidas, por ejemplo, trisomía 21 y segmentos de cromosomas que se multiplican en enfermedades como el cáncer, por ejemplo, trisomía parcial 8 en la leucemia mieloide aguda. El método comprende un enfoque estadístico que tiene en cuenta la variabilidad acumulada que se deriva de la variabilidad relacionada con el proceso, intercromosómica o de inter-secuenciación. El método es aplicable para determinar la CNV de cualquier aneuploidía fetal, y se sabe o se sospecha que las CNV están asociadas con una variedad de afecciones médicas.

[0040] A menos que se indique lo contrario, la práctica de la presente invención implica técnicas convencionales comúnmente utilizadas en biología molecular, microbiología, purificación de proteínas, ingeniería de proteínas, secuenciación de proteínas y ADN, y campos de ADN recombinante, que están dentro de los conocimientos de la técnica. Tales técnicas son conocidas por los expertos en la técnica y se describen en numerosos textos estándar y trabajos de referencia.

[0041] Los intervalos numéricos son inclusive de los números que definen el rango. Se pretende que cada limitación

numérica máxima dada a lo largo de esta especificación incluya todas las limitaciones numéricas inferiores, como si dichas limitaciones numéricas más bajas estuvieran expresamente escritas en este documento. Cada limitación numérica mínima dada a lo largo de esta especificación incluirá cada limitación numérica superior, como si tales limitaciones numéricas superiores estuvieran expresamente escritas en este documento. Cada rango numérico dado a lo largo de esta especificación incluirá cada rango numérico más estrecho que caiga dentro de un rango numérico más amplio, como si dichos rangos numéricos más estrechos estuvieran expresamente escritos aquí.

5.1 DEFINICIONES

[0042] Como se usa en este documento, los términos singulares "un", "una", "el" y "ella" incluyen la referencia plural a menos que el contexto indique claramente lo contrario. A menos que se indique lo contrario, los ácidos nucleicos se escriben de izquierda a derecha en la orientación de 5' a 3' y las secuencias de aminoácidos se escriben de izquierda a derecha en la orientación de amino a carboxi, respectivamente.

[0043] El término "evaluación" en el presente documento se refiere a la caracterización de la situación de una aneuploidía cromosómica por uno de los tres tipos de llamadas: "normal", "afectadas", y "no-llamadas". Por ejemplo, en presencia de trisomía, la llamada "normal" está determinada por el valor de un parámetro, por ejemplo, una dosis de cromosoma de prueba que está por debajo de un umbral de confiabilidad definido por el usuario, la llamada "afectada" está determinada por un parámetro, por ejemplo una dosis de cromosoma de prueba, que está por encima de un umbral de confiabilidad definido por el usuario, y el resultado de "no llamada" se determina mediante un parámetro, por ejemplo, una dosis de cromosoma de prueba, que se encuentra entre los umbrales de confiabilidad definidos por el usuario para hacer una llamada "normal" o "afectada".

[0044] El término "variación del número de copia" en el presente documento se refiere a la variación en el número de copias de una secuencia de ácido nucleico que es 1 kb o más grandes presentes en una muestra de prueba en comparación con el número de copias de la secuencia de ácido nucleico presente en una muestra cualificada. Una "variante del número de copias" se refiere a la secuencia de ácido nucleico de 1 kb o más en la que se encuentran las diferencias en los números de copias al comparar una secuencia de interés en la muestra de prueba con la presente en una muestra calificada. Las variantes/variaciones del número de copias incluyen supresiones, incluidas microdelecciones, inserciones, incluidas microinserciones, duplicaciones, multiplicaciones, inversiones, translocaciones y variantes complejas de sitios múltiples. La CNV abarca aneuploidías cromosómicas y aneuploidías parciales.

[0045] El término "aneuploidía" aquí se refiere a un desequilibrio de material genético causado por una pérdida o ganancia de un cromosoma, o una parte de un cromosoma.

[0046] El término "aneuploidía cromosómica" aquí se refiere a un desequilibrio de material genético causado por una pérdida o ganancia de un cromosoma completo, e incluye la aneuploidía de la línea germinal y aneuploidía mosaico.

[0047] El término "aneuploidía parcial" en el presente documento 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 los desequilibrios que resultan de translocaciones, deleciones e inserciones.

[0048] El término "pluralidad" se usa en el presente documento en referencia a un número de moléculas de ácido nucleico o secuencia de etiquetas que es suficiente para identificar diferencias significativas en la variación en el número de copias (e. G. Dosis de cromosomas) en muestras de ensayo y las muestras cualificados utilizando en Los métodos descritos en este documento. En algunas realizaciones, al menos aproximadamente 3×10^6 etiquetas de secuencia, al menos aproximadamente 5×10^6 etiquetas de secuencia, al menos aproximadamente 8×10^6 etiquetas de secuencia, al menos aproximadamente 10×10^6 etiquetas de secuencia, al menos aproximadamente 15×10^6 etiquetas de secuencia, al menos aproximadamente 20×10^6 etiquetas de secuencia, al menos aproximadamente 30×10^6 etiquetas de secuencia, al menos aproximadamente 40×10^6 etiquetas de secuencia, o al menos aproximadamente 50×10^6 etiquetas de secuencia que comprenden entre 20 y 40 pb de lectura Se obtienen para cada muestra de prueba.

[0049] Los términos "polinucleótido", "ácido nucleico" y "moléculas de ácido nucleico" se usan indistintamente y se refieren a una secuencia unida covalentemente de nucleótidos (es decir, ribonucleótidos para ARN y desoxirribonucleótidos para ADN) en la que la posición 3' de la pentosa de un nucleótido está unida por un grupo fosfodiéster a la posición 5' de la pentosa del siguiente, incluye secuencias de cualquier forma de ácido nucleico, incluidas, entre otras, moléculas de ARN, ADN y ADNcf. El término "polinucleótido" incluye, sin limitación, polinucleótido de cadena simple y doble.

[0050] El término "parte" se usa en el presente documento en referencia a la cantidad de información de la secuencia de moléculas de ácido nucleico fetal y materno en una muestra biológica que en cantidad suma a menos que la información de la secuencia de <1 genoma humano.

[0051] El término "muestra de ensayo" en el presente documento se refiere a una muestra que comprende una mezcla de ácidos nucleicos que comprenden al menos una secuencia de ácido nucleico cuyo número de copias es sospechoso de haber sufrido variación. Los ácidos nucleicos presentes en una muestra de prueba se conocen como "ácidos nucleicos de prueba".

[0052] El término "muestra calificado" aquí se refiere a una muestra que comprende una mezcla de ácidos nucleicos que están presentes en un número de copias conocido al que se comparan los ácidos nucleicos en una muestra de ensayo, y es una muestra que es normal, es decir no aneuploide, para la secuencia de interés, por ejemplo, una muestra calificada utilizada para identificar un cromosoma normalizador para el cromosoma 21 es una muestra que no es una muestra de trisomía 21.

[0053] El término "ácido nucleico calificado" se usa indistintamente con "secuencia calificada" es una secuencia contra la que se compara la cantidad de una secuencia de prueba o ensayo de ácido nucleico. Una secuencia calificada es una presente en una muestra biológica, preferiblemente en una representación conocida, es decir. Se conoce la cantidad de una secuencia calificada. Una "secuencia de interés calificada" es una secuencia calificada para la cual se conoce la cantidad en una muestra calificada, y es una secuencia que se asocia con una diferencia en la representación de la secuencia en un individuo con una condición médica.

[0054] El término "secuencia de interés" en el presente documento se refiere a una secuencia de ácido nucleico que está asociada con una diferencia en la secuencia de representación en individuos sanos frente a enfermos. Una secuencia de interés puede ser una secuencia en un cromosoma que está tergiversado, es decir, sobre o subrepresentado, en una enfermedad o condición genética. Una secuencia de interés también puede ser una porción de un cromosoma o un cromosoma. Por ejemplo, una secuencia de interés puede ser un cromosoma que está sobrerrepresentado en una condición de aneuploidía, o un gen que codifica un supresor de tumores que está poco representado en un cáncer. Las secuencias de interés incluyen secuencias que están representadas de forma excesiva o insuficiente en la población total, o una subpoblación de células de un sujeto. Una "secuencia de interés calificada" es una secuencia de interés en una muestra calificada. Una "secuencia de prueba de interés" es una secuencia de interés en una muestra de prueba.

[0055] El término "secuencia normalizadora" aquí se refiere a una secuencia que muestra una variabilidad en el número de etiquetas de secuencia que se asignan a la misma entre las muestras y series de secuenciación que mejor se aproximan a las de la secuencia de interés para las que se utiliza como un parámetro de normalización, y eso puede diferenciar mejor una muestra afectada de una o más muestras no afectadas. Un "cromosoma de normalización" es un ejemplo de una "secuencia de normalización".

[0056] El término "diferenciabilidad" aquí se refiere a la característica de un cromosoma de normalización que permite distinguir una o más muestras no afectadas es decir, normales, de una o más muestras afectadas, es decir aneuploides.

[0057] El término "dosis de secuencia" en el presente documento se refiere a un parámetro que relaciona la densidad de etiqueta de secuencia de una secuencia de interés a la densidad de la etiqueta de una secuencia de normalización. Una "dosis de secuencia de prueba" es un parámetro que relaciona la densidad de la etiqueta de secuencia de una secuencia de interés, por ejemplo, el cromosoma 21, con la de una secuencia de normalización, por ejemplo, el cromosoma 9, determinado en una muestra de prueba. De manera similar, una "dosis de secuencia calificada" es un parámetro que relaciona la densidad de la etiqueta de secuencia de una secuencia de interés con la de una secuencia de normalización determinada en una muestra calificada.

[0058] El término "densidad de etiqueta de secuencia" en el presente documento se refiere al número de lecturas de secuencia que se asignan a una secuencia de referencia del genoma por ejemplo, la densidad de etiqueta de secuencia para el cromosoma 21 es el número de lecturas de secuencia generado por el método de secuenciación que es mapeado en el cromosoma 21 del genoma de referencia. El término "relación de densidad de etiqueta de secuencia" en el presente documento se refiere a la proporción del número de etiquetas de secuencia que se asignan a un cromosoma del genoma de referencia, por ejemplo el cromosoma 21, a la longitud del genosoma de referencia cromosoma 21.

[0059] El término "parámetro" en el presente documento se refiere a un valor numérico que caracteriza un conjunto de datos cuantitativos y/o una relación numérica entre los conjuntos de datos cuantitativos. Por ejemplo, una relación (o función de una relación) entre el número de etiquetas de secuencia asignadas a un cromosoma y la longitud del cromosoma al que se asignan las etiquetas, es un parámetro.

[0060] Los términos "valor umbral" y "valor umbral calificado" en el presente documento se refieren a cualquier número que se calcula usando un conjunto de datos de aceptación establecidos y sirve como un límite de diagnóstico de una variación del número de copias por ejemplo una aneuploidía, en un organismo. Si los resultados obtenidos de la práctica de los métodos descritos en este documento exceden un umbral, se puede diagnosticar a un sujeto con una variación en el número de copias, por ejemplo, trisomía 21.

[0061] El término "lectura" se refiere a una secuencia de ADN de longitud suficiente (p. ej., al menos aproximadamente 30 pb) que se puede utilizar para identificar una secuencia o región más grande, p. ej. que se pueden alinear y específicamente asignar a un cromosoma o región genómica o gen.

[0062] El término "etiqueta de secuencia" se usa aquí de forma intercambiable con el término "etiqueta de secuencia mapeada" para referirse a una lectura de secuencia que ha sido asignada específicamente es decir asignada, a una secuencia más grande, por ejemplo un genoma de referencia, por la alineación. Las etiquetas de secuencia asignadas se asignan de forma única a un genoma de referencia, es decir, se asignan a una única ubicación para el genoma de referencia. Las etiquetas que se pueden asignar a más de una ubicación en un genoma de referencia, es decir, las etiquetas que no se asignan de forma única, no se incluyen en el análisis.

[0063] Como se usa en el presente documento, los términos "alineado", "alineación", o "alinear" se refieren a una o más secuencias que se identifican como un partido en términos de la orden de sus moléculas de ácido nucleico a una secuencia conocida a partir de un genoma de referencia. Dicha alineación se puede hacer manualmente o mediante un algoritmo de computadora, siendo ejemplos de ello el programa de computadora Efficient Local Alignment of Nucleotide Data (ELAND) distribuido como parte de la tubería de Illumina Genomics Analysis. La coincidencia de una secuencia leída en alineación puede ser una coincidencia de secuencia del 100% o menos del 100% (coincidencia no perfecta).

[0064] Tal como se utiliza aquí, el término "genoma de referencia" se refiere a cualquier secuencia del genoma particular conocido, ya sea parcial o completa, de cualquier organismo o virus que puede ser utilizado para hacer referencia a secuencias identificadas a partir de un sujeto. Por ejemplo, un genoma de referencia utilizado para sujetos humanos, así como muchos otros organismos, se encuentra en el Centro Nacional de Información Biotecnológica en www.ncbi.nlm.nih.gov. Un "genoma" se refiere a la información genética completa de un organismo o virus, expresado en secuencias de ácidos nucleicos.

[0065] Los términos "genoma de secuencias diana artificiales" y "genoma de referencia artificial" en el presente documento se refieren a una agrupación de secuencias conocidas que abarcan alelos de sitios polimórficos conocidos. Por ejemplo, un "genoma de referencia de SNP" es un genoma de secuencias diana artificiales que comprende una agrupación de secuencias que abarcan alelos de SNP conocidos.

[0066] El término "secuencia clínicamente relevante" en el presente documento se refiere a una secuencia de ácido nucleico que se sabe o se sospecha que estar asociado o implicado con una condición genética o enfermedad. La determinación de la ausencia o presencia de una secuencia clínicamente relevante puede ser útil para determinar un diagnóstico o confirmar un diagnóstico de una afección médica, o proporcionar un pronóstico para el desarrollo de una enfermedad.

[0067] El término "derivado" cuando se usa en el contexto de un ácido nucleico o una mezcla de ácidos nucleicos, en este documento se refiere a los medios por los cuales el (los) ácido(s) nucleico(s) se obtiene(n) de la fuente de la que se originan. Por ejemplo, en una realización, una mezcla de ácidos nucleicos que se deriva de dos genomas diferentes significa que los ácidos nucleicos, por ejemplo, ADNcf, fueron liberados naturalmente por las células a través de procesos naturales como la necrosis o la 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.

[0068] El término "muestra mixta" aquí se refiere a una muestra que contiene una mezcla de ácidos nucleicos, que se derivan de diferentes genomas.

[0069] El término "muestra materna" aquí se refiere a una muestra biológica obtenida de un sujeto embarazado, por ejemplo, una mujer.

[0070] El término "muestra materna original" en este documento se refiere a una muestra biológica obtenida de un sujeto embarazado, por ejemplo, una mujer, que sirve como fuente de la cual se extrae una porción para amplificar los ácidos nucleicos diana polimórficos. La "muestra original" puede ser cualquier muestra obtenida de un sujeto embarazado, y las fracciones procesadas de la misma, por ejemplo, una muestra de ADNcf purificada extraída de una muestra de plasma materno.

[0071] El término "fluido biológico" se refiere aquí a un líquido tomado de una fuente biológica e incluye, por ejemplo, sangre, suero, plasma, esputo, fluido de lavado, líquido cefalorraquídeo, orina, semen, sudor, lágrimas, saliva, y similares. Como se usa en el presente documento, los términos "sangre", "plasma" y "suero" abarcan expresamente las fracciones o porciones procesadas de los mismos. Del mismo modo, donde se toma una muestra de una biopsia, un hisopo, frotis, etc., la "muestra" abarca expresamente una fracción o porción transformadas derivadas de la biopsia, un hisopo, frotis, etc.

[0072] Los términos "ácidos nucleicos maternos" y "ácidos nucleicos fetales" en el presente documento se refieren a los ácidos nucleicos de una mujer embarazada y a los ácidos nucleicos del feto que son transportados por la hembra

embarazada, respectivamente.

[0073] Tal como se utiliza aquí, el término "correspondiente a" se refiere a una secuencia de ácido nucleico, por ejemplo, un gen o un cromosoma, que está presente en el genoma de diferentes temas, y que no necesariamente tienen 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 cromosoma.

[0074] Tal como se utiliza aquí, el término "sustancialmente libre de células" abarca preparaciones de la muestra deseada de la que se eliminan los componentes que normalmente están asociados con ella. Por ejemplo, una muestra de plasma se vuelve esencialmente libre de células mediante la eliminación de células sanguíneas, por ejemplo, células rojas, que normalmente están asociadas con ella. En algunas realizaciones, se procesan muestras sustancialmente libres para eliminar las células que de otro modo contribuirían al material genético deseado que se va a analizar para una CNV.

[0075] Tal como se utiliza aquí, el término "fracción fetal" se refiere a la fracción de ácidos nucleicos fetales presentes en una muestra que comprende ácido nucleico fetal y materno.

[0076] Tal como se utiliza aquí, el término "cromosoma" se refiere al portador del gen herencia de soporte de una célula viva que se deriva de la cromatina y que comprende ADN y proteínas componentes (especialmente histonas). El sistema de numeración de cromosomas del genoma humano individual convencional reconocido internacionalmente se emplea aquí.

[0077] Tal como se utiliza aquí, el término "longitud de polinucleótido" se refiere a que el número absoluto de moléculas de ácido nucleico (nucleótidos) en una secuencia o en una región de un genoma de referencia. El término "longitud del cromosoma" se refiere a la longitud conocida del cromosoma dado en pares de bases, por ejemplo, proporcionado en el ensamblaje NCBI36/hg18 del cromosoma humano encontrado en la red mundial en genome.ucsc.edu/cgi-bin/hgTracks?hgid=167155613&chromInfoPage=

[0078] El término "sujeto" se refiere a un sujeto humano así como a un sujeto no humano tal como un mamífero, un invertebrado, un vertebrado, un hongo, una levadura, una bacteria y un virus. Aunque los ejemplos en este documento se refieren a seres humanos y el lenguaje se dirige 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 la medicina veterinaria, ciencias animales, laboratorios de investigación y otros.

[0079] El término "condición" en este documento se refiere a "condición médica" como un término amplio que incluye todas las enfermedades y trastornos, pero pueden incluir lesiones y situaciones de salud normales, como el embarazo, que pueden afectar a la salud de una persona, se benefician de la asistencia médica, o tienen implicaciones para tratamientos médicos.

[0080] El término "cromosoma aneuploides" aquí se refiere a un cromosoma que está implicado en una aneuploidía.

[0081] El término "aneuploidía" aquí se refiere a un desequilibrio de material genético causado por una pérdida o ganancia de un todo cromosoma, o parte de un cromosoma.

[0082] Los términos "biblioteca" y "biblioteca de secuenciación" en el presente documento se refieren a una colección o pluralidad de moléculas de plantilla que comparten secuencias comunes en sus extremos 5' y secuencias comunes en sus extremos 3'.

[0083] Los términos "extremos romos" y de "reparación de extremos" se usan aquí de forma intercambiable para referirse a un proceso enzimático que se traduce en ambas cadenas de una molécula de ADN de doble cadena para terminar en un par de bases, y no incluye la purificación de productos de extremos romos de la enzima de extremos romos.

[0084] El término "cola de dA" aquí se refiere a un proceso enzimático que añade al menos una base de adenina al extremo 3' del ADN, y no incluye la purificación del producto dA de cola de la enzima de cola de dA.

[0085] El término "adaptador de ligadura" aquí se refiere a un proceso enzimático que se liga una secuencia adaptadora de ADN a los fragmentos de ADN, y no incluye purificar el producto ligado al adaptador de la enzima de ligadura.

[0086] El término "recipiente de reacción" se refiere aquí a un recipiente de cualquier forma, tamaño, capacidad o material que puede ser utilizado para el procesamiento de una muestra durante un laboratorio procedimiento por ejemplo, investigación o clínico.

[0087] El término "pasos consecutivos" se usa aquí en referencia a las sucesivas etapas enzimáticas de extremo romo de cola de dA y ADN ligado al adaptador que no están interpuestos por etapas de purificación.

[0088] Tal como se utiliza aquí, el término "purificado" se refiere a material (por ejemplo, un polinucleótido aislado) que está en un estado relativamente puro, por ejemplo, al menos aproximadamente el 80% puro, al menos aproximadamente el 85% puro, al menos aproximadamente 90% puro, al menos aproximadamente el 95% puro, al menos aproximadamente el 98% puro, o incluso al menos aproximadamente el 99% puro.

[0089] Los términos "extraído", "recuperado", "aislado" y "separado", se refieren a un compuesto, proteína, célula, ácido nucleico o de aminoácidos que se retira de al menos un componente con el que está asociado de forma natural y se encuentra en la naturaleza.

[0090] El término "SNP en tándem" aquí se refiere a dos o más SNP que están presentes dentro de una secuencia de ácido nucleico diana polimórfico.

[0091] Los términos "ácido nucleico diana polimórfico", "secuencia polimórfica", "secuencia de ácido nucleico diana polimórfico" y "ácido nucleico polimórfico" se usan indistintamente en este documento para referirse a una secuencia de ácido nucleico por ejemplo, una secuencia de ADN, que comprende uno o más sitios polimórficos.

[0092] El término "sitio polimórfico" en el presente documento se refiere a un polimorfismo de nucleótido único (SNP), una pequeña escala deleción multi-base o de inserción, un polimorfismo de multi-nucleótidos (MNP) o una repetición corta en tándem (STR).

[0093] El término "pluralidad de ácidos nucleicos diana polimórficos" aquí se refiere a una serie de secuencias de ácido nucleico que comprende cada uno al menos un sitio polimórfico tal 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 a partir de los ácidos nucleicos diana polimórficos para identificar y/o cuantificar los alelos fetales presentes en muestras maternas que comprenden ácidos nucleicos fetales y maternos.

[0094] El término "enriquecer" en el presente documento se refiere al proceso de amplificación de ácidos nucleicos diana polimórficos contenidos en una porción de una muestra materna, y combinar el producto amplificado con el resto de la muestra maternal de la que se eliminó la porción.

[0095] El término "densidad de etiqueta de secuencia" en el presente documento se refiere al número de lecturas de secuencia que se asignan a una secuencia de referencia del genoma por ejemplo, la densidad de etiqueta de secuencia para el cromosoma 21 es el número de lecturas de secuencia generado por el método de secuenciación que es mapeado en el cromosoma 21 del genoma de referencia. El término "relación de densidad de etiqueta de secuencia" en el presente documento se refiere a la proporción del número de etiquetas de secuencia que se asignan a un cromosoma del genoma de referencia, por ejemplo el cromosoma 21, a la longitud del genosoma de referencia cromosoma 21.

[0096] Tal como se utiliza aquí, el término "amplificación en fase sólida", como se usa en el presente documento se refiere a cualquier reacción de amplificación de ácido nucleico aplicada en o en asociación con un soporte sólido tal que la totalidad o una parte de los productos amplificados se inmovilizan en el soporte sólido a medida que se forman. En particular, el término abarca la reacción en cadena de la polimerasa en fase sólida (PCR en fase sólida) y la amplificación isotérmica en fase sólida, que son reacciones análogas a la amplificación en fase de solución estándar, excepto que uno o ambos de los cebadores de amplificación directa e inversa se inmovilizan en el soporte sólido. La PCR en fase sólida cubre sistemas como las emulsiones, en donde un cebador se ancla a una cuenta y el otro se encuentra en solución libre, y la formación de colonias en matrices de gel de fase sólida donde un cebador se ancla a la superficie y el otro se encuentra en solución libre. El término fase sólida, o superficie, se usa para significar una matriz plana en la que los cebadores se unen a una superficie plana, por ejemplo, portaobjetos de microscopio de vidrio, sílice o plástico o dispositivos de células de flujo similares; perlas, en donde uno o dos cebadores se unen a las perlas y se amplifican las perlas; o una serie de cuentas sobre una superficie después de que las cuentas hayan sido amplificadas.

[0097] Como se usa en el presente documento, el término "grupo de cromosomas" aquí se refiere a un grupo de dos o más cromosomas.

[0098] Un "polimorfismo de nucleótido único" (SNP) se produce en un sitio polimórfico ocupado por un único nucleótido, que es el sitio de variación entre secuencias alélicas. El sitio suele ir precedido y seguido por secuencias altamente conservadas del alelo (por ejemplo, secuencias que varían en menos de 1/100 o 1/1000 miembros de las poblaciones). Un SNP generalmente surge debido a la sustitución de un nucleótido por otro en el sitio polimórfico. Una transición es el reemplazo de una purina por otra purina o una pirimidin por otra pirimidin. Una transversión es el reemplazo de una purina por una pirimidin o viceversa. Los SNP también pueden surgir de una eliminación de un nucleótido o una inserción de un nucleótido en relación con un alelo de referencia. Los polimorfismos de un solo nucleótido (SNP) son posiciones en las cuales dos bases alternativas ocurren con una frecuencia apreciable (>1%) en la población humana, y son el tipo más común de variación genética humana.

[0099] Tal como se utiliza aquí, el término "repeticón corta en tándem" o "STR" tal como se utiliza aquí, se refiere a una clase de polimorfismos que se produce cuando un patrón de dos o más nucleótidos se repiten y las secuencias repetidas son directamente adyacentes entre sí. El patrón puede variar en longitud de 2 a 10 pares de bases (pb) (por ejemplo (CATG)_n en una región genómica) y está típicamente en la región intrón no codificadora. Al examinar varios loci STR y al contar cuántas repeticiones de una secuencia STR específica en un locus dado, es posible crear un perfil genético único de un individuo.

[0100] Tal como se utiliza aquí, el término "miniSTR" aquí se refiere a la repetición en tándem de cuatro o más pares de bases que se extiende por menos de aproximadamente 300 pares de bases, menos de aproximadamente 250 pares de bases, menos de aproximadamente 200 pares de bases, menos de aproximadamente 150 pares de bases, menos de unos 100 pares de bases, menos de unos 50 pares de bases, o menos de unos 25 pares de bases. "miniSTRs" son STR que son amplificables a partir de plantillas ADNcf.

[0101] El término "tándem SNPs" aquí se refiere a dos o más SNPs que están presentes dentro de una secuencia de ácido nucleico diana polimórfico.

[0102] Tal como se utiliza aquí, el término "biblioteca enriquecida" en el presente documento se refiere a una biblioteca de secuenciación que comprende secuencias de ácido nucleico diana polimórfico amplificado. Un ejemplo de una biblioteca enriquecida es una biblioteca de secuenciación que comprende secuencias de ADNcf de origen natural y secuencias de ácido nucleico diana amplificadas. Una "biblioteca no enriquecida" en el presente documento se refiere a una biblioteca de secuenciación que no comprende, es decir, una biblioteca generada a partir de secuencias de ADNcf que ocurren naturalmente. Una "biblioteca polimórfica de ácidos nucleicos diana" es una biblioteca generada a partir de ácidos nucleicos diana amplificados.

[0103] Como se usa en este documento, el término "secuencias ADNcf de origen natural" en el presente documento se refiere a fragmentos ADNcf ya que están presentes en una muestra, y en contraste con fragmentos de ADN genómico que se obtienen por métodos de fragmentación descritos en este documento.

5.2 DESCRIPCION

[0104] La descripción se refiere a métodos para determinar la presencia o ausencia de una aneuploidía por ejemplo aneuploidía cromosómica o parcial, y/o fracción fetal en muestras maternas que comprenden ácidos nucleicos fetales y maternos por secuenciación masiva en paralelo. El método comprende un nuevo protocolo para preparar bibliotecas de secuenciación que mejora inesperadamente la calidad del ADN de la biblioteca mientras acelera el proceso de análisis de muestras para diagnósticos prenatales. Los métodos permiten determinar las variaciones en el número de copias (CNV) de cualquier secuencia de interés en una muestra de prueba que comprende una mezcla de ácidos nucleicos que se conocen o se sospecha que difieren en la cantidad de una o más secuencias de interés, y/o determinan la fracción de una de al menos dos poblaciones de ácidos nucleicos contribuyó a la muestra por diferentes genomas.

Métodos de secuenciación

[0105] En una realización, el método descrito aquí emplea la tecnología de secuenciación de próxima generación (NGS) en la que plantillas de ADN clonalmente amplificadas o moléculas de ADN individuales se secuencian de forma masiva en paralelo 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 secuencia de alto rendimiento, NGS proporciona información cuantitativa digital, en el sentido de que cada secuencia leída es una "etiqueta de secuencia" contable que representa una plantilla de ADN clonal individual o una molécula de ADN individual. Esta cuantificación permite a NGS expandir el concepto de PCR digital de conteo de moléculas de ADN libres de células (Fan et al., Proc Natl Acad Sci EE.UU. 105: 16266-16271 [2008]; Chiu et al., Proc Natl Acad Sci EE.UU. 2008; 105: 20458-20463 [2008]). Las tecnologías de secuenciación de NGS incluyen pirosecuenciación, secuenciación por síntesis con terminadores de tinte reversibles, secuenciación por ligadura de sonda de oligonucleótidos y secuenciación en tiempo real.

[0106] Algunas de las tecnologías de secuenciación están disponibles comercialmente, tales como la plataforma de secuenciación por hibridación de Affymetrix Inc. (Sunnyvale, CA) y las plataformas de secuenciación por síntesis a partir de 454 Life Sciences (Bradford, CT), Illumina/Solexa (Hayward, CA) y Helicos Biosciences (Cambridge, MA), y la plataforma de secuenciación por ligación de Applied Biosystems (Foster City, CA), como se describe a continuación. Además de la secuenciación de una sola molécula realizada mediante la secuenciación por síntesis de Helicos Biosciences, otras tecnologías de secuenciación de una sola molécula están abarcadas por el método de la invención e incluyen la tecnología SMRT™ de Pacific Biosciences, la tecnología Ion Torrent™ y la secuenciación de nanoporos en desarrollo. por ejemplo, por Oxford Nanopore Technologies.

[0107] Mientras que el método automatizado de Sanger se considera una tecnología de "primera generación", la secuenciación de Sanger que incluye la secuenciación automática de Sanger también puede emplearse mediante el método de la invención. Los métodos de secuenciación adicionales que comprenden el uso de tecnologías de

obtención de imágenes de ácido nucleico en desarrollo, por ejemplo, microscopía de fuerza atómica (AFM) o microscopía electrónica de transmisión (TEM), también se incluyen en el método de la invención. Tecnologías de secuenciación ejemplares se describen a continuación.

[0108] En una realización, la tecnología de secuenciación de ADN que se utiliza en el método de la invención es la Secuenciación de Moléculas Únicas Verdaderas de Helicos (tSMS) (por ejemplo, como se describe en Harris TD et al., Science 320: 106-109 [2008]). En la técnica tSMS, una muestra de ADN se divide en cadenas de aproximadamente 100 a 200 nucleótidos, y se agrega una secuencia políA al extremo 3' de cada cadena de ADN. Cada hebra se marca mediante la adición de un nucleótido de adenosina marcado de forma fluorescente. Las cadenas de ADN se hibridan luego con 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 tener una densidad de aproximadamente 100 millones de plantillas/cm². La célula de flujo se carga en un instrumento, por ejemplo, el secuenciador Heliscope™, y un láser ilumina la superficie de la célula de flujo, revelando la posición de cada plantilla. Una cámara CCD puede asignar la posición de las plantillas en la superficie de la célula de flujo. La plantilla de la etiqueta fluorescente se escinde y se lava. La reacción de secuenciación comienza con la introducción de una ADN polimerasa y un nucleótido marcado con fluorescencia. El ácido nucleico oligo-T sirve como cebador. La polimerasa incorpora los nucleótidos marcados al cebador en una forma dirigida por la plantilla. Se eliminan la polimerasa y los nucleótidos no incorporados. Las plantillas que tienen incorporación directa del nucleótido marcado con fluorescencia se diferencian por la obtención de imágenes de la superficie de la célula de flujo. Después de la obtención de imágenes, una etapa de escisión elimina la etiqueta fluorescente, y el proceso se repite con otros nucleótidos marcados con fluorescencia hasta que se alcanza la longitud de lectura deseada. La información de secuencia se recopila con cada paso de adición de nucleótidos.

[0109] En una realización, la tecnología de secuenciación de ADN que se utiliza en el método de la invención es la secuencia 454 (Roche) (por ejemplo, como se describe en Margulies, M. y otros, Nature 437: 376-380 [2005]). La secuenciación 454 implica dos pasos. En el primer paso, el ADN se corta en fragmentos de aproximadamente 300-800 pares de bases, y los fragmentos son romos. Los adaptadores de oligonucleótidos se ligan a los extremos de los fragmentos. Los adaptadores sirven como cebadores para la amplificación y secuenciación de los fragmentos. Los fragmentos pueden unirse a las perlas de captura de ADN, por ejemplo, perlas recubiertas con estreptavidina usando, por ejemplo, el Adaptador B, que contiene la etiqueta 5'-biotina. Los fragmentos unidos a las perlas se amplifican por PCR dentro de gotitas de una emulsión de aceite-agua. El resultado son múltiples copias de fragmentos de ADN amplificados clonalmente en cada cuenta. En el segundo paso, las cuentas se capturan en pozos (tamaño pico-litro). La pirosecuenciación se realiza en cada fragmento de ADN en paralelo. La adición de uno o más nucleótidos genera una señal de luz que es grabada por una cámara CCD en un instrumento de secuenciación. La intensidad de la señal es proporcional al número de nucleótidos incorporados. La pirosecuenciación hace uso del pirofosfato (PPI) que se libera tras la adición de nucleótidos. El PPI se convierte en ATP por la ATP sulfúrilasa en presencia de adenosina 5' fosfosulfato. La luciferasa usa ATP para convertir la luciferina en oxiluciferina, y esta reacción genera luz que se discierne y analiza.

[0110] En una realización, la tecnología de secuenciación de ADN que se utiliza en el método de la invención es la tecnología SOLiD™ (Applied Biosystems). En la secuenciación por ligación de SOLiD™, el ADN genómico se divide en fragmentos, y los adaptadores se unen a los extremos 5' y 3' de los fragmentos para generar una biblioteca de fragmentos. Alternativamente, pueden introducirse adaptadores internos ligando los adaptadores a los extremos 5' y 3' de los fragmentos, circularizando los fragmentos, dirigiendo el fragmento circularizado para generar un adaptador interno, y uniendo adaptadores a los extremos 5' y 3' de los extremos resultantes. Fragmentos para generar una biblioteca emparejada. A continuación, se preparan poblaciones de perlas clonales en microrreactores que contienen perlas, cebadores, plantilla y componentes de PCR. Después de la PCR, las plantillas se desnaturalizan y las perlas se enriquecen para separar las perlas con plantillas extendidas. Las plantillas en las perlas seleccionadas se someten a una modificación 3' que permite la unión a un portaobjetos de vidrio. La secuencia se puede determinar por hibridación secuencial y ligación de oligonucleótidos parcialmente aleatorios con una base determinada central (o un par de bases) que se identifica por un fluoróforo específico. Después de registrar un color, el oligonucleótido ligado se escinde y se elimina, y luego se repite el proceso.

[0111] En una realización, la tecnología de secuenciación de ADN que se utiliza en el método de la invención es la tecnología de secuenciación de una sola molécula, en tiempo real (SMRT™) de Pacific Biosciences. En la secuenciación SMRT, la incorporación continua de nucleótidos marcados con colorante se visualiza durante la síntesis de ADN. Las moléculas de ADN polimerasa individuales se unen a la superficie inferior de los identificadores individuales de longitud de onda de modo cero (identificadores de ZMW) que obtienen información de la secuencia mientras que los nucleótidos fosforizados se incorporan a la cadena de cebadores en crecimiento. Una ZMW es una estructura de confinamiento que permite la observación de la incorporación de un solo nucleótido por la ADN polimerasa en el contexto de nucleótidos fluorescentes que se difunden rápidamente en una salida de la ZMW (en microsegundos). Se requieren varios milisegundos para incorporar un nucleótido en una hebra en crecimiento. Durante este tiempo, la etiqueta fluorescente se excita y produce una señal fluorescente, y la etiqueta fluorescente se separa. La identificación de la fluorescencia correspondiente del tinte indica qué base se incorporó. El proceso se repite.

[0112] En una realización, la tecnología de secuenciación de ADN que se utiliza en el método de la invención es secuenciación de 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 de Nanoporo están siendo desarrolladas industrialmente por varias compañías, incluida Oxford Nanopore Technologies (Oxford, Reino Unido). La secuenciación de nanoporos es una tecnología de secuenciación de una sola molécula mediante la cual una única molécula de ADN se secuencia directamente cuando 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 produce una ligera corriente eléctrica debido a la conducción de iones a través del nanoporo. La cantidad de corriente que fluye es sensible al tamaño y forma del nanoporo. A medida que una molécula de ADN pasa a través de un nanoporo, cada nucleótido en la molécula de ADN obstruye el nanoporo en un grado diferente, cambiando la magnitud de la corriente a través del nanoporo en diferentes grados. Por lo tanto, 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.

[0113] En una realización, la tecnología de secuenciación de ADN que se utiliza en el método de la invención es la matriz de transistor de efecto de campo sensible (chemFET) a los productos químicos (por ejemplo, como se describe en la Publicación de la Solicitud de Patente de Estados Unidos N° 20090026082). En un ejemplo de la técnica, las moléculas de ADN se pueden colocar en cámaras de reacción, y las moléculas de plantilla se pueden hibridar con un cebador de secuenciación unido a una polimerasa. La incorporación de uno o más trifosfatos en una nueva cadena de ácido nucleico en el extremo 3' del cebador de secuenciación puede discernirse por un cambio en la corriente por parte de un chemFET. Una matriz puede tener múltiples sensores chemFET. En otro ejemplo, pueden unirse ácidos nucleicos individuales a las perlas, y los ácidos nucleicos pueden amplificarse en la perla, y las perlas individuales pueden transferirse a cámaras de reacción individuales en una matriz chemFET, teniendo cada cámara un sensor chemFET, y los ácidos nucleicos pueden secuenciarse.

[0114] En una realización, la tecnología de secuenciación de ADN que se utiliza en el método de la invención es el método de la Halcyon Molecular que utiliza microscopía electrónica de transmisión (TEM). El método, denominado Individual Molecule Placement Rapid Nano Transfer (IMPRNT), comprende utilizar imágenes de microscopía electrónica de transmisión de resolución de un solo átomo de ADN de alto peso molecular (150 kb o más) etiquetadas selectivamente con marcadores de átomos pesados y disponer estas moléculas en películas ultra-delgadas en matrices paralelas ultra-densas (3nm de hebra a hebra) con espaciado constante de base a base. El microscopio electrónico se utiliza para obtener imágenes de las moléculas en las películas para determinar la posición de los marcadores de los átomos pesados y para extraer información de la secuencia de bases del ADN. El método se describe con más detalle en la publicación de patente PCT WO 2009/046445. El método permite secuenciar genomas humanos completos en menos de diez minutos.

[0115] En una realización, la tecnología de secuenciación de ADN es la secuenciación de molécula única de Ion Torrent, que combina la tecnología de semiconductores con una química de secuenciación simple para traducir directamente la información codificada químicamente (A, C, G, T) en información digital (0, 1) en un chip semiconductor. En la naturaleza, cuando un nucleótido se incorpora a una cadena de ADN por una polimerasa, se libera un ión de hidrógeno como un subproducto. Ion Torrent utiliza una matriz de alta densidad de pozos micromecanizados para realizar este proceso bioquímico de forma masivamente paralela. Cada pozo contiene una molécula de ADN diferente. Debajo de los pozos hay una capa sensible a los iones y, por debajo, un sensor de iones. Cuando un nucleótido, por ejemplo una C, se agrega a una plantilla de ADN y luego se incorpora a una hebra de ADN, se liberará un ion hidrógeno. La carga de ese ion cambiará el pH de la solución, que puede identificarse mediante el sensor de iones de Ion Torrent. El secuenciador, esencialmente el medidor de pH de estado sólido más pequeño del mundo, llama a la base y va directamente de la información química a la información digital. El secuenciador de la máquina de genoma personal de Ion (PGMTM) luego inunda secuencialmente el chip con un nucleótido tras otro. Si el siguiente nucleótido que inunda el chip no coincide. No se registrará ningún cambio de voltaje y no se llamará ninguna base. Si hay dos bases idénticas en la cadena de ADN, el voltaje será doble y el chip registrará dos bases idénticas llamadas. La identificación directa permite el registro de la incorporación de nucleótidos en segundos.

[0116] Otros métodos de secuenciación incluyen PCR digital y secuenciación por hibridación. La reacción en cadena de la polimerasa digital (PCR digital o dPCR) se puede usar para identificar y cuantificar directamente los ácidos nucleicos en una muestra. La PCR digital se puede realizar en una emulsión. Los ácidos nucleicos individuales se separan, por ejemplo, en un dispositivo de cámara microfluídica, y cada lata nucleica se amplifica individualmente por PCR. Los ácidos nucleicos pueden separarse, por lo que hay un promedio de aproximadamente 0,5 ácidos nucleicos/pocillo, o no más de un ácido nucleico/pocillo. Se pueden utilizar diferentes sondas para distinguir los alelos fetales y los 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 polinucleotídicas con una pluralidad de sondas polinucleotídicas, en donde cada una de la pluralidad de sondas polinucleotídicas puede unirse opcionalmente a un sustrato. El sustrato podría ser una superficie plana que comprende una serie de secuencias de nucleótidos conocidas. El patrón de hibridación con la matriz se puede usar para determinar las secuencias polinucleotídicas presentes en la muestra. En otras realizaciones, cada sonda está atada a un cordón, por ejemplo, un cordón magnético o similar. La hibridación a las perlas puede identificarse y

usarse para identificar la pluralidad de secuencias de polinucleótidos dentro de la muestra.

[0117] En una realización, el método emplea una secuenciación masiva paralela de millones de fragmentos de ADN utilizando la secuenciación de Illumina por síntesis y la química de la secuenciación basada en terminador reversible (por ejemplo, como se describe en Bentley et al., Nature 6: 53-59 [2009]). El ADN de plantilla puede ser ADN genómico, por ejemplo, ADNcf. En algunas realizaciones, el ADN genómico de células aisladas se utiliza como plantilla y se fragmenta en longitudes de varios cientos de pares de bases. En otras realizaciones, se utiliza ADNcf como plantilla, y no se requiere fragmentación ya que ADNcf existe como fragmentos cortos. Por ejemplo, el ADNcf fetal circula en el torrente sanguíneo como fragmentos de <300 pb, y se estima que el ADNcf materno circula como fragmentos de entre aproximadamente 0,5 y 1 Kb (Li et al., Clin Chem, 50: 1002-1011 [2004]). La tecnología de secuenciación de Illumina se basa en la unión de ADN genómico fragmentado a una superficie plana, ópticamente transparente sobre la cual se unen los anclajes de oligonucleótidos. El ADN de plantilla se repara en el extremo para generar extremos romos fosforilados en 5', y la actividad de la polimerasa del fragmento Klenow se usa para agregar una sola base A al extremo 3' de los fragmentos de ADN fosforilados romos. Esta adición prepara los fragmentos de ADN para la ligadura de los adaptadores de oligonucleótidos, que tienen un saliente de una sola base T en su extremo 3' para aumentar la eficiencia de la ligadura. Los oligonucleótidos adaptadores son complementarios a los anclajes de las células de flujo. En condiciones de dilución limitante, el ADN de plantilla monocatenaria modificada por adaptador se agrega a la célula de flujo y se inmoviliza por hibridación a los anclajes. Los fragmentos de ADN adjuntos se extienden y se amplifican en puente para crear una célula de flujo de secuenciación de densidad ultraalta con cientos de millones de agrupaciones, cada una de las cuales contiene ~ 1.000 copias de la misma plantilla. En una realización, el ADN genómico fragmentado aleatoriamente, por ejemplo, el ADNcf, se amplifica utilizando la PCR antes de someterse a la amplificación por grupos. Alternativamente, se utiliza una preparación de biblioteca genómica libre de amplificación, y el ADN genómico fragmentado aleatoriamente, por ejemplo, ADNcf se enriquece utilizando la amplificación de agrupación sola (Kozarewa et al., Nature Methods 6: 291-295 [2009]). Las plantillas se secuencian utilizando una robusta tecnología de secuenciación por síntesis de ADN de cuatro colores que emplea terminadores reversibles con tintes fluorescentes removibles. La identificación de fluorescencia de alta sensibilidad se logra mediante la excitación con láser y la óptica de reflexión interna total. Lecturas de secuencias cortas de aproximadamente 20-40 pb, *p. ej.*, 36 pb, están alineadas con un genoma de referencia con repetición enmascarada y las diferencias genéticas se denominan mediante el uso de software de pipeline de análisis de datos especialmente desarrollado. Después de completar la primera lectura, las plantillas pueden regenerarse in situ para permitir una segunda lectura desde el extremo opuesto de los fragmentos. De este modo, según el método, se utiliza la secuenciación de un solo extremo o un extremo de los fragmentos de ADN. Se realiza la secuenciación parcial de los fragmentos de ADN presentes en la muestra, y se cuentan las etiquetas de secuencia que comprenden lecturas de longitud predeterminada, por ejemplo, 36 pb, que se asignan a un genoma de referencia conocido.

[0118] La longitud de la secuencia de lectura está asociada con la tecnología de secuenciación particular. Los métodos NGS proporcionan lecturas de secuencia que varían en tamaño desde decenas hasta cientos de pares de bases. En algunas realizaciones del método descrito en el presente documento, las lecturas de la secuencia son de aproximadamente 20 pb, aproximadamente 25 pb, aproximadamente 30 pb, aproximadamente 35 pb, aproximadamente 40 pb, aproximadamente 45 pb, aproximadamente 50 pb, aproximadamente 60 pb, aproximadamente 65 pb, aproximadamente 70 pb, aproximadamente 75 pb, aproximadamente 80 pb, aproximadamente 85 pb, aproximadamente 90 pb, aproximadamente 95 pb, aproximadamente 100 pb, aproximadamente 110 pb, aproximadamente 120 pb, aproximadamente 130 pb, aproximadamente 140 pb, aproximadamente 150 pb, aproximadamente 250 pb, aproximadamente 300 pb, aproximadamente 350 pb, aproximadamente 400 pb, aproximadamente 450 pb o aproximadamente 500 pb. Se espera que los avances tecnológicos permitan lecturas de un solo extremo de más de 500 pb, permitiendo lecturas de más de aproximadamente 1.000 pb cuando se generan lecturas de extremos emparejados. En una realización, las lecturas de secuencia son 36 pb. Otros métodos de secuenciación que pueden emplearse por el método de la invención incluyen los métodos de secuenciación de una sola molécula que pueden secuenciar moléculas de ácidos nucleicos >5000 pb. La cantidad masiva de salida de secuencia se transfiere mediante una tubería de análisis que transforma la salida de imágenes primarias del secuenciador en cadenas de bases. Un paquete de algoritmos integrados realiza los pasos principales de la transformación de datos primarios: análisis de imágenes, puntuación de intensidad, llamada de base y alineación.

[0119] En una realización, se lleva a cabo la secuenciación parcial de fragmentos de ADN presentes en la muestra, y las etiquetas de secuencia que comprende lecturas de longitud predeterminadas, por ejemplo 36 pb, que mapa para un genoma de referencia conocida se cuentan. Solo las lecturas de secuencia que se alinean de forma única con el genoma de referencia se cuentan como etiquetas de secuencia. En una realización, el genoma de referencia es la secuencia del genoma de referencia humano NCBI36/hg18, que está disponible en la web 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 (Laboratorio Europeo de Biología Molecular) y DDBJ (el Banco de Datos de ADN de Japón). En otra realización, el genoma de referencia comprende la secuencia de genoma de referencia humano NCBI36/hg18 y un genoma de secuencias diana artificial, que incluye secuencias diana polimórficas, por ejemplo, un genoma de SNP que comprende las SEQ ID NO: 1-56. En otra realización más, el genoma de referencia es un genoma de secuencia diana artificial que comprende secuencias diana polimórficas, por ejemplo, secuencias de SNP de SEQ ID NO: 1-56.

[0120] El mapeo de las etiquetas de secuencia se logra comparando la secuencia de la etiqueta con la secuencia del genoma de referencia para determinar el origen cromosómico de la molécula de ácido nucleico secuenciado (por ejemplo, ADNcf), y no se necesita información específica de la secuencia genética. Hay una serie de algoritmos informáticos disponibles para alinear secuencias, incluyendo, entre otros, BLAST (Altschul et al., 1990), BLITZ (MPsrch) (Sturrock y 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, EE.UU.). En una realización, un extremo de las copias expandidas clonalmente de las moléculas de ADNrc de plasma se secuencia y procesa mediante análisis de alineación bioinformática para el Analizador del Genoma de Illumina, que utiliza el software Efficient Large Scale Alignment of Nucleotide Databases (ELAND). El análisis de la información de secuenciación para la determinación de aneuploidía puede permitir que un pequeño grado de desajuste (0-2 desajustes por etiqueta de secuencia) tenga en cuenta los polimorfismos menores que pueden existir entre el genoma de referencia y los genomas en la muestra mixta. El análisis de la información de secuenciación para la determinación de la fracción fetal puede permitir un pequeño grado de desajuste en función de la secuencia polimórfica. Por ejemplo, se puede permitir un pequeño grado de desajuste si la secuencia polimórfica es un STR. En los casos en que la secuencia polimórfica es un SNP, todas las secuencias que coinciden exactamente con cualquiera de los dos alelos en el sitio de SNP se cuentan primero y se filtran de las lecturas restantes, por lo que se puede permitir un pequeño grado de falta de coincidencia.

Preparación de la biblioteca de secuenciación

[0121] Secuenciadores de ADN de próxima generación, como los 454-FLX (Roche; en la dirección web 454.com), SOLiD™3 (Applied Biosystems; en la dirección web solid.appliedbiosystems.com), y el Genome Analyzer (Illumina; <http://www.illumina.com/pages.ilmn?ID=204>) han transformado el panorama de la genética a través de su capacidad para producir cientos de megabases de información de secuencia en una sola ejecución.

[0122] Los métodos de secuenciación requieren la preparación de bibliotecas de secuenciación. La preparación de la biblioteca de secuenciación implica la producción de una colección aleatoria de fragmentos de ADN modificados por adaptadores, que están listos para ser secuenciados. Las bibliotecas de secuenciación de polinucleótidos pueden prepararse a partir de ADN o ARN, incluidos equivalentes, análogos de ADN o ADNc, que es complementario o copia de ADN producido a partir de una plantilla de ARN, por ejemplo, por la acción de la transcriptasa inversa. Los polinucleótidos pueden originarse en la forma de ADN de doble cadena (ADNds) (por ejemplo, fragmentos de ADN genómico, productos de PCR y amplificación) o polinucleótidos que pueden haberse originado en forma de cadena simple, como ADN o ARN, y se han convertido en la forma de ADN de doble cadena. A modo de ejemplo, las moléculas de ARNm pueden copiarse en ADNc de doble cadena adecuados para usar en la preparación de una biblioteca de secuenciación. La secuencia precisa de las moléculas de polinucleótido primarias generalmente no es material para el método de preparación de bibliotecas, y puede ser conocida o desconocida. En una realización, las moléculas de polinucleótido son moléculas de ADN. Más particularmente, las moléculas de polinucleótido representan el complemento genético completo de un organismo, y son moléculas de ADN genómico, por ejemplo, moléculas de ADNnc, que incluyen secuencias tanto de intrones como de exones (secuencia codificadora), así como secuencias reguladoras no codificantes, tales como secuencias promotoras y potenciadoras. Aún más particularmente, las moléculas de polinucleótido primarias son moléculas de ADN genómico humano, por ejemplo, moléculas de ADNcf presentes en la sangre periférica de un sujeto preñado. La preparación de bibliotecas de secuenciación para algunas plataformas de secuenciación NGS requiere que los polinucleótidos sean de un rango específico de tamaños de fragmentos, por ejemplo, 0-1200 pb. Por lo tanto, puede requerirse la fragmentación de polinucleótidos, por ejemplo, ADN genómico. ADNcf existe como fragmentos de <300 pares de bases. Por lo tanto, la fragmentación de ADNcf no es necesaria para generar una biblioteca de secuenciación utilizando muestras de ADNcf. Fragmentación de moléculas de polinucleótidos por medios mecánicos *p. ej.* nebulización, sonicación e hidrosial, da como resultado fragmentos con una mezcla heterogénea de extremos romos y extremos sobresalientes en 3' y 5'. Ya sea que los polinucleótidos estén fragmentados a la fuerza o existan naturalmente como fragmentos, se convierten en ADN de extremos romos que tienen 5-fosfatos y 3'-hidroxilo.

[0123] Típicamente, los extremos de los fragmentos se reparan al final, es decir, *con* extremos romos utilizando métodos o kits conocidos en la técnica. Los fragmentos de extremos romos se pueden fosforilar mediante tratamiento enzimático, por ejemplo, utilizando polinucleótido quinasa. En algunas realizaciones, se agrega un único desoxinucleótido, por ejemplo, desoxiadenosina (A) a los extremos 3' de los polinucleótidos, por ejemplo, por la actividad de ciertos tipos de ADN polimerasa, como la Taq polimerasa o Klenow exo menos polimerasa. Los productos con cola de dA son compatibles con el saliente "T" presente en el extremo 3' de cada región dúplex de adaptadores a los que se ligan en una etapa posterior. Cola de dA evita la auto-ligación de ambos polinucleótidos de extremos romos, de manera que existe un sesgo hacia la formación de las secuencias ligadas al adaptador. Los polinucleótidos con cola de dA se ligan a secuencias de polinucleótidos adaptadores de doble cadena. Se puede usar el mismo adaptador para ambos extremos del polinucleótido, o se pueden utilizar dos conjuntos de adaptadores. Los métodos de ligación son conocidos en la técnica y utilizan enzimas ligasa, como la ADN ligasa, para enlazar covalentemente el adaptador al polinucleótido de cola dA. El adaptador puede contener un resto 5'-fosfato para facilitar la ligadura al 3'-OH diana. El polinucleótido con cola de dA contiene un resto 5'-fosfato, ya sea residual del proceso de cizallamiento, o agregado mediante una etapa de tratamiento enzimático, y se ha reparado

al final, y opcionalmente se ha extendido por una base o bases sobresalientes, para dar una 3'-OH adecuada para la ligadura. Los productos de la reacción de ligadura se purifican para eliminar los adaptadores no ligados, adaptadores que pueden estar ligados entre sí, y para seleccionar un rango de tamaños de plantillas para la generación de agrupaciones, que puede ir precedida de una amplificación, por ejemplo, una amplificación por PCR. La purificación de los productos de ligadura se puede obtener mediante métodos que incluyen electroforesis en gel e inmovilización reversible en fase sólida (SPRI).

[0124] Los protocolos estándar, por ejemplo, los protocolos para la secuenciación usando, por ejemplo, la plataforma Illumina, instruir a los usuarios para purificar los productos de reparación de extremos antes de cola de dA, y para purificar los productos de cola de dA antes de las etapas de adaptador de ligadura de la preparación de la biblioteca. La purificación de los productos reparados en el extremo y los productos con cola de dA eliminan las enzimas, tampones, sales y similares para proporcionar condiciones de reacción favorables para la etapa enzimática posterior. En una forma de realización, las etapas de reparación de extremos, ligadura de dA y ligadura del adaptador excluyen las etapas de purificación. Por lo tanto, en una realización, el método de la invención abarca la preparación de una biblioteca de secuenciación que comprende los pasos consecutivos de reparación de extremo, cola de dA y ligadura de adaptador. En las formas de realización para preparar bibliotecas de secuenciación que no requieren la etapa de seguimiento dA, por ejemplo, los protocolos para la secuenciación utilizando las plataformas de Roche 454 y SOLID™3, las etapas de reparación de extremos y ligadura del adaptador excluyen la etapa de purificación de los productos reparados antes de la adaptación de la ligación de adaptador.

[0125] En un paso siguiente de una realización del método, se prepara una reacción de amplificación. La etapa de amplificación introduce en las moléculas de plantilla ligadas por el adaptador las secuencias de oligonucleótidos requeridas para la hibridación con la célula de flujo. Los contenidos de una reacción de amplificación son conocidos por un experto en la técnica e incluyen sustratos apropiados (Como los dNTP), enzimas (por ejemplo, una polimerasa ADN) y componentes de tampón necesarios para una reacción de amplificación. Opcionalmente, se puede omitir la amplificación de polinucleótidos ligados al adaptador. En general, las reacciones de amplificación requieren al menos dos cebadores de amplificación, es decir, oligonucleótidos de cebador, que pueden ser idénticos, e incluyen una "porción específica del adaptador", capaz de hibridar con una secuencia de unión a cebador en la molécula de polinucleótido que se amplificará (o su complemento si la plantilla se ve como una sola hebra) durante el paso de recocido. Una vez formados, la biblioteca de plantillas preparada de acuerdo con los métodos descritos anteriormente se puede usar para la amplificación de ácido nucleico en fase sólida. El término "amplificación en fase sólida", como se usa en el presente documento, se refiere a cualquier reacción de amplificación de ácido nucleico llevada a cabo en o en asociación con un soporte sólido, de manera tal que todos o una parte de los productos amplificados se inmovilicen en el soporte sólido a medida que se forman. En particular, el término abarca la reacción en cadena de la polimerasa en fase sólida (PCR en fase sólida) y la amplificación isotérmica en fase sólida, que son reacciones análogas a la amplificación en fase de solución estándar, excepto que uno o ambos de los cebadores de amplificación directa e inversa se inmovilizan en el soporte sólido. La PCR en fase sólida cubre sistemas como las emulsiones, en donde un cebador se ancla a una perla y el otro se encuentra en solución libre, y la formación de colonias en matrices de gel de fase sólida donde un cebador se ancla a la superficie y el otro se encuentra en solución libre. Después de la amplificación, y las bibliotecas de secuenciación se pueden analizar mediante electroforesis capilar microfluídica para garantizar que la biblioteca esté libre de dímeros adaptadores o ADN de cadena sencilla. La biblioteca de moléculas de polinucleótido de plantilla es particularmente adecuada para uso en métodos de secuenciación en fase sólida. Además de proporcionar plantillas para la secuenciación en fase sólida y la PCR en fase sólida, las plantillas de biblioteca proporcionan plantillas para la amplificación del genoma completo.

[0126] En una realización, la biblioteca de polinucleótidos ligados al adaptador se somete a secuenciación masiva en paralelo, que incluye técnicas para la secuenciación de millones de fragmentos de ácidos nucleicos, por ejemplo, el uso de la unión de ADN genómico fragmentado aleatoriamente a una superficie plana, ópticamente transparente y amplificación de fase sólida para crear una célula de flujo de secuenciación de alta densidad con millones de agrupaciones. Los arreglos agrupados se pueden preparar usando un proceso de termociclado, como se describe en la patente WO9844151, o un proceso por el cual la temperatura se mantiene constante, y los ciclos de extensión y desnaturalización se realizan usando cambios de reactivos. El método de Solexa/Illumina al que se hace referencia aquí se basa en la unión de ADN genómico fragmentado aleatoriamente a una superficie plana, ópticamente transparente. Los fragmentos de ADN adjuntos se extienden y se amplifican en puente para crear una célula de flujo de secuenciación de densidad ultraalta con millones de agrupaciones que contienen miles de copias de la misma plantilla (WO 00/18957 y WO 98/44151). Las plantillas de grupos se secuencian utilizando una robusta tecnología de secuenciación por síntesis de ADN de cuatro colores que emplea terminadores reversibles con tintes fluorescentes removibles. Alternativamente, la biblioteca puede amplificarse en perlas en las que cada perla contiene un cebador de amplificación directa e inversa.

[0127] La secuenciación de las bibliotecas amplificadas se puede llevar a cabo usando cualquier técnica de secuenciación adecuada como se describe en el presente documento. En una realización, la secuenciación es una secuenciación masivamente paralela que usa secuenciación por síntesis con terminadores de tinte reversibles. En otras realizaciones, la secuenciación es una secuenciación masivamente paralela utilizando la secuenciación por ligación. En otras realizaciones, la secuenciación es la secuenciación de una sola molécula.

Determinación de la aneuploidía

[0128] La precisión requerida para determinar correctamente si una aneuploidía está presente o ausente en una muestra, se basa en parte en la variación del número de etiquetas de secuencia que se mapean en el genoma de referencia entre las muestras dentro de una secuenciación (variabilidad inter-cromosómica), y la variación del número de etiquetas de secuencia que se asignan al genoma de referencia en diferentes ejecuciones de secuenciación (variabilidad entre secuencias). Por ejemplo, las variaciones pueden ser particularmente pronunciadas para las etiquetas que se asignan a secuencias de referencia ricas en GC o pobres en GC. En una realización, el método utiliza la información de secuenciación para calcular la dosis de cromosoma, lo que explica de manera intrínseca la variabilidad acumulada derivada de la variabilidad inter-cromosómica, inter-secuenciación y dependiente de la plataforma. Las dosis de cromosomas se determinan a partir de la información de secuenciación, es decir, el número de etiquetas de secuencia, para la secuencia de interés, por ejemplo, el cromosoma 21, y el número de etiquetas de secuencia para una secuencia de normalización. La identificación de una secuencia de normalización se realiza en un conjunto de muestras calificadas que se sabe que no contienen una aneuploidía de la secuencia de interés. El diagrama de flujo proporcionado en la Figura 1 muestra una realización del método **100** mediante el cual se identifican secuencias normalizantes, por ejemplo, cromosomas normalizadores, y se determina la presencia o ausencia de una aneuploidía.

[0129] En el paso **110**, se obtiene un conjunto de muestras maternas calificadas para identificar secuencias de normalización cualificadas, por ejemplo, cromosomas de normalización, y para proporcionar valores de varianza para el uso en la determinación de la identificación estadísticamente significativa de una aneuploidía en muestras de ensayo. En la etapa **110**, una pluralidad de muestras biológicas calificadas se obtienen de una pluralidad de sujetos que se sabe que comprenden células que tienen un número de copias normal para cualquier secuencia de interés, por ejemplo, un cromosoma de interés tal como un 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 para tener un número normal de copias de cromosomas en relación con el cromosoma de interés. Las muestras maternas biológicas calificadas pueden ser muestras de fluidos biológicos, por ejemplo, muestras de plasma, o cualquier muestra adecuada como se describe anteriormente que contenga una mezcla de moléculas de ADNfc fetal y maternal. La muestra es una muestra materna que se obtiene de una mujer embarazada, por ejemplo, una mujer embarazada. Cualquier muestra biológica materna se puede usar como una fuente de ácidos nucleicos fetales y maternos que están contenidos en las células o que están "libres de células". En algunas realizaciones, es ventajoso obtener una muestra materna que comprenda ácidos nucleicos libres de células, por ejemplo, ADNcf. Preferiblemente, la muestra biológica materna es una muestra de fluido biológico. Un fluido biológico incluye, como ejemplos no limitativos, sangre, plasma, suero, sudor, lágrimas, esputo, orina, esputo, flujo de oído, linfa, saliva, líquido cefalorraquídeo, estragos, suspensión de médula ósea, flujo vaginal, lavado transcervical, líquido cerebral, ascitis, leche, secreciones de vías respiratorias, intestinales y genitourinarias, muestras de líquido amniótico y leucoforesis. En algunas realizaciones, la muestra de fluido biológico es una muestra que se puede obtener fácilmente mediante procedimientos no invasivos, por ejemplo, sangre, plasma, suero, sudor, lágrimas, esputo, orina, esputo, flujo del oído y saliva. En algunas realizaciones, la muestra biológica es una muestra de sangre periférica, o el plasma y/o sus fracciones de suero. En otra realización, la muestra es una mezcla de dos o más muestras biológicas, por ejemplo, una muestra biológica puede comprender dos o más muestras de un fluido biológico. Tal como se usa en el presente documento, los términos "sangre", "plasma" y "suero" abarcan expresamente las fracciones o partes procesadas de los mismos. En algunas realizaciones, la muestra biológica se procesa para obtener una fracción de muestra, por ejemplo, plasma, que contiene la mezcla de ácidos nucleicos maternos y fetales. En algunas realizaciones, la mezcla de ácidos nucleicos maternos y fetales se procesa adicionalmente a partir de la fracción de muestra, por ejemplo, plasma, para obtener una muestra que comprende una mezcla purificada de ácidos nucleicos maternos y fetales, por ejemplo, ADNcf. Los ácidos nucleicos libres de células, incluido el ADN libre de células, pueden obtenerse mediante diversos métodos conocidos en la técnica a partir de muestras biológicas que incluyen, entre otras, 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 ADNfc de las células, fraccionamiento, centrifugación (p. ej., Centrifugación en gradiente de densidad), precipitación específica de ADN o clasificación de células y/o métodos de separación de células de alto rendimiento pueden ser utilizados. Los kits disponibles comercialmente para la separación manual y automática de ADNcf están disponibles (Roche Diagnostics, Indianapolis, IN, Qiagen, Valencia, CA, Macherey-Nagel, Duren, DE). En algunos casos, puede ser ventajoso fragmentar las moléculas de ácido nucleico en la muestra de ácido nucleico. La fragmentación puede ser aleatoria o específica, como se logra, por ejemplo, mediante la digestión con endonucleasas de restricción. Los métodos para la fragmentación aleatoria son bien conocidos en la técnica e incluyen, por ejemplo, la digestión con DNAsa limitada, el tratamiento con álcali y la cizalladura física. En una realización, los ácidos nucleicos de la muestra se obtienen a partir de ADNcf, que no está sujeto a fragmentación. En otras realizaciones, los ácidos nucleicos de la muestra se obtienen como ADN genómico, que se somete a fragmentación en fragmentos de aproximadamente 500 o más pares de bases, y a los cuales se pueden aplicar fácilmente los métodos de NGS. Se prepara una biblioteca de secuenciación a partir de ADN naturalmente fragmentado o fragmentado por la fuerza. En una realización, la preparación de la biblioteca de secuenciación comprende los pasos consecutivos de la reparación del extremo, la cola de dA y el ligamiento con adaptador de los fragmentos de ADN. En otra realización, la preparación de la biblioteca de secuenciación comprende los pasos consecutivos de la reparación final y la ligadura adaptadora de los fragmentos de ADN.

[0130] En el paso **120**, al menos una porción de cada uno de todos los ácidos nucleicos calificados contenidos en las muestras maternas calificados están secuenciados. Antes de la secuenciación, la mezcla de ácidos nucleicos fetales y maternos, por ejemplo, ADNcf purificado, se modifica para preparar una biblioteca de secuenciación para generar lecturas de secuencia de entre 20 y 40 pb, por ejemplo, 36 pb, que están alineadas con un genoma de referencia, por ejemplo, hg18. En algunas realizaciones, las lecturas de secuencias comprenden aproximadamente 20 pb, aproximadamente 25 pb, aproximadamente 30 pb, aproximadamente 35 pb, aproximadamente 40 pb, aproximadamente 45 pb, aproximadamente 50 pb, aproximadamente 55 pb, aproximadamente 65 pb, aproximadamente 70 pb, aproximadamente 75 pb, aproximadamente 80 pb, aproximadamente 85 pb, aproximadamente 90 pb, aproximadamente 95 pb, aproximadamente 100 pb, aproximadamente 110 pb, aproximadamente 120 pb, aproximadamente 130, aproximadamente 140 pb, aproximadamente 150 pb, aproximadamente 200 pb, aproximadamente 250 pb, aproximadamente 300 pb, aproximadamente 350 pb, aproximadamente 400 pb, aproximadamente 450 pb, aproximadamente 500 pb. Se espera que los avances tecnológicos permitan lecturas de un solo extremo de más de 500 pb, permitiendo lecturas de más de aproximadamente 1.000 pb cuando se generan lecturas de extremos emparejados. En una realización, las lecturas de secuencia comprenden 36 pb. Las lecturas de secuencia se alinean con un genoma de referencia humano, y las lecturas que se asignan de forma única al genoma de referencia humano se cuentan como etiquetas de secuencia. En una realización, al menos aproximadamente 3×10^6 etiquetas de secuencia calificados, al menos aproximadamente 5×10^6 etiquetas de secuencia calificados, al menos aproximadamente 8×10^6 etiquetas de secuencia calificados, al menos aproximadamente 10×10^6 etiquetas de secuencia calificados, al menos aproximadamente 15×10^6 etiquetas de secuencia calificados, al menos aproximadamente 20×10^6 etiquetas de secuencia calificados, al menos aproximadamente 30×10^6 etiquetas de secuencia calificados, al menos aproximadamente 40×10^6 etiquetas de secuencia calificados, o al menos aproximadamente 50×10^6 las etiquetas de secuencia calificada que comprenden entre 20 y 40 pb de lectura se obtienen de lecturas que se asignan de forma única a un genoma de referencia.

[0131] En el paso **130**, todas las etiquetas obtenidas a partir de la secuenciación de los ácidos nucleicos en las muestras maternas calificados se cuentan para determinar una densidad de etiqueta de secuencia calificada. En una realización, la densidad de la etiqueta de secuencia se determina como el número de etiquetas de secuencia calificados asignadas a la secuencia de interés en el genoma de referencia. En otra realización, la densidad de etiquetas de secuencia calificada se determina como el número de etiquetas de secuencia calificados asignadas a una secuencia de interés normalizada a la longitud de la secuencia de interés calificada a la que se asignan. Las densidades de etiqueta de secuencia que se determinan como una relación de la densidad de etiqueta con respecto a la longitud de la secuencia de interés se denominan en este documento relaciones de densidad de etiqueta. No se requiere la normalización de la longitud de la secuencia de interés, y puede incluirse como un paso para reducir el número de dígitos en un número para simplificarlo para la interpretación humana. Ya que todas las etiquetas de secuencia calificados se mapean y se cuentan en cada una de las muestras calificados, se determina la densidad de la etiqueta de secuencia para una secuencia de interés, por ejemplo, el cromosoma de interés, en las muestras calificados, así como las densidades de la etiqueta de secuencia para secuencias adicionales a partir de las cuales secuencias normalizadas *p. ej.* cromosomas, se identifican posteriormente. En una realización, la secuencia de interés es un cromosoma que está asociado con una aneuploidía cromosómica, por ejemplo, el cromosoma 21, y la secuencia de normalización calificada es un cromosoma que no está asociado con una aneuploidía cromosómica y cuya variación en la densidad de la etiqueta de la secuencia se aproxima mejor que del cromosoma 21. Por ejemplo, una secuencia de normalización calificada es una secuencia que tiene la variabilidad más pequeña. En algunas realizaciones, la secuencia de normalización es una secuencia que distingue mejor una o más muestras calificados de una o más muestras afectadas, es decir, la secuencia de normalización es una secuencia que tiene la mayor diferenciabilidad. El nivel de diferenciabilidad se puede determinar como una diferencia estadística entre las dosis de cromosomas en una población de muestras calificados y las dosis de cromosomas en una o más muestras de prueba. 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 una translocación cromosómica desequilibrada, y la secuencia normalizada es un segmento cromosómico que no está asociado con la aneuploidía parcial y cuya variación en la densidad de la etiqueta de secuencia se aproxima mejor a la del segmento cromosómico asociado con la aneuploidía parcial.

[0132] En el paso **140**, en base a las densidades de etiquetas calificados calculadas, una dosis de secuencia calificada para una secuencia de interés se determina como la relación 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 cuales se identifican posteriormente las secuencias normalizadoras. En una realización, las dosis para el cromosoma de interés, por ejemplo, el cromosoma 21, se determinan como una proporción de la densidad de la etiqueta de secuencia del cromosoma 21 y la densidad de la etiqueta de secuencia para cada uno de los cromosomas restantes, es decir, los cromosomas 1-20, cromosoma 22, cromosoma X y cromosoma Y (véase Ejemplos 3-5 y Figuras 9-15).

[0133] En el paso **145**, una secuencia de normalización, por ejemplo, un cromosoma de normalización, se identifica por una secuencia de interés, por ejemplo el cromosoma 21, en una muestra calificada en base a las dosis de secuencia calculadas. El método identifica secuencias que inherentemente tienen características similares y que son propensas a variaciones similares entre las muestras y las ejecuciones de secuenciación, y que son útiles para

determinar las dosis de secuencia en muestras de prueba. En algunas realizaciones, la secuencia de normalización es la que mejor diferencia una muestra afectada, es decir, una muestra aneuploide, de una o más muestras calificadas. En otras realizaciones, una secuencia de normalización es una secuencia que muestra una variabilidad en el número de etiquetas de secuencia que se asignan a ella entre muestras y ejecuciones de secuenciación que se aproximan mejor a la secuencia de interés para la cual se utiliza como parámetro de normalización, y/o que puede diferenciar mejor una muestra afectada de una o más muestras no afectadas.

[0134] En algunas realizaciones, se identifica más de una secuencia de normalización. Por ejemplo, la variación, por ejemplo, el coeficiente de variación en la dosis de cromosoma para el cromosoma de interés 21 es mínimo cuando se utiliza la densidad de la etiqueta de secuencia del cromosoma 14. En otras realizaciones, se identifican dos, tres, cuatro, cinco, seis, siete, ocho o más secuencias de normalización para uso en la determinación de una dosis de secuencia para una secuencia de interés en una muestra de prueba.

[0135] En una realización, la secuencia de normalización para el cromosoma 21 se selecciona a partir 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 de normalización para el cromosoma 21 se selecciona del cromosoma 9, cromosoma 1, cromosoma 2, cromosoma 11, cromosoma 12 y cromosoma 14. Alternativamente, la secuencia de normalización 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 de normalización para el cromosoma 21 es un grupo de cromosomas seleccionados del cromosoma 9, cromosoma 1, cromosoma 2, cromosoma 11, cromosoma 12 y cromosoma 14.

[0136] En una realización, la secuencia de normalización para el cromosoma 18 se selecciona cromosoma 8, cromosoma 2, cromosoma 3, cromosoma 4, cromosoma 5, cromosoma 6, cromosoma 7, cromosoma 9, cromosoma 10, cromosoma 11, cromosoma 12, cromosoma 12 y el cromosoma 14. Preferiblemente, la secuencia de normalización para el cromosoma 18 es cromosoma 8, el cromosoma 2, el cromosoma 3, el cromosoma 5, el cromosoma 6, el cromosoma 12 y el cromosoma 14 seleccionados. Alternativamente, la secuencia de normalización para el cromosoma 18 es un grupo de los 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 de normalización para el cromosoma 18 es un grupo de cromosomas seleccionados del cromosoma 8, cromosoma 2, cromosoma 3, cromosoma 5, cromosoma 6, cromosoma 12, y el cromosoma 14.

[0137] En una realización, la secuencia de normalización para el cromosoma X se selecciona entre el cromosoma 1, el cromosoma 2, el cromosoma 3, el cromosoma 4, el cromosoma 4, el cromosoma 7, el cromosoma 8, el cromosoma 8, el cromosoma 8, el cromosoma 12, el cromosoma 12, el cromosoma 12 y el cromosoma 15. y cromosoma 16. Preferiblemente, la secuencia de normalización para el cromosoma X se selecciona entre el cromosoma 2, el cromosoma 3, el cromosoma 4, el cromosoma 5, el cromosoma 6 y el cromosoma 8. Alternativamente, la secuencia de normalización para el cromosoma X es un grupo de los 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 de normalización para el cromosoma X es un grupo de cromosomas seleccionados entre el cromosoma 2, el cromosoma 3, el cromosoma 4, el cromosoma 5, el cromosoma 6 y el cromosoma 8.

[0138] En una realización, la secuencia de normalización 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 de normalización para el cromosoma 13 se selecciona de entre el 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 de normalización para el cromosoma 13 es un grupo de cromosomas seleccionados entre el cromosoma 2, el cromosoma 3, el cromosoma 4, el cromosoma 5, el cromosoma 6 y el cromosoma 8.

[0139] La variación en la dosis de cromosoma para el cromosoma Y es mayor que 30 independientemente de que se utiliza el cromosoma normalizador en la determinación de 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 pueden usar como la secuencia de normalización para el cromosoma Y. En una realización, el al menos un cromosoma de normalización es un grupo de cromosomas que consiste en los cromosomas 1-22 y el cromosoma X. En otra realización, el al menos un cromosoma normalizador es un grupo de cromosomas seleccionados entre el cromosoma 2, el cromosoma 3, el cromosoma 4, el cromosoma 5 y el cromosoma 6.

[0140] Basado en la identificación de la(s) secuencia(es) de normalización en muestras calificadas, una dosis de secuencia se determina para una secuencia de interés en una muestra de prueba que comprende una mezcla de ácidos nucleicos derivados de genomas que difieren en una o más secuencias de interés.

[0141] En el paso **115**, una muestra de ensayo por ejemplo, muestra de plasma, que comprende ácidos nucleicos fetales y maternos por ejemplo ADNcf, se obtiene de un sujeto embarazado por ejemplo, una mujer embarazada, para la que debe determinarse la presencia o ausencia de una aneuploidía fetal.

[0142] Una biblioteca de secuenciación se preparó como se describe para el paso 120, y en el paso **125**, al menos una porción de los ácidos nucleicos de prueba en la muestra de ensayo se secuencian para generar millones de lecturas de secuencia que comprende entre 20 y 500 pb por ejemplo de 36 pb. como en el paso **120**, las lecturas generadas a partir de la secuenciación de los ácidos nucleicos en la muestra de prueba se mapean de forma única a un genoma de referencia humano y se cuentan. Como se describe en el paso **120**, al menos aproximadamente 3×10^6 etiquetas de secuencia calificadas, al menos aproximadamente 5×10^6 etiquetas de secuencia calificadas, al menos aproximadamente 8×10^6 etiquetas de secuencia calificadas, al menos aproximadamente 10×10^6 etiquetas de secuencia calificadas, al menos aproximadamente 15×10^6 etiquetas de secuencia calificadas, al menos aproximadamente 20×10^6 etiquetas de secuencia calificadas, al menos aproximadamente 30×10^6 etiquetas de secuencia calificadas, al menos aproximadamente 40×10^6 etiquetas de secuencia calificadas, o al menos aproximadamente 50×10^6 etiquetas de secuencia calificadas que comprenden entre 20 y 40 lecturas de pb se obtienen de lecturas que se asignan de forma única al genoma de referencia humano.

[0143] En el paso **135**, todas las etiquetas obtenidas a partir de la secuenciación de los ácidos nucleicos en las muestras de ensayo se cuentan para determinar una densidad de etiqueta de secuencia de prueba. En una realización, el número de etiquetas de secuencia de prueba mapeadas a una secuencia de interés se normaliza a la longitud conocida de una secuencia de interés a la que se mapean para proporcionar una densidad de etiqueta de secuencia de prueba. 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 el número de dígitos en un número para simplificarlo para la interpretación humana. Ya que todas las etiquetas de secuencia de prueba mapeadas se cuentan en la muestra de prueba, se determina la densidad de etiqueta de secuencia para una secuencia de interés, por ejemplo, una secuencia clínicamente relevante, como el cromosoma 21, en las muestras de prueba, al igual que las densidades de etiqueta de secuencia para secuencias adicionales que corresponden a al menos una secuencia de normalización identificada en las muestras calificadas.

[0144] En el paso **150**, basándose en la identidad de al menos una secuencia de normalización en las muestras calificadas, una dosis de secuencia de prueba se determina para una secuencia de interés en la muestra de ensayo. La dosis de secuencia, por ejemplo, la dosis de cromosoma, para una secuencia de interés en una muestra de prueba es una relación de la densidad de la etiqueta de secuencia determinada para la secuencia de interés en la muestra de prueba y la densidad de la etiqueta de secuencia de al menos una secuencia de normalización determinada en la muestra de prueba, en donde la secuencia de normalización en la muestra de prueba corresponde a la secuencia de normalización identificada en las muestras calificadas para la secuencia particular de interés. Por ejemplo, si se determina que la secuencia de normalización identificada para el cromosoma 21 en las muestras calificadas es el cromosoma 14, entonces la dosis de secuencia de prueba para el cromosoma 21 (secuencia de interés) se determina como la relación de la densidad de la etiqueta de secuencia para el cromosoma 21 y la densidad de la etiqueta de secuencia para el cromosoma 14 se determinó en la muestra de prueba. De manera similar, se determinan las dosis de cromosomas para los cromosomas 13, 18, X, Y y otros cromosomas asociados con aneuploidías cromosómicas. como se describió anteriormente, una secuencia de interés puede ser parte de un cromosoma, por ejemplo, un segmento de cromosoma. Por consiguiente, la dosis para un segmento cromosómico se puede determinar como la proporción de la densidad de la etiqueta de secuencia determinada para el segmento en la muestra de prueba y la densidad de la etiqueta de secuencia para el segmento de cromosoma normalizado en la muestra de prueba, en donde el segmento de normalización en la muestra de prueba corresponde al segmento de normalización identificado en las muestras calificadas para el segmento de interés particular.

[0145] En el paso **155**, los valores de umbral son derivados de valores de desviación estándar establecidos para 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, el tipo de aneuploidía. Preferiblemente, los umbrales se eligen de la distribución empírica para cada tipo de aneuploidía, por ejemplo, trisomía 21. Los posibles valores de umbral que se establecieron para clasificar la trisomía 13, la trisomía 18, la trisomía 21 y las aneuploidías de monosomía X como se describe en los Ejemplos, que describen el uso de método para determinar las aneuploidías cromosómicas mediante la secuenciación del ADNcf extraído de una muestra materna que comprende una mezcla de ácidos nucleicos fetales y maternos.

[0146] En el paso **160**, la variación del número de copias de la secuencia de interés por ejemplo, aneuploidía cromosómica o parcial, se determina en la muestra de ensayo mediante la comparación de la dosis de secuencia de prueba para la secuencia de interés a al menos un valor umbral establecido a partir de la secuencia cualificado dosis.

[0147] En el paso 160, la dosis calculada para una secuencia de prueba de interés se compara con el conjunto como los valores de umbral que se eligen de acuerdo con un umbral definido por el usuario de fiabilidad para clasificar la muestra como "normal", "afectada" o "no llamada" en el paso 165. Las muestras "sin llamada" son muestras para las que no se puede hacer un diagnóstico definitivo con fiabilidad.

[0148] Otra realización descrita en el presente documento es un método para proporcionar el diagnóstico prenatal de una aneuploidía cromosómica fetal en una muestra biológica que comprende moléculas de ácido nucleico fetal y materno. El diagnóstico se realiza sobre la base de recibir los datos de la secuenciación de al menos una parte de la mezcla de las moléculas de ácido nucleico materno y fetal derivadas de una muestra de prueba biológica, por ejemplo, una muestra de plasma materno, computando a partir de los datos de secuenciación una dosis de cromosoma normalizante para uno o más cromosomas de interés, determinando una diferencia estadísticamente significativa entre la dosis de cromosoma normalizante para el cromosoma de interés en la muestra de prueba y un valor umbral establecido en una pluralidad de muestras calificadas (normales), y proporcionando el diagnóstico prenatal basado en la diferencia estadística. como se describe en el paso 165 del método, se realiza un diagnóstico de normal o afectado. Se proporciona una "no llamada" en caso de que el diagnóstico de normal o afectado no se pueda realizar con confianza.

Determinación de la NVC para diagnósticos prenatales.

[0149] El ADN fetal libre de células y el ARN que circula en la sangre materna se pueden usar para el diagnóstico prenatal no invasivo temprano (NIPD) de un número creciente de afecciones genéticas, tanto para el manejo del embarazo como para ayudar a la toma de decisiones reproductivas. La presencia de ADN libre de células que circula en el torrente sanguíneo se conoce desde hace más de 50 años. Más recientemente, se descubrió la presencia de pequeñas cantidades de ADN fetal circulante en el torrente sanguíneo materno durante el embarazo (Lo et al., Lancet 350: 485-487 [1997]). Se cree que se origina a partir de células placentarias moribundas, el ADN fetal libre de células (ADNcf) consiste en fragmentos cortos que suelen tener menos de 200 pb de longitud (Chan et al., Clin Chem 50: 88-92 [2004]), que pueden ser discernidos desde las 4 semanas de gestación (Illanes et al., Early Human Dev 83: 563-566 [2007]), y se sabe que se eliminó de la circulación materna a las pocas horas de la administración (Lo et al., Am J Hum Genet 64: 218-224 [1999]). Además de ADNcf, los fragmentos de RNA fetal libre de células (ARNcf) también pueden discernirse en el torrente sanguíneo materno, que se originan a partir de genes que se transcriben en el feto o la placenta. La extracción y el posterior análisis de estos elementos genéticos fetales a partir de una muestra de sangre materna ofrece nuevas oportunidades para la NIPD.

[0150] El presente método es un método independiente de polimorfismo que para su uso en NIPD y que no requiere que el ADNcf fetal se distinga de la ADNcf materna 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. Por el momento, la aneuploidía conocida más común compatible con la vida es la trisomía 21, es decir, el síndrome de Down (SD), que se debe a la presencia de una parte o la totalidad del cromosoma 21. En raras ocasiones, la SD puede ser causada por un defecto hereditario o esporádico, por el cual se presenta un defecto adicional. Una copia de todo o parte del cromosoma 21 se une a otro cromosoma (generalmente el cromosoma 14) para formar un solo cromosoma aberrante. La SD se asocia con deterioro intelectual, graves dificultades de aprendizaje y el exceso de mortalidad causado por problemas de salud a largo plazo, como las enfermedades del corazón. Otras aneuploidías con importancia clínica conocida incluyen el síndrome de Edward (trisomía 18) y el síndrome de Patau (trisomía 13), que con frecuencia son fatales en los primeros meses de vida. Las anomalías asociadas con la cantidad de cromosomas sexuales también son conocidas e incluyen monosomía X, por ejemplo, síndrome de Turner (XO) y síndrome de triple X (XXX) en partos femeninos y síndrome de Klinefelter (XXY) y síndrome de XYY en partos masculinos, todos los cuales están asociados con diversos fenotipos incluyendo la esterilidad y la reducción de las habilidades intelectuales. El método de la divulgación se puede utilizar para diagnosticar estas y otras anomalías cromosómicas antes del nacimiento.

[0151] Según las realizaciones de la presente divulgación, la trisomía 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 XXY (Síndrome de Klinefelter), XYY o XXX trisomías. Se apreciará que varias otras trisomías y trisomías parciales se pueden determinar en ADNcf fetal de acuerdo con las enseñanzas de la presente invención. Estos incluyen, pero no se limitan a, trisomía parcial 1q32-44, trisomía 9 p con trisomía, mosaicismo de trisomía 4, trisomía 17p, trisomía parcial 4q26-qter, trisomía 9, trisomía parcial 2p, trisomía parcial 1q y/o trisomía parcial 6p/monosomía 6q.

[0152] El método de la presente descripción también se pueden utilizar para determinar monosomía X cromosómica, y monosomías parciales tales como, monosomía 13, monosomía 15, monosomía 16, monosomía 21, y monosomía 22, que son conocidas por estar involucradas en aborto involuntario. La monosomía parcial de los cromosomas que participan típicamente en la aneuploidía completa también se puede determinar por el método de la invención. La

monosomía 18p es un trastorno cromosómico raro en el que se elimina todo (o parte) del brazo corto (p) del cromosoma 18 (monosómico). El trastorno se caracteriza típicamente por estatura baja, grados variables de retraso mental, retrasos en el habla, malformaciones del cráneo y la región facial (craneofacial) y/o anomalías físicas adicionales. Los defectos craneofaciales asociados pueden variar mucho en rango y severidad de un caso a otro.

Las afecciones causadas por los cambios en la estructura o el número de copias del cromosoma 15 incluyen el síndrome de Angelman y el síndrome de Prader-Willi, que implican una pérdida de la actividad genética en la misma parte del cromosoma 15, la región 15q11-q13. Se apreciará que varias translocaciones y microdeleciones pueden ser asintomáticas en el progenitor portador, pero pueden causar una enfermedad genética importante en la descendencia. Por ejemplo, una madre sana que lleva la microdelección 15q11-q13 puede dar a luz a un niño con síndrome de Angelman, un trastorno neurodegenerativo grave. Por lo tanto, la presente invención se puede usar para identificar tal supresión en el feto. La monosomía parcial 13q es un trastorno cromosómico raro que se produce cuando falta una parte del brazo largo (q) del cromosoma 13 (monosómico). Los bebés que nacen con monosomía parcial 13q pueden presentar bajo peso al nacer, malformaciones de la cabeza y la cara (región craneofacial), anomalías esqueléticas (especialmente de las manos y los pies) y otras anomalías físicas. El retraso mental es característico de esta condición. La tasa de mortalidad durante la infancia es alta entre las personas nacidas con este trastorno. Casi todos los casos de monosomía parcial 13q ocurren al azar sin ninguna razón aparente (esporádica). El síndrome de deleción 22q11,2, también conocido como síndrome de DiGeorge, es un síndrome causado por la eliminación de una pequeña pieza del cromosoma 22. La eliminación (22 q11.2) ocurre cerca de la mitad del cromosoma en el brazo largo de uno de los par de cromosomas. Las características de este síndrome varían ampliamente, incluso entre los miembros de la misma familia, y afectan a muchas partes del cuerpo. Los signos y síntomas característicos pueden incluir defectos de nacimiento, como cardiopatía congénita, defectos en el paladar, más comúnmente relacionados con problemas neuromusculares con el cierre (insuficiencia velo-faríngea), problemas de aprendizaje, diferencias leves en la características faciales, e infecciones recurrentes. Las microdeleciones en la región cromosómica 22q11,2 se asocian con un riesgo 20 a 30 veces mayor de esquizofrenia.

En una realización, el método de la invención se usa para determinar monosomías parciales que incluyen pero no se limitan a monosomía 18p, monosomía parcial del cromosoma 15 (15q11-q13), monosomía parcial 13q y monosomía parcial del cromosoma 22 también puede determinarse usando método. El Ejemplo 6 y la Figura 16 ilustran el uso del método de la divulgación para determinar esa presencia de una eliminación parcial del cromosoma 11.

[0153] El método de la divulgación también se puede utilizar para determinar cualquier aneuploidía si uno de los padres es un portador conocido de dicha anomalía. Estos incluyen, pero no se limitan a, mosaico para un pequeño cromosoma marcador supernumerario (SMC); t(11;14)(p15;p13) translocación; translocación desequilibrada t(8;11)(p23.2;p15.5); microdelección 11q23; supresión del síndrome de Smith-Magenis 17p11.2; deleción 22q13.3; microdelección Xp22.3; supresión 10p14; microdelección 20p, síndrome de Di-George [del(22)(q11.2q11.23)], síndrome de Williams (deleciones 7q11,23 y 7q36); supresión 1p36; micro-eliminación 2p; neurofibromatosis tipo 1 (17q11,2 microdelección), deleción Yq; síndrome de Wolf-Hirschhorn (WHS, microdelección 4p16.3); 1p36,2 microdelección; supresión 11q14; 19q13,2 microdelección; Rubinstein-Taybi (16 p13,3 microdelección); microdelección 7p21; síndrome de Miller-fieker (17p13.3), deleción 17p11.2; y microdelección 2q37.

Determinación de la NVC de trastornos clínicos.

[0154] Además de la determinación temprana de defectos de nacimiento, los métodos descritos en el presente documento se pueden aplicar a la determinación de cualquier anomalía en la representación de secuencias genéticas dentro del genoma. Se ha demostrado que el plasma sanguíneo y el ADN sérico de pacientes con cáncer contienen cantidades medibles de ADN tumoral, que pueden recuperarse y usarse como fuente sustituta de ADN tumoral. Los tumores se caracterizan por aneuploidía, o números inapropiados de secuencias de genes o incluso cromosomas completos. La determinación de una diferencia en la cantidad de una secuencia dada, es decir, una secuencia de interés, en una muestra de un individuo puede, por lo tanto, usarse en el diagnóstico de una condición médica, por ejemplo, cáncer.

[0155] Las realizaciones de la divulgación proporcionan un método para evaluar la variación del número de copias de una secuencia de interés, por ejemplo, una secuencia clínicamente relevante, en una muestra de prueba que comprende una mezcla de ácidos nucleicos derivados de dos genomas diferentes, y que son conocidos o se sospecha que difieren en la cantidad de una o más secuencias de interés. La mezcla de ácidos nucleicos se deriva de dos o más tipos de células. En una realización, la mezcla de ácidos nucleicos se deriva de células normales y cancerosas derivadas de un sujeto que padece una afección médica, por ejemplo, cáncer.

[0156] Se cree que muchos tumores sólidos, como el cáncer de mama, progresan desde el inicio hasta la metástasis a través de la acumulación de varias aberraciones genéticas. [Sato et al., Cancer Res., 50: 7184-7189 [1990]; Jongsma et al., J Clin Pathol: Mol Path 55: 305-309 [2002]]. Dichas aberraciones genéticas, a medida que se acumulan, pueden conferir ventajas proliferativas, inestabilidad genética y la capacidad concomitante para desarrollar rápidamente la resistencia al fármaco, y una mayor angiogénesis, proteólisis y metástasis. Las aberraciones genéticas pueden afectar a los "genes supresores de tumores" recesivos o a los oncogenes de acción dominante. Se cree que las supresiones y la recombinación que conducen a la pérdida de heterocigosidad (LOH) desempeñan un papel importante en la progresión del tumor al descubrir alelos supresores de tumores mutados.

[0157] ADNcf se ha encontrado en la circulación de pacientes diagnosticados con tumores malignos que incluyen pero no se limitan a cáncer de pulmón (Pathak et al. Clin Chem 52: 1833-1842 [2006]), cáncer de próstata (Schwartzentbach et al. Clin Cancer Res 15: 1032-8 [2009]), y cáncer de mama (Schwartzentbach et al. Disponible en línea en breast-cancer-research.com/COntent/11/5/R71 [2009]). La identificación de las inestabilidades genómicas asociadas con los cánceres que pueden determinarse en la circulación del ADNcf en pacientes con cáncer es una herramienta potencial de diagnóstico y pronóstico. En una realización, el método de la divulgación evalúa la CNV de una secuencia de interés en una muestra que comprende una mezcla de ácidos nucleicos derivados de un sujeto que se sospecha o se sabe que tiene cáncer, por ejemplo, carcinoma, sarcoma, linfoma, leucemia, tumores de células germinales y blastoma. En una realización, la muestra es una muestra de plasma derivada (procesos) de sangre periférica y que comprende una mezcla de ADNcf derivado de células normales y cancerosas. En otra realización, la muestra biológica que se necesita para determinar si una CNV está presente se deriva de una mezcla de células cancerosas y no cancerosas de otros fluidos biológicos que incluyen, entre otros, suero, sudor, lágrimas, esputo, orina, esputo. flujo del oído, linfa, saliva, líquido cefalorraquídeo, estragos, suspensión de la médula ósea, flujo vaginal, lavado transcervical, líquido cerebral, ascitis, leche, secreciones de las vías respiratorias, tracto intestinal y genitourinario, y muestras de leucoforesis, o en biopsias de tejidos, hisopos o frotis.

[0158] La secuencia de interés es una secuencia de ácido nucleico que se sabe o se sospecha que desempeñar un papel en el desarrollo y/o progresión del cáncer. Los ejemplos de una secuencia de interés incluyen secuencias de ácidos nucleicos que se amplifican o eliminan en células cancerosas como se describe a continuación.

[0159] Los genes de acción dominante asociados con tumores sólidos humanos típicamente ejercen su efecto por sobreexpresión o expresión alterada. La amplificación de genes es un mecanismo común que conduce a la regulación positiva de la expresión génica. La evidencia de los estudios citogenéticos indica que se produce una amplificación significativa en más del 50% de los cánceres de mama humanos. En particular, la amplificación del receptor 2 del factor de crecimiento epidérmico humano protooncogénico (HER2) ubicado en el cromosoma 17 (17(17q21-q22)) produce una sobreexpresión de los receptores HER2 en la superficie celular, lo que lleva a una señalización excesiva y desregulada en cáncer de mama y otras neoplasias malignas (Park et al., Clinical Breast Cancer 8: 392-401 [2008]). Se ha encontrado que una variedad de oncogenes se amplifican en otras neoplasias malignas humanas. Ejemplos de la amplificación de oncogenes celulares en tumores humanos incluyen amplificaciones de: c-myc en la línea celular HL60 de leucemia promielocítica, y en líneas celulares de carcinoma de pulmón de células pequeñas, N-myc en neuroblastomas primarios (etapas III y IV), líneas celulares de neuroblastoma, línea celular de retinoblastoma y tumores primarios, líneas de carcinoma de pulmón de células pequeñas y tumores, L-myc en líneas de células de carcinoma de pulmón de células pequeñas y tumores, c-myb en leucemia mieloide aguda y en líneas de células de carcinoma de colon, c-erbB en células de carcinoma epidermoide, y gliomas primarios, cK-ras-2 en carcinomas primarios de pulmón, vejiga y recto, N-ras en la línea celular de carcinoma mamario (Varmus H., Ann Rev Genetics 18: 553-612 (1984) [citado en Watson et al. al., Molecular Biology of the Gene (4ª ed.; Benjamin/Cummings Publishing co. 1987]).

[0160] Deleciones cromosómicas que implican genes supresores de tumores pueden desempeñar un papel importante en el desarrollo y progresión de tumores sólidos. El gen supresor de tumores más ampliamente caracterizado es el gen supresor de tumores del retinoblastoma (Rb-1), ubicado en el cromosoma 13q14. El producto del gen Rb-1, una fosfoproteína nuclear de 105 kDa, aparentemente desempeña un papel importante en la regulación del ciclo celular (Howe et al., Proc Natl Acad Sci (EE.UU.) 87: 5883-5887 [1990]). La expresión alterada o pérdida de la proteína Rb es causada por la inactivación de ambos alelos genéticos ya sea a través de una mutación puntual o una deleción cromosómica. Se ha encontrado que las alteraciones del gen Rb-1 están presentes no solo en los retinoblastomas sino también en otras neoplasias malignas como los osteosarcomas, el cáncer de pulmón de células pequeñas (Rygaard et al., Cancer Res 50: 5312-5317 [1990]) y el cáncer de mama. Los estudios de polimorfismo de longitud de fragmentos de restricción (RFLP, por sus siglas en inglés) han indicado que estos tipos de tumores con frecuencia pierden heterocigosidad a 13q, lo que sugiere que uno de los alelos del gen Rb-1 se ha perdido debido a una deleción cromosómica macroscópica (Bowcock et al., Am J Hum. Genet. 46: 12 [1990]). Las anomalías del cromosoma 1, incluidas las duplicaciones, las eliminaciones y las translocaciones desequilibradas que involucran el cromosoma 6 y otros cromosomas asociados, indican que las regiones del cromosoma 1, en particular 1q21-1q32 y 1p11-13, pueden albergar oncogenes o genes supresores de tumores que son patogénicamente relevantes tanto para la enfermedad crónica como para la avanzada. Fases de las neoplasias mieloproliferativas (Caramazza et al., Eur J Hematol 84: 191-200 [2010]). Las neoplasias mieloproliferativas también se asocian con deleciones del cromosoma 5. La pérdida completa o las deleciones intersticiales del cromosoma 5 son las anomalías cariotípicas más comunes en los síndromes mielodisplásicos (SMD). Los pacientes con del(5q)/5q-MDS aislada tienen un pronóstico más favorable que aquellos con defectos cariotípicos adicionales, que tienden a desarrollar neoplasmas mieloproliferativos (MPN) y leucemia mieloide aguda. La frecuencia de las eliminaciones desequilibradas del cromosoma 5 ha llevado a la idea de que 5q alberga uno o más genes supresores de tumores que tienen funciones fundamentales en el control del crecimiento de las células madre/progenitoras hematopoyéticas (HSCs/HPC). El mapeo citogenético de las regiones comúnmente eliminadas (CDR) centradas en 5q31 y 5q32 identificaron genes supresores de tumores candidatos, que incluyen la subunidad ribosomal RPS14, el factor de transcripción Egr1/Krox20 y la proteína de remodelación del citoesqueleto, alfa-catenina (Eisenmann et al., OnCOgene 28): 3429-3441 [2009]). Los estudios citogenéticos y alelotipados de tumores frescos y líneas celulares tumorales han demostrado que la pérdida alélica de varias regiones distintas en el cromosoma 3p, incluidas 3p25,

3p21-22, 3p21.3, 3p12-13 y 3p14, son las anomalías genómicas más tempranas y más frecuentes involucradas en un amplio espectro de cánceres epiteliales principales de pulmón, mama, riñón, cabeza y cuello, ovario, cuello uterino, colon, páncreas, esófago, vejiga y otros órganos. Varios genes supresores de tumores se han mapeado en la región del cromosoma 3p, y se cree que las deleciones intersticiales o la hipermetilación del promotor preceden a la pérdida del 3p o todo el cromosoma 3 en el desarrollo de carcinomas (Angeloni D., Briefings Functional Genomics 6:19- 39 [2007]).

[0161] Los recién nacidos y los niños con síndrome de Down (SD) a menudo presentan leucemia congénita transitoria y tienen un mayor riesgo de leucemia mieloide aguda y leucemia linfoblástica aguda. El cromosoma 21, que alberga alrededor de 300 genes, puede estar involucrado en numerosas aberraciones estructurales, por ejemplo, translocaciones, deleciones y amplificaciones, en leucemias, linfomas y tumores sólidos. Además, se han identificado genes ubicados en el cromosoma 21 que juegan un papel importante en la tumorigénesis. Las aberraciones somáticas tanto del cromosoma 21 como del cromosoma se asocian con leucemias, y genes específicos que incluyen RUNX1, TMRSS2 y TFF, que se encuentran en 21q, desempeñan un papel en la tumorigénesis (gen de Fonatsch C, cromosomas, cáncer, 49: 497-508 [2010]).

[0162] En una realización, el método se refiere a un medio para evaluar la asociación entre la amplificación génica y el alcance de la evolución del tumor. La correlación entre la amplificación y/o la eliminación y la etapa o grado de un cáncer puede ser importante desde el punto de vista del pronóstico, ya que dicha información puede contribuir a la definición de un grado tumoral de base genética que predice mejor el curso futuro de la enfermedad con tumores más avanzados que tienen el peor pronóstico. Además, la información sobre la amplificación temprana y/o los eventos de eliminación puede ser útil para asociar esos eventos como predictores de la progresión posterior de la enfermedad. La amplificación de genes y las deleciones identificadas por el método pueden asociarse con otros parámetros conocidos como el grado del tumor, la histología, el índice de marcación Brd/Urd, el estado hormonal, la afectación nodal, el tamaño del tumor, la duración de la supervivencia y otras propiedades tumorales disponibles en estudios epidemiológicos y bioestadísticos. Por ejemplo, el ADN tumoral que se analizará con el método podría incluir hiperplasia atípica, carcinoma ductal in situ, cáncer en etapa I-III y ganglios linfáticos metastásicos para permitir la identificación de asociaciones entre amplificaciones y deleciones y etapa. Las asociaciones realizadas pueden posibilitar una efectiva intervención terapéutica. Por ejemplo, las regiones amplificadas consistentemente pueden contener un gen sobreexpresado, cuyo producto puede ser atacado terapéuticamente (por ejemplo, el receptor del factor de crecimiento de quinasa de tirosina, p185^{HER2}).

[0163] El método puede ser utilizado para identificar eventos de amplificación y/o deleción que están asociados con la resistencia a fármacos mediante la determinación de la variación del número de copias de ácidos nucleicos a partir de cánceres primarios a los de células que han producido metástasis en otros sitios. Si la amplificación y/o eliminación de genes es una manifestación de inestabilidad cariotípica que permite un rápido desarrollo de la resistencia al fármaco, se esperaría más amplificación y/o eliminación en tumores primarios de pacientes quimiorresistentes que en tumores en pacientes quimiosensibles. Por ejemplo, si la amplificación de genes específicos es responsable del desarrollo de la resistencia al fármaco, se esperaría que las regiones que rodean a esos genes se amplifiquen consistentemente en células tumorales de derrames pleurales de pacientes quimiorresistentes pero no en los tumores primarios. El descubrimiento de asociaciones entre la amplificación y/o eliminación de genes y el desarrollo de resistencia farmacológica puede permitir la identificación de pacientes que se beneficiarán o no de la terapia adyuvante.

Determinación simultánea de aneuploidía y fracción fetal

[0164] En otra realización, el método permite la determinación simultánea de la fracción del componente de ácido nucleico fetal menor, es decir, fracción fetal, en una muestra que comprende una mezcla de ácidos nucleicos fetales y maternos. En particular, el método permite la determinación de la fracción de ADNcf contribuido por un feto a la mezcla de ADNcf 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 contribución relativa de un alelo polimórfico derivado del genoma fetal a la contribución del correspondiente alelo polimórfico derivado del genoma materno. Las secuencias polimórficas se pueden usar junto con pruebas de diagnóstico clínicamente relevantes como control positivo de la presencia de ADNcf para resaltar los resultados falsos negativos o falsos positivos derivados de niveles bajos de ADNcf por debajo del límite de identificación. El método descrito es útil en un rango de edades gestacionales.

[0165] Las formas de realización ejemplares del método para determinar simultáneamente la fracción fetal y la presencia o ausencia de una aneuploidía se representan en las **Figuras 2-5** de la siguiente manera.

[0166] La **Figura 2** proporciona un diagrama de flujo de una realización del método de la divulgación **200** para determinar simultáneamente una aneuploidía fetal y la fracción de ácidos nucleicos fetales en una muestra biológica materna. En la etapa **210**, se obtiene de un sujeto una muestra de prueba que comprende una mezcla de ácidos nucleicos maternos y fetales. Las muestras de prueba incluyen muestras descritas en la etapa **110** de la realización del método **100**. En algunas realizaciones, la muestra de prueba es una muestra de sangre periférica obtenida de una hembra embarazada, por ejemplo, una mujer. En la etapa **220**, la mezcla de ácidos nucleicos presente en la muestra se enriquece para ácidos nucleicos diana polimórficos, comprendiendo cada uno un sitio polimórfico. En

algunas realizaciones, los ácidos nucleicos que están enriquecidos son ADNcf. Los ácidos nucleicos diana son segmentos de material genético que se sabe que comprenden al menos un sitio polimórfico. En algunas realizaciones, los ácidos nucleicos diana comprenden un SNP. En otras realizaciones, el ácido nucleico diana comprende un STR. En otras realizaciones más, los ácidos nucleicos diana comprenden un STR en tándem. El enriquecimiento de una mezcla de ácidos nucleicos maternos y fetales comprende la amplificación de secuencias diana de una porción de ácidos nucleicos contenidos en la muestra materna original, y la combinación parcial o total del producto amplificado con el resto de la muestra materna original. En la etapa **230**, al menos una porción de la mezcla enriquecida se secuencía, se identifican las diferencias de secuencia derivadas de la naturaleza polimórfica de las secuencias diana, y la contribución relativa de las secuencias polimórficas derivadas del genoma fetal, es decir, la fracción fetal, se determina en la etapa **240**. En algunas realizaciones, la muestra de prueba 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 ADNcf fetal y materno purificado.

Secuencias polimorfas

[0167] Los sitios polimórficos que están contenidos en los ácidos nucleicos diana incluyen, sin limitación, polimorfismos de un solo nucleótido (SNP), SNP en tándem, delecciones o inserciones multibásicas a pequeña escala, llamadas IN-DELS (también llamados polimorfismos de inserción de eliminación o DIP), polimorfismos multinucleótidos (MNP) y repeticiones cortas en tándem (STR). Los sitios polimórficos que se incluyen en el método de la enfermedad se ubican en cromosomas autosómicos, lo que permite 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 este documento se puede usar para determinar simultáneamente la fracción fetal y la presencia o ausencia de una aneuploidía en una muestra materna.

[0168] En una realización, la mezcla de ácidos nucleicos fetales y maternos en la muestra está enriquecida en ácidos nucleico diana que comprenden al menos un SNP. En algunas realizaciones, cada ácido nucleico diana comprende un al menos un SNP. Las secuencias de ácido nucleico diana que comprenden SNP están disponibles en bases de datos de acceso público que incluyen, entre otras, la base de datos de SNP humano en la dirección de Internet de wi.mit.edu, página NCBI dbSNP en la dirección de Internet ncbi.nlm.nih.gov, en la dirección de Internet lifesciences.perkinelmer.com, la base de datos de SNP de Celera Human en la dirección de Internet de gran alcance celera.com, la base de datos de SNP del Grupo de Análisis del Genoma (GAN) en la dirección de Internet gan.iarc.fr. En una realización, los SNP elegidos para enriquecer el ADNcf fetal y materno se seleccionan del grupo de 92 SNP de identificación individual (IISNPs) descritos por Pakstis *et al.* (Pakstis *et al.* Hum Genet 127: 315-324 [2010]), que ha demostrado tener una variación muy pequeña en la frecuencia entre las poblaciones ($F_{st} < 0,06$), y es altamente informativo en todo el mundo con una heterocigosidad promedio $\geq 0,4$. Los SNP que están abarcados por el método de la divulgación incluyen los SNP vinculados y no vinculados. Cada ácido nucleico diana comprende al menos un sitio polimórfico, por ejemplo, un solo SNP, que difiere del presente en otro ácido nucleico diana para generar un panel de sitios polimórficos, por ejemplo SNP, que contienen un número suficiente de sitios polimórficos de los 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 SNP puede configurarse para que incluya al menos un SNP informativo.

[0169] En una realización, las fuentes que son seleccionadas para la amplificación se seleccionan de entre rs560681, rs7041158, rs70478, rs7409098, rs7409098, rs740908, rs7409098, rs740598, rs7409098, rs740598, rs1277284, rs258684, rs1347696, rs508485, rs9788670, rs8137254, rs3143, rs2182957, rs3739005 y rs530022.

[0170] En otras realizaciones, cada ácido nucleico diana comprende dos o más SNPs es decir, cada diana de ácido nucleico comprende SNPs en tándem. Preferiblemente, cada ácido nucleico diana comprende dos SNP en tándem. Los SNP en tándem se analizan como una sola unidad como haplotipos cortos, y se proporcionan aquí como conjuntos de dos SNP. Para identificar secuencias SNP en tándem adecuadas, se puede buscar en la base de datos del consorcio Internacional HapMap (The International HapMap Project, Nature 426: 789-796 [2003]). La base de datos está disponible en la web en hapmap.org. En una realización, los SNP en tándem que se dirigen a la amplificación se seleccionan de los siguientes conjuntos de pares en 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.

[0171] En otra realización, la mezcla de ácidos nucleicos fetales y maternos en la muestra está enriquecida en ácidos nucleicos diana que comprenden al menos un STR. Los loci STR se encuentran en casi todos los cromosomas del genoma y pueden amplificarse utilizando una variedad de cebadores de reacción en cadena de la polimerasa (PCR). Las repeticiones de tetranucleótidos se han preferido entre los científicos forenses debido a su fidelidad en la amplificación por PCR, aunque también se utilizan algunas repeticiones de tri- y pentanucleótidos. Se compila una lista completa de referencias, hechos e información de secuencias sobre los STR, los cebadores de PCR publicados, los sistemas multiplex comunes y los datos de población relacionados en STRBase, a la que se puede acceder a través del Internet en ibm4.carb.nist.gov:8800/dna/home.htm. También se puede acceder a la información de secuencia de GenBank® (<http://www2.ncbi.nlm.nih.gov/cgi-bin/genbank>) para los loci STR de uso común a través de STRBase. La naturaleza polimórfica de las secuencias de ADN repetidas en tándem que se extienden por todo el genoma humano las ha convertido en importantes marcadores genéticos para los estudios de mapeo de genes, análisis de ligamiento y pruebas de identidad humana. Debido al alto polimorfismo de los STR, la mayoría de los individuos serán heterocigotos, es decir, la mayoría de las personas poseerán dos alelos (versiones) de cada uno heredado de cada padre, con un número diferente de repeticiones. Por lo tanto, la secuencia de STR fetal no hereditaria por maternidad diferirá en el número de repeticiones de la secuencia materna. La amplificación de estas secuencias de STR dará como resultado dos productos principales de amplificación correspondientes a los alelos maternos (y el alelo fetal de herencia materna) y un producto secundario correspondiente al alelo fetal de herencia no materna. Esta técnica se informó por primera vez en 2.000 (Pertl et al., Human Genetics 106: 45-49 [2002]) y posteriormente se desarrolló utilizando la identificación simultánea de múltiples regiones STR diferentes mediante 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 se puede determinar mediante la secuenciación de ácidos nucleicos diana polimórficos que comprenden STR, que varían entre los individuos en el número de unidades repetidas en tándem entre alelos. En una realización, la determinación simultánea de aneuploidía y fracción fetal comprende la secuenciación de al menos una porción de los ácidos nucleicos maternos y fetales presentes en una muestra materna que se ha enriquecido para secuencias polimórficas que comprenden STR. Dado que el tamaño del ADNcf fetal es <300 pb, las secuencias polimórficas comprenden miniSTR, que se puede amplificar para generar amplicones que tienen longitudes aproximadamente del tamaño de los fragmentos de ADN fetal circulantes. El método puede usar uno o una combinación de cualquier número de miniSTR informativos para determinar la fracción de ácido nucleico fetal. Por ejemplo, se puede utilizar cualquiera o una combinación de cualquier número de miniSTR, por ejemplo, los miniSTR descritos en la Tabla 22. En una realización, la fracción de ácido nucleico fetal en una muestra materna se realiza utilizando un método que incluye la determinación del número de copias del ácido nucleico materno y fetal presente en la muestra materna amplificando al menos una miniSTR autosómica elegida de CSF1PO, FGA, TH01, TPOX, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, Penta D, Penta E, D2S1338, D1S1677, D2S441, D4S2364, D10S1248, D14S1434, D22S1045, D22S1045, D20S1082, D20S482, D18S853, D17S1301, D17S974, D14S1434, D12ATA63, D11S4463, D10S1435, D10S1248, D9S2157, D9S1122, D8S1115, D6S1017, D6S474, D5S2500, D5S2500, D4S2408, D4S2364, D3S4529, D3S3053, D2S1776, D2S441, D1S1677, D1S1627, y D1GATA113. En otra realización, el al menos un miniSTR autosómico es el grupo de miniSTR CSF1PO, FGA, D13S317, D16S539, D18S51, D2S1338, D21S11 y D7S820.

[0172] El enriquecimiento de la muestra para los ácidos nucleicos diana se realiza mediante métodos que comprenden específicamente la amplificación de secuencias de ácido nucleico diana que comprenden el sitio polimórfico. La amplificación de las secuencias diana se puede realizar mediante cualquier método que utilice PCR o variaciones del método que incluyen, entre otros, la PCR asimétrica, la amplificación dependiente de helicasa, la PCR de arranque en caliente, la qPCR, la PCR en fase sólida y la PCR de toma de contacto. Alternativamente, la replicación de secuencias de ácido nucleico diana puede obtenerse mediante métodos independientes de la enzima, por ejemplo, síntesis química en fase sólida utilizando las fosforamiditas. La amplificación de las secuencias diana se logra utilizando pares de cebadores, cada uno capaz de amplificar una secuencia de ácido nucleico diana 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 diana amplificados que comprenden al menos dos, al menos tres, al menos 5, al menos 10, al menos 15, al menos 20, al menos 25, al menos 30, al menos 30, al menos 35, al menos 40 o más sitios polimórficos en la misma reacción de secuenciación. Cualquier panel de conjuntos de cebadores puede configurarse para amplificar al menos una secuencia polimórfica informativa.

Amplificación de secuencias polimórficas.

[0173] Un número de cebadores de ácidos nucleicos están ya disponibles para amplificar fragmentos de ADN que contienen polimorfismos de SNP y sus secuencias se pueden obtener, por ejemplo, de las bases de datos identificadas anteriormente. También se pueden diseñar cebadores adicionales, por ejemplo, utilizando un método similar al publicado por Vieux, EF, Kwok, PY y Miller, RD en BioTechniques (junio de 2002), vol. 32. Suplemento: "SNPs: Discovery of Marker Disease, pp. 28-32. En una realización, 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 se eligen para amplificar un ácido nucleico diana que comprende al menos un SNP informativo en una porción de una mezcla de ADNcf fetal y materno. Los

conjuntos son de: los primos incluyen los cebadores directos e inversos que abarcan al menos uno de los siguientes: los tipos de pasajes se incluyen entre los siguientes: los elementos que se incluyen entre los siguientes: los elementos que se incluyen entre los siguientes: rs9951171, rs338882, rs10776839, rs9905977, rs1277284, rs258684, rs1347696, rs508485, rs9788670, rs8137254, rs3182957, rs3182957, rs3182957 y rs3182957 y los conjuntos de cebadores que se utilizan para amplificar los SNP descritos en este documento se proporcionan en el Ejemplo 7 y en las Tablas 10 y 11, y se describen como SEQ ID NO: 57-112. En otra realización, el grupo de 13 conjuntos de cebadores SEQ ID NO: 57-82 se usa para amplificar un ácido nucleico diana que comprende al menos un SNP, por ejemplo, un solo SNP, en una porción de una mezcla de ADNcf fetal y materno.

[0174] En otra realización, al menos un conjunto de cebadores se usa para amplificar un ácido nucleico diana que comprenden cada uno al menos un tándem SNP por ejemplo, un conjunto de dos SNP en tándem, en una porción de una mezcla de ADNcf fetal y materna. 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 utilizados para amplificar las secuencias de destino que comprenden los SNP en tándem están diseñados para abarcar ambos sitios de SNP. Conjuntos ejemplares de cebadores utilizados para amplificar los SNP en tándem descritos en este documento se proporcionan en el Ejemplo 12 y se describen como SEQ ID NO: 197-310.

[0175] La amplificación de los ácidos nucleicos diana se realiza usando cebadores específicos de secuencia que permiten la amplificación específica de secuencia. Por ejemplo, los cebadores de PCR están diseñados para discriminar contra la amplificación de genes o parálogos similares que se encuentran en otros cromosomas aprovechando las diferencias de secuencia entre el ácido nucleico diana y cualquier parálisis de otros cromosomas. Los cebadores de PCR directos o inversos están diseñados para ser recogidos cerca del sitio de SNP y para amplificar una secuencia de ácido nucleico de longitud suficiente como para abarcarse en las lecturas generadas por métodos de secuenciación masivamente paralelos. Algunos métodos de secuenciación masivamente paralelos requieren que la secuencia de ácido nucleico tenga una longitud mínima (pb) para permitir la amplificación de puente que puede usarse opcionalmente antes de la secuenciación. Por lo tanto, los cebadores de PCR utilizados para amplificar los ácidos nucleicos diana están diseñados para amplificar secuencias que son de longitud suficiente para amplificarse en puente e identificar los SNP que están comprendidos en las lecturas de secuencia. En algunas realizaciones, el primero de los dos cebadores en el conjunto de cebadores que comprende el cebador directo y el inverso para amplificar el ácido nucleico diana está diseñado para identificar un único SNP presente dentro de una secuencia leída de aproximadamente 20 pb, aproximadamente 25 pb, aproximadamente 30 pb, aproximadamente 35 pb, aproximadamente 40 pb, aproximadamente 45 pb, aproximadamente 50 pb, aproximadamente 55 pb, aproximadamente 60 pb, aproximadamente 65 pb, aproximadamente 70 pb, aproximadamente 75 pb, aproximadamente 80 pb, aproximadamente 85 pb, aproximadamente 90 pb, aproximadamente 95 pb, aproximadamente 100 pb, aproximadamente 110 pb, aproximadamente 120 pb, aproximadamente 140 pb., aproximadamente 150 pb, aproximadamente 200 pb, aproximadamente 250 pb, aproximadamente 300 pb, aproximadamente 350 pb, aproximadamente 400 pb, aproximadamente 450 pb, o aproximadamente 500 pb. Se espera que los avances tecnológicos en tecnologías de secuenciación masivamente paralelas permitan lecturas de un solo extremo de más de 500 pb. En una realización, uno de los cebadores de PCR está diseñado para amplificar los SNP que se incluyen en lecturas de secuencia de 36 pb. El segundo cebador está diseñado para amplificar el ácido nucleico diana como un amplicón de longitud suficiente para permitir la amplificación del puente. En una realización, los cebadores de PCR a modo de ejemplo están diseñados para amplificar los ácidos nucleicos diana que contienen un solo SNP seleccionado 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. En otras realizaciones, los cebadores directo e inverso están diseñados para amplificar ácidos nucleicos diana, cada uno de los cuales comprende un conjunto de dos SNP en tándem, cada uno de los cuales está presente dentro de una secuencia leída de aproximadamente 20 pb, aproximadamente 25 pb, aproximadamente 30 pb, aproximadamente 35 pb, aproximadamente 40 pb, aproximadamente 45 pb, aproximadamente 50 pb, aproximadamente 55 pb, aproximadamente 60 pb, aproximadamente 65 pb, aproximadamente 70 pb, aproximadamente 75 pb, aproximadamente 80 pb, aproximadamente 85 pb, aproximadamente 90 pb, aproximadamente 95 pb, aproximadamente 100 pb, aproximadamente 110 pb, aproximadamente 120 pb, aproximadamente 130 pb, aproximadamente 150 pb, aproximadamente 150 pb, aproximadamente 200 pb, aproximadamente 250 pb, aproximadamente 300 pb, aproximadamente 350 pb,

aproximadamente 400 pb, aproximadamente 450 pb, o aproximadamente 500 pb. En una realización, al menos uno de los cebadores está diseñado para amplificar el ácido nucleico diana que comprende un conjunto de dos SNP en tándem como un amplicón de longitud suficiente para permitir la amplificación del puente.

[0176] Los SNP, SNP únicos o en tándem, están contenidos en amplicones de ácido nucleico diana amplificados de al menos aproximadamente 100 pb, al menos aproximadamente 150 pb, al menos aproximadamente 200 pb, al menos aproximadamente 250bp, al menos aproximadamente 300 pb, al menos aproximadamente 350 bp, o al menos aproximadamente 400 pb. En una realización, los ácidos nucleicos diana que comprenden un sitio polimórfico, por ejemplo, un SNP, se amplifican como amplicones de al menos aproximadamente 110 pb, y que comprenden un SNP dentro de 36 pb desde el extremo 3' o 5' del amplicón. En otra realización, los ácidos nucleicos diana que comprenden dos o más sitios polimórficos, por ejemplo, dos SNP en tándem, se amplifican como amplicones de al menos aproximadamente 110 pb, y que comprenden el primer SNP dentro de 36 pb desde el extremo 3' del amplicón, y/o el segundo SNP dentro de 36 pb desde el extremo 5' del amplicón.

[0177] En una realización, 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 se eligen para amplificar un ácido nucleico diana que comprende al menos un SNP en tándem informativo en una porción de una mezcla de ADNcf fetal y materno.

Amplificación de STR

[0178] Un número de cebadores de ácidos nucleicos están ya disponibles para amplificar fragmentos de ADN que contienen los RTS y sus secuencias se pueden obtener, por ejemplo, de las bases de datos identificadas anteriormente. Se han utilizado amplicones de PCR de varios tamaños para discernir las respectivas distribuciones de tamaño de las especies de ADN materno y fetal circulantes, y han demostrado que las moléculas de ADN fetal en el plasma de las mujeres embarazadas son generalmente más cortas que las moléculas de ADN maternas (Chan et al., Clin Chem. 50: 8892 [2004]). El fraccionamiento por tamaño del ADN fetal circulante ha confirmado que la longitud promedio de los fragmentos de ADN fetal circulante es <300 pb, mientras que el ADN materno se ha estimado entre aproximadamente 0,5 y 1 Kb (Li et al., Clin Chem, 50: 1002-1011 [2004]). Estos hallazgos son consistentes con los de Fan *et al.*, quienes determinaron mediante el uso de NGS que el ADNcf fetal rara vez es >340 pb (Fan et al., Clin Chem 56: 1279-1286 [2010]). El método de la divulgación abarca la determinación de la fracción de ácido nucleico fetal en una muestra materna que se ha enriquecido con ácidos nucleicos diana, comprendiendo cada uno un miniSTR que comprende cuantificar al menos un alelo fetal y uno materno en un miniSTR polimórfico, que puede amplificarse para generar amplicones que tienen longitudes aproximadamente del tamaño de los fragmentos de ADN fetal circulantes.

[0179] En una realización, el método comprende determinar el número de copias de al menos un alelo fetal y al menos materno al menos en un miniSTR polimórfico que se amplifica para generar amplicones que son menos de aproximadamente 300 pb, menos de aproximadamente 250 pb. bp, menos de aproximadamente 200 bp, menos de aproximadamente 150 bp, menos de aproximadamente 100 bp, o menos de aproximadamente 50 bp. En otra realización, los amplicones que se generan al amplificar los miniSTR son menores que aproximadamente 300 pb. En otra realización, los amplicones que se generan al amplificar los miniSTR son menores que aproximadamente 250 pb. En otra realización, los amplicones que se generan al amplificar los miniSTR son menores que aproximadamente 200 pb. La amplificación del alelo informativo incluye el uso de cebadores miniSTR, que permiten la amplificación de los amplicones de tamaño reducido para discernir los alelos STR que son menos de aproximadamente 500 pb, menos de aproximadamente 450 pb, menos de aproximadamente 400 pb, menos de aproximadamente 350 pb, menos de unos 300 pares de bases (pb), menos de unos 250 pb, menos de unos 200 pb, menos de unos 150 pb, menos de unos 100 pb o menos de unos 50 pb. Los amplicones de tamaño reducido generados utilizando los cebadores miniSTR se conocen como miniSTR que se identifican de acuerdo con el nombre del marcador correspondiente al lugar al que se asignaron. En una realización, los cebadores miniSTR incluyen cebadores miniSTR que han permitido la reducción máxima de tamaño en el tamaño del amplicón para los 13 loci coDIS STR además del D2S1338, Pentady pentaE que se encuentran en los kits STR disponibles comercialmente (Butler et al., J Forensic Sci 48: 1054-1064 [2003]), loci miniSTR que no están vinculados a los marcadores coDIS como lo describen Coble y Butler (Coble y Butler, J Forensic Sci 50: 43-53 [2005]), y otros miniSTR que se han caracterizado en el NIST. La información sobre los miniSTRs caracterizados en NIST se puede acceder a través de Internet en cstl.nist.gov/biotech/strbase/newSTRs.htm. Se puede usar cualquier par o una combinación de dos o más pares de cebadores miniSTR para amplificar al menos un miniSTR. Por ejemplo, al menos un conjunto de cebadores se selecciona de los conjuntos de cebadores proporcionados en la Tabla 22 (Ejemplo 11) y se describe como SEQ ID NO: 113-196 se pueden usar para amplificar secuencias diana polimórficas que comprenden un STR.

[0180] El enriquecimiento de la muestra se obtiene amplificando los ácidos nucleicos diana contenidos en una porción de la mezcla de ácidos nucleicos maternos y fetales en la muestra original, y combinando al menos una porción o todo el producto amplificado con el resto de la muestra original no amplificada. El enriquecimiento comprende amplificar los ácidos nucleicos diana que están contenidos en una porción de muestra de fluido biológico. En una realización, la muestra que se enriquece es la fracción plasmática de una muestra de sangre

(consulte la **Figura 3**). Por ejemplo, una porción de una muestra de plasma materno original se usa para amplificar secuencias de ácido nucleico diana. Posteriormente, parte o la totalidad del producto amplificado se combina con la muestra de plasma original no amplificada restante, lo que lo enriquece (véase Ejemplo 10). En otra realización, la muestra que se enriquece es la muestra de ADNcf purificado que se extrae del plasma (consulte la **Figura 4**). Por ejemplo, el enriquecimiento comprende amplificar los ácidos nucleicos diana que están contenidos en una porción de una muestra original de mezcla purificada de ácidos nucleicos maternos y fetales, por ejemplo, ADNcf que se ha purificado de una muestra de plasma materno, y posteriormente combinar algunos o todos los amplificados producto con la muestra purificada original sin amplificar restante (véase Ejemplo 9). En otra realización más, la muestra que está enriquecida es una muestra de biblioteca de secuenciación preparada a partir de una mezcla purificada de ácidos nucleicos maternos y fetales (véase la **Figura 5**). Por ejemplo, el enriquecimiento comprende amplificar los ácidos nucleicos diana que están contenidos en una porción de una muestra original de mezcla purificada de ácidos nucleicos maternos y fetales, por ejemplo, el ADNcf que se ha purificado de una muestra de plasma materno, preparando una primera biblioteca de secuenciación de secuencias de ácido nucleico no amplificadas, preparando una segunda biblioteca de secuenciación de ácidos nucleicos diana polimórficos amplificados y, posteriormente, combinando parte o la totalidad de la segunda biblioteca de secuenciación con alguna o toda la primera biblioteca de secuenciación (véase Ejemplo 8). La cantidad de producto amplificado que se utiliza para enriquecer la muestra original se selecciona para obtener información de secuencia suficiente para determinar la presencia o ausencia de aneuploidía y la fracción fetal de la misma serie de secuencia. Al menos aproximadamente el 3%, al menos aproximadamente el 5%, al menos aproximadamente el 7%, al menos aproximadamente el 10%, al menos aproximadamente el 15%, al menos aproximadamente el 20%, al menos aproximadamente el 25%, al menos aproximadamente el 30% o más del número total de etiquetas de secuencia obtenidas de la secuenciación se asigna para determinar la fracción fetal.

[0181] En una realización, la etapa de enriquecer la mezcla de ácidos nucleicos fetales y maternos para los ácidos nucleicos diana polimórficos comprende la amplificación de los ácidos nucleicos diana en una porción de una muestra de ensayo, por ejemplo una muestra de ensayo de plasma, y la combinación de todos o una parte del producto amplificado con la muestra de prueba de plasma restante. La realización del método **300** se representa en el diagrama de flujo que se muestra en la **Figura 3**. En el paso **310**, una muestra de prueba, por ejemplo, una muestra de fluido biológico, tal como una muestra de sangre, se obtiene de una mujer embarazada, y en el paso **320** una porción del ADNcf contenida en la fracción plasmática de la muestra de sangre se utiliza para amplificar ácidos nucleicos diana que comprenden sitios polimórficos, por ejemplo, SNP. En una realización, al menos aproximadamente el 1%, al menos aproximadamente el 1,5%, al menos aproximadamente el 2%, al menos aproximadamente el 10% del plasma materno se usó para amplificar los ácidos nucleicos diana. En el paso **330**, una porción o la totalidad de los ácidos nucleicos diana amplificados se combinan con la mezcla de ADNcf fetal y materno presente en la muestra materna, y el ADNcf combinado y los ácidos nucleicos amplificados se purifican en el paso **340**, y se usan para preparar una biblioteca que se secuenció en el paso **350**. La biblioteca se preparó a partir de ADNcf purificado y que comprende al menos aproximadamente el 10%, al menos aproximadamente el 15%, al menos aproximadamente el 20%, al menos aproximadamente el 25%, al menos aproximadamente el 30%, al menos aproximadamente el 35%, al menos aproximadamente el 40%, al menos aproximadamente el 45%, o al menos aproximadamente el 50% del producto amplificado. En el paso **360**, se analizan los datos de las ejecuciones de secuenciación y se realiza la determinación simultánea de la fracción fetal y la presencia o ausencia de aneuploidía.

[0182] En una realización, el paso de enriquecer la mezcla de ácidos nucleicos fetales y maternos para los ácidos nucleicos diana polimórficos comprende una pluralidad de ácidos nucleicos diana polimórficos en una porción de una mezcla de ácidos nucleicos fetales y maternos purificados a partir de una muestra de ensayo materna. En una realización, una porción de una mezcla de ácidos nucleicos maternos y fetales, por ejemplo, el ADNc, purificado de una muestra de plasma materno se usa para amplificar secuencias de ácido nucleico polimórfico, y una porción del producto amplificado se combina con la mezcla no amplificada de ácidos nucleicos fetales y maternos purificados, *p. ej.*, ADNcf (véase **Figura 4**). La realización del método **400** se representa en el diagrama de flujo que se muestra en la **Figura 4**. En el paso **410**, una muestra de prueba, por ejemplo, una muestra de fluido biológico, como una muestra de sangre, que comprende una mezcla de ácidos nucleicos maternos y fetales se obtiene de una mujer embarazada, y la mezcla de ácidos nucleicos maternos y fetales se purifica a partir de la fracción plasmática en el paso **420**. Como se describió anteriormente, los métodos para la separación de ADN libre de células del plasma son bien conocidos. En el paso **430**, una porción del ADNcf contenido en la muestra purificada se usa para amplificar ácidos nucleicos diana que comprenden sitios polimórficos, por ejemplo, SNP. Al menos aproximadamente el 5%, al menos aproximadamente el 10%, al menos aproximadamente el 15%, al menos aproximadamente el 20%, al menos aproximadamente el 25%, al menos aproximadamente el 30%, al menos aproximadamente el 35%, al menos aproximadamente el 40%, al menos aproximadamente el 45%, o al menos aproximadamente el 50% de ADNcf purificado se utiliza para amplificar los ácidos nucleicos diana. Preferiblemente, la amplificación de las secuencias diana se puede realizar mediante cualquier método que use PCR o variaciones del método que incluyen pero no se limitan a PCR asimétrica, amplificación dependiente de helicasa, PCR de arranque en caliente, qPCR, PCR en fase sólida y PCR de toma de contacto. En el paso **440**, una porción, por ejemplo, al menos aproximadamente el 0,01% del producto amplificado se combina con la muestra de ADNc purificada no amplificada, y la mezcla de ácidos nucleicos maternos y fetales amplificados y no amplificados se secuencian en el paso **450**. En una realización, se realiza la secuenciación utilizando cualquiera de las tecnologías NGS. En el paso **460**, se analizan los datos de las

secuencias de secuenciación y se realiza 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.

[0183] En otra realización, el paso **220** de enriquecer la mezcla de ácidos nucleicos fetales y maternos para los ácidos nucleicos diana polimórficos (**Figura 2**) comprende la combinación de al menos una porción de una primera biblioteca de secuenciación de moléculas de ácido nucleico fetal y materno no amplificados con al menos una parte de una segunda biblioteca de secuenciación de ácidos nucleicos diana polimórficos amplificados. Por lo tanto, la muestra que se enriquece es la muestra de la biblioteca que se prepara para la secuenciación (**Figura 5**). El enriquecimiento de la muestra de la biblioteca para los ácidos nucleicos diana se realiza mediante métodos que comprenden amplificar específicamente las secuencias de ácido nucleico que comprenden el sitio polimórfico. En el paso **510**, una muestra de prueba, por ejemplo, una muestra de fluido biológico tal como una muestra de sangre, que comprende una mezcla de ácidos nucleicos maternos y fetales, se obtiene de una mujer embarazada, y la mezcla de ácidos nucleicos maternos y fetales se purifica de la fracción plasmática en el paso **520**. En el paso **530**, una porción del ADNcf contenida en la muestra purificada se usa para amplificar ácidos nucleicos diana que comprenden sitios polimórficos, por ejemplo, SNP. Al menos aproximadamente el 5%, al menos aproximadamente el 10%, al menos aproximadamente el 15%, al menos aproximadamente el 20%, al menos aproximadamente el 25%, o al menos aproximadamente el 30% del ADNcf purificado se usa para amplificar secuencias de ácido nucleico diana. Preferiblemente, la amplificación de las secuencias diana se puede realizar mediante cualquier método que use PCR o variaciones del método que incluyen, entre otros, la PCR asimétrica, la amplificación dependiente de la helicasa, la PCR de arranque en caliente, la qPCR, la PCR en fase sólida y la PCR de toma de contacto. En el paso **540**, los ácidos nucleicos diana amplificados que comprenden los sitios polimórficos, por ejemplo, SNP, se usan para preparar una biblioteca de secuenciación de ácido nucleico diana. De manera similar, la porción de ADNcf no amplificado purificado se usa para preparar una biblioteca de secuenciación primaria en el paso **550**. En el paso **560**, una porción de la biblioteca diana se combina con la biblioteca primaria generada a partir de la mezcla no amplificada de ácidos nucleicos, y la mezcla de los ácidos nucleicos fetales y maternos comprendidos en las dos bibliotecas se secuencian en el paso **570**. La biblioteca enriquecida comprende al menos aproximadamente el 5%, al menos aproximadamente el 10%, al menos aproximadamente el 15%, al menos aproximadamente el 20%, o al menos aproximadamente el 25% de la biblioteca diana. En el paso **580**, se analizan los datos de las secuencias de secuenciación y se realiza 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 la aneuploidía a partir de la secuenciación de bibliotecas enriquecidas

[0184] La presencia o ausencia de aneuploidía se determina a partir de secuenciación de la biblioteca enriquecida para secuencias polimórficas diana como se describe para la biblioteca no enriquecida descrita en el método **100**.

Determinación de la fracción fetal a partir de la secuenciación de bibliotecas enriquecidas

[0185] La determinación de la fracción fetal en los pasos **240 (Figura 2)**, **360 (Figura 3)**, **480 (Figura 4)**, y **580 (Figura 5)** se basa en el número total de etiquetas que se asignan al primer alelo y el número total de etiquetas que se asignan al 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 del genoma de referencia humano NCBI36/hg18, o el genoma de referencia comprende la secuencia del genoma de referencia humano NCBI36/hg18 y un genoma de secuencias diana artificial, que incluye las secuencias polimórficas diana. En una realización, el genoma artificial de la diana abarca secuencias polimórficas que comprenden los SNPs rs560681, rs1109037, rs13182883, rs13218440, rs1277284, rs258684, rs1347696, rs508485, rs9788670, rs8137254, rs3143, rs2182957, rs3739005 y rs530022. En una realización, el genoma artificial incluye las secuencias diana polimórficas de las SEQ ID NO: 1-56. En otra realización, el genoma artificial incluye las secuencias diana polimórficas de las SEQ ID NO: 1-26 (véase Ejemplo 7). En otra forma de realización, el genoma artificial de diana abarca secuencias polimórficas que comprenden STR seleccionados 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 más, el genoma diana artificial abarca secuencias polimórficas que comprenden uno o más SNP en tándem (SEQ ID NO: 1-56). La composición del genoma de las secuencias diana artificiales variará dependiendo de las secuencias polimórficas que se utilizan para determinar la fracción fetal. Por consiguiente, un genoma de secuencias diana artificiales no está limitado a las secuencias SNP o STR ejemplificadas en este documento.

[0186] El sitio informativo polimórfico por ejemplo SNP se identifica por la diferencia en las secuencias alélicas y la cantidad de cada uno de los posibles alelos. El ADNcf fetal está presente en una concentración que es <10% del ADNcf materno. Por lo tanto, la presencia de una contribución menor de un alelo a la mezcla de ácidos nucleicos maternos y fetales en relación con la contribución principal del alelo materno puede ser asignada al feto. Los alelos que se derivan del genoma materno se denominan aquí alelos principales, y los alelos que se derivan del genoma fetal se denominan aquí alelos menores. Los alelos que están representados por niveles similares de etiquetas de

secuencia mapeadas representan alelos maternos. Los resultados de un ejemplo de amplificación múltiple de ácidos nucleicos diana que comprenden SNP y derivados de una muestra de plasma materno se muestran en la **Figura 18**. Los SNP informativos se distinguen del cambio de un solo nucleótido en un sitio polimórfico predeterminado, y los alelos fetales se distinguen por su contribución menor relativa a la mezcla de ácidos nucleicos maternos y fetales en la muestra en comparación con la contribución principal a la mezcla por los ácidos nucleicos maternos, es decir, las secuencias de SNP son informativas cuando la madre es heterocigótica y está presente un tercer alelo paterno, lo que permite una comparación cuantitativa entre el alelo heredado maternalmente y el alelo heredado paternalmente para calcular la fracción fetal. Por consiguiente, la abundancia relativa de ADNcf fetal en la muestra materna se determina como un parámetro del número total de etiquetas de secuencia únicas mapeadas a la secuencia de ácido nucleico diana en un genoma de referencia para cada uno de los dos alelos del sitio polimórfico predeterminado. En una realización, la fracción de ácidos nucleicos fetales en la mezcla de ácidos nucleicos maternos y fetales se calcula para cada uno de los alelos informativos (alelo_x) de la siguiente manera:

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

y la fracción fetal para la muestra se calcula como el promedio de la fracción fetal de todos los alelos informativos. **[0187]** Opcionalmente, 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) como sigue:

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

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

[0188] La fracción fetal porcentual 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 promedio determinada para al menos 3 alelos informativos.

[0189] De manera similar, la fracción fetal se puede calcular a partir del número de etiquetas asignadas a los alelos SNP en tándem, como se hace para los SNP únicos, pero teniendo en cuenta las etiquetas asignadas a los dos alelos SNP en tándem x e y están presentes en cada una de las secuencias de ácido nucleico polimórficas amplificadas que se amplifican para enriquecer las muestras, es decir

$$\% \text{ de fracción fetal alelo}_{x+y} = ((\sum \text{etiquetas de secuencia fetal para alelo}_{x+y}) / (\sum \text{etiquetas de secuencia materna para alelo}_{x+y})) \times 100$$

[0190] Opcionalmente, la fracción de ácidos nucleicos fetales en la mezcla de ácidos nucleicos maternos y fetales se calcula para cada uno de los alelos informativos (alelo_{x-y}) de la siguiente manera:

$$\% \text{ de fracción fetal alelo}_{x+y} = ((2 \times \sum \text{etiquetas de secuencia fetal para alelo}_{x-y}) / (\sum \text{etiquetas de secuencia materna para alelo}_{x-y})) \times 100,$$

[0191] Para compensar la presencia de 2 conjuntos de alelos fetales en tándem, uno de ellos enmascarado por el fondo materno. Las secuencias SNP en tándem son informativas cuando la madre es heterocigótica y está presente un tercer haplotipo paterno, lo que permite una comparación cuantitativa entre el haplotipo heredado maternalmente y el haplotipo heredado paternalmente para calcular la fracción fetal.

[0192] La fracción fetal puede determinarse a partir de bibliotecas de secuenciación que comprenden secuencias amplificadas polimórficas diana que comprenden STR contando el número de etiquetas asignadas a un alelo mayor (materna) y menor (fetal). Las etiquetas comprenden secuencias de longitud suficiente para abarcar los alelos STR. Los alelos STR informativos pueden dar como resultado una o dos secuencias de etiquetas principales correspondientes a los alelos maternos (y el alelo fetal de herencia materna) y una secuencia de etiquetas secundaria correspondiente al alelo fetal de herencia no materna. La fracción fetal se calcula como una proporción del número de etiquetas asignadas a los alelos fetales y maternos.

Determinación de la fracción fetal por secuenciación masivamente paralela

[0193] Además de usar el presente método para determinar simultáneamente la fracción fetal y la aneuploidía, la fracción fetal se puede determinar independientemente de la determinación de aneuploidía como se describe aquí, pero se puede determinar de forma independiente y/o junto con otros métodos utilizados para la determinación de la aneuploidía, como los métodos descritos en las publicaciones de solicitud de patente US2010/0112575A1, US2009/0087847A1; US2009/0029377A1; US2008/0220422A1; US2008/0138809A1, US2008/0153090A1, y la patente de EE.UU. 7.645.576. El método para determinar la fracción fetal también se puede combinar con ensayos para determinar otras afecciones prenatales asociadas con la madre y/o el feto. Por ejemplo, el método se puede

utilizar junto con los análisis prenatales, por ejemplo, como se describe en las publicaciones de solicitud de patente de EE.UU. Números US2010/0112590A1, US2009/0162842A1, US2007/0207466A1 y US2001/0051341A1.

[0194] La Figura 6 muestra un diagrama de flujo de una realización del método de la enfermedad para determinar la fracción de ácidos nucleicos fetales en una muestra biológica materna mediante secuenciación masivamente paralela de ácidos nucleicos diana polimorfos amplificados por PCR independientemente de la determinación simultánea de aneuploidía. El método comprende la secuenciación de una biblioteca de secuenciación de ácido nucleico diana polimórfica como sigue. En el paso **610** se obtiene de un sujeto una muestra materna 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. Otras muestras maternas pueden ser de mamíferos, por ejemplo, vaca, caballo, perro o gato. Si el sujeto es un humano, la muestra se puede tomar en el primer o segundo trimestre del embarazo. Ejemplos de muestras biológicas maternas son las descritas anteriormente. En el paso **620**, la mezcla de ácidos nucleicos fetales y maternos se procesa adicionalmente de la muestra de fracción por ejemplo, plasma, para obtener una muestra que comprende una mezcla purificada de ácidos nucleicos fetales y maternos por ejemplo ADNcf, como se describe para forma de realización 100. En el paso **630**, una porción de la mezcla purificada de ADNcf fetal y materno se usa para amplificar una pluralidad de ácidos nucleicos diana polimórficos, cada uno de los cuales comprende un sitio polimórfico. Los sitios polimórficos que están contenidos en los ácidos nucleicos diana incluyen, sin limitación, polimorfismos de nucleótido único (SNP), SNP en tándem, eliminaciones o inserciones de múltiples bases a pequeña escala, llamadas IN-dELS (también llamados polimorfismos de inserción de delección o DIP), Polimorfismos Multi-Nucleótidos (MNP), repeticiones cortas en tándem (STR), polimorfismo de longitud de fragmentos de restricción (RFLP), o un polimorfismo que comprende cualquier otro cambio de secuencia en un cromosoma. Las secuencias polimórficas ejemplares y los métodos para amplificarlas son como se describen para las realizaciones que se muestran en las Figuras 2-5. En algunas realizaciones, los sitios polimórficos se ubican en cromosomas autosómicos, lo que permite la determinación de la fracción fetal independientemente del sexo del feto. Los polimorfismos asociados con cromosomas distintos de los cromosomas 13, 18, 21 e Y también se pueden usar en los métodos descritos en este documento.

[0195] En el paso **640**, una porción o la totalidad de las secuencias polimórficas amplificadas se utilizan para preparar una biblioteca de secuenciación para la secuenciación de un modo paralelo como se describe. En una realización, la biblioteca se prepara para la secuenciación por síntesis utilizando la química de secuenciación basada en el terminador reversible de Illumina, como se describe en el Ejemplo 13. En el paso **640**, la información de secuencia que se necesita para determinar la fracción fetal se obtiene usando un método NGS. En el paso **650**, la fracción fetal se determina según el número total de etiquetas que se asignan al primer alelo y la cantidad total de etiquetas que se asignan al segundo alelo en un sitio polimórfico informativo, por ejemplo, un SNP, contenido en un genoma de referencia artificial, por ejemplo, un SNP genoma de referencia. Los genomas diana artificiales son como se describen aquí. Se identifican los sitios polimórficos informativos, y la fracción fetal se calcula como se describe.

[0196] La determinación de la fracción fetal de acuerdo con el presente se puede usar junto con pruebas de diagnóstico clínicamente relevantes como control positivo de la presencia de ADNcf para resaltar resultados falsos negativos o falsos positivos derivados de niveles bajos de ADNcf a continuación. El límite de identificación. En una realización, la información de la fracción fetal se puede usar para establecer umbrales y estimar el tamaño mínimo de la muestra en la detección de aneuploidía. Tal uso se describe en el Ejemplo 16 a continuación. La información de la fracción fetal se puede utilizar junto con la información de secuenciación. Por ejemplo, los ácidos nucleicos de una muestra libre de células, por ejemplo, una muestra materna de plasma o suero, pueden usarse para enumerar secuencias en una muestra. Las secuencias se pueden enumerar usando cualquiera de las técnicas de secuencia descritas anteriormente. El conocimiento de la fracción fetal se puede utilizar para establecer los umbrales de "corte" para llamar a los estados "aneuploidía", "normal" o "marginal/sin llamada" (incierto). Luego, se pueden realizar cálculos para estimar el número mínimo de secuencias requeridas para lograr una sensibilidad adecuada (es decir, la probabilidad de identificar correctamente un estado de aneuploidía).

[0197] Los presentes métodos se pueden aplicar para determinar la fracción de cualquier población de ácidos nucleicos en una mezcla de ácidos nucleicos aportados por diferentes genomas. Además de determinar la fracción contribuida a una muestra por dos individuos, por ejemplo, el feto y la madre que lleva el feto contribuyen con los diferentes genomas, los métodos pueden usarse para determinar la fracción de un genoma en una mezcla derivada de dos células diferentes de un individuo, por ejemplo, los genomas son contribuidos a la muestra por células cancerosas aneuploides y células euploides normales del mismo sujeto.

Composiciones y kits

[0198] También se describen en este documento composiciones y kits o sistemas de reactivos útiles para practicar los métodos descritos en el presente documento.

[0199] Las composiciones descritas en este documento pueden ser incluidas en kits para las mezclas de secuenciación masivamente paralelas de moléculas de ácidos nucleicos fetales y maternos por ejemplo ADNcf, presente en una muestra materna por ejemplo una muestra de plasma. Los kits comprenden una composición que comprende al menos un conjunto de cebadores para amplificar al menos un ácido nucleico diana polimórfico en

dichas moléculas de ácido nucleico materno y fetal. Los ácidos nucleicos polimórficos pueden comprender, sin limitación, polimorfismos de un solo nucleótido (SNP), SNP en tándem, deleciones o inserciones de múltiples bases a pequeña escala, llamadas IN-dELS (también denominados polimorfismos de inserción o DIP), polimorfismos de múltiples nucleótidos (MNP), repeticiones cortas en tándem (STR), polimorfismo de la longitud del fragmento de restricción (RFLP), o un polimorfismo que comprende cualquier otro cambio de secuencia en un cromosoma. Los métodos de secuenciación son métodos NGS de moléculas de ácido nucleico único o moléculas de ácido nucleico amplificadas clonalmente como se describe en el presente documento. Los métodos NGS son métodos de secuenciación masivamente paralelos que incluyen pirosecuenciación, secuenciación por síntesis con terminadores de colorante reversibles, secuenciación en tiempo real, secuenciación por ligadura de la sonda oligonucleótida o secuenciación de una sola molécula.

[0200] En una realización, la composición incluye cebadores para amplificar ácidos nucleicos diana polimórficos que comprenden cada uno al menos un SNP. El al menos un SNP se selecciona entre SNP 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. Los conjuntos correspondientes de cebadores para amplificar los SNP se proporcionan como SEQ ID NO: 57-112.

[0201] En otra realización, la composición comprende cebadores para amplificar ácidos nucleicos diana polimórficos que comprenden cada uno al menos un SNP en tándem. Los SNP en tándem ejemplares incluyen 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, la composición incluye cebadores para amplificar los SNP en tándem ejemplares descritos en el presente documento, y la composición comprende los cebadores ejemplares correspondientes de SEQ ID NO: 197-310.

[0202] En otra realización, la composición comprende cebadores para amplificar ácidos nucleicos diana polimórficos que comprenden cada uno al menos un STR. Los STR 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. En una realización, la composición incluye cebadores para amplificar los STR en tándem ejemplares descritos en el presente documento, y la composición comprende los cebadores ejemplares correspondientes de las SEQ ID NO: 113-196.

[0203] Los kits pueden contener una combinación de reactivos que incluye los elementos necesarios para realizar un ensayo de acuerdo con los métodos descritos en el presente documento. El sistema de reactivos se presenta en una forma comercial empaquetada, como una composición o mezcla donde la compatibilidad de los reactivos permitirá, en una configuración de dispositivo de prueba, o más típicamente como un kit de prueba, es decir, una combinación empaquetada de uno o más contenedores, dispositivos o dispositivos similares que contienen los reactivos necesarios y, de preferencia, incluyen instrucciones escritas para la realización de los ensayos. El kit descrito en el presente documento puede adaptarse para cualquier configuración de ensayo y puede incluir composiciones para realizar cualquiera de los diversos formatos de ensayo descritos en el presente documento. Los kits para determinar la fracción fetal comprenden composiciones que incluyen conjuntos de cebadores para amplificar ácidos nucleicos polimórficos presentes en una muestra materna como se describe y, cuando corresponde, los reactivos para purificar el ADNcf, están dentro del alcance de la divulgación. En una realización, un kit diseñado para permitir la cuantificación de secuencias polimórficas fetales y maternas, por ejemplo, STR y/o SNP y/o SNP en tándem, en una muestra de plasma de ADNcf, incluye al menos un conjunto de oligo-nucleótidos específicos de alelo específicos para un SNP seleccionado y/o región de repeticiones en tándem. Preferiblemente, el kit incluye una pluralidad de conjuntos de cebadores para amplificar un panel de secuencias polimórficas. Un kit puede comprender otros reactivos e/o información para genotipar o cuantificar alelos en una muestra (por ejemplo, tampones, nucleótidos, instrucciones). Los kits también incluyen una pluralidad de contenedores de tampones y reactivos apropiados.

Productos informáticos

[0204] La determinación de aneuploidía y/o la determinación de la fracción fetal se deriva computacionalmente de la gran cantidad de información de secuencia que se obtiene de acuerdo con los métodos descritos en el presente documento. En una realización, se describe en este documento un medio legible por computadora que tiene almacenado en él instrucciones legibles por computadora para determinar la presencia o ausencia de aneuploidía a partir de la información obtenida de la secuenciación de ácidos nucleicos fetales y maternos en una muestra materna. En una realización, el medio legible por computadora utiliza información de secuencia obtenida de una pluralidad de moléculas de ácido nucleico maternas y fetales en una muestra de plasma materno para identificar una serie de etiquetas de secuencia mapeadas para un cromosoma de interés y para un cromosoma de normalización. Utilizando el número de etiquetas de secuencia mapeadas identificadas para un cromosoma de interés y el número de etiquetas de secuencia mapeadas identificadas para al menos un cromosoma normalizador, el medio legible por computadora calcula una dosis de cromosoma para un cromosoma de interés; y compara la dosis de cromosoma con al menos un valor umbral, y de ese modo identifica la presencia o ausencia de aneuploidía fetal. Ejemplos de cromosomas de interés incluyen, sin limitación de los cromosomas 21, 13, 18 y X.

[0205] En otra realización, se da a conocer en el presente documento un sistema de procesamiento de ordenador que se adapta o configura para determinar la presencia o ausencia de aneuploidía de la información obtenida de la secuenciación Ácidos nucleicos fetales y maternos en una muestra materna. El sistema de procesamiento informático está adaptado o configurado para (a) utilizar información de secuencia obtenida de una pluralidad de moléculas de ácido nucleico maternas y fetales en una muestra de plasma materno para identificar una serie de etiquetas de secuencia mapeadas para un cromosoma de interés; (b) utilizar la información de secuencia obtenida de una pluralidad de moléculas de ácido nucleico materno y fetal en una muestra de plasma materno para identificar una serie de etiquetas de secuencia mapeadas para al menos un cromosoma normalizador; (c) utilizar el número de etiquetas de secuencia mapeadas identificadas para un cromosoma de interés en el paso (a) y el número de etiquetas de secuencia mapeadas identificadas para al menos un cromosoma normalizador en el paso (b) para calcular una dosis de cromosoma para un cromosoma de interés; y (d) comparar la dosis de cromosoma con al menos un valor de umbral y, de ese modo, identificar la presencia o ausencia de aneuploidía fetal. Los ejemplos de cromosomas de interés incluyen, sin limitación, los cromosomas 21, 13, 18 y X.

[0206] En otra realización, se da a conocer en el presente documento el aparato adaptado o configurado para determinar la presencia o ausencia de aneuploidía de la información obtenida de la secuenciación de ácidos nucleicos fetales y maternos en una muestra materna. El aparato está adaptado o configurado para comprender (a) un dispositivo de secuenciación adaptado o configurado para secuenciar al menos una porción de las moléculas de ácido nucleico en una muestra de plasma materno que comprende moléculas de ácido nucleico materno y fetal, generando así información de secuencia; y (b) un sistema de procesamiento informático configurado para realizar los pasos de: (i) utilizar la información de secuencia generada por el dispositivo de secuenciación para identificar un número de etiquetas de secuencia mapeadas para un cromosoma de interés; (ii) usar la información de secuencia generada por el dispositivo de secuenciación para identificar un número de etiquetas de secuencia mapeadas para al menos un cromosoma normalizador; (iii) usar el número de etiquetas de secuencia mapeadas identificadas para un cromosoma de interés en el paso (i) y el número de etiquetas de secuencia mapeadas identificadas para al menos un cromosoma normalizador en el paso (ii) para calcular una dosis de cromosoma para un cromosoma de interés; y (iv) comparar dicha dosis de cromosoma con al menos un valor umbral, y de ese modo identificar la presencia o ausencia de aneuploidía fetal. Los ejemplos de cromosomas de interés incluyen, sin limitación, los cromosomas 21, 13, 18 y X.

Ejemplo 1

Procesamiento de muestras y extracción de ADNcf

[0207] Muestras de sangre periférica se obtuvieron de las mujeres embarazadas en su primer o segundo trimestre del embarazo y que fueron considerados en riesgo de aneuploidía fetal. El consentimiento informado se obtuvo de cada participante antes de la extracción de sangre. La sangre se extrajo antes de la amniocentesis o de las muestras de vellosidades coriónicas. El análisis del cariotipo se realizó con vellosidades coriónicas o muestras de amniocentesis para confirmar el cariotipo fetal.

[0208] Sangre periférica extraída de cada sujeto se recogió en tubos ACD. Un tubo de muestra de sangre (aproximadamente 6-9 ml/tubo) se transfirió a un tubo de centrifugación de baja velocidad de 15 ml. La sangre se centrifugó a 2640 rpm, 4°C durante 10 min usando la centrífuga Beckman Allegra 6 R y el modelo con rotor GA 3.8.

[0209] Para la extracción de plasma libre de células, la capa superior de plasma se transfirió a un tubo de centrífuga de 15 ml de alta velocidad y se centrifugó a 16.000 x g, 4°C durante 10 min usando centrífuga Beckman Coulter Avanti J-E, y rotor JA-14. Los dos pasos de centrifugación se realizaron dentro de las 72 h posteriores a la extracción de sangre. El plasma libre de células que comprende ADNcf se almacenó a -80°C y se descongeló solo una vez antes de la amplificación de ADNcf en plasma o para la purificación de ADNcf.

[0210] El ADN libre de células purificado (cfDNA) se extrajo del plasma libre de células utilizando el Mini kit de ADN en sangre QIAamp (Qiagen) esencialmente de acuerdo con las instrucciones del fabricante. Se añadieron un mililitro

de tampón AL y 100 µl de solución de proteasa a 1 ml de plasma. La mezcla se incubó durante 15 minutos a 56°C. Se añadió un mililitro de etanol al 100% a la digestión de plasma. La mezcla resultante se transfirió a mini columnas QIAamp que se ensamblaron con VacVálvulas y VacConectores que se proporcionan en el conjunto de columna QIAvac 24 Plus (Qiagen). Se aplicó vacío a las muestras y el ADNcf retenido en los filtros de la columna se lavó a vacío con 750 µl de tampón AW1, seguido de un segundo lavado con 750 µl de tampón AW24. La columna se centrifugó a 14.000 RPM durante 5 minutos para eliminar cualquier tampón residual del filtro. El ADNcf se eluyó con tampón AE mediante centrifugación a 14.000 RPM, y la concentración se determinó utilizando la Plataforma de cuantificación Qubit™ (Invitrogen).

Ejemplo 2

Preparación y secuenciación de bibliotecas de secuenciación primarias y enriquecidas

a. Preparación de bibliotecas de secuenciación - protocolo abreviado

[0211] Todas las bibliotecas de secuenciación, es decir, las bibliotecas primarias y enriquecidas, se prepararon a partir de aproximadamente 2 ng de ADNcf purificado que se extraía de plasma materno. La preparación de la biblioteca se realizó utilizando los reactivos del conjunto de reactivos de ADN de preparación de muestra de ADN NEBNext™, serie 1 (número de pieza E6000L; New England Biolabs, Ipswich, MA), para Illumina® de la siguiente manera. Debido a que el ADN plasmático libre de células es de naturaleza fragmentada, no se realizó una fragmentación adicional por nebulización o sonicación en las muestras de ADN plasmático. Los salientes de aproximadamente 2 ng de fragmentos de ADNcf purificados contenidos en 40 µl se convirtieron en extremos romos fosforilados de acuerdo con el módulo de reparación del extremo NEBNext® incubando en un tubo de microcentrífuga de 1,5 ml el ADNcf con 5 µl de tampón de fosforilación 10X, 2 µl de mezcla de solución de desoxinucleótido (10 mM cada dNTP), 1 µl de una dilución 1:5 de ADN polimerasa I, 1 µl de ADN polimerasa T4 y 1 µl de quinasa de polinucleótido T4 proporcionada en el conjunto de reactivos de ADN de preparación de muestras de ADN NEBNext™ 1 durante 15 minutos a 20°C. Las enzimas se inactivaron por calor mediante la incubación de la mezcla de reacción a 75°C durante 5 minutos. La mezcla se enfrió a 4°C, y la cola de dA del ADN de extremos romos se realizó utilizando 10 µl de la mezcla maestra de cola de dA que contiene el fragmento Klenow (3' a 5' exo menos) (NEBNext™ DNA Sample Prep DNA Reagent Set 1), e incubación durante 15 minutos a 37°C. Posteriormente, el fragmento Klenow se inactivó por calor mediante la incubación de la mezcla de reacción a 75°C durante 5 minutos. Después de la inactivación del fragmento Klenow, se usó 1 µl de una dilución 1:5 de Oligo Mix del Adaptador Genómico de Illumina (Núm. de Parte 1000521; Illumina Inc., Hayward, CA) para ligar los adaptadores de Illumina (Adaptadores Y Sin Índice) al ADN con cola de dA utilizando 4 µl de la ligasa de ADN T4 provista en el conjunto de reactivos de ADN 1 de preparación de muestra de ADN NEBNext™, incubando la mezcla de reacción durante 15 minutos a 25°C. La mezcla se enfrió a 4°C y el ADNcf ligado con el adaptador se purificó a partir de adaptadores no dilatados, dímeros adaptadores y otros reactivos utilizando perlas magnéticas proporcionadas en el sistema de purificación Agencourt AMPure XP PCR (Parte N° A63881; Beckman Coulter Genomics, Danvers, MA). Se realizaron dieciocho ciclos de PCR para enriquecer selectivamente el ADNcf ligado con adaptador (25 µl) utilizando la mezcla maestra de alta fidelidad Phusion® (25 µl; Finnzymes, Woburn, MA) y los cebadores de PCR de Illumina (0,5 µm cada uno) complementarios a adaptadores (n.º de pieza 1000537 y 1000537). El ADN ligado con el adaptador se sometió a PCR (98°C durante 30 segundos; 18 ciclos de 98°C durante 10 segundos, 65°C durante 30 segundos y 72°C durante 30; extensión final a 72°C durante 5 minutos y manténgalo a 4°C) usando Illumina Genomic PCR Primers (números de pieza 100537 y 100538) y la mezcla maestra PhusionHFPCR incluida en el conjunto de reactivos de ADN de preparación de muestras de ADN NEBNext™, conjunto 1, de acuerdo con las instrucciones del fabricante. El producto amplificado se purificó utilizando el sistema de purificación por PCR Agencourt AMPure XP (Agencourt Bioscience corporation, Beverly, MA) de acuerdo con las instrucciones del fabricante disponibles en www.beckmangenomics.com/products/AMPureXPProtocol_000387v001.pdf. El producto amplificado purificado se eluyó en 40 µl de Qiagen EB Buffer, y se analizó la concentración y la distribución de tamaños de las bibliotecas amplificadas utilizando el kit Agilent DNA 1000 para el bioanalizador 2100 (Agilent technologies Inc., Santa Clara, CA).

b. Preparación de bibliotecas de secuenciación de protocolo de longitud completa.

[0212] El protocolo de longitud completa descrito es esencialmente el protocolo estándar proporcionado por Illumina, y sólo difiere del protocolo Illumina en la purificación de la biblioteca amplificada: el protocolo Illumina instruye que la biblioteca amplificada se purificó usando electroforesis en gel, mientras que el protocolo aquí descrito usa perlas magnéticas para el mismo paso de purificación. Aproximadamente 2 ng de ADNcf purificado que se había extraído del plasma materno se usaron para preparar una biblioteca de secuenciación primaria utilizando el conjunto de reactivos 1 de ADN de preparación de muestra NEBNext™ (Nº de pieza E6000L; New England Biolabs, Ipswich, MA) para Illumina® esencialmente de acuerdo con las instrucciones del fabricante. Todos los pasos, excepto la purificación final de los productos ligados con adaptador, que se realizaron con perlas magnéticas y reactivos de Agencourt en lugar de la columna de purificación, se realizaron de acuerdo con el protocolo que acompaña a los reactivos NEBNext™ para la preparación de muestras para una biblioteca de ADN genómico secuenciada utilizando el Illumina® GAI. El protocolo NEBNext™ sigue esencialmente el proporcionado por Illumina, que está disponible en grcf.jhml.edu/hts/protocols/11257047_ChIP_Sample_Prep.pdf.

[0213] Los salientes de aproximadamente 2 ng de fragmentos de ADNcf purificados contenidos en 40 µl se convirtieron en extremos romos fosforilados de acuerdo con el módulo de reparación de extremo NEBNext® incubando los 40 µl de ADNcf con 5 µl de tampón de fosforilación 10X, 2 µl de mezcla de solución de desoxinucleótidos (10 mM cada dNTP), 1 µl de una dilución 1:5 de ADN polimerasa I, 1 µl de ADN polimerasa T4 y 1 µl de quinasa de polinucleótido T4 proporcionada en el conjunto de reactivos de ADN de preparación de muestra de NEBNext™ un tubo de microcentrifuga de 200 µl en un termociclador durante 30 minutos a 20°C. La muestra se enfrió a 4°C y se purificó utilizando una columna QIAQuick proporcionada en el kit de purificación por PCR QIAQuick (QIAGEN Inc., Valencia, CA) de la siguiente manera. La reacción de 50 µl se transfirió a un tubo de microcentrifuga de 1,5 ml y se agregaron 250 µl de tampón Qiagen PB. Los 300 µl resultantes se transfirieron a una columna QIAQuick, que se centrifugó a 13.000 RPM durante 1 minuto en una microcentrifuga. La columna se lavó con 750 µl de Qiagen Buffer PE y se volvió a centrifugar. El etanol residual se eliminó mediante una centrifugación adicional durante 5 minutos a 13.000 RPM. El ADN se eluyó en 39 µl de Qiagen Buffer EB por centrifugación. La cola de dA de 34 µl del ADN de extremos romos se logró utilizando 16 µl de la mezcla maestra de colas de dA que contiene el fragmento Klenow (3' a 5' menos) (incubación de muestra de ADN NEBNext™ Conjunto 1 de reactivos de ADN), e incubación durante 30 minutos a 37°C de acuerdo con el NEBNext® dA-Tailing Module del fabricante. La muestra se enfrió a 4°C y se purificó utilizando una columna provista en el Kit de purificación de PCR MinElute (QIAGEN Inc., Valencia, CA) de la siguiente manera. Los 50 µl de reacción se transfirieron a 1,5 ml de tubo de microcentrifuga y se agregaron 250 µl de Qiagen Buffer PB. Los 300 µl se transfirieron a la columna MinElute, que se centrifugó a 13.000 RPM durante 1 minuto en una microcentrifuga. La columna se lavó con 750 µl de Qiagen Buffer PE y se volvió a centrifugar. El etanol residual se eliminó mediante una centrifugación adicional durante 5 minutos a 13.000 RPM. El ADN se eluyó en 15 µl de Qiagen Buffer EB por centrifugación. Diez microlitros del eluido de ADN se incubaron con 1 µl de una dilución 1:5 de la Illumina Genomic Adapter Oligo Mix (Parte N° 1.000.521), 15 µl de 2X Quick Ligation Reaction Buffer, y 4 µl Quick T4 DNA Ligase, durante 15 minutos a 25°C según el módulo de ligadura rápida NEBNext®. La muestra se enfrió a 4°C y se purificó utilizando una columna MinElute de la siguiente manera. Ciento cincuenta microlitros de Qiagen Buffer PE se agregaron a la reacción de 30 µl, y todo el volumen se transfirió a una columna MinElute a una columna MinElute, que se centrifugó a 13.000 RPM durante 1 minuto en una microcentrifuga. La columna se lavó con 750 µl de Qiagen Buffer PE y se volvió a centrifugar. El etanol residual se eliminó mediante una centrifugación adicional durante 5 minutos a 13.000 RPM. El ADN se eluyó en 28 µl de Qiagen Buffer EB por centrifugación. Veintitrés microlitros del eluato de ADN ligado al adaptador se sometieron a 18 ciclos de 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 Illumina Genomic PCR Primers (números de pieza 100537 y 1000538) y la mezcla maestra de Phusion HF PCR proporcionada en el conjunto de reactivos de ADN de preparación de muestras de ADN NEBNext™, de acuerdo con las instrucciones del fabricante. El producto amplificado se purificó utilizando el sistema de purificación por PCR Agencourt AMPure XP (Agencourt Bioscience Corporation, Beverly, MA) de acuerdo con las instrucciones del fabricante disponibles en www.beckmangencomics.com/products/AMPureXPProtocol_000387v001.pdf. El sistema de purificación por PCR Agencourt AMPure XP elimina los dNTP, cebadores, dímeros de cebadores, sales y otros contaminantes no incorporados, y recupera los amplicones mayores de 100 pb. El producto amplificado purificado se eluyó de las perlas de Agencourt en 40 µl de Qiagen EB Buffer y se analizó la distribución de tamaños de las bibliotecas utilizando el kit Agilent DNA 1.000 para el bioanalizador 2100 (Agilent technologies Inc., Santa Clara, CA).

c. Análisis de las bibliotecas de secuenciación preparadas de acuerdo con los protocolos abreviados (a) y completos (b)

[0214] Los electropherograms generados por el Bioanalizador se muestran en la **Figura 7**. La **Figura 7 (A)** muestra el electroferograma de la ADN de biblioteca preparada a partir de ADNcf purificada a partir de muestra de plasma M24228 utilizando el protocolo de longitud completa descrito en (a), y la **Figura 7 (B)** muestra el electroferograma del ADN de la biblioteca preparado a partir de ADNcf purificado de la muestra de plasma M24228 utilizando el protocolo de longitud completa descrito en (b). En ambas figuras, los picos 1 y 4 representan el marcador inferior de 15 pb y el marcador superior de 1.500, respectivamente; los números sobre los picos indican los tiempos de migración para los fragmentos de la biblioteca; y las líneas horizontales indican el umbral establecido para la integración. El electroferograma en la **Figura 7 (A)** muestra un pico menor de fragmentos de 187 pb y un pico mayor de fragmentos de 263 pb, mientras que el electroferograma en la **Figura 7 (B)** muestra solo un pico a 265 pb. La integración de las áreas de los picos dio como resultado una concentración calculada de 0,40 ng/µl para el ADN del pico de 187 pb en la **Figura 7 (A)**, una concentración de 7,34 ng/µl para el ADN del pico de 263 pb en la **Figura 7 (A)**, y una concentración de 14,72 ng/µl para el ADN del pico de 265 pb en la **Figura 7 (B)**. Se sabe que los adaptadores de Illumina que se ligaron al ADNcf tienen 92 pb, que cuando se sustraen del 265 pb, indican que el tamaño máximo de pfcNA es de 173 pb. Es posible que el pico menor a 187 pb represente fragmentos de dos cebadores que se ligaron de extremo a extremo. Los fragmentos de dos cebadores lineales se eliminan del producto de la biblioteca final cuando se usa el protocolo abreviado. El protocolo abreviado también elimina otros fragmentos más pequeños de menos de 187 pb. En este ejemplo, la concentración de ADNcf ligado con adaptador purificado es el doble que la del ADNcf ligado con adaptador producido utilizando el protocolo de longitud completa. Se ha observado que la concentración de los fragmentos de ADNcf ligados con el adaptador es siempre mayor que la obtenida utilizando el protocolo de longitud completa (datos no mostrados).

[0215] Por lo tanto, una ventaja de la preparación de la biblioteca de secuenciación utilizando el protocolo abreviado

es que la biblioteca obtenida consistentemente comprende sólo un pico principal en el intervalo 262-267bp mientras que la calidad de la biblioteca preparada usando el protocolo de longitud completa varía como se refleja por el número y la movilidad de picos distintos de los que representan el ADNcf. Los productos que no son de ADNcf ocuparían espacio en la célula de flujo y disminuirían la calidad de la amplificación del grupo y la imagen posterior de las reacciones de secuenciación, lo que subyace a la asignación general del estado de aneuploidía. Se demostró que el protocolo abreviado no afecta la secuenciación de la biblioteca (consulte la **Figura 8**).

[0216] Otra ventaja de la preparación de la biblioteca de secuenciación utilizando el protocolo abreviado es que los tres pasos enzimáticos de extremos romos, cola del dA, y el adaptador de la ligación, tardan menos de una hora para completar para apoyar la validación y la aplicación de un servicio de diagnóstico de aneuploidías rápida

[0217] Otra ventaja es que los tres pasos enzimáticos de extremos romos, cola de dA y ligadura del adaptador, se realizan en el mismo tubo de reacción, evitando así múltiples transferencias de muestras que potencialmente conducirían a la pérdida de material y, lo que es más importante, mezclar muestras y contaminación de la muestra.

Ejemplo 3

Secuenciación masiva paralela y determinación de aneuploidía

[0218] Muestras de sangre periférica se obtuvieron de mujeres embarazadas y ADNcf se purificó de la fracción de plasma como se describe en el Ejemplo 1. Todas las bibliotecas de secuenciación se prepararon utilizando el protocolo de preparación de biblioteca abreviado descrito en el Ejemplo 2. El ADN amplificado se secuenció utilizando Genoma Analizador de Illumina II para obtener lecturas de un solo extremo de 36 pb. Solo se necesitan aproximadamente 30 pb de información de secuencia aleatoria para identificar una secuencia como perteneciente a un cromosoma humano específico. Las secuencias más largas pueden identificar de forma única objetivos más particulares. En el presente caso, se obtuvo un gran número de lecturas de 36 pb, que cubren aproximadamente el 10% del genoma. La secuenciación del ADN de la biblioteca se realizó utilizando el Analizador de Genoma II (Illumina Inc., San Diego, CA, EE.UU.) según los protocolos estándar del fabricante. Las copias del protocolo para la secuenciación del genoma completo utilizando la tecnología Illumina/Solexa se pueden encontrar en BioTechniques.RTM. Protocol Guide 2007, publicado en diciembre de 2006: p. 29, y en la web en biotechniques.com/default.asp?page=Protocol&subsection=article_display&id=112378. La biblioteca de ADN se diluyó a 1nM y se desnaturalizó. El ADN de biblioteca (5 pM) se sometió a amplificación de grupo de acuerdo con el procedimiento descrito en Cluster Station User Guide y Cluster Station Operations Guide de Illumina, disponible en Internet en illumina.com/systems/genome_analyzer/cluster_station.ilmn. Al finalizar la secuenciación de la muestra, el "Software de control del secuenciador" de Illumina transfirió los archivos de imagen y de llamada base a un servidor Unix que ejecuta el software Illumina Analyzer Pipeline de Illumina versión 1.51. El programa Illumina "Gerald" se ejecutó para alinear secuencias con el genoma humano de referencia que se deriva del genoma hg18 proporcionado por el Centro Nacional de Información Biotecnológica (NCBI36/hg18, disponible en la web en <http://genome.ucsc.edu/cgi-bin/hgGateway?org=Human&db=hg18&hgsid=166260105>). Los datos de secuencia generados por el procedimiento anterior que se alinearon de forma única con el genoma se leyeron desde la salida de Gerald (archivos export.txt) por un programa (c2c.pl) que se ejecuta en una computadora que ejecuta el sistema operativo Linux. Las alineaciones de secuencia con desajustes de base se permitieron e incluyeron en los recuentos de alineación solo si se alineaban de forma única con el genoma. Se excluyeron las alineaciones de secuencia con coordenadas de inicio y final idénticas (duplicados).

[0219] Entre aproximadamente 5 y 15 millones de etiquetas de 36 pb con 2 o menos desajustes se asignaron de forma única al genoma humano. Todas las etiquetas mapeadas se contaron e incluyeron en el cálculo de las dosis de cromosomas en las muestras de prueba y de calificación. Las regiones que se extienden desde la base 0 hasta la base 2 x 10⁶, la base 10 x 10⁶ hasta la base 13 x 10⁶ y la base 23 x 10⁶ hasta el final del cromosoma Y, se excluyeron específicamente del análisis porque las etiquetas derivadas de fetos masculinos o femeninos se asignan a estas regiones del cromosoma Y.

[0220] Se observó que alguna variación en el número total de etiquetas de secuencia mapeadas a cromosomas individuales a través de muestras secuenciadas en la misma serie (variación inter-cromosómica), pero se observó una variación substancialmente mayor entre las diferentes secuencias de secuenciación (variación de ejecución de inter-secuenciación).

Ejemplo 4

Dosis y varianza para los cromosomas 13, 18, 21, X e Y

[0221] Para examinar la extensión de la variación inter-cromosómica y de inter-secuenciación en el número de etiquetas de secuencia mapeadas para todos los cromosomas, plasma de ADNpc de la sangre periférica obtenido de 48 pacientes embarazadas voluntarias se extrajo y se secuenció como se describe en el Ejemplo 1, y se analizó como sigue.

[0222] Se determinó el número total de etiquetas de secuencias que fueron asignadas a cada cromosoma (densidad de etiqueta de secuencia). Alternativamente, el número de etiquetas de secuencia mapeadas puede normalizarse a la longitud del cromosoma para generar una relación de densidad de etiqueta de secuencia. La normalización a la longitud del cromosoma no es un paso necesario, y se puede realizar 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 etiquetas de secuencia pueden ser las longitudes proporcionadas en Internet en gnome.ucsc.edu/goldenPath/stats.html#hg18.

[0223] La densidad de etiqueta de secuencia resultante para cada cromosoma estaba relacionado con la densidad de etiqueta de secuencia de cada uno de los cromosomas restantes para derivar una dosis de cromosoma cualificado, que se calculó como la relación de la densidad de etiqueta de secuencia para el cromosoma de interés por ejemplo, cromosoma 21, y la densidad de la etiqueta de secuencia de cada uno de los cromosomas restantes, es decir, los cromosomas 1-20, 22 y X. La Tabla 1 proporciona un ejemplo de la dosis de cromosoma calificada calculada para los cromosomas de interés 13, 18, 21, X e Y, determinada en una de las muestras calificadas. Las dosis de cromosomas se determinaron para todos los cromosomas en todas las muestras, y las dosis promedio para los cromosomas de interés 13, 18, 21, X e Y en las muestras calificadas se proporcionan en las Tablas 2 y 3, y se representan en las Figuras 9-13. Las Figuras 9-13 también representan las dosis de cromosomas para las muestras de prueba. Las dosis de cromosomas para cada uno de los cromosomas de interés en las muestras calificadas proporcionan una medida de la variación en el número total de etiquetas de secuencia mapeadas para cada cromosoma de interés en relación con la de cada uno de los cromosomas restantes. Así, las dosis de cromosomas calificadas pueden identificar el cromosoma o un grupo de cromosomas, es decir, un cromosoma normalizador, que tiene una variación entre las muestras más cercana a la variación del cromosoma de interés y que serviría como secuencias ideales para normalizar los valores para una evaluación estadística adicional. Las figuras 14 y 15 representan las dosis promedio calculadas de cromosomas determinadas en una población de muestras calificadas para los cromosomas 13, 18 y 21, y los cromosomas X e Y.

[0224] En algunos casos, el mejor cromosoma normalizador puede no tener la menor variación, pero puede tener una distribución de dosis calificadas que distinga mejor una muestra de prueba o muestras de las muestras calificadas, es decir, el mejor cromosoma normalizador puede no tener la más baja variación, pero puede tener la mayor diferenciabilidad. Por lo tanto, la diferenciabilidad explica la variación en la dosis de cromosoma y la distribución de las dosis en las muestras calificadas.

[0225] Las tablas 2 y 3 proporcionan el coeficiente de variación como la medida de la variabilidad y los valores de la prueba t de Student como una medida de la diferenciabilidad para los cromosomas 18, 21, X e Y, en donde, cuanto más pequeño sea el valor de la prueba T, mayor la diferenciabilidad. La diferenciabilidad para el cromosoma 13 se determinó como la proporción de la diferencia entre la dosis media del cromosoma en las muestras calificadas y la dosis para el cromosoma 13 en la única muestra de prueba T13, y la desviación estándar de la media de la dosis calificada.

[0226] Las dosis de cromosomas cualificados también sirven como la base para determinar los valores de umbral cuando la identificación de aneuploidías en muestras de ensayo como se describe en lo siguiente.

TABLA 1

Dosis de cromosomas calificada para los cromosomas 13, 18, 21, X e Y (n = 1; muestra # 11342, 46 XY)					
Cromosoma	chr 21	chr 18	chr 13	chr X	chrY
chr1	0,149901	0,306798	0,341832	0,490969	0,003958
chr2	0,15413	0,315452	0,351475	0,504819	0,004069
chr3	0,193331	0,395685	0,44087	0,633214	0,005104
chr4	0,233056	0,476988	0,531457	0,763324	0,006153
chr5	0,219209	0,448649	0,499882	0,717973	0,005787
chr6	0,228548	0,467763	0,521179	0,748561	0,006034
chr7	0,245124	0,501688	0,558978	0,802851	0,006472
chr8	0,256279	0,524519	0,584416	0,839388	0,006766
chr9	0,309871	0,634203	0,706625	1,014915	0,008181
chr10	0,25122	0,514164	0,572879	0,822817	0,006633
chr11	0,257168	0,526338	0,586443	0,8423	0,00679
chr12	0,275192	0,563227	0,627544	0,901332	0,007265
chr13	0,438522	0,897509	1	1,436285	0,011578
chr14	0,405957	0,830858	0,925738	1,329624	0,010718
chr15	0,406855	0,832697	0,927786	1,332566	0,010742
chr16	0,376148	0,769849	0,857762	1,231991	0,009931
chr17	0,383027	0,783928	0,873448	1,254521	0,010112
chr18	0,488599	1	1,114194	1,600301	0,0129
chr19	0,535867	1,096742	1,221984	1,755118	0,014148
chr20	0,467308	0,956424	1,065642	1,530566	0,012338
chr21	1	2,046668	2,280386	3,275285	0,026401
chr22	0,756263	1,547819	1,724572	2,476977	0,019966
chrX	0,305317	0,624882	0,696241	1	0,008061
chrY	37,87675	77,52114	86,37362	124,0572	1

TABLA 2

Dosis de cromosomas calificada, varianza y diferenciabilidad para los cromosomas 21, 18 y 13								
	21 (n=35)				18 (n=40)			
	Promedio	DesEst	CV	Prueba T	Promedio	DesEst	CV	Prueba T
chr1	0,15335	0,001997	1,30	3,18E-10	0,31941	0,008384	2,62	0,001675
chr2	0,15267	0,001966	1,29	9,87E-07	0,31807	0,001756	0,55	4,39E-05
chr3	0,18936	0,004233	2,24	1,04E-05	0,39475	0,002406	0,61	3,39E-05
chr4	0,21998	0,010668	4,85	0,000501	0,45873	0,014292	3,12	0,001349
chr5	0,21383	0,005058	2,37	1,43E-05	0,44582	0,003288	0,74	3,09E-05
chr6	0,22435	0,005258	2,34	1,48E-05	0,46761	0,003481	0,74	2,32E-05
chr7	0,24348	0,002298	0,94	2,05E-07	0,50765	0,004669	0,92	9,07E-05
chr8	0,25269	0,003497	1,38	1,52E-06	0,52677	0,002046	0,39	4,89E-05
chr9	0,31276	0,003095	0,99	3,83E-09	0,65165	0,013851	2,13	0,000559
chr10	0,25618	0,003112	1,21	2,28E-10	0,53354	0,013431	2,52	0,002137
chr11	0,26075	0,00247	0,95	1,08E-09	0,54324	0,012859	2,37	0,000998
chr12	0,27563	0,002316	0,84	2,04E-07	0,57445	0,006495	1,13	0,000125
chr13	0,41828	0,016782	4,01	0,000123	0,87245	0,020942	2,40	0,000164
chr14	0,40671	0,002994	0,74	7,33E-08	0,84731	0,010864	1,28	0,000149
chr15	0,41861	0,007686	1,84	1,85E-10	0,87164	0,027373	3,14	0,003862
chr16	0,39977	0,018882	4,72	7,33E-06	0,83313	0,050781	6,10	0,075458
chr17	0,41394	0,02313	5,59	0,000248	0,86165	0,060048	6,97	0,088579

chr18	0,47236	0,016627	3,52	1,3E-07				
chr19	0,59435	0,05064	8,52	0,01494	1,23932	0,12315	9,94	0,231139
chr20	0,49464	0,021839	4,42	2,16E-06	1,03023	0,058995	5,73	0,061101
chr21					2,03419	0,08841	4,35	2,81E-05
chr22	0,84824	0,070613	8,32	0,02209	1,76258	0,169864	9,64	0,181808
chrX	0,27846	0,015546	5,58	0,000213	0,58691	0,026637	4,54	0,064883

TABLA 3

Dosis de cromosomas calificada, varianza y diferenciabilidad para los cromosomas 13, X e Y								
	13 (n=47)				X (n=19)			
	Promedio	DesEst	CV	Dif	Promedio	DesEst	CV	Prueba T
chr1	0,36536	0,01775	4,86	1,904	0,56717	0,025988	4,58	0,001013
chr2	0,36400	0,009817	2,70	2,704	0,56753	0,014871	2,62	9,6E-08
chr3	0,45168	0,007809	1,73	3,592	0,70524	0,011932	1,69	6,13E-11
chr4	0,52541	0,005264	1,00	3,083	0,82491	0,010537	1,28	1,75E-15
chr5	0,51010	0,007922	1,55	3,944	0,79690	0,012227	1,53	1,29E-11
chr6	0,53516	0,008575	1,60	3,758	0,83594	0,013719	1,64	2,79E-11
chr7	0,58081	0,017692	3,05	2,445	0,90507	0,026437	2,92	7,41E-07
chr8	0,60261	0,015434	2,56	2,917	0,93990	0,022506	2,39	2,11E-08
chr9	0,74559	0,032065	4,30	2,102	1,15822	0,047092	4,07	0,000228
chr10	0,61018	0,029139	4,78	2,060	0,94713	0,042866	4,53	0,000964
chr11	0,62133	0,028323	4,56	2,081	0,96544	0,041782	4,33	0,000419
chr12	0,65712	0,021853	3,33	2,380	1,02296	0,032276	3,16	3,95E-06
chr13					1,56771	0,014258	0,91	2,47E-15
chr14	0,96966	0,034017	3,51	2,233	1,50951	0,05009	3,32	8,24E-06
chr15	0,99673	0,053512	5,37	1,888	1,54618	0,077547	5,02	0,002925
chr16	0,95169	0,080007	8,41	1,613	1,46673	0,117073	7,98	0,114232
chr17	0,98547	0,091918	9,33	1,484	1,51571	0,132775	8,76	0,188271
chr18	1,13124	0,040032	3,54	2,312	1,74146	0,072447	4,16	0,001674
chr19	1,41624	0,174476	12,32	1,306	2,16586	0,252888	11,68	0,460752
chr20	1,17705	0,094807	8,05	1,695	1,81576	0,137494	7,57	0,08801
chr21	2,33660	0,131317	5,62	1,927	3,63243	0,235392	6,48	0,00675
chr22	2,01678	0,243883	12,09	1,364	3,08943	0,34981	11,32	0,409449
chrX	0,66679	0,028788	4,32	1,114				
chr2-6	0,46751	0,006762	1,45	4,066				
chr3-6	0,50332	0,005161	1,03	5,260				
chr_tot					1,13209	0,038485	3,40	2,7E-05
	Y (n=26)							
	Promedio	DesEst	CV	Prueba T				
Chr 1-22, X	0,00734	0,002611	30,81	1,8E-12				

[0227] Los ejemplos de diagnósticos de T21, T13, T18 y un caso de síndrome de Turner obtenidos usando los cromosomas normalizadores, las dosis de cromosomas y la diferenciabilidad para cada uno de los cromosomas de interés se describen en el Ejemplo 3.

Ejemplo 5

Diagnóstico de la aneuploidía fetal mediante el uso de cromosomas normalizadores

[0228] Para aplicar el uso de dosis de cromosomas para evaluar la aneuploidía en una muestra de prueba biológica, se obtuvieron muestras de sangre materna de voluntarias embarazadas y se preparó el ADNcf, y se secuenció una

biblioteca de secuenciación según el Protocolo abreviado descrito en el Ejemplo 2 y analizado.

Trisomia 21

[0229] La Tabla 4 proporciona la dosis calculada para el cromosoma 21 en una muestra de prueba ejemplar (n° 11403). El umbral calculado para el diagnóstico positivo de la aneuploidía T21 se estableció en >2 desviaciones estándar de la media de las muestras calificadas (normales). Se proporcionó un diagnóstico para T21 basado en que la dosis de cromosoma en la muestra de prueba fue mayor que el umbral establecido. Los cromosomas 14 y 15 se usaron como cromosomas normalizadores en cálculos separados para mostrar que cualquiera de los cromosomas tiene la variabilidad más baja *p. ej.* el cromosoma 14, o un cromosoma que tiene la mayor diferenciabilidad *p. ej.* el cromosoma 15, se puede utilizar para identificar la aneuploidía. Trece muestras de T21 fueron identificadas usando las dosis de cromosomas calculadas y las muestras de aneuploidía eran T21 por cariotipo.

TABLA 4

Dosis de cromosoma para una aneuploidía T21 (muestra # 11403, 47 XY +21)			
Cromosoma	Densidad de la etiqueta de secuencia	Dosis de cromosoma para Chr 21	Límite
Chr21	333.660	0,419672	0,412696
Chr14	795.050		
Chr21	333.660	0,441038	0,433978
Chr15	756.533		

Trisomia 18

[0230] La Tabla 5 proporciona la dosis calculada para el cromosoma 18 en una muestra de prueba (n° 11390). El umbral calculado para el diagnóstico positivo de la aneuploidía T18 se estableció en 2 desviaciones estándar de la media de las muestras calificadas (normales). Se proporcionó un diagnóstico para T18 basado en que la dosis de cromosoma en la muestra de prueba fue mayor que el umbral establecido. El cromosoma 8 se utilizó como cromosoma normalizador. En este caso, el cromosoma 8 tuvo la variabilidad más baja y la mayor diferenciabilidad. Se identificaron ocho muestras de T18 utilizando dosis de cromosomas, y se confirmó que eran T18 por cariotipo.

[0231] Estos datos muestran que un cromosoma normalizador puede tener tanto la variabilidad más baja como la mayor diferencia.

TABLA 5

Dosis de cromosoma para una aneuploidía T21 (muestra n° 11403, 47 XY +21)			
Cromosoma	Densidad de la etiqueta de secuencia	Dosis de cromosoma para Chr 21	Umbral
Chr18	602.506	0,585069	0,530867
Chr8	1.029.803		

Trisomia 13

[0232] La tabla 6 proporciona la dosis calculada para el cromosoma 13 en una muestra de prueba (n° 51236). El umbral calculado para el diagnóstico positivo de aneuploidía T13 se estableció en 2 desviaciones estándar de la media de las muestras calificadas. Se proporcionó un diagnóstico para T13 basado en que la dosis de cromosoma en la muestra de prueba fue mayor que el umbral establecido. La dosis de cromosoma para el cromosoma 13 se calculó utilizando el cromosoma 5 o el grupo de cromosomas 3, 4, 5 y 6 como el cromosoma normalizador. Se identificó una muestra de T13.

TABLA 6

Dosis de cromosoma para una aneuploidía T21 (muestra n° 11403, 47 XY +21)			
Cromosoma	Densidad de la etiqueta de secuencia	Dosis de cromosoma para Chr 21	Umbral
Chr13	692.242	0,541343	0,52594
Chr5	1.278.749		
Chr13	692.242	0,530472	0,513647
Chr3-6 [promedio]	1.304.954		

[0233] La densidad de etiqueta de secuencia para los cromosomas 3-6 es los recuentos promedio de etiqueta para los cromosomas 3-6.

[0234] Los datos muestran que la combinación de los cromosomas 3, 4, 5 y 6 proporciona una variabilidad menor que la del cromosoma 5, y la mayor diferenciabilidad que cualquiera de los otros cromosomas.

[0235] Por lo tanto, se puede usar un grupo de cromosomas como el cromosoma normalizador para determinar las dosis de cromosomas e identificar aneuploidías.

Síndrome de Turner (monosomía X)

[0236] La Tabla 7 proporciona la dosis calculada para los cromosomas X e Y en una muestra de prueba (n° 51238). El umbral 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 muestras calificadas (normales).

TABLA 7

Dosis de cromosoma para una aneuploidía de Turners (XO) (muestra n.º 51238, 45 X)			
Cromosoma	Densidad de la etiqueta de secuencia	Dosis de cromosoma para Chr X y Chr Y	Umbral
ChrX	873.631	0,786642	0,803832
Chr4	1.110.582		
ChrY	1.321	0,001542101	0,00211208
Chr_Total (1-22, X) (promedio)	856.623.6		

[0237] Se comprobó que una muestra que tiene una dosis del cromosoma X menor que la del umbral conjunto tenía menos de un cromosoma X. Se determinó que la misma muestra tenía una dosis de cromosoma Y que era menor que el umbral establecido, lo que indica que la muestra no tenía un cromosoma Y. Por lo tanto, la combinación de dosis de cromosomas para X e Y se usó para identificar las muestras del Síndrome de Turner (monosomía X).

[0238] Por lo tanto, el método descrito en el presente documento permite la determinación de CNV de cromosomas. En particular, el método permite la determinación de aneuploidías cromosómicas por encima y por debajo de la representación mediante la secuenciación masiva en paralelo del plasma de la madre ADNcf y la identificación de cromosomas normalizados para el análisis estadístico de los datos de secuenciación. La sensibilidad y la fiabilidad del método permiten realizar pruebas precisas de aneuploidía en el primer y segundo trimestre.

Ejemplo 6

Determinación de la aneuploidía parcial.

[0239] Se aplicó el uso de dosis de secuencia para la evaluación de la aneuploidía parcial en una muestra de ensayo biológica de ADNcf que fue preparada a partir de plasma de sangre, y se secuenció como se describe en el Ejemplo 1. Se confirmó el cariotipo que la muestra se derivó de un sujeto con una eliminación parcial del cromosoma 11.

[0240] El análisis de los datos de secuenciación para la aneuploidía parcial (eliminación parcial del cromosoma 11, es decir, q21-q23) se realizó como se describe para las aneuploidías cromosómicas en los ejemplos anteriores. El mapeo de las etiquetas de secuencia al cromosoma 11 en una muestra de prueba reveló una notable pérdida de conteos de etiquetas entre los pares de bases 81000082-103000103 en el brazo q del cromosoma en relación con los conteos de etiquetas obtenidos para la secuencia correspondiente en el cromosoma 11 en las muestras calificadas (datos no mostrados). Las etiquetas de secuencia asignadas a la secuencia de interés en el cromosoma 11 (81000082-103000103bp) en cada una de las muestras calificadas, y las etiquetas de secuencia asignadas a todos los segmentos de 20 megabases en todo el genoma en las muestras calificadas, es decir, densidades de etiquetas de secuencia calificadas, se utilizaron para determinar las dosis de secuencias calificadas como relaciones de densidades de etiquetas en todas las muestras calificadas. La dosis media de secuencia, la desviación estándar y el coeficiente de variación se calcularon para los 20 segmentos de megabase en todo el genoma, y la secuencia de 20 megabase que tuvo la menor variabilidad fue la secuencia de normalización identificada en el cromosoma 5 (13000014-33000033bp) (véase tabla 8), que se utilizó para calcular la dosis para la secuencia de interés en la muestra de prueba (consulte la Tabla 9). La Tabla 8 proporciona la secuencia de dosis para la secuencia de interés en el cromosoma 11 (81000082-103000103bp) en la muestra de prueba que se calculó como la proporción de etiquetas de secuencia asignadas a la secuencia de interés y las etiquetas de secuencia asignadas a la secuencia de normalización identificada. La Figura 16 muestra las dosis de secuencia para la secuencia de interés en las 7

muestras calificadas (O) y la dosis de secuencia para la secuencia correspondiente en la muestra de prueba (◇). La media se muestra mediante la línea continua y el umbral calculado para el diagnóstico positivo de aneuploidía parcial que se estableció en 5 desviaciones estándar de la media se muestra mediante la línea discontinua. El diagnóstico de aneuploidía parcial se basó en que la dosis de secuencia en la muestra de prueba era inferior al umbral establecido. La muestra de prueba se verificó mediante cariotipo para tener una eliminación q21-q23 en el cromosoma 11.

[0241] Por lo tanto, además de la identificación de aneuploidías cromosómicas, el método de la descripción se puede utilizar para identificar aneuploidías parciales.

TABLA 8

Secuencia de normalización calificada, dosis y varianza para la secuencia Chr11: 81000082-103000103 (muestras calificadas n=7)			
	Chr11: 81000082-103000103		
	Promedio	DesEst	CV
Chr5: 13000014-33000033	1,164702	0,004914	0,42

TABLA 9

Dosis de secuencia de secuencia de interés (81000082-103000103) en el cromosoma 11 (muestra de prueba 11206)			
Segmento de cromosoma	Densidad de la etiqueta de secuencia	Dosis del segmento de cromosoma para Chr 11 (q21-q23)	Umbral
Chr11: 81000082-103000103	27.052	1,0434313	1,1401347
Chr5: 13000014-33000033	25.926		

Ejemplo 7

Determinación simultánea de aneuploidía y fracción fetal mediante secuenciación masivamente paralela: selección de SNP autosómicos para la determinación de fracción fetal

[0242] Un conjunto de 28 SNP autosómicos fueron seleccionados de una lista de 92 SNP (Pakstis et al, Hum Genet 127: 315-324 [2010]), y a partir de secuencias de SNP disponibles en Applied Biosystems en dirección de Internet appliedbiosystems.com, y validado para su uso en amplificación por PCR multiplexada y para secuenciación masivamente paralela para determinar la fracción fetal con o sin determinar simultáneamente la presencia o ausencia de aneuploidía. Los cebadores se diseñaron para hibridar con una secuencia cercana al sitio de los SNP en el ADNcf para asegurar que se incluya en la lectura de 36 pb generada a partir de la secuenciación masivamente paralela en el Analizador GII de Illumina, y para generar amplicones de longitud suficiente para sufrir una amplificación de puente durante la formación de racimos. Por lo tanto, los cebadores se diseñaron para generar amplicones que tenían al menos 110 pb, que cuando se combinaron con los adaptadores universales (Illumina Inc., San Diego, CA) utilizados para la amplificación de grupos, dieron como resultado moléculas de ADN de al menos 200 pb. Se identificaron secuencias de los cebadores, y conjuntos de cebadores, es decir, cebadores directo e inversos, fueron sintetizados por Integrated DNA Technologies (San Diego, CA), y se almacenaron como una solución de 1µM para ser utilizado para la amplificación de secuencias diana polimórficas como se describe en los Ejemplos 5-8. La Tabla 10 proporciona los números de ID de acceso de RefSNP (rs), los cebadores utilizados para amplificar la secuencia de ADNc de destino y las secuencias de los amplicones que comprenden los posibles alelos de SNP que se generarían utilizando los cebadores. Los SNP dados en la Tabla 10 se usaron para la amplificación simultánea de 13 secuencias diana en un ensayo multiplexado para determinar simultáneamente la fracción fetal y la presencia o ausencia de una aneuploidía en muestras de ADNcf derivadas de mujeres embarazadas. El panel proporcionado en la Tabla 10 es un panel de SNP ejemplar. Se pueden emplear menos o más SNP para enriquecer el ADN fetal y materno para ácidos nucleicos diana polimórficos. Los SNP adicionales que se pueden usar incluyen los SNP que figuran en la Tabla 11. Los SNP en la Tabla 11 se validaron en amplificaciones de PCR multiplex, y se secuenciaron con el analizador Genomell A como se describe anteriormente. Los alelos SNP en las Tablas 10 y 11 se muestran en negrita y están subrayados.

TABLA 10

Panel SNP para la Determinación de Fracción Fetal					
SNP ID	Chr	Amplicon: Alelo 1	Amplicon: Alelo 2	Primera secuencia de cebador, nombre y SEQ ID NO:	Secuencia de cebador inversa, nombre y SEQ ID NO:
rs560681	1	CACATGCACAGCCA GCAACCTGTGACG AGGAGTTCCACCA GTTTCTTCTGAGAA CATCTGTTACAGTTT CTCTCCATCTCTATT TACTCAGGTCACAG GACCTTGGGG (SEQ ID NO:1)	CACATGCACAGCCA GCAACCTGTGACG AGGAGTTCCACCA GTTTCTTCTGAGAA CATCTGTTACAGTTT CTCTCCATCTCTATT TACTCAGGTCACAG GACCTTGGGG (SEQ ID NO:2)	CACATGCA CAGCCAGC AACCC (rs560681_C 1_1_F; SEQ ID NO:57)	CCCCAAGGTC CTGTGACCTGA GT (rs560681_C1_1_ R; SEQ ID NO:58)
rs1109037	2	TGAGGAAAGTGAGGC TCAGAGGGTAAGAA ACTTTGTCACAGAGC TGGTGGTGAGGGTG GAGATTTTACACTCC CTGCCTCCCACACCA GTTTCTCCAGAGTGG AAAGACTTTTCATCTC GCACTGGCA (SEQ ID NO:3)	TGAGGAAAGTGAGGC TCAGAGGGTAAGAA ACTTTGTCACAGAGC TGGTGGTGAGGGTG GAGATTTTACACTCC CTGCCTCCCACACCA GTTTCTCCGAGTGG AAAGACTTTTCATCTC GCACTGGCA (SEQ ID NO:4)	TGAGGAAG TGAGGCTC AGAGGGT (rs110937_C 2_1_F; SEQ ID NO:59)	TGCCAGTGCG AGATGAAAGT CTTT (rs110937_C2_1_ R; SEQ ID NO:60)
rs9866013	3	GTGCCTTCAGAACCT TTGAGATCTGATTCT ATTTTAAAGCTTCT TAGAAGAGAGATTG CAAAGTGGGTGTTT CTCTAGCCAGACAG GGCAGGCAAAATAGG GGTGGCTGGTGGGA TGGGA (SEQ ID NO:5)	GTGCCTTCAGAACCT TTGAGATCTGATTCT ATTTTAAAGCTTCT TAGAAGAGAGATTG CAAAGTGGGTGTTT CTCTAGCCAGACAG GGCAGGTAATAGG GGTGGCTGGTGGGA TGGGA (SEQ ID NO:6)	GTGCCTTC AGAACCTT TGAGATCT GAT (rs9866013_C3_1_F; SEQ ID NO:61)	TCCCATCCAC CAGCCACCC (rs9866013_C3_1_ R; SEQ ID NO:62)

rs13182883	5	AGGTGTGTCCTCTT TTGTGAGGGGAGGG GTCCCTTCTGGCCTA GTAGAGGGCCTGGC CTGCAGTGAGCATTC AAATCCTCAAGCAA CAGGGTGGGGAGGT GGGACAAAGG (SEQ ID NO:7)	AGGTGTGTCCTCTT TTGTGAGGGGAGGG GTCCCTTCTGGCCTA GTAGAGGGCCTGGC CTGCAGTGAGCATTC AAATCCTCGAGCAA CAGGGTGGGGAGGT GGGACAAAGG (SEQ ID NO:8)	AGGTGTGT CTCTCTTT GTGAGGGG (rs13182883_C5_1_F; SEQ ID NO:63)	CCTTTGTCCCA CCTCCCCACC (rs13182883_C5_1_R; SEQ ID NO:64)
rs13218440	6	CCTCGCCTACTGTGC TGTTTCTAACCATCA TGCTTTTCCCTGAAT CTCTTGAGTCTTTTT CTGCTGTGGACTGA AACTTGATCCTGAG ATTCACTCTAGTCC CTCTGAGCAGCCTCC TGGAACTACTCAGCT GGGATGG (SEQ ID NO:9)	CCTCGCCTACTGTGC TGTTTCTAACCATCA TGCTTTTCCCTGAAT CTCTTGAGTCTTTTT CTGCTGTGGACTGA AACTTGATCCTGAG ATTCACTCTAGTCC CTCTGGGAGCCTCC TGGAACTACTCAGCT 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)
rs7041158	9	AATTGCAATGGTGA GAGGTGATGGTAA AATCAAACGGAAT TGTTATTTGTCAAT CTGATGGACTGGAA CTGAGGATTTTCAAT TTCTCTCCAAACCCA AGACACTTCTCACTG G (SEQ ID NO:11)	AATTGCAATGGTGA GAGGTGATGGTAA AATCAAACGGAAT TGTTATTTGTCAAT CTGATGGACTGGAA CTGAGGATTTTCAAT TTCTTTTCCAAACCCA AGACACTTCTCACTG G (SEQ ID NO:12)	AATTGCAA	CCAGTGAGAA
				TGGTGAGA GGTTGATG GT (SEQ ID NO:67)	GTGTCTTGGGT TGG (SEQ ID NO:68)

rs740598	10	GAAATGCCTTCTCAG GTAATGGAAGTTA TCCAAATATTTTCG TAAAGTATTTCAAATA GCAATGGCTCGTCTA TGGTTAGTCTCAGAG CCACATTCTCAGAAC TGCTCAAACC (SEQ ID NO:13)	GAAATGCC TTCTCAGG TAATGGAA GGT (SEQ ID NO:69)	GGTTTGAGCA GTTCTGAGAAT GTGGCT (SEQ ID NO:70)
rs10773760	12	ACCCAAAACACTGG AGGGCCCTCTTCTCA TTTTCGGTAGACTGC AAGTGTAGCCGTC GGGACCAAGTTCTGT CTGGAAGTTCTGTC AATTGCAGTTAAGTC CAAGTATGCCACAT AGCAGATAAGGG (SEQ ID NO:15)	ACCCAAA CACTGGAG GGGCCT (SEQ ID NO:71)	CCCTTATCTGC TATGTGGCATA CTTGG (SEQ ID NO:72)
rs4530059	14	GCACCAGAAATTAA ACAACGCTGACAAT AAATATGCAGTCGA TGATGACTTCCAGA GCTCCAGAAAGCAAC TCCAGCACACAGAG AGGCGCTGATGTGC CTGTCAGGTGC (SEQ ID NO:17)	GCACCAGA ATTTAAAC AACGCTGA CAA (SEQ ID NO:73)	GCACCTGACA GGCACATCAG CG (SEQ ID NO:74)

rs7205345	16	TGACTGTATACCCCA GGTGACCCCTTGGGT CATCTCTATCATAGA ACTTATCTCACAGAG TATAAGAGCTGATTT CTGTGTCTGCCCTTC ACACTAGACTTCCAC ATCCTTAGTGC (SEQ ID NO:19)	TGACTGTATACCCCA GGTGACCCCTTGGGT CATCTCTATCATAGA ACTTATCTCACAGAG TATAAGAGCTGATTT CTGTGTCTGCCCTTC ACACTAGACTTCCAC ATCCTTAGTGC (SEQ ID NO:20)	TGACTGTA TACCCAG GTGCACCC (SEQ ID NO:75)	GCACTAAGGA TGTGGAAAGTCT AGTGTG (SEQ ID NO:76)
rs8078417	17	TGACGTGGTCAACCA GGGGACGCCCTGGCG CTGCGAGGGAGGCC CCGAGCCTCGTGCCC CCGTGAAGCTTCAG CTCCCCTCCC ^{GG} CT GTCCTTGAGGCTCTT CTCACACT (SEQ ID NO:21)	TGACGTGGTCAACCA GGGGACGCCCTGGCG CTGCGAGGGAGGCC CCGAGCCTCGTGCCC CCGTGAAGCTTCAG CTCCCCTCCC ^{GG} CT GTCCTTGAGGCTCTT CTCACACT (SEQ ID NO:22)	TGTACGTG GTCACAG GGGACG (SEQ ID NO:77)	AGTGTGAGAA GAGCCTCAAG GACAGC (SEQ ID NO:78)
rs576261	19	CAGTGGACCCCTGCT GCACCTTTCTCC ^{CC} TCCCATCAACCTCTT TTGTGCCCTCCCC ^{CT} CGGTACCACTTCT CTGTCACCAACCCTG GCCTCACAACTCTCT	CAGTGGACCCCTGCT GCACCTTTCTCC ^{CC} TCCCATCAACCTCTT TTGTGCCCTCCCC ^{CT} CGGTACCACTTCT CTGTCACCAACCCTG GCCTCACAACTCTCT	CAGTGGAC CCTGCTGC ACCTT (SEQ ID NO:79)	GTGGCAAAGG AGAGAGTTGT GAGG (SEQ ID NO:80)
		CCTTTGCCAC (SEQ ID NO:23)	CCTTTGCCAC (SEQ ID NO:24)		

rs2567608	20	CAGTGGCATAAGTAG TCCAGGGGCTCCTCC TCAGCACCTCCAGC ACCTCCAGGAGGC AGCAGGCAGGCAG AGAACCCGCTGGAA GAATCGGCGGAAGT TGTCGGAGAGG (SEQ ID NO:25)	CAGTGGCA TAGTAGTC CAGGGGCT (SEQ ID NO:81)	CCTCTCCGACA ACTTCCGCCG (SEQ ID NO:82)
-----------	----	--	--	--

TABLA 11

Panel SNP para la Determinación de Fracción Fetal					
SNP ID	Chr	Amplicon: Alelo 1	Amplicon: Alelo 2	Primera secuencia de cebador, nombre y SEQ ID NO:	Secuencia de cebador inversa, nombre y SEQ ID NO:
rs430046	18	AGGTCGGGGCC GCTGAATGCCAAGC TGGGAATCTTAAAT GTTAAGGAACAAG GTCAATACAAATGAAT GGTGTGATGTATAA GCTTGGGAGCTGAT TTCTGAGGGTACGT GCTGGGTTTAAATGG GAGGA (SEQ ID NO:27)	AGGTCGGGGGGCC TGAATGCCAAGCTGG GAATCTTAAATGTTA AGGAACAAGGTGATA CAATGAATGTTGTGA TGTAAAGCTTGGGA GGTGA'TTTTGGAGGG TAGGTGCTGGGTTTA ATGGGAGGA (SEQ ID NO:28)	AGGTC TGG GGCGCGCT GAAT (rs430046_C 1_1_F; SEQ ID NO:83)	TCCTCCCATTA AACCAGCAC CT (rs430046_C1_1_ R; SEQ ID NO:84)
rs9951171	18	ACGGTCTGTCTGT TAGGGAGAAAAG TCCTCGTTGTCTCT CTGGGATGCAACAT GAGAGAGCAGCAC ACTGAGGCTTTATG GATTGCCCTGCCAC AAGTGAACAGG (SEQ ID NO:29)	ACGGTCTGTCTGT AGGGAGAAAAGTCC TCGTTGTCTCTGG ATGCAACATGAGAGA GCAGCACACTGAGGC TTTATGG'GTTGCCCTG CCACAAGTGAACAGG (SEQ ID NO:30)	ACGGTTCT GTCTCTGTA GGGGAGA (rs9951171_C1_1 _R; SEQ ID ID NO:85)	CCTGTTCACCTT GTGGCAGGGC A (rs9951171_C1_1 _R; SEQ ID NO:86)
rs335882	5	GCGCAGTCAGATG GGCGTCTGGCGTC TGCTCTCTCTCTC CTGCTCTGGCTT CA'TTTCTCTCTCT CTGTCTCACCTTCT TTGCTGGCTCTGTG CACACACAGTTTG GGACAAAGG CTGGA (SEQ ID NO:31)	GCGCAGTCAGATGGG CGTGTGGCGTCTGT CTTCTCTCTCTCTG TCTCTGGCTTCATTT TCTCTCTCTCTCTCT ACCTTCTTCTGTGTG CTGTGCATACACAG TTTGGGACAAGGG CTGGA (SEQ ID NO:32)	GCGCAGTC AGATGGGC GTGC (rs335882_C 1_1_F; SEQ ID NO:87)	TCCAGCCCTTG TCCCAAAACGT GT (rs335882_C1_1_ R; SEQ ID NO:88)

rs10776839	9	GCCGGACCTGGGAA ATCCCAAAATGCC AAACATTCCCGCT CACATGATCCAGA GAGAGGGGAGCCCA GTGTCCCAAGCTTG CAGCTGAGAGGCC CGAGGTTGCCGTCA GATCAGAGCCCCA GTTGCCCG (SEQ ID NO:33)	GCCGGACCTGGGAAA TCCCAAAATGCCAAA CATTCGGCTCACA TGATCCAGAGAGAG GGGACCCAGTGTTC CAGCTTGAGCTGAG GAGCCCGAGTTGCC GTCAGATCAGAGCCC CAGTTGCCCG (SEQ ID NO:34)	GCCGGACC TGCAGAAAT CCCAA (rs10776839 C1_F; SEQ ID NO:89)	CGGGCAACTG GGGCTCTGATC (rs10776839_C1_ _R; SEQ ID NO:90)
rs9905977	17	AGCAGCCTCCCTCG ACTAGCTCACACTA CGATAAGGAAAT TCATGAGCTGGTGT CCAAGGAGGCTG GGTGACTCGTGGCT CAGTCAGCATCAAG ATCCTTTCGCTTT CCCCCTCTGCC (SEQ ID NO:35) AGCAGCCTCCCTCGA CTAGCTCACACTACG ATAAGGAAATTCAT GAGCTGGTGTCCAG GAGGGCTGGGTGACT CGTGGCTCAGTCAGC GTCAAGATTCCTTC GTCTTCCCTCTGCC (SEQ ID NO:36)	AGCAGCCT CCCTCGAC TAGCT (rs9905977_C1_ _F; SEQ ID NO:91)	GGCAGAGGGG AAAGACGAA GGA (rs9905977_C1_ _R; SEQ ID NO:92)	
rs1277284	4	TGGCAITGCTGTGA ATACATAGCCAT GGTTTTTATAGGC AATTAAGATGAAT AGCTTCTAAACTAT AGATAAGTTTCATT ACCCAGGAAGCT GAACTATAGCTACT TTACCCAAATCAT TAGAATGGTCTT (SEQ ID NO:37) TGGCATTGCTGTAA TATACATAGCCATGG TTTTTTATAGGCAAT TAAGATGAATAGCTT CTAAACTATAGATAA GTTTCATTACCCAG GAAGCTGAACATATAG CTACTTCCCAAAA TCATTAGAAATGGTGC TT (SEQ ID NO:38)	TGGCATTG CCTGTAAT ATACATAG (rs1277284_C4_ _F; SEQ ID NO:93)	AAGCACCAT CTAATGATTTT GG (rs1277284_C4_ _R; SEQ ID NO:94)	

rs258684	7	ATGAAGCCTTCCAC CAACTGCCTGTATG ACTCATCTGGGAC TTCTGCTCTACT CAAAGTGGCTTAGT CACTGCCAATGTAT TTCCATATGAGGA CGATGATTACTAAG GAAATATAGAAAC AACAACCTGATC (SEQ ID NO:39)	ATGAAGCCTTCCACC AACTGCCTGTATGAC TCATCTGGGACTTC TGCTCTATACTCAA GTGGCTTAGTCACTG CCAATGTATTCCAT ATGAGGGACGGTGAT TACTAAGGAAATATA GAAACAACAACCTGAT C (SEQ ID NO:40)	ATGAAGCC TTCCACCA ACTG (rs258684_C 7_1_F; SEQ ID NO:95)	GATCAGTTGTT GTTTCTATATT TCCTT (rs258684_C7_1_ R; SEQ ID NO:96)
rs1347696	8	ACAACAGAAATCAG GTGATTGGAGAAA AGATCACAGGCCTA GGCACCCAAAGGCTT GAAGGATGAAAGA ATGAAAAGATGGAC GGAACAAAATTAG GACCTTAATCTTT GTTCAAGTTCAG (SEQ ID NO:41)	ACAACAGA ATCAGGTG ATTGGA (rs1347696_ C8_4_F; SEQ ID NO:97)	CTGAACCTGAA CAAAGAAATA AGGTC (rs1347696_C8_4 _F; SEQ ID NO:98)	
rs508485	11	TTGGGGTAAATTTT CATTGTCATATGTG GAATTTAAATATAC CATCATCTACAAAG AATCCACACAGATT AAATATCTTAAAGTT AAACACTTAAATA AGTGTTCGGTGAT ATTTGATGACAGA TAAACAGAGTCTAA TTCCCACCCC (SEQ ID NO:43)	TTGGGGTA AATTTTCA TTGTCA (rs508485_C 11_1_F; SEQ ID NO:99)	GGGGTGGGAA TTAGACTCTG (rs508485_C11_1 _R; SEQ ID NO:100)	

rs9788670	15	TGCAATTCAAATCA GGAAGTATGACCA AAAGACAGAGATC	TGCAATTCAAATCAG GAAATATGACCAAAA GACAGAGATCTTTT	TGCAATTC AAATCAGG AAGTATG	GCAACATCGA GGTTTGTCTCAG (rs9788670_c15_
		TTTTTTGGATGATC CCTAGCCTAGCAAT GCCTGGCAGCCATG CAGGTGCAATGTCA ACCTTAAATAATGT ATTGCAAACTCAGA GCTGACAAACCTCG ATGTTGC (SEQ ID NO:45)	TGGATGATCCCTAGC CTAGCAATGCCTGGC AGCCATGCAGGTGCA ATGTCAACCTTAAAT AATGTATTGCAAAAT CAGAGCTGACAAACC TCGATGTTGC (SEQ ID NO:46)	(rs9788670_c 15_2_F: SEQ ID NO:101)	2_R: SEQ ID NO:102)
rs8137254	22	CTGTGCTCTGCGAA TAGCTGCAGAAAGTA ACTTGGGGACCCAA AATAAAGCAGAAAT GCTAATGTCAAAGTC CTGAGAACCAAGC CCTGGGACTCTGGT GCCATTTCGGATTC TCCATGAGCATGGT (SEQ ID NO:47)	CTGTGCTCTGCGAAT AGCTGCAGAAAGTAAC TTGGGGACCCAAAT AAAGCAGAAATGCTAA TGTCAAAGTCTTGAGA ACCAAAGCCCTGGGAC TCTGGTGCCATTITGG ATTCTCCATGAGCAT GGT (SEQ ID NO:48)	CTGTGCTC TGCGAATA GCTG (rs8137254_c 22_2_F: SEQ ID NO:103)	ACCATGCTCAT GGAGAATCC (rs8137254_c22_2_R: SEQ ID NO:104)
rs3143	19	TTTTTCCAGCCAAC TCAAGGCCAAAAA AAATTCTTAAATAT AGTTATTATGCGAG GGGAGGGGAAGCA AAGGAGCACAGGT AGTCCACAGAATA AGACACAAGAAAC CTCAAGCTGTG (SEQ ID NO:49)	TTTTTCCAGCCAACCTC AAGGCCAAAAAAAT TTCTTAATATAGTTAT TATGCGAGGGGAGGG GAAGCAAAAGGAGCA CAGGTAGTCCACAGA ATAGGACACAAGAAA CCTCAAGCTGTG (SEQ ID NO:50)	TTTTTCCA GCCAACTC AAGG (rs3143_c19_2_F: SEQ ID NO:105)	CACAGCTTGA GGTTTCTTGTG (rs3143_c19_2_R: SEQ ID NO:106)

rs2182957	13	TCTTCTCGTCCCCT AAGCAAAACAACAT CCGCTTGCTTCTGT CTGTGTAACACACAG TGAATGGGTGTGCA CGCTTGATGGGCT CTGAGCCCTGTGTG CACAAACACAGAAA (SEQ ID NO:51)	TCTTCTCGTCCCCTAA GCAAAACAACATCCGC TTGCTTCTGTCTGTGT AACCACAGTGAATGG GTGTGACGCTGTGTG GGGCTCTGAGCCCC TGTTCACAAAACACAG AAA (SEQ ID NO:52)	TCTTCTCG TCCCCTAA GCAA (rs2182957_c 13_1_F; SEQ ID NO:107)	TTTCTGTTTG TGCAACACAG (rs2182957_c13_ 1_R; SEQ ID NO:108)
rs3739005	2	CACATGGGGGCATT AAGAAATCGCCAG GGAGGAGGAGGGA GAACGCTGCTTTT CACATTGCAATTG AATTTTCGAGTTCC CAGGATGTGTTTTT GTGCTCATCGATGT (SEQ ID NO:53)	CACATGGGGGCATTA AGAAATCGCCAGGGA GGAGGAGGAGGAAAC GCGTCTTTTTCACATT TGCAATTTGAATTTTG AGTTCCAGGATGTG TTTTTGTGCTCATCGA TGT (SEQ ID NO:54)	CACATGGG GGCATTA GAAT (rs3739005_c 2_2_F; SEQ ID NO:109)	ACATCGATGA GCACAAAAAC AC (rs3739005_c2_2_ _R; SEQ ID NO:110)
rs530022	1	GGGCTCTGAGGTGT GTGAAATAAAAC AAATGTCCATGTCT GTCCCTTTTATGGCA TTTGGGACTTTAC ATTTCAAACATTTC AGACATGTATCACA ACACGAAGGAATA ACAGTTCCAGGGAT ATCT (SEQ ID NO:55)	GGGCTCTGAGGTGTG TGAAATAAAACAAA TGTCATGTCTGTCTCT TTATGGCATTTTGGG ACTTACATTTCAA CATTTACAGACATGTA TCACAACACGAGGGA ATAACAGTTCCAGGG ATATCT (SEQ ID NO:56)	GGGCTCTG AGGTGTGT GAAA (rs530022_c1 _2_F; SEQ ID NO:111)	AGATATCCCTG GAACTGTATT CC (rs530022_c1_2_ _R; SEQ ID NO:112)

Ejemplo 8**Determinación simultánea de aneuploidía y fracción fetal: enriquecimiento de los ácidos nucleicos maternos y fetales en una muestra de la biblioteca de secuenciación de ADNcf.**

[0243] Para determinar simultáneamente la fracción fetal y la presencia o ausencia de una aneuploidía en una muestra materna, se enriqueció una biblioteca de secuenciación primaria de ácidos nucleicos maternos y fetales para las secuencias de ácido nucleico diana polimórficas, y se secuenció como sigue.

[0244] ADNcf purificada se preparó a partir de una muestra de plasma materno como se describe en el Ejemplo 1. Se utilizó una primera porción de la muestra ADNcf purificado para preparar una biblioteca de secuenciación primaria utilizando el protocolo abreviado descrito en el Ejemplo 2. Una segunda porción de la ADNcf purificada la muestra se usó para amplificar secuencias de ácido nucleico diana polimórficas, es decir, SNP, y preparar una biblioteca de secuencia de destino de la siguiente manera. El ADNcf contenido en 5 µl de ADNcf purificado se amplificó en un volumen de reacción de 50 µl que contenía 7,5 µl de una mezcla de cebador 1 µm (Tabla 5), 10 µl de NEB 5X Mastermix y 27 µl de agua. El ciclo térmico se realizó con el Gene Amp9700 (Applied Biosystems). Usando las siguientes condiciones de ciclado: incubación a 95°C durante 1 minuto, seguido de 30 ciclos a 95°C durante 20 segundos, 68°C durante 1 minuto y 68°C durante 30s, seguido de una incubación final a 68°C durante 5 minutos. Se agregó una retención final a 4°C hasta que las muestras se retiraron para combinar con la porción no amplificada de la muestra de ADNc purificada. El producto amplificado se purificó utilizando el sistema de purificación por PCR Agencourt AMPure XP (Pieza N° A63881; Beckman Coulter Genomics, Danvers, MA). Se agregó una retención final a 4°C hasta que las muestras se retiraron para preparar la biblioteca diana. El producto amplificado se analizó con un bioanalizador 2100 (Agilent Technologies, Sunnyvale, CA) y se determinó la concentración del producto amplificado. Una quinta parte del producto amplificado purificado se usó para preparar una biblioteca de secuenciación diana de ácidos nucleicos polimórficos amplificados como se describe en el Ejemplo 2. Las bibliotecas de secuenciación principal y la diana se diluyeron cada una a 10 nM, y la biblioteca diana se combinó en una proporción de 1:9 con la biblioteca de secuenciación para proporcionar una biblioteca de secuenciación enriquecida. La secuenciación de la biblioteca enriquecida se realizó como se describe en el Ejemplo. El análisis de los datos de secuenciación para determinar la aneuploidía se realizó como se describe en el Ejemplo 3 utilizando el genoma humano hg18 como genoma de referencia. El análisis de los datos de secuenciación para determinar la fracción fetal se realizó de la siguiente manera concomitante con el análisis para determinar la aneuploidía, se analizaron los datos de secuenciación para determinar la fracción fetal. Tras la transferencia de la imagen y los archivos de llamada base al servidor Unix que ejecuta la versión 1,51 del software Illumina "Genoma Analyzer Pipeline", como se describe en el Ejemplo 2c., las lecturas de 36 pb se alinearon con un "genoma SNP" utilizando el programa BOWTIE. El genoma de SNP se identificó como la agrupación de las secuencias de ADN polimórficas, es decir, las SEQ ID NO: 1-56, que abarcan los alelos de los 13 SNP descritos en la Tabla 10 en el Ejemplo 7. Sólo se utilizaron las lecturas que se asignaron de forma única al genoma de SNP para el análisis de la fracción fetal. Las lecturas que coincidían perfectamente con el genoma de SNP se contaron como etiquetas y se filtraron. De las lecturas restantes, solo las lecturas con uno o dos desajustes se contaron como etiquetas y se incluyeron en el análisis. Se contaron las etiquetas asignadas a cada uno de los alelos SNP y se determinó la fracción fetal. Aproximadamente un millón del número total de etiquetas de secuencia obtenidas de la secuenciación de la biblioteca enriquecida correspondió a las etiquetas que se asignan al genoma de referencia de SNP. **La Figura 17** muestra un gráfico de la proporción del número de etiquetas de secuencia asignadas a cada cromosoma y el número total de etiquetas asignadas a todos los cromosomas (1-22, X e Y) obtenidos de la secuenciación de una biblioteca de ADNcf no enriquecida (•), y biblioteca ADNcf enriquecida con 5% (j) o 10% (♦) biblioteca SNP multiplex amplificada. El gráfico indica que la combinación de una biblioteca de secuencias polimórficas amplificadas con una biblioteca de secuencias no amplificadas de la muestra materna no afecta la información de secuenciación utilizada para determinar la aneuploidía. Los ejemplos de determinación de la fracción fetal para muestras obtenidas de sujetos que portan un feto con una aneuploidía cromosómica se dan en las Tablas 12, 13 y 14 a continuación.

a. Determinación de la fracción fetal

[0245] La fracción fetal se calculó como:

$$\% \text{ fracción fetal alelo}_x = ((\Sigma \text{Etiquetas de secuencia fetal para alelo}_x) / (\Sigma \text{Etiquetas de secuencia materna para alelo}_x)) \times 100$$

donde alelo_x es un alelo informativo.

TABLA 12

Determinación simultánea de aneuploidía y fracción fetal: determinación de fracción fetal			
Muestra ID (Cariotipo)	SNP	RECuento DE ETIQUETAS SNP	FRACCIÓN FETAL (%)
11409 (47, XY+21)	rs13182883,1 Chr, 5 longitud=111 alelo=A	261	4,41
	rs13182883,2 Chr, 5 longitud=111 alelo=G	5918	
	rs740598,1 Chr,10 longitud=114 alelo=A	5545	7,30
	rs740598,2 Chr,10 longitud=114 alelo=G	405	
	rs8078417,1 Chr, 17 longitud=110 alelo=C	8189	6,74
	rs8078417,2 Chr, 17 longitud=110 alelo=T	121470	
	rs576261,1 Chr,19 longitud=114 alelo=A	58342	7,62
	rs576261,2 Chr,19 longitud=114 alelo=C	4443	
Fracción Fetal (Media±D.E.) = 6,53±1,45			
Muestra ID			
95133 (47, XX+18)	rs1109037,1 Chr,2 longitud=126 alelo=A	12229	2,15
	rs1109037,2 Chr,2 longitud=126 alelo=G	263	
	rs13218440, 1 Chr, 6 longitud=139 alelo=A	55949	3,09
	rs13218440,2 Chr, 6 longitud=139 alelo=G	1729	
	rs7041158,1 Chr, 9 longitud=1117 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	
Fetal Fraction (Mean±D.E.) = 2,9±0,9			
Muestra ID			
51236 (46,XY+13)	rs13218440,1 Chr, 6 longitud=139 alelo=A	1119	1,65
	rs13218440,2 Chr, 6 longitud=139 alelo=G	67756	
	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

Fracción fetal (Media+D.E.) = 2,9+0,9			
Muestra ID			
	rs9866013.2 Chr. 31 longitud=121 alelo=T	5024	
Fracción fetal (Media+D.E.) = 2,7+1,7			
Muestra ID			
54430 (45,XO)	rs1109037.1 Chr.2 longitud=126 alelo=A	19841	1,80
	rs1109037.2 Chr.2 longitud=126 alele=G	357	
	rs9866013.1 Chr. 31 longitud=121 alelo=C	12931	3,81
	rs9866013.2 Chr. 31 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.E.) = 3,9+1,2			

b. Determinación de la aneuploidía.

[0246] La determinación de la aneuploidía de los cromosomas 21, 13, 18 y X se realizó utilizando las dosis de cromosoma como se describe en el Ejemplo 4. La dosis de cromosoma calificada, la varianza y la diferenciabilidad para los cromosomas 21, 18, 13, X e Y se dan en las Tablas X y Y. La clasificación de los cromosomas normalizados identificados por las dosis de cromosomas determinadas a partir de la secuenciación de la biblioteca enriquecida fue la misma que la determinada por la secuenciación de una biblioteca primaria (no enriquecida) del Ejemplo 4. La **Figura 17** muestra la secuenciación de una biblioteca que ha sido enriquecida por secuencias diana de polimorfos. *p. ej.* SNP, no se ve afectada por la inclusión de los productos SNP amplificados.

TABLA 13

Dosis de cromosomas calificada, varianza y diferenciabilidad para los cromosomas 21 y 18								
	21 (n = 35)				18 (n = 40)			
	Promedio	Des. Est.	CV	Prueba t	Promedio	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
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 14

Qualified Chromosome Dose, Variance and Differentiability for chromosomes 13, X and Y								
	13 (n=47)				X (n=20)			
	Avg	Stdev	CV	Diff	Avg	Stdev	CV	T Test
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
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)							
	Promedio	Des. Est.	CV	Prueba T				
Chr 1-22, X	0,00728	0,00227	31,19	1,30E-13				

[0247] La dosis del cromosoma 21 se determinó utilizando el cromosoma 14 como el cromosoma de normalización; La dosis del cromosoma 13 se determinó utilizando el grupo de los cromosomas 3, 4, 5 y 6 como el cromosoma normalizador; la dosis del cromosoma 18 se determinó utilizando el cromosoma 8 como cromosoma normalizador; y la dosis del cromosoma X se determinó utilizando el cromosoma 4 como cromosoma normalizador. Los umbrales se calcularon para ser 2 desviaciones estándar por encima y por debajo de la media determinada en las muestras calificadas.

[0248] La tabla 12 muestra los datos para la determinación de la fracción fetal en muestras ejemplares. Los valores de dosis de cromosomas calculados para los cromosomas 21, 18, 13, X e Y en muestras de prueba ejemplares correspondientes se dan en las Tablas 15, 16, 17 y 18, respectivamente.

Trisomia 21

[0249] La Tabla 8 proporciona la dosis calculada para el cromosoma 21 en la muestra de prueba (11409). El cromosoma 14 se utilizó como cromosomas normalizadores. El umbral calculado para el diagnóstico positivo de la aneuploidia T21 se estableció en 2 desviaciones estándar de la media de las muestras calificadas (normales). Se proporcionó un diagnóstico para T21 basado en que la dosis de cromosoma en la muestra de prueba fue mayor que el umbral establecido. Las doce muestras de T21 que se confirmaron como T21 por cariotipo se identificaron en una población de 48 muestras de sangre.

TABLA 15

Dosis de cromosoma para una aneuploidía T21			
Cromosoma	Densidad de etiqueta de secuencia	Dosis de cromosoma para Chr 21	Umbral
Chr21	264.404	0,439498	0,410634
Chr14	601.605		

Trisomia 18

[0250] La Tabla 9 proporciona la dosis calculada para el cromosoma 18 en una muestra de prueba (95133). El cromosoma 8 se utilizó como cromosoma normalizador. En este caso, el cromosoma 8 tuvo la variabilidad más baja y la mayor diferenciabilidad. El umbral calculado para el diagnóstico positivo de la aneuploidía T18 se estableció en >2 desviaciones estándar de la media de las muestras calificadas (no T18). Se proporcionó un diagnóstico para T18 basado en la dosis de cromosoma en la muestra de prueba es mayor que el umbral establecido. Se identificaron ocho muestras de T18 utilizando dosis de cromosomas, y se confirmó que eran T18 mediante cariotipo.

TABLA 16

Dosis de cromosoma para una aneuploidía T18			
Cromosoma	Densidad de etiqueta de secuencia	Dosis de cromosoma para Chr 18	Umbral
Chr18	604.291	0,550731	0,533297
Chr8	1.097.253		

Trisomia 13

[0251] Las Tablas 10 y 11 proporcionan la dosis calculada para el cromosoma 13 en una muestra de prueba (51236). El umbral calculado para el diagnóstico positivo de 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 utilizando la densidad de la etiqueta de secuencia para el cromosoma 4 como el cromosoma normalizador, mientras que la dosis dada en la Tabla 11 se determinó utilizando el promedio de las relaciones de densidades de la etiqueta de secuencia para el grupo de cromosomas 3, 4, 5, y 6 como el cromosoma normalizante. Se proporcionó un diagnóstico para T13 basado en que la dosis de cromosoma en la muestra de prueba fue mayor que el umbral establecido. Se identificó una muestra de T13 utilizando dosis de cromosomas, y se confirmó que eran T13 mediante cariotipo.

[0252] Los datos muestran que la combinación de los cromosomas 3, 4, 5, y 6 proporcionan una variabilidad (1,06) que es similar a la del cromosoma 4 (1,01), lo que demuestra que un grupo de cromosomas se puede utilizar como cromosoma normalizador para determinar las dosis de cromosomas e identificar aneuploidías.

TABLA 17

Dosis de cromosoma para una aneuploidía T13			
Cromosoma	Densidad de la etiqueta de secuencia	Dosis de cromosoma para Chr 13	Umbral
Chr13	669.872	0,538140	0,536044
Chr4	1.244.791		

TABLA 18

Dosis cromosómica para una aneuploidía T13			
Cromosoma	Densidad de la etiqueta de secuencia	Dosis de cromosoma para Chr 13	Umbral
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)

[0253] Se identificaron tres muestras con una dosis de cromosoma inferior a la del umbral 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 umbral establecido, lo que indicaba que las muestras no tenían un cromosoma Y.

[0254] Las dosis calculadas para los cromosomas X e Y en la muestra de ensayo de monosomía X ejemplar (54430) se dan en la Tabla 12. El cromosoma 4 fue seleccionado como el cromosoma de normalización para calcular la dosis para el cromosoma X; y todos los cromosomas, es decir, 1-22, e Y, se utilizaron como los cromosomas normalizadores. El umbral 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 muestras calificadas (no monosomía X).

TABLA 19

Dosis de cromosoma para un síndrome de Turner (monosomía X)			
Cromosoma	Densidad de la etiqueta de secuencia	Dosis de cromosoma para Chr X	Umbral
ChrX	904.049	0,777990	0,797603
Chr4	1.162.031		
ChrY	390	0,0004462	0,002737754
Chr (1-22, X) (Promedio)	874.108,1		

[0255] De este modo, 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 ADNcf fetal y materno que se ha enriquecido para una pluralidad de secuencias polimórficas, comprendiendo cada una de ellas un SNP. En este ejemplo, la mezcla de ácidos nucleicos maternos y fetales se enriqueció mediante la combinación de una porción de una biblioteca de secuenciación que se construyó a partir de secuencias polimorfas maternas y fetales amplificadas con una biblioteca de secuenciación que se construyó a partir de la mezcla original de ADNcf fetal y materna no amplificada.

Ejemplo 9

Determinación simultánea de aneuploidía y fetal fracción: Enriquecimiento de los ácidos nucleicos maternos y fetales en una muestra de ADNcf purificada

[0256] Para enriquecer el ADNcf fetal y materno contenido en una muestra purificada de ADNcf extraída de una muestra de plasma materno, se usó una porción del ADNcf purificado para amplificar secuencias de ácido nucleico diana polimórficas, comprendiendo cada una de ellas un SNP seleccionado del panel de SNP dado en la Tabla 6.

[0257] Se obtuvo plasma libre de células de una muestra de sangre materna, y ADNcf se purificó a partir de la muestra de plasma como se describe en el Ejemplo 1. Se determinó que la concentración final era 92,8pg/μl.

[0258] El ADNcf contenido en 5 μl de ADNcf purificado se amplificó en un volumen de reacción de 50 μl que contenía 7,5 μl de una mezcla de cebador 1uM (Tabla 5), 10 μl de NEB 5X Mastermix y 27 μl de agua. El ciclo térmico se realizó con el Gene Amp9700 (Applied Biosystems). Usando las siguientes condiciones de ciclado: incubación a 95°C durante 1 minuto, seguido de 30 ciclos a 95°C durante 20 segundos, 68°C durante 1 minuto y 68°C durante 30s, seguido de una incubación final a 68°C durante 5 minutos. Se agregó una retención final a 4°C hasta que las muestras se retiraron para combinar con la porción no amplificada de la muestra de ADNc purificada. El producto amplificado se purificó utilizando el sistema de purificación por PCR Agencourt AMPure XP (Pieza N° A63881; Beckman Coulter Genomics, Danvers, MA), y la concentración se cuantificó utilizando el Nanodrop 2.000 (Thermo Scientific, Wilmington, DE). El producto de amplificación purificado se diluyó 1:10 en agua y se agregaron 0,9 μl (371 pg) a 40 μl de muestra de ADNcp purificada para obtener un pico del 10%. El ADNcf enriquecido fetal y materno presente en la muestra de ADNcf purificada se usó para preparar una biblioteca de secuenciación, y se secuenció como se describe en el Ejemplo 2.

[0259] La Tabla 13 proporciona los recuentos de etiquetas obtenidos para cada uno de los cromosomas 21, 18, 13, X e Y, es decir, densidad de la 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 etiquetas SNP. Los datos muestran que la información de secuenciación puede obtenerse a partir de la secuenciación de una biblioteca única construida a partir de una muestra de ADNcf materna purificada que se ha enriquecido para secuencias que comprenden SNP para determinar simultáneamente la presencia o ausencia de aneuploidía y la fracción fetal. En el ejemplo dado, los datos muestran que la fracción de ADN fetal en la muestra de plasma AFR105 fue cuantificable a

partir de los resultados de la secuenciación de cinco SNP informativos y se determinó que era del 3,84%. Se proporcionan densidades de etiquetas de secuencia para los cromosomas 21, 13, 18, X e Y. La muestra AFR105 fue la única muestra que se sometió al protocolo de enriquecimiento de ADNcf purificado para secuencias polimórficas amplificadas. Por lo tanto, no se proporcionaron los coeficientes de variación y las pruebas de diferenciabilidad. Sin embargo, el ejemplo muestra que el protocolo de enriquecimiento proporciona los recuentos de etiquetas necesarios para determinar la aneuploidía y la fracción fetal a partir de un solo proceso de secuenciación.

TABLA 20

Determinación simultánea de aneuploidía y fracción fetal: Enriquecimiento de los ácidos nucleicos maternos y fetales en una muestra de ADNcf purificada					
Aneuploidía					
	Cromosoma 21	Cromosoma 18	Cromosoma 13	Cromosoma X	Cromosoma Y
Densidad de la etiqueta de secuencia	178763	359529	388204	572330	2219
Cariotipo	Inafectado	Inafectado	Inafectado	Inafectado	Inafectado
Fracción fetal					
SNP		DENSIDAD DE ETIQUETA SNP		FRACCIÓN 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=110 alelo=A		94503		1,74	
rs2567608.2 Chr.20 longitud=110 alelo=G		1649			
rs7041158.1 Chr.9 longitud=117 alelo=C		107		5,61	
rs7041158.2 Chr.9 longitud= 17 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.E.) = 3,8+1,6					

Ejemplo 10

Determinación simultánea de aneuploidía y fracción fetal: enriquecimiento de ácidos nucleicos maternos y fetales en una muestra de plasma

[0260] Para enriquecer el ADNcf fetal y materno contenido en una muestra de plasma original derivada de una mujer embarazada, una porción de la muestra de plasma original se usó para amplificar secuencias de ácido nucleico diana polimórficas, comprendiendo cada una de ellas un SNP elegido del panel de SNP que figura en la Tabla 14, y una porción del producto amplificado se combinó con la muestra de plasma original restante.

[0261] ADNcf contenido en 15 µl de plasma libre de células se amplificó en un volumen de reacción de 50 µl que contenía 9 µl de una mezcla 1 µm de cebadores (15 plexTabla 5), 1 µl de polimerasa ADN de Phusion, 25 µl del tampón de PCR en sangre 2X Phusion que contiene trifosfatos de desoxinucleótidos (dNTPs: dATP, dCTP, dGTP y dTTP). El ciclo térmico se realizó con el Gene Amp9700 (Applied Biosystems) utilizando las siguientes condiciones de ciclo: incubación a 95°C durante 3 minutos, seguido de 35 ciclos a 95°C durante 20 segundos, 55°C durante 30 segundos y 70°C durante 1 minuto, seguido de una incubación final a 68°C durante 5 minutos. Se agregó una retención final a 4°C hasta que las muestras se retiraron para combinarlas con la porción no amplificada del plasma libre de células. El producto amplificado se diluyó 1:2 con agua y se analizó utilizando el Bioanalizador. Se diluyeron 3 µl adicionales de producto amplificado con 11,85 µl de agua para obtener una concentración final de 2 ng/µl. Se combinaron 2,2 µl del producto amplificado diluido con la muestra de plasma restante. El ADNcf 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.

[0262] Los resultados se dan en la Tabla 21. En el ejemplo dado, los datos muestran que la fracción de ADN fetal en muestra de plasma SAC2517 fue cuantificable a partir de los resultados de la secuenciación de un SNP informativo y se determinó que era 9,5%. En el ejemplo dado, se demostró que la muestra SAC2517 por cariotipo no se ve afectada por las aneuploidías de los cromosomas 21, 13, 18, X e Y. Se proporcionan densidades de etiquetas de

secuencia para los cromosomas 21, 13, 18, X e Y. La muestra SAC2517 fue la única muestra que se sometió al protocolo de enriquecimiento de ADNcf para secuencias polimórficas amplificadas. Por lo tanto, no se pudieron determinar los coeficientes de variación y las pruebas de diferenciabilidad. El ejemplo demuestra que enriquecer la mezcla de ADNcf fetal y materno presente en una muestra de plasma para las secuencias de ácido nucleico que comprenden al menos un SNP informativo puede usarse para proporcionar la secuencia necesaria y los recuentos de etiquetas de SNP para determinar la aneuploidía y la fracción fetal de un solo proceso de secuenciación.

TABLA 21

Determinación simultánea de aneuploidía y fracción fetal: enriquecimiento de ácidos nucleicos maternos y fetales en una muestra de plasma					
Aneuploidía					
	Cromosoma 21	Cromosoma 18	Cromosoma 13	Cromosoma X	Cromosoma Y
Densidad de la etiqueta de secuencia	183851	400582	470526	714055	2449
Cariotipo	Inafectado	Inafectado	Inafectado	Inafectado	Inafectado
Fracción Fetal					
SNP	RECuentos de Etiqueta		Fracción Fetal (%)		
rs10773760.1 Chr.12 longitud=128 alelo=A	8536		9,49		
rs10773760.2 Chr.12 longitud =128 alelo=G	89924				

Ejemplo 11**Determinación simultánea de aneuploidía y fracción fetal en muestras maternas enriquecidas para secuencias polimórficas que comprenden STR**

[0263] Para determinar simultáneamente la presencia o ausencia de una aneuploidía y la fracción fetal en una mezcla de ADNcf fetal y materno obtenida de una muestra materna, la mezcla se enriquece para secuencias polimórficas que comprenden STR, se secuencia y se analizan los datos. El enriquecimiento puede ser de una biblioteca de secuenciación como se describe en el Ejemplo 8, de una muestra de ADNcf purificada como se describe en el Ejemplo 9, o de una muestra de plasma como se describe en el Ejemplo 10. En cada caso, la información de secuenciación se obtiene de la secuenciación de una sola biblioteca, que permite determinar simultáneamente la presencia o ausencia de una aneuploidía y la fracción fetal. Preferiblemente, la biblioteca de secuenciación se prepara utilizando el protocolo abreviado proporcionado en el Ejemplo 2.

[0264] Los STR que se amplifican se seleccionan de los STR codis y no codis descritos en la Tabla 22, y la amplificación de las secuencias polimórficas de STR se obtiene utilizando los conjuntos correspondientes de cebadores proporcionados. Algunos de los STR se han descrito y/o analizado previamente para determinar la fracción fetal, se enumeran en la Tabla 22 y se describen en las solicitudes provisionales de EE.UU. 61/296.358 y 61/360.837.

TABLA 22

miniSTR de CODIS y NON-CODIS				
STR Locus (Nombre de marcador)	Ubicación de cromosoma	Rango de tamaño (bp)	Accesión de GenBank	Secuencias de cebador (delantera/inversa)
Codis miniSTR loci*				
CSF1PO	5q33.1	89-129	X14720	ACAGTAACTGCCTTCATAGATAG (CSF1PO_F; SEQ ID NO:113) GTGTCAGACCCTGTTCTAAGTA (CSF1PO_R; SEQ ID NO:114)
FGA	4q31.3	125-281	M64982	AAATAAAATTAGGCATATTTACAAGC (FGA_F; SEQ ID NO:115) GCTGAGTGATTTGTCTGTAATTG(FGA_R; SEQ ID NO:116)
TH01	11p15.5	51-98	D00269	CCTGTTCTCCCTTATTTCCC(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) GTCCTTGTGAGCGTTTATTTGC(TPOX_R; SEQ ID NO:120)

miniSTR de CODIS y NON-CODIS				
STR Locus (Nombre de marcador)	Ubicación de cromosoma	Rango de tamaño (bp)	Accesión de GenBank	Secuencias de cebador (delantera/inversa)
Codis miniSTR loci*				
vWA	12p13.31	88-148	M25858	AATAATCAGTATGTGACTTGGATTGA(vWA_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	GGGTGATTTTCCTCTTTGGT(D5S818_F; SEQ ID NO:125) AACATTTGTATCTTTATCTGTATCCTTATTTAT(D5S818_R; SEQ ID NO:126)
D7S820	7q21.11	136-176	AC004848	GAACACTTGTCATAGTTTAGAACGAAC(D7S820_F; SEQ ID NO:127) TCATTGACAGAATTGCACCA(D7S820_R; SEQ ID NO:128)
D8S1179	8q24.13	86-134	AF216671	TTTGTATTTTCATGTGTACATTCGTATC(D7S820_F; SEQ ID NO:129) ACCTATCCTGTAGATTATTTCACTGTG(D7S820_R; SEQ ID NO:130)
D13S317	13q31.1	88-132	AL353628	TCTGACCCATCTAACGCCTA(D13S317_F; SEQ ID NO:131) CAGACAGAAAGATAGATAGATGATTGA(D13S317_R; SEQ ID NO:132)
D16S539	16q24.1	81-121	AC024591	ATACAGACAGACAGACAGGTG(D16S539_F; SEQ ID NO:133) GCATGTATCTATCATCCATCTCT(D16S539_R; SEQ ID NO:134)
D18S51	18q21.33	113-193	AP001534	TGAGTGACAAATTGAGACCTT(D18S51_F; SEQ ID NO:135) GTCTTACAATAACAGTTGCTACTATT(D18S51_R; SEQ ID NO:136)
D21S11	21q21.1	153-221	AP000433	ATTCCCCAAGTGAATTGC(D21S11_F; SEQ ID NO:137) GGTAGATAGACTGGATAGATAGACGA(D21S11_R; SEQ ID NO:138)
D2S1338	2q35	90-142	AC01036	TGGAACAGAAATGGCTTGG(D2S1338_F; SEQ ID NO:139) GATTGCAGGAGGGAAGGAAG(D2S1338_R; SEQ ID NO:140)
Penta D	21q22.3	94-167	AP001752	GAGCAAGACACCATCTCAAGAA(Penta D_F; SEQ ID NO:141) GAAATTTTACATTTATGTTTATGATTCTCT(Penta D_R; SEQ ID NO:142)
Penta E	15q26.2	80-175	AC027004	GGCGACTGAGCAAGACTC(Penta E_F; SEQ ID NO:143) GGTTATTAATTGAGAAAACCTCCTTACA(Penta E_R; SEQ ID NO:144)
miniSTR loci no codis*				
D22S1045	22q12.3	82-115	AL022314 (17) F;	ATTTTCCCGATGATAGTAGTCT(D22S1045_F; SEQ ID NO:145) GCGAATGTATGATTGGCAATATTTT(D22S1045_R; SEQ ID NO:146)
D20S1082	20q13.2	73-101	AL158015	ACATGTATCCAGAACTTAAAGTAAAC(D20S1082_F; SEQ ID NO:147) GCAGAAGGGAAAATTGAAGCTG(D20S1082_R; SEQ ID NO:148)

miniSTR de CODIS y NON-CODIS				
STR Locus (Nombre de marcador)	Ubicación de cromosoma	Rango de tamaño (bp)	Accesión de GenBank	Secuencias de cebador (delantera/inversa)
Codis miniSTR loci*				
D20S482	20p13	85-126	AL121781 (14) F;	CAGAGACACCGAACCAATAAGA(D20S482_ SEQ ID NO:149) GCCACATGAATCAATTCCTATAATAAA (D20S482_R; SEQ ID NO:150)
D18S853	18p11.31	82-104	AP005130 (11)	GCACATGTACCCTAAAACTTAAAT(D18S853_F; SEQ ID NO:151) GTCAACCAAACTCAACAAGTAGTAA(D18S853_R; SEQ ID NO:152)
D17S1301	17q25.1	114-139	AC016888 (12)	AAGATGAAATTGCCATGTAAAAATA(D17S1301_F; SEQ ID NO:153) GTGTGTATAACAAAATTCCTATGATGG(D17S1301_R; SEQ ID NO:154)
D17S974	17p13.1	114-139	AC034303 (10)	GCACCCAAAAGTGAATGTCATA(D17S974_F; SEQ ID NO:155) GGTGAGAGTGAGACCCTGTC(D17S974_R; SEQ ID NO:156)
D14S1434	14q32.13	70-98	AL121612 (13)	TGTAATAACTCTACGACTGTCTGTCTG(D14S1434_F; SEQ ID NO:157) GAATAGGAGGTGGATGGATGG(D14S1434_R; SEQ ID NO:158)
D12ATA63	12q23.3	76-106	AC009771 (13)	GAGCGAGACCCTGTCTCAAG(D12ATA63F; SEQ ID NO:159) GGAAAAGACATAGGATAGCAATTT(D12ATA63_R; SEQ ID NO:160)
D11S4463	11q25	88 - 116	AP002806 (14)	TCTGGATTGATCTGTCTGTCC(D11S4463 F; SEQ ID NO:161) GAATTAATACCATCTGAGCACTGAA(D11S4463_R; SEQ ID NO:162)
D10S1435	10p15.3	82-139	AL354747 (11)	TGTTATAATGCATTGAGTTTTATTCTG(D10S1435_F; SEQ ID NO:163) GCCTGTCTCAAAAATAAAGAGATAGACA(D10S1435_R; SEQ ID NO:164)
D10S1248	10q26.3	79 - 123	AL391869 (13)	TTAATGAATTGAACAAATGAGTGAG(D10S1248_F; SEQ ID NO:165) GCAACTCTGGTTGTATTGTCTTCAT(D10S1248_R; SEQ ID NO:166)
D9S2157	9q34.2	71 - 107	AL162417 (10)	CAAAGCGAGACTCTGTCTCAA(D9S2157_F; SEQ ID NO:167) GAAAATGCTATCCTCTTTGGTATAAAT(D9S2157_R; SEQ ID NO:168)
D9S1122	9q21.2	93 - 125	AL161789 (12)	GGGTATTTCAAGATAACTGTAGATAGG(D9S1122_F; SEQ ID NO:168) GCTTCTGAAAGCTTCTAGTTTACC(D9S1122_R; SEQ ID NO:170)
D8S1115	8p11.21	63 - 96	AC090739 (9)	TCCACATCCTCACCAACAC(D8S1115_F; SEQ ID NO:171) GCCTAGGAAGGCTACTGTCAA(D8S1115_R; SEQ ID NO:172)
D6S1017	6p21.1	81 - 110	AL035588 (10)	CCACCCGTCCATTTAGGC(D6S1017_F; SEQ ID NO:173) GTGAAAAAGTAGATATAATGGTTGGTG(D6S1017_R; SEQ ID NO:174)
D6S474	6q21	107-136	AL357514 (17)	GGTTTTCCAAGAGATAGACCAATTA(D6S474_F; SEQ ID NO:175) GTCCTCTCATAAATCCCTACTCATATC(D6S474_R; SEQ ID NO:176)

miniSTR de CODIS y NON-CODIS				
STR Locus (Nombre de marcador)	Ubicación de cromosoma	Rango de tamaño (bp)	Accesión de GenBank	Secuencias de cebador (directo/inversa)
Codis miniSTR loci*				
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(D4 S2408_F; SEQ ID NO:179) GTGAAATGACTGAAAAATAGTAACCA(D4S 2408_R; SEQ ID NO:180)
D4S2364	4q22.3	67-83	AC022317 (9)	CTAGGAGATCATGTGGGTATGATT(D4S2364 U_F; SEQ ID NO:181) GCAGTGAATAAATGAACGAATGGA(D4S236 4_R; SEQ ID NO:182)
D3S4529	3p12.1	111- 139	AC117452 (13)	CCCAAATTACTTGAGCCAAT(D3 S452_F; SEQ ID NO:183) GAGACAAAATGAAGAAACAGACAG(D3S45 2_R; SEQ ID NO:184)
D3S3053	3q26.31	84-108	AC069259 (9)	TCTTTGCTCTCATGAATAGATCAGT(D3S305 3_F; SEQ ID NO:185) GTTTGTGATAATGAACCCACTCAG(D3S3053 _R; SEQ ID NO:186)
D2S1776	2q24.3	127 - 161	AC009475 (11)	TGAACACAGATGTTAAGTGTGTATATG(D2S 1776_F; SEQ ID NO:187) GTCTGAGGTGGACAGTTATGAAA(D2S1776_ R; SEQ ID NO:188)
D2S441	2p14	78 - 110	AC079112 (12)	CTGTGGCTCATCTATGAAAACCTT(D2S441_F; SEQ ID NO:189) GAAGTGGCTGTGGTGTATGAT(D2S441_R; SEQ ID NO:190)
D1S1677	1q23.3	81 - 117	AL513307 (15)	TTCTGTTGGTATAGAGCAGTGTTT(D1S1677 F; SEQ ID NO:191)
				GTGACAGGAAGGACGGAATG(D1S1677_R; SEQ ID NO:192)
D1S1627	1p21.1	81-100	AC093119 (13)	CATGAGGTTTGCAAATACTATCTTAAC(D1S 1627_F; SEQ ID NO:193) GTTTTAATTTCTCCAAATCTCCA(D1S1627_ R; SEQ ID NO:194)
D1GATA113	1p36.23	81 - 105	Z97987 (11)	TCTTAGCCTAGATAGATACTTGCTTCC(D1G ATA113_F; SEQ ID NO:195) GTCAACCTTTGAGGCTATAGGAA(D1GATA 113_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)				

[0265] Los miniSTR proporcionados en la Tabla 22 se han utilizado con éxito para determinar la fracción fetal en muestras de ADNcf de plasma obtenidas de mujeres embarazadas con fetos masculinos o femeninos, utilizando electroforesis capilar (consulte la Tabla 24 en el Ejemplo 15) para identificar y cuantificar los alelos de fetalidad y la maternidad. Por lo tanto, se espera que las secuencias polimórficas que comprenden otros STR, por ejemplo, los STR restantes de la Tabla 22 se puedan usar para determinar la fracción fetal mediante métodos de secuenciación masivamente paralelos.

[0266] La secuenciación de la biblioteca enriquecida para secuencias polimórficas de STR se realiza usando una tecnología NGS, por ejemplo, secuenciación masivamente paralela por síntesis. Las lecturas de secuencia de longitudes de al menos 100 pb se alinean con un genoma de referencia, por ejemplo, la secuencia del genoma humano de referencia NCBI36/hg18, y con un genoma de STR, y el número de etiquetas de secuencia asignadas al genoma humano de referencia y el genoma de referencia de STR obtenido para alelos informativos se utiliza para determinar la presencia o ausencia de aneuploidía y la fracción fetal, respectivamente. El genoma de referencia de

STR incluye las secuencias de amplicones amplificados a partir de los cebadores dados.

Ejemplo 12

5 **Determinación simultánea de aneuploidía y fracción fetal mediante secuenciación masiva en paralelo de muestras maternas enriquecidas para secuencias polimórficas que comprenden SNP en tándem**

[0267] Para determinar simultáneamente la aneuploidía y la fracción fetal en muestras maternas que comprenden ácidos nucleicos maternos y fetales, las muestras de plasma, las muestras de ADNcf purificadas y las muestras de la biblioteca de secuenciación se enriquecen para las secuencias de ácido nucleico diana polimórficas, comprendiendo cada una de ellas un par de SNP en tándem seleccionados entre rs7277033-rs2110153; rs2822654-rs1882882; rs368657-rs376635; rs2822731-rs2822732; rs1475881-rs7275487; rs1735976-rs2827016; rs447340-rs2824097; rs418989-rs13047336; rs987980-rs987981; rs4143392-rs4143391; rs1691324-rs1305434; rs11909758-rs9980111; rs2826842-rs232414; rs1980869-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 utilizados para amplificar las secuencias de destino que comprenden los SNP 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 par de SNP en tándem, es decir, cada uno de los sitios de SNP en el par en tándem se engloba dentro de los 36 pb generados por el método de secuenciación. La secuenciación de extremos emparejados se puede usar para identificar todas las secuencias que abarcan los sitios SNP en tándem. Los conjuntos ejemplares de cebadores que se utilizan para amplificar los SNP en tándem descritos en este documento son: rs7277033-rs2110153_F: TCCTGGAAACAAAGTATT (SEQ ID NO: 197) y rs7277033-rs2110153_R: AACCTTACAACAAAGCTAGAA (SEQ ID NO:198), conjunto rs2822654-rs1882882_F: ACTAAGCCTTGGGGATCCAG (SEQ ID NO: 199) y rs2822654-rs1882882_R: TGCTGTGGAAATACTAAAGG (SEQ ID NO:200), conjunto 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:TGCTTGA-GAAAGCTAACACTT (SEQ ID NO:206), rs1735976-rs2827016_F:CAGTGTGGGAAATTGTCTG (SEQ ID NO:207) y rs1735976-rs2827016_R: GGCAGTGGGAGATTATTGTA (SEQ ID NO: 208), rs447349-rs2824097_F: TCCTGTTGTAAAGTACACAT (SEQ ID NO: 209) y rs447349-rs2824097_R: GGGCCGTAATTACTTTTG (SEQ ID NO: 210), rs418989-rs13047336_F: ACTCAGTAGG CACTTTGTGTC (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: TGAAACTCAAAGAGAAAAG (SEQ ID NO:219) y rs11909758-rs9980111_R: ACAGATTTCTACTTAAATTT (SEQ ID NO:220), rs2826842-rs232414_F: TGAAACTCAAAGAGAAAAG (SEQ ID NO:221) y rs2826842-rs232414_R: ACAGATTTCTACTTAAATTT (SEQ ID NO:22), rs2826842-rs232414_F: GCAAAGGGGTACTCTATGTA (SEQ ID NO:223) y rs2826842-rs232414_R: TATCGGGTCATCTTGTTAAA (SEQ ID NO:224), rs1980969-rs1980970_F: TCTAACAAAGCTCTGTCCAAA (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: ATTTCACTATTCTTCATTTT (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 ID NO:235) y rs455921-rs2898102_R: TTTTGTCTTGTCTGCTGCTGA (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 ID NO:241) y rs2174536-rs458076_R: TTTCTAAATTTCCATTGAACAG (SEQ ID NO:242), rs11088023-rs11088024_F: GAAATTGGCAATCTGATTCT (SEQ ID NO:243) y rs11088023-rs11088024_R: CAACTTGTCTTTATTGATGT (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: TATGGCCCCAAACTATTCT (SEQ ID NO:250), rs8134080-rs2831524_F: ACAAGTACTGGGCAGATTGA (SEQ ID NO:251) y rs8134080-rs2831524_R: GCCAGGTTTAGCTTCAAGT (SEQ ID NO:252), rs4817219-rs4817220_F: TTTTATATCAGGAGAAACACTG (SEQ ID NO:253) y rs4817219-rs4817220_R: CCAGAATTTTGGAGGTTTAAT (SEQ ID NO:254), rs2250911-rs2250997_F: TGCTATTCCTCCTTTATCTCCA (SEQ ID NO:255) y rs2250911-rs2250997_R: TTCTTTTGCCTCTCCCAAAG (SEQ ID NO:256), rs2831899-rs2831900_F: ACCCTGGCACAGTGTTGACT (SEQ ID NO:257) y rs2831899-rs2831900_R: TGGGCCTGAGTTGAGAAGAT (SEQ ID NO:258), rs2831902-rs2831903_F: AATTTGTAAGTATGTGCAACG (SEQ ID NO:259) y rs2831902-rs2831903_R: TTTTCCCATTTTCCAACCTCT (SEQ ID NO:260), rs11088086-rs2251447_F: AAAAGATGAGACAGGCAGGT (SEQ ID NO:261) y rs11088086-rs2251447_R: ACCCTGTGAATCTCAAAAT (SEQ ID NO:262), rs2832040-rs11088088_F: GCACTTGCTTCTATTGTTGT (SEQ ID NO:263) y rs2832040-rs11088088_R: CCCTTCCTCTCTTCCATTCT (SEQ ID NO:264), rs2832141-rs2246777_F: AGCACTGCAGTA (SEQ ID NO:265) y rs2832141-rs2246777_R: ACAGATACCAAAGAACTGCAA (SEQ ID NO:266), rs2832959-rs9980934_F: TGGACACSTTTCAACTTAGA (SEQ ID NO:267) y rs2832959-rs9980934_R: GAACAGTAATGTTGAACTTTT (SEQ ID NO:268), rs2833734-rs2833735_F: TCTTGCAAAAAGCTTAGCACA (SEQ ID NO:269) y rs2833734-rs2833735_R: AAAAAGATCTCAAAGGGTCCA (SEQ ID NO:270), rs933121-rs933122_F: GCTTTTGCTGAACATCAAGT (SEQ ID NO:271) y rs933121-rs933122_R: CTTCCAGCAGCATAGTCT (SEQ ID NO:272), rs2834140-rs12626953_F: AAATCCAGGATGTGCAGT (SEQ ID NO:273) y rs2834140-rs12626953_R: ATGATGAGGTGAGTGGTGT (SEQ ID NO:274), rs2834485-rs3453_F: CATCACAGATCATAGTAAATGG (SEQ ID NO:275) y rs2834485-rs3453_R: AATTATTATTTTGAGGCAAT (SEQ ID NO:276), rs9974986-rs2834703_F: CATGAGGCAAACACTTTCC (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-rs1984015_F: TAGGAGAACAGAAGATCAGAG (SEQ ID NO:281) y rs1984014-rs1984015_R: AAAGACTATTGCTAAATGCTTG (SEQ ID NO:282), rs7281674-rs2835316_F: TAAGCGTAGGGCTGTGTGTG (SEQ ID NO:283) y rs7281674-rs2835316_R: GGACGGATAGACTCCAGAAGG (SEQ ID NO:284), rs13047304-rs13047322_F: GAATGACCTGGCACTTTTATCA (SEQ ID NO:285) y rs13047304-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: TTAGCTTGCATGTACCTGTGT (SEQ ID NO:292), rs2836660-rs2836661_F: AGCTAGATGGGGTGAAATTTT (SEQ ID NO:293) y rs2836660-rs2836661_R: TGGGCTGAGGGGAGATTC (SEQ ID NO:294), rs465612-rs8131220_F: ATCAAGCTAATTAATGTTATCT (SEQ ID NO:295) y rs465612-rs8131220_R: AATGAATAAGGTCCTCAGAG (SEQ ID NO:296), rs9980072-rs8130031_F: TTTAATCTGATCATTGCCCTA (SEQ ID NO:297) y rs9980072-rs8130031_R: AGCTGTGGGTGACCTTGA (SEQ ID NO:298), rs418359-rs2836926_F: TGTCACCATTGTGTATTA (SEQ ID NO:299) y rs418359-rs2836926_R: TCAGACTTGAAGTCCAGGAT (SEQ ID NO:300), rs7278447-rs7278858_F: GCTTCAGGGGTGTTAGTTTT (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: CCCAGCTGCAGTGTCTAC (SEQ ID NO:306), rs2837296-rs2837297_F: TCTTGTTCGAATCACAGGAG (SEQ ID NO:307) y rs2837296-rs2837297_R: 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 fijación de GC, para garantizar la hibridación específica de cebadores ricos en AT (Ghanta *et al.*, PLOS ONE 5 (10): doi10.1371/journal.pone.0013184 [2010], disponible en Internet en plosone.org). Un ejemplo de una secuencia de fijación de GC que puede incluir 5' del cebador directo o 3' del cebador inverso es GCCGCCTGCAGCCCGCGCCCCCGTGCCCCCGCCCCGCGCCGCCCCGGCGCC (SEQ ID NO: 311).

[0268] Preparación de la muestra y el enriquecimiento de la biblioteca de secuenciación ADNcf, una muestra ADNcf purificada, y una muestra de plasma se lleva a cabo de acuerdo con el método descrito en los Ejemplos 8, 9 y 10, 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 e incluye la secuenciación de extremos pareados. El análisis de los datos de secuenciación para la determinación de la aneuploidía fetal se realiza como se describe en los Ejemplos 4 y 5. Concomitante al análisis para determinar la aneuploidía, los datos de secuenciación se analizan para determinar la fracción fetal de la siguiente manera. Tras la transferencia de la imagen y los archivos de llamada base al servidor Unix que ejecuta el software Illumina "Genome Analyzer Pipeline" versión 1.51 como se describe, las lecturas de 36 pb se alinean con un "genoma SNP en tándem" mediante el programa BOWTIE. El genoma de SNP en tándem se identifica como la agrupación de las secuencias de ADN que abarcan los alelos de los 58 pares de SNP en tándem descritos anteriormente. Solo las lecturas que se asignan de forma única al genoma de SNP en tándem se utilizan para el análisis de la fracción fetal. Las lecturas que coinciden perfectamente con el genoma SNP en tándem se cuentan como etiquetas y se filtran. De las lecturas restantes, solo las lecturas que tienen uno o dos desajustes se cuentan como etiquetas y se incluyen en el análisis. Las etiquetas asignadas a cada uno de los alelos SNP en tándem se cuentan, y la fracción fetal se determina esencialmente como se describe en el **Ejemplo 6** anterior, pero se tienen en cuenta las etiquetas asignadas a los dos alelos x e y SNP en tándem presentes en cada

una de las secuencias de ácido nucleico diana polimórficas amplificadas que se amplifican para enriquecer las muestras maternas, es decir. % fracción fetal $\text{alelo}_{x+y} = ((\sum \text{etiquetas de secuencias fetales para } \text{alelo}_{x+y}) / (\sum \text{etiquetas de secuencias para } \text{alelo}_{x+y})) \times 100$ Opcionalmente, la fracción de ácidos nucleicos fetales en la mezcla de ácidos nucleicos fetales y maternos se calculan para cada uno de los alelos informativos (alelo_{x+y}) de la siguiente manera:

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

para compensar la presencia de 2 conjuntos de alelos fetales en tándem, uno de ellos enmascarado por el fondo materno. Las secuencias SNP en tándem son informativas cuando la madre es heterocigótica y está presente un tercer haplotipo paterno, lo que permite una comparación cuantitativa entre el haplotipo heredado maternalmente y el haplotipo heredado paternalmente para calcular la fracción fetal mediante el cálculo de una relación de haplotipo (HR). El porcentaje de fracción fetal se calcula para al menos 1, al menos 2, al menos 3, al menos 4, al menos 5, al menos 6, al menos 7, al menos 8, al menos 9, al menos 10, al menos 11, al menos 12, al menos 13, al menos 14, al menos 15, al menos 16, al menos 17, al menos 18, al menos 19, al menos 20, al menos 25, al menos 30, al menos 40 o más conjuntos informativos de alelos en tándem. En una realización, la fracción fetal es la fracción fetal promedio determinada para al menos 3 conjuntos informativos de alelos en tándem.

Ejemplo 13

Determinación de la fracción fetal por secuenciación masivamente paralela de una biblioteca de destino que comprende ácidos nucleicos polimórficos que comprenden SNP

[0269] Para determinar la fracción de ADNcf fetal en una muestra materna, las secuencias de ácido nucleico polimórficas diana que comprenden cada una SNP se amplificaron y se usaron para preparar una biblioteca diana para la secuenciación masivamente paralela.

[0270] ADNcf se extrajo como se describe en el Ejemplo 1. Se preparó una biblioteca de secuenciación de la siguiente manera. El ADNcf contenido en 5 µl de ADNcf purificado se amplificó en un volumen de reacción de 50 µl que contenía 7,5 µl de una mezcla de cebador 1 µm (Tabla 10), 10 µl de NEB 5X Mastermix y 27 µl de agua. El ciclo térmico se realizó con el Gene Amp9700 (Applied Biosystems) utilizando las siguientes condiciones de ciclo: incubación a 95°C durante 1 minuto, seguido de 20-30 ciclos a 95°C durante 20 segundos, 68°C durante 1 minuto y 68°C durante 30s, seguido de una incubación final a 68°C durante 5 minutos. Se agregó una retención final a 4°C hasta que las muestras se retiraron para combinar con la porción no amplificada de la muestra de ADNc purificada. El producto amplificado se purificó utilizando el sistema de purificación por PCR Agencourt AMPure XP (Pieza N° A63881; Beckman coulter Genomics, Danvers, MA), y la concentración se cuantificó utilizando el Nanodrop 2.000 (Thermo Scientific, Wilmington, DE). Se agregó una retención final a 4°C hasta que las muestras se retiraron para preparar la biblioteca diana. El producto amplificado se analizó con un bioanalizador 2100 (Agilent Technologies, Sunnyvale, CA) y se determinó la concentración del producto amplificado. Se preparó una biblioteca de secuenciación de ácidos nucleicos diana amplificados utilizando el protocolo abreviado descrito en el Ejemplo 2, y se secuenció de manera masivamente paralela utilizando secuenciación por síntesis con terminadores de colorantes reversibles y de acuerdo con el protocolo de Illumina. El análisis y el recuento de las etiquetas asignadas a un genoma de referencia que consta de 26 secuencias (13 pares cada uno que representa dos alelos) que comprenden un SNP, es decir, la SEQ ID NO: 1-56 se realizó como se describe.

[0271] La Tabla 23 proporciona los recuentos de etiquetas obtenidos a partir de la secuenciación de la biblioteca diana y la fracción fetal calculada derivada de los datos de secuenciación.

TABLA 23

Determinación de la fracción fetal mediante secuenciación masivamente paralela de una biblioteca de ácidos nucleicos polimórficos que comprenden SNP		
SNP	RECuentos de ETIQUETA DE SNP	Fracción fetal (%)
rs10773760.1 Chr.12 longitud=128 alelo=A	236590	1,98
rs10773760.2 Chr.12 longitud=128 alelo=G	4680	
rs13182883.1 Chr.5 longitud=111 alelo=A	3607	4,99
rs13182883.2 Chr.5 longitud=111 alelo=G	72347	
rs4530059.1 Chr.14 longitud=110 alelo=A	3698	1,54
rs4530059.1 Chr.14 longitud=110 alelo=G	239801	
rs8078417.1 Chr.17 longitud=110 alelo=C	1E+06	3,66
rs8078417.2 Chr.17 longitud=110 alelo=T	50565	
Fracción Fetal (Media+D.E.) = 12.4+6.6		

[0272] Los resultados muestran que las secuencias de ácidos nucleicos polimórficos que comprenden cada uno al menos un SNP pueden amplificarse a partir de ADNcf derivado de una muestra de plasma materno para construir una biblioteca que puede secuenciarse de forma masiva paralela para determinar la fracción de ácidos nucleicos fetales en la muestra materna. Los métodos de secuenciación masivamente paralelos para determinar la fracción fetal se pueden usar en combinación con otros métodos para proporcionar el diagnóstico de la aneuploidía fetal y otras pruebas prenatales.

Ejemplo 14

Determinación de la fracción fetal por secuenciación masivamente paralela de una biblioteca de destino que comprende ácidos nucleicos polimórficos que comprenden STR o SNP en tándem

[0273] La fracción fetal puede determinarse independientemente de la determinación de aneuploidía utilizando una biblioteca de destino que comprende SNP o STR en tándem como se describe para la biblioteca de destino SNP del Ejemplo 13. Para preparar una biblioteca de destino SNP en tándem, una porción de biblioteca de ADNcf purificada que comprende los ácidos nucleicos maternos y fetales se utiliza para amplificar las secuencias diana utilizando una mezcla de cebadores, por ejemplo, las Tablas 10 y 11. Para preparar una biblioteca STR, se utiliza una porción de una biblioteca de ADNcp purificada que comprende los ácidos nucleicos maternos y fetales una mezcla de cebadores, por ejemplo, la Tabla 22. La biblioteca diana de SNP en tándem se secuencia como se describe en el Ejemplo 12.

[0274] Las bibliotecas diana se secuenciaron como se ha descrito, y la fracción fetal se determina a partir del número de etiquetas de secuencias asignadas al genoma de referencia STR o SNP en tándem que comprende respectivamente todos los posibles alelos STR o SNP en tándem abarcados por los cebadores. Los alelos informativos se identifican y la fracción fetal se determina utilizando el número de etiquetas mapeadas a los alelos de las secuencias polimórficas.

Ejemplo 15

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

[0275] Para determinar la fracción fetal en muestras maternas comprenden ADNcf fetal y materna, se recogieron muestras de sangre periférica de mujeres embarazadas con fetos de voluntarios ya sea hombre o mujer. Las muestras de sangre periférica se obtuvieron y procesaron para proporcionar un ADNf purificado como se describe en el Ejemplo 1

[0276] Se analizaron diez microlitros de las muestras de ADNc utilizando el kit de amplificación por PCR AmpF1STR® MiniFiler™ (Applied Biosystems, Foster City, CA) de acuerdo con las instrucciones del fabricante. Brevemente, el ADNcf contenido en 10 µl se amplificó en un volumen de reacción de 25 µl que contiene 5 µL de cebadores marcados con fluorescencia (Conjunto de cebador AmpF/STR® MiniFiler™), y la mezcla maestra AmpF/STR® MiniFiler™, que incluye AmpliTaq Gold® ADN polimerasa y tampón asociado, sal (MgCl₂ 1,5 mM), y 200 µM trifosfatos de desoxinucleótidos (dNTPs: dATP, dCTP, dGTP y dTTP). Los cebadores marcados con fluorescencia son cebadores directos que están marcados con tintes 6FAM™, VIC™, NED™ y PET™. El ciclo térmico se realizó con el Gene Amp9700 (Applied Biosystems) utilizando las siguientes condiciones de ciclo: incubación a 95°C durante 10 minutos, seguido de 30 ciclos a 94°C durante 20 segundos, 59°C durante 2 minutos y 72°C durante 1 minuto, seguido de una incubación final a 60°C durante 45 minutos. Se agregó una retención final a 4°C hasta que las muestras se retiraron para su análisis. El producto amplificado se preparó diluyendo 1 µl de producto amplificado en 8,7 µl de Hi-di™ formamida (Applied Biosystems) y 0,3 µl de GeneScan™-500 LIZ_estandar de tamaño interno (Applied Biosystems) y se analizó con un analizador genético ABI PRISM3130x1 (Applied Biosystems) utilizando recopilación de datos HID_G5_POP4 (Applied Biosystems) y una matriz de capilares de 36 cm. Todo el genotipado se realizó con el software GeneMapper_ID v3.2 (Applied Biosystems) utilizando escaleras y cubetas y paneles alélicos provistos por el fabricante.

[0277] Todas las mediciones de genotipificación se realizaron en el Analizador genético Applied Biosystems 3130 xl, utilizando una "ventana" de $\pm 0,5$ -nt en torno al tamaño obtenido para cada alelo para permitir la detección y la correcta asignación de alelos. Se determinó que cualquier alelo de muestra cuyo tamaño estaba fuera de la ventana $\pm 0,5$ -nt es OL, es decir, "Off Ladder". Los alelos OL son alelos de un tamaño que no está representado en la escalera alélica AmpF/STR® MiniFiler™ o un alelo que no corresponde a una escalera alélica, pero cuyo tamaño está justo fuera de una ventana debido a un error de medición. El umbral de altura de pico mínimo de >50 RFU se estableció en base a los experimentos de validación realizados para evitar la tipificación cuando es probable que los efectos estocásticos interfieran con la interpretación precisa de las mezclas. El cálculo de la fracción fetal se basa en promediar todos los marcadores informativos. Los marcadores informativos se identifican por la presencia de picos en el electroferograma que se encuentran dentro de los parámetros de los contenedores preestablecidos para los

STR que se analizan.

[0278] Los cálculos de la fracción fetal se realizaron utilizando la altura de pico promedio para alelos mayores y menores en cada locus STR determinado a partir de inyecciones por triplicado. Las reglas aplicadas al cálculo son:

1. los datos de alelos off-ladder (OL) para alelos no se incluyen en el cálculo; y
2. solo se incluyen en el cálculo las alturas de los picos derivadas de >50 RFU (unidades de fluorescencia relativa)
3. si solo está presente un contenedor, el marcador se considera no informativo; y
4. si se llama un segundo recipiente, pero los picos del primero y el segundo son dentro del 50-70% de sus unidades de fluorescencia relativa (RFU) en altura de pico, la fracción minoritaria no se mide y el marcador se considera no informativo.

[0279] 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 expresó como un porcentaje se calculó primero para cada locus informativo como

$$\text{fracción fetal} = (\Sigma \text{altura de pico de alelo menor} / \Sigma \text{altura de pico de alelo mayor}) \times 100,$$

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

[0280] La tabla 8 proporciona los datos obtenidos al analizar el ADNcf de una paciente embarazada de un feto masculino.

TABLA 24

Fracción fetal determinada en ADNcf de una embarazada por análisis de STR								
STR	Alelo 1	Alelo 2	Alelo 3	Altura de Alelo 1	Altura de alelo 2	Altura de 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								

[0281] Los resultados muestran que ADNcf se puede utilizar para determinar la presencia o ausencia de ADN fetal como se indica por la detección de un componente menor en uno o más alelos STR, para determinar la fracción fetal por ciento, y para determinar el sexo del feto como se indica por la presencia o ausencia del alelo de Amelogenin.

Ejemplo 16

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

[0282] Los recuentos de coincidencias de secuencia con diferentes cromosomas se manipulan para generar una puntuación que variará con el número de copias cromosómicas que se puede interpretar para identificar la amplificación o eliminación cromosómica. Por ejemplo, tal puntuación podría generarse comparando la cantidad relativa de una secuencia de etiquetas en un cromosoma que experimenta cambios en el número de copias a un cromosoma que se sabe que es un euploide. Los ejemplos de puntuaciones que se pueden usar para identificar la amplificación o la eliminación incluyen, entre otros, los siguientes: recuentos del cromosoma de interés dividido por recuentos de otro cromosoma de la misma serie experimental, los recuentos del cromosoma de interés divididos por el total número de recuentos de la prueba experimental, comparación de recuentos de la muestra de interés versus una muestra de control separada. Sin pérdida de generalidad, se puede suponer que las puntuaciones aumentarán a medida que aumente el número de copias. El conocimiento de la fracción fetal se puede usar para establecer umbrales de "corte" para llamar estados de "aneuploidía", "normal" o "marginal" (incierto). Luego, se realizan cálculos para estimar el número mínimo de secuencias requeridas para lograr una sensibilidad adecuada (es decir, la probabilidad de identificar correctamente un estado de aneuploidía).

[0283] La Figura 19 es una gráfica de dos poblaciones diferentes de puntuaciones. El eje x es la puntuación y el eje y es la frecuencia. Las puntuaciones en muestras de cromosomas sin aneuploidía pueden tener una distribución que se muestra en la Figura 19A. La Figura 19B ilustra una distribución hipotética de una población de puntuaciones en muestras con un cromosoma amplificado. Sin pérdida de generalidad, los gráficos y las ecuaciones muestran el caso de una puntuación univariada donde la condición de aneuploidía representa una amplificación del número de copias. Los casos multivariados y/o las anomalías de reducción/eliminación son simples extensiones o reordenamientos de las descripciones dadas y pretenden caer dentro del alcance de esta técnica.

[0284] La cantidad de "solapamiento" entre las poblaciones puede determinar lo bien que se pueden discriminar los casos normales y aneuploidía. En general, el aumento de la fracción fetal, ff, aumenta el poder de discriminación al separar los dos centros de población (moviendo "C2", el "centro de puntuaciones de aneuploidía", y aumentando "d", lo que hace que las poblaciones se superpongan menos. Además, un aumento en el valor absoluto de la magnitud, m, (por ejemplo, tener cuatro copias del cromosoma en lugar de una trisomía) de la amplificación también aumentará la separación de los centros de población que conducen a un mayor poder (es decir, una mayor probabilidad de identificar correctamente los estados de aneuploidía).

[0285] Al aumentar el número de secuencias generadas, N, reduce las desviaciones estándar "sdevA" y/o "sdevB", la extensión de las dos poblaciones de puntuaciones, lo que también hace que las poblaciones se superpongan menos.

Configuración de umbrales y estimación del tamaño de la muestra

[0286] El siguiente procedimiento se puede usar para establecer "c", el valor crítico para llamar a los estados "aneuploidía", "normal" o "marginal" (incierto). Sin pérdida de generalidad, las pruebas estadísticas unilaterales se utilizan a continuación.

[0287] Primero, se decide una tasa de falsos positivos aceptable, FP (a veces también llamado "error de tipo I" o "especificidad"), que es la probabilidad de un falso positivo o una falsa llamada de aneuploidía. Por ejemplo, FP puede ser al menos, o aproximadamente 0,001, 0,002, 0,003, 0,004, 0,005, 0,006, 0,007, 0,009, 0,001, 0,02, 0,03, 0,04, 0,05, 0,06, 0,07, 0,08, 0,09, o 0,1.

[0288] Segundo, el valor de "c" se puede determinar resolviendo la ecuación: $FP = \int_C^{\infty} f_1(x) dx$.

$$FP = \int_C^{\infty} f_1(x) dx$$

(Ecuación 1)

[0289] Una vez que se ha determinado un valor crítico, c , las secuencias de números mínimos requeridas para lograr un cierto $TP =$ tasa positiva verdadera puede ser estimada. La verdadera tasa positiva puede ser, por ejemplo, aproximadamente 0,5, 0,6, 0,7, 0,8 o 0,9. En una realización, la tasa positiva verdadera puede ser 0,8. En otras palabras, N es el número mínimo de secuencias requeridas para identificar el $100 \cdot TP$ de aneuploidía del tiempo. $N =$ número mínimo tal que $TP = \int_c^\infty f_2(x, ff) dx > 0,8$. N se determina resolviendo

$$\min_N s.t. \{TB \geq \int_c^\infty f_2(x, N) dx\}$$

(Ecuación 2)

[0290] En las pruebas estadísticas clásicas f_1 y f_2 son a menudo F , distribuciones F no centrales (un caso especial de t y distribuciones t no centrales) aunque esa no es una condición necesaria para esta aplicación.

Configuración de "niveles" de umbrales para dar más control de errores

[0291] Los umbrales también se pueden establecer en etapas utilizando los métodos anteriores. Por ejemplo, se puede establecer un umbral para la llamada de alta confianza de "aneuploidía", digamos ca , usando FP 0,001 y un umbral "marginal", digamos cb , usando FP 0,05. En este caso si Puntuación, S :

$(S > ca)$	luego llamar "Trisomía"
$(cb > S \leq ca)$	luego llamar "Marginal"
$(S < cb)$	luego llamar "Normal"

Algunas generalizaciones triviales que caen dentro del alcance de esta técnica

[0292] Diferentes valores para los umbrales tales como TP , FP , etc. pueden ser utilizados. Los procedimientos se pueden ejecutar en cualquier orden. Por ejemplo, uno puede comenzar con N y resolver para c , etc. Las distribuciones pueden depender de ff para que $f_1(x, N, ff)$, $f_2(x, N, ff)$ y/u otras variables. Las ecuaciones integrales anteriores se pueden resolver por referencia a tablas o por métodos computacionales iterativos. Se puede estimar un parámetro no central y se puede leer la potencia de las tablas estadísticas estándar. El poder estadístico y los tamaños de muestra pueden derivarse del cálculo o estimación de los cuadrados medios esperados. Se pueden usar distribuciones teóricas de forma cerrada como f , t , t no central, normal, etc. o estimaciones (kernel u otro) para modelar las distribuciones f_1 , f_2 . El ajuste de umbral empírico y la selección de parámetros utilizando curvas características del operador del receptor (ROC) se pueden usar y agrupar con la fracción fetal. Se pueden usar varias estimaciones de la distribución (varianza, desviación media absoluta, rango intercuartil, etc.). Se pueden utilizar varias estimaciones del centro de distribución (media, mediana, etc.). Se pueden utilizar pruebas estadísticas a dos caras en lugar de a un solo lado. La prueba de hipótesis simple se puede reformular como regresión lineal o no lineal. Los métodos combinatorios, la simulación (p. ej., monte carlo), la maximización (p. ej., la expectativa de maximización), los métodos iterativos u otros pueden usarse de forma independiente o junto con los anteriores para establecer la potencia estadística o los umbrales.

Ejemplo 17

Demostración de detección de aneuploidía

[0293] Los datos de secuenciación obtenidos para las muestras descritas en los Ejemplos 4 y 5, y mostrados en las figuras 9-13 se analizaron adicionalmente para ilustrar la sensibilidad del método para identificar con éxito aneuploidías en muestras maternas. Las dosis de cromosomas normalizados para los cromosomas 21, 18, 13, X e Y se analizaron como una distribución relativa a la desviación estándar de la media (eje Y) y se muestran en la Figura 20. El cromosoma de normalización utilizado se muestra como el denominador (eje X).

[0294] La Figura 20 (A) muestra la distribución de las dosis de cromosomas en relación con la desviación estándar de la media de la dosis del cromosoma 21 en las muestras (o) no afectadas (normal) y las muestras de trisomía 21 (T_{21} ; Δ) cuando se usa el cromosoma 14 como el cromosoma normalizador para el cromosoma 21. La Figura 20 (B) muestra la distribución de las dosis de cromosomas en relación con la desviación estándar de la media para la dosis del cromosoma 18 en las muestras no afectadas (o) y las muestras de trisomía 18 (T_{18} ; Δ) cuando usando el cromosoma 8 como el cromosoma normalizador para el cromosoma 18. La figura 20 (C) muestra la distribución de las dosis de cromosomas en relación con la desviación estándar de la media para la dosis del cromosoma 13 en las muestras no afectadas (o) y las muestras de trisomía 13 (T_{13} ; Δ), utilizando la densidad media de la etiqueta de secuencia del grupo de los cromosomas 3, 4, 5 y 6 como el cromosoma normalizador para determinar la dosis de cromosoma para el cromosoma 13. La Figura 20 (D) muestra la distribución de las dosis de cromosoma

relacionadas con la desviación estándar de la media para la dosis del cromosoma X en las muestras femeninas no afectadas (o), las muestras masculinas no afectadas (Δ) y las muestras de monosomía X (XO; +) cuando se utiliza el cromosoma 4 como cromosoma normalizador para el cromosoma X. La figura 20 (E) muestra la distribución de las dosis de cromosomas en relación con la desviación estándar de la media de la dosis del cromosoma Y en las muestras masculinas no afectadas (o), las muestras femeninas no afectadas (Δ) y las muestras de monosomía X (+), cuando se utiliza la densidad de etiqueta de secuencia promedio del grupo de cromosomas 1-22 y X como el cromosoma normalizador para determinar la dosis de cromosoma para el cromosoma Y.

[0295] Los datos muestran que la trisomía 21, trisomía 18, trisomía 13 eran claramente distinguibles de las muestras no afectadas (normales). Las muestras de monosomía X fueron fácilmente identificables por tener una dosis de cromosoma X claramente inferior a la de las muestras femeninas no afectadas (Figura 20 (D)) y por tener dosis de cromosoma Y claramente menores que las muestras masculinas no afectadas (Figura 20). (E)).

[0296] Por lo tanto, el método aquí descrito es sensible y específico para determinar la presencia o ausencia de aneuploidías cromosómicas en una muestra de sangre materna.

LISTADO DE SECUENCIAS

[0297]

<110> VERINATA HEALTH, INC.

<120> NOVEL PROTOCOL FOR PREPARING SEQUENCING LIBRARIES

<130> P070843EP

<150> 61/455,849

<151> 2010-10-26

<150> 61/407,017

<151> 2010-10-26

<150> 61/360,837

<151> 2010-07-01

<150> 61/296,358

<151> 2010-01-19

<160> 427

<170> PatentIn versión 3.5

<210> 1

<211> 111

<212> ADN

<213> Homo sapiens

<400> 1

cacatgcaca gccagcaacc ctgtcagcag gagttccac cagtttcttt ctgagaacat 60

ctgttcaggt ttctctccat ctctatttac tcagggtcaca ggaccttggg g 111

<210> 2

<211> 111

<212> ADN

<213> Homo sapiens

<400> 2

cacatgcaca gccagcaacc ctgtcagcag gagttccac cagtttcttt ctgagaacat 60

ctgttcaggt ttctctccat ctctgtttac tcagggtcaca ggaccttggg g 111

ES 2 704 701 T3

5	<210> 3		
	<211> 126		
	<212> ADN		
	<213> Homo sapiens		
	<400> 3		
10	tgaggaagtg aggctcagag ggtaagaaac tttgtcacag agctggtggt gaggggtggag		60
	attttacact ccctgcctcc cacaccagtt tctccagagt ggaaagactt tcatctcgca		120
	ctggca		126
15	<210> 4		
	<211> 126		
	<212> ADN		
	<213> Homo sapiens		
	<400> 4		
25	tgaggaagtg aggctcagag ggtaagaaac tttgtcacag agctggtggt gaggggtggag		60
	attttacact ccctgcctcc cacaccagtt tctccgaggt ggaaagactt tcatctcgca		120
	ctggca		126
30	<210> 5		
	<211> 121		
	<212> ADN		
	<213> Homo sapiens		
	<400> 5		
40	gtgccttcag aacctttgag atctgattct atttttaag cttcttagaa gagagattgc		60
	aaagtgggtt gtttctctag ccagacaggg caggtaaata ggggtggctg gtgggatggg		120
	a		121
45	<210> 6		
	<211> 121		
	<212> ADN		
	<213> Homo sapiens		
	<400> 6		
55	gtgccttcag aacctttgag atctgattct atttttaag cttcttagaa gagagattgc		60
	aaagtgggtt gtttctctag ccagacaggg caggtaaata ggggtggctg gtgggatggg		120
	a		121
60	<210> 7		
	<211> 111		
	<212> ADN		
	<213> Homo sapiens		
	<400> 7		

ES 2 704 701 T3

<400> 7

5	aggtgtgtct ctcttttgtg aggggagggg tcccttctgg cctagtagag ggcctggcct	60
	gcagtgagca ttcaaatcct caaggaacag ggtggggagg tgggacaaag g	111
10	<210> 8 <211> 111 <212> ADN <213> Homo sapiens	
15	<400> 8	
20	aggtgtgtct ctcttttgtg aggggagggg tcccttctgg cctagtagag ggcctggcct	60
	gcagtgagca ttcaaatcct cgaggaacag ggtggggagg tgggacaaag g	111
25	<210> 9 <211> 139 <212> ADN <213> Homo sapiens	
30	<400> 9	
35	cctcgctac tgtgtgttt ctaaccatca tgcttttccc tgaatctctt gagtcttttt	60
	ctgctgtgga ctgaaacttg atcctgagat tcacctctag tccctctgag cagcctcctg	120
	gaatactcag ctgggatgg	139
40	<210> 10 <211> 139 <212> ADN <213> Homo sapiens	
45	<400> 10	
50	cctcgctac tgtgtgttt ctaaccatca tgcttttccc tgaatctctt gagtcttttt	60
	ctgctgtgga ctgaaacttg atcctgagat tcacctctag tccctctggg cagcctcctg	120
	gaatactcag ctgggatgg	139
55	<210> 11 <211> 117 <212> ADN <213> Homo sapiens	
60	<400> 11	
	aattgcaatg gtgagagggt gatggtaaaa tcaaacggaa cttgttattt tgtcattctg	60
	atggactgga actgaggatt ttcaatttcc tctccaaccc aagacacttc tcaactgg	117

ES 2 704 701 T3

5	<210> 12 <211> 117 <212> ADN <213> Homo sapiens		
10	<400> 12		
15	aattgcaatg gtgagagggt gatggtaaaa tcaaacggaa cttgttattt tgtcattctg atggactgga actgaggatt ttcaatttcc ttccaaccc aagacacttc tcaactgg	60 117	
20	<210> 13 <211> 114 <212> ADN <213> Homo sapiens		
25	<400> 13		
30	gaaatgcctt ctcaggtaat ggaaggttat ccaaataattt ttcgtaagta tttcaaatag caatggctcg tctatgggta gtctcacagc cacatttctca gaactgctca aacc	60 114	
35	<210> 14 <211> 114 <212> ADN <213> Homo sapiens		
40	<400> 14		
45	gaaatgcctt ctcaggtaat ggaaggttat ccaaataattt ttcgtaagta tttcaaatag caatggctcg tctatgggta gtctcgagc cacatttctca gaactgctca aacc	60 114	
50	<210> 15 <211> 128 <212> ADN <213> Homo sapiens		
55	<400> 15		
60	acccaaaaca ctggaggggc ctcttctcat ttctggtaga ctgcaagtgt tagccgtcgg gaccagcttc tgtctggaag ttctgcaaat tgcagttaag tccaagtatg ccacatagca gataaggg	60 120 128	
60	<210> 16 <211> 128 <212> ADN <213> Homo sapiens		
	<400> 16		

ES 2 704 701 T3

	acccaaaaca ctggaggggc ctcttctcat ttctggtaga ctgcaagtgt tagccgtcgg	60
5	gaccagcttc tgtctggaag ttctgtcaaat tgcagttagg tccaagtatg ccacatagca	120
	gataaggg	128
10	<210> 17 <211> 110 <212> ADN <213> Homo sapiens	
15	<400> 17	
	gcaccagaat ttaaacaacg ctgacaataa atatgcagtc gatgatgact tcccagagct	60
20	ccagaagcaa ctccagcaca cagagaggcg ctgatgtgcc tgtcagggtgc	110
25	<210> 18 <211> 110 <212> ADN <213> Homo sapiens	
30	<400> 18	
	gcaccagaat ttaaacaacg ctgacaataa atatgcagtc gatgatgact tcccagagct	60
35	ccagaagcaa ctccagcaca cggagaggcg ctgatgtgcc tgtcagggtgc	110
40	<210> 19 <211> 116 <212> ADN <213> Homo sapiens	
	<400> 19	
45		
	tgactgtata ccccagggtgc acccttgggt catctctatc atagaactta tctcacagag	60
50	tataagagct gattttctgtg tctgcctctc aactagact tccacatcct tagtgc	116
55	<210> 20 <211> 116 <212> ADN <213> Homo sapiens	
	<400> 20	
60		
	tgactgtata ccccagggtgc acccttgggt catctctatc atagaactta tctcacagag	60
	tataagagct gattttctgtg tctgcctctc aactagact tccacatcct tagtgc	116

ES 2 704 701 T3

<210> 21
 <211> 110
 <212> ADN
 <213> Homo sapiens
 5
 <400> 21

 10 tgtacgtggt caccagggga cgcctggcgc tgcgaggag gcccgagcc tcgtgcccc 60
 gtgaagcttc agctcccctc ccgggctgtc cttgaggctc ttctcacact 110

 15 <210> 22
 <211> 110
 <212> ADN
 <213> Homo sapiens
 20 <400> 22

 tgtacgtggt caccagggga cgcctggcgc tgcgaggag gcccgagcc tcgtgcccc 60
 25 gtgaagcttc agctcccctc cctggctgtc cttgaggctc ttctcacact 110

 30 <210> 23
 <211> 114
 <212> ADN
 <213> Homo sapiens
 <400> 23
 35
 cagtggaccg tgetgcacct ttctccctc cccatcaacc tcttttgtgc ctccccctcc 60
 gtgtaccacc ttctctgtca ccaaccctgg cctcacaact ctctcctttg ccac 114
 40
 45 <210> 24
 <211> 114
 <212> ADN
 <213> Homo sapiens
 <400> 24
 50
 cagtggaccg tgetgcacct ttctccctc cccatcaacc tcttttgtgc ctccccctcc 60
 gtgtaccacc ttctctgtca ccaccctgg cctcacaact ctctcctttg ccac 114
 55
 60 <210> 25
 <211> 110
 <212> ADN
 <213> Homo sapiens
 <400> 25

ES 2 704 701 T3

	cagtggcata gtagtccagg ggctcctcct cagcacctcc agcaccttcc aggaggcagc	60
5	agcgcaggca gagaacccgc tggaagaatc ggcggaagtt gtcggagagg	110
10	<210> 26 <211> 110 <212> ADN <213> Homo sapiens <400> 26	
15	cagtggcata gtagtccagg ggctcctcct cagcacctcc agcaccttcc aggaggcagc	60
	agcgcaggca gagaacccgc tggaaggatc ggcggaagtt gtcggagagg	110
20	<210> 27 <211> 129 <212> ADN <213> Homo sapiens	
25	<400> 27	
	aggtctgggg gccgctgaat gccaaagctgg gaatcttaaa tgttaaggaa caaggtcata	60
30	caatgaatgg tgtgatgtaa aagcttgga ggtgatttct gagggtaggt gctgggttta	120
	atgggagga	129
35	<210> 28 <211> 129 <212> ADN <213> Homo sapiens	
40	<400> 28	
	aggtctgggg gccgctgaat gccaaagctgg gaatcttaaa tgttaaggaa caaggtcata	60
45	caatgaatgg tgtgatgtaa aagcttgga ggtgattttt gagggtaggt gctgggttta	120
	atgggagga	129
50	<210> 29 <211> 107 <212> ADN <213> Homo sapiens	
55	<400> 29	
	acggttctgt cctgtagggg agaaaagtc tcgttggtcc tctgggatgc aacatgagag	60
60	agcagcacac tgaggcttta tggattgcc tgccacaagt gaacagg	107
	<210> 30 <211> 107	

ES 2 704 701 T3

	<212> ADN		
	<213> Homo sapiens		
5	<400> 30		
	acggttctgt cctgtagggg agaaaagtcc tcgttggtcc tctgggatgc aacatgagag	60	
10	agcagcacac tgaggcttta tgggttgccc tgccacaagt gaacagg	107	
	<210> 31		
	<211> 127		
15	<212> ADN		
	<213> Homo sapiens		
	<400> 31		
20	gcgcagtcag atgggcgtgc tggcgtctgt cttctctctc tcctgctctc tggcttcatt	60	
	tttctctcct tctgtctcac cttctttcgt gtgcctgtgc acacacacgt ttgggacaag	120	
	ggctgga	127	
25	<210> 32		
	<211> 127		
	<212> ADN		
30	<213> Homo sapiens		
	<400> 32		
35	gcgcagtcag atgggcgtgc tggcgtctgt cttctctctc tcctgctctc tggcttcatt	60	
	tttctctcct tctgtctcac cttctttcgt gtgcctgtgc atacacacgt ttgggacaag	120	
	ggctgga	127	
40	<210> 33		
	<211> 130		
	<212> ADN		
	<213> Homo sapiens		
45	<400> 33		
	gccggacctg cgaaatccca aaatgccaaa cattcccgcc tcacatgata ccagagagag	60	
50	gggacctcagt gttcccagct tgcagctgag gagcccagag ttgccgtcag atcagagccc	120	
	cagttgcccg	130	
55	<210> 34		
	<211> 130		
	<212> ADN		
	<213> Homo sapiens		
60	<400> 34		

ES 2 704 701 T3

	gccggacctg cgaaatccca aaatgccaaa cattcccgcc tcacatgac ccagagagag	60
	gggacccagt gttcccagct tgcagctgag gagcccgagt ttgccgtcag atcagagccc	120
5	cagttgcccg	130
10	<210> 35 <211> 121 <212> ADN <213> Homo sapiens	
15	<400> 35	
	agcagcctcc ctcgactagc tcacactacg ataaggaaaa ttcattgagct ggtgtccaag	60
	gagggctggg tgactcgtgg ctcatgcagc atcaagattc ctttcgtctt tcccctctgc	120
20	c	121
25	<210> 36 <211> 121 <212> ADN <213> Homo sapiens	
	<400> 36	
30	agcagcctcc ctcgactagc tcacactacg ataaggaaaa ttcattgagct ggtgtccaag	60
	gagggctggg tgactcgtgg ctcatgcagc gtcaagattc ctttcgtctt tcccctctgc	120
35	c	121
40	<210> 37 <211> 138 <212> ADN <213> Homo sapiens	
	<400> 37	
45	tggcattgcc tgtaatatat atagccatgg ttttttatag gcaatttaag atgaatagct	60
	tctaaactat agataagttt cattacccca ggaagctgaa ctatagctac tttacccaaa	120
50	atcattagaa tgggtgctt	138
55	<210> 38 <211> 138 <212> ADN <213> Homo sapiens	
	<400> 38	
60		

ES 2 704 701 T3

	tggcattgcc tgtaatat acatagccatgg ttttttatag gcaatttaag atgaatagct	60
	tctaaactat agataagttt cattacccca ggaagctgaa ctatagctac tttccccaaa	120
5	atcattagaa tggtgctt	138
10	<210> 39 <211> 136 <212> ADN <213> Homo sapiens	
15	<400> 39	
	atgaagcctt ccaccaactg cctgtatgac tcatctggg acttctgctc tataactcaa	60
	gtggcttagt cactgccaat gtatttccat atgagggacg atgattacta aggaaatata	120
20	gaaacaacaa ctgatc	136
25	<210> 40 <211> 136 <212> ADN <213> Homo sapiens	
	<400> 40	
30	atgaagcctt ccaccaactg cctgtatgac tcatctggg acttctgctc tataactcaa	60
	gtggcttagt cactgccaat gtatttccat atgagggacg gtgattacta aggaaatata	120
35	gaaacaacaa ctgatc	136
40	<210> 41 <211> 118 <212> ADN <213> Homo sapiens	
	<400> 41	
45	acaacagaat caggtgattg gagaaaagat cacaggccta ggcacccaag gcttgaagga	60
	tgaaagaatg aaagatggac ggaacaaaat taggacctta attctttgtt cagttcag	118
50		
55	<210> 42 <211> 118 <212> ADN <213> Homo sapiens	
	<400> 42	
60		

ES 2 704 701 T3

5	acaacagaat caggtgattg gagaaaagat cacaggccta ggcacccaag gcttgaagga	60
	tgaaagaatg aaagatggac ggaagaaaat taggacotta attccttgtt cagttcag	118
10	<210> 43 <211> 150 <212> ADN <213> Homo sapiens	
15	<400> 43	
	ttggggtaaa ttttcattgt catatgtgga atttaaataat accatcatct acaaagaatt	60
20	ccacagagtt aaatatctta agttaaacac ttaaaataag tgtttgctg atattttgat	120
	gacagataaa cagagtctaa ttcccacccc	150
25	<210> 44 <211> 150 <212> ADN <213> Homo sapiens	
30	<400> 44	
	ttggggtaaa ttttcattgt catatgtgga atttaaataat accatcatct acaaagaatt	60
35	ccacagagtt aaatatctta agttaaacac ttaaaataag tgtttgctg atattttgat	120
	gatagataaa cagagtctaa ttcccacccc	150
40	<210> 45 <211> 145 <212> ADN <213> Homo sapiens	
45	<400> 45	
	tgcaattcaa atcaggaagt atgacaaaa gacagagatc ttttttgat gatccctagc	60
50	ctagcaatgc ctggcagcca tgcaggtgca atgtcaacct taaataatgt attgcaaact	120
	cagagctgac aaacctgat gttgc	145
55	<210> 46 <211> 145 <212> ADN <213> Homo sapiens	
60	<400> 46	

ES 2 704 701 T3

5	tgcaattcaa atcaggaagt atgaccacaaa gacagagatc ttttttggat gatccctagc	60
	ctagcaatgc ctggcagcca tgcaggtgca atgtcaacct taaataatgt attgcaaatt	120
10	cagagctgac aaacctcgat gttgc	145
	<210> 47	
	<211> 124	
15	<212> ADN	
	<213> Homo sapiens	
	<400> 47	
20	ctgtgctctg cgaatagctg cagaagtaac ttggggaccc aaaataaagc agaatgctaa	60
	tgtcaagtcc tgagaaccaa gccctgggac tctggtgcc tttcggattc tccatgagca	120
25	tggt	124
	<210> 48	
	<211> 124	
	<212> ADN	
30	<213> Homo sapiens	
	<400> 48	
35	ctgtgctctg cgaatagctg cagaagtaac ttggggaccc aaaataaagc agaatgctaa	60
	tgtcaagtcc tgagaaccaa gccctgggac tctggtgcc ttttggattc tccatgagca	120
	tggt	124
40	<210> 49	
	<211> 118	
	<212> ADN	
	<213> Homo sapiens	
45	<400> 49	
50	tttttccagc caactcaagg ccaaaaaaaaaa tttcttaata tagttattat gcgaggggag	60
	gggaagcaaa ggagcacagg tagtccacag aataagacac aagaaacctc aagctgtg	118
	<210> 50	
55	<211> 118	
	<212> ADN	
	<213> Homo sapiens	
	<400> 50	
60	tttttccagc caactcaagg ccaaaaaaaaaa tttcttaata tagttattat gcgaggggag	60
	gggaagcaaa ggagcacagg tagtccacag aataggacac aagaaacctc aagctgtg	118

ES 2 704 701 T3

5	<210> 51 <211> 110 <212> ADN <213> Homo sapiens <400> 51	
10	tcttctcgtc ccctaagcaa acaacatccg cttgcttctg tctgtgtaac cacagtgaat gggtgtgcac gcttgatggg cctctgagcc cctgttgac aaaccagaaa	60 110
15	<210> 52 <211> 110 <212> ADN <213> Homo sapiens <400> 52	
20	tcttctcgtc ccctaagcaa acaacatccg cttgcttctg tctgtgtaac cacagtgaat gggtgtgcac gcttgatggg cctctgagcc cctgttgac aaaccagaaa	60 110
25	<210> 53 <211> 110 <212> ADN <213> Homo sapiens <400> 53	
30	cacatggggg cattaagaat cgcccaggga ggaggaggga gaacgcgtgc ttttcacatt tgcatttgaa ttttcgagtt cccaggatgt gtttttgtgc tcatcgatgt	60 110
35	<210> 54 <211> 110 <212> ADN <213> Homo sapiens <400> 54	
40	cacatggggg cattaagaat cgcccaggga ggaggaggga gaacgcgtgc ttttcacatt tgcatttgaa ttttcgagtt cccaggatgt gtttttgtgc tcatcgatgt	60 110
45	<210> 55 <211> 128 <212> ADN <213> Homo sapiens <400> 55	
50	gggctctgag gtgtgtgaaa taaaaacaaa tgtccatgtc tgtcctttta tggcattttg ggactttaca tttcaaacat ttcagacatg tatcacaaca cgaaggaata acagttccag ggatatct	60 120 128

ES 2 704 701 T3

5	<210> 56 <211> 128 <212> ADN <213> Homo sapiens <400> 56	
10	gggctctgag gtgtgtgaaa taaaaacaaa tgtccatgtc tgtcctttta tggcattttg 60 ggactttaca tttcaaacat ttcagacatg tatcacaaca cgagggaata acagttccag 120	
15	ggatatct 128	
20	<210> 57 <211> 21 <212> ADN <213> Secuencia artificial	
25	<220> <223> Descripción de secuencia artificial: Cebador sintético <400> 57 cacatgcaca gccagcaacc c 21	
30	<210> 58 <211> 23 <212> ADN <213> Secuencia artificial	
35	<220> <223> Descripción de secuencia artificial: Cebador sintético <400> 58 cccaaggtc ctgtgacctg agt 23	
40	<210> 59 <211> 23 <212> ADN <213> Secuencia artificial	
45	<220> <223> Descripción de secuencia artificial: Cebador sintético <400> 59 tgaggaagtg aggctcagag ggt 23	
50	<210> 60 <211> 24 <212> ADN <213> Secuencia artificial	
55	<220> <223> Descripción de secuencia artificial: Cebador sintético <400> 60 tgccagtgcg agatgaaagt cttt 24	
60	<210> 61 <211> 27 <212> ADN	

<213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 5
 <400> 61
 gtgccttcag aaccttgag atctgat 27
 10
 <210> 62
 <211> 20
 <212> ADN
 <213> Secuencia artificial
 15
 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 <400> 62
 tcccatcca ccagccacc 20
 20
 <210> 63
 <211> 25
 <212> ADN
 <213> Secuencia artificial
 25
 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 <400> 63
 aggtgtgtct ctctttgtg agggg 25
 30
 <210> 64
 <211> 21
 <212> ADN
 <213> Secuencia artificial
 35
 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 <400> 64
 cctttgtccc acctccccc c 21
 40
 <210> 65
 <211> 26
 <212> ADN
 <213> Secuencia artificial
 45
 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 <400> 65
 cctgcctac tggctgttt ctaacc 26
 50
 <210> 66
 <211> 25
 <212> ADN
 <213> Secuencia artificial
 55
 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 60
 <400> 66
 ccatccagc tgagtattcc aggag 25

5 <210> 67
 <211> 26
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 10 <400> 67
 aattgcaatg gtgagagggt gatggt 26
 <210> 68
 <211> 24
 <212> ADN
 15 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 20 <400> 68
 ccagtgagaa gtgtcttggg ttgg 24
 <210> 69
 <211> 27
 25 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 30 <400> 69
 gaaatgcctt ctcaggtaat ggaagg 27
 <210> 70
 <211> 27
 35 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 40 <400> 70
 ggtttgagca gttctgagaa tgtggct 27
 45 <210> 71
 <211> 22
 <212> ADN
 <213> Secuencia artificial
 50 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 <400> 71
 acccaaaaaca ctggaggggc ct 22
 55 <210> 72
 <211> 27
 <212> ADN
 <213> Secuencia artificial
 60 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 72
 cccttatctg ctatgtggca tacttgg 27

 5 <210> 73
 <211> 27
 <212> ADN
 <213> Secuencia artificial

 10 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 73
 gcaccagaat taaacaacg ctgacaa 27

 15 <210> 74
 <211> 22
 <212> ADN
 <213> Secuencia artificial

 20 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 74
 gcacctgaca ggcacatcag cg 22
 25
 <210> 75
 <211> 24
 <212> ADN
 <213> Secuencia artificial
 30
 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 75
 35 tgactgtata ccccaggtgc accc 24

 <210> 76
 <211> 27
 <212> ADN
 40 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 45 <400> 76
 gcactaagga tgtggaagtc tagtgtg 27

 <210> 77
 <211> 22
 50 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 55
 <400> 77
 tgtacgtggt caccagggga cg 22

 <210> 78
 60 <211> 26
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 78
 5 agtgtgagaa gagcctcaag gacagc 26

 <210> 79
 <211> 21
 <212> ADN
 10 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 79
 15 cagtggaccc tgctgcacct t 21

 <210> 80
 <211> 24
 20 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 25
 <400> 80
 gtggcaaagg agagagttgt gagg 24

 <210> 81
 30 <211> 24
 <212> ADN
 <213> Secuencia artificial

 <220>
 35 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 81
 cagtggcata gtagtcaggg ggct 24

 <210> 82
 40 <211> 21
 <212> ADN
 <213> Secuencia artificial

 <220>
 45 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 82
 50 cctctccgac aactccgcc g 21

 <210> 83
 <211> 20
 <212> ADN
 <213> Secuencia artificial
 55
 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 83
 60 aggtctgggg gccgctgaat 20

 <210> 84
 <211> 23

<212> ADN
 <213> Secuencia artificial

 5 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 84
 tcctccatt aaaccagca cct 23

 10 <210> 85
 <211> 23
 <212> ADN
 <213> Secuencia artificial

 15 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 85
 acggttctgt cctgtagggg aga 23
 20
 <210> 86
 <211> 22
 <212> ADN
 <213> Secuencia artificial
 25
 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 86
 cctgttctact tgtggcaggg ca 22
 30
 <210> 87
 <211> 20
 <212> ADN
 <213> Secuencia artificial
 35
 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 87
 gcgcagtcag atgggcgtgc 20
 40
 <210> 88
 <211> 23
 <212> ADN
 <213> Secuencia artificial
 45
 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 88
 tccagccctt gtccaaacg tgt 23
 50
 <210> 89
 <211> 21
 <212> ADN
 <213> Secuencia artificial
 55
 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 89
 gccggacctg cgaaatcca a 21
 60

5 <210> 90
 <211> 21
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

10 <400> 90
 cgggcaactg gggctctgat c 21
 <210> 91
 <211> 21
 <212> ADN
 <213> Secuencia artificial

15 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

20 <400> 91
 agcagctcc ctgactagc t 21

25 <210> 92
 <211> 23
 <212> ADN
 <213> Secuencia artificial

30 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 92
 ggcagagggg aaagacgaaa gga 23

35 <210> 93
 <211> 24
 <212> ADN
 <213> Secuencia artificial

40 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 93
 tggcattgcc tgtaataac atag 24

45 <210> 94
 <211> 23
 <212> ADN
 <213> Secuencia artificial

50 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 94
 aagcaccatt ctaatgattt tgg 23

55 <210> 95
 <211> 20
 <212> ADN
 <213> Secuencia artificial

60 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 95
 atgaagcctt ccaccaactg 20

 5 <210> 96
 <211> 27
 <212> ADN
 <213> Secuencia artificial

 10 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 96
 gatcagttgt tgttctata ttcctt 27

 15 <210> 97
 <211> 22
 <212> ADN
 <213> Secuencia artificial

 20 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 97
 acaacagaat caggtgattg ga 22
 25
 <210> 98
 <211> 25
 <212> ADN
 <213> Secuencia artificial
 30
 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 98
 35 ctgaactgaa caaagaatta aggtc 25

 <210> 99
 <211> 22
 <212> ADN
 40 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 45 <400> 99
 ttgggtaaa tttcattgt ca 22

 <210> 100
 <211> 20
 50 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 55
 <400> 100
 ggggtgggaa ttagactctg 20

 60 <210> 101
 <211> 23
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 101
 5 tgcaattcaa atcaggaagt atg 23

 <210> 102
 <211> 20
 <212> ADN
 10 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 102
 15 gcaacatcga ggttgtcag 20

 <210> 103
 <211> 20
 20 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 25
 <400> 103
 ctgtgctctg cgaatagctg 20

 <210> 104
 <211> 20
 30 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 35
 <400> 104
 accatgctca tggagaatcc 20

 <210> 105
 <211> 20
 40 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 45
 <400> 105
 ttttccagc caactcaagg 20
 50
 <210> 106
 <211> 21
 <212> ADN
 <213> Secuencia artificial
 55
 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 106
 60 cacagcttga ggttcttgt g 21

 <210> 107
 <211> 20

<212> ADN
 <213> Secuencia artificial

 <220>
 5 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 107
 tcttctcgtc ccctaagcaa 20

 10 <210> 108
 <211> 20
 <212> ADN
 <213> Secuencia artificial

 15 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 108
 20 ttctgggtt gtgcaacagg 20

 <210> 109
 <211> 20
 <212> ADN
 <213> Secuencia artificial
 25
 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 109
 30 cacatggggg cattaagaat 20

 <210> 110
 <211> 22
 <212> ADN
 35 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 110
 40 acatcgatga gcacaaaaac ac 22

 <210> 111
 <211> 20
 45 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 50
 <400> 111
 gggctctgag gtgtgtgaaa 20

 <210> 112
 <211> 24
 55 <212> ADN
 <213> Secuencia artificial

 <220>
 60 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 112
 agatatccct ggaactgta ttcc 24

5
 <210> 113
 <211> 23
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

10
 <400> 113
 acagtaactg ccttcataga tag 23
 <210> 114
 <211> 22
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

20
 <400> 114
 gtgtcagacc ctgttctaag ta 22
 <210> 115
 <211> 26
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

30
 <400> 115
 aaataaaaatt aggcataattt acaagc 26
 <210> 116
 <211> 23
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

40
 <400> 116
 gctgagtgat ttgtctgtaa ttg 23
 <210> 117
 <211> 21
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

50
 <400> 117
 cctgttcctc ccttattcc c 21
 <210> 118
 <211> 21
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

60

<400> 118
 gggaacacag actccatggt g 21
 5 <210> 119
 <211> 22
 <212> ADN
 <213> Secuencia artificial
 <220>
 10 <223> Descripción de secuencia artificial: Cebador sintético
 <400> 119
 cttaggaac cctcactgaa tg 22
 15 <210> 120
 <211> 22
 <212> ADN
 <213> Secuencia artificial
 20 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 <400> 120
 gtcctgtca gcgtttatt gc 22
 25 <210> 121
 <211> 26
 <212> ADN
 <213> Secuencia artificial
 30 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 <400> 121
 35 aataatcagt atgtgacttg gattga 26
 <210> 122
 <211> 22
 <212> ADN
 40 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 45 <400> 122
 ataggatgga tggatagatg ga 22
 <210> 123
 <211> 21
 50 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 55 <400> 123
 cagagcaaga ccctgtctca t 21
 <210> 124
 60 <211> 21
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de secuencia artificial: Cebador sintético

5 <400> 124
 tcaacagagg cttgcatgta t 21

10 <210> 125
 <211> 20
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de secuencia artificial: Cebador sintético

15 <400> 125
 ggggtgatttt cctcttttgt 20

20 <210> 126
 <211> 33
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de secuencia artificial: Cebador sintético

25 <400> 126
 aacatttgta tctttatctg tctccttatt tat 33

30 <210> 127
 <211> 27
 <212> ADN
 <213> Secuencia artificial

35 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 127
 gaacacttgt catagtttag aacgaac 27

40 <210> 128
 <211> 20
 <212> ADN
 <213> Secuencia artificial

45 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 128
 tcattgacag aattgcacca 20

50 <210> 129
 <211> 27
 <212> ADN
 <213> Secuencia artificial

55 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

60 <400> 129
 ttgtatttc atgtgtacat tcgtatc 27

<210> 130
 <211> 28

<212> ADN
 <213> Secuencia artificial

 <220>
 5 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 130
 acctatcctg tagattattt tcaactgtg 28

 10 <210> 131
 <211> 20
 <212> ADN
 <213> Secuencia artificial

 15 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 131
 tctgacccat ctaacgccta 20
 20
 <210> 132
 <211> 27
 <212> ADN
 <213> Secuencia artificial
 25
 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 132
 30 cagacagaaa gatagataga tgattga 27

 <210> 133
 <211> 21
 <212> ADN
 35 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 133
 40 atacagacag acagacaggt g 21

 <210> 134
 <211> 23
 45 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 50
 <400> 134
 gcatgtatct atcatccatc tct 23

 <210> 135
 <211> 21
 55 <212> ADN
 <213> Secuencia artificial

 <220>
 60 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 135
 tgagtgacaa attgagacct t 21

5 <210> 136
 <211> 26
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

10 <400> 136
 gtcttacaat aacagttgct actatt 26

15 <210> 137
 <211> 18
 <212> ADN
 <213> Secuencia artificial

20 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 137
 attccccaag tgaattgc 18

25 <210> 138
 <211> 26
 <212> ADN
 <213> Secuencia artificial

30 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 138
 ggtagataga ctggatagat agacga 26

35 <210> 139
 <211> 20
 <212> ADN
 <213> Secuencia artificial

40 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

45 <400> 139
 tggaacaga aatggcttgg 20

50 <210> 140
 <211> 20
 <212> ADN
 <213> Secuencia artificial

55 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 140
 gattgcagga gggaaggaag 20

60 <210> 141
 <211> 22
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 141
gagcaagaca ccatctcaag aa 22

5 <210> 142
<211> 30
<212> ADN
<213> Secuencia artificial

10 <220>
<223> Descripción de secuencia artificial: Cebador sintético

<400> 142
gaaattttac atttatgttt atgattctct 30

15 <210> 143
<211> 18
<212> ADN
<213> Secuencia artificial

20 <220>
<223> Descripción de secuencia artificial: Cebador sintético

<400> 143
ggcgactgag caagactc 18

25 <210> 144
<211> 27
<212> ADN
<213> Secuencia artificial

30 <220>
<223> Descripción de secuencia artificial: Cebador sintético

35 <400> 144
ggttattaat tgagaaaact ccttaca 27

40 <210> 145
<211> 23
<212> ADN
<213> Secuencia artificial

45 <220>
<223> Descripción de secuencia artificial: Cebador sintético

<400> 145
attttcccg atgatagtag tct 23

50 <210> 146
<211> 26
<212> ADN
<213> Secuencia artificial

55 <220>
<223> Descripción de secuencia artificial: Cebador sintético

<400> 146
gcgaatgtat gattggcaat attttt 26

60 <210> 147
<211> 27
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de secuencia artificial: Cebador sintético

5 <400> 147
acatgtatcc cagaacttaa agtaaac 27

<210> 148
<211> 22
10 <212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de secuencia artificial: Cebador sintético

15 <400> 148
gcagaaggga aaattgaagc tg 22

<210> 149
20 <211> 22
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de secuencia artificial: Cebador sintético

25 <400> 149
cagagacacc gaaccaataa ga 22

<210> 150
30 <211> 27
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de secuencia artificial: Cebador sintético

<400> 150
gccacatgaa tcaattccta taataaa 27

40 <210> 151
<211> 25
<212> ADN
<213> Secuencia artificial

45 <220>
<223> Descripción de secuencia artificial: Cebador sintético

<400> 151
gcacatgtac cctaaaactt aaaat 25

50 <210> 152
<211> 26
<212> ADN
55 <213> Secuencia artificial

<220>
<223> Descripción de secuencia artificial: Cebador sintético

60 <400> 152
gtcaacaaaa actcaacaag tagtaa 26

<210> 153

<211> 25
 <212> ADN
 <213> Secuencia artificial

5 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 153
 aagatgaaat tgccatgtaa aaata 25

10 <210> 154
 <211> 27
 <212> ADN
 <213> Secuencia artificial

15 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 154
 20 gtgtgtataa caaaattcct atgatgg 27

<210> 155
 <211> 22
 <212> ADN
 25 <213> Secuencia artificial

<220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 155
 30 gcacccaaaa ctgaatgtca ta 22

<210> 156
 <211> 20
 35 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 156
 40 ggtgagagtg agaccctgtc 20

<210> 157
 <211> 27
 45 <212> ADN
 <213> Secuencia artificial

<220>
 50 <223> Descripción de secuencia artificial: Cebador sintético

<400> 157
 tgtaataact ctacgactgt ctgtctg 27

<210> 158
 <211> 21
 <212> ADN
 55 <213> Secuencia artificial

<220>
 60 <223> Descripción de secuencia artificial: Cebador sintético

<400> 158
 gaataggagg tggatggatg g 21

 5 <210> 159
 <211> 20
 <212> ADN
 <213> Secuencia artificial

 10 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 159
 gagcgagacc ctgtctcaag 20

 15 <210> 160
 <211> 24
 <212> ADN
 <213> Secuencia artificial

 20 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 160
 ggaaaagaca taggatagca attt 24
 25
 <210> 161
 <211> 21
 <212> ADN
 <213> Secuencia artificial
 30
 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 161
 35 tctggattga tctgtctgtc c 21

 <210> 162
 <211> 26
 <212> ADN
 40 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 45 <400> 162
 gaattaaata ccatctgagc actgaa 26

 <210> 163
 <211> 27
 50 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 55
 <400> 163
 tgttataatg cattgagttt tattctg 27

 <210> 164
 60 <211> 28
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de secuencia artificial: Cebador sintético

5 <400> 164
 gcctgtctca aaaataaaga gatagaca 28

10 <210> 165
 <211> 25
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de secuencia artificial: Cebador sintético

15 <400> 165
 ttaatgaatt gaacaaatga gtgag 25

20 <210> 166
 <211> 25
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de secuencia artificial: Cebador sintético

25 <400> 166
 gcaactctgg ttgtattgtc tcat 25

30 <210> 167
 <211> 21
 <212> ADN
 <213> Secuencia artificial

35 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 167
 caaagcgaga ctctgtctca a 21

40 <210> 168
 <211> 27
 <212> ADN
 <213> Secuencia artificial

45 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 168
 gaaaatgcta tcctcttgg tataaat 27

50 <210> 169
 <211> 27
 <212> ADN
 <213> Secuencia artificial

55 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

60 <400> 169
 gggatttca agataactgt agatagg 27

<210> 170
 <211> 24

<212> ADN
 <213> Secuencia artificial

 <220>
 5 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 170
 gcttctgaaa gcttctagtt tacc 24

 10 <210> 171
 <211> 19
 <212> ADN
 <213> Secuencia artificial

 15 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 171
 tccacatcct caccaacac 19
 20
 <210> 172
 <211> 21
 <212> ADN
 <213> Secuencia artificial
 25
 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 172
 30 gcctaggaag gctactgtca a 21

 <210> 173
 <211> 18
 <212> ADN
 35 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 173
 40 ccacccgtcc atttaggc 18

 <210> 174
 <211> 27
 45 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 50
 <400> 174
 gtgaaaaagt agatataatg gttggtg 27

 <210> 175
 <211> 25
 55 <212> ADN
 <213> Secuencia artificial

 <220>
 60 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 175
 gggtttccaa gagatagacc aatta 25

5 <210> 176
 <211> 27
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

10 <400> 176
 gtcctctcat aaatccctac tcatatc 27

15 <210> 177
 <211> 27
 <212> ADN
 <213> Secuencia artificial

20 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 177
 ctgttggtac ataatagga ggtaggt 27

25 <210> 178
 <211> 19
 <212> ADN
 <213> Secuencia artificial

30 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 178
 gtcgtgggcc ccataaatc 19

35 <210> 179
 <211> 27
 <212> ADN
 <213> Secuencia artificial

40 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 179
 aaggtagata acagttcaat agaaagc 27

45 <210> 180
 <211> 26
 <212> ADN
 <213> Secuencia artificial

50 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

55 <400> 180
 gtgaaatgac tgaaaaatag taacca 26

60 <210> 181
 <211> 24
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 181
ctaggagatc atgtgggtat gatt 24

5 <210> 182
<211> 24
<212> ADN
<213> Secuencia artificial

10 <220>
<223> Descripción de secuencia artificial: Cebador sintético

<400> 182
gcagtgaata aatgaacgaa tgga 24

15 <210> 183
<211> 21
<212> ADN
<213> Secuencia artificial

20 <220>
<223> Descripción de secuencia artificial: Cebador sintético

<400> 183
25 cccaaaatta cttgagccaa t 21

<210> 184
<211> 24
<212> ADN
30 <213> Secuencia artificial

<220>
<223> Descripción de secuencia artificial: Cebador sintético

35 <400> 184
gagacaaaat gaagaaacag acag 24

<210> 185
<211> 25
40 <212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de secuencia artificial: Cebador sintético

45 <400> 185
tctttgctct catgaataga tcagt 25

<210> 186
<211> 24
50 <212> ADN
<213> Secuencia artificial

<220>
55 <223> Descripción de secuencia artificial: Cebador sintético

<400> 186
gtttgtgata atgaaccac tcag 24

60 <210> 187
<211> 27
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de secuencia artificial: Cebador sintético

5 <400> 187
tgaacacaga tgtaagtgt gtatatg 27

<210> 188
<211> 23
10 <212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de secuencia artificial: Cebador sintético

15 <400> 188
gtctgaggtg gacagttatg aaa 23

<210> 189
20 <211> 23
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de secuencia artificial: Cebador sintético

25 <400> 189
ctgtggctca tctatgaaa ctt 23

<210> 190
30 <211> 22
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de secuencia artificial: Cebador sintético

35 <400> 190
gaagtggtg ttgtgttatg at 22

40 <210> 191
<211> 24
<212> ADN
<213> Secuencia artificial

45 <220>
<223> Descripción de secuencia artificial: Cebador sintético

<400> 191
50 ttctgttgt atagagcagt gttt 24

<210> 192
<211> 20
<212> ADN
55 <213> Secuencia artificial

<220>
<223> Descripción de secuencia artificial: Cebador sintético

60 <400> 192
gtgacaggaa ggacggaatg 20

<210> 193

<211> 27
 <212> ADN
 <213> Secuencia artificial

5 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 193
 catgagggtt gcaaatacta tcttaac 27

10 <210> 194
 <211> 24
 <212> ADN
 <213> Secuencia artificial

15 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 194
 gttttaattt tctccaaatc tcca 24

20 <210> 195
 <211> 27
 <212> ADN
 <213> Secuencia artificial

25 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 195
 tcttagccta gatagatact tgcttc 27

30 <210> 196
 <211> 23
 <212> ADN
 <213> Secuencia artificial

35 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 196
 gtcaacctt gagctatag gaa 23

40 <210> 197
 <211> 19
 <212> ADN
 <213> Secuencia artificial

45 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 197
 tcttggaac aaaagtatt 19

50 <210> 198
 <211> 21
 <212> ADN
 <213> Secuencia artificial

55 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

60 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 198
 aaccttaca caaagctaga a 21

 5 <210> 199
 <211> 20
 <212> ADN
 <213> Secuencia artificial

 10 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 199
 actaagcctt ggggatccag 20

 15 <210> 200
 <211> 21
 <212> ADN
 <213> Secuencia artificial

 20 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 200
 tgctgtggaa atactaaaag g 21
 25 <210> 201
 <211> 20
 <212> ADN
 <213> Secuencia artificial
 30 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 201
 ctccagaggt aatcctgtga 20
 35 <210> 202
 <211> 21
 <212> ADN
 <213> Secuencia artificial
 40 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 202
 tggtgtgaga tggatcttag g 21
 45 <210> 203
 <211> 22
 <212> ADN
 <213> Secuencia artificial
 50 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 203
 gtataatcca tgaatcttgt tt 22
 55 <210> 204
 <211> 22
 <212> ADN
 <213> Secuencia artificial
 60 <210> 204
 <211> 22
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 204
 5 ttcaaattgt atataagaga gt 22

 <210> 205
 <211> 20
 <212> ADN
 10 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 205
 15 gcaggaaagt tattttaat 20

 <210> 206
 <211> 21
 20 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 25
 <400> 206
 tgctgagaa agctaact t 21

 <210> 207
 30 <211> 20
 <212> ADN
 <213> Secuencia artificial

 <220>
 35 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 207
 cagtgtttgg aaattgtctg 20

 40 <210> 208
 <211> 20
 <212> ADN
 <213> Secuencia artificial

 45 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 208
 50 ggcactggga gattattga 20

 <210> 209
 <211> 20
 <212> ADN
 55 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 209
 60 tcctgtgtt aagtacacat 20

 <210> 210
 <211> 18
 <212> ADN

<213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 5
 <400> 210
 gggccgtaat tacttttg 18
 <210> 211
 10 <211> 21
 <212> ADN
 <213> Secuencia artificial
 <220>
 15 <223> Descripción de secuencia artificial: Cebador sintético
 <400> 211
 actcagtagg cactttgtgt c 21
 20 <210> 212
 <211> 18
 <212> ADN
 <213> Secuencia artificial
 25 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 <400> 212
 30 tcttcacca caccaatc 18
 <210> 213
 <211> 19
 <212> ADN
 <213> Secuencia artificial
 35 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 <400> 213
 40 tggcttttca aaggtaaaa 19
 <210> 214
 <211> 21
 <212> ADN
 45 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 50 <400> 214
 gcaacgttaa catctgaatt t 21
 <210> 215
 55 <400> 215
 000
 <210> 216
 <211> 22
 60 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de secuencia artificial: Cebador sintético

 5 <400> 216
 atztatatg tcatgatcta ag 22

 10 <210> 217
 <211> 19
 <212> ADN
 <213> Secuencia artificial

 15 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 20 <400> 217
 agagattaca ggtgtgagc 19

 <210> 218
 <211> 19
 <212> ADN
 <213> Secuencia artificial

 25 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 218
 atgacctca actgcctct 19

 30 <210> 219
 <211> 20
 <212> ADN
 <213> Secuencia artificial

 35 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 219
 tgaaactcaa aagagaaaag 20

 40 <210> 220
 <211> 20
 <212> ADN
 <213> Secuencia artificial

 45 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 220
 acagatttct acttaaaatt 20

 50 <210> 221
 <211> 20
 <212> ADN
 <213> Secuencia artificial

 55 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 221
 tgaaactcaa aagagaaaag 20

 60 <210> 222
 <211> 20

<212> ADN
 <213> Secuencia artificial

5

<220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 222
 acagatttct acttaaaatt 20

10

<210> 223
 <211> 20
 <212> ADN
 <213> Secuencia artificial

15

<220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 223
 gcaaaggggt actctatgta 20

20

<210> 224
 <211> 20
 <212> ADN
 <213> Secuencia artificial

25

<220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 224
 tatcgggtca tctgttaaa 20

30

<210> 225
 <211> 22
 <212> ADN
 <213> Secuencia artificial

35

<220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 225
 tctaacaaag ctctgtccaa aa 22

40

<210> 226
 <211> 21
 <212> ADN
 <213> Secuencia artificial

45

<220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 226
 ccacactgaa taactggaac a 21

50

<210> 227
 <211> 21
 <212> ADN
 <213> Secuencia artificial

55

<220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 227
 gcaagcaagc tctctacctt c 21

60

5 <210> 228
 <211> 22
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de secuencia artificial: Cebador sintético

10 <400> 228
 tgttctcca aaattcacat gc 22

15 <210> 229
 <211> 21
 <212> ADN
 <213> Secuencia artificial

20 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 229
 atttcactat tccttcattt t 21

25 <210> 230
 <211> 22
 <212> ADN
 <213> Secuencia artificial

30 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 230
 taattgtgc acactaaatt ac 22

35 <210> 231
 <211> 21
 <212> ADN
 <213> Secuencia artificial

40 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

45 <400> 231
 aaaaagccac agaatcagt c 21

50 <210> 232
 <211> 22
 <212> ADN
 <213> Secuencia artificial

55 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 232
 ttcttatatc tcactgggca tt 22

60 <210> 233
 <211> 22
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 233
 ggatggtaga agagaagaaa gg 22

5 <210> 234
 <211> 22
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 234
 ggatggtaga agagaagaaa gg 22

15 <210> 235
 <211> 20
 <212> ADN
 <213> Secuencia artificial

20 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 235
 25 tgcaaagatg cagaaccaac 20

 <210> 236
 <211> 22
 <212> ADN
 30 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

35 <400> 236
 tttgttcct tgcctggct ga 22

 <210> 237
 <211> 20
 40 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

45 <400> 237
 tgcaaagatg cagaaccaac 20

 <210> 238
 <211> 21
 50 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

55 <400> 238
 gcctccagct ctatccaagt t 21

60 <210> 239
 <211> 22
 <212> ADN
 <213> Secuencia artificial

<220>
<223> Descripción de secuencia artificial: Cebador sintético

5 <400> 239
ccttaataatc ttccatgtc ca 22

<210> 240
<211> 22
10 <212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de secuencia artificial: Cebador sintético

15 <400> 240
attgtagtg cctctctgc tt 22

<210> 241
20 <211> 20
<212> ADN
<213> Secuencia artificial

<220>
25 <223> Descripción de secuencia artificial: Cebador sintético

<400> 241
gagaagtgag gtcagcagct 20

30 <210> 242
<211> 22
<212> ADN
<213> Secuencia artificial

35 <220>
<223> Descripción de secuencia artificial: Cebador sintético

<400> 242
40 ttctaaatt tcattgaac ag 22

<210> 243
<211> 20
<212> ADN
<213> Secuencia artificial

45 <220>
<223> Descripción de secuencia artificial: Cebador sintético

<400> 243
50 gaaattggca atctgattct 20

<210> 244
<211> 21
<212> ADN
55 <213> Secuencia artificial

<220>
<223> Descripción de secuencia artificial: Cebador sintético

60 <400> 244
caactgtcc ttattgatg t 21

<210> 245

<211> 22
 <212> ADN
 <213> Secuencia artificial

5 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 245
 ctatgttgat aaaacattga aa 22

10 <210> 246
 <211> 20
 <212> ADN
 <213> Secuencia artificial

15 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 246
 gcctgtctgg aatatagttt 20

20 <210> 247
 <211> 22
 <212> ADN
 <213> Secuencia artificial

25 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 247
 cagggcatat aatctaagct gt 22

30 <210> 248
 <211> 21
 <212> ADN
 <213> Secuencia artificial

35 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 248
 caatgactct gagtgagca c 21

40 <210> 249
 <211> 18
 <212> ADN
 <213> Secuencia artificial

45 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 249
 actctctccc tcccctct 18

50 <210> 250
 <211> 20
 <212> ADN
 <213> Secuencia artificial

55 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

60 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 250
 tatggcccca aaactattct 20

 5 <210> 251
 <211> 20
 <212> ADN
 <213> Secuencia artificial

 10 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 251
 acaagtactg ggcagattga 20

 15 <210> 252
 <211> 20
 <212> ADN
 <213> Secuencia artificial

 20 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 252
 gccaggttta gcttcaagt 20
 25 <210> 253
 <211> 22
 <212> ADN
 <213> Secuencia artificial
 30 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 253
 35 ttttatatca ggagaaacac tg 22

 <210> 254
 <211> 21
 <212> ADN
 40 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 45 <400> 254
 ccagaatttt ggaggtttaa t 21

 <210> 255
 <211> 22
 50 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 55 <400> 255
 tgtcattcct ccttatctc ca 22

 <210> 256
 60 <211> 20
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 256
 5 ttctttgcc tctccaaag 20

 <210> 257
 <211> 20
 <212> ADN
 10 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 257
 15 accctggcac agtgtgact 20

 <210> 258
 <211> 20
 20 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 25
 <400> 258
 tgggcctgag ttgagaagat 20

 <210> 259
 <211> 21
 30 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 35
 <400> 259
 aatttgaag tatgtgaac g 21

 <210> 260
 <211> 20
 40 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 260
 50 ttttcccat ttccaactct 20

 <210> 261
 <211> 20
 <212> ADN
 55 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 261
 60 aaaagatgag acaggcaggt 20

 <210> 262
 <211> 20

<212> ADN
 <213> Secuencia artificial

 <220>
 5 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 262
 acccctgtga atctcaaaat 20

 10 <210> 263
 <211> 21
 <212> ADN
 <213> Secuencia artificial

 15 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 263
 gcacttgctt ctattgttg t 21
 20
 <210> 264
 <211> 20
 <212> ADN
 <213> Secuencia artificial
 25
 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 264
 cccttcctct cttccattct 20
 30
 <210> 265
 <211> 13
 <212> ADN
 35 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 265
 agcactgcag gta 13
 40
 <210> 266
 <211> 21
 45 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 50
 <400> 266
 acagatacca aagaactgca a 21

 <210> 267
 55 <211> 20
 <212> ADN
 <213> Secuencia artificial

 <220>
 60 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 267
 tggacacctt tcaacttaga 20

5 <210> 268
 <211> 22
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de secuencia artificial: Cebador sintético

10 <400> 268
 gaacagtaat gttgaacttt tt 22

15 <210> 269
 <211> 21
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de secuencia artificial: Cebador sintético

20 <400> 269
 tcttgcaaaa agcttagcac a 21

25 <210> 270
 <211> 21
 <212> ADN
 <213> Secuencia artificial

30 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 270
 aaaaagatct caaagggtcc a 21

35 <210> 271
 <211> 20
 <212> ADN
 <213> Secuencia artificial

40 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 271
 gcttttgctg aacatcaagt 20

45 <210> 272
 <211> 19
 <212> ADN
 <213> Secuencia artificial

50 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

55 <400> 272
 cctccagca gcatagtct 19

60 <210> 273
 <211> 18
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 273
 aaatccagga tgtgcagt 18

5 <210> 274
 <211> 19
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 274
 atgatgaggt cagtgggtgt 19

15 <210> 275
 <211> 22
 <212> ADN
 <213> Secuencia artificial

20 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 275
 25 catcacagat catagtaa t gg 22

<210> 276
 <211> 21
 <212> ADN
 30 <213> Secuencia artificial

<220>
 <223> Descripción de secuencia artificial: Cebador sintético

35 <400> 276
 aattattatt ttgcaggcaa t 21

<210> 277
 <211> 20
 40 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de secuencia artificial: Cebador sintético

45 <400> 277
 catgaggcaa acacctttcc 20

<210> 278
 <211> 22
 50 <212> ADN
 <213> Secuencia artificial

<220>
 55 <223> Descripción de secuencia artificial: Cebador sintético

<400> 278
 gctggactca ggataaagaa ca 22

60 <210> 279
 <211> 20
 <212> ADN
 <213> Secuencia artificial

<220>
<223> Descripción de secuencia artificial: Cebador sintético

5 <400> 279
tggaagcctg agctgactaa 20

<210> 280
<211> 20
10 <212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de secuencia artificial: Cebador sintético

15 <400> 280
ccttcttttc cccagaatc 20

<210> 281
20 <211> 21
<212> ADN
<213> Secuencia artificial

<220>
25 <223> Descripción de secuencia artificial: Cebador sintético

<400> 281
taggagaaca gaagatcaga g 21

30 <210> 282
<211> 22
<212> ADN
<213> Secuencia artificial

35 <220>
<223> Descripción de secuencia artificial: Cebador sintético

<400> 282
aaagactatt gctaaatgct tg 22

40 <210> 283
<211> 20
<212> ADN
<213> Secuencia artificial

45 <220>
<223> Descripción de secuencia artificial: Cebador sintético

<400> 283
50 taagcgtagg gctgtgtgtg 20

<210> 284
<211> 21
<212> ADN
55 <213> Secuencia artificial

<220>
<223> Descripción de secuencia artificial: Cebador sintético

60 <400> 284
ggacggatag actccagaag g 21

<210> 285

<211> 23
 <212> ADN
 <213> Secuencia artificial

5 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 285
 gaatgacctt ggcacttta tca 23

10 <210> 286
 <211> 27
 <212> ADN
 <213> Secuencia artificial

15 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 286
 aaggatagag atatacagat gaatgga 27

20 <210> 287
 <211> 18
 <212> ADN
 <213> Secuencia artificial

25 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 287
 catgcaccgc gcaaatac 18

30 <210> 288
 <211> 19
 <212> ADN
 <213> Secuencia artificial

35 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 288
 atgcctcacc cacaacac 19

40 <210> 289
 <211> 20
 <212> ADN
 <213> Secuencia artificial

45 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 289
 tccaagcct tctcactcac 20

50 <210> 290
 <211> 20
 <212> ADN
 <213> Secuencia artificial

55 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

60 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 290
 ctgggacggt gacattttct 20

 5 <210> 291
 <211> 21
 <212> ADN
 <213> Secuencia artificial

 10 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 291
 cccaggaaga gtggaaagat t 21

 15 <210> 292
 <211> 21
 <212> ADN
 <213> Secuencia artificial

 20 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 292
 ttagcttgca tgtacctgtg t 21
 25 <210> 293
 <211> 20
 <212> ADN
 <213> Secuencia artificial

 30 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 293
 agctagatgg ggtgaatttt 20
 35 <210> 294
 <211> 18
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 45 <400> 294
 tgggctgagg ggagattc 18

 <210> 295
 <211> 22
 50 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 55 <400> 295
 atcaagctaa ttaatgttat ct 22

 <210> 296
 <211> 20
 60 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de secuencia artificial: Cebador sintético

 5 <400> 296
 aatgaataag gtcctcagag 20

 <210> 297
 <211> 21
 <212> ADN
 10 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 15 <400> 297
 ttaatctga tcattgcctt a 21

 <210> 298
 <211> 18
 20 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 25

 <400> 298
 agctgtgggt gaccttga 18

 <210> 299
 30 <211> 20
 <212> ADN
 <213> Secuencia artificial

 <220>
 35 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 299
 tgtccacca ttgtgtatta 20

 40 <210> 300
 <211> 20
 <212> ADN
 <213> Secuencia artificial

 45 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 300
 tcagacttga agtccaggat 20
 50

 <210> 301
 <211> 20
 <212> ADN
 <213> Secuencia artificial
 55

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 301
 60 gcttcagggg tgtagtttt 20

 <210> 302
 <211> 20

<212> ADN
 <213> Secuencia artificial

 <220>
 5 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 302
 ctttgtgaaa agtcgtccag 20

 10 <210> 303
 <211> 18
 <212> ADN
 <213> Secuencia artificial

 15 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 303
 ccatcatgga aagcatgg 18
 20
 <210> 304
 <211> 20
 <212> ADN
 <213> Secuencia artificial
 25
 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 304
 30 tcatctccat gactgcacta 20

 <210> 305
 <211> 20
 <212> ADN
 35 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 305
 40 gagatgacgg agtagctcat 20

 <210> 306
 <211> 18
 45 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético
 50
 <400> 306
 cccagctgca ctgtctac 18

 <210> 307
 <211> 20
 55 <212> ADN
 <213> Secuencia artificial

 <220>
 60 <223> Descripción de secuencia artificial: Cebador sintético

 <400> 307
 tcttgtcca atcacaggac 20

ES 2 704 701 T3

<210> 308
 <211> 20
 <212> ADN
 5 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 10 <400> 308
 atgctgttag ctgaagctct 20

 <210> 309
 <211> 20
 15 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 20 <400> 309
 tgaaagctcc taaagcagag 20

 <210> 310
 <211> 20
 25 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 30 <400> 310
 ttgaagagat gtgctatcat 20

 <210> 311
 <211> 54
 35 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Descripción de secuencia artificial: Cebador sintético

 40 <400> 311
 gccgcctgca gccgcgcccc cccgtgcccc cgccccgccg ccggccccggg cgcc 54

 45 <210> 312
 <211> 601
 <212> ADN
 <213> Homo sapiens

 50 catagtgaca ggtatatgcc caactaactg tggaaaacag ttctttcttt caaccttact 60 <400>
 catcaccctc acggtctgtt tatgaggctc tcctccacca gccagaaagg atgacgtgcc 120 312
 55 atacctgcaa aacttataca gcatcaacag aatgaatctt tccaacaagc cgaaacattg 180
 agtattgtgg cacagaatat gccccaccca ttactcaatc tagatatacct ttatttccac 240
 cgtctcatga ttttcttttt cctggaaaac aaaagtattt ctttcatagc ccagctagca 300
 60 ygataaatca gcgagtcaga attctagctt tgttgtaagg ttttgcgaaat atctgatcct 360
 cttattttgt acttttctat ttcttaggca aatctgagta tttcaccag ttttcottaa 420
 ctaggcattg aaaactcagt ttttttctta caaaccttca tgtcttctctg ctcatattgca 480
 cagtcttatac ttgcacctcc tataaaatgg agaaacttga cattaaaacg taatttttat 540
 tacattttga gggattccca gagaattttt ccccaatctc cttaggtagg gacttcttta 600
 c 601

ES 2 704 701 T3

5

10 <210> 313
<211> 601
<212> ADN
<213> Homo sapiens

15 <400> 313

	gtgggaacta tagtaaagaa gtccctacct aaggagattg gggaaaaatt ctctgggaat	60
	ccctcaaaat gtaataaaaa ttacgtttta atgtcaagtt tctccatttt ataggaggtg	120
20	caagataaga ctgtgcaaat gagcaggaag acatgaaggt ttgtaagaaa aaaactgagt	180
	tttcaatgcc tagttaagga aaactgggtg aaatactcag atttgcctag gaaatagaaa	240
25	agtacaaaat aagaggatca gatattcgca aaaccttaca acaaagctag aattctgact	300
	ygctgattta tcgtgctagc tgggctatga aagaaatact tttgttttcc aggaaaaaga	360
	aaatcatgag acggtggaat aaaaggatat ctagattgag taatgggtgg ggcataattct	420
30	gtgccacaat actcaatgtt tcggcttgtt ggaaagattc attctgttga tgctgtataa	480
	gttttgacag tatggcacgt catcctttct ggctgggtgga ggagagcctc ataaacagac	540
	cgtgaggggtg atgagtaagg ttgaaagaaa gaactgtttt ccacagttag ttgggcatat	600
35	a	601

40 <210> 314
<211> 650
<212> ADN
<213> Homo sapiens

45 <400> 314

	tttattgggtc ctgactggta caaatactga taaaaaggat ttttaagatca tattcatact	60
	tttgggggaat gagagccaca attaatatac aatgtctgcc atgagattgg atgcaagagt	120
50	atggcactca tactattcct acttctgtct aattacacta tttgtttctg tgtgcaaaaa	180
	tctttggtag gtggtggatg tgcccaagac acagggaaga aaaagaagta aacagggaag	240
55	tacaacacag actctgaaat ggggcatcat ggaagacgga gctttgtcgt cttggtcttt	300
	gctgtatatt cacttcctac aacagtgcga aataccttgt ggatgcttaa atatatataa	360
	tgaatgcata aatgaaaaga gtaataaag agtgatatatg aaagtatgta gataaaattc	420
60	ttcactaagc cttggggatc cagctgcttm aggactaaga ccgtatctag ctcccttttag	480
	tatttccaca gcatgccatg gagatacatg tttctgatta tatatgatac atggaaatta	540
	tatgttggtg aatgagtgat tgagtaaatg tgtactaggg cagctaataca taaatatctc	600
65	tactattgct aaaatgactg gatttatcca ttcttctga gagtttatac	650

ES 2 704 701 T3

5	<210> 315 <211> 626 <212> ADN <213> Homo sapiens <400> 315	
10		
15	ctgcttaagg actaagaccr tatctagctc ctttttagtat ttccacagca tgccatggag atacatgttt ctgattatat atgatacatg gaaattatat gttgttgaat gagtgattga gtaaatgtgt actagggcag ctaatcataa atatttctac tattgctaaa atgactggat ttatccattc cttctgagag tttatactga ttgcttatat tgtatcaaat accgtaactg agggaatgt ttactcaaac taatagcacc attcaaattt atgcaaaca taacactata tctttaaaat gttttcacta aaagctgcat aaagagtgtt ttcaacaaca atagaataat tttacaatct tttttcttgc ttaatggcca tttgtgcctt ctgacatgct gctagccatt caaaggtcac actaccttga agttgaagat caagacaaat gattagactc ataaaagaca aatcacgtct ttctggacag gtgattatta ataattaatt agcattttaa catgtattat ttaagttctt ttttaagttat aaagtctttg atttgctaaa cagttttaa atgaataaa acataaaata ataatagtta ccattt	60 120 180 240 300 360 420 480 540 600 626
35	<210> 316 <211> 1113 <212> ADN <213> Homo sapiens <400> 316	
40		
45		
50		
55		
60		
65		

ES 2 704 701 T3

caagagctgc atctcactcc aatTTTTctt ctccctataa ccttatctag attcccagtt	60
gagggaaacg atgacctaat tcctctcagt ttaaattgcaa cacaggagca aattccaaat	120
atctatgctg gtcttgctgg gattgcagaa cccaggggtg gttatcctcc tccagaggta	180
atcctgtgat cagcactaac rccacatacc agccctttca tcagcttggt ggagaagcat	240
ctttacttcc caccaagcag tgacctagat accatctcac accagttaga atcaggatca	300
ttaaaaagtc aagaaaaaac agatgctgaa gaggatgtgg agaaatagga atgcttttac	360
actgttagtg ggaatgtaa ttagttcaac cattgtcaaa gacagtgtgg cgatccctca	420
cagatctaga accagaaata ccatttgacc cagcaatccc attactgggt ctatacccaa	480
aggattataa attactctac tataaagaca catgcacaca tatgtttatt gcagcaccat	540
tcacaatagc aaagaattgc aaccaaccct aatgcccatc aatgacagac tggataaaga	600
aaatctggca catatacacc atggaatact acgcagccat aaaaaaggat gagtttatgt	660
cctttacagg gacatggatg aagctggaaa ccatcattct cagcaaaacta acacaggaac	720
agaaaaccaa acacatgttc tcactcacia gtgggagttg aacaatgaga acacatggac	780
acagggaggg gaacatcaca caccactgct tgtcaggggg tggggggcta ggggaaggat	840
agcattagga gaaataccta atgtagatga agggttgatg ggtgcagcaa accaccatgg	900
catgtgtata cctgtgtaac aaacctccat gttctgcacg tgtatcccag aacttaaagt	960
acaatacaaa aaaaaaaaaa agtghtaatcc agtttacatt ttcaaggcca aagtgggtac	1020
aatgtatatc atcttggggt aagaagagaa aaggaaaaat tottgcttta aatcttagaa	1080
gtctgggtttt tttccctggt ttgtacccca tcc	1113

ES 2 704 701 T3

5	<210> 317 <211> 1113 <212> ADN <213> Homo sapiens	
10	<400> 317	
	ttcccagttg agggaaacga tgacctaat cctctcagtt taaatgcaac acaggagcaa	60
	attccaaata tctatgctgg tcttgctggg attgcagaac cccaggggtg ttatcctcct	120
15	ccagaggtaa tcctgtgatc agcactaacg ccacatacca gccctttcat cagcttggtg	180
	gagaagcatc ttacttccc rccaagcagt gacctagata ccatctcaca ccagttagaa	240
	tcaggatcat taaaaagtca agaaaaaaca gatgctgaag aggatgtgga gaaataggaa	300
20	tgcttttaca ctgttagtgg gaatgtaaat tagttcaacc attgtcaaag acagtgtggc	360
	gatccctcac agatctagaa ccagaaatac catttgaccc agcaatocca ttactgggtc	420
25	tatacccaaa ggattataaa ttactctact ataaagacac atgcacacat atgtttattg	480
	cagcaccatt cacaatagca aagaattgca accaacccta atgcccatca atgacagact	540
	ggataaagaa aatctggcac atatacacca tgggaatacta cgcagccata aaaaaggatg	600
30	agtttatgtc ctttacaggg acatggatga agctggaaac catcattctc agcaaaactaa	660
	cacaggaaca gaaaaccaa cacatgttct cactcacaag tgggagttga acaatgagaa	720
	cacatggaca cagggagggg aacatcacac accactgctt gtcagggggg ggggggctag	780
35	gggaaggata gcattaggag aaatacctaa tgtagatgaa gggttgatgg gtgcagcaaa	840
	ccaccatggc atgtgtatac ctgtgttaaca aacctccatg ttctgcacgt gtatcccaga	900
	acttaaagta caatacaaaa aaaaaaaaaa gtgtaatcca gtttacattt tcaaggtcaa	960
40	agtgggtaca atgctatcta tcttgggcta agaagagaaa aggaaaaatt cttgctttaa	1020
	atcttagaag tctgggtttt ttccctgttt tgtaccccat cctcttggtc tctctagata	1080
45	tatttaagac tcacatagga cttgtctttt cta	1113
50		
55		
60		
65		

ES 2 704 701 T3

5	<210> 318 <211> 1001 <212> ADN <213> Homo sapiens		
10	<400> 318		
		tcatcaacta aatagttgat gaggggaaat tgttctgtat atgttcatac ttcagctaata	60
15		caattaaaaa tgatgaaata ataagattac cattttgcaa acccctaata caatggttga	120
		tccaggcaat gatcatcaat ggccactaaa atcacacaaa aggagataac cagaatatgt	180
20		gctttgtgat ggaagcatta aatacaacta atgagatatt gtttataaga aagaaaggaa	240
		gcaagaaagc aatcacacca agctctgtat ctagctacca catttaagga aaaaaagaga	300
		cagaagagca tgttaaatgt taccaagaag atacagtcag tcggaaaaaa tacagacaag	360
25		aaaatacaga gcaaaacaac ccagcttctt cagcaaatca atataaaaaa attttaagaa	420
		agagttaaag tataaactga gagacttcag aaacatatta tccaagtata atccatgaat	480
		cttgttttaa tatagatcaa rtaaaccact ataccaaaaa catcaaaaga caactgggta	540
30		aattttttaa atgactagct atttgatggt aaggaagtaa tgttactctc ttatatataa	600
		tttgaaataa tctagcgagg agcagcaaat gtgcggtat gaggaagaaa cacaattggc	660
35		cattcttgaa tcattagctg gatggtggct atatgggggt agattttact actctctaata	720
		tttacaataa tttaaaatgt tccataataa attggtgagt tatcaaaaga aatatttcta	780
		tataatagct aaaattattt ataaaagtta gtggtctcat aactttattt atttatttac	840
40		ttattttgag accgagtctc cctctgttat gcaggctgga gtgcagtggc tccatctcgg	900
		ctcactgcaa acttcacctc ctggattgaa gcgattctcc tgcctcagcc ccccgagta	960
		gctgggatta caggcttgca cccccaagcc cagctaattt t	1001
45	<210> 319 <211> 601 <212> ADN <213> Homo sapiens		
50			
55			
60			

ES 2 704 701 T3

<400> 319

5	agctaccaca ttttaaggaaa aaaagagaca gaagagcatg ttaaattgta ccaagaagat	60
	acagtcagtc ggaaaaaata cagacaagaa aatacagagc aaaacaaccc agcttcttca	120
	gcaaatcaat ataaaaaaat ttttaagaaag agttaaaagta taaactgaga gacttcagaa	180
10	acatattatc caagtataat ccatgaatct tgttttaaata tagatcaaata aaaccaactat	240
	acaaaaaaca tcaaaagaca actgggtaaa ttttttaaata gactagctat ttgatgttaa	300
	rgaagtaatg ttactctctt atatacaatt tgaaataatc tagcgaggag cagcaaatgt	360
15	gcggtctatga ggaagaaaca caattggcca ttcttgaatc attagctgga tgggtgctat	420
	atgggggtag attttactac tctctaattt tacatatatt taaaatgttc cataataaat	480
	tgttgagtta tcaaaagaaa tattttctata taatagctaa aattatttat aaaagttagt	540
20	ggtctcataa ctttatttat ttatttactt attttgagac cgagtctccc tctgttatgc	600
	a	601

25
 <210> 320
 <211> 601
 <212> ADN
 <213> Homo sapiens

30
 <400> 320

35	ccaactgata taattagata aacttagtca atatatattga atccacatt ccagcagcta	60
	ttttctccat ttgcttttat tgctgtttgt ggtgagtttg atatataatt ttaaggtgtt	120
	aacatcccta acttatgtat gggtagagct cataaatacg aacctgtgtc atgcaactca	180
40	tatatgactg tgttcaaaat aatgtgtatt agactgtaaa acgattttta tattttaaat	240
	aactttcctg catttgctgg tttcagcagg aaagttattt ttaataactt ccctgtattt	300
	sttggtttca gtattaatta atctcattaa tgctaaactt tgtgatacta ggttaaaaaa	360
45	catattcaag atagcttcag aatgttttgt atacaaatag gtctggctaa atataagtgt	420
	tagctttctc aagcatctaa atgctggcgg gcttttaaaa aaccagggct ttaaggagaa	480
	aacacctgct ctgtggtttt gtagcagata tgaagtattc aaatttctta ataaatagaa	540
50	aaagaaatat ataacagaaa caggttgacac ttgtctttct cattaagcag gtggttagta	600
	c	601

55
 <210> 321
 <211> 501
 <212> ADN
 <213> Homo sapiens

60
 <400> 321

ES 2 704 701 T3

	agctcataaa tacgaacctg tgtcatgcaa ctcatatatg actgtgttca aaataatgtg	60
5	tattagactg taaaacgatt ttaatatattt aaataacttt cctgcatttg tcggtttcag	120
	caggaaagtt atttttaata acttccctgt atttggtggg ttcagtatta attaattctca	180
	ttaatgctaa actttgtgat cctagggttaa aaaacatatt caagatagct tcagaatgtt	240
10	tggtatacaa rtaggtctgg ctaaataataa gtgtagctt tctcaagcat ctaaagtctg	300
	gcgggctttt aaaaaaccag ggctttaagg agaaaacacc tgctctgtgg tttttagtca	360
	gatatgaagt attcaaattt cttaataaat agaaaaagaa atatataaca gaaacagggt	420
15	gcacttgtct ttctcattaa gcagggtggt agtaccatta tttgcattct catagcotta	480
	atatacattt tccttctcta g	501

20 <210> 322
<211> 601
<212> ADN
<213> Homo sapiens

25 <400> 322

30

35

40	ttttgagttt ctactttagt gtotttagtgc tttctcgata tgggagaatt catgtcctcc	60
	attcagaagt atgcactaag taagaggat catgtctggg tcttgattag gtactaatct	120
45	tgaaatatta tcctacaata ggtagagca cgtatatctc ctgataatat attgaatatg	180
	atagatttaa ataattgggt aactaaatac taaagcaaat tgctgcacgt atcatttatt	240
	attcattgtg tagaaagtgc ctgactcagt gtttggaat tgtotgactt ttcctcatat	300
50	rtagtgtggg ttcattgttat tgtatataag acctgacatg aactctgttt acaataatct	360
	cccagtgcc taaagaccat aataaataat ataaccaatt ggtttcttta tgctgtcatt	420
55	tattagggca tatggcatta gtggaggatt acctgtatt acccatagtg cttagagtat	480
	gaatcacaca tgcacctga aggaaaagag gtgcaatgta ataagaaacc agatattgaa	540
	aatgcaagtt ttgttatgtt attctgggta tgtaaacctt tattcctgcc ctccatatgc	600
60	a	601

<210> 323
<211> 501

ES 2 704 701 T3

<212> ADN
<213> Homo sapiens

<400> 323

5

	aagaggtatc atgtctggtt cttgattagg tactaatctt gaaatactat cctacagtag	60
10	gtagagcac gtatatctcc tgataatata ttgaatatga tagatttaaa taattgggta	120
	actaaatact aaagcaaatt gctgcacgta tcattttatta ttcattgtgt agaaagtgcc	180
	tgactcagtg tttggaaatt gtctgacttt tctcatata tagtgtggtt tcatgttatt	240
15	gtatataaga mctgacatga accctgttta caataatctc ccagtgccat aaagaccata	300
	ataaataata taaccaattg gtttctttat gctgtcattt attagggcat atggcattag	360
	tggaggatta ctttgtatta cccatagtgc ttagagtatg aatcacacat gcaccttgaa	420
20	ggaaaagagg tgcaatgtaa taagaaacca gatattgaaa atgcaagttt tgttatgtta	480
	ttctgggtat gttaaccttt a	501

25 <210> 324
<211> 854
<212> ADN
<213> Homo sapiens

30 <400> 324

	tttcagcact gagagccaga gtggaattgt ctcccttcatt gccactgcct tcacgttttg	60
35	tgtgtcgtat ctgttttgtg atcactgaga cccaagaacc ccgacttgc cgacatacta	120
	tgtggccccc agagaggact tgagctctct gggtttcac attaccatca attaaataaa	180
	caggacagta gcttcttcct tggattgtta atttaaggct ctggataata catgtaaccg	240
40	cottatgata gagcagaatt gtaagtaggc tcatggtaga atcgttcaat gacatttccc	300
	tttccttttg gagaaacaga aattcacagg tctaattctt ttcctattaa tagttcctgr	360
	ccattattcc agaactgtcc taaaggaatt ctttctcctt aaggacacca cctcccagga	420
45	gggtatttaa agatttgcac aggcggggca cgggtggctca tgcttgtaat cccagcagtt	480
	tgggaggcca aggcgggtgg atcacttgtg ctccaggggtt caagaccggc ctggccaaca	540
50	tggtgaaacc ctatctctac taaaaacaca aaagttagct gggcctggct atgcatgcct	600
	gtaattccag ctactcggga ggctgaggct ggagaatagc ttgaaccagg gaggtggaga	660
	taacagtgag ctgagatgcc actatgacac tccagcctgg gtgacagagc aagactctct	720
55	ctcaaaaaaa aaaaaagatt tttatagtcc agtattcaac gttcatagta cacctttctt	780
	atcctagtaa atcttctttt atcaaggat atgatcccat atagtagtta actcttactc	840
	ttactttatg acaa	854

60

65

ES 2 704 701 T3

5			
	<210> 325		
	<211> 501		
	<212> ADN		
10	<213> Homo sapiens		
	<400> 325		
15	aaatacttac tattaatat gagaaactgt ggtgtttatc ggtaagatcc acgaaggaag	60	
	aagttttaaa gaaaaataact ttaaccgtgg aaaaaaaaaa ctttaaatgtc tattatcgaa	120	
	taggggccgt aattactttt gcaaaataaa aaaacaaaca agactagcta tagtgtaa	180	
20	gtaatctgta tgctttttta tgaacaatt aagtaggttg cccatttaca attagcctga	240	
	ttttctcctg ygtgggtatta tgtgtactta acaacaggac ccagtggaaa ttcactcatt	300	
	taacaaagtc tgccctacatg gtttcaaata tgggcctaac ttgaaaattc agtcataatt	360	
25	aatctaagg actaaaacaa atctgtataa aaagattctg ctaaataagg gaaaattcaa	420	
	gtctagggct acattctgaa agatattgaa gtagaacctc tgcagcaaga ctaggcttgg	480	
30	aaagtgcggg gaggaggga a	501	
	<210> 326		
	<211> 601		
35	<212> ADN		
	<213> Homo sapiens		
	<400> 326		
40	ccacatcaga aacatgagga aattctacat ggtaaaaaca gcaacaacca aaaaatactt	60	
	aaagtcaaca aaccaggaag agacatctct gaatatagga atgccaaacc ttttaacacaa	120	
45	taaaacacag attatatttc agaaggctat attatatgtg tataccaaca tcaatatgtc	180	
	cagagtagct gcacagagtt ccataatttta gtctttataa gttccctctc tcaccctact	240	
	cagtaggcac tttgtgtcta gaaacttctg tgtcaacagt tttccctctc tctggaattc	300	
50	mtcaggacag aagtgattgg tgtgggtgga gaggggtgtg ctaagagtga agttatatga	360	
	aagtaggatg gaggttagca agtagttaa gtccagaaag gcaataaggt gttaaggaag	420	
	aacttttcca ttttacaggt ctgagcaagc aggaaatcaa ctctacaaac tttgaaactt	480	
55	ggtaaatatg aaaacattct caataccatt tgtcatttaa taaatacaaa ttatactatt	540	
	ttactgcttg catctagaag tttgtcaaag atctogtctt aattattcat tgtgtcggcg	600	
60	a	601	
	<210> 327		
	<211> 601		

ES 2 704 701 T3

<212> ADN
<213> Homo sapiens

<400> 327

5

10

15

20

25

```

gacgagatct ttgacaaact tctagatgca agcagtaaaa tagtataatt tgtatttatt      60
aatgacaaa tggtaggtgag aatgttttca tatttaccaa gtttcaaagt ttgttagagtt      120
gatttcctgc ttgctcagac ctgtaaaatg gaaaagtctt tccttaacac cttattgcct      180
ttctggactt taactacttg ctaacctcca tcctactttc atataacttc actottagca      240
caaccctctt ccaccacacc aatcacttct gtcctgatga attccagaga gagggaaaac      300
ygttgacaca gaagtttcta gacacaaagt gcctactgag tagggtagag aggggaactt      360
ataaagacta aaatatggaa ctctgtgcag ctactctgga catattgatg ttggtatata      420
catataatat agccttctga aatataatct gtgttttatt gtgttaaagg tttggcattc      480
ctatattcag agatgtcttt tcctggtttg ttgactttaa gtattttttg gttgttgctg      540
tttttaccat gtagaatttc ctcatgtttc tgatgtggaa agtataagaa tatcagccag      600
a                                                                                   601

```

<210> 328
<211> 811
<212> ADN
<213> Homo sapiens

<400> 328

35

40

45

50

55

60

```

taaataatct ctaattagta taatgggtgt tcttagtgca gtgggtactt ttaaagtgct      60
ttgtggcttt tgatgaaaat tgtcttagta tttaaaactt tttcttacc aattttttgt      120
tcccatcgaa ttagcaatgc tgtaaagaaa ggcattctat tccatttttt gttgctataa      180
aggaatactt gaggtgggt aatttataaa gatgaaaagt ttatttggct cgcaattctg      240
gatggctgga aggttaagta ctgggccaca gcatctggtg ggggcctcga gctgcttcta      300
gtcataatgg aaggtgaagg gtgtaaagat catgtgacaa gggaggaaag aagagaagga      360
aggaggtgct ggttctttct atcaaccaat tcgcaagaga actaatagag aaagaactca      420
cttagccctg tgggaacaca ttaatctatt cataagggat ctggctgtat gatacaaaaca      480
cctccatta ggccccacct ccaaattgta tccattggg gatcaaattt caaaaagaga      540
tttgaagga acaaacaac catatctaag ccatagtaaa aggaatggct tttcaaaggt      600
aaaatttact ragtgtatta atattttacc aatttccagc caggagagta tgaatgttgc      660
attattacat tgctttgaaa caaagcatta gtcttaattc agaagtttaa attcagatgt      720
taacgttgca tttttaataa tgacacaacca gtactaaaat cctcattgaa atgacaaata      780
attttatttc gaatccctta tagaggttca c                                                                                   811

```

<210> 329
<211> 811

65

ES 2 704 701 T3

<212> ADN

<213> Homo sapiens

<400> 329

5

tgtcttagta	tttaaaactt	tttcttacct	aattttttgt	tcccatcgaa	ttagcaatgc	60
tgtaaagaaa	ggcatccttat	tccatttttt	gttgctataa	aggaatactt	gaggctgggt	120
aatttataaa	gatgaaaagt	ttatttggtc	cgcaattctg	gatggctgga	aggttaagta	180
ctgggccaca	gcctctgggt	ggggcctcga	gctgcttcta	gtcataatgg	aagggtgaagg	240
gtgtaaagat	catgtgacaa	gggaggaaaag	aagagaagga	aggaggtgct	ggttctttct	300
atcaaccaat	tgcgaagaga	actaatagag	aaagaactca	cttagccctg	tgggaacaca	360
ttaatctatt	cataagggat	ctggctgtat	gatacaaaaca	cctcccatta	ggccccacct	420
ccaaattgta	tcccattggg	gatcaaattt	caaaaagaga	tttggaagga	acaacaaaac	480
catatctaag	ccatagtaaa	aggaatggct	tttcaaaggt	aaaatttact	aagtgtatta	540
atattttacc	aatttcacgc	caggagagta	tgaatgttgc	attattacat	tgctttgaaa	600
caaagcatta	ktcttaattc	agaagtttaa	attcagatgt	taacgttgca	tatttaataa	660
tgcacaacca	gtactaaaat	cctcattgaa	atgacaaata	attttatttc	gaatccotta	720
tagaggttca	caatgtttta	acaatgtagt	tttgactaaa	tagaagtagt	caaaacctgt	780
cagattggaa	atagtattta	taaaacataa	a			811

30

35 <210> 330
 <211> 601
 <212> ADN
 <213> Homo sapiens

40 <400> 330

gctcatcaat	tttgacttaa	gaaaattcta	gcaacattta	tagattttgc	caaaattcag	60
cttcttccca	aatcaatcta	taagaaggct	cttccttaaa	cataattttt	atatctatga	120
actgcactag	catttactat	atatttttat	cactctcacc	attactggat	aataaataaa	180
agctcattaa	aagagttaac	aaaacatatt	tatttttaggc	atcctgaaaa	aaagattcaa	240
ttttattatc	atttctacaa	taagtattga	agaaaggaga	atttaaatta	cttcatatac	300
stgataaagg	aaaacatatg	caaggcaaat	aaacatctta	gatcatgaca	tataaaataa	360
tagattatta	ctaaagatta	aaatactttc	ttaagaatta	aagcaattct	aaaagcaata	420
gtaaataaca	ttctttctag	tgatcagaca	ctggatacta	tgtttgagat	agacagtgaa	480
ttgggaatgt	tgttttacag	aagctcctac	cttgcaagga	caggcaagtt	taaatgtcag	540
ctagaaaact	atcttgagtt	ttcagtaatg	taagattttc	ctattcaatt	tcacacttta	600
a						601

ES 2 704 701 T3

<210> 331
<211> 601
<212> ADN
<213> Homo sapiens

<400> 331

10	agaaaattct agcaacattt atagattttg ccaaaattca gtttcttccc aaatcaatct	60
	ataagaaggc tcttccttaa acataatttt tataatctatg aactgcacta gcatttacta	120
15	tatatatttta tcaactctcac cttacttgga taataaataa aagctcatta aaagagttaa	180
	caaaacatat ttatttttagg catcctgaaa aaaagattca attttattat cttttctaca	240
	ataagtattg aagaaggag aatttaaatt acttcatata cctgataaag gaaaacatat	300
20	rcaaggcaaa taaacatctt agatcatgac atataaaata atagattatt actaaagatt	360
	aaaatacttt cttaagaatt aaagcaattc taaaagcaat agtaaataac attctttcta	420
	gtgatcagac actggatact atgtttgaga tagacagtga attgggaatg ttgttttaca	480
25	gaagctccta ccttgcaagg acaggcaagt ttaaatgtca gctagaaaac tatcttgagt	540
	tttcagtaat gtaagatttt cctattcaat ttcacacttt aaattttata tatatataaa	600
30	a	601

<210> 332
<211> 1110
<212> ADN
<213> Homo sapiens

<400> 332

40	tgtagaagtt cttatcactt cctggccttt tggctaagat caagtgtgaa atgtagaagt	60
	tcctctaagc ttacttccc tcaaaaacta gttttatctt gtcagcagga ttcacttaaa	120
	aagacaaatt cagattatga atttttttot tttttacagg gtctgctctg ttgccaggc	180
45	tggagtgcag aggcacaatc tcggctcact gcagcctccg ctcctgggt tcaagcaatt	240
	ctcttgctc agcctcccga gtaactggga ttacaggcat gtgccaccac ccagctaatt	300
	tttgtatttt tagtagagat ggggtttcac cacattggtc aggctggtct cgaactgctg	360
50	gcctcaagtg atccacttgc ctggcctcc caaagtgcag agattacagg tgtgagccac	420
	cgtgcccagc ctcataaccg tttcaactac tttttcactt gacaagcaga tgtgaagtta	480
	acaaagtcac ccataattga aataaagata gtatattcct ggggyaggca gaggcagttg	540
55	aggatcatga aataactatg ttggcatagt tatttaggtg ttgatactgt tattatgcca	600

60

65

ES 2 704 701 T3

5

10

ttgaaagtta aacagagAAC cctctgggta catgttttat accaatgcac actatcttat	660
tagtccctct cataatgtgc agtcatcatt actgttacgg gttgaggtgt ccccatcctc	720
tatgggacac ctctatgttg aagtctcaga ttccctagaa tctcagaatg tgaccttgtt	780
tggaaacaga tttgctacag acgcaattag ttgagatgcg cttatatggg taggtcctaa	840
ttcagtgact ggtgtcctta aaaaaatgga aatgtacaca cgggtggtaga catgcataga	900
gggaagagag atggagaaaa tggtcaccta caagccaaag acaggggtct ggagcagatc	960
cttccctcac agccctcaga aggaaccaat cttgccaata ccttgatttt ggacttccac	1020
ctccagaact ataacacatt tctgttcttc aagcaatttg tagccatttg ttacagctaa	1080
tacaatcaca catagaaatg acttgtaa	1110

<210> 333
<211> 691
<212> ADN
<213> Homo sapiens

<400> 333

35

taaaacatgt acccagaggg ttctctgttt aactttcaat ggcataataa cagtatcaac	60
acctaaataa ctatgccaac atagttatct catgatcctc aactgcctct gcctacccca	120
ggaatatact atctttatct caaatatggg tgactttgtt aacttcacat ctgcttgcca	180
agtgaaaaag tagttgaaac rgttatgagg ctgggcacgg tggctcacac ctgtaatctc	240
tgcaactttg gaggccgagg caagtggatc acttgaggcc agcagttcga gaccagcctg	300
accaatgtgg tgaaacccca tctctactaa aaatacaaaa attagctggg tgggtggcaca	360
tgctgtaat ccagttact cgggaggctg aggcaagaga attgcttgaa ccagaggaggc	420
ggaggctgca gtgagccgag attgtgcctc tgcaactccag cctgggcaac agagcagacc	480
ctgtaaaaaa gaaaaaatt cataatctga atttgtcttt ttaagtgaat cctgctgaca	540
agataaaaact agtttttgag ggaagttaaag cttagaggaa cttctacatt tcacacttga	600
tcttagccaa aaggccagga agtgataaga acttctacat ttttaagttat tcacaagata	660
actattaatg aacctgaaat agtttgtaaa g	691

<210> 334
<211> 640
<212> ADN
<213> Homo sapiens

<400> 334

65

ES 2 704 701 T3

5

10

aaaccttttt cctgtttttac tattactaaa ggtggcacia cagcaacctc aacaactttg 60

caccatgcc aactgatgt ttacaccag cacagcattt ttggtctcta tttttattct 120

15

cctctgaatg taatgaggat tcctagatgg ctagccaatt cgaatattta aggcaactga 180

aagttagaat gtttctgaaa catagtgttg ttgccagaga gtacgaaagt tttcaagaat 240

20

atcgggcaat tctgaaagta caaagaagcc agattaaatg aaataaactt ggcgaaagttt 300

tagcaagggtg actctcatat aatgatcatt atcattacca cagttaaaag aaaagagttg 360

tttatgaaag gccatgtgtc tgcaatgaaa ctcaaaagag aaaagttaac aggtgcaara 420

25

ggtagtttta ttataaaagg agggtaggca acaagaatat gttaattttt tcttcctttt 480

catgagtaag gacaagagtt tcatatatgt gaatatTTTT atttaatttt aagtagaaat 540

ctgtttttta aatatgggta tatgcttatt tgtgtaagtg taagaaacag aagtaagtac 600

30

agcaaaccag aaataggcca aacactcctg agcataattt 640

<210> 335

<211> 919

35

<212> ADN

<213> Homo sapiens

<400> 335

40

45

50

55

60

65

ES 2 704 701 T3

5

10

tacacccagc acagcatttt tggctctctat ttttattctc ctctgaatgt aatgaggatt 60

cctagatggc tagccaattc gaatatTTaa ggcaactgaa agttagaatg tttctgaaac 120

15

atagtgttgt tgccagagag tacgaaagtt ttcaagaata tcgggcaatt ctgaaagtac 180

aaagaagcca gattaaatga aataaactg gcgaagtttt agcaagggtga ctctcatata 240

atgatcatta tcattaccac agttaaaga aaagagttgt ttatgaaagg ccatgtgtct 300

20

gcaatgaaac tcaaaagaga aaagttaaca ggtgcaaaag gtagttttat tataaaagga 360

gggtaggcaa caagaatatg tttaatTTTT ctctcttttc atgagtaagg acaagagtkt 420

catatatgtg aatatTTTTa tttaatTTTT agtagaaatc tgTTTTTaaa atatgggtat 480

25

atgcttattt gtgtaagtgt aagaaacaga agtaagtaca gcaaaccaga aataggccaa 540

acactcctga gcataatttt acttggtaga ttattcctga aacttaagga atcatctttg 600

aactcttttc ctcaactgac ttccaggatt caccatgcac ttgtgatttt cctttcattt 660

30

cactctccgt tcctcctcag tctTTTTtTc tccccaggt cTTTTttgtt catcttaaac 720

tctaaatttt agaatatccc aggggtctgc ctccggcctt ctcttttata tctacactgg 780

35

cctcatacat aatcttaacc aagtcattat tttaaatacc tacaatatac tgaaaacttc 840

taaatttgta ttttaattct tgacttcttc catacagtct agatttgat gtcacataggc 900

tgacatcatt ggctgatac 919

40

<210> 336

<211> 1001

<212> ADN

45

<213> Homo sapiens

<400> 336

ttactaaata ttctccaaca aatatatact tagtatatac tattagtgat gcatgctttc 60

50

55

60

65

5

10

aaatatttgg actatatcaa tgaatgaaac aaaaaattat ttgcccttaa ggagcttaga 120

ttctaacaga tggattcaga tgatttttat gccttatttc gtaggtttaa aagagcaatg 180

15

gggaaaaggg aagaagagag ggattgaaaa tattgagaag gttgggagac ttagcaattt 240

taagtaaggt agtgagggtta ggttttattg gcaaagtgat ttttcagcag agactgggaa 300

agatgaacgt ggtatcctgg aggaaagcct cccaggcaga gttaagctgc taacaaaagt 360

20

gcccttaggc tggagtgggc ttgtttgatt aaggaacaaa gaggtcagca tggttgcact 420

agagagaaaa aatcagatgg cgtaaggaga tgaatcaga aagatacgag gctaggcaaa 480

ggggactact atgtaatgaa yatgacctgg cagtactgac atctcctgag ggactgttag 540

25

aagtgcagac tcttgatatct tttctcaagt ctatgaaatc tagacttcat tttacaaga 600

tgacccgata tttacataca cattaaagtt ccagaagcac tgatataaca cattgttaaga 660

tcgcacagga cttcaattct ttttctggtt tttagaggca gtcctttggg gtgttttgtg 720

30

tagagtataa tgacctgaaa tatctaggat cactctagct actatcttga ggaaagagtg 780

caataaggcg gaacagttca gaggcaatgg tggctcttcta aatgaaagac acacagcact 840

35

caaaccaggc agttgaggag ggatgggaag aagttgtcaa attctagaca tatttttaaag 900

gtagtgtcca gagaatttcc ttagatgcgt aggaacatgg aggataggac ataggggtgga 960

aataaacgaa ataaagaaac tgaagctgat tctgacattt t 1001

40

<210> 337

<211> 1576

<212> ADN

45

<213> Homo sapiens

<400> 337

50

55

60

65

ES 2 704 701 T3

5

10

15

20

25

30

35

40

45

50

55

60

65

```

atacctttta agtgacatcc tagtgaatct ccattttgtca cgagacctca agctttccag      60
ttctggcaca aagtgattac tcataccatc acttcaaaat gatgattatc ttcattttatt      120
ttagttatat tgaacaaaat atacatttaa aaaatctaata tactaaatat tctccaacaa      180
atatatactt agtatatact attagtgatg catgctttca aatattttgga ctatatcaat      240
gaatgaaaca aaaaatttatt tgcccttaag gagcttagat tctaacagat ggattcagat      300
gatttttatg ccttattttcg taggttttaa agagcaatgg ggaaaaggga agaagagagg      360
gattgaaaat attgagaagg ttgggagact tagcaatttt aagtaaggta gtgagggtag      420
gtttttattg caaagtgatt tttcagcaga gactgggaaa gatgaacgtg gtatcctgga      480
ggaaagcctc ccaggcagag ttaagctgct aacaaaagtg cccttaggct ggagtgggct      540
tgtttgatta aggaacaaaag aggtcagcat ggttgcaacta gagagaaaaa atcagatggc      600
gtaaggagat gaaatcagaa agatacgagg ctaggcaaag gggtagctcta tgtaatgaac      660
atgacctggc agtactgaca tctcctgagg gactgttaga agtgcagact cttgtatctt      720

ttctcaartc tatgaaatct agacttcatt ttaacaagat gacccgatat ttacatacac      780
attaaagttc cagaagcact gatataacac attgtaagat cgcacaggac ttcaattctt      840
tttctggttt ttagaggcag tcctttgggg tgttttgtgt agagtataat gacctgaaat      900
atctaggatc actctagcta ctatcttgag gaaagagtgc aataaggcgg aacagttcag      960
aggcaatggt ggtcttctaa atgaaagaca cacagcactc aaaccaggca gttgaggagg     1020
gatgggaaga agttgtcaaa ttctagacat attttaaagg tagtgtccag agaatttcct     1080
tagatgcgta ggaacatgga ggataggaca tagggtgga ataaacgaaa taaagaaact     1140
gaagctgatt ctgacatttt agacctaaaa tctcaactaa aagttgccaa gatgggaaaa     1200
actaggtgca tcttggtttg tgagtggaaa tcagccttgt gaattaagac ttaaactgat     1260
gtctttaatc ccgtagaaat accatgaagg cagtagaaga tggctaaaga gaggtctaga     1320
ctgtaggtac aaatttaaaa gtcacttgca tttggatgct taaagtcagg atattgtgaa     1380
gtcaacagag gaataaataa atgcagagag gggaaagaaa aggcccatag actgagccat     1440
tgtctggttt atttacatat tagtatatat tttcttaaag atgtttgcta tataataatg     1500
agttacctaa agtgtgactt ttctaaattt atggggaatt ttctacattg tgttatggca     1560
ctactaaaaa taataa                                     1576

```

<210> 338

<211> 1275

<212> ADN

<213> Homo sapiens

<400> 338

5

10

gtaaaactaa ttataattaa aatcaaaata tttactgaac ctacttactc ctataatttg 60

cgttgctggt taaaaccag ctataaaaat tttgatcaaa aattttttatt ttgtaaatga 120

15

tctgacacag cataaatgtt aatcacattt ctttatttta tttgcagatt aatttgagta 180

atttgaaaaa ttattaatgt tacttaatta ctctcaacac cttacagtgt ctctgttaag 240

cactattggt gatactgaat ttaagttaca tttacaact atcagaaaat agttttttaa 300

20

gtaaaaatta tgatttgag tttaccaact aaatcttggt agctttcact gcctctattg 360

agaagagcag cagttcttat cttctctctt tttctctctt aattaacaag agattatttg 420

tatcatagcc ataaaatcag ttcaggtatt acatgaacga caccctgac tgcaatggtg 480

25

tagtttattg tattagtcca tttctatgct gctgataaag acatacataa gactgggtaa 540

tttataaaga aatagaagtt taacggactc acagttccat gtggctgggg aagcctcaca 600

atcatgatcg aaggcaaaag gcacatctta catggcaaca ggcaagagag aatgagagcc 660

30

aagtgaag agaaaccct tataaaacct tcagacctca tgagacttat tcactaccac 720

aagaacagta tgtgagaaac agtcccatga tccagttatc tcccactggg tccctccac 780

35

cacacaaggg aattatggga actgcaattc aagatgaaat gtgggtggaa gcacaacgga 840

actatatcat gatcaaagca ttattgtttt ctctgataag ctgatctaga aagtgtgct 900

tgtgatcagc tttggtgacc atgatcagtg aaatgggttaa ggaaatctac agattttgta 960

ggtttgtgcc ttgacagacg accggtatct gtttctcttt tcatgatgaa gtatctaaca 1020

40

aagctctgtc caaaattttg aatttctcgt taaawgcac atgattatag aacagaggtt 1080

acaatcaatt attcagtcac acaatcactc tcatcagtc ttaaggtgca tacctggtgt 1140

tccagttatt cagtgtggt taacaaacta cctggaactt aatggcttga aatagtcacc 1200

45

attacattat gattgtccat tctctgcac aataattagg atttggcaaa gagggaatgg 1260

tttgtttaca gacag 1275

50

<210> 339

<211> 1275

<212> ADN

<213> Homo sapiens

55

<400> 339

60

65

5

10

15

20

25

30

35

40

45

50

gtaaaactaa ttataattaa aatcaaaaata tttactgaac ctacttactc ctataaatttg	60
cgttgctggt taaaacccag ctataaaaaat tttgatcaaa aattttttatt ttgtaaatga	120
totgacacag cataaatggt aatcacattt ctttatttta tttgcagatt aatttgagta	180
atttgaaaaa ttattaatgt tacttaatta ctctcaacac cttacagtgt ctctgtgaag	240
cactattggt gatactgaat ttaagttaca ttttaacaact atcagaaaat agttttttaa	300
gtaaaaatta tgatttggag tttaccaact aaatcttggt agcttttact gcctctattg	360
agaagagcag cagttcttat ctctctcctt tttcttcttt aattaacaag agattatttg	420
tatcatagcc ataaaatcag ttcaggtatt acatgaacga caccctgac tgcaatggtg	480
tagtttattg tattagtcca ttttcatgct gctgataaag acatacataa gactgggtaa	540
tttataaaga aatagaagtt taacggactc acagttccat gtggctgggg aagcctcaca	600
atcatgatcg aaggcaaaag gcacatctta catggcaaca ggcaagagag aatgagagcc	660
aagtgaagg agaaacccct tataaaacct tcagacctca tgagacttat tcactaccac	720
aagaacagta tgtgagaaac agtcccatga tccagttatc tcccactggg tccctccac	780
cacacaaggg aattatggga actgcaattc aagatgaaat gtgggtggaa gcacaacgga	840
actatatcat gatcaaagca ttattgtttt ctctgataag ctgatctaga aagtgtgct	900
tgtgatcagc tttggtgacc atgatcagt aaatgggttaa ggaaatctac agattttgta	960
ggtttgtgcc ttgacagacg accggtatct gtttctcttt tcatgatgaa gtatctaaca	1020
aagctctgtc caaaattttg aatttctcgt taaatgcac atgattatag aacagaggtt	1080
acaatcaatt attcagtcac acaatcactc tcatcagtca ttaaggtgcr tacctggtgt	1140
tccagttatt cagtgtggta taacaaacta cctggaactt aatggcttga aatagtcacc	1200
attacattat gattgtocat tctctgcac aataattagg atttggcaaa gagggaatgg	1260
tttgtttaca gacag	1275

<210> 340

<211> 1001

<212> ADN

<213> Homo sapiens

<400> 340

60

65

5

10

15

20

25

30

35

40

45

50

55

60

65

gaaacaaaaa attgcttttt atatattgat atttttgcac ggattttctta ggatttttcta 60
 tgtacatgac catgtcatct gcaaatgaaa tagttttatt tctttatcaa tccggatgaa 120
 tttattaaaa ttatcttgcc taatttccca aatagggcct ccatgttgaa cataagtggc 180
 ggcaagggcg atctgttgct aatctcagtg gatgatattc agtgttttac aatgatcttc 240
 gacagctctg gctgttaaat tatcatagtc tgtatggcct aaacaaacaa aatacttatg 300
 attatggggg aggctgggat atccaagatc aagttgctgg caggcttagc aacctgccac 360
 tgggaagccc tgcttcccag ttttcagatg gccaccttct tatagtatct tcaccaaaga 420
 tagggcagag agagcaagca agctctctac cttctcatat aagggcacta atcccaccat 480
 gaaggcgcca ctgtcatgac stgattatgt cacaaagacc ccggggcaaa tattaccact 540
 gtgaggagta cagtttttagc atgtgaattt tgggaagaaca caaacattta gtacagagtg 600
 actattaagt atgttattaa ctatggagtt tttgtaggca ttttttaaca cattgagaaa 660
 gtttctctta ttctactttt tgttgagaag tttttatgat gacaaggcat tacattttat 720
 ccaatgactt ttctgtgtgt attgagatga ctgatttggt ctgccaattt aaatccattg 780
 ttgattctct ctaggatttt ttttatttca gttattaaat ttttcaacag gagaattact 840
 gtcttggttt tttttttgta atttctgtcc ctttactggg attccatatt taataaggca 900
 tcataatagt actctctttt agtttcttaa agatgggttt ctttagtttt taacatattt 960
 atgtctattt agaagtcttt gttaagtctg acatctgagc t 1001

<210> 341
 <211> 1001
 <212> ADN
 <213> Homo sapiens
 <400> 341

ggattttctta ggatttttcta tgtacatgac catgtcatct gcaaatgaaa tagttttatt 60
 tctttatcaa tccggatgaa tttattaaaa ttatcttgcc taatttccca aatagggcct 120
 ccatgttgaa cataagtggc ggcaagggcg atctgttgct aatctcagtg gatgatattc 180
 agtgttttac aatgatcttc gacagctctg gctgttaaat tatcatagtc tgtatggcct 240
 aaacaaacaa aatacttatg attatggggg aggctgggat atccaagatc aagttgctgg 300

5

10

15

20

25

30

35

40

<210> 342
<211> 701
<212> ADN
<213> Homo sapiens

<400> 342

cagggtotagc aacctgccac tgggaagccc tgcttcccag ttttcagatg gccaccttct 360
tatagtatct tcaccaaaga tagggcagag agagcaagca agctctctac cttctcatat 420
aagggcacta atcccaccat gaaggcgcca ctgtcatgac ctgattatgt cacaaagacc 480
ccggggcaaa tattaccact stgaggagta cagtttttagc atgtgaattt tggaagaaca 540
caaacattta gtacagagtg actattaagt atgttattaa ctatggagtt tttgtaggca 600
ttttttaaca cattgagaaa gtttctctta ttctactttt tgttgagaag tttttatgat 660
gacaaggcat tacattttat ccaatgactt ttctgtgtgt attgagatga ctgatttgtt 720
ctgccaattt aaatccattg ttgattctct ctaggatttt ttttatttca gttattaaat 780
ttttcaacag gagaattact gtcttgttct tttttttgta atttctgtcc ccttactggg 840
attccatatt taataaggca tcataatagt actcttcttt agtttcttaa agatgggttt 900
ctttagtttt taacatattt atgtctattt agaagtcttt gttaagtctg acatctgagc 960
tctctcaaag tttctgctga tttttttttt cctatgtttg g 1001

45

50

55

60

65

ggaaaccctg gcctcttgat cacactttcc tggagtttag tcccctctgc aatatgtacc 60
tgggagtcac aagaaatgcc agttacaaaa acttctgtga cagatatact agcactcaac 120
tggaaaccgg ggagagtcac aattctgtct ttccagccat atgtaactga aatggagatc 180
ttttcacctc gagccagggg tgatgggaaa gggagctggg catggctcaa tgtttagcct 240
tttcttggtc ttcaagattt catagacatt cttaaataca tgtttctttc aatgaagttt 300
gcccttagga caattcacag ctacattagg tactttttta ataatacttt tgaccatccg 360
tggttatttc attgaagaaa atctatagag cacctcagcc atcattccag aagtgactat 420
cctcctcagt aatggttctt attctaattt taaatatcat tgatgtagaa cattctatct 480
cactattcct tcattttatt rttatgggaa attatatata gttctccaga tttttaaaagc 540
cttgctaaca tgttttaagt cacacaaata ttcttctgtg ggaaaatgac agtaatttag 600
tgtgcaacaa ttatatagaa ctatttttca aacttataaa cgaagtgaaa ttctaaataa 660
aatcatttat caaacacaaa aatttgagcc agaataagga a 701

ES 2 704 701 T3

<210> 343
<211> 701
<212> ADN
<213> Homo sapiens

5

<400> 343

	aatgccagtt acaaaaactt cctgtacaga tatcctagca ctcaactgga aaccggggag	60
10	agtcacaatt ctgtctttcc agccatatgt aactgaaatg gagatctttt caccctgagc	120
	caggggtgat gggaaagga gctgggtcatg gctcaatgtt tagccttttc ttgggtcttca	180
	agatttcata gacattctta aatacatgtt tctttcaatg aagtttgccc ttaggacaat	240
15	tcacagctac attaggtact ttttaaataa tacttttgac catccgtggt tatttcattg	300
	aagaaaatct atagagcacc tcagccatca ttccagaagt gactatcctc ctcagtaatg	360
20	gttcttattc taatttttaa tatcattgat gtagaacatt ctatttcaat attccttcat	420
	tttattatta tgggaaatta tatacagttc tccagatgtt taaagccttg ctaacatgtt	480
	ttaagtcaca caaatattct yctgtgggaa aatgacagta atttagtgtg caacaattat	540
25	atagaactat ttttcaaact tataaacgaa gtgaaattct aaataaaatc atttatcaaa	600
	cacaaaaatt tgagccagaa taaggaatgt aaattacaat ttaaacacag attataaact	660
	atcttacttt taaaatgtta aaattcctaa cttgtttgaa a	701

30

<210> 344
<211> 768
<212> ADN
<213> Homo sapiens

35

<400> 344

40	ctaaaatcta ccattatatg atatccttcc caatacataa attaaaaaaa aaaacactgt	60
	agaggaaaaa gcaatatttt gaaatgatat gcttttcttt gtttgtcttc aaacaattac	120
	atcttcatca taatggttgt attagtctgt ttttacctg ctataaagaa ttgcctgaga	180
45	ctgagtaaca tataaagaaa aaagttttta ttgaccacag tttcacaggc ttaataggaa	240
	gcatgactgg gaaacttaga atcatggcag aagaggaagg ggaagcaagg atcttcttca	300
50	catggtagca ggagagagag cacaagggg gacacgctac acactttcaa acaacgagat	360
	ctcctgagaa ctctatcggg agaacagcaa gaggaagtt caccctatg attcaatcag	420
	ctcccacggg gcttctcccc tgacacatga ggaattacaa ttggatgaga gatttgggtg	480
55	gggacacaca gacaaacct atcaactgtc atggacttaa acaattgtct ttgaattgtc	540
	ttttttcata cttttatttg catcttityca ctaaaaagat gacacaaagt aatcctagtt	600
	tacatttttt accatgtaat tccatattac tttttcctga aagttactta tttttaaatc	660
60	tcaaagctct tcatacttat ggtttgatct gcacttacaa ctggatctca gaaagattga	720
	attctcccat cataccaagt tcatgtctct cactcttaat atttggtc	768

<210> 345

ES 2 704 701 T3

<211> 701
<212> ADN
<213> Homo sapiens

5 <400> 345

	aaatgatatg cttttctttg tttgtcttca aacaattaca tottcatcat aatggttgta	60
10	ttagtctgtt ttacactgc tataaagaat tgcctgagac tgagtaacat ataaagaaaa	120
	aagttttaat tgaccacagt ttcacaggct taataggaag catgactggg aaacttagaa	180
	tcatggcaga agaggaaggg gaagcaagga tcttcttcac atggtagcag gagagagagc	240
15	acaaaggggg acacgctaca cactttcaaa caacgagatc tcctgagaac tctatcggga	300
	gaacagcaag agggaagtgc acccctatga ttcaatcagc tcccaccggg cttctccct	360
	gacacatgag gaattacaat tggatgagag atttgggtgg ggacacacag acaaaccata	420
20	tcaactgtca tggacttaaa caattgtctt tgaattgtct tttttcatac ttttatttgc	480
	atcttttcac taaaaagatg rcacaaagta atcctagttt acatttttta ccatgtaatt	540
	ccatattact ttttctgaa agttacttat ttttaaactc caaagctctt catacttatg	600
25	gtttgatctg cacttacaac tggatctcag aaagattgaa ttctcccatc ataccaagtt	660
	catgtctctc actcttaata tttgttccca agacaacaat t	701

30 <210> 346
<211> 6758
<212> ADN
<213> Homo sapiens

35 <400> 346

40

45

50

55

60

65

ES 2 704 701 T3

agagtgggcc attgttctga ctagtctggg gctcccaaaa gaactggtat ctgtctcacc	60
tgactcagaa caatgataag gctgtagatc tttttggaag tctatgaaaa caggcacaat	120
gaaggcagca tgtagagat ataattccac aggaagatgc caggtaaaac aaaagagaaa	180
aagcaggaac aagctgatta ggaaatttgt gatgactaaa agtatatata caagcccaaa	240
taagatactc caaagatggt tgatagggtc tagatctcta gatatactgc tcaatgaaag	300
tgtccccctg aacaaagcca gtctgcaaag actgggtgag atgatttttt ttaaattgtca	360
agtctcagca acaacaaaaa tgacaagaca tgcacagaag caagaaaata taacacaatc	420
aaagaaaaaa aagccacaga aatcagtcct agagaaaacy gatctatgag ctgcctgaca	480
ataattataa aataactatc ataaaaatgc ccagtggagat ataagaaaac acagacaact	540
aatgaatca ggaaatgat gcatgaacaa aatgggcata tcaacagaga tggaaatgac	600
aaagataaac aaacagaaat tttggagctt aaaaatacag taagtaaagt gaataattca	660
ctaaaaatat tcaatagcag actagatcag gcagaagaaa atatcaatga acttgaagac	720
agatcatcaa gtcagaggaa caacagcaac aaaaaagaat gaaaaaagtg aagacagcct	780
aagggactta ggagtcagta ccaaggaaat caatatatac gttatagatg tatcagaaga	840
aaaagggaga aaaatgaaaa gaaagcatat ttgaaaaaat aatagctgaa gaattctcaa	900
tttcaaagag agaaattgat atacaaattc aagaagttca aaagactcta gccataataa	960

ES 2 704 701 T3

atctaaagag	actcacacta	agacatatta	tcatcaaact	gtcaaaatca	aagacaaaga	1020
attgtgaaat	ctgccaaagga	aaagtgactc	atcacacata	agagatataa	cataagattg	1080
tcacaggatt	tctgaacaga	cactttgcag	gtcagaggga	agtaggggtga	catattccag	1140
gtgctgaaag	aagaaaaacac	cctgccaacc	aagaatatgg	catccagaaa	aactttccta	1200
gaagaatgaa	ggagaaatth	agactttccc	aaataaacaa	aagctgaggg	agttcattac	1260
taccagacct	gctctgcaaa	atgctaaaga	gaaaccttca	ggtgaaacaa	aaagatgcta	1320
gacagtaaca	caaaaccact	cataaataac	ttcttcagta	aaaataatac	atcgacaaat	1380
atggtaacct	gtattaatac	tgggtgcacaa	attcactttc	aaatthttata	aataagaatt	1440
taaaggatga	aaacatctaa	aactaactat	aatctatat	aatgaatata	caatatataa	1500
aaaaatthgt	gatcacaata	acataaaatg	ggggaggtag	agctgtatag	gggtagagct	1560
thttgtatgca	attgaaatta	ccatcagtht	aaactgaact	gthtataacat	taagatgtht	1620
tatgtaattg	caatggtaac	tatattctat	agaatatatt	aaaaagaaaa	agaaaatagg	1680
aagggaatca	aagcatgtcc	ttgtaaaaaa	gtcaatgaaa	gcaaaagaaa	ggcagaaaga	1740
gtgaaaagga	ggaataaaaa	gthtataagac	ataaaaaaaa	tgaaaatagt	aatagtcctg	1800
ccatatcagt	aattacatta	aatataaatg	gattaaactc	cctaatacaa	tcatagattg	1860
gthttgcaaga	actaacttht	caattaaaga	cacacagctg	acggtgaagg	gagaaaaaaa	1920
acttccatgc	agtgaccaaa	atagaggagg	gtggctgtat	tactgtcaga	caaaataaaa	1980
thttaagtcaa	aaactgttac	aagagtaaaa	gaagggcatt	atacagthta	aaaagthaat	2040
tcgccaggca	gacacaacaa	ttataaatat	caatacataa	aaataagagc	tcctaaatat	2100
atgcagcaaa	cagacataat	tgaagaaaga	aataaatagc	taaaatggta	gaagacttht	2160
atacccccac	ttacaataat	gtataaaata	acaagacaga	atgtaaataa	aatgttagag	2220
aatttgagca	acactgtaga	ccaattggac	ctaataaata	tactcagaat	aatccatcca	2280
accaaagcag	aaacagaata	tacattctth	tcaagtacac	atttgacatt	ctctgggatt	2340
aactacatgt	tatgcaacaa	acaagtctca	acaatgttht	aaagtctgat	attacacaaa	2400
gtattgttht	tgatgacgat	ggaagaacc	tagaagccaa	tagcaaaaag	aaaatagaaa	2460
atccacacat	atgtggaat	taaactacat	gcaattaagc	aaagggccaa	agaagaagaa	2520
gaaaaaagaa	aacaccgtga	aacaaataaa	aacaaaaata	cagcatatga	aatgcatgg	2580
gatgcagcaa	aagtgatgg	aagagaaatg	thttatagtht	taaatgcaaa	cottaaaaaa	2640
gaagaagaa	aacaaaaata	ctcaaathta	caactthtaca	agtcaagaag	gtagagaaaa	2700
aagaacaaac	tataccaaaa	gctaacacag	aaagaaaaga	ataaagatta	aaaacaaaaa	2760
caatthtaaaa	aatagcagaa	ctaaaagtht	gthtctthgaa	aagatcaaca	gaattgacaa	2820

ES 2 704 701 T3

tttcttagct	acattaagaa	aaatacaaga	ctcaaataac	acaaatcagt	ggtgaaaggg	2880
ggtattataa	ctgatgccac	agaaatacaa	aaggatcata	agggactact	acaaattgta	2940
tgacaacaaa	ttgagtaacc	taggatacct	tgataaattc	caaaaaatgc	acaatatact	3000
gaatcatgaa	tacatgaccc	ttataaatca	agactaaatc	ataaagaaat	agaaaatatc	3060
aacagaccaa	taattagtaa	ggagaataaa	ctagtaatca	gaaacctccc	aacaaagaaa	3120
agcttaggac	caaatggctt	tactggagaa	ttctaccaac	cattaaaagg	ataattaaga	3180
ccaatcttcc	tcaaactttt	aaaacaaatg	ttaaagagga	ggaaactctt	tcaatctcat	3240
tcataaggtc	agcattatcc	ttataccaaa	accagacaaa	gacactatta	aaaaaactta	3300
gaccaatata	cctgatgaat	ttcgatgcaa	gaatcctcag	caaaatacta	tcaaacaatt	3360
caacagcata	cttaaagat	tatatgctgt	aatcaagatg	catttattct	ttgaatgcaa	3420
gtgtaattca	acacataaaa	ttcaatcaat	gtaatacacc	acattaacag	aatgagagac	3480
aaaaaccaca	taattatata	aactgatgca	gaaaaaaatc	tgacacagtt	caacaccttt	3540
tgtgataaaa	acactcaaca	aactaggaaa	agaaggaaac	aactttaaca	catcatatgc	3600
tcactgatga	aaatctacaa	gttctttata	aaagatcagg	aacaagacaa	taatctgcat	3660
tgttaccact	tctattatac	gtagtattgg	aagttctaata	cagagcaaat	taggcaagaa	3720
aaataaataa	aaggcatcca	aagtggaaag	gaagtaaaat	aatctctttt	tacagatgat	3780
ataaccttag	aattagaaaa	tcctaaaaat	ttcacatacc	aagaaaaagc	gtgttaaaat	3840
taataagtaa	attcagcaag	ttgactgata	caaaatcaac	acagaaagct	cagttgtgtg	3900
tctgtgtgtc	tcatacacta	acaatgaaca	atctgaaaag	gagattaaga	aaacaatttc	3960
atttacaata	gcatcaggaa	aaaaaataaa	tacttaggaa	caaacttaac	caaggggttg	4020
gaattcctgt	atactgaaaa	ctacaaatat	tgccaaaaga	aaataaagga	gacacaaata	4080
agtgatatgt	ttttaatatg	tccacccaaa	gtgatcttca	gattcaatga	aatccctatc	4140
aaagttataa	tggcattttt	ctgcagggaat	gtaaaaatct	atcctaaaat	tcatatagaa	4200
tctctaggta	ccttgagggc	caaacaattt	tgagaaaaaa	aaaagaacaa	aattggagga	4260
ctcacacttc	cagattacaa	gaatatttac	aaattacata	tttacaaaaa	aaattacaaa	4320
gccacaataa	tcaaaacaac	gtgggatttg	cataaaggca	gatatataga	ccagtgggaat	4380
agtattgaga	gtccagaaat	aaacccttag	gtatatcatc	aaatgacatt	tgacaaagtg	4440
ctggtaccac	tcaatgggaa	tgggacaatt	tgttcaacaa	atagagcaaa	gaaaactaaa	4500
catccatgtg	caaaagaata	aatctggacc	ottataattac	actatagaca	aaattaattc	4560
aaaatggatt	aaagatctaa	atgaaagatc	taaaactata	aaactcctag	gagaaaacag	4620
aggaaaaatt	tcattgcta	ttggcaacat	tttgtgatgt	gacaccaaaa	gcagagtcaa	4680
taaaagcaaa	aattagacag	atggaaatcc	atcatagttt	ataacttttg	gtcattaaag	4740

ES 2 704 701 T3

aacagtcaac agagtgaaaa ggcaatctat aaaatggggg aaaaacagaa aatatgtgca	4800
aatcacagat atctgatagg ggattcatat ccagaataaa taaagaactc ctatatctca	4860
acaacaaaaa atctaatacca atcaaaaaat gggccaaggg agtgaagata cttttctcca	4920
aagatgttat acaaatggcc aggaagcata tgaaaagatg ttcaatgtca ctaatcatca	4980
gagaaatgca aatcaaaacc acagtgcaat atcacttcac attcattaga atggcctctg	5040
tcatgaacaa cagaaaataa caagtgttga tgagtgtgta gagaaattga gacctttata	5100
taattttggc agaaattcaa aatgggtgcaa ccactataaa aaatgatatg gaggtcctca	5160
aaaaattaaa aatagaacta ccatatgatc cacaatccca cctctgggta catattcaaa	5220
agaattgaaa gcagggtgtt gaagatatat ttgcacactc tttatagcag cactgttcac	5280
aatagccaag agatgaaagt aacccaaagg ttcatgaagc aatgaataaa caaaatatat	5340
tatgtacata gagtaaaata ctgtgcagct ttaaagagaa aggaaatctt atactatgct	5400
acaacatgaa tggaacttta gggcattata gtaagtaaaa taagccagtt ttttttaaag	5460
gacaataaaa cactatacga ttctacttaa gtattttaatg ttgtcaaatt tataaatata	5520
gaatgtagaa tagtggttac cctgagctgg gggaaagggg caaaggggaa ttgttatttt	5580
aatgggtata gtttcagttc tgcaaaatga aaagggtctg gaaatctgtt tcacaatgtt	5640
gtaaatataa ttactctgaa attgtacact taaaaatggt taagatgaca aatagagttg	5700
tgatgtcttc ttttgttatt atatagaaaa actttttcat atgataatag tctttgtttt	5760
taagctgact ttgctgatat taatataatc cttccatttt tctttaaaat gctatatgct	5820
ttcacataat tttgctttac gttgatgtat ttatacataa ggtgggtttc ttatagatac	5880
cacgttgtgt gtctttttta tctaagttga tagacttgcc ttttgttagg gtattttaat	5940
aatttatatt taatgtaatt attgatatag ttgagtgtgt tgatttttgt tttctatttg	6000
ctccatctgt tgttggttct cattattcct ctgtttctac cttcttttgt actaattatt	6060
atattttatt atttttcacc tcaactgttg gottattagc cacattgctt ttaaaatttt	6120
taatgattgc tctagggttt ataataaaca aaatgttagc attttctacc atcaaatatt	6180
tttacctat tcatgtatac ttcaatttct ttcttcccat cctttgaact atatcttcat	6240
acattttact ctacatttgt tataactcag tgctttgaaa gtcaattatt tttgtctttg	6300
acagtcaatg atttttaaag agtttaacag tgaaaaaaa tggctttcat ctttttccat	6360
tagatttcat actccttctg cctgaagaat ttcttttaat agaccttgta ctgcgggtct	6420
caggcaagaa attctctcag cctttgttgg ttgaaaaaac tgcttattac accttgttt	6480
ttgaaagata ttttcactag gtatagaagt ctgggttgac agttctcatt gtttgtcaca	6540
gcatttttaa gatgccatt caattgtctt gtcttgtata attttggatt agtctggtgt	6600

ES 2 704 701 T3

5
 atttcttacc ttgttcctc tctgtgcaat gcttcaacca tcccacttca ggctgccttt 6660
 aagatgtttt cttttccctt aatctttagt ttttagctgg ttgacagtga cgcataaag 6720
 10 tgtagtgtat gaggttgctt ttattgtcac tgttggtg 6758

15
 <210> 347
 <211> 6758
 <212> ADN
 <213> Homo sapiens
 <400> 347

20

ES 2 704 701 T3

agagtgggcc attgttctga ctagtctggg gctcccaaaa gaactggtat ctgtctcacc	60
tgactcagaa caatgataag gctgtagatc tttttggaag tctatgaaaa caggcacaat	120
gaaggcagca tgtagagat ataattccac aggaagatgc caggtaaaac aaaagagaaa	180
aagcaggaac aagctgatta ggaaatttgt gatgactaaa agtatatata caagcccaaa	240
taagatactc caaagatgtt tgataggttc tagatctcta gatatactgc tcaatgaaag	300
tgtcccoctg aacaaagcca gtctgcaaag actgggtgag atgatttttt ttaaattgtca	360
agtctcagca acaacaaaaa tgacaagaca tgcacagaag caagaaaata taacacaatc	420
aaagaaaaaa aagccacaga aatcagtcct agagaaaact gatctatgag ctgcctgama	480
ataattataa aataactatc ataaaaatgc ccagtggagat ataagaaaac acagacaact	540
aatgaatca ggaaatgat gcatgaacaa aatgggcata tcaacagaga tggaaatgac	600
aaagataaac aaacagaaat tttggagctt aaaaatacag taagtaaagt gaataattca	660
ctaaaaatat tcaatagcag actagatcag gcagaagaaa atatcaatga acttgaagac	720
agatcatcaa gtcagaggaa caacagcaac aaaaaagaat gaaaaaagtg aagacagcct	780
aagggactta ggagtcagta ccaaggaaat caatatatac gttatagatg tadcagaaga	840
aaaagggaga aaaatgaaaa gaaagcatat ttgaaaaaat aatagctgaa gaattctcaa	900
tttcaaagag agaaattgat atacaaattc aagaagttca aaagactcta gccataataa	960
atctaaagag actcacacta agacatatta tcatcaaact gtcaaaatca aagacaaaga	1020
attgtgaaat ctgccaagga aaagtgactc atcacacata agagatatata cataagattg	1080
tcacaggatt tctgaacaga cactttgcag gtcagaggga agtaggggtga catattccag	1140
gtgctgaaag aagaaaacac cctgccaacc aagaatatgg catccagaaa aactttccta	1200
gaagaatgaa ggagaaatth agactttccc aaataaacia aagctgaggg agttcattac	1260
taccagacct gctctgcaaa atgctaaaga gaaaccttca ggtgaaacia aaagatgcta	1320
gacagtaaca caaaaccact cataaataac ttcttcagta aaaataatac atcgacaaat	1380
atggtaacct gtattaatac tgggtgcacia attcactttc aaattttata aataagaatt	1440
taaaggatga aaacatctaa aactaactat aaatctatat aatgaatata caatatataa	1500

aaaaatttgt gatcacaata acataaaatg ggggaggtag agctgtatag gggtagagct	1560
tttgtatgca attgaaatta ccatcagttt aaactgaact gttataacat taagatgttt	1620
tatgtaattg caatggtaac tatattctat agaatatatt aaaaagaaaa agaaaatagg	1680
aagggaatca aagcatgtcc ttgtaaaaaa gtcaatgaaa gcaaaagaaa ggcagaaaga	1740
gtgaaaagga ggaataaaaa gttataagac ataaaaaaa tgaaaatagt aatagtcttg	1800
ccatatcagt aattacatta aatataaatg gattaaactc cctaatacaa tcatagattg	1860
gtttgcaaga actaacttta caattaaaga cacacagctg acggtgaagg gagaaaaaa	1920
acttccatgc agtgaccaa atagaggagg gtggctgtat tactgtcaga caaaataaaa	1980
tttaagtcaa aaactgttac aagagtaaaa gaagggcatt atacagttaa aaaagtaaat	2040
tgcgccaggca gacacaacaa ttataaatat caatacataa aaataagagc tcctaaatat	2100
atgcagcaaa cagacataat tgaagaaaga aataaatagc taaaatggta gaagacttta	2160
atacccccac ttacaataat gtataaaata acaagacaga atgtaaataa aaatgtagag	2220
aatttgagca aactgtaga ccaattggac ctaataaata tactcagaat aatccatcca	2280
accaaagcag aacagaata tacattcttt tcaagtacac atttgacatt ctctgggatt	2340
aactacatgt tatgcaacaa acaagtctca acaatgttta aaagtctgat attacacaaa	2400
gtattgtttc tgatgacgat ggaaagaacc tagaagccaa tagcaaaaag aaaatagaaa	2460
atccacacat atgtggaaat taaactacat gcaattaagc aaagggccaa agaagaagaa	2520
gaaaaaagaa aacaccgtga aacaaataaa aacaaaaata cagcatatga aaatgcatgg	2580
gatgcagcaa aagtgatggt aagagaaatg tttatagtta taaatgcaaa ccttaaaaaa	2640
gaagaaagaa aacaaaaata ctcaaattaa caactttaca agtcaagaag gtagagaaaa	2700
aagaacaaac tatacaaaaa gctaacacag aaagaaaaga ataaagatta aaaacaaaaa	2760
caatttaaaa aatagcagaa ctaaaagttg gttctttgaa aagatcaaca gaattgacaa	2820
tttcttagct acattaagaa aaatacaaga ctcaaataac acaaatcagt ggtgaaaggg	2880
ggtattataa ctgatgccac agaaatacaa aaggatcata agggactact acaaattgta	2940
tgacaacaaa ttgagtaacc taggatacct tgataaatc caaaaaatgc acaatatact	3000
gaatcatgaa tacatgaccc ttataaatca agactaaatc ataaagaaat agaaaatatc	3060
aacagaccaa taattagtaa ggagaataaa ctagtaatca gaaacctccc aacaaagaaa	3120
agcttaggac caaatggctt tactggagaa ttctaccaac cattaaaagg ataattaaga	3180
ccaatcttcc tcaaactttt aaaacaaatg ttaaagagga ggaaactctt tcaatctcat	3240
tcataaggtc agcattatcc ttataccaaa accagacaaa gacactatta aaaaaactta	3300
gaccaatatc cctgatgaat ttcgatgcaa gaatcctcag caaaatacta tcaaacaatt	3360
caacagcata cttaaatgat tatatgctgt aatcaagatg catttattct ttgaatgcaa	3420

ES 2 704 701 T3

gtgtaattca acacataaaa ttcaatcaat gtaatacacc acattaacag aatgagagac	3480
aaaaaccaca taatttatatc aactgatgca gaaaaaaatc tgacacagtt caacaccttt	3540
tgtgataaaa acactcaaca aactagggaaa agaaggaaac aactttaaca catcatatgc	3600
tcactgatga aaatctacaa gttctttata aaagatcagg aacaagacaa taatctgcat	3660
tgttaccact tctattatac gtagtattgg aagttctaata cagagcaaata taggcaagaa	3720
aaataaataa aaggcatcca aagtggaaag gaagtaaaat aatctctttt tacagatgat	3780
ataaccttag aattagaaaa tcctaaaaat ttcacatacc aagaaaaagc gtgttaaaat	3840
taataagtaa attcagcaag ttgactgata caaaatcaac acagaaagct cagttgtgtg	3900
tctgtgtgtc tcatacacta acaatgaaca atctgaaaag gagattaaga aaacaatttc	3960
atttacaata gcatcaggaa aaaaaataaa tacttaggaa caaacttaac caaggggttg	4020
gaattcctgt atactgaaaa ctacaaatat tgccaaaaga aaataaagga gacacaaata	4080
agtgatatgt ttttaatatg tccacccaaa gtgatcttca gattcaatga aatccctatc	4140
aaagttataa tggcattttt ctgcaggaat gtaaaaattt atcctaaaat tcatatagaa	4200
tctctaggta ccctgagggc caaacaattt tgagaaaaaa aaaagaacaa aattggagga	4260
ctcacacttc cagattacaa gaatatattac aaattacata ttacaaaaaa aaattacaaa	4320
gccacaataa tcaaaacaac gtgggatttg cataaaggca gatatataga ccagtggaat	4380
agtattgaga gtccagaaat aaacccttag gtatatcatc aaatgacatt tgacaaagtg	4440
ctggtaccac tcaatgggaa tgggacaatt tgttcaacaa atagagcaaa gaaaactaaa	4500
catccatgtg caaaagaata aatctggacc cttatattac actatagaca aaattaattc	4560
aaaatggatt aaagatctaa atgaaagatc taaaactata aaactcctag gagaaaacag	4620
aggaaaaatt tcatgctaata ttggcaacat tttgtgatgt gacacccaaa gcagagtcaa	4680
taaaagcaaa aattagacag atggaaatcc atcatagttt ataacttttg gtcattaaag	4740
aacagtcaac agagtgaaaa ggcaatctat aaaatggggg aaaaacagaa aatatgtgca	4800
aatcacagat atctgatagg ggattcatat ccagaataaa taaagaactc ctatatctca	4860
acaacaaaaa atctaatacca atcaaaaaat gggccaaggg agtgaagata catttctcca	4920
aagatgttat acaaatggcc aggaagcata tgaaaagatg ttcaatgtca ctaatcatca	4980
gagaaatgca aatcaaaacc acagtgcaat atcacttcac attcattaga atggcttctg	5040
tcatgaacaa cagaaaaataa caagtgttga tgagtgtgta gagaaattga gacctttata	5100
taattttggc agaaattcaa aatggtgcaa ccactataaa aaatgatatg gaggtcctca	5160
aaaaattaaa aatagaacta ccatatgatc cacaatccca cctctgggta catattcaaa	5220
agaattgaaa gcaggtgtt gaagatatat ttgcacactc tttatagcag cactgttcac	5280

ES 2 704 701 T3

5	aatagccaag agatgaaagt aacccaaagg ttcattgaagc aatgaataaa caaaatatat	5340
	tatgtacata gagtaaaata ctgtgcagct ttaaagagaa aggaaatctt atactatgct	5400
	acaacatgaa tggaacttta gggcattata gtaagtaaaa taagccagtt ttttttaaag	5460
10	gacaaataaa cactatacga ttctacttaa gtatttaatg ttgtcaaatt tataaatata	5520
	gaatgtagaa tagtgggttac cctgagctgg gggaaagggg caaaggggaa ttgttatatt	5580
	aatgggtata gtttcagttc tgcaaaatga aaagggtctg gaaatctgtt tcacaatgtt	5640
15	gtaaatataa ttactctgaa attgtacact taaaaatggg taagatgaca aatagagttg	5700
	tgatgtcttc ttttgttatt atatagaaaa actttttcat atgataatag tctttgtttt	5760
	taagctgact ttgtgatata taatataatc cttccatttt tctttaaaat gctatatgct	5820
20	ttcacataat tttgctttac gttgatgtat ttatacataa ggtgggtttc ttatagatac	5880
	cacgttgtgt gtctttttta tctaagttga tagacttgcc ttttgttagg gtatttaaat	5940
	aatttatatt taatgtaatt attgatatag ttgagtggtg tgatttttgt tttctatttg	6000
25	ctccatctgt tgttgggtct cattattcct ctgtttctac cttcttttgt actaattatt	6060
	atattttatt atttttcatt tcaactgttg gcttattagc cacattgctt ttaaaatttt	6120
30	taatgattgc tctagggttt ataataaaca aaatgttagc attttctacc atcaaatatt	6180
	tttacctat tcatgtatac ttcaatttct ttcttcccat cctttgaact atatcttcat	6240
	acattttact ctacatttgt tataactcag tgctttgaaa gtcaattatt tttgtctttg	6300
35	acagtcaatg attttttaaag agtttaacag tgaaaaaaaaa tggctttcat ctttttccat	6360
	tagatttcat actcctctg cctgaagaat ttcttttaat agaccttgta ctgcgggtct	6420
	caggcaagaa attctctcag cctttgttgg ttgaaaaaac tgcttattac acctttgttt	6480
40	ttgaaagata ttttcaactag gtatagaagt ctgggttgac agttctcatt gtttgcaca	6540
	gcatttttaa gatgccatt caattgtctt gtcttgata attttggatt agtctggtgt	6600
	atttcttacc tttgttctc tctgtgcaat gcttcaacca tcccacttca ggctgccttt	6660
45	aagatgtttt cttttccctt aatctttagt ttttagctgg ttgacagtga cgcactaag	6720
	tgtagtgtat gaggttgctt ttattgtcac tgttgttg	6758

50
 <210> 348
 <211> 501
 <212> ADN
 <213> Homo sapiens
 55
 <400> 348

60

65

ES 2 704 701 T3

5	gaccatgtta tgacatttta gtgcttgcta agcagtaaact actgacttac tttcctgcta	60
	cactcttcag agcagaaaga gaaatctaca aaaagggcaa tgtagttggg atccaccaca	120
	gccttgagac tgggccatgt ttctacagct taccacatt ttacccccac tttctctgag	180
10	aaacaatgca aactggagaa caaggtcaga gaagttatct tggatggtag aagagaagaa	240
	aggagaagaa rggataagca gaaaatcaaa aagggcataa aaaaattact ggggaaaata	300
15	attcttagtc actcaccatt tcttatgttt gtgaaaacag aaacgaggag caagtgttgt	360
	tgtaagaatt gttcttgccc ctccccctcc accaccaca tctgtcaagc tatccctggt	420
	tcactgtttc ctctgcactc tctattaact tctttgtcct cctcttttct tttcctacag	480
20	caaagacttt ttgtcatggt t	501

<210> 349
 <211> 501
 <212> ADN
 <213> Homo sapiens

<400> 349

30	tgacttactt tcctgctaca ctcttcagag cagaaagaga aatctacaaa aagggaatg	60
	tagttgggat ccaccacagc cttagagactg ggccatgttt ctacagctta cccacatttt	120
35	acccccactt tctctgagaa acaatgcaaa ctggagaaca aggtcagaga agttatcttg	180
	gatggtagaa gagaagaaag gagaagaaag gataagcaga aaatcaaaaa gggcataaaa	240
40	aaattactgg rgaaaataat tottagtcac tcaccatttc ttatgtttgt gaaaacagaa	300
	acgaggagca agtgttggtg taagaattgt tcttgcccct cccctccac caccacatc	360
	tgtcaagcta tcctgtttc actgtttcct ctgcactctc tattaacttc tttgtcctcc	420
45	tcttttcttt tcctacagca aagacttttt gtcatgtttt gtttcttttt ctattgtttc	480
	tttccctttt ctaatccttg a	501

<210> 350
 <211> 1148
 <212> ADN
 <213> Homo sapiens

<400> 350

60

65

ES 2 704 701 T3

5	tatgagattt aatgttaaga aataaaatgt aggatctaaa acgtaatcta tagcataatc	60
	tcaaaaatgg tttagaaatg acataataat acagacattt gtgggtggta ggattatgca	120
	tatttttata tattttttaa tatatttttc aaaagcttcc tataaagaat gtaattcttt	180
10	cccaattcca aatctagctt aaacataatt ttacaaaaat tattctctca gaatgtaaac	240
	tagtaccacc tctatggaaa acattatgga gatttcctaa agagttaaaa gtagatctac	300
	catttgatcc agcaatctta atactgggta tctacccgga ggaaaagaag tcattgtatg	360
15	aaaaagacac ttgtacacat atgtttacag gaccacaatt cacaatgca aagatgcaga	420
	accaacctaa gtggccastg actaatgaga ggataaagaa gatgtggcat atatatatca	480
	gggactacta ctcagccatt acaaggaaca aaataatgtc ttttgcaaca acttggatag	540
20	agctggaggg cattattcta agtaaagtaa ttcaggaatt ggaaaaccaa aaaccgtatg	600
	ttctctctta taagtgggaa ctaagttagg aataagcaaa ggcacacaga gggacatatt	660
25	ggactttaga gactcacgag gaggagggtg ataggggact agggattaaa agaaaaacta	720
	gacattaggt acaaggtacc ctacttaagt gcactaaaat ctcagaattc accactacgt	780
	aattcaacta agtaacaaga aaccacttgt acccctaaaag ctactgaaat aaaaattatt	840
30	ctctcaaaaa ttttaagccc taaacttcag ttcctattgt ttatatttac taagaaaaac	900
	aacagaaaac actgttttaa aaatgggtga tttttttaag gttaaaggta tataagacag	960
	ctgcctaagg aaacgcagat acccctgtac cttgttggtg ttgttgtttt tcactttttt	1020
35	aaaaaacata gagatgggat ctccttatgc tgcccaggct tgtctcaaac tcctgagctc	1080
	aagcaatcct ctgacctcag actctcaaag ttttgggact acaggcgaca gtcaccatgc	1140
40	cagccaat	1148

<210> 351
 <211> 1148
 <212> ADN
 <213> Homo sapiens

<400> 351

ES 2 704 701 T3

5	tatgagattt aatgttaaga aataaaatgt aggatctaaa acgtaatcta tagcataatc	60
	tcaaaaatgg tttagaaatg acataataat acagacattt gtgggtggta ggattatgca	120
10	tatTTTTata tattttttaa tatatttttc aaaagcttcc tataaagaat gtaattcttt	180
	cccaattcca aatctagctt aaacataatt ttacaaaaat tattctctca gaatgtaaac	240
	tagtaccacc tctatggaaa acattatgga gatttcctaa agagttaaaa gtagatctac	300
15	catttgatcc agcaatctta atactgggta tctaccggga ggaaaagaag tcattgtatg	360
	aaaaagacac ttgtacacat atgtttacag gaccacaatt cacaaatgca aagatgcaga	420
	accaacctaa gtggccactg actaatgaga ggataaagaa gatgtggcat atatayatca	480
20	gggactacta ctcagccatt acaaggaaca aaataatgtc ttttgcaaca acttggatag	540
	agctggaggc cattattcta agtaaagtaa ttcaggaatt ggaaaaccaa aaaccgtatg	600
	ttctctctta taagtgggaa ctaagttagg aataagcaaa ggcacacaga gggacatatt	660
25	ggactttaga gactcacgag gaggagggtg ataggggact agggattaaa agaaaaacta	720
	gacattagggt acaaggtacc ctacttaagt gcactaaaat ctcagaattc accactacgt	780
	aattcaacta agtaacaaga aaccacttgt accccaaaag ctactgaaat aaaaattatt	840
30	ctctcaaaaa ttttaagccc taaacttcag ttcctattgt ttatatttac taagaaaaac	900
	aacagaaaac actgttttaa aaatgggtgga tttttttaag gttaaaggta tataagacag	960
	ctgcctaagg aaacgcagat acccctgtac cttgttggtg ttgttggttt tcaacttttt	1020
35	aaaaaacata gagatgggat ctccttatgc tgcccaggct tgtctcaaac tcctgagctc	1080
	aagcaatcct ctgacctcag actctcaaag ttttgggact acaggcgaca gtcaccatgc	1140
40	cagccaat	1148

45 <210> 352
 <211> 1148
 <212> ADN
 <213> Homo sapiens

50 <400> 352

55

60

65

ES 2 704 701 T3

5	tatgagattt aatgttaaga aataaaatgt aggatctaaa acgtaatcta tagcataatc	60
	tcaaaaatgg tttagaaatg acataataat acagacattt gtgggtggta ggattatgca	120
10	tattttttata tattttttaaa tataatttttc aaaagcttcc tataaagaat gtaattcttt	180
	cccaattcca aatctagctt aaacataatt ttacaaaaat tattctctca gaatgtaaac	240
	tagtaccacc tctatggaaa acattatgga gatttcctaa agagttaaaa gtagatctac	300
15	catttgatcc agcaatctta atactgggta tctaccggga ggaaaagaag tcattgtatg	360
	aaaaagacac ttgtacacat atgtttacag gaccacaatt cacaaatgca aagatgcaga	420
	accaacctaa gtggccactg actaatgaga ggataaagaa gatgtggcat atatayatca	480
20	gggactacta ctcagccatt acaaggaaca aaataatgtc ttttgcaaca acttgगतag	540
	agctggaggc cattattcta agtaaagtaa ttcaggaatt ggaaaaccaa aaaccgtatg	600
	ttctctctta taagtgggaa ctaagttagg aataagcaaa ggcacacaga gggacatatt	660
25	ggacttttaga gactcacgag gaggagggtg ataggggact agggattaaa agaaaaacta	720
	gacattaggt acaaggtacc ctacttaagt gcactaaaat ctcagaattc accactacgt	780
	aattcaacta agtaacaaga aaccacttgt accccaaaag ctactgaaat aaaaattatt	840
30	ctctcaaaaa ttttaagccc taaacttcag ttctatttgt ttatatttac taagaaaaac	900
	aacagaaaac actgttttta aaatgggtga tttttttaag gttaaaggta tataagacag	960
35	ctgcctaagg aaacgcagat acccctgtac cttgttgttg ttgttgtttt tcaacttttt	1020
	aaaaaacata gagatgggat ctccttatgc tgcccaggct tgtctcaaac tcctgagctc	1080
	aagcaatcct ctgacctcag actctcaaag ttttgggact acaggcgaca gtcaccatgc	1140
40	cagccaat	1148

45 <210> 353
 <211> 1148
 <212> ADN
 <213> Homo sapiens

50 <400> 353

55	tatgagattt aatgttaaga aataaaatgt aggatctaaa acgtaatcta tagcataatc	60
	tcaaaaatgg tttagaaatg acataataat acagacattt gtgggtggta ggattatgca	120

60

65

ES 2 704 701 T3

5	tatattttata tatatttttaa tatatttttc aaaagcttcc tataaagaat gtaattcttt	180
	cccaattcca aatctagctt aaacataatt ttacaaaaat tattctctca gaatgtaaac	240
	tagtaccacc tctatggaaa acattatgga gatttcctaa agagttaaaa gtagatctac	300
10	catttgatcc agcaatctta atactgggta tctacccgga ggaaaagaag tcattgtatg	360
	aaaaagacac ttgtacacat atgtttacag gaccacaatt cacaaatgca aagatgcaga	420
	accaacctaa gtggccactg actaatgaga ggataaagaa gatgtggcat atatatatca	480
15	gggactactr ctcagccatt acaaggaaca aaataatgtc ttttgcaaca acttggatag	540
	agctggaggc cattattcta agtaaagtaa ttcaggaatt ggaaaacca aaaccgtatg	600
	ttctctctta taagtgggaa ctaagttagg aataagcaaa ggcacacaga gggacatatt	660
20	ggactttaga gactcacgag gaggagggta ataggggact agggattaaa agaaaaacta	720
	gacattaggt acaaggtacc ctacttaagt gcactaaaat ctcagaattc accactacgt	780
	aattcaacta agtaacaaga aaccacttgt accccaaaag ctactgaaat aaaaattatt	840
25	ctctcaaaaa ttttaagccc taaacttcag ttctctattgt ttatatttac taagaaaaac	900
	aacagaaaaac actgttttaa aaatgggtga tttttttaag gttaaaggta tataagacag	960
30	ctgcctaagg aaacgcagat acccctgtac cttgttggtt ttgttggttt tcactttttt	1020
	aaaaaacata gagatgggat ctcttatgc tgcccaggct tgtctcaaac tcctgagctc	1080
	aagcaatcct ctgacctcag actctcaaag ttttgggact acaggcgaca gtcaccatgc	1140
35	cagccaat	1148

40 <210> 354
 <211> 611
 <212> ADN
 <213> Homo sapiens

45 <400> 354

50

55

60

65

ES 2 704 701 T3

5
 caaaacctca accttcacaga taagtctaag ggtgagaact tcacacaaga tgaataagaa 60
 ccaatttctt ccagggcgat gttgaacctg gaaatgaaag ccaatctctc ttggaaggcc 120
 10 tggttttag tagaatgtcagt ctttgtttca agctgtggga gaatgagaag caagacttta 180
 gggaaagagg aataaaatag atgtgcagaa ataacagagt gagaaagtct tcaggggtgtc 240
 gctagcccta attgcaggca tccctgaatc ctagaccttg gattgcaaga gactccttaa 300
 15 tatcttccca tgtccacatt tgcctcacat agtttgaatg tggcttctat tatatacaga 360
 tacaagattc aatccaacc tctaygatga ctggtcttgt gaataagcag aagaggcact 420
 aacaatatga cgtgagggat tcagggaaga gcactttctt gagcacatat cttccctggt 480
 20 ctgccagctg tagtttatga aattccacaa tgaggatgaa atggaatcac catttacaga 540
 gtactctcca gatgtctaac cctaagctag gtaccttcaa aatattatct agtttagata 600
 25 atcaaccctt t 611

<210> 355
 <211> 601
 <212> ADN
 <213> Homo sapiens
 <400> 355

35
 ttctctagtc caaaggggtg attatctaaa ctagataata ttttgaaggc acctagctta 60
 gggtttagaca tctggagagt actctgtaaa tgggtgattcc atttcatcct cattgtggaa 120
 40 tttcataaac tacagctggc agaccagga agatatgtgc tcaagaaagt gctcttccct 180
 gaatccctca cgtcatattg ttagtgcttc ttctgcttat tcacaagacc agtcatcata 240
 gaggttggat ttgaatcttg tatctgtata taatagaagc cacattcaaa ctatgtgaag 300
 45 yaaatgtgga catgggaaga tattaaggag tctcttgcaa tccaaggctc aggattcagg 360
 gatgcctgca attagggcta gcgacacct gaagactttc tcaactctgtt atttctgcac 420
 50 atctatttta ttctcttttc cctaaagtct tgcttctcat tctccacag cttgaaacaa 480
 agactgacat ttctacaaac caggccttcc aagagagatt ggctttcatt tcaggttca 540
 acatcgccct ggaagaaatt ggttcttatt catcttgtgt gaagttctca cccttagact 600
 55 t 601

<210> 356
 <211> 527
 <212> ADN
 <213> Homo sapiens
 <400> 356

65

ES 2 704 701 T3

5	gctctagaat atggcattcc agaagtggga tgctacaaat agtctcattg agagtcaact	60
	tgcacaatgt atcgtcctac cttacatca atttctgaaa caacttctct ttgcacttcc	120
10	cctatagtta catgcataat aaattctgac aactcttatg aagtcattgga ataaactttct	180
	tcttatgttt cctatcaatg tcattagccc tttatcttgt ttgagtttcc atcagcaatg	240
	ttttcaagtc ccaagatcat tcatgtatcc acaagcaatg atacgccaga tttggacaaa	300
15	taatactgaa tactatctta ttttactgac catgatcaag gcagtgtgga ttgctgcaa	360
	gtccaagaga agtgaggtca gcagctgcaa gccacctccg tcatttagaa aagcttcattg	420
	atgtagtggtg tcgtttcgat gtgacactgt ctacacagagt taaaatgatg tgmaaggaac	480
20	tgttcaatgg aaatttagaa atttctcttt ttctcaattt tagtgta	527

25 <210> 357
 <211> 601
 <212> ADN
 <213> Homo sapiens

30 <400> 357

	gaacaagatt ttcctgcttt taaaaatact acattaaagc tgaaaattta ggccaaaatt	60
35	ttcaagtggg aatagttaca ggcaattcat ctttctgggc agaaaagggt gttactgcag	120
	ctatttctgc ctgaaactgg gtggcactac tacttttttt tttttttttt taactgagca	180
	gacattttcc ttactactaaa attgagaaaa agagaaattt ctaaatttcc attgaacagt	240
40	tccttgacaa tcatttttaac tctgtgagac agtgtcacat cgaaacgaca cactacatca	300
	ygaagctttt ctaaattgacg gaggtggctt gcagctgctg acctcacttc tcttggactt	360
	ggcagcaatc cactactgct tgatcatggc agtgaaaata agatagtatt cagtattatt	420
45	tgtccaaatc tggcgtatca ttgcttgtgg atacatgaat gatcttggga cttgaaaaca	480
	ttgctgatgg aaactcaaac aagataaagg gctaatagaca ttgataggaa acataagaag	540
	aaagttattc catgacttca taagagttgt cagaatttat tatgcatgta actacagggg	600
50	a	601

55 <210> 358
 <211> 1001
 <212> ADN
 <213> Homo sapiens

60 <400> 358

65

ES 2 704 701 T3

5 gcttaatacc tgagtgatgg aatattctgt tcaacaaacc cctctgacat aggtttgcct 60
 atataataaa cctgttcatg tactcctgaa cctaaaagtt taaaaaagat tatgtagaaa 120
 10 acccaaagga atctataaaa agtctactag agctagagtg attttaacaa gatttcaata 180
 cacaaattca aatgtctttc tatatattaa tgacaatcaa caataaaatt ttaaaacatt 240
 attaaagtat aatgaaaata tcaactgttt agggagaaat gtaacaagaa tggatgaagga 300
 15 cctatacact aaaaagcttc aatatgttgt tgagattaac tgaagaaggt ctaaatagat 360
 ttttttttca tgtctcggaa gacttaatat gtgaagatac caattcttcc ccaaatagat 420
 aacaggtgaa atgcaatccc aatcaaaaac ccagcaatta ttttaagggg gaaattggca 480
 20 atctgattct aaaattcata yggaaaaaaa caatggagtt agaataacta aaacaagtcc 540
 gaaaaagaaa aagaaatgga ggactaatgc tacctgattt caagtcttat cgtataaatc 600
 tacatcaata aaggacaagt tggatattggg ttaaagatag ataaatacat cagtgaata 660
 25 gaatattgaa tccagaataa atccacacat atatggataa aaataccaga caattcagtg 720
 gagatgggtt tgtttttaca acaaatgtta ctggaacaaa ttgatatatg tattagtcag 780
 30 atatggctgc cataacaaag aaccacaaac aggtgggtta aataatggaa ataaatttcc 840
 tcagaattct ggagtatgga agcccaagat caagttgctg ggaggattcg tttcttctga 900
 gtgtctcttt ttttgatgac agatgactat cttttaccaa tgtcttcaat tggttttccc 960
 35 tctgtgtgtg cctaggtcct attctccaat tcctataagg a 1001

<210> 359
 <211> 1001
 <212> ADN
 <213> Homo sapiens

<400> 359

ES 2 704 701 T3

5	ctaaaagttt aaaaaagatt atgtagaaaa cccaaaggaa tctataaaaa gtctactaga	60
	gctagagtga ttttaacaag atttcaatac acaaattcaa atgtctttct atatattaat	120
	gacaatcaac aataaaattt taaaacatta ttaaagtata atgaaaatat caactgttta	180
10	gggagaaatg taacaagaat ggtgaaggac ctatacacta aaaagcttca atatgttggt	240
	gagattaact gaagaaggtc taaatagatt tttttttcat gtctcggaag acttaatatg	300
	tgaagatacc aattcttccc caaatgatca acagggtgaaa tgcaatccca atcaaaaatcc	360
15	cagcaattat ttttaagggg aaattggcaa tctgattcta aaattcatat ggaaaaaac	420
	aatggagtta gaataactaa aacaagtccg aaaaagaaaa agaaatggag gactaatgct	480
	acctgatttc aagtcttata rtataaatct acatcaataa aggacaagtt ggtattgggt	540
20	taaagataga taaatacatc agtggaatag aatattgaat ccagaataaa tccacacata	600
	tatggataaa aataccagac aattcagtgg agatgggttt gtttttacia caaatgttac	660
	tggaacaaat tgatatatgt attagtcaga tatggctgcc ataacaaaga accacaaaca	720
25	ggtgggttaa ataattggaa taaatttcct cagaattctg gagtatggaa gcccagatc	780
	aagtgtctgg gaggattcgt ttcttctgag tgtctctttt ttgatgaca gatgactatc	840
30	ttttaccaat gtcttcactt ggttttccct ctgtgtgtgc ctaggctcta ttctccaatt	900
	cctataagga aaccagtcac attggattag ggcccactct aatggcccca ttttacttgc	960
	attatctctt taaagacact atctccagat gtagccacac t	1001

35

<210> 360
 <211> 1058
 <212> ADN
 <213> Homo sapiens
 <400> 360

40

45	catgattagc tatgctactt tccactgctc ttagtatact gagaggcagc ataagtaaaa	60
	ctaaaatata tgaagatagc aatagactat ttaaagtaga agaagtatgc tatttttggt	120
	ttgttttcat ttccaaggaa atatgcaaag gtttattgag tatttcagct tctcttacag	180
50	taggtttttt ttggattctt tctgtgttg tctatgttga taaaacattg aaatgccata	240
	tagctcaaag gtcattcact taagaaatct aagtactgat aacatcttag ccccgattct	300
	tcataggcat tgtaagcct attataattt tggtwcagag agaaggtaaa ctatattcca	360
55	gacaggcata taaagcaatt tctcctataa ttggagttca cgaaaaattc acatatttct	420
	ttttaatagt aactctcaca gcaagaacat atgtttgtaa ataatacatc acagaatctt	480

60

65

ES 2 704 701 T3

5	attggcagac aaggaaatc ctaaaatatt ttttactgcc acatcaatta agatatataa	540
	aataccttat atagaagatg ttgacacca ggccaaacaa atcaaacaag aatagaagca	600
10	ctgacagtct tatttcaaaa ttggtttaac ttgtatttac aggatattgt agtaccttat	660
	aaagttgatt gctgattggc cgtcttttac agaattctgt cagattgtta ttatttcttg	720
	taaagattga ttcaaacaaa taaaattgt caggattgga tatgtcctat agtgaggtgt	780
15	agttatgtca catgagattt ttaattacaa agaaatggaa aataaatga gaatagaatt	840
	gagactcccc tgtcacctca caaatatgtt gaaatacaat gaaatttcca aagatgttaa	900
	agcatataaa gttgaataat tcttattatg tattaaactt acagaaattt aatttcttta	960
20	ctttataaga ggtagtgaaa atataaaatt aattatgaag acagagtagt cttagtcaga	1020
	catggcccta taaagcatat tccattcgt tacatcaa	1058
25	<210> 361 <211> 1058 <212> ADN <213> Homo sapiens	
30	<400> 361	
35		
40		
45		
50		
55		
60		
65		

ES 2 704 701 T3

5	catgattagc tatgctactt tccactgctc ttagtatact gagaggcagc ataagtaaaa	60
	ctaaaatata tgaagatagc aatagactat ttaaagtaga agaagtatgc tatttttggt	120
10	ttgttttcat ttogaaggaa atatgcaaag gtttattgag tatttcagct tctcttacag	180
	taggtttttt ttggattcct tctgtgtttg tctatgttga taaaacattg aaatgccaya	240
	tagctcaaag gtcattcact taagaaatct aagtactgat aacatcttag ccccgattct	300
15	tcataggcat tgttaagcct attataattt tggtagagag agaaggtaaa ctatattcca	360
	gacaggcata taaagcaatt tctcctataa ttggagttca cgaaaaattc acatatttct	420
	ttttaatagt aactctcaca gcaagaacat atgtttgtaa ataatacatc acagaatctt	480
20	attggcagac aaggaaattc ctaaaatatt ttttactgcc acatcaatta agatatataa	540
	aataccttat atagaagatg tttgcaccca ggccaaacaa atcaaacaag aatagaagca	600
	ctgacagtct tatttcaaaa ttggtttaac ttgtatttac aggatattgt agtaccttat	660
25	aaagttgatt gctgattggc cgtcttttac agaattctgt cagattgtta ttatttcttg	720
	taaagattga ttcaaacaaa taaaaattgt caggattgga tatgtcctat agtgagggtg	780
30	agttatgtca catgagattt ttaattacaa agaaatggaa aataaaatga gaatagaatt	840
	gagactcccc tgtcacctca caaatatgtt gaaatacaat gaaatttcca aagatgttaa	900
	agcatataaa gttgaataat tcttattatg tattaaactt acagaaattt aatttcttta	960
35	ctttataaga ggtagtgaaa atataaaatt aattatgaag acagagtagt cttagtcaga	1020
	catggcccta taaagcatat tccattcgt tacatcaa	1058

40
 <210> 362
 <211> 956
 <212> ADN
 <213> Homo sapiens

45
 <400> 362

50

55

60

65

ES 2 704 701 T3

5	aaaacaagga acaaacaac aaaaatgtta caaccgaaca acagactttt gagtcatgtt	60
	tcaggccaag aggtgatgag ttactgtagt tgcttgagct ggttggtgaa atattacctg	120
10	gcaacaaaac tgaaatagaa ggtggcttag taaaatgcag attcagaatg agtgccttaa	180
	ggttaaggca tataagacca aactgatttt ctttttcacg aggtcttcag gtaaggccat	240
	tgtagaagat accttgtttg cgaacttcag taaattactt cacttgtctc atattttcat	300
15	tttcaggatg gaggcttgag attgaattgt agtgcaatta ggtaaatttt taccattttt	360
	aaatataata ttaaaatatt aattataaat taccttattt gaatctggaa taatatttat	420
	tgcagggcac ataacttaag ctgtaaacgt cctgtyagaa gacaacatat tcatcttgct	480
20	aagggtataag ctatatgact ggcactgtgc tcaactcaga gtcattgaat gaacagtatt	540
	tatttaatct atgaatgaga gcaactcaag tatacagaaa gatatctcaa aagattcagc	600
	cttacattgc tcataacttc aatgacttag atgaaaacct cctgaacatt tttatcagtt	660
25	gtataggtac cccaaatcat aagggaatgt ttatcaatta gatgatgaaa tggggatgca	720
	actacatcat ggcaggctaa agcaatagaa tgactttgac aagaggaaat tacatagagg	780
	cacctgagtc tcctaaacca atttcaaagg tatgagaggg gggatgata aataaatagt	840
30	tgatagatga aaaaactcag aagttatagt tgacagcaat tttaatatata tatgaaaaat	900
	gtggttggac ttttagggaa aaaaacctaa taaaatctaa tggaaattag tgggtcc	956
35	<210> 363 <211> 956 <212> ADN <213> Homo sapiens	
40	<400> 363	
45	caaccgaaca acagactttt gagtcatgtt tcaggccaag aggtgatgag ttactgtagt	60
	tgcttgagct ggttggtgaa atattacctg gcaacaaaac tgaaatagaa ggtggcttag	120
	taaaatgcag attcagaatg agtgccttaa ggttaaggca tataagacca aactgatttt	180
50	ctttttcacg aggtcttcag gtaaggccat tgtagaagat accttgtttg cgaacttcag	240
	taaattactt cacttgtctc atattttcat tttcaggatg gaggcttgag attgaattgt	300
	agtgcaatta ggtaaatttt taccattttt aaatataata ttaaaatatt aattataaat	360
55	taccttattt gaatctggaa taatatttat tgcagggcac ataacttaag ctgtaaacgt	420
	cctgtcagaa gacaacatat tcatcttgct aaggtrtaag ctatatgact ggcactgtgc	480
60		
65		

ES 2 704 701 T3

5	tcaactcaga gtcattgaat gaacagtatt tatttaatct atgaatgaga gcacttcaag	540
	tatacagaaa gatatctcaa aagattcagc ottacattgc tcataacttc aatgacttag	600
10	atgaaaacct cctgaacatt tttatcagtt gtataggtag cccaaatcat aagggaatgt	660
	ttatcaatta gatgatgaaa tggggatgca actacatcat ggcaggctaa agcaatagaa	720
	tgactttgac aagaggaaat tacatagagg cacctgagtc tcctaaacca atttcaaagg	780
15	tatgagaggg gggatgata aataaatagt tgatagatga aaaaactcag aagttatagt	840
	tgacagcaat tttaatataa tatgaaaaat gtgggtggac ttttagggaa aaaaacctaa	900
	taaaatctaa tggaaattag tgggtccactc atttctccac ctaggatggt aaaaat	956
20		
	<210> 364	
	<211> 601	
	<212> ADN	
25	<213> Homo sapiens	
	<400> 364	
30	gtaaaacaca tagatcgctg tacccttggt cagtaagcta caacatactc gtatctcctg	60
	aaatcctggg cttaaatoga ggtctcaaag gctttgtttt gttttgttgt atggttgtat	120
	ggtgagtgtg tgtgtgtgtg tgtgtgtgtg tgtttattct cctgaaattc tcctcctcac	180
35	ttgacttaag ctaaaagata aacgtcctct tcctttcagc cacagatggt gatggataaa	240
	ttgaatgtca ttcacattat tcccttaaaa taaactctct cctccctc tccogtctca	300
	wccttggtccc tttctttata taatgggtaa tgcgttaatg tcagcagaat agttttgggg	360
40	ccataatggc aagtatcacg tggatgggtt agcattgttt ttagaatgct gtgaatttgg	420
	gtatatgtga gttttgggga aagttttgca actatatggt tgttaattaa atgaggacta	480
45	taaagtaata taaaattatg tttctggaac atattttgga agctataaag tcactgtat	540
	ttattatcca cagacataat gtcattgttc aggtcctgca accttcttat aatcaacata	600
	c	601
50		
	<210> 365	
	<211> 601	
55	<212> ADN	
	<213> Homo sapiens	
	<400> 365	
60		
65		

ES 2 704 701 T3

5	agtaagctac aacataactcg tatctcctga aatcctgggc ttaaatacgag gtctcaaagg	60
	ctttgttttg ttttggttga tgggtgtatg gtgagtgtgt gtgtgtgtgt gtgtgtgtgt	120
10	gtttattctc ctgaaattct cctcctcact tgacttaagc taaaagataa acgtcctctt	180
	cctttcagcc acagatgggt atggataaat tgaatgtcat tcacattatt cccttaaaat	240
	aaactctctc cctccctct cccgtctcat ccttgtccct ttctttatat aatgggtaat	300
15	kcgtaaagt cagcagaata gttttggggc cataatggca agtatcacgt ggatgggtta	360
	gcattgtttt tagaatgctg tgaatttggg tatatgtgag ttttggggaa agttttgcaa	420
20	ctatatgttt gttaattaaa tgaggactat aaagtaatat aaaattatgt ttctggaaca	480
	tattttggaa gctataaagt catctgtatt tattatccac agacataatg tcattgttca	540
	ggtcctgcaa ccttcttata atcaacatac gtggggccag ggattttatg tatcttggcc	600
25	t	601

	<210> 366
	<211> 1079
30	<212> ADN
	<213> Homo sapiens

	<400> 366
--	-----------

35

40

45

50

55

60

ES 2 704 701 T3

5

10	gaatttatgg tctgatggag aagggaatca ttaaagttct atgtagtgag atatcccaa	60
	ggggtgtatt aggcctacca cactggaat ctggatagat gaagacagag tggcaggga	120
	gtcgtattaa ggttctgttt ctgctgggag ccacaggtcc tcaggaagca acaagtactg	180
15	ggcagattga tactgtagct rggctctagc tctatacctc tagaataaag gttacaaact	240
	agcaacttga aagctaaacc tggccacag atatgtttta tttggctcct acaactgttt	300
	aaaaaatatt accaacatctt aaaactggga agttttatga aaaaaccag acttctggat	360
20	tctgttgaaa aaaaaaatca gaagatctgg caatactgag ctgacattcc tatatgacaa	420
	caattggctg gatctatgca gcttctctcc aaaaagcaaa gaatgtgttc ttgcttaaca	480
	cagtcccccac cactccctca tattctccaa tcctggacct gagcgtcatt tgctatgtat	540
25	cgccatttgc catgaagttt tactctctac agaaatataa tttttttgta gaagactatg	600
	ctttaatcaa gatcaggata atataaagtg agatctgaaa gtggaaaaaa gataaatgtc	660
	caacaatgat agactggatt aagaaaatgt ggcacatata caccgtggag tactatgcag	720
30	ccaaaaaaa cgatgagttc atgtcctttg tagggacatg gatgaagctg gaaaccacca	780
	ttctcagcaa actatcgcaa ggacaaaaaa ccaaacgcgc catgttctca ctcataggtg	840
	ggaattgaac aatgagaaca cttgggcaca ggaaggggaa catcacacac cgggccctgt	900
35	tgtggggtgg ggggaggagg gagggatagc atttgagat atacctaag ttaaatgact	960
	agtttctggg tgcagcacac catcatggca catgtataca tatgtaacta acctgcacat	1020
	tgtgcacatg taccctaaaa cttaaagtat aattttttaa aaaagatatt ttcttatct	1079

40

45 <210> 367
<211> 501
<212> ADN
<213> Homo sapiens

50 <400> 367

55	ataaattttc tcttccctca agaatttatg gtctgatgga gaagggaatc attaaagttc	60
----	---	----

60

ES 2 704 701 T3

	tatgtagtga gatatcccca aggggtgtat taggcttacc accactggaa tctggataga	120
5	tgaagacaga gtggcagggg agtcgtatta aggttctgtt tctgctggga gccacaggtc	180
	ctcaggaagc aacaagtact gggcagattg atactgtagc tgggctctag ctctatacct	240
	ctagaataaa kgttacaaac tagcaacttg aaagctaaac ctggcccaca gatatgtttt	300
10	atttggtctct tacactgttt taaaaaatat taccaacatt taaaactggg aagttttatg	360
	aaaaaaccca gacttctgga ttctgttgaa aaaaaaatc agaagatctg gcaatactga	420
	gctgacattc ctatatgaca acaattggct ggatctatgc agcttctctc caaaaagcaa	480
15	agaatgtgtt cttgcttaac a	501

<210> 368
 <211> 1001
 <212> ADN
 <213> Homo sapiens

<400> 368

	tgaagaagcc gcctggcttc ttgtttcttc tcatagcaaa atgcaatgag aaagagataa	60
30	tttgagaaaa gaaccgttta aacaaaaaga aaccaagaca taatgatttt ggaaattctc	120
	agtttattca gactgcaaaa gatattaaaa taaagaaact cagtaacagg gatagataat	180
	ctaaagaaaa agcctaggac acggctgtag taaccttctg tttttatacc tcagcaattt	240
35	gctaatgcct caaaaagatc aaaagtactc aaatataaag ggctctttga agagattaga	300
	tttcctcaat caaaccaaag agcatcgagg aagcttaagg ttactgtccc tcacatatct	360
	cagcagaagg caaaaataga agactgatta tctaagaaag atctctgaaa gagtctcata	420
40	ttatggagtg aaccctgtg gcatacatgg gagaccact tggttcttga gaattttata	480
	tcaggagaaa cactgtcagt ytgtattgaa aggaacagag aaaatacgaa attaaagaag	540
45	actattaaac ctccaaaatt ctggcaggaa agaagcttac acagctactc agttgcaaa	600
	atctgccact tttcatatac atgaaaggac tcagaggagg aagccacagg tttagaagga	660
	aaagctaaaa gcaacatcgt attagtcttg gatctaggaa cctaatttct ctagcagaat	720
50	ctagaaatgg cttgggacaa gtgattgttt ttttacctag gattttctcc ctottgaaaa	780
	caggactgtc tgtaactatt atcctatgcc tgcctacca tcatatttca gaaacaggta	840
	acttatgttt tcactttcaa agattcacia taaagagaaa ttgtacctca gaatggatta	900
55	taccagagct ttcctcatgc ataaattaaa taatttaggt tatgtgattt gaagcttttg	960
	agtgggtgag gtgacatttt ggatgctgag ttggtgccgt a	1001

<210> 369
 <211> 1001
 <212> ADN
 <213> Homo sapiens

ES 2 704 701 T3

<400> 369

5	tcttctcata gcaaaatgca atgagaaaga gataatgtga gaaaagaacc gtttaaacia	60
	aaagaaacca agacataatg attttgaaa ttctcagttt attcagactg caaaagatat	120
10	taaaataaag aaactcagta acagggatag ataattctaa gaaaaagcct aggacacggc	180
	tgtagtaacc ttctgttttt atacctcagc aatttgctaa tgcctcaaaa agatcaaaag	240
	tactcaaata taaagggctc tttgaagaga ttagatttcc tcaatcaaac caaagagcat	300
15	cgaggaagct taagggttact gtccctcaca tatctcagca gaaggcaaaa atagaagact	360
	gattatctaa gaaagatctc tgaaagagtc tcatattatg gagtgaaccc ctgtggcata	420
	catgggagac ccacttggtt cttgagaatt ttatatcagg agaaacactg tcagtctgta	480
20	ttgaaaggaa cagagaaaat rcgaaattaa agaagactat taaacctcca aaattctggc	540
	aggaaagaag cttacacagc tactcagttg caaagatctg ccacttttca tatacatgaa	600
25	aggactcaga ggaggaagcc acagggttag aaggaaaagc taaaagcaac atcgtattag	660
	tcttgatct aggaacctaa tttctctagc agaattctaga aatggcttgg gacaagtgat	720
	tgttttttta cctaggattt tctccctctt gaaaacagga ctgtctgtaa ctattatcct	780
30	atgcctgccc taccatcata tttcagaaac aggttaactta tgttttctact ttcaaagatt	840
	cacaataaag agaaattgta cctcagaatg gattatacca gagctttcct catgcataaa	900
	ttaaataatt taggttatgt gatttgaagc ttttgagtgg gtgaggtgac attttggatg	960
35	ctgagttggt gccgtagtga gtccagaatt ctgcggaact t	1001

40 <210> 370
 <211> 601
 <212> ADN
 <213> Homo sapiens

45 <400> 370

50

55

60

65

ES 2 704 701 T3

5 ctctagactc ctctgtatt ttaatttagc cactttttta gggcctacaa ttttagatct 60
ccacagggct cttgaaactt cttgaacctc atcagtaaca tgtccattag tggcatgacc 120
caagagttct agaacatcta ttcagcaagt gtgtatctgg taagtgaata ttccttctat 180
10 gtgttccctt ttgcatcaaa ctacacactg tcattcctcc tttatctcca aaagcttgaa 240
aattcctcac ttgtatctca ttctttctct cttagaaaac tgatcacctc tgatgaatta 300
raacggaatg accaagcttt gggagaggca aaagaatctc ggtgttaaag actcagagtt 360
15 taagaagcaa caaaaagatt atacagatgt gaatatgtga ccttcctcca ccagggcatg 420
ttgccttgga gtaagataat ctaagcacac acttcatagc ctgagaacaa ttttggaagt 480
ctttgcttta tggatattta cataaagcaa atatggatat ttacctaaag gctggaccaa 540
20 ggcctaattc ctctagagcc ccttgatcat gaacaccatt cctgtcatga ttcttaaggt 600
c 601

25 <210> 371
<211> 601
<212> ADN
<213> Homo sapiens

30 <400> 371

acaagctcca gccatggacg caattccttc tagaagcaaa atttatctct agactcctcc 60
35 tgtatttttaa tttagccact tttttagggc ctacaatttt agatctccac agggctcttg 120
aaacttcttg aacctcatca gtaacatgtc cattagtggc atgacccaag agttctagaa 180
catctattca gcaagtgtgt atctggtaag tgaatattcc ttctatgtgt tcccttttgc 240
40 atcaaaactac aactgtcat tctccttta tctccaaaag cttgaaaatt cctcacttgt 300
rtctcattct ttctctctta gaaaactgat cacctctgat gaattagaac ggaatgacca 360
agctttggga gaggcaaaag aatctcgggtg ttaaagactc agagttaaag aagcaacaaa 420
45 aagattatac agatgtgaat atgtgacctt cctccaccag ggcatgttgc cttggagtaa 480
gataatctaa gcacacactt catagcctga gaacaatttt ggaagtcttt gctttatgga 540
50 tatttacata aagcaaatat ggatatttac ctaaaggctg gaccaaggcc taattcctct 600
a 601

55 <210> 372
<211> 701
<212> ADN
<213> Homo sapiens

60 <400> 372

65

ES 2 704 701 T3

5 gaagatgcac tctaatgttt tttcccagaa gctctgtagg tttagctttt acctttctgg 60
 gtttgttttg ttttgttttt tgagatggag tcccactcgt gtcacccagg ctggagtaca 120
 atggtgcaat ctgggttcac tgcaacctcc acctccggg ttcaagcaat tcccctgtct 180
 ccacctctcg agtagctggg atgggaggcg cctgccacca tacctggcta attttcatat 240
 10 ttttagtaaa gatagggttt caccatgtta gccaggctgg tctcgaactc ctgacctcaa 300
 gtgatccacc cgctcagct tcccaaagtg ctgggattac aggcgtgagc cactgcgccc 360
 agccctagct ttttggctta tgattcctcc caaattaatt tctgtgaacc attaccttaa 420
 15 gatgttgaga tttaatgtcc agaattctcat ttgttcacct ttgaaaatta agaaaccctg 480
 gcacagtgtt gactggagcc wcttacotta atagaaaata aagctcacat atatccataa 540
 20 tgaaaagcag agaccagcac aacctatgct acctgacagt tttaaaatcc aaggccagga 600
 tcttctcaac tcaggcccac tcaactaact cacaacatac ttcttctttc ctcagcatct 660
 actacttgtg ctgggacctt ggtcttccca ttgttcatgt c 701

25
 <210> 373
 <211> 701
 <212> ADN
 30 <213> Homo sapiens
 <400> 373

35 agatggagtc cactcgtgt caccaggct ggagtacaat ggtgcaatct cggttcactg 60
 caacctccac ctccgggtt caagcaattc cctgtctcc acctctcgag tagctgggat 120
 gggaggcgcc tgccaccata cctggctaatt tttcatattt ttagtaaaga tagggtttca 180
 40 ccatgttagc caggctggc tcgaactcct gacctcaagt gatccaccg cctcagcttc 240
 ccaaagtgtt gggattacag gcgtgagcca ctgcgccag ccctagcttt ttggtctatg 300
 attcctccca aattaatttc tgtgaaccat taccttaaga tgttgagatt taatgtccag 360
 45 aatctcatth gttcaccttt gaaaattaag aaacctggc acagtgttga ctggagccac 420
 ttaccttaat agaaaataaa gtcacatat atccataatg aaaagcagag accagcacia 480
 50 ccatagtcac ctgacagttt waaaatccaa ggccaggatc ttctcaactc aggccactc 540
 acttactcca caacatactt cttctttcct cagcatctac tacttgtgtt gggaccttgg 600
 tcttccatt gttcatgtca ttcttttct cagagttccc attcttttct cctgaaata 660
 55 aagaaatttc aaaatatacc atgtttcatg aaaaagacaa a 701

60 <210> 374
 <211> 701
 <212> **ADN**
 <213> Homo sapiens
 <400> 374
 65

ES 2 704 701 T3

	gatttccacc ctcaggtgat ggggatggtt gaacatccaa cacctgaaac aggacagacg	60
	atattgacag tacttgtag ttgcatataa tcacagacca gtggaaacag atgaaccaca	120
5	cagggccaca gcggggtttc actggggaac agagtgaaca atcaggaggt gtgggaggca	180
	ggttagtag tttaaagagg ttgaggtgtc cccctggatc ccatgggagg atcacattgg	240
	ctcatttgaa ttatcatag gactggcagg gaactgaaat cttctactca gggataagca	300
10	gaaactgtcc ctggtttcct tgataaaaag ggttggttga taggggacct tatccatggg	360
	aggaaagtga ggagggaat ttgtggctaa gccattcaag gccctcccag tttactaga	420
	tgtcaaggca gcacacgtaa tattgggact taatttttagc cacataacta ataaatttgt	480
15	aagtatgtgc aacggctcac rcttgcttcc agaatggcac ctaaaaaaca gatttacctc	540
	tccccaaatt cagatatgga attaaatgta atgtcaggaa aattgtctaa gagttggaaa	600
20	tgggaaaaaa atgttctttt ggtggagtta tggactccag aggttatcag attctattga	660
	ataacgtact tttgattgta ttgttaacaa ttaggctatt t	701

25 <210> 375
 <211> 1001
 <212> ADN
 <213> Homo sapiens

30 <400> 375

35

40

45

50

55

60

ES 2 704 701 T3

	gcatataatc acagaccagt ggaacagat gaaccacaca gggccacagc ggggtttcac	60
5	tggggaacag agtgaacaat caggaggtgt gggaggcagg tttagtagtt taaagaggtt	120
	gaggtgtccc cctggatccc atgggaggat cacattggct catttgaatt atcatacggg	180
	ctggcagggg actgaaatct tctactcagg gataagcaga aactgtccct ggtttccttg	240
10	ataaaaaggg ttgtttgata ggggacctta tccatgggag gaaagtgagg agggaaattt	300
	gtggctaagc cattcaaggc cctcccagtt ttactagatg tcaaggcagc acacgtaata	360
	ttgggactta attttagcca cataactaat aaatttgtaa gtatgtgcaa cggtccacac	420
	ttgcttccag aatggcacct aaaaaacaga ttacctctc cccaaattca gatatggaat	480
15	taaatgtaat gtcaggaaaa ytgtetaaga gttggaaatg ggaaaaaat gttcttttgg	540
	tggagttatg gactccagag gttatcagat tctattgaat aacgtacttt tgattgtatt	600
	tgtaacaatt aggctatttg tgaactcggg aggggtagaa atcgagttgt agaaaatgga	660
20	tggtaatgca agtgattttt gaccatatca atgcaaatga attctgttgg tagaaatatt	720
	catttccaca ctgtagatga cctaaacat atgtcattac attatatattt attgccttat	780
	agactattaa ocaattttga atcatacagt agcaaattha tttcagcatt ottgtgtgta	840
25	tgtgtttata tatacacgtg catatgtatt taagatatat aattgtatat tcttcaaatt	900
	cttctttgaa caggtttgaa cctcttatta gtttcctcat taaggaattt aataagacct	960
	ttaatgcatg tttgtatttt catgagagtc attattttac c	1001

30

35

<210> 376
<211> 695
<212> ADN
<213> Homo sapiens

40

<400> 376

45

	tgctccttca ttagtgcaat ggaacagcaa atcaggatac tttcacagtt ctcttaagtg	60
	agcctagaag tggggagctg cttgttcaca aacttgaagc ctgaatatgt taatattctt	120
	tcagtggcgg gacgcgggtg ctcatgcctg taatcccaac actttgggag gccgaggtag	180
50	gcagatcaac ctgaagtcag gagttcgagg ccagcctggc caacatggtg aaacccaccc	240
	tgttggtctg tactaaaaat agaaaaatta gctgggcatg gtggcgcagc cctgtaatcc	300
	cagctactca ggaggctgtg gcagaagaat cgctgcacc tgggaggcag aggttgcttt	360
55	gagttgatat cgtgtcactg cactccagcc tgggcaacag agtgagatcc tttcagaaac	420
	ctgctgtctg tatttgata caattaaaaa aaaaaaaaag atgagacagg caggtgcgaa	480
	agaaataaaa gtcamaactg atccagttgg gaaactcaga attgacagtt acgtgtcctt	540

60

ES 2 704 701 T3

5

10 tcattttattg atatttttgag attcacaggg gtttaaactt tattttttcca agactgaata 600
 gttccccacct cccttccata tataaaattt gagtagctgg ggagatttaa aagaggctcc 660
 ccataaactc agaagttaaa agagacaagg gtccc 695

15

<210> 377
 <211> 601
 <212> ADN
 <213> Homo sapiens
 <400> 377

25

 aacccccacct gttggtctgt actaaaaata gaaaaattag ctgggcatgg tggcgcatgc 60
 ctgtaatccc agctactcag gaggctgtgg cagaagaatc gcctgcacct gggaggcaga 120
 30 ggttgctttg agttgatatc gtgtcactgc actccagcct gggcaacaga gtgagatcct 180
 ttcagaaacc tgctgtctgt atttggatac aattaaaaaa aaaaaaaga tgagacaggg 240
 aggtgcgaaa gaaataaaag tcacaactga tccagttggg aaactcagaa ttgacagtta 300
 35 sgtgtccttt cattttattga tattttgaga ttcacagggg tttaaacttt attcttccaa 360
 gactgaatag ttccccacctc ccttccatat ataaaatttg agtagctggg gagattttaa 420
 agaggctccc cataaactca gaagttaaaa gagacaaggg tcccagtaaa taaaaaatga 480
 40 ttgggggtga ggaggcagat tttctgtcct cagtgaagtt tgttggttgg ttggttggtt 540
 ggttggttaa ttggttggtt tttgagtcag ggtctcactt tgtcacccaa gctggagtgc 600
 45 a 601

45

<210> 378
 <211> 663
 <212> ADN
 <213> Homo sapiens
 <400> 378

55

60

65

ES 2 704 701 T3

5

10	tgtagcaaca ggagggatga gacccaaagg tctgaaaagc cagtatttta agaagtcttg	60
	gaaaatgttg aggttgaaaa atctaacagg agtgcttgct tcagcagcaa tttagagtag	120
	attagcatgg cctctgcgcc aggatgacat gcacattcct aaaagtgttc cgtgttttaa	180
15	aaaaaagaga gagacagaat ctaaggggat gtgtacattt gctagagcta ctataacaaa	240
	gtaccagagg cagggtcact tcaacaacag aaatttattt ctcacagttc tggaggctag	300
	acgtccaaga ttaaggtgtt gactgggttg aattcagccc ataacaggaa ataaggagtt	360
20	aaataaagca cttgcttcta ttgtttgtac ctaaacttaa cagaayacag taagtaacaa	420
	gtcattggga tgcagaaaag aaaaaagaga gtgaagggaag gagagaagggt gaaggagaa	480
	tggaagagag gaagggaggg aggaaagaaa agtttgatga atgattgcag tctaaactgg	540
25	ttcaaacaag agatcttgtt taattaagga attcatccca tctctgccta ttaggaggag	600
	gaaaaagtct aaaatagaag atggtgaaag ttggatgacc ccaggcatta aggccattca	660
30	tct	663

<210> 379
 <211> 662
 <212> ADN
 <213> Homo sapiens

<400> 379

40	ttaagaagtc ttggaaaatg tggaggttga aaaatctaac aggagtgtt gcttcagcag	60
	caatttagag tagattagca tggcctctgc gccaggatga catgcacatt cctaaaagtg	120
45	ttccgtgttt taaaaaaaag agagagacag aatctaaggg gatgtgtaca tttgctagag	180
	ctactataac aaagtaccag aggcagggtc acttcaacaa cagaaattta tttctcacag	240
	ttctggaggc tagacgtcca agattaagggt gttgactggg ttgaattcag ccataaacag	300
50	gaaataagga gttaaataaa gcacttgctt ctattgtttg tacctaaact taacagaaca	360
	cagtaagtaa caagtcattg ggatgcagaa aagaaaaaag agagtgaagg aaggagaraa	420
55	ggtgaaggga gaatggaaga gaggaaggga gggaggaaag aaaagtttga tgaatgattg	480
	cagtctaaac tggttcaaac aagagatctt gtttaattaa ggaattcatc ccatctctgc	540
	ctattaggag gaggaaaaag tctaaaatag aagatgggtga aagttggatg accccaggca	600
60	ttaaggccat tcatctttta ctgttatgct tggatcatgc aaatgtgtct ggtagctaca	660
	ag	662

65

ES 2 704 701 T3

<210> 380
<211> 615
<212> ADN
<213> Homo sapiens

5

<400> 380

10	ttccatacat tccttcacac ccattgccct taacctttca aattcctgct taaaactaat	60
	cccattttta tggctgacct caccctgtat caaaaactcc gacatccctt tacgacagag	120
	agcacaaaact agtgggtccaa aatgtcatgg gggctcttctc agagttgttt tttcaatcag	180
15	gaaatttcac ataaaaatat ggatttctga tttctctttt aaaaacagaa aaacgagcca	240
	ccagtgggag cactgcagggt atctgtgtga gaccygtact tcacaactcc tgctttccct	300
	ccataaagta gcttgcattt tccacattga ctttgcagtt ctttggatc tgtattggtt	360
20	ttaagataat ttctactata tcacatatct cctcacagta caaagatata attttctttc	420
	ccttttcttt ttaaaaaatt tgtattttta atttttgtgg gtacacagta gatatttatg	480
	gggcatatga ggtattttat aggcataata tatgtactag ggtaagtggg gtattcatca	540
25	cctcaagcat ttatcctttc tttgtgtaaa atatagcatt ttctgaacac tatgaatact	600
	taagtacaag gatca	615

30

<210> 381
<211> 994
<212> ADN
<213> Homo sapiens

35

<400> 381

40

45

50

55

60

ES 2 704 701 T3

	tcaaagtgt acaaatttcc ttctctcata aactagcaga cattctatcc cctcattatt	60
	gtaacacatt tctaatatct ttctcaaatt gtcttctgt attacaatgc actcaccttg	120
5	gcttagaatg tctgagacaa gaaaatctat tcaccattcc cacagatgac tcctcactc	180
	tcctcccaag tcttccatac attccttcca caccattgcc cttaaccttt caaattcctg	240
	cttaaaacta atcccatttt tatggctgac ctcaccctgt atcaaaaact ccgacatccc	300
10	tttacgacag agagcacaaa ctagtgtgcc aaaatgtcat gggggtcttc tcagagttgt	360
	tttttcaatc aggaatttcc acataaaaat atggatttct gatttctctt ttaaaaacag	420
	aaaaacgagc caccagtggg agcactgcag gtatctgtgt gagacctgta cttcacaact	480
15	cctgctttcc ctccataaag yagcttgcatt ttccacatt gactttgcag ttctttggta	540
	tctgtattgg ttttaagata atttctacta tatcacatat ctctcacag tacaagata	600
	tcattttctt tcccttttct ttttaaaaaa tttgtatttt taatttttgt ggggtacacag	660
20	tagatattta tggggcatat gaggtatttt ataggcatat aatatgtact agggtaagtg	720
	gggtattcat cacctcaagc atttatcctt tctttgtgta aaatatagca ttttctgaac	780
	actatgaata cttaagtaca aggatcaagt cataggattt ggaattgatt tttaaaatat	840
25	gttgaccaa gtgctcttat catcaaactt aacatcacta atgaaggatg aacatccaa	900
	atctgaaaat ccaaaatcca aaatgctcca taatctaaaa cttgttgagc accaacatga	960
	tgcctaaagg aaatgctcct ggagcatttc agat	994

30

35

<210> 382
<211> 1001
<212> ADN
<213> Homo sapiens

40

<400> 382

45

ctatgagaaa tattttttaa gtggttagga acaattcata gcactgacat gttatcagta	60
aaaatagaag aaaataaatt aatattatga aatattaatt atatttcatt aattatgtaa	120
tatgaattat gtttttagctc aaatatttcc caagggacaa ttaagtaa gaaaaataca	180
cacagattaa aataataaat agagaaggag atattaatga ggtacaaaaa gaaaaaatac	240
atgtaatcac atgaaatgct attatttgaa agattaacaa aacttgtaaa ctacctgcta	300
acttgatcaa agaaaaaaat cgagaaacca tatgcgcaat taatagtaag agggaaataa	360
acattgaaac agaagacatt tgaaatacca tataagactg ggtttcagag ctctatgtac	420
gtaaattgat aatgtcctgg agaagtgcag atgacaaaaa tggacacott tcaacttaga	480

60

ES 2 704 701 T3

5
10
15
20
25
30
35
40
45
50
55
60
65

aatcataaac agattcattt ycttaaagtt aatgaaaaga attaacagac cctcctcaaa 540
aaagacatat atgcggccta caatcatatg aaaaaaagtt caacattact gttcattaga 600
gaaatgcaaa tcaaaaccac aatgagatac catctcacac cagtcagaat ggctattatt 660
aagaagtcaa aaaataaaag atgctggcga ggttgtggag aaaaaagaat gctttttatac 720
acttgggtggg aatgtaaatt agttcagtca ttgtggaaga ctttgatgat tcctagaaga 780
cctaaatata gaactactat ttgacccaac aatcccatata ctgggtatat actcaaatga 840
ctataaatca ttctattata aagacacatg catgggatatg ttcattacag cactatgcac 900
aatagcaaa gacttgaatc aacatgaatg tccatcaatg atagactaga taaagaaaat 960
gtggtacaca tataccatgg aatactatgc agccataaaa a 1001

<210> 383
<211> 1001
<212> ADN
<213> Homo sapiens
<400> 383

tcagtaaaaa tagaagaaaa taaattaata ttatgaaata ttaattatat ttcattaatt 60
atgtaatatg aattatgttt tagctcaaat atttccaag ggacaattaa gtaaatgaaa 120
aatacacaca gattaaaata ataaatagag aaggagatat taatgaggta caaaaagaaa 180
aaatacatgt aatcacatga aatgctatta ttgaaagat taacaaaact tgtaaacctac 240
ctgctaactt gatcaaagaa aaaaatcgag aaaccatatg cgcaattaat agtaagaggg 300
aaataaacat tgaacacaga gacatttgaa ataccatata agactgggtt tcagagctct 360
atgtacgtaa attgataatg tcctggagaa gtgcagatga ccaaaatgga cacctttcaa 420
cttagaaatc ataaacagat tcatttcctt aaagttaatg aaaagaatta acagaccctc 480
ctcaaaaaag acatatatgc rgcctacaat catatgaaaa aaagttcaac attactgttc 540
attagagaaa tgcaaatcaa aaccacaatg agataccatc tcacaccagt cagaatggct 600
attattaaga agtcaaaaaa taaaagatgc tggcgagggt gtggagaaaa aagaatgctt 660
ttatacactt ggtgggaatg taaattagtt cagtcattgt ggaagacttt gatgattcct 720
agaagacctt aatacagaac tactatattga cccaacaatc ccattactgg gtatatactc 780
aatgactat aatcattctt attataaaga cacatgcatt gatatgttca ttacagcact 840
atgcacaata gcaaagactt ggaatcaaca tgaatgtcca tcaatgatag actagataaa 900
gaaaatgtgg tacacatata ccatggaata ctatgcagcc ataaaaatga aggagatcat 960
gccctttgca gggacacgaa tagagggtga ggccattatc c 1001

ES 2 704 701 T3

5	<210> 384 <211> 501 <212> ADN <213> Homo sapiens <400> 384	
10		
15	agttgcttga aagcaaagtt ctgcagtag ctctctatct agaaggaggc attttattta	60
	tgtaaggaag tcacctaaaa gaaaattcat ttgttatggt gtggctttaa gagttactta	120
20	cttttaaatgg aatccccag ataataataa attctgaaaa aaaaaaatca gaatcatggc	180
	atgttaaaac tggatacatt cctagaaata gatggaaact gctcttgcaa aaagcttagc	240
	acatgttaaa rcatttttaga aacaatttgc caaagtttat ttagtctagt gatttcgaca	300
25	ggttaaatgg accctttgag atcttttttc ctcaagtaca aaggctcact tgcttaatga	360
	acacagtccc agaaaagcag ggggctgaac cttggctcta ccatcttacc taagattcta	420
	gagttagcaa agggtttcca caagccaaa ttattatggt taatcttttc aattatctgt	480
	gaagcattag gttggtgcaa a	501
30	<210> 385 <211> 501 <212> ADN <213> Homo sapiens <400> 385	
35		
40	gaggcatttt atttatgtaa ggaagtcacc taaaagaaaa ttcatttggt atggtgtggc	60
	tttaagagtt acttactttt aatggaatcc ccagataat aataaattct gaaaaaaaaa	120
45	aatcagaatc atggcatggt aaaactggat acattcctag aaatagatgg aaactgctct	180
	tgcaaaaagc ttagcacatg ttaaagcatt ttagaaacaa tttgccaaag tttatttagt	240
	ctagtgattt ygacagggtta aatggaccct ttgagatctt ttttcctcaa gtacaaaggc	300
50	tcacttgctt aatgaacaca gtcccagaaa agcagggggc tgaaccttgg ctctaccatc	360
	ttacctaaaga ttctagagtt agcaaagggt ttccacaagc ccaaattatt atgtttaatc	420
	ttttcaatta tctgtgaagc attaggttgg tgcaaaagta actgcagggt ttgacattaa	480
	aactggcaaa aactgcaata a	501
55	<210> 386 <211> 703 <212> ADN <213> Homo sapiens <400> 386	
60		

ES 2 704 701 T3

5

10 gacaccagtt agcatattgt cgcgggggag aggggtggga aaggcgagag aacagcatgt 60
 ggtccagagg ccatacccag atggaggctg cagtcagctc ccagtc aaa ggcaaagccc 120
 aagtcaaagc catgcttccc tcttgcccac ctgctccaat gccaccaca gagagtgcgc 180
 15 cacagctcac aggatgcagg tctggttgaa tcttaacaat aactttgtaa gggaggtgtc 240
 attagctcca ttctcctggc aggaggatga ggctcaaggc agctaaaggc ttttgctgaa 300
 catcaagtgg tgagccagga ctcaawgcca gatcttcttg tttccctgtt aggtgtatgt 360
 20 agcacaactg gtatctgcag actatgctgc tggaagggt agccgtcact gttatcacag 420
 cgactgctgc ctgagatatg ccaggtagctg ctgcaagaag tttacaaata taagctcact 480
 25 tgatcttcat aacatactac ctaggtacaa tcattatatt tatttgacag atacagagac 540
 agagggggaca cagaaaggat tagtaacttg ccccaaacca cacagccagc aaggtgtaag 600
 tgagcacctg cagtctagat gagacaccac tcaaaacgtc atttttctgg cagccccgtg 660
 30 cagttaccac agtggtcacc ccagtgggtca gctaaaggcc aag 703

<210> 387
 <211> 704
 <212> ADN
 <213> Homo sapiens

35 <400> 387

40 gcatattgtc gcggggggaga ggggtgggaa agggcgagaga acagcatgtg gtccagaggc 60
 catacccaga tggagggtgc agtcagctcc ccagtc aaag gcaaagccca agtcaaagcc 120
 atgcttccct cttgcccacc tgctccaatg ccacccacag agagtgcgcc acagctcaca 180
 45 ggatgcagggt ctggttgaat cttaacaata actttgtaag ggaggtgtca ttagctccat 240
 tctcctggca ggaggatgag gctcaaggca gctaaaggct tttgctgaac atcaagtgggt 300
 50 gagccaggac tcaatgccag atcttcttgt ttccctgtta ggtgtwtgta gcacaactgg 360
 tatctgcaga ctatgctgct ggaagggtta gccgtcactg ttatcacagc gactgctgcc 420
 tgagatatgc caggtagctg tgcaagaagt ttacaaatat aagctcactt gatcttcata 480
 55 acatactacc taggtacaat cattatatatt atttgacaga tacagagaca gaggggacac 540
 agaaaggatt agtaacttgc ccaaaccac acagccagca aggtgtaagt gagcacctgc 600
 agtctagatg agacaccact caaaacgtca tttttctggc agccccgtgc agttaccaca 660
 60 gtggtcaccc cagtgggtcag ctaaaggcca agcccaccgt ttct 704

<210> 388
 <211> 975

ES 2 704 701 T3

<212> ADN
<213> Homo sapiens

<400> 388

5	gacttaagac aagggggtct taatttgatt atttttttct gttttatatg atttctatga	60
	aaactacaac aaaataaagt taattctatt taagtgactt tttaatgaat tgcctttgtt	120
	agaaaaaaaa ttaagtgttt ttgtctcact ctgtcaccga ggctggagca cagtgggtgtg	180
10	atcatggctt actgcagcca tgacctcccg ggctcagggtg atcctccac ctcagcttcc	240
	caaatagatg ggactacagt tgtgtgccac aacgcctggc taatttttgt atttttttgt	300
	agagacaggg tctcaccagg ttgccaggc tgatcttgaa ctcttggtc caagcgatcc	360
15	accacacctca gcctccctga gtgctgggat tacaggcatg agccagcgca cccagccaga	420
	attacatittt tttaaatggt actgtcctag aaaatccagg atgtgcagtg atcaygtatg	480
	aatgcatgga cctgcacaca caggagtga caaaagacc acccctgcca ggtcaccact	540
20	catactcac cccagccac gctagctcac actcctccc acacaccact gacctcatca	600
	ttgctaggta cccacttgac ttctcaacag gttcaagaca attggccttc ctgctctctt	660
	ctagaaacac cctcttttct gggctttgtg taacacctgg tctttctccc ctctctggcc	720
25	acttctcagc ttttcttttt ctttctttct tttttttttt tttttttttg ccacttcctc	780
	ttcctctaca tcaagcttgt ccaaccaca gcccaggaca gctttgaatg cagcctaaca	840
	caaattcgta agctttctta aaacattatg agatgtgtgt gtgtgtgtgt gtgtgtgtgt	900
30	gtgtgtgtgt gtgtgtgtgt gtttagctca tcagctatcg ttattgttag tgtattttat	960
	gtgtggccca agaca	975

35

40 <210> 389
<211> 976
<212> ADN
<213> Homo sapiens

45 <400> 389

50

55

60

65

ES 2 704 701 T3

		gacttttttaa tgaattgoc t t t g t t a g a a a a a a t t a a g t g t t t t t g t c t c a c t c t g t c	60
5		accacaggctg gagcacagtg gtgtgatcat ggcttactgc agccatgacc tcccgggctc	120
		aggtgatcct cccacotcag ctccccaat agatgggact acagtttgtgt gccacaacgc	180
		ctggctaatt tttgtat t t t g t a g a g a c a g g t c t c a c c a g g t t g c c c a g g c t g a t c	240
10		ttgaactcct tggctcaagc gatccacca cctcagcctc cctgagtgtc gggattacag	300
		gcatgagcca ggcgacccag ccagaattac attttttttaa atgggtactgt cctagaaaat	360
		ccaggatgtg cagtgatcac gtatgaatgc atggacctgc acacacagga gtgaacaaaa	420
15		gacccacccc tgccagg t c a c c a c t c a t a t c t c a c c c c a g c c c a c g c t a g c t c a c r c t c c	480
		tccccacaca ccactgacct catcattgct aggtacccac ttgacttctc aacaggttca	540
		agacaattgg ccttcctcgt ctcttctaga aacaccctct tttctgggct ttgtgtaaca	600
20		cctgggtcttt ctcccctctc tggccacttc tcagcttttc tttttctttc tttctttttt	660
		tttttttttt ttttgccact tcctcttcct ctacatcaag ctgtgtccaac ccacagccca	720
		ggacagcttt gaatgcagcc taacacaaat tcgtaaagctt tcttaaaaca ttatgagatg	780
25		tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgttta gctcatcagc	840
		tatcgttatt gttagtgtat tttatgtgtg gcccaagaca tttcttcttc cagtgtggcc	900
30		cagggaagcc aaaagattgg acaccctgc tctacaacat ctcaatatag gcctttttca	960
		tgtttcattc tagatt	976
35	<210> 390 <211> 801 <212> ADN <213> Homo sapiens		
40	<400> 390		
		atccagacgg tccccatact ccctgctctg tctagatggg gtccacattc cctgctccgt	60
45		ctagactgtg cccatattcg ctgctggctg caaatgagag gagttgacag cagcctcccc	120
		tttacaaggc aggaggtgcc actgttcgcc attgtctcca cctagggctt cacttgcttt	180
		ctatctgcag acatcagagg gacccacatc tctctgttct gacacgctgt gtgttgatgg	240
50		cagagtttaa ttatccacat gcaatcttac tttccttatt cccaagtccg tggggctgcc	300
		tcataaaagc attgtaagaa ctgataacca tcttctagaa gtatcatagt gatattaaga	360
		acacacatca cagatcatag taaatggctt taat t t t t t t a r c g a a a t c t c a c t a c t g c a a	420
55		atgcattgtt gtcctagcta atgaatgc atagattatgc ctgcaaaata ataattgaga	480
		ttctattttt aagaagctta gaacagtaca tgggtgcatag caaagactct gtgtatgtga	540
		agccagattt taaaaatag taacaagtgt ctgaaaatat gtggctcaat ttgtctcccg	600
60		gttacttttc cctctccccc tttaaaatgt agaggaagga gaagaagaga taagaggttt	660
		gtgagtgaag acaagggccc ttttaaggcct gggaagacta acgcatagg gatctccctc	720
		tgccttaaaa ggcacaggaa tcttagtggg gaaaaagaag tgggtgataaa tagccagtcc	780
65		gtgtgcctgg aatatcaaag t	801

ES 2 704 701 T3

5	<210> 391 <211> 801 <212> AND <213> Homo sapiens <400> 391	
10	ccctgctccg tctagactgt gcccatattc gctgctggct gcaaatgcga ggagttgaca	60
	gcagcctccc ctttacaagg caggaggtgc cactgttcgc cattgtctcc acctagggtc	120
	tcacttgctt tctatctgca gacatcagag ggaccacat ctctctgttc tgacacgctg	180
15	tgtgttgatg gcagagttta attatccaca tgcaatotta ctttccttat tcccaggtcc	240
	gtggggctgc ctcatcaaag cattgtaaga actgataacc atcttctaga agtatcatag	300
	tgatattaag aacacacatc acagatcata gtaaatggct ttaatttttt agcgaaatct	360
20	cactactgca aatgcattgt tgtcctagct aatgaatgca yagagtattg cctgcaaaat	420
	aataattgag attctatttt taagaagctt agaacagtac atggtgcata gcaaagactc	480
	tgtgtatgtg aagccagatt ttaaaatatg gtaacaagtg tctgaaaata tgtggctcaa	540
25	tttgtctccc ggtaactttt cctctctccc ctttaaaatg tagaggaagg agaagaagag	600
30	ataagaggtt tgtgagtga gacaagggcc ctttaaggcc tgggaagact aacgccatag	660
	ggatctccct ctgccttaaa aggcacagga atcttagtgg ggaaaaagaa gtggtgataa	720
	atagccagtc cgtgtgcctg gaatatcaaa gtcagtgcgt gccagggatc aactgcggg	780
35	tcacgtgcac tctgggtctc t	801
40	<210> 392 <211> 601 <212> ADN <213> Homo sapiens <400> 392	
45		
	ttggcctggg gctgattcct ccaaagcaat gtgtctcttc gcagagtctc ttagagctgc	60
	aaggcagtat gggatcatca gagaggatgc taggaagctt cagaaatgga ggtcctggta	120
50	gaaaggggcc tttggcgtgg cctctgaaga gtccaaatgt gggacaagac cctccgaaag	180
	cggtggcctg gggagccaca ggtggggcag ccagcacgga agaggggtgc tttgctacca	240
	ttgggaaaac ttatcctcca catcctcatg aggcacacac ctttcctacc ttaccgtcc	300
55	ycagtggcct ccctgttgcc ttcttattca agactaagac cctctagaat gttctttatc	360
	ctgagtccag ctgattgtct atactaatat cagtacgggg ttagatgag gacaaccagt	420
	gtgcctggct gccaggcacc cctcccccac accccaggag tttctggaac attccaactc	480
60	tgcttgaggg tatccatgca gcactacta ctgtgagcag gtggtctgat ctgtggaaa	540
	cttctatgat tcacctgagg gtaactgccc tttgtgattt gaaagaatga tgctaacaga	600
65	a	601

ES 2 704 701 T3

<210> 393
<211> 601
<212> ADN
<213> Homo sapiens

5

<400> 393

	gcagagtctc ttagagctgc aaggcagtat gggatcatca gagaggatgc taggaagctt	60
10	cagaaatgga ggtcctggta gaaaggggtcc tttggcgtgg cctctgaaga gtccaaatgt	120
	gggacaagac cctccgaaag cgggtggcctg gggagccaca ggtggggcag ccagcacgga	180
	agaggggtggc tttgctacca ttgggaaaac ttatcctcca catcctcatg aggcaaacac	240
15	ctttcctacc ttaccgctcc tcagtggcct ccctgttgcc ttcttattca agactaagac	300
	yctctagaat gttctttatc ctgagtccag ctgattgtct atactaatat cagtacgggg	360
	tgtagatgag gacaaccagt gtgcctggct gccaggcacc ccctcccaa accccaggag	420
20	tttctggaac attccaactc tgcttgaggg tatccatgca gcactacta ctgtgagcag	480
	gtggtctgat ctgtggaaaa cttctatgat tcacctgagg gtaactgcc tttgtgattt	540
25	gaaagaatga tgctaacaga aagtgttgtc atttctgaac ttttctgaac tctgcagcga	600
	g	601

30

<210> 394
<211> 1001
<212> ADN
<213> Homo sapiens

35

<400> 394

40

45

50

55

60

65

ES 2 704 701 T3

	agatttggat ggggacacaa aaccaaacca tatcataggt taaattgtgt ctcccacccc	60
5	aaaaatgtgt atgttgaagt cctaaccttc agtactcaga atgtgacatt atttggaat	120
	agggtcattg cagatggagt tagttaagat gaggtcatta ggatgagtcc ctaatccaat	180
	atgactggtg ctcttacaaa aaggggaagt ttggacacag agccatgcac atgggtggga	240
10	agaatcccaa atgaacggat aggcagaggg ttggagagat gcatcaacaa ggaacaccaa	300
	agattgccag caacccccag aagctggggg agaggcctgg aacagattct ccctcacagc	360
	ctgagaggaa ccaagctggc tgacaccttg atctcaggtt accggccttg agaactgaga	420
15	gaccctgggt ttctgttgtt taagcctctc aggggtgcagc actttattat ggaagcctga	480
	gctgactaat acaggtgtct ytatatctca ctgagggaaa gtgacaggaa agtaagaacc	540
	atztatgtcc aagagtccag aggagtcaac cagattctgg gggaaaagaa ggtacaatgc	600
20	tggcctctcc atgcagccta gtccccaaca cttgtagggc ccagggcaag atctaaagca	660
	ctctctcacc tatgcatcta tatgctgtaa ctgagataaa caaactatta aataatatat	720
25	gtgtcttgcc tctcaatctg acaattacac ctttataata gcaacatagg aaaataacta	780
	aaactatggt ttttaggcaa ccaaatacca gcaaaatgta ataattccta ttattagata	840
	tgtttaagtg ttctgctggt gggtcagcat ctttggtaga gtcataaaat taaaatgtac	900
30	ataattaatt aaatattata tgtttattcc ctaacattta tttctgtcat ttcttttttc	960
	tttttttcag acagtctcac tcttttgccc aggccggagt g	1001

35
 <210> 395
 <211> 1001
 <212> ADN
 <213> Homo sapiens
 40
 <400> 395

45	ttgtgtctcc cccccaaaa atgtgtatgt tgaagtcta accttcagta ctcagaatgt	60
	gacattatit ggaatatagg tcatcgcaga tggagttagt taagatgagg tcattaggat	120
	gagtccttaa tccaatatga ctggtgctct taaaaaagg ggaagtttgg acacagagcc	180
50	atgcacatgg gtgggaagaa tcccaaatga acggataggc agagggttgg agagatgcat	240
	caacaaggaa caccaaagat tgccagcaac cccagaagc tgggggagag gcctggaaca	300
	gattctccct cacagcctga gaggaaccaa gctggctgac accttgatct caggttaccg	360

55

60

65

ES 2 704 701 T3

	gccttgagaa ctgagagacc ctgggtttct gttgtttaag cctctcaggg tgcagcactt	420
	tattatggaa gcctgagctg actaatacag gtgtctctat atctcactga gggaaagtga	480
5	caggaaagta agaaccattt rtgtccaaga gtccagagga gtcaaccaga ttctggggga	540
	aaagaaggta caatgctggc ctctccatgc agcctagtcc ccaacacttg tagggcccag	600
10	ggcaagatct aaagcactct ctcacctatg catctatatg ctgtaactca gataaacaaa	660
	ctattaaata atatatgtgt cttgcctctc aatctgacaa ttacaccttt ataatagcaa	720
	cataggaaaa taactaaaac tatgggtttt aggcaaccaa ataccagcaa aatgtaataa	780
15	ttcctattat tagatatgtt taagtgttct gctggtgggt cagcatcttt ggtagagtca	840
	taaaattaaa atgtacataa ttaattaaat attatatgtt tattccctaa cttttatttc	900
	tgtcatttct tttttctttt tttcagacag tctcactctt ttgccaggc cggagtgcag	960
20	tggcgtgatc tcagctcact gcaacctccg cctcccaggt t	1001

25 <210> 396
 <211> 1218
 <212> ADN
 <213> Homo sapiens

30 <400> 396

35

40

45

50

55

60

65

ES 2 704 701 T3

	gataaagaaa ggtcatcctc aatttcaatt tactttatat attctttgag aggtaaccgt	60
	gtcttatctc cccccaaat tccttttaaa aggaaatttc caaagatgct ctattctgtg	120
5	aataaagcat tgtgccacag ccgagaggat ccagcaatga acatgagatt gcccttgatt	180
	cataaggtct acaagctagt aaggatagag aacacttta aataaaaaaa aatagttttt	240
10	ggtatattta tatttgttat ttggtataat tgagttttct acattctcat atatgtattt	300
	catattttga agaatatgca gaaaataatc aagcttccaa ataaacattt ttttttaaga	360
	actgcacaag tgagaattta ggagaacaga agatcagagg gctgcacrgg ctaaactaga	420
15	caatgagccc atgcaagtaa gttaagagga gaagcgggta agtatgcacc tgctttgtct	480
	aggtgaccag caagcattta gcaatagtct tttcaaaaca acagctcctt atattgtcaa	540
	atctcaagaa gtaatattta tggttaaaaa aatctcagac ccaacagaaa atccatgagg	600
20	gagatggttt tggaacgca gaattttcag ctatgatatc cttttataaa caagcagata	660
	ctttcccaa atataattca atgcctcagt ctacctctg ctgaaaccac taacaccacc	720
	actaaagctc gactatatgg gaaaatttag gtgtcacttt caaaatatgt cctagcataa	780
25	aggcaattaa aaaatgtaaa gcaccaaaga tgcaagagag acataaatga ataaaatatc	840
	tggcacgaaa gttttcaaaa gcttggaat ctgattcaaa aaaaaataaa atcagccaag	900
30	cagtgttagt aagttagcca atcaggtttc aagaaggcag aaagacaaaa tcaacatcac	960
	cagcatttga caccgctact gggggaaaaa agggggatgg agttcgttta tggccttttt	1020
	aaaaatgcca ttacttggac aagagtcata acagagaagc actgcttatt tcagttctgt	1080
35	taactgtaaa tatcagagcc aacaccaga aaaagttcac cattagccaa ttggttttgc	1140
	ctggccaatt ggagatggta ataggcctgc tatggatgac attctttctg atataagttg	1200
40	tttcttgctt tttctccc	1218

<210> 397
 <211> 1218
 <212> ADN
 <213> Homo sapiens

<400> 397

ES 2 704 701 T3

	gataaagaaa ggtcatcctc aatttcaatt tactttatat attotttgag aggtaacogt	60
5	gtcttatctc ccccaaaat tcctttttaa aggaaatttc caaagatgct ctattctgtg	120
	aataaagcat tgtgccacag ccgagaggat ccagcaatga acatgagatt gcccttgatt	180
	cataaggtct acaagctagt aaggatagag aacactttaa aataaaaaaa aatagttttt	240
10	ggatatattta tatttgtgtat ttggtataat tgagttttct acattctcat atatgtattt	300
	catattttga agaatatgca gaaaataatc aagcttccaa ataaacattt ttttttaaga	360
	actgcacaag tgagaattta ggagaacaga agatcagagg gctgcacggg ctaaactaga	420
15	caatgagccc atgcaagtaa gttaagagga gaagcgggta agtatgcacc tgctttgtct	480
	aggwgaccag caagcattta gcaatagtct tttcaaaaca acagctcctt atattgtcaa	540
	atctcaagaa gtaatattta tgggttaaaaa aatctcagac ccaacagaaa atccatgagg	600
20	gagatggttt tggaaacgca gaattttcag ctatgatatc cttttataaa caagcagata	660
	ctttcccaa atataattca atgcctcagt ctacctctg ctgaaaccac taacaccacc	720
	actaaagctc gactatatgg gaaaatttag gtgtcacttt caaaatatgt cctagcataa	780
25	aggcaattaa aaaatgtaaa gcaccaaaga tgcaagagag acataaatga ataaaatata	840
	tggcacgaaa gttttcaaaa gcttgggaat ctgattcaaa aaaaaataaa atcagccaag	900
30	cagtgttagt aagttagcca atcaggtttc aagaaggcag aaagacaaaa tcaacatcac	960
	cagcatttga caccgctact gggggaaaaa agggggatgg agttcgttta tggccttttt	1020
	aaaaatgcca ttacttggac aagagtcata acagagaagc actgcttatt tcagttctgt	1080
35	taactgtaaa tatcagagcc aacaccacaga aaaagttcac cattagccaa ttggttttgc	1140
	ctggccaatt ggagatggta ataggcctgc tatggatgac attctttctg atataagttg	1200
	tttcttgctt tttctccc	1218

40

<210> 398
 <211> 1072
 <212> ADN
 <213> Homo sapiens

45

<400> 398

50

55

60

65

ES 2 704 701 T3

5	cacttaaaag ctctggaaac ctacgagatt atctttaaaa tcgtggggac caaatggctg	60
	gccaaggact tgtttctgta caggtgcgat tgcttctctg ctgtgttcct ttttattacc	120
	caagtaaccg gtatttcagc tcacaagatg agaaaatgac aaacaggcaa aataagcgta	180
10	gggctgtgtg tgcaacagtt watcataaag ccatcaccag gagacgtcac tgggcgcctt	240
	ctggagtgta tccgtcctaa ctttgctttc tttctttttt tttttaaat taagttctag	300
	ggtacatatg cacaacgtgc aggtttgtca cacatgtata catgtgccat gttgggtgtgc	360
15	tgcacccatt aactcgtcat ttacattagg tgtatctcct agtgctatcc ctccccactc	420
	ccccgacccc atgacaggcc ccagtgtgtg atgttccccct tctgtgtgcc aagtgttctc	480
20	attgttcaat cccacctat gagtgagaac atgccatgtt tggttttttg tccttgogat	540
	agtttgctga gaatgatggg ttccagcttc atccatgtcc ctacaaagga catgaactca	600
	tcctttttta tggctacata gtattccatg gtgtatatgt gccacatttt cttaatccag	660
25	tctatcatcg atggacattt gggttgggtc caagtctttg ctattgtgac tagtgttgca	720
	ataaatatac gtgtggatgt gtctttatag cagtttgatt tataatcctt tgggtatata	780
30	cccagtaacg ggatggctgg gtcaaagtgt atttctagtt ctagatcctt gaggaatcgc	840
	cacactgact tccacaatgg ttgaactagt taacagtccc accaacagtg tgaaagtgtt	900
	cctattttctc cacatcctct ccagcacccc attttgactt tgctataagg gaactttagc	960
35	atctgaacgt gcggacagct tcattgctgg cttgttacgt aacagtgttt tgtgaccatc	1020
	tcatgtcata cccacacatc gaaaccagca gtttaaatgg ccagctgttt gc	1072

<210> 399

<211> 1072

<212> ADN

40 <213> Homo sapiens

<400> 399

45

50

55

60

ES 2 704 701 T3

5	agattatctt taaaatcgtg gggaccaaag ggctggccaa ggacttggtt ctgtacaggt	60
	gcgattgctt ctctgctgtg ttccctttta ttacccaagt aaccgggtatt tcagctcaca	120
	agatgagaaa atgacaaaca ggcaaaataa gcgtagggct gtgtgtgcaa cagtttatca	180
10	taaagccatc accaggagac rtcactgggc gccttctgga gtctatccgt cctaactttg	240
	ctttctttct tttttttttt aaatttaagt tctagggtag atatgcacaa cgtgcagggt	300
	tgtcacacat gtatacatgt gccatgttgg tgtgctgcac ccattaactc gtcatttaca	360
15	ttaggtgtat ctccatagtc tatccctccc cactcccccg accccatgac agggcccagc	420
	gtgtgatgtt ccccttcctg tgtccaagtg ttctcattgt tcaatcccca cctatgagtg	480
20	agaacatgcc atgtttgggt ttttgtcctt gcgatagttt gctgagaatg atggtttcca	540
	gcttcatcca tgtccctaca aaggacatga actcatcctt ttttatggct acatagtatt	600
25	ccatgggtgta tatgtgccac attttcttaa tccagtctat catcgatgga catttgggtt	660
	ggttccaagt ctttgctatt gtgactagtg ttgcaataaa tatacgtgtg gatgtgtctt	720
	tatagcagtt tgatttataa tcctttgggt atatacccag taacgggatg gctgggtcaa	780
30	atggtatttc tagttctaga tccttgagga atcgccacac tgacttccac aatggttgaa	840
	ctagttaaca gtcccaccaa cagtgtgaaa gtgttcctat ttctccacat cctctccagc	900
	accccatttt gactttgcta taagggaact ttagcatctg aacgtgcgga cagcttcatt	960
35	gctggcttgt tacgtaacag tgtttgtga ccatctcatg tcataccac acatcgaaac	1020
	cagcagttta aatggccagc tgtttgcttg tgaaaactcc cctcggtgg ct	1072
40	<210> 400 <211> 948 <212> ADN <213> Homo sapiens	
45	<400> 400	
50		
55		
60		
65		

ES 2 704 701 T3

5	aaatcttctt tgctgaagtg tcttttcaaa tttttgcctt ttaaaaaaat tgagttgtct	60
	taatattgag tcgtaagggt ctttatatat tctggctata tgcctttgt cagatatatg	120
	tcttgcaaat attttctccc agtctgtggc ttaccttttc cattttttaa ctgtgtttta	180
10	taaaaaaaag aagttttttt agatcaaagt ccattttaat cattttttct tttatagttc	240
	atgctttttg tgtctcattt aagaaatctt tccctactcc aatgtcacia atatattctc	300
	tgagaagctt aacagttttt gcaactaaat ttaggtctat gatccgtttt gacttaattt	360
15	ttccatatgg tgtcatgtaa cagttgagat ttttttccta tgcaggcaga tattcaatgg	420
	ttcaagtacc atttattgaa atggctatct tttctccact gaatgacctt ggcactttta	480
	tcaaacatca actggccaca yacaggtgag tctacttctg gacacttacc ctgttccatt	540
20	catctgtata tctctatcct tacaccaaca cgcatagtct tgaatactag ggcaagttaa	600
	ttttaagatg tctcctggat atgtaaaaat tatatctgag ttgaaactaca gtttatttat	660
	atatccaggc agcaaataaa tgtgagaatc tggaggtgag ggaagagatc agagatacca	720
25	ccttggaac catcaattta gagatgattc ttaaggcagg ggactaaggg aactctgtta	780
	ggacacagac atagagaagg gaaggggctg cggcctgaac accccacctg catgtctact	840
30	cacatacttt cgtcggcctg tgtaaacgaa gtgctgggtc tcccagcct ctctcatctg	900
	taagcagtgc caacaacgtc caacacagtt ccatccaatt tggatctg	948

35
 <210> 401
 <211> 920
 <212> ADN
 <213> Homo sapiens
 40
 <400> 401

45

50

55

60

65

ES 2 704 701 T3

5	aattttttgcc ttttaaaaaa attgagttgt cttaatatgg agtcgtaagg ttctttatat	60
	attctggcta tatgtccttt gtcagatata tgtcttgcaa atattttctc ccagtctgtg	120
	gcttaccttt tccattttta aactgtgttt tataaaaaaa agaagttttt ttagatcaaa	180
10	gtccatttta atcatttttt cttttatagt tcatgctttt tgtgtctcat ttaagaaatc	240
	tttccctact ccaatgtcac aaatatattc tctgagaagc ttaacagttt ttgcaactaa	300
	atthaggtct atgatccgtt ttgacttaat ttttccatat ggtgtcatgt aacagttgag	360
15	atttttttcc tatgcaggca gatattcaat ggttcaagta ccattttattg aaatggctat	420
	cttttctcca ctgaatgacc ttggcacttt tatcaaacat caactggcca cacacaggtg	480
	agtctacttc tggacactta yccgtttcca ttcatctgta tatctctatc cttacaccaa	540
20	cacgcatagt cttgaatact agggcaagtt aattttaaga tgtctcctgg atatgtaaaa	600
	attatatctg agttgaacta cagttttattt atatatccag gcagcaaata aatgtgagaa	660
	tctggagggtg agggaagaga tcagagatac caccttgga accatcaatt tagagatgat	720
25	tcttaaggca ggggactaag ggacactctg taggacacag acatagagaa ggggaagggc	780
	tgcggcctga acacccacc tgcattgtca ctacataact ttctgctggc tgtgttaacg	840
30	aagtgtctggg tctccccagc ctctctcatc tgtaagcagt gccacaacg tccaacacag	900
	ttccatccaa tttgatctg	920

35 <210> 402
 <211> 701
 <212> ADN
 <213> Homo sapiens

40 <400> 402

45

50

55

60

65

ES 2 704 701 T3

5	tgtgctgctt ccattccata ggcacctgat cctaagtgtt aaccaatccc agaactctcc	60
	ccttattttct tgctgcatgt tttgaattga tggataaac aatgtgattc gagcgcttta	120
	actcagccta tgagcctctc tattctgtga ctgctggaat aggctgcttg gccatgttct	180
10	tggaagctac caccatatca rggtaatctt ccacacaaca ttccagcccc tgctttcccc	240
	tctggcctta tctagggcca ttccccaaact cagggtgaatg cagactccaa atgtactgag	300
	ctgtgtgcag gggccagggtg caaatgcttt ctgtgcactc gcacatgctg ttctacctgg	360
15	gaagtccctt cctcctttca cctattttta ccttaaacct cagacatcat ctaccctgga	420
	aagtccttcc tgacctcacg catctaagta ggtccccccc ataatcccta tccatgcctt	480
	ctatagtact taacatgggt acctttaatt gttcatttac ttagctctct gctctccac	540
20	actgtgaact ccttacaac agggaaatgt atctctgaat gaatctttca tctccatgta	600
	acacatgctt ccaaccctac ctagcacaca atctggcata taacaggcac tcaataaacc	660
25	ttcaatgaat gccttgatca agtacaagga acataagcaa a	701

30 <210> 403
<211> 701
<212> ADN
<213> Homo sapiens

35 <400> 403

	ttaaccaatc ccagaactct ccccttattt ctgctgcat gttttgaatt gatgtgataa	60
	acaatgtgat tcgagcgtct taactcagcc tatgagcctc tctattctgt gactgctgga	120
40	ataggctgct tggccatgtt cttggaagct accaccatat cagggtaat tcccacacaa	180
	cattccagcc cctgctttcc yctctggcct tatctagggc cattccocaa ctgaggtgaa	240
	tgcagactcc aaatgtactg agctgtgtgc aggggccagg tgcaaatgct ttctgtgcat	300
45	ctgcacatgc tgttctacct gggaaagcct ttccctcttt cacctatttt taccttaaac	360
	ctcagacatc atctaccctg gaaagtcctt cctgacctca cgcactaag taggtcccc	420
	ccataatccc tatccatgcc ttctatagta cttaacatgg tgacctttaa ttgttcattt	480
50	acttagctct ctgctctccc acactgtgaa ctcccttaca acagggaaatg tcatctctga	540
	atgaatcttt catctccatg taacacatgc ctccaaccct acctagcaca caatctggca	600
55	tataacaggc actcaataaa ccttcaatga atgccttgat caagtacaag gaacataagc	660
	aaatttcctg tggaaaaaaa gaattgtatt aagtctcttg g	701

60 <210> 404
<211> 601
<212> ADN
65 <213> Homo sapiens

ES 2 704 701 T3

<400> 404

5	atgttcaactt acacatcttt ctttcaactta attgaatcct ttatTTTTgt cttagaatct	60
	tctgaatatt gaaaacagag aactatactg gaagaacata gtgtatttaag actcatggag	120
	agggagatgt gatactgtgt cactgaggtc gttccagtca taggagaaat gttaccactg	180
10	gattgaggtc tggtagatctt taaaagatga tttaattcta tgatatgtgt tcaacttgca	240
	ctaggatagt ttttactttc acctttgttc catgcaccgc gcaaatacct ggggaaccctt	300
	rttgcccaac tcaagagcca gagtcctctg tcatcatctt gcctctctcc taagtgcacag	360
15	gactgagtc agacttgggtg tttgtgggtg aggcagtgtg actgacaggc aggcctcagt	420
	ttatttagcg agtgtgagcc ctggcaggaa gattctcttt ctctgcttgc cagggtgagg	480
	aggcctcatt aagcagtttg aacttgtggt tttggcgtgt ctagtccctg tgacaggtggc	540
20	ttggtatcct cacaggcatt tctttggcct cacccttggg gtgactgttc acttgtgttt	600
	g	601

25

<210> 405

<211> 601

<212> ADN

30 <213> Homo sapiens

<400> 405

35	tcttctgaat attgaaaaca gagaactata ctggaagaac atagtgtatt aagactcatg	60
	gagagggaga tgtgatactg tgtcactgag gtcgttccag tcataggaga aatgttacca	120
	ctggattgag gtctggtaca ttttaaaaga tgatttaatt ctatgatatg tgttcaactt	180
40	gcactaggat agtTTTTact ttcacctttg ttccatgcac cgcgcaaata cctgggaacc	240
	cttggtgccc aactcaagag ccagagtcct ctgtcatcat tttgcctctc tcctaagtga	300
	saggactgag tgcagacttg gtgtttgtgg gtgaggcatg tggactgaca ggcaggcttc	360
45	agtttattta gcgagtgtga gccctggcag gaagattctc tttctctgct tgccagggtg	420
	aggaggcctc attaagcagt ttgaacttgt ggttttggcg tgtctagtcc tgggtcaggt	480
	ggcttggtat cctcacaggc atttcttttg cctcaccctt ggggtgactg ttcacttgtg	540
50	tttgagcggc tgggactcag taggttcact ggagtaggta tttctttaga gccactggcg	600
	g	601

55

<210> 406

<211> 701

<212> ADN

60 <213> Homo sapiens

<400> 406

ES 2 704 701 T3

	cagctccttg gcaagcctgc tccttcccca gcaaattgaa acaccattct gaacacctgg	60
5	gcattgtctc tgatgtccct tttcatctcc ctactctcac acaatccagc tgcctctctg	120
	ccttccacgg atattaagaa cgtccaccat ctcttgagtc caagcccttc tcaactacct	180
	ctttcttgaa ctaatttctt yctgtttttt tccagtcctc cttctctgtc atgtctctcc	240
10	tctgcacact tccattttct gggttcagaaa atgtcacctg cccagtcaca cttgccttat	300
	ggctgttggtg tcataaatac agttgacact tgaacaacat gggtttgaac tgcattggatt	360
	cacttataca catatttttt caatacaaat atatttaaaa attttgagaga tttgcaacaa	420
15	tttgaaaaaa cttgcagatg aacagcatag catagaaata ttgaaaaatt aagaaaaagg	480
	tatgtcatga atgcataaaa catatgcaga tactagtcta ttttaacctt tactgcata	540
	aaatatacac aaatctatta taaaaggta aagtttatca aagcttatgc acacaaacac	600
20	ttatagacca tatagggagc cattcagtag agagaaatgt aagcgaacgt aaagggtgtg	660
	tatttaataca caactgcata cacactgtac cactgcacta a	701

25 <210> 407
<211> 701
<212> ADN
<213> Homo sapiens

30 <400> 407

	gggcattgtc tctgatgtcc cttttcatct cctactctc acacaatcca gctgcctctc	60
35	tgccttccac ggatattaag aaggtccacc atctcctgag tccaagccct tctactcac	120
	ctctttcttg aactaatttc tttctgtttt tttccagtc ccccttctgt tcatgtctct	180
	cctctgcaca cttccatttt stgggtcaga aaatgtcacc gtcccagtc cacttgcctt	240
40	atggctgttg tgcataaat acagttgaca cttgaacaac atgggtttga actgcatgga	300
	ttcacttata cacatatttt ttcaatacaa atatatttaa aaattttgga gatttgcaac	360
	aatttgaaaa aacttgcaga tgaacagcat agcatagaaa tattgaaaaa ttaagaaaaa	420
45	ggatgtcat gaatgcataa aacatatgca gatactagtc tattttaacc tttactgcca	480
	taaaatatac acaaacttat tataaaagggt taaagtttat caaagcttat gcacacaaac	540
	acttatagac catataggga gccattcagt agagagaaat gtaagcgaac gtaaagggtg	600
50	gctatttaat cacaactgca tacacactgt accactgcac taatttcaga gccacctcct	660
	gttgtgattg tggtagagccc aagtgttggtg aggatctgct t	701

55 <210> 408
<211> 501
<212> ADN
<213> Homo sapiens

60 <400> 408

65

ES 2 704 701 T3

5	caggtagggg aagcaaatga acacaaattc aaactcggaa ttcaaaacca gcctctgtgt	60
	attcctgagg accatactgt ctgctaagtg tagagaaagg cacatcctgg ttcaacagca	120
	gagaaagcaa acaggaggca ctttctgtga gtcattctca ccacagggcc ctctcttttg	180
10	tgatccagcg atacttggtc acagtcaaag ccaggaaga gtggaaagat taacctttgt	240
	gagccaaacc rtgtgacact tgattacttg acagaactaa tccttctgtc ctgatgacag	300
	aaattcaact acacaggtag atgcaagcta atatctgttg taatgcctcc cagtttctct	360
15	ggagaattcc ttagtttctc ggacattctc gaaatgcaa gttttggcaa cgagtctctg	420
	aattaacctc tgaaaatctc acccagccaa gatggccttc ttgagaagac tgaagaacat	480
20	ggttggttcc aggtgagct g	501

<210> 409
 <211> 604
 <212> ADN
 <213> Homo sapiens

<400> 409

30	cactttctgt gattcatctc caccacaggg ccctctcttt tttgatccag cgatacttgt	60
	tcacagtcaa agcccaggaa gattggaaag attaaccttt gtgagccaaa ccgtgtgaca	120
35	cttgattact tgacagaact aatccttctg tcctgatgac agaamttaa ctacacagg	180
	acatgcaagc taatatctgt tgtaatgcct ccagtttct ctggagaatt ccttagtttc	240
	ctggacatct ctgaaatgca aagttttggc aacgagtctc tgaattaacc tctgaaaatc	300
40	tcaccagacc aagatggcct tcttgagaag actgaagaac atggttggtt tcaggctgag	360
	ctggaagtgg ttacctccc aggagaggtt cccacagtg gtgtttaagg catgggtgg	420
	accaacacca ggaagactca gacatcacac caccacctt caactcagtc acatccacct	480
45	acattttctg aaaacaaaag gcagtctccc caaaaagcac tgagactctt gtgtaggtaa	540
	tctgagcaga caccaacttc ccagggcttc cttttatcca ggagagcttg gctgttcttt	600
50	ttaa	604

<210> 410
 <211> 701
 <212> ADN
 <213> Homo sapiens

<400> 410

60

ES 2 704 701 T3

5

10

15

20

25

30

35

40

45

50

55

60

ctccttccgc catggttgta agtttctga cgctcccag tcatgcttcc cgtacagcct 60
gcagaactgc gagtcaatga aatccctttt ctccacaaat taccagctct caggtagttc 120
cttacagcag cgtgggaaca gactcaagag ctgaagcaag caaggccgtt agcaaggagc 180
gggctgggga gagcactcca ggcagagga acagccaggg ccagggcctt gagacagacg 240
tgagccagga tatctgagga acagcagaga agccagtgtg gccgcagcta aatgaggaac 300
aatgtgtgag ttccctgggg cggccaaaac aaacaccacg gacggggggc ttcaaccaca 360
gacaccgatt tcctcacagc tctggaggcg aaaagtccaa gaaaactgca cggagtatct 420
atgaggccct gatggagacc tgacctggtc cacacccatg gcctggcaag ctagatgggg 480
tgaattttca cctgccacag ycgcaagtca aagccaccgg cttctctctt ctccctccca 540
ttgctcctga cagccagggt taatatattg cctcatgtaa acagggaggc atccaccoga 600
gaatctcccc tcagcccaca taagctctgc agagagggct gtgttgctcc agttcccacc 660
tggacatgag cactttgaag ggcagcttcc ctccgggggt c 701

<210> 411
<211> 612
<212> ADN
<213> Homo sapiens

<400> 411

gggctgggga gagcactcca ggcagagga acagccaggg ccagggcctt gagacagacg 60
tgagccagga tatctgagga acagcagaga agccagtgtg gccgcagcta aatgaggaac 120
aatgtgtgag ttccctgggg cggccaaaac aaacaccacg gacggggggc ttcaaccaca 180
gacaccgatt tcctcacagc tctggaggcg aaaagtccaa gaaaactgca cggagtatct 240
atgaggccct gatggagacc tgacctggtc cacacccatg gcctggcaag ctagatgggg 300
tgaattttca cctgccacag tgcgaagtca aagccaccgg cttctctctt ctccctccca 360
ttgctcctga cagccagggt taatatattg cctcatgtaa acagggaggc aycaccoga 420
gaatctcccc tcagcccaca taagctctgc agagagggct gtgttgctcc agttcccacc 480
tggacatgag cactttgaag ggcagcttcc ctccgggggt ctggtgagc tcagggtagg 540
cgtcagtctg catggattgg atggaggaag gctgtgctg gcaggagatg aactgcct 600
tgggctgtgt gg 612

ES 2 704 701 T3

<210> 412
<211> 1001
<212> ADN
<213> Homo sapiens

5

<400> 412

	ttggggaagg aagcactggg gggaaggaag cactgggctt gggacagggc tgggcgctgc	60
10	ctcttcactg gaccatgaca aggttgttac ctcaccaagg agaggtgcaa aaagcttagg	120
	ggcttggatt tctagatttc agtgccaact atgccactta ctggctttat ccttggggaa	180
	tttatctact ctgtgaccct cagttttttt atcttaatta ttaatacata cctcataatg	240
15	tgactgtgag gattcactta ataatatatg gaaaaccata gaatagtgcc cagcatctag	300
	gaagtggcac agcccccttc agaagctagt gaaacctgca gaccactttt cagagtgata	360
	ttattatttt tttctagggt tactgagtta taattgaaaa aataaaaatg gaatatagat	420
20	gtacaacatg aagctctgat gcatatatcc attgtgaaat gatgaccaca atcaagctaa	480
	ttaatgttat ctatcacttc wcatagttca accttttttt gtggtgagag tactgaagat	540
	ctactctctt agcaattttc aaatctaaaa tacattatta ttaacacagt cactgtgccg	600
25	tacgttagct ctgaggacct tattcatttt atacctaaaa gtctgtatcc tttaaccaac	660
	ctctccta at tcccactgt catccctact gccacctctg gtaaccagcc ttctgctctg	720
	tttctgagtc caaccttctt agattccaca tatgagttag atcatgctgt gcagtgtttg	780
30	tttttctgtg tctggcttgc tttcacttag cataatgtcc tccagggtcca cccatgttgt	840
	tgcaaatggc agaatcttct tcttggtaaa gactgaataa tatccctgtg tgtgcgtgca	900
	tgtgtgtgtg tgtttgtgtg tgtgtgtgta tcacattttc ttcattocatt catccatcaa	960
35	tggacactaa gcactaaggt tgattccgta tcttggctat t	1001

40

<210> 413
<211> 2480
<212> ADN
<213> Homo sapiens

45

<400> 413

50	aacattactt ggggaaggaa gcactggggg gaaggaagca ctgggcttgg gacagggctg	60
----	---	----

55

60

65

ggcgctgcct cttcactgga ccatgacaag gttgttacct caccaaggag aggtgcaaaa 120
 agcttagggg cttggatttc tagatttcag tgccaactat gccacttact ggctttatcc 180
 ttggggaatt tatctactct gtgacctca gtttttttat ctttaattatt aatacatacc 240
 tcataatgtg actgtgagga ttcacttaat aatatatgga aaaccataga atagtgccca 300
 gcatctagga agtgccacag ccccttcag aagctagtga aacctgcaga ccacttttca 360
 gagtgatatt attatttttt totaggttta ctgagttata attgaaaaa taaaaatgga 420
 atatagatgt acaacatgaa gctctgatgc atatatccat tgtgaaatga tgaccacaat 480
 caagctaatt aatgttatct atcacttcac atagttcaac ctttttttgt ggtgagagta 540
 ctgaagatct actctcttag caattttcaa atctaaaata cattattatt aacacagtca 600
 ctgtgccrta cgttagctct gaggacctta ttcattttat acctaaaagt ctgtatcctt 660
 taaccaacct ctctaattt cccactgtca tccctactgc cacctctggt aaccagcctt 720
 ctgctctgtt tctgagtcca acctctcttag attccacata tgagtgagat catgctgtgc 780
 agtgtttgtt tttctgtgtc tggcttgctt tcaacttagca taatgtcctc cagggtccacc 840
 catgttggtg caaatggcag aatcttcttc ttgttaaaga ctgaataata tccctgtgtg 900
 tgcgtgcatg tgtgtgtgtg tttgtgtgtg tgtgtgtatc acattttctt catccattca 960
 tccatcaatg gacactaagc actaagggtg attccgtatc ttggctattg tgaataatgc 1020
 tgcaataaac atatgagtcc agatacctct tcaagatact gatttcattt cctttaaata 1080
 tatgcccaga agtgggattg ctggatcata tggtagttct atatttagta tcttgaggaa 1140
 tttccatact gtttttcata atgattgtag caatctatat tcccatcaac agtgtacaag 1200
 ggttccattt tctacatggc cttaccaacg tttgttatca cttatctttt tgataataga 1260
 tattctagca ggtgtgaggt ggtatctcat tgtggtttta atttgcatth tctgtatgat 1320
 tagtggtgta gagcatcttt tcatattccc attggtaatt cgtatatctt cctttgagaa 1380
 atattttatt agatcttttg cccattgtta gctgagttat atgtgagttg gttttggttt 1440
 gttgttggtt tttgtttttg ctattgagct gagttccttg tatatttttg atattaaatc 1500
 cttctcagct gtatgggtga cagatacatt cttgcattct gtaagttgca tctgtaggtt 1560
 gcaacagagt ctctttactc tgttgattgc ttgctttact gtgtgaaagc ttttttagct 1620
 tgatgtaatt gtgtttgtct atttttgctt ttgttgcttg tacttttagt gtcatatcca 1680
 aaaagttatt gccagacca gtgtcatccc ctatgttttc ttctagtaat tttaaagtth 1740
 caggctctat gtctatgtct ttaatccatt ttgagttaat ttttgtgtag ggtttaagat 1800
 aagaatccaa ttttattttt attttttgta tatggatatc caatttcccc aacaccattt 1860
 attgaaaatt ctatcctttc tttgttggtt attaacatca gaataatatt tttaaataca 1920

ES 2 704 701 T3

5

10

15

20

25

30

35

40

45

50

55

60

65

taaaattcag aagatgacaa aggaaaccaa ttacattgaa atgcatacag agttataatt 1980
ctgaaagagc aatatatgtg cctctttgtg aacacatcat atatcaaact gcagtgaccg 2040
ttctaacaac tattgcaatt tcaaagtcac gttgagtagg aggagtactt tgagattctg 2100
aaacaacggt cttgtgctat gaaatatcca tgattttgat tggatgatgg atcccagggtc 2160
ttgttaatgc tgctgtaatc tgggtgcttc attccatagt tgaataaaat gcttgatatac 2220
tggttgaaat tagtaaaaat aaaaacgtat ttttttccat ccaagttcat tctcagaccc 2280
tgaagagtca cttctctgga ttctgcagca aagttcccag ctggggcagc aagatttagg 2340
caattgaaaa gaacatacac cttgtttctca gtggcaaacc acatggaaag ctttaaatgt 2400
cagagaagaa ttctgccatt ttgctgactt ttttgtagtt ctcctaataa acaagtgtta 2460
agtgacaagc ttttcagagg 2480

<210> 414
<211> 601
<212> ADN
<213> Homo sapiens

<400> 414

ccccccccc ccccgcatat ctcagggtgg catttttgaa cttaactaga taacaaaaca 60
cagctaagac aagtcctttt ctccagcaaa gatggcaatg ctctaataac tctgagcata 120
ttaaagattc tccaagactc tagcctctgc tgcaaaaaca catacaaata cctactacta 180
ctgctgctgt gatgatgatg atgacagcaa tagtgagaat attttaataa tgccaggcac 240
ggtggcaact gctttccaaa tattatcata tttaatctga tcattgccct atgaggtagg 300
ragtattctg attcccattt tataaataag gaacccgagg cttagagagc atcgggtgact 360
tgttcaaggt caccacagc tgtcaagtga cagaacttcg ataaaaatcc agactccttt 420
aatggagtat ggagggagggt cagaaaacat aggaagtaag ggattgtgat tgacaatgtg 480
tccttgcaaa gggacaggtt aagagacaca agggcagctg tctgaggtgt gccattcacc 540
agcttcagga gagaagtggc aggctacctc cagctatcca gccctatcca gccaaaggaag 600
c 601

<210> 415
<211> 601

ES 2 704 701 T3

<212> ADN
<213> Homo sapiens

<400> 415

5

10

15

20

25

30

35

40

45

50

55

60

caaaacacag ctaagacaag tccttttctc cagcaaagat ggcaatgctc taataactct	60
gagcatatta aagattctcc aagactctag cctctgctgc aaaaacacat acaaatacct	120
actactactg ctgctgtgat gatgatgatg acagcaatag tgagaatatt ttaaataatgc	180
caggcacggt ggcaactgct ttccaaatat tatcatatct aatctgatca ttgccctatg	240
aggtagggag tattctgatt cccattttat aaataaggaa cccgaggctt agagagcatc	300
rgtgacttgt tcaagggtcac ccacagctgt caagtgcacg aacttcgata aaaatccaga	360
ctcctttaat ggagtatgga gggagggtcag aaaacatagg aagtaaggga ttgtgattga	420
caatgtgtcc ttgcaaaggg acagggttaag agacacaagg gcagctgtct gaggtgtgcc	480
attcaccagc ttcaggagag aagtggcagg ctacctccag ctatccagcc ctatccagcc	540
aagggaagctt gggagacatg ttagttcccg ccttcatttc catcagcaac ctcaaagcca	600
c	601

<210> 416
<211> 5823
<212> ADN
<213> Homo sapiens

<400> 416

ES 2 704 701 T3

tatttcaggc tttcttcttt ctatggataa gaaagctcct cagggtggcaa caaaggccat	60
ttcttttgaa gcaggcatgg catgtgacga aaaaaagaca tctcagaaaa gagccaagaa	120
taagactgga gagccactgt cagagaacag aaactgggct taatcaagga acatctcttg	180
ttcccagagt aggaggctgg caatattttc tcaactgaaat ttcagaattg ttatggacca	240
gtgactgctc tatgtgttca atttgttccc ttttcaaattg gaagcattta ttgcagacga	300
cctgcctctg tcccaccatt gtgtattagg tttgtagagy gtagacaact tgccttttta	360
gtttgtaggt ttctgtatca agagaagatg tgtgtggggc taacctagat tacaggatcc	420
tggacttcaa gtctgatata atgactggat gagactttga ctgtcctaga attgggatga	480
acatattttg ccggtgggag ggcgtgagta attgcggtta gagggcagac tgtccctcac	540
acctattcct tttcatgggt ccttcccaaa ctgcctctgg aggtggccac acaaattggt	600
ttggccattg tgaccatggg aaacttgatg cagaggctgg aaaaagcact tgcattgttc	660
tgtctcctct cttgttcctc tacaatcaca agaaatgtct aggcaggtct gagcaggccc	720
aggtcatct gccatggaag aagaatggca catggaagag ggtcacattg tccaaccaa	780
gacgatccta gaccagccag gcccagttc atggttcaag acacatgaac atagttgcac	840
gaaccaagat tagttgtgta tggccagac tagcagcagc acccatccaa cctacagact	900
ctgagaaata aatactagtt gtcttaagct tccaagtttc agtgtgagca ttaggtagta	960
acagttaatg aataagacag ataactattt tatctgtctg gatacttata caatgatttc	1020
tattttttat tgatacataa tattttacat attgctgggg tacatgtgac attttctac	1080
atacatagaa tgtgtaatga tccagtcagg atatctgagg tgtccatcac tttgagaatt	1140
tctcacttct gtgtgttggg aacaattcaa gtcgtctctt ctagtatttt taaaatatac	1200
aatacattgt taactgtagt cttttttatt gaatgacagg acttgtacct tttatctaac	1260

ES 2 704 701 T3

tgtatgtttg tatctattaa gctagttctc tttatccctg cccctccta cccactcaact	1320
cttcccaacc tctaacatgt atcatcctat tctatatctc catgagatca acttctttag	1380
ctccacata tgagcaaaaa catatgatgt ttgtctttct gtgccgggtt tatttcactt	1440
atgacctoca tttccatoca tgttactata aatgacagga tttcattctt tttgtggcca	1500
aacagtattt cattgtgtat atatactaca ttttctttat ccattoatcc attgatgaac	1560
acttacgttg attccatato tttgctattg tgaatggtgc tgcaataaac atgcacgtgc	1620
agttatccct ttgatacaact gatttatttt cctttggata aatacccagt agtgagattg	1680
ctggatcata cggtagttct acttttagtt tttgagacat ttccatactt ttccagtgtt	1740
tgtattaatt tacattccca tcaacaatgt ataagatttc cctttcctcc acatcctcac	1800
cagcatctgc tattttttgt ctttttaata atagtcattc taactgggggt gagaggatat	1860
ctcgctatgg ttttgatttg catttccctg atatttaaatg atattgagca tttcttcata	1920
taacctattg gccatttggt tgtctttttt tttttttttt ttttttttga gaattgtcta	1980
ctcatttttg gctttttaaa agatttattt tttgttgttg ttgagtttag tgcataatcct	2040
ggatattagt ctcttatctg atgaagagtt tgccaatatt ttctccatt caacaggttg	2100
tctcttcatt ctgttgactg tttcctttgc tgtgcagaag cactttatat acagtcccat	2160
ttgtctattt ttttagtgct tatgcattta aggtctcagc cacaaaatct ttgcctagac	2220
cagtcctaaa gtgtttcccc tatattttct tctagtagtt ttattgtttc atgtcttata	2280
tttaagtcta taatccattg tgagttgatt tttgtatatg gtgagatagg ggccttgttt	2340
cattcctctg catatagata ttttaatttc tcagcaccat ttattgaagg tgtccttccc	2400
tattgtatgt tcttggtgcc tttgtcaaaa ttcagttggc tataaatatg tgaatttatt	2460
tctgggttct ctatgtggtt ccattagtct atgtgtctat ttttatacca atatcatgct	2520
gttttgatta ccatagcctt gtaatatatt ttgaagtcag gtagtgtgat gcctccagct	2580
ttgttctttt tgctcaggat tgctgtgcat actctggctt tttggttaca tacaaatttc	2640
aggatttttg tatttctgtg aaaaatggca ttagtatttt gataggaatt gcactggatc	2700
tgtatattgt cctggacaac atggtcattt taacaatatt aattcttcta atctatgagt	2760
atgagacgtc ttcccacttg tttgtgtcct cttcaatttc tttcattggt gtttcataat	2820
ctcccttcta caggcctttc acctccttgg ttaaattaat tcctaggtat tttttttag	2880
ctactgtaaa tgggactgcc ttctttctca gctagttcat ttttggtgca tagaaaccct	2940
atttttgtat gttcattttc tatcctgcaa cattacaaa tttgcttata agctttaagt	3000
gtgtattttg ctttgcttgt agagtcttct gggttctcta aatgtaagac gatgtcatct	3060
gcaaacgggg acaatttgac ttctcttaa aaatctgtat gccttttatt cctttctctt	3120

gectgattgc tctggctcta cctccagtac tatactgaat aaaagtggta aaagtgagca 3180
tccttccttg tcttgcctcta gttcttagag gaaatacttt cagtttttcc ccactcagta 3240
tgatgttagc tgtgggtcat atatagcctt tattatgtta agatatgttt cttctgtacc 3300
tggtttggtg acagcttttt atcataaaaag gatgtagaat tttatcaa atgtttttctg 3360
catctgttga gataatcata tggtttttgt cattccttct actggttga tgtatcatgt 3420
ttattgattt gtgtatgtta aaccatcctt gtgtccttgg tataaattat acttggtcat 3480
ggtgtattat ctttttggca tcctgtcgaa ttgtttgcta gctttttgtt ttgttctttt 3540
tgagaatttt tatgtctagg ttccttagaa aactggcct gtagttctct ttttgtgtgt 3600
gtgtccttgt ctagtttggg gtcagggaaa tgggtggtct gtagaatgag ttgttttttc 3660
tttgattttt ttgcaagagt ttgaggagaa tgggtattag ttcttcttta tgtggttggg 3720
caaattggca gtgaattcat tcagtcatga gcttttcttt ttttgggagg gttctcatta 3780
ctgagttaat cacactgctc attactgac tggtcagatt ttctatttct tctggaatct 3840
cagtagttgt atgtttccag caatttatcc atttcctcta ggttttctag tttggtagta 3900
tatagctatt cataatagtc tctgatgac ttttgtattt ctgtgatatc agttgtaatg 3960
tctttttcat ttcctatttt atttggtct tttcttgttt agtctagcaa ggggtttatc 4020
tattttatct ttttgaagaa ccaacttttt gtttcattga ccctttctac gtctttagtc 4080
tttatttcat ttagatttgc tctgaacttt actatgtctt tccttcta at tttgggtttg 4140
gtttgttctt tctagtttcc ttgaggtgca tcattgaatt gtttctttga tatctatcta 4200
ctcatttgat gtaggtgttt attgctatac actctccct cctagagctc cttttgttgt 4260
gtcccatagg tcttggtatg ttgtttctat tttcatttgt ttcaaacatt ttatttccat 4320
attaattttt atcattcagg aggagcatat tatttaattc ccatgtattt gtatagtttc 4380
caaagttcct cttatttcta tttttactcc atttggtctt gagaagatac ttcatatgat 4440
ttcaattttt aaaaatttgt caagacttgt ttttgtcct aacatatggt ctatcctgga 4500
gaatgttcca tgtctgatg agaaaaatgt gtactcagca gttgttgagt aacatgttct 4560
acaaatatct gttagatcca tttggtctaa agtctagttt aaatccaatg agtttttgtt 4620
aattttgtct agatgatcat gatctgagac tgaggtgaag tccccaacaa ttatcgtgtt 4680
ggagtctacc tctcttttta aatctagaaa tatttgcttt ataaatccgg gtggtctagt 4740
gttgggtgca tatatttagt tgttatttcc tcattagatt gatctcttta ctattatata 4800
ataactgttt actgcttctg gcataaagtc tgttttatgt aagtacagcc attcctgctt 4860
gagtttagta ccatgttgac aaagggatgc atagagagtt ggtaaagcat gatttctggg 4920
tgtctgtgtg aaggtgtttt gagaagagtt tagcatgagt ctgtggagtg agtggaaga 4980
ttctcctca atgtcagcag gaaccatcca tccactgggg gccaggtag aaaaagatga 5040

ES 2 704 701 T3

5

10	agaaatggtg aattctctct ctctcctgga gctgggtcac ccttcttctg cccttgaaca	5100
	ggacatcaca actccaggct ctccagcctt tggactccaa gactgacacc agtgcccttc	5160
	cccaattacc ccaggccctc aggcctttgg cctaggattg agacttacac catcagcttc	5220
15	cctggttctg aggcttctgg acttgcactg ggccatacta ccagcatccc agggctctcca	5280
	gcttgcaag agcctgttgt gggacttttc agcctccata atcaagtaag ccaatttccc	5340
	tggatatctat atagatatac aatcatgttt tgcttaccag cctgaaaaat gtatcgctag	5400
20	atgagtctgt cattgcataa acatcatagt gtacttacac aaacctagat tctatagcct	5460
	actacacacc tagctataa acatgtacag catgttactg tactgaatat tgtaggcaac	5520
	tgtaacacaa tggatgaatat ttgcatattt aaacatatct tatcattaaa aagatacagt	5580
25	aaacataagg tataaaagac aaaaaccggc acacctatat agggcactta ccataaatgc	5640
	agcttgcaag actagaagtc actcagggcg agtcagtgag cgaacgtgaa ggccataggtt	5700
	attactgtcc actacggtag actttatcaa cactgtacac aggctacact aaatttatct	5760
30	tttaaaaatt tgctctccaa taataaatta atcttcgcat cctttttttg ttgttctactg	5820
	tgg	5823

35

<210> 417
 <211> 5823
 <212> ADN
 <213> Homo sapiens

 <400> 417

40

45

50

55

60

65

ES 2 704 701 T3

tatttcaggc tttcttcttt ctatggataa gaaagctcct caggtaggcaa caaaggccat	60
ttcttttgaa gcaggcatgg catgtgacga aaaaaagaca tctcagaaaa gagccaagaa	120
taagactgga gagccactgt cagagaacag aaactgggct taatcaagga acatctcttg	180
ttcccagagt aggaggctgg caatattttc tcaactgaaat ttcagaattg ttatggacca	240
gtgactgctc tatgtgttca atttgttccc ttttcaaagt gaagcattta ttgcagacga	300
cctgcctctg tcccaccatt gtgtattagg tttgtagagt gtagacaact tgccttttta	360
gtttgtaggt ttctgtatca agagaagatg tgtgtrggcc taacctagat tacaggatcc	420
tggacttcaa gtctgatata atgactggat gagactttga ctgtcctaga attgggatga	480
acatattttg ccggtgggag ggcgtgagta attgcggtta gagggcagac tgtccctcac	540
acctattcct tttcatggtg ccttcccaaa ctgcctctgg aggtggccac acaaattggct	600
ttggccattg tgaccatggg aaacttgatg cagaggctgg aaaaagcact tgcattgttc	660
tgtctcctct cttgttcctc tacaatcaca agaaatgtct aggcaggctc gagcaggccc	720
aggctcatct gccatggaag aagaatggca catggaagag ggtcacattg tccaaccaa	780
gacgatccta gaccagccag gccccagttc atggttcaag acacatgaac atagttgcac	840

gaaccaagat tagttgtgta tggcccagac tagcagcagc acccatccaa cctacagact 900
 ctgagaaata aatactagtt gtcttaagct tccaagtttc agtgtgagca ttaggtagta 960
 acagttaatg aataagacag ataatcattt tatctgtctg gatacttata caatgatttc 1020
 tattttttat tgatacataa tattttacat attgctgggg tacatgtgac attttgctac 1080
 atacatagaa tgtgtaatga tccagtcagg atatctgagg tgtccatcac ttgagaatt 1140
 tctcacttct gtgtgttggg aacaattcaa gtctgtctct ctagttattt taaaatatac 1200
 aatacattgt taactgtagt cttttttatt gaatgacagg acttgtacct tttatctaac 1260
 tgtatgtttg tatctattaa gctagtcttc tttatccctg cccctcccta cccactcact 1320
 cttcccaacc tctaacatgt atcatcctat tctatatctc catgagatca acttctttag 1380
 ctcccacata tgagcaaaaa catatgatgt ttgtctttct gtgcccggtt tatttcaactt 1440
 atgacotcoa tttccatcoa tgttactata aatgacagga tttcattctt ttgtgtggcca 1500
 aacagtattt cattgtgtat atatactaca ttttctttat ccattcatcc attgatgaac 1560
 acttacgttg attccatctc ttgctattg tgaatgggtg tgcaataaac atgcacgtgc 1620
 agttatccct ttgatacact gatttatttt cctttggata aatacccagt agtgagattg 1680
 ctggatcata cggtagttct acttttagtt tttagacat ttccatactt ttccagtgtt 1740
 tgtattaatt tacattccca tcaacaatgt ataagatttc cctttcctcc acatcctcac 1800
 cagcatctgc tattttttgt ctttttaata atagtcattc taactggggt gagaggatat 1860
 ctcgctatgg ttttgatttg catttccctg atatttaatg atattgagca tttcttcata 1920
 taacctattg gccatttggt tgtctttttt tttttttttt ttttttttga gaattgtcta 1980
 ctcatTTTTG gctTTTTaaa agatttattt ttgtttgttg ttgagtttag tgcatatcct 2040
 ggatattagt ctcttatctg atgaagagtt tgccaatatt ttctcccat caacagggtg 2100
 tctcttcatt ctgttgactg ttctctttgc tgtgcagaag cactttatat acagtccat 2160
 ttgtctattt tttagtagtc tatgcattta aggtctcagc cacaaaatct ttgcctagac 2220
 cagtcctaaa gtgtttcccc tatattttct tctagtagtt ttattgtttc atgtcttata 2280
 ttttaagtcta taatccattg tgagttgatt ttgttatatg gtgagatagg ggccttgttt 2340
 cattcctctg catatagata tttaattttc tcagcaccat ttattgaagg tgccttccc 2400
 tattgtatgt tottggtgoc ttgtcaaaa ttcagttggc tataaatatg tgaatttatt 2460
 tctgggttct ctatgtggtt ccattagtct atgtgtctat ttttatacca atatcatgct 2520
 gttttgatta ccatagcctt gtaatatatt ttgaagtcag gtagtgtgat gcctccagct 2580
 ttgttctttt tgetcaggat tgetgtgcat actctggctt ttgtgttaca taaaaatttc 2640
 aggatttttg tatttctgtg aaaaatggca ttagtatttt gataggaatt gcactggatc 2700

ES 2 704 701 T3

tgtatattgt cctggacaac atggtcattt taacaatatt aattccttcta atctatgagt	2760
atgagacgtc ttcccacttg tttgtgtcct cttcaatttc tttcattggg gtttcataat	2820
ctcccttcta caggcctttc acctccttgg ttaaattaat tcctaggtat ttttttgtag	2880
ctactgtaaa tgggactgcc ttctttctca gctagttcat ttttggtgca tagaaaccct	2940
atttttgtat gttoattttc tctcctgcaa cattaccaa tttgcttato agctttaagt	3000
gtgtattttg ctttgcttgt agagtcttct gggttctcta aatgtaagac gatgtcatct	3060
gcaaacgggg acaatttgac ttctctctaa aaatctgtat gccttttatt cctttctctt	3120
gcctgattgc tctggctcta cctccagtac tatactgaat aaaagtggta aaagtgagca	3180
tccttccttg tcttgctcta gttcttagag gaaatacttt cagtttttcc ccactcagta	3240
tgatgttagc tgtgggtcat atatagcctt tattatgtta agatatgttt cttctgtacc	3300
tggtttgttg acagcttttt atcataaaag gatgtagaat tttatcaa atgtttttctg	3360
catctgttga gataatcata tgggttttgt cattccttct actgttgtga tgtatcatgt	3420
ttattgattt gtgtatgtta aaccatcctt gtgtccttgg tataaattat acttgggtcat	3480
gggtgtattat ctttttggca tcctgtcgaa ttgtttgcta gctttttgtt ttgttctttt	3540
tgagaatttt tatgtctagg ttcccttagaa acactggcct gtagttctct ttttgtgtgt	3600
gtgtccttgt ctagtttggg gtcagggaat tgggtgtcct gtagaatgag ttgttttttc	3660
tttgattttt ttgcaagagt ttgaggagaa tgggtattag ttcttcttta tgtggttggg	3720
caaattggca gtgaattcat tcagtcata gcttttcttt ttttgggagg gttctcatta	3780
ctgagttaat cacactgtc attactgac tgttcagatt ttctatttct tctggaatct	3840
cagtagttgt atgtttccag caatttatcc atttcctcta ggttttctag tttggtagta	3900
tatagctatt cataatagtc tctgatgac ttttgtattt ctgtgatatc agttgtaatg	3960
tccttttcat ttctattttt atttgggtct tttcttgttt agtctagcaa ggggtttatc	4020
tattttatct ttttgaagaa ccaacttttt gtttcattga ccctttctac gtcttttagtc	4080
tttatttcat ttagatttgc tctgaacttt actatgtcct tccttcta at tttgggtttg	4140
gtttgttctt ttctagttcc ttgaggtgca tcattgaatt gtttctttga tatctatcta	4200
ctcatttgat gtaggtgttt attgtatata actctccct cctagagctc cttttgttgt	4260
gtcccatagg tcttggtagt ttgtttctat tttcatttgt ttcaaacatt ttatttccat	4320
attaattttt atcattcagg aggagcata tatttaattc ccatgtattt gtatagtttc	4380
caaagtccct cttatttcta tttttactcc attgtgtgtc gagaagatac ttcatatgat	4440
ttcaattttt aaaaatttgt caagacttgt tttttgtcct aacatatggg ctatcctgga	4500
gaatgttcca tgtgctgatg agaaaaatgt gtactcagca gttgttgagt aacatgttct	4560
acaaatatct gttagatcca tttggtctaa agtctagttt aaatccaatg agtttttgtt	4620

ES 2 704 701 T3

5

10	aattttgtct agatgatcat gatctgagac tgaggtgaag tccccaacaa ttatcgtgtt	4680
	ggagtctacc tctcttttta aatctagaaa tatttgcttt ataatccgg gtggtctagt	4740
	gttgggtgca tatatttagt tgttatttcc tcattagatt gatctcttta ctattatata	4800
15	ataactgttt actgcttctg gcataaagtc tgttttatgt aagtacagcc attcctgctt	4860
	gagtttagta ccatgttgac aaagggatgc atagagagtt ggtaaagcat gatttctggg	4920
	tgtctgtgtg aaggtgtttt gagaagagtt tagcatgagt ctgtggagtg agtgggaaga	4980
20	ttctccctca atgtcagcag gaaccatcca tccactgggg gccaggtag aaaaagatga	5040
	agaaatggtg aattctctct ctctcctgga gctgggtcac ccttcttctg ccctgaaca	5100
25	ggacatcaca actccaggct ctccagcctt tggactccaa gactgacacc agtgcccctc	5160
	cccaattacc ccaggccctc aggccttttg cctaggattg agacttacac catcagcttc	5220
	cctggttctg aggccttctg acttgcactg ggcatacta ccagcatccc agggtctcca	5280
30	gcttgccagag agcctgttgt gggacttttc agcctccata atcaagtaag ccaatttccc	5340
	tggatatctat atagatatac aatcatgttt tgcttaccag cctgaaaaat gtatcgctag	5400
	atgagtctgt cattgcataa acatcatagt gtacttacac aaacctagat tctatagcct	5460
35	actacacacc tagtctataa acatgtacag catgttactg tactgaatat tgtaggcaac	5520
	tgtaacacaa tggatgaatat ttgcataatt aaacatatct tatcattaaa aagatacagt	5580
	aaacataagg tataaaagac aaaaaccggc acacctatat agggcactta ccataaatgc	5640
40	agcttgccagg actagaagtc actcagggtg agtcagtgag cgaacgtgaa ggcctagggt	5700
	attactgtcc actacggtag actttatcaa cactgtacac aggctacact aaatttattt	5760
	tttaaaaatt tgctctocaa taataaatta atcttgcgat cctttttttg ttgttccactg	5820
45	tgg	5823

50 <210> 418
 <211> 707
 <212> ADN
 <213> Homo sapiens

55 <400> 418

60

65

ES 2 704 701 T3

5

10	aacgggtgtca gctggagtga actcctgtgt gtgcaaggcc tgggtctcct ggtcagacta	60
	ctttctatgg gaaaggcata gtgtatagtc tatatactat acataggggt gctgggagga	120
	actgggggttt tcacagccag ctttggtttt cattaggttt gtttagtttc cattgcttca	180
15	ggggtgttag ttttgtgttc mcaactagat tataaactcc tcttgcatto ctgatggcag	240
	tgacttgaag gcattttattt gaagaataat agacatacag aaaggggccc atgtcataaa	300
	ggtacagctg gacgactttt cacaagtga gcacatttgt atgatcgatg ttgagaccaa	360
20	gagcattcag tggacaactc ctttccagtt actccacccc actcccagtg accatcattc	420
	tgacttctaa ctgtgtagac atgttttgtc tgttttgtac tttaaaaca tatctactct	480
25	attttaggtg gctagacaat gtgttttaca atgctggcca tgacagtgtt tgaaagaata	540
	aaatggaatc aaatagaatg ggcagtatca gagtgtgttg cctgcctaag aaatgttttg	600
	tgacattttg gctttgggtc tatttacaca ttaaatctaa gagcaccaga atgtggtgtc	660
30	aaaatgtgtt tggggatgaa gatattctaa agtcctgtag taagcaa	707

35 <210> 419
<211> 712
<212> ADN
<213> Homo sapiens

40 <400> 419

45	cagactactt tctatgggaa aggcatagtg tatagtctat atactataca taggggtgct	60
	gggaggaact ggggttttca cagccagctt tgggtttcat taggtttgtt tagtttccat	120
	tgcttcaggg gtgttagttt tgtgttccca actagattat aaactcctct tgcattcctg	180
50	atggcagtga cttgaaggca tttatttgaa gaataataga catacagaaa ggggcrctatg	240
	tcataaagggt acagctggac gacttttcac aaagttagca catttgtatg atcgatgttg	300
	agaccaagag cattcagtgg acaactcctt tccagttact ccaccccact cccagtgacc	360
55	atcattctga cttctaactg tgtagacatg ttttgcttgt tttgtacttt acaaacatat	420
	ctactctatt ttaggtggct agacaatgtg ttttacaatg ctggccatga cagtgtttga	480
	aagaataaaa tggaatcaaa tagaatgggc agtatcagag tgtgttgcct gcctaagaaa	540
60	tgttttgtga cattttggct ttgggtctat ttacacatta aatctaagag caccagaatg	600
	tgggtgtcaaa atgtgttttg ggatgaagat attctaaagt cctgtagtaa gcaatgcaaa	660
	acgttctgga ggtgtttatt aaacatttgt ttgtagaatg gagaggaaga ca	712

65

ES 2 704 701 T3

	<210> 420	
	<211> 1210	
	<212> ADN	
5	<213> Homo sapiens	
	<400> 420	
10	aaagcagcac tgctctgcat tcagccttgc tacgtctcct tcagatgggc gcactagata	60
	ctgagtgatg atcatgcctt gtctaggatc tcaccaagac agttcatgaa agagacagtg	120
	cagctcatgg aggagatggg gcagctcaca gagaggatgg tgccatcatg gaaagcatgg	180
15	ggcagtcatg gagatgacgg rgtagctcat ggagaatata atgccatcat ggaaggcata	240
	gtgcagtcac ggagatgatg gtgcagctca tggagaagat ggtgccatca tgggaaggcat	300
	ggtgcaatca tggagtagac agtgcagctg ggccaagatt ctccctgact aagctcttct	360
20	caggcacctc tgagccgtcg tottaactag gcctccagct tggcttgtga aaactgcaga	420
	ctctcagcac aaatgatttg cctcctacat taagagactt aaataaacac ttgcatggct	480
25	gtgtttatth aaacagctca aggctgtgtc cctgggatga caatgactcc agcccctaaa	540
	attcctgctt gtgaaagctc attgctgaca gaaggatcta ccatttgctt cagccaacac	600
	ctgggtggcag gcagataggc cctgagcccc atttaagagc agttccttta gaaagcttgc	660
30	aattgtaaat cttttctctg ccattttgag atgtaaatct tctaccacct agaactgtct	720
	tctcaaggac ctgtgagctg actcactgaa atgcaaacat tcagggagat aactccactc	780
	ctgtcccat gcgacggcga ggccctgact ttgggtggca ccttgcctct atttgacca	840
35	ccacctcctg tcctaaagac atgagacggt tgtctctcct ctggataagt gcctattaac	900
	caacccaggt gtctcgtgca catgaaccag tccagcctag cacctggcac tgcctttccc	960
	tcagcacact ccagtctgta aaagtctcct tatggttgtt ttggcaaagt tgagcttagt	1020
40	taatgctaga ccccttctct actgcaatag ttactgctga ataaagtcta tccttaccac	1080
	tttaactagt gttgggcttt gtttctcttt cataagctca tggagaagac aatgcagttc	1140
45	catcaagttt ctggctctta cactgctaac agtcagctct ggggtccctg agagggacag	1200
	actcacacca	1210
50	210> 421	
	<211> 1194	
	<212> ADN	
	<213> Homo sapiens	
55	<400> 421	
60		
65		

ES 2 704 701 T3

5

10	gcattcagcc ttgctacgtc tccttcagat gggcgcacta gatactgagt gatgatcatg	60
	ccttgctctag gatctcacca agacagttca tgaaagagac agtgcagctc atggaggaga	120
	tggtgcagct cacagagagg atggtgccat catggaaagc atggggcagt catggagatg	180
15	acggagtagc tcatggagaa kataatgcc a tcatggaagg catagtgcag tcatggagat	240
	gatggtgcag ctcatggaga agatggtgcc atcatggaag gcatggtgca atcatggagt	300
	agacagtgc a gctgggcca gattctccct gactaagctc ttctcaggca cctctgagcc	360
20	gtcgtcttaa ctaggcctcc agcttggtt gtgaaaactg cagactctca gcacaaatga	420
	tttgctcct acattaagag acttaaataa acacttgc at ggctgtgttt atttaaacag	480
25	ctcaaggctg tgtccctggg atgacaatga ctccagcccc taaaattcct gcttgtgaaa	540
	gctcattgct gacagaagga tctaccattt gttccagcca acacctggtg gcaggcagat	600
	agggcctgag ccccatTTaa gagcagttcc tttagaaagc ttgcaattgt aaatcttttc	660
30	tctgcccatt tgagatgtaa atcttctacc acctagaact gtcttctcaa ggacctgtga	720
	gctgactcac tgaaatgcaa acattcaggg agataactcc actcctgtcc ccattgcgacg	780
	gcgaggccct gactttggtg ggcaccttgc tcttatttgc accaccacct cctgtcctaa	840
35	agacatgaga cgtttgtctc tcctctggat aagtgcctat taaccaaccc aggtgtcctg	900
	gtcacatgaa ccagtccagc ctagcacctg gactgcctt tccctcagca cactccagtc	960
40	tgtaaaagtc tccttatggt tgttttggca aagttgagct tagttaatgc tagaccctt	1020
	ctctactgca atagttactg ctgaataaag tctatcotta ccactttaac tagtggtggg	1080
	ctttgtttct ctttcataag ctcatggaga agacaatgca gttccatcaa gtttctggct	1140
45	cttacactgc taacagtcag ctctggggtc cctgagaggg acagactcac acca	1194

50 <210> 422
<211> 1194
<212> ADN
<213> Homo sapiens

55 <400> 422

60

65

ES 2 704 701 T3

5

10	gcattcagcc ttgctacgtc tccttcagat gggcgcacta gatactgagt gatgatcatg	60
	ccttgtctag gatctcacca agacagttca tgaaagagac agtgcagctc atggaggaga	120
	tggtgcagct cacagagagg atggtgccat catggaaagc atggggcagt catggagatg	180
15	acggagtagc tcatggagaa kataatgcc a tcatggaagg catagtgcag tcatggagat	240
	gatggtgcag ctcatggaga agatggtgcc atcatggaag gcatggtgca atcatggagt	300
	agacagtgc a gctgggcca gattctccct gactaagctc ttctcaggca cctctgagcc	360
20	gtcgtcttaa ctaggcctcc agcttggtt gtgaaaactg cagactctca gcacaaatga	420
	tttgccctct acattaagag acttaaataa acacttgc at ggctgtgttt atttaaacag	480
	ctcaaggctg tgtccctggg atgacaatga ctccagcccc taaaattcct gcttgtagaa	540
25	gctcattgct gacagaagga tctaccattt gttccagcca acacctggtg gcaggcagat	600
	agggccctgag ccccatthaa gagcagttcc tttagaaagc ttgcaattgt aaatcttttc	660
30	tctgccatt tgagatgtaa atcttctacc acctagaact gtcttctcaa ggacctgtga	720
	gctgactcac tgaaatgcaa acattcaggg agataactcc actcctgtcc ccatgcgacg	780
	gcgaggccct gactttggtg ggcaccttgc tcttatttgc accaccacct cctgtcctaa	840
35	agacatgaga cgtttgtctc tcctctggat aagtgcctat taaccaacct aggtgtcctg	900
	gtcacatgaa ccagtccagc ctagcacctg gcactgcctt tccctcagca cactccagtc	960
	tgtaaaagtc tccttatggt tgttttggtg aagttgagct tagttaatgc tagacccttc	1020
40	ctctactgca atagttactg ctgaataaag tctatcctta ccactttaac tagtggtggg	1080
	ctttgtttct ctttcataag ctcatggaga agacaatgca gttccatcaa gtttctggct	1140
45	cttacactgc taacagtcag ctctgggtgc cctgagaggg acagactcac acca	1194

<210> 423
 <211> 1118
 <212> ADN
 <213> Homo sapiens

<400> 423

55

60

65

5

10	accaagacag ttcattgaaag agacagtgc gctcatggag gagatgggtgc agctcacaga	60
	gaggatgggtg ccatcatgga aagcatgggg cagtcattgga gatgacggag tagctcatgg	120
	agaagataat gccatcatgg aaggcatagt gcagtcattgga agatgatgggt gcagctcatg	180
15	gagaagatgg tgccatcatg raaggcatgg tgcaatcatg gagtagacag tgcagctggg	240
	ccaagattct ccctgactaa gctcttctca ggcacctctg agccgtcgtc ttaactaggc	300
	ctccagcttg gcttgtagaa actgcagact ctgacacaa atgatttgcc tctacatta	360
20	agagacttaa ataaacactt gcatggctgt gtttatttaa acagctcaag gctgtgtccc	420
	tgggatgaca atgactccag cccctaaaat tctgtcttgt gaaagctcat tgctgacaga	480
25	aggatctacc atttgttoca gccaacacct ggtggcaggc agataggccc tgagcccat	540
	ttaagagcag ttcctttaga aagcttgcaa ttgtaaatct tttctctgcc catttgagat	600
	gtaaatcttc taccacctag aactgtcttc tcaaggacct gtgagctgac tcaactgaaat	660
30	gcaaacattc agggagataa ctccactcct gtcccatgac gacggcgagg ccctgacttt	720
	ggtgggcacc ttgctcttat ttgcaccacc acctcctgtc cttaaagacat gagacgtttg	780
	tctctcctct ggataagtgc ctattaacca acccaggtgt cctggtcaca tgaaccagtc	840
35	cagcctagca cctggcactg cctttccctc agcacactcc agtctgtaaa agtctcctta	900
	tggttgtttt ggcaaagtgt agcttagtta atgctagacc ccttctctac tgcaatagtt	960
	actgctgaat aaagtctatc cttaccactt taactagtgt tgggctttgt ttctctttca	1020
40	taagctcatg gagaagacaa tgcagttcca tcaagtttct ggctcttaca ctgctaacag	1080
	tcagctctgg ggtccctgag agggacagac tcacacca	1118

45 <210> 424
 <211> 601
 <212> ADN
 <213> Homo sapiens

50 <400> 424

	gtaggggac tgtctatact ggctgcactc tggccagtgc tgtcccaacg ctgaccctc	60
	tggaagctaa tctggcttat aatgaggatg ctttctttag aggggactct ccatgcacag	120
55	cagaaaatcc caatggagtg gttcttccct atgtcccaa gggactggga atattctttc	180
	agtaacaatg gccattggg ggaagaagga tgaaagtggg gtgagagacg tgaaatttgg	240
	agaggtccct caaagattgt gatgtgcctc tcttgttcca atcacaggac aggggtataa	300
60	yggtcttctt ttgaaacacg gggatgaatt taactattca cttccaggt agattcatca	360
	gggtctagag cttcagctaa cagcatgagg aagattocaa atgtgcccc atcagcatag	420
65	gaactgggta tgttgagtct atggtctcat aaaaccagaa gaaggacaag ggattgtggc	480

ES 2 704 701 T3

5

10	tccaggcttg ggagcacott ttccttacca tgggctacag tatttattta gggtaaagga	540
	aggaaactcc tgaggtgcta tggggtgcca gcaatttgga gcatcagtaa ttcaatgtcc	600
	c	601

15

<210> 425
<211> 601
<212> ADN
<213> Homo sapiens

20

<400> 425

25	acgctgaccc	ctctggaagc	taatctggct	tataatgagg	atgctttctt	tagaggggac	60
	tctccatgca	cagcagaaaa	tccaatgga	gtggttcttc	cctatgtccc	caagggactg	120
	ggaatattct	ttcagtaaca	atggccatt	gggggaagaa	ggatgaaagt	ggggtgagag	180
30	acgtgaaatt	tggagaggtc	cctcaaagat	tgtgatgtgc	ctctcttgtt	ccaatcacag	240
	gacaggggta	taacggcttt	cctttgaaac	acggggatga	atttaactat	tcacttccca	300
	rgtagattca	tcagggtcta	gagcttcagc	taacagcatg	aggaagattc	caaatgtgcc	360
35	cccatcagca	taggaactgg	gtatgttgag	tctatggtct	cataaaacca	gaagaaggac	420
	aagggattgt	ggctccaggc	ttgggagcac	cttttcctta	ccatgggcta	cagtatttat	480
	ttagggtaaa	ggaaggaaac	tcctgaggtg	ctatggggtg	ccagcaattt	ggagcatcag	540
40	taattcaatg	tcccttcagc	catgtgtatt	caactcctgc	tgtgggtgtg	gacttggtgc	600
	a						601

45 <210> 426
 <211> 701
 <212> ADN
 <213> Homo sapiens

50 <400> 426

55

60

65

5

10	ttcctgggca tcgtcatatt ctgtaaaaca aggaagctca gccagtggtg ttctaacaatg	60
	acctcctttc tacatcctta ggtgttggtta tgcgtgaatc acgtcccccc aaaagacatg	120
	ttcatgtcct aacccccagg acctcagaat gtgtgatctg gtttggaat aaggtcatca	180
15	cagatgaaat tagctaagac aaggtcatat tggaataggg ttggccctta atccactgtg	240
	actggtgtcc ttttaagaag aggacacaga cacaggaggg gagaggcca tgggatgatg	300
	cagggtggaga ctggagtgtc acagctgcaa gcaatacat ttctgtgtg tgaagccacc	360
20	catttggtgg tactacgtta aaacagctct aggaattaa tacagatgtt gcctgtatct	420
	ttgtttctca tattactact cattgtttta atgatgactg ttttattcat taagtgtgaa	480
25	gctcctaaag cagagggacc rtatctttat gtcccaactc tccttaaggc cttgcctatg	540
	atagcacatc tcttcaatag aattgtccta actttaacag agacaacttg gggtatttaa	600
	tatggagaac aaagggttaa gctggtgcca gatgggtttc attttctcta aatctggaac	660
30	caaaggcagc aagtctatgg ggtggacgga gttcttagct c	701

35
 <210> 427
 <211> 701
 <212> ADN
 <213> Homo sapiens

40	<400> 427	
	caaggaagct cagcccagtg tgttctaaca tgacctcctt tctacatcct taggtgtgtg	60
	tatgcgtgaa tcacgtccc ccaaaagaca tgttcatgtc ctaaccccca ggacctcaga	120
45	atgtgtgatc tggtttgga ataaggtcat cacagatgaa attagctaag acaaggtcat	180
	attggaatag ggttgccct taatccactg tgactggtgt ccttttaaga agaggacaca	240
	gacacaggag gggagagggc catgggatga tgcagggtga gactggagtg ctacagctgc	300
50	aagcaaatac atttctgtgc tgtgaagcca ccatttggg ggtactacgt taaaacagct	360
	ctaggaaatt aatacagatg ttgcctgtat tttgtttct catattacta ctcatgttt	420
	taatgatgac tgttttatcc attaatgtga aagctcctaa agcagaggga ccatattttt	480
55	atgtcccaac tctccttaag sccttgccca tgatagcaca tctcttcaat agaattgtcc	540
	taactttaac agagacaact tgggttattt aatatggaga acaaagggtt aagctggtgc	600
60	cagatgggtt tcattttctc taaatctgga accaaaggca gcaagtctat ggggtggacg	660
	gagttcttag ctcaaccctt tggtgaggta agaagaagga t	701

65

REIVINDICACIONES

- 5 **1.** Un método para preparar una biblioteca de secuenciación a partir de una muestra de prueba que comprende moléculas de ácido nucleico, en donde el método comprende los pasos consecutivos de la reparación de extremos, la cola de dA y el adaptador que ligan dichos ácidos nucleicos, y en el que dichos pasos consecutivos excluyen la purificación de los productos reparados de forma final antes de la etapa de la cola de dA y excluyen la purificación de los productos de cola de dA antes del paso de ligadura del adaptador.
- 10 **2.** El método de la Reivindicación 1, en el que dichos pasos consecutivos se realizan en ausencia de polietilenglicol.
- 3.** El método de la Reivindicación 1 o la reivindicación 2, en el que dichos pasos consecutivos se realizan en menos de 1 hora.
- 15 **4.** El método de cualquier reivindicación precedente, en el que dichos ácidos nucleicos son moléculas de ADN.
- 5.** El método de cualquier reivindicación precedente, en el que dichos ácidos nucleicos:
 - (i) son moléculas de ADN genómico; o
 - (ii) son moléculas de ADN genómico humano.
- 20 **6.** El método de la Reivindicación 5, en el que dicho ADN genómico se somete a fragmentación antes de los pasos consecutivos de reparación final, cola de dA y adaptador que liga dichos ácidos nucleicos.
- 25 **7.** El método de cualquiera de las Reivindicaciones 1-5, en el que dichos ácidos nucleicos son moléculas de ADN libres de células (ADNcf) y no se someten a fragmentación antes de los pasos consecutivos de reparación final, colas de dA y ligadura con adaptador de dichos ácidos nucleicos .
- 8.** Uso de una biblioteca de secuenciación preparada según el método de cualquiera de las Reivindicaciones 1-7 en un método de secuenciación de ácido nucleico.
- 30 **9.** El uso de la Reivindicación 8, en el que dicho método de secuenciación es un método de secuenciación de próxima generación (NGS).
- 10.** El uso de la Reivindicación 8, en el que dicho método de secuenciación es un método de secuenciación masivamente paralelo, y en el que
 - (i) dicha secuenciación es una secuenciación masivamente paralela que utiliza secuenciación por síntesis con terminadores de tinte reversibles; o
 - (ii) dicha secuenciación es una secuenciación masivamente paralela utilizando la secuenciación por ligadura.
- 40 **11.** El uso de la Reivindicación 8, en el que dicha secuencia comprende una amplificación.
- 12.** El uso de la Reivindicación 8, en el que dicha secuenciación es una secuenciación de una sola molécula.
- 45 **13.** El método de cualquiera de las Reivindicaciones 1-7 o el uso de cualquiera de las reivindicaciones 8-12, en el que la muestra:
 - (a) es una muestra de plasma derivada de sangre periférica que comprende una mezcla de ADNcf derivado de células normales y cancerosas;
 - (b) se deriva de una mezcla de células cancerosas y no cancerosas de un fluido biológico seleccionado de suero, sudor, lágrimas, esputo, orina, flujo de oído, linfa, saliva, líquido cefalorraquídeo, lavados, suspensión de médula ósea, flujo vaginal, Lavado transcervical, líquido cerebral, ascitis, leche, secreciones de las vías respiratorias, intestinales y geniutnarias, y muestras de leucoforesis; o
 - (c) se selecciona de biopsias de tejido, o frotis.
- 50
- 55
- 60
- 65

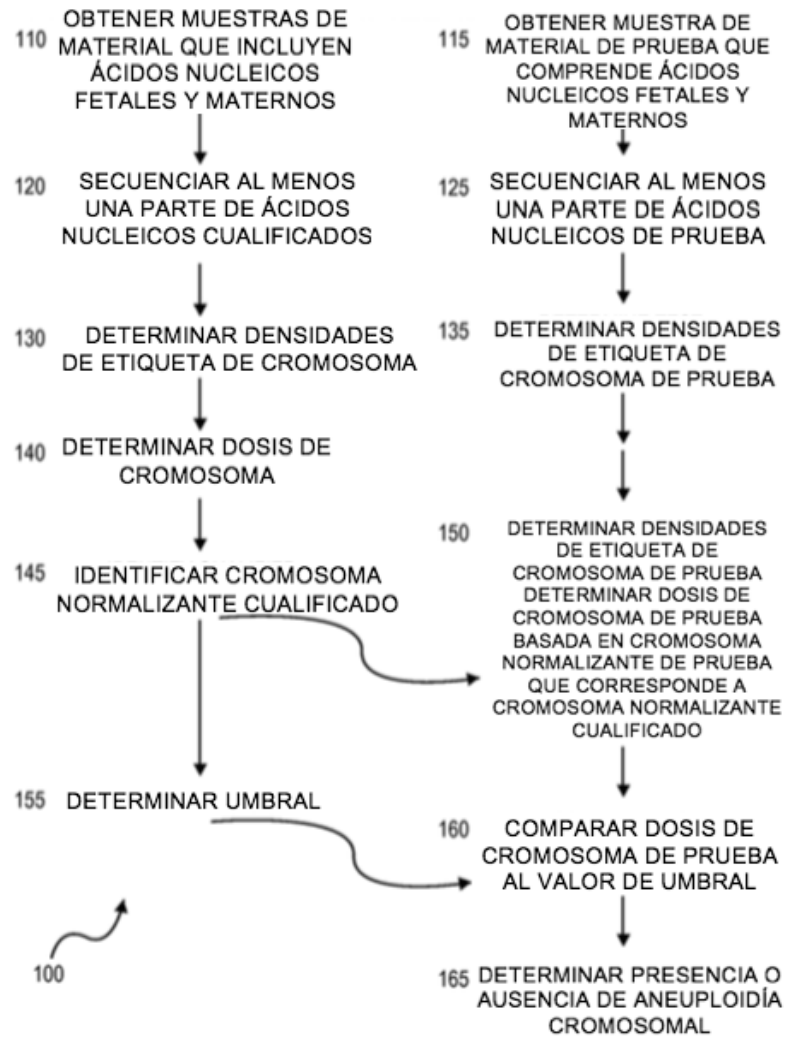


FIG. 1

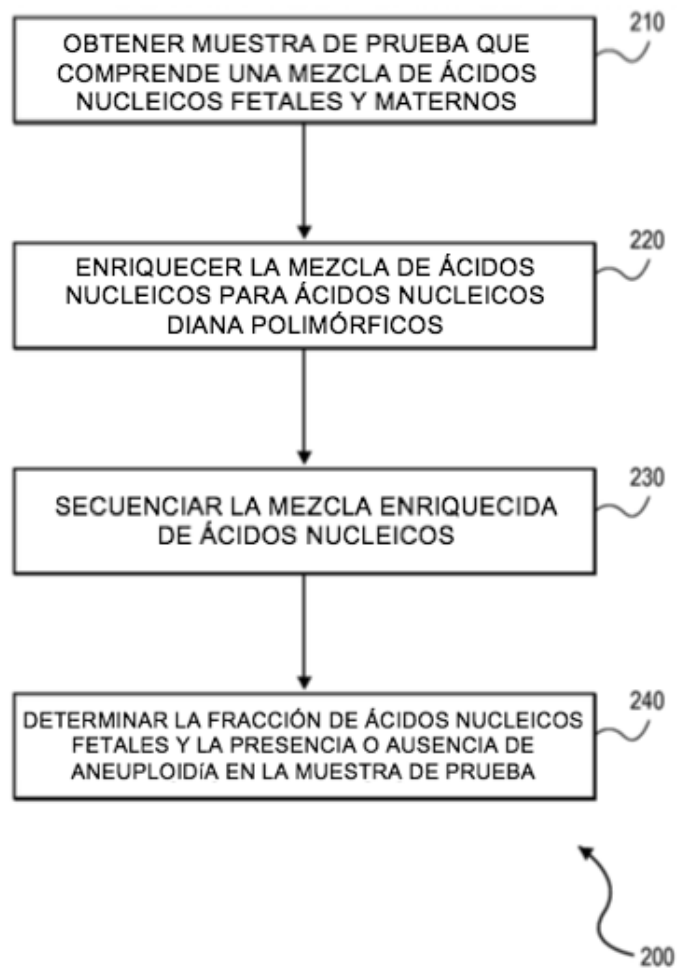
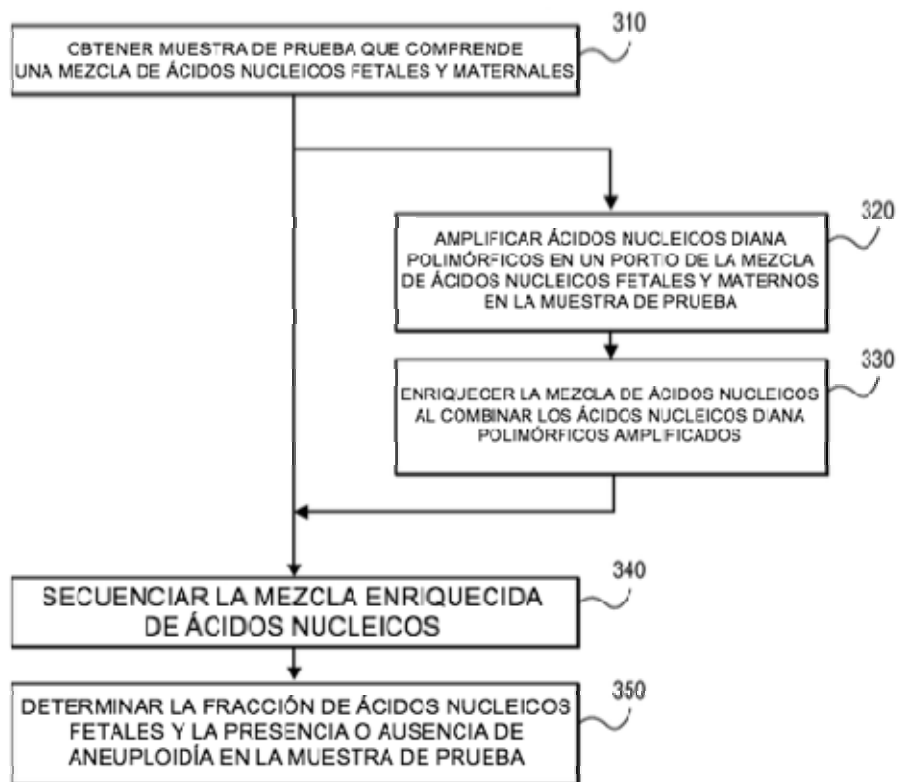


FIG. 2



300

FIG. 3

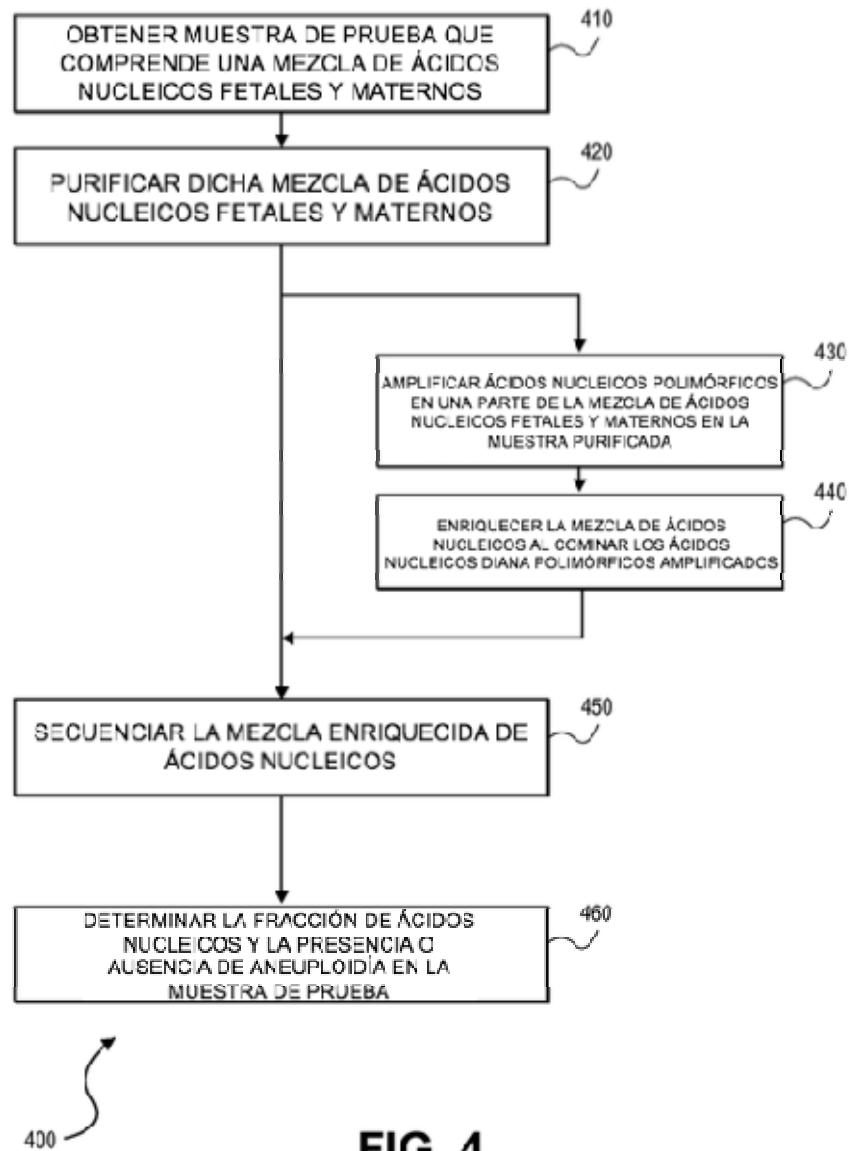


FIG. 4

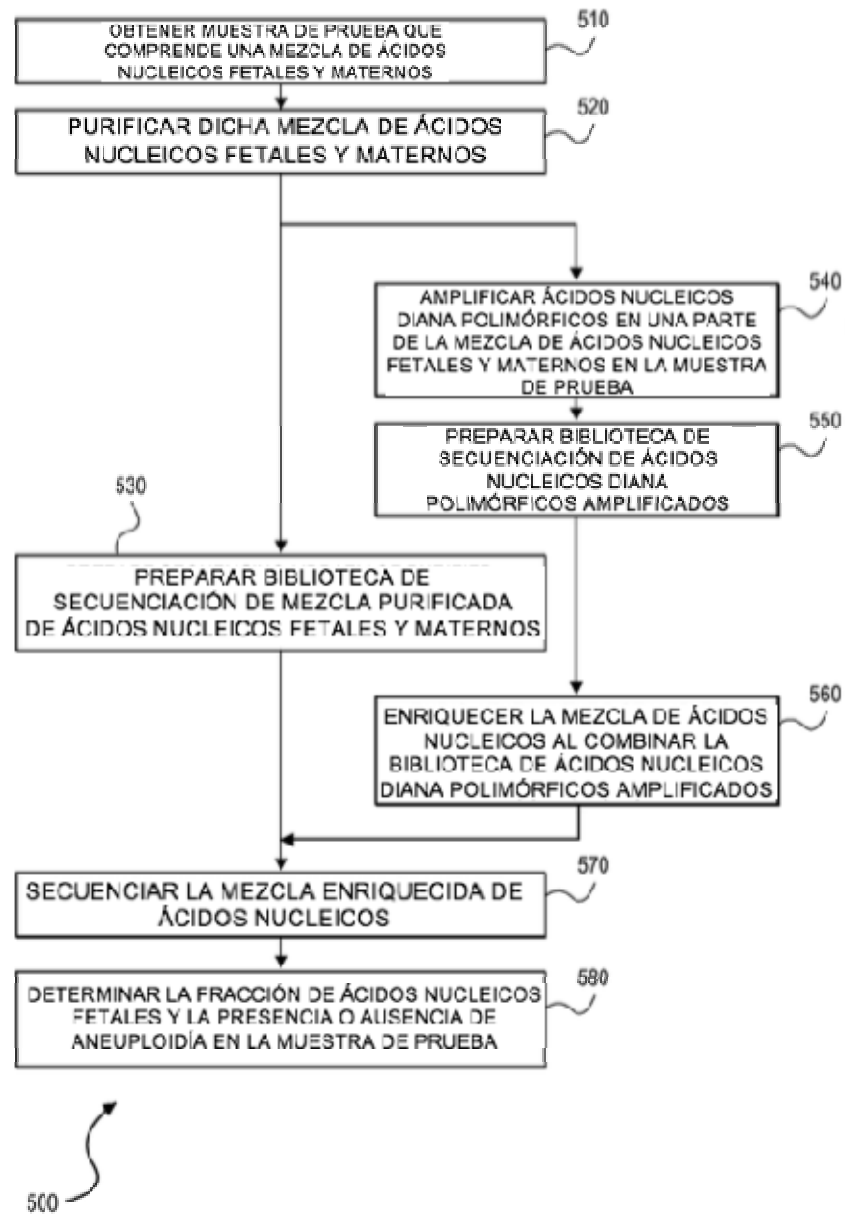


FIG. 5

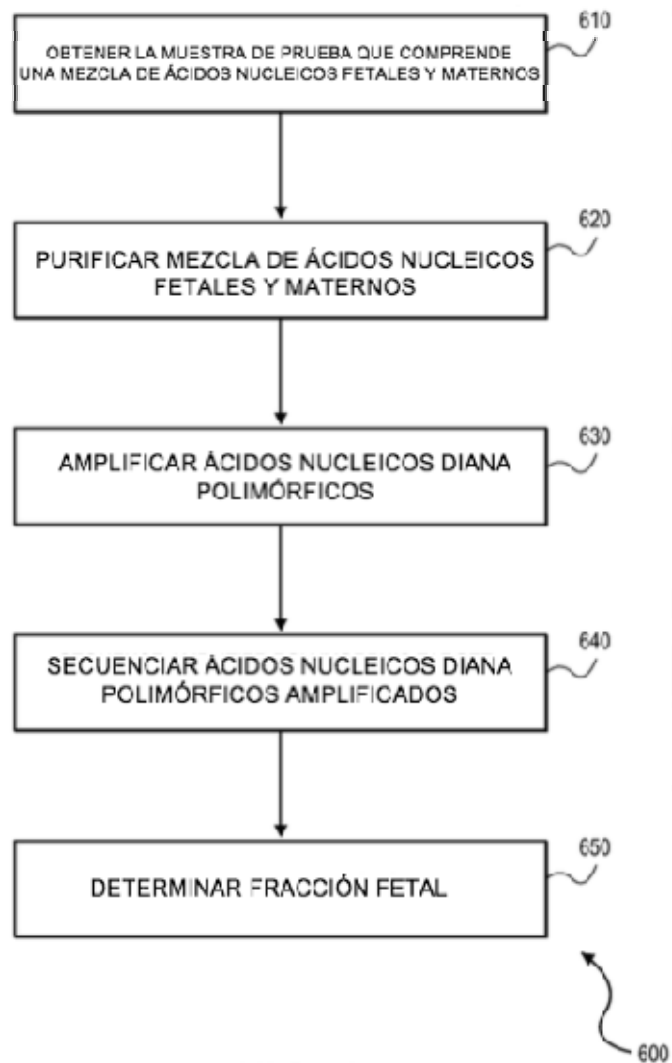


FIG. 6

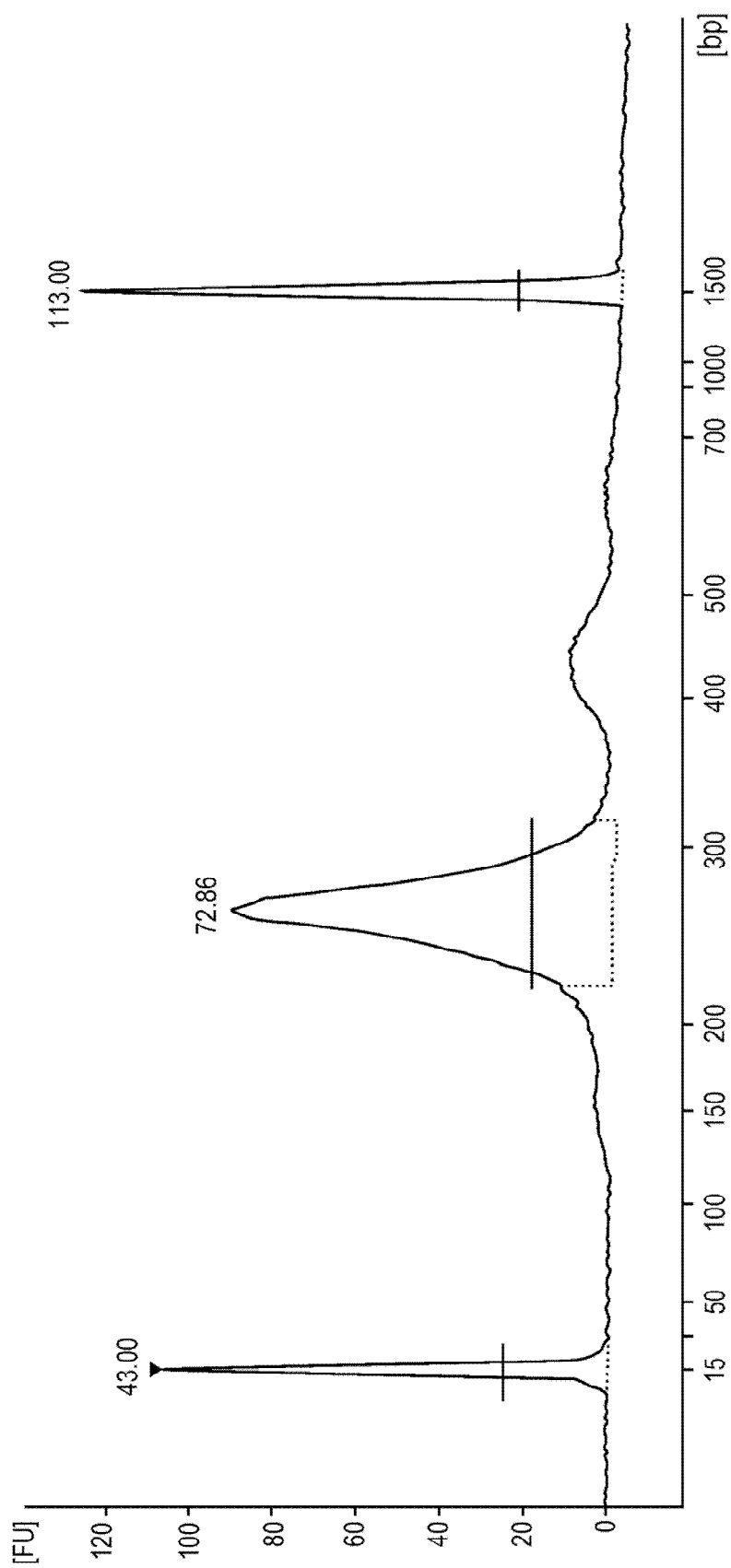


FIG. 7A

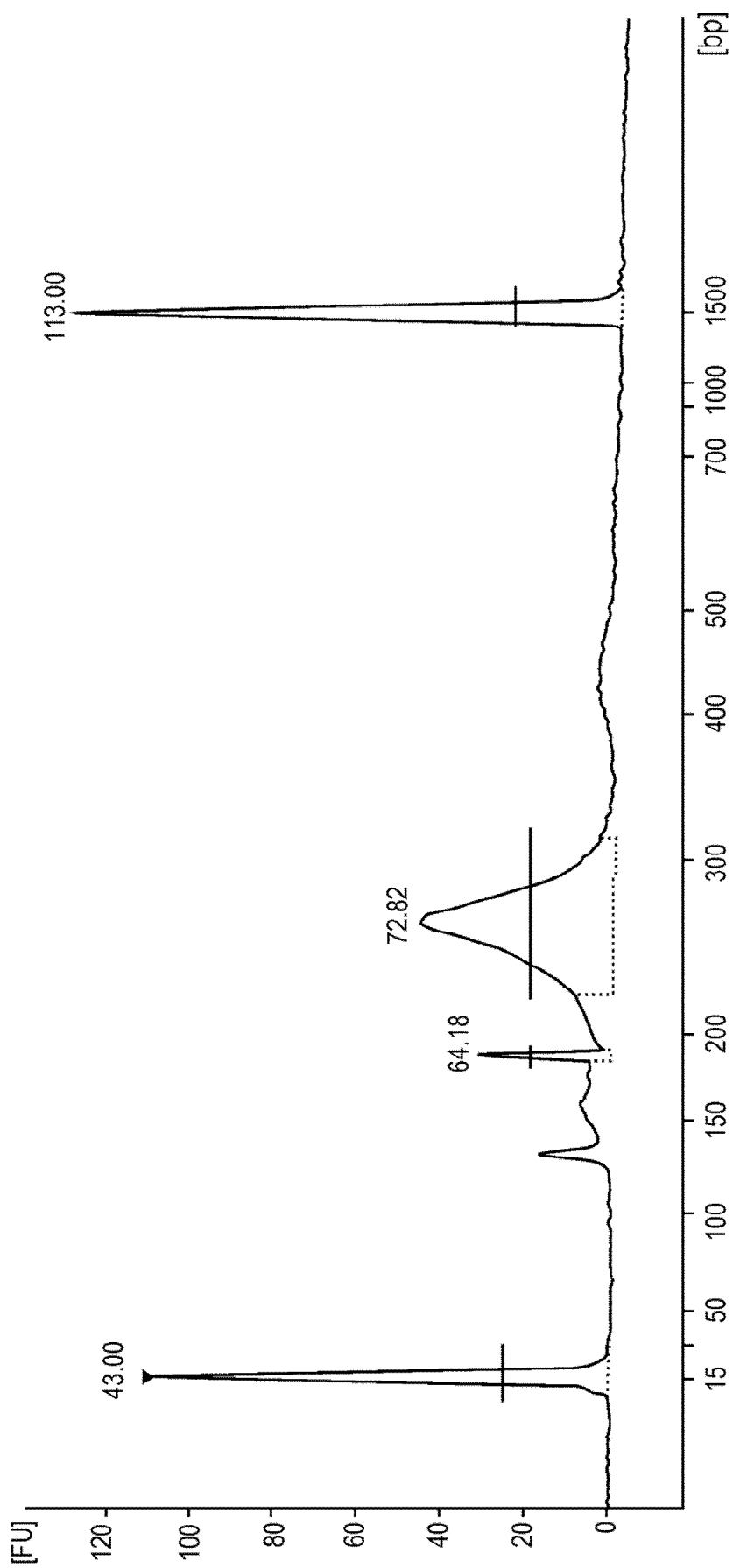


FIG. 7B

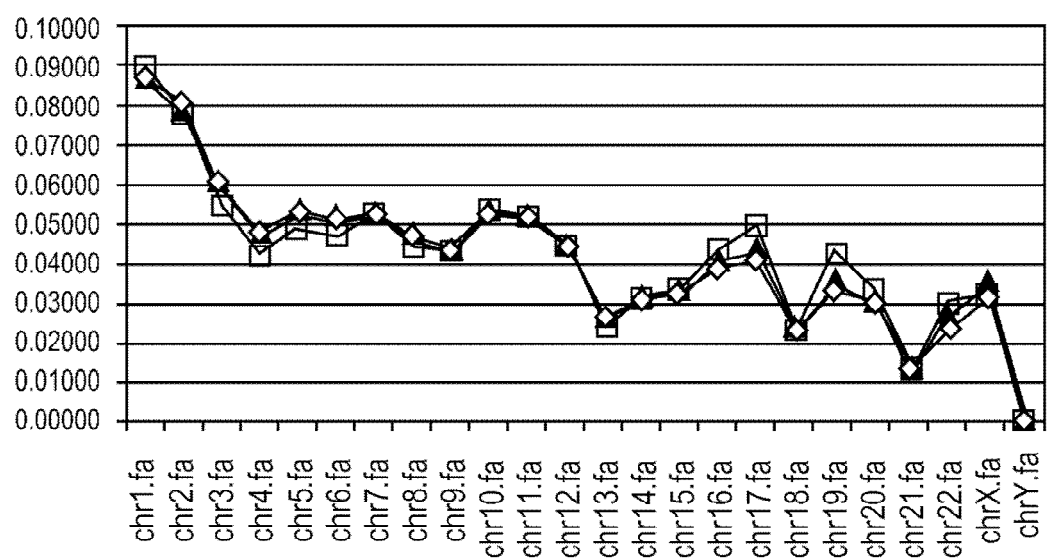


FIG. 8

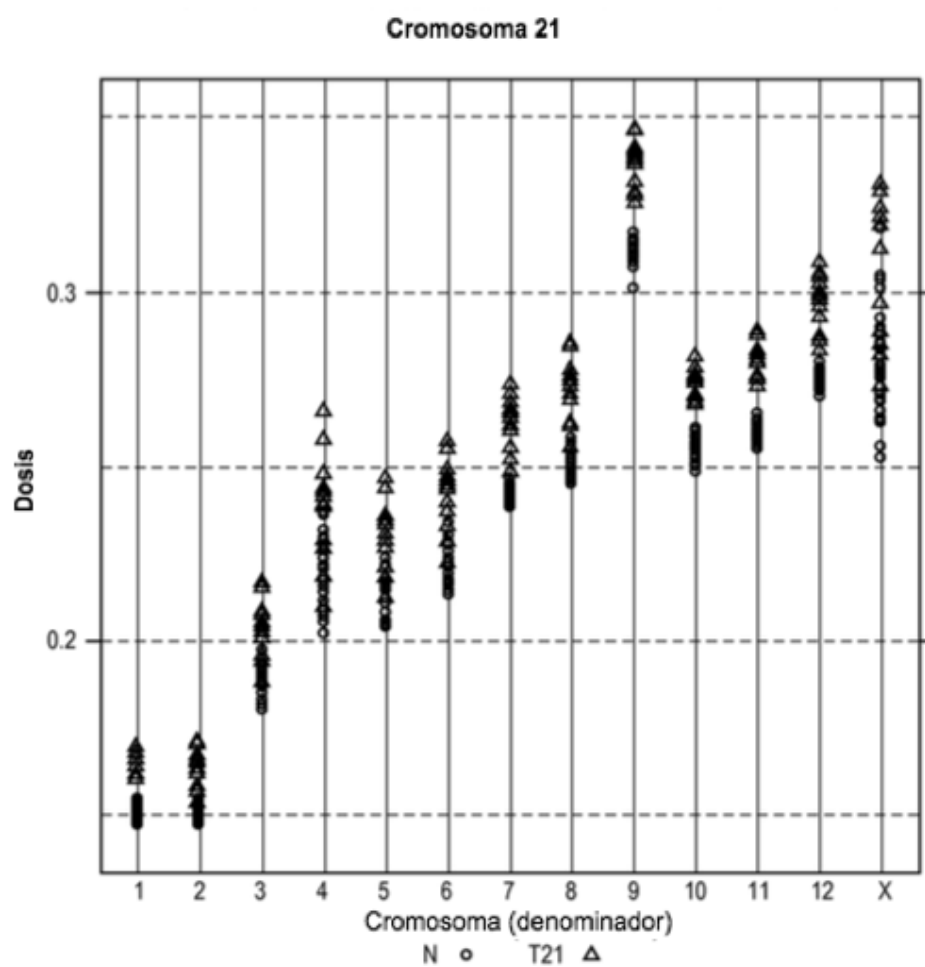


FIG. 9A

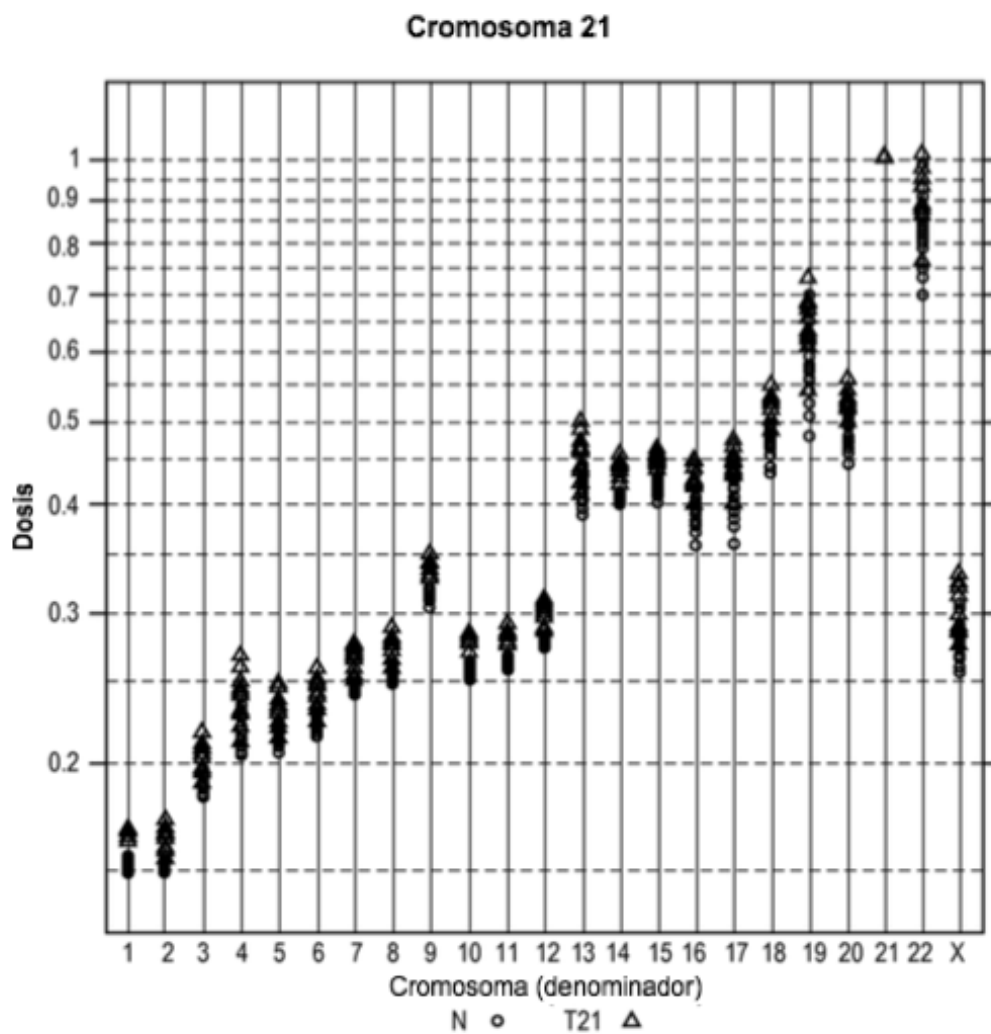


FIG. 9B

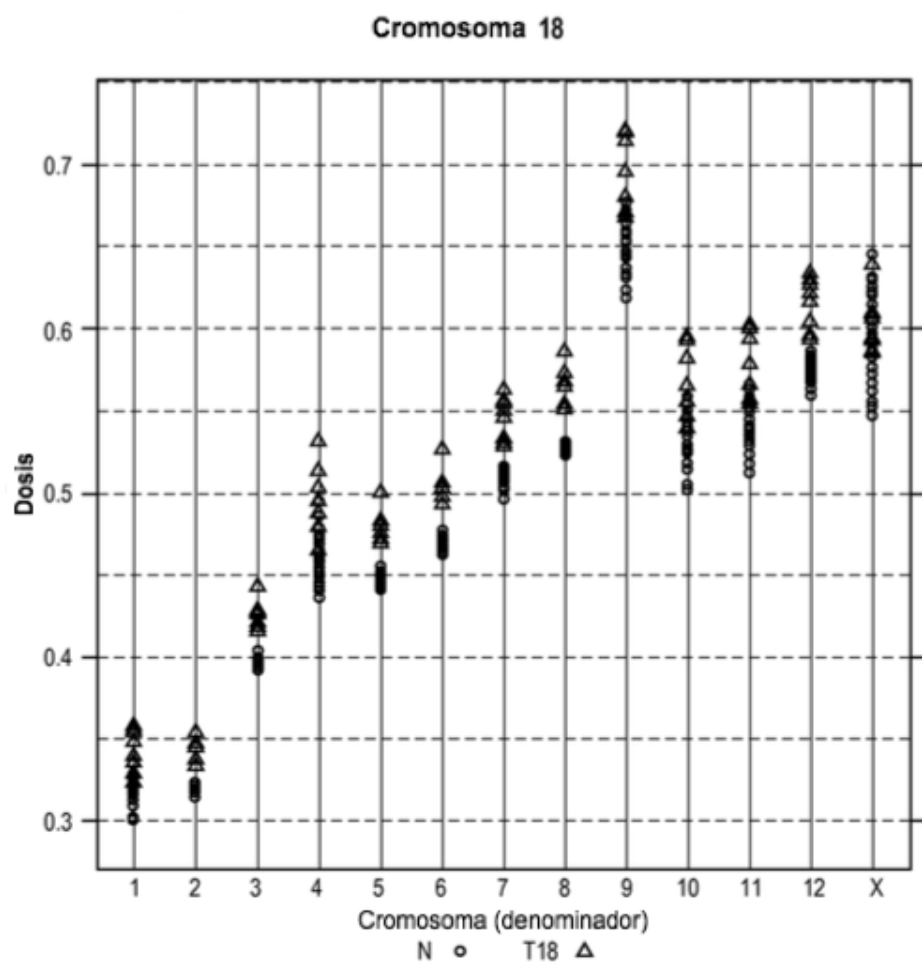


FIG. 10A

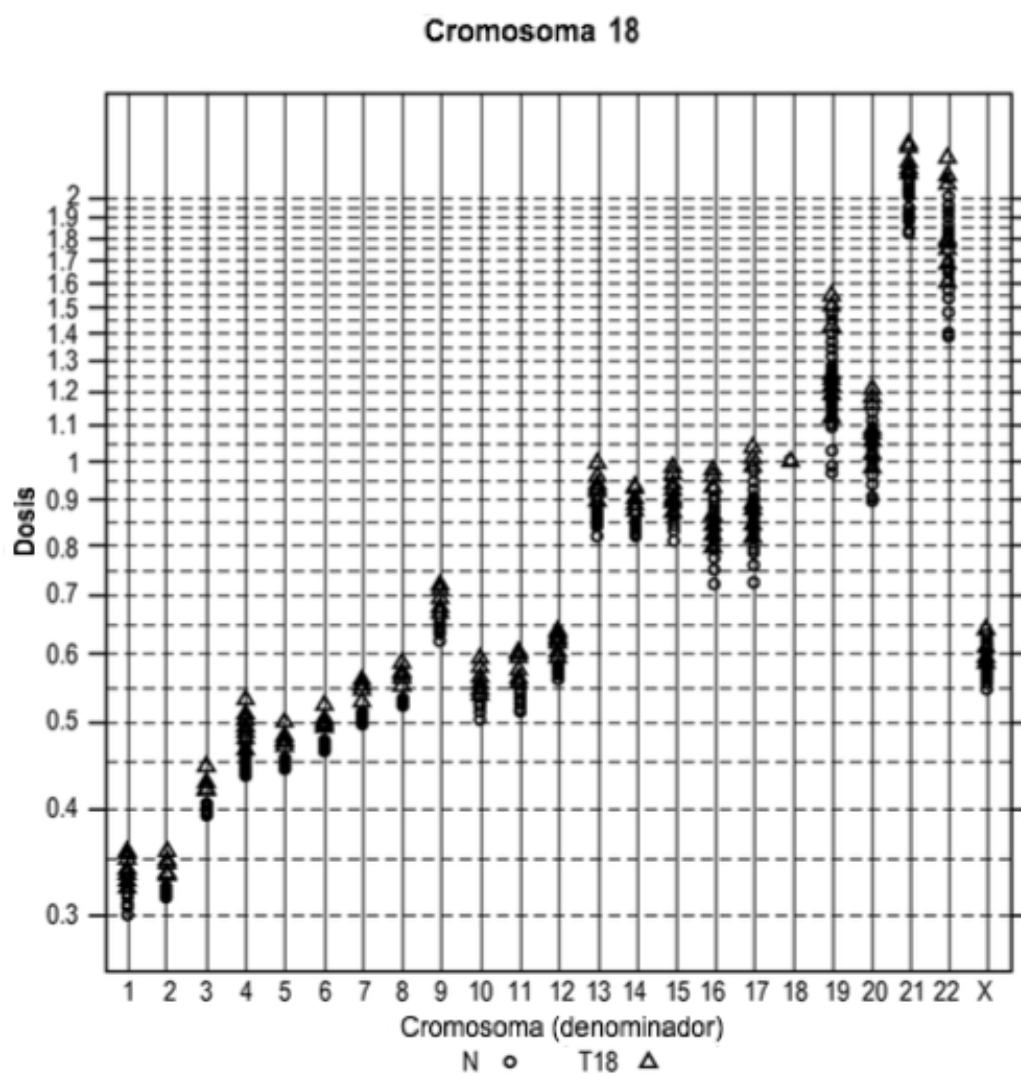


FIG. 10B

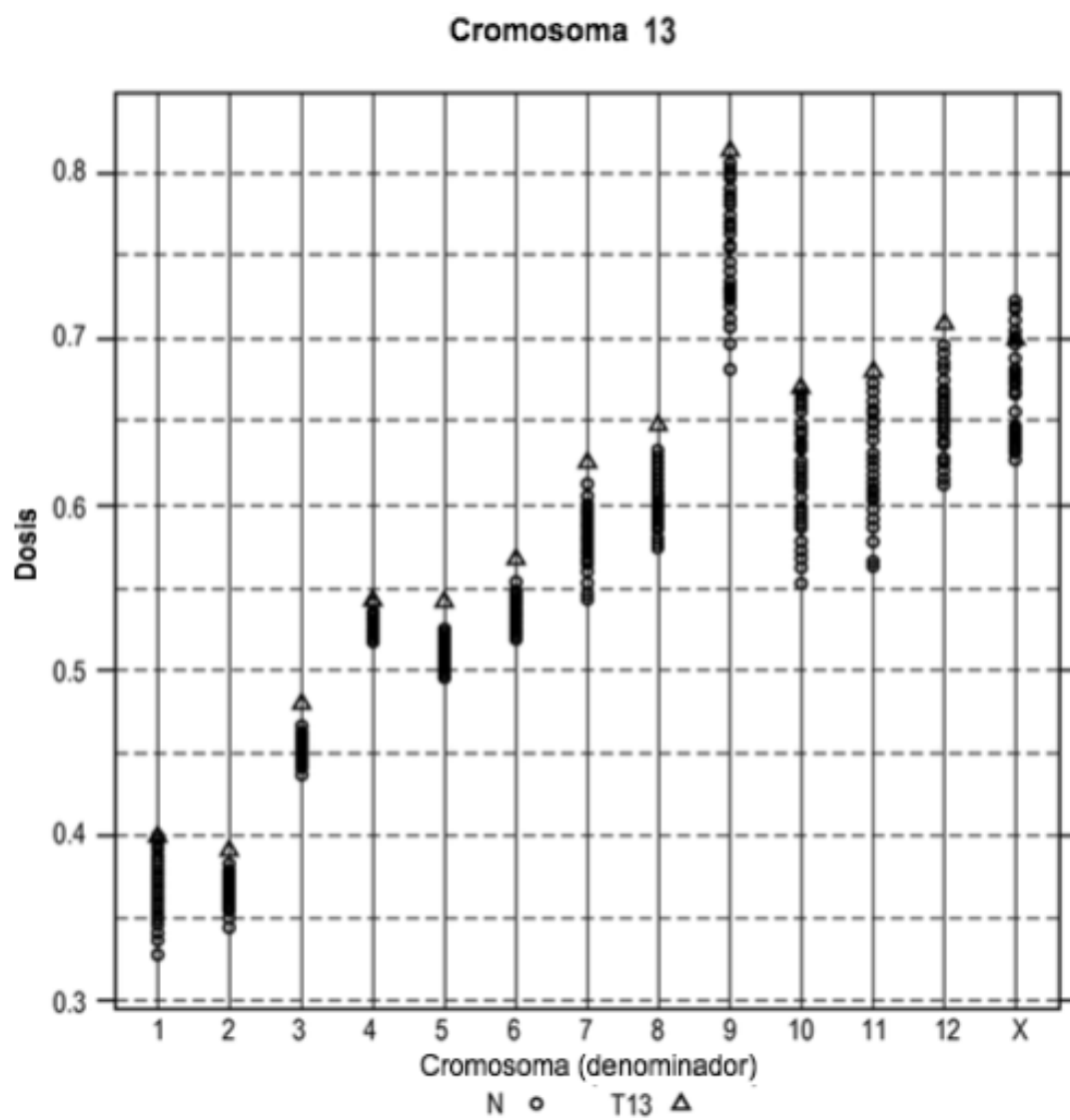


FIG. 11A

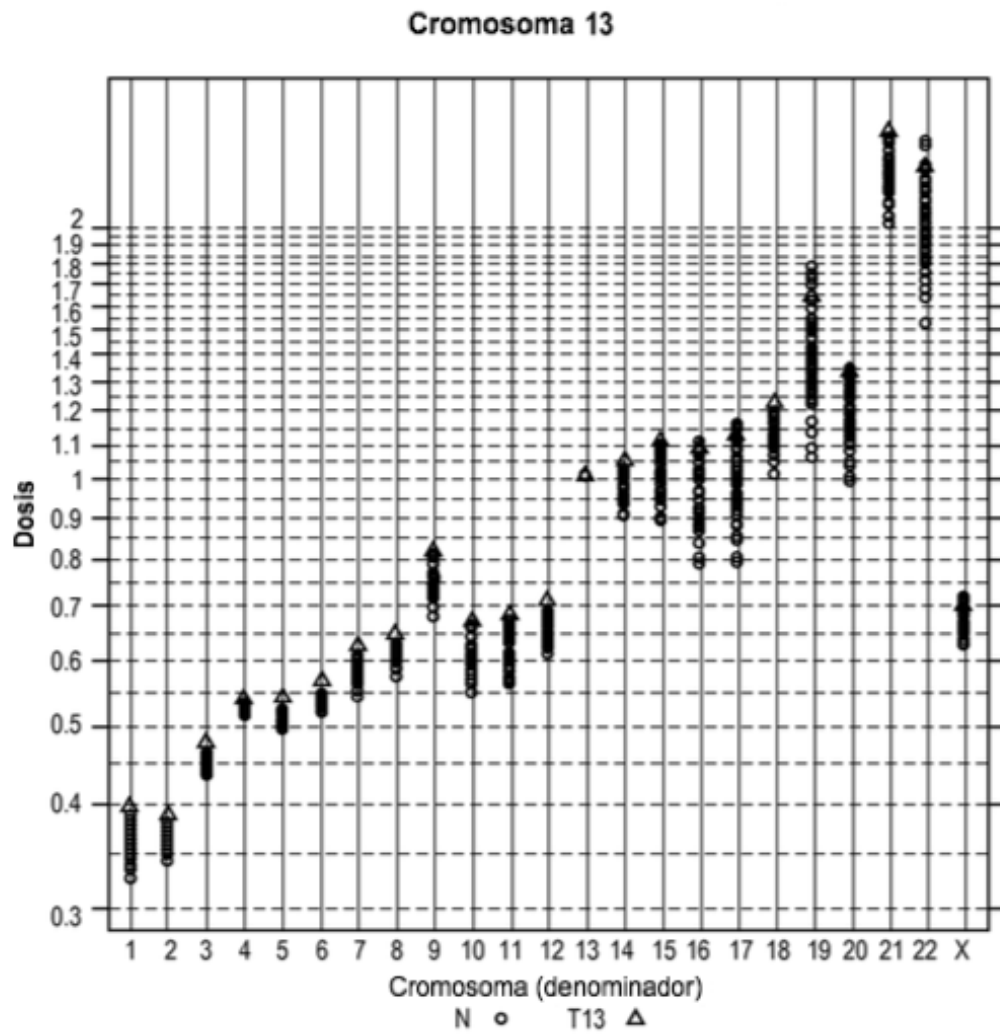


FIG. 11B

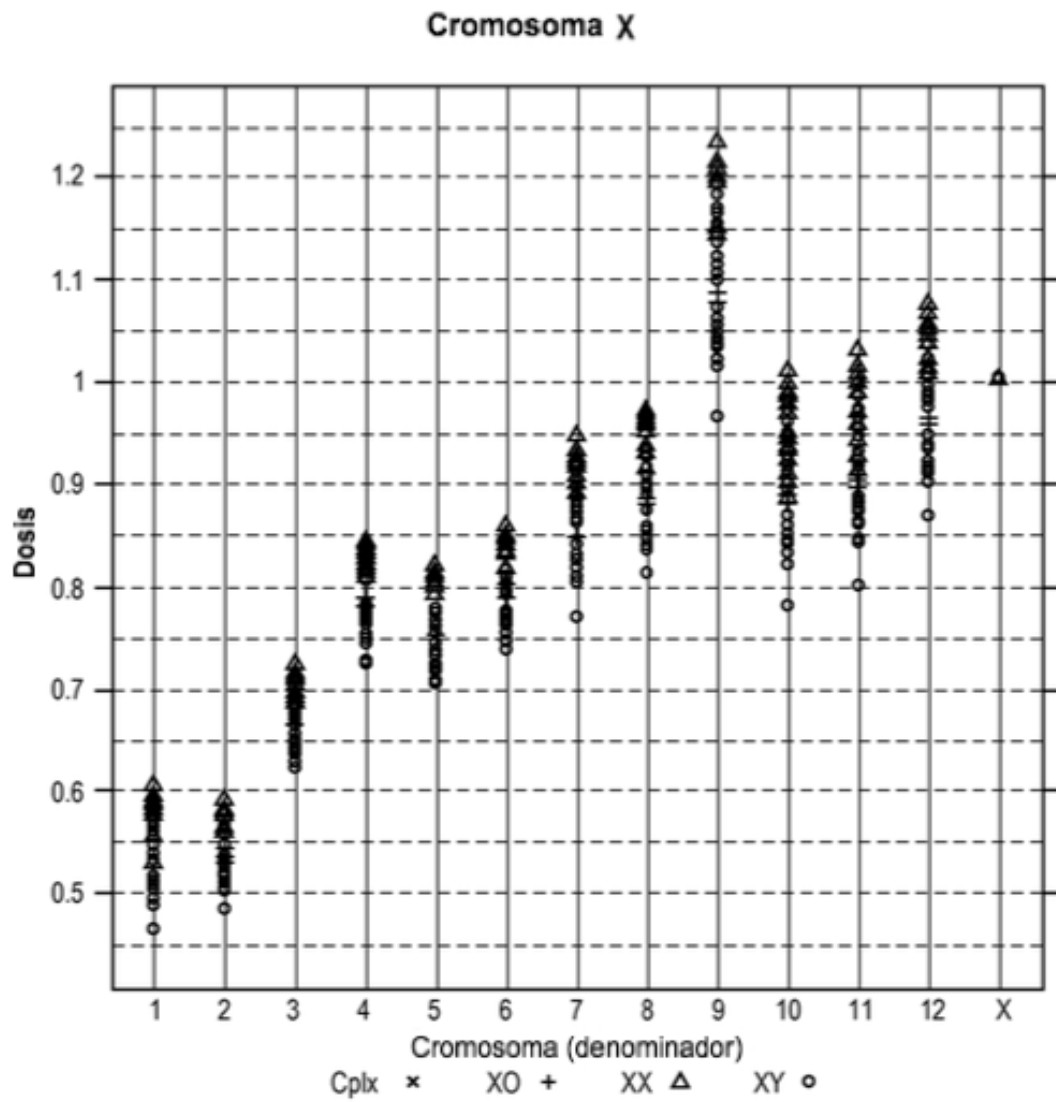


FIG. 12A

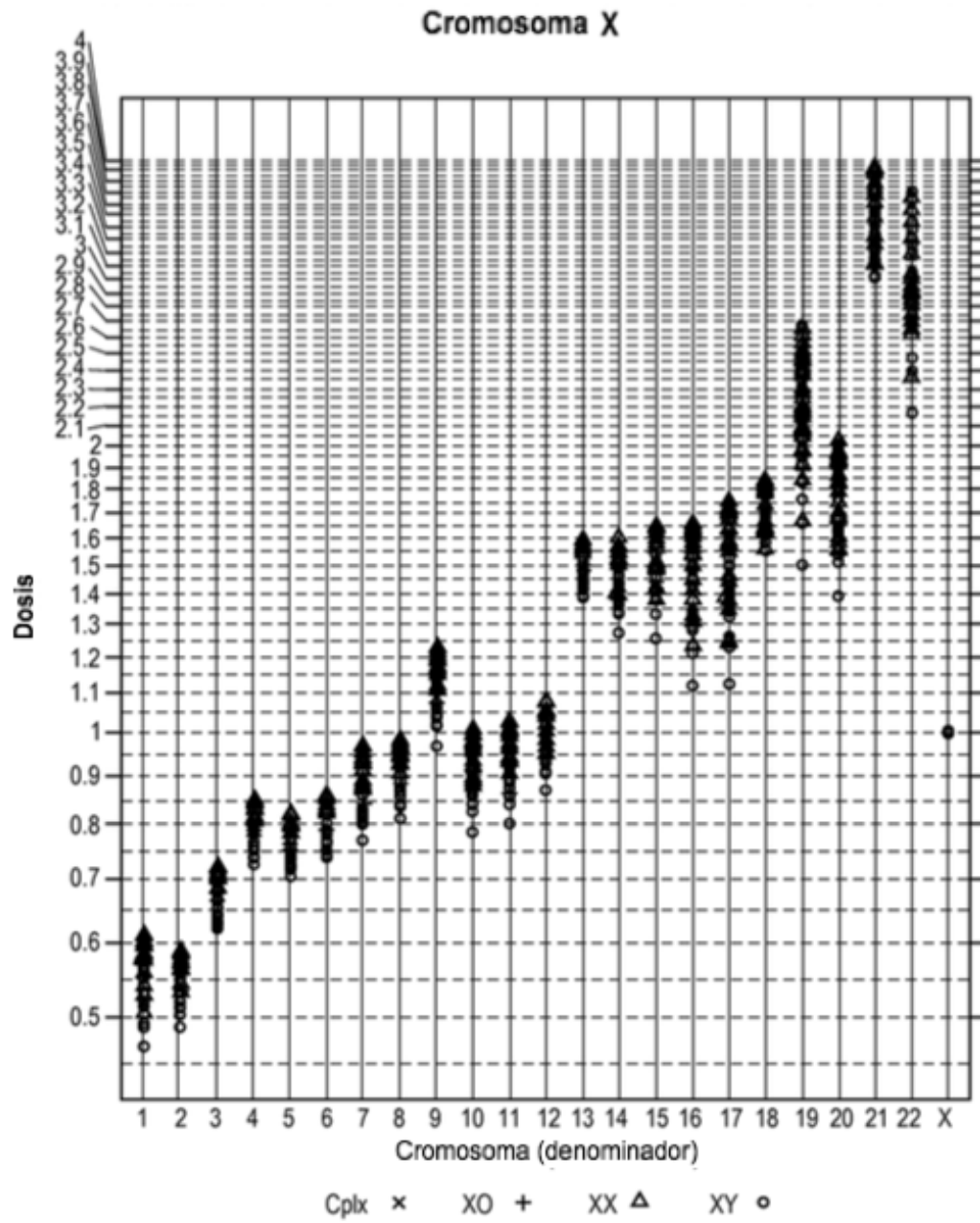


FIG. 12B

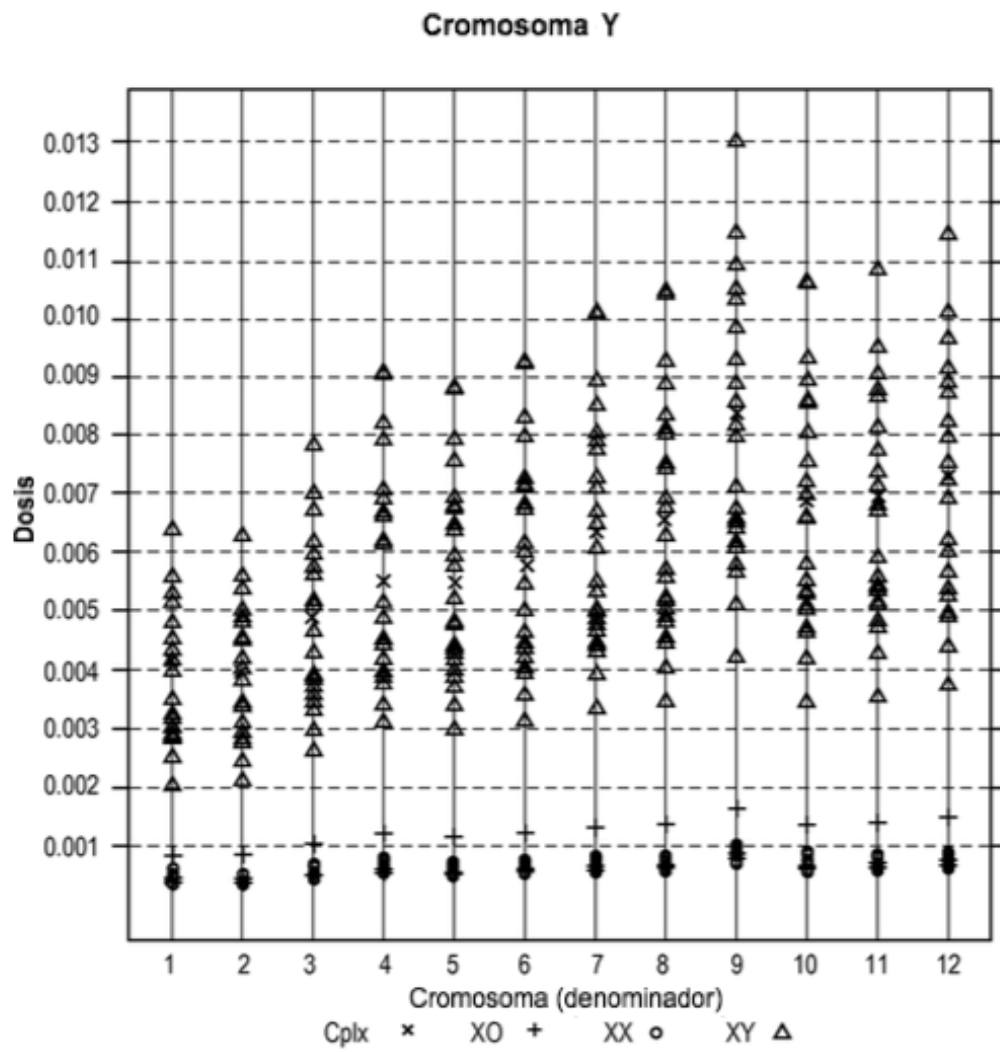


FIG. 13A

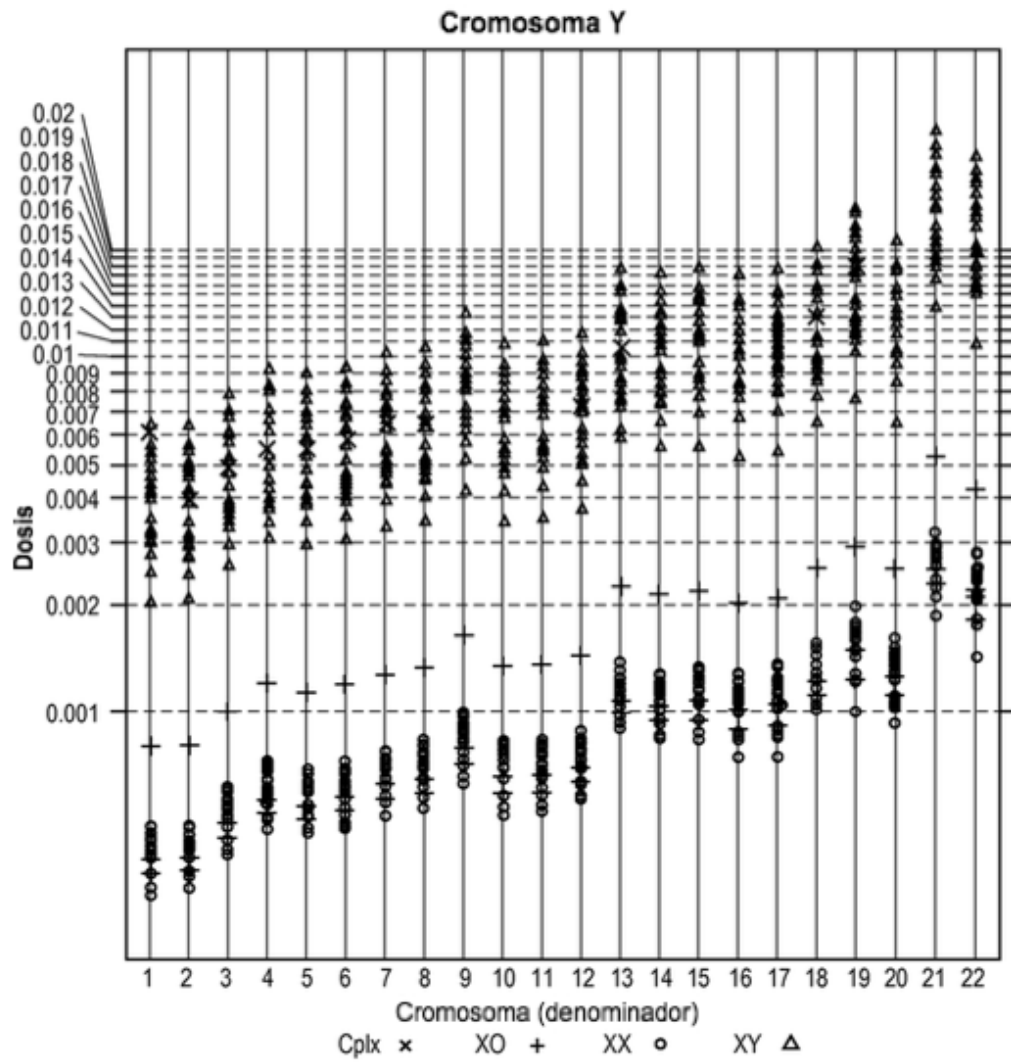
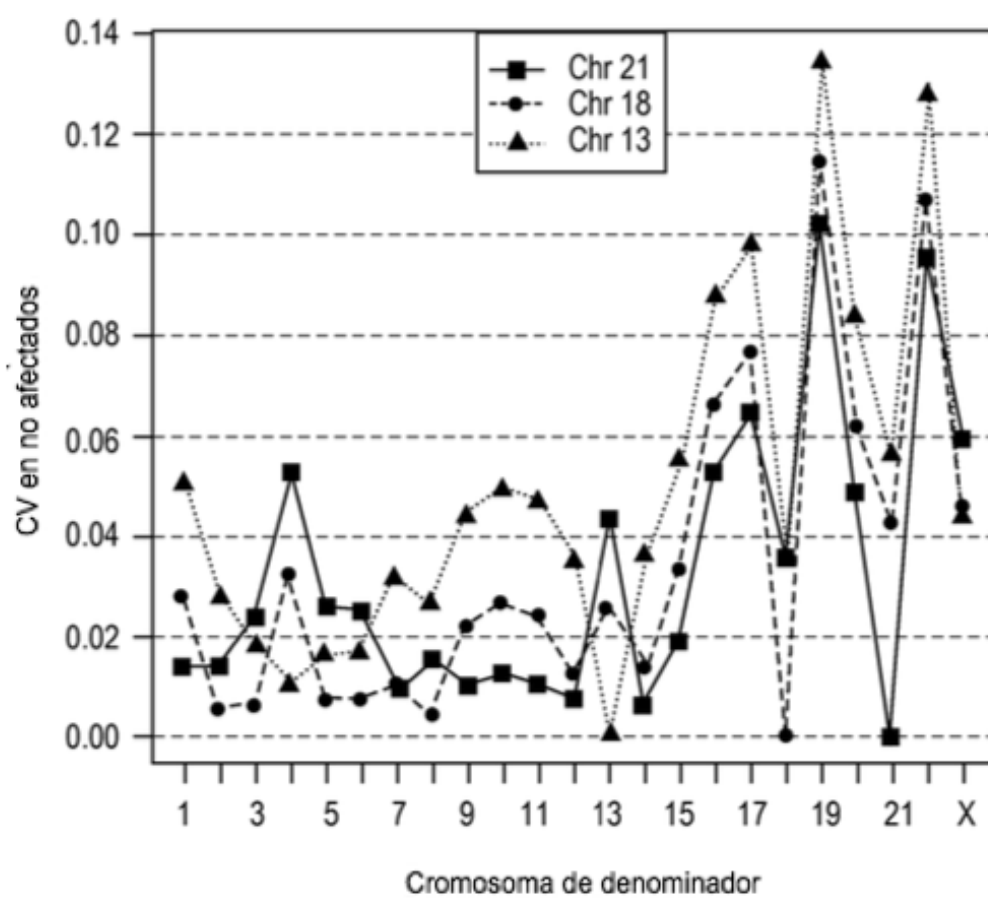
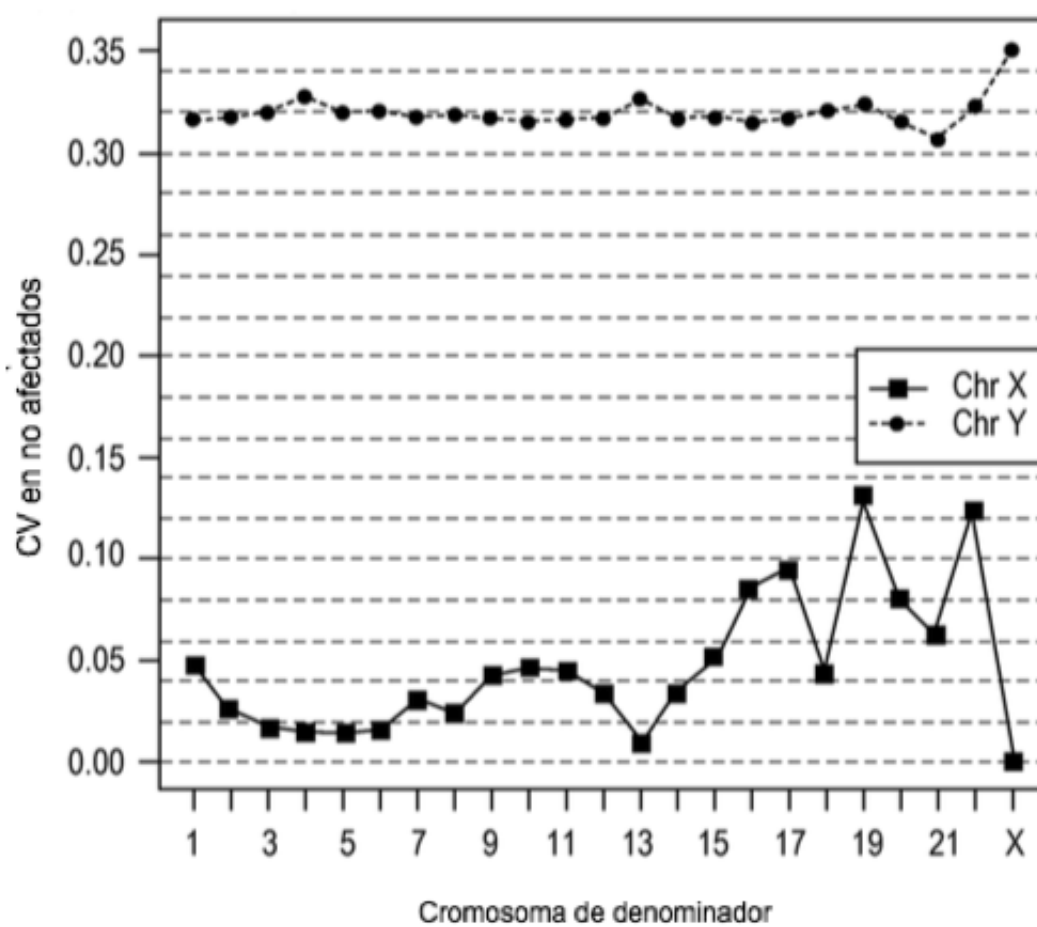


FIG. 13B

**FIG. 14**

**FIG. 15**

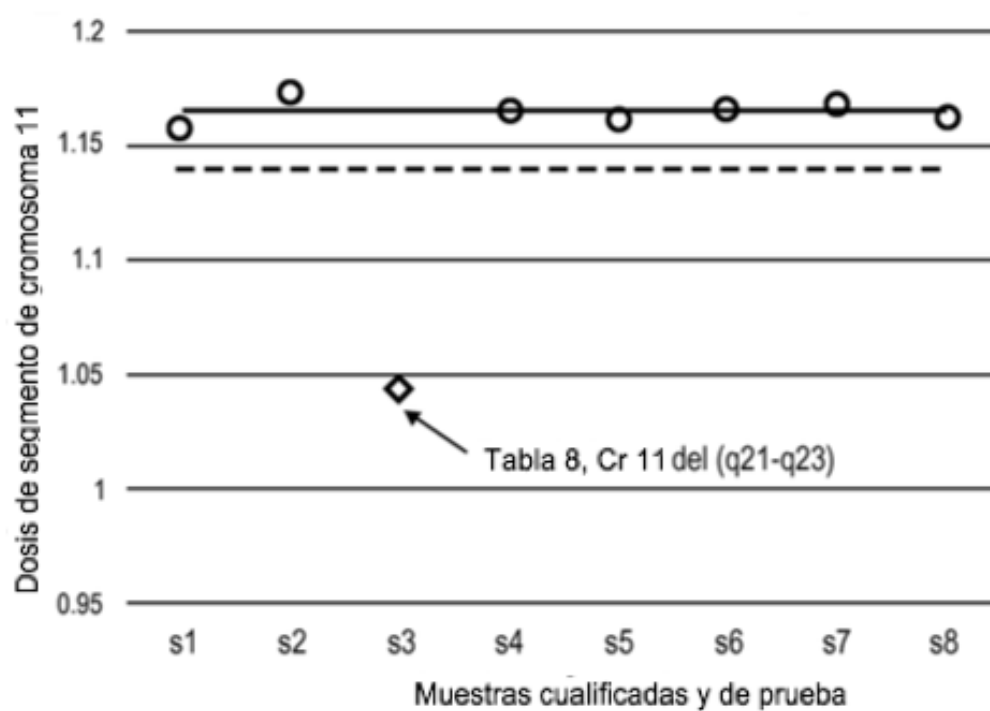


FIG. 16

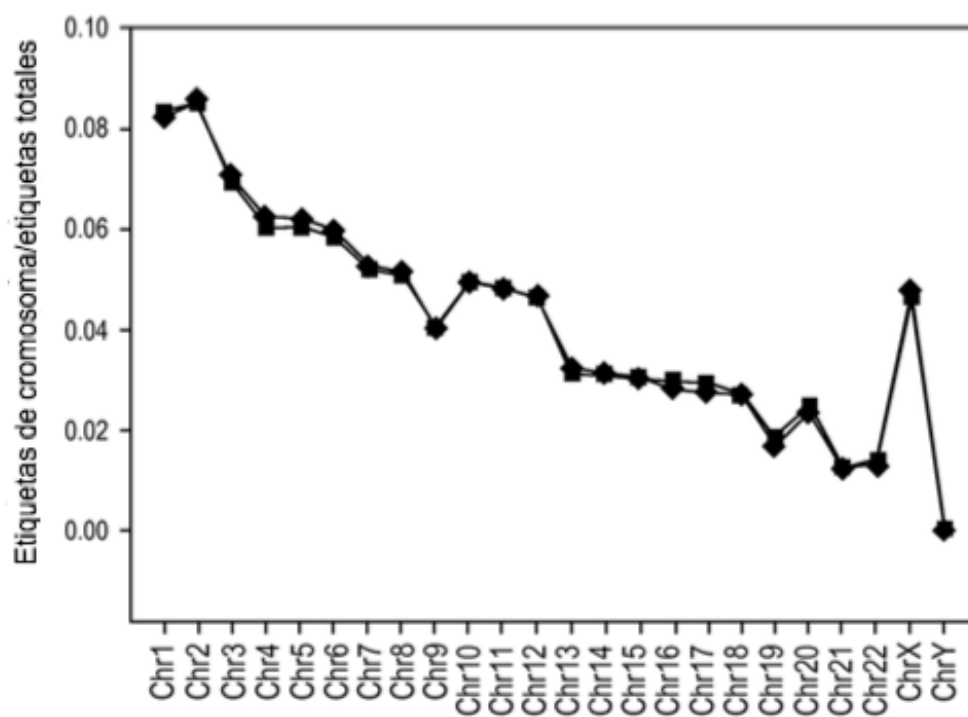


FIG. 17

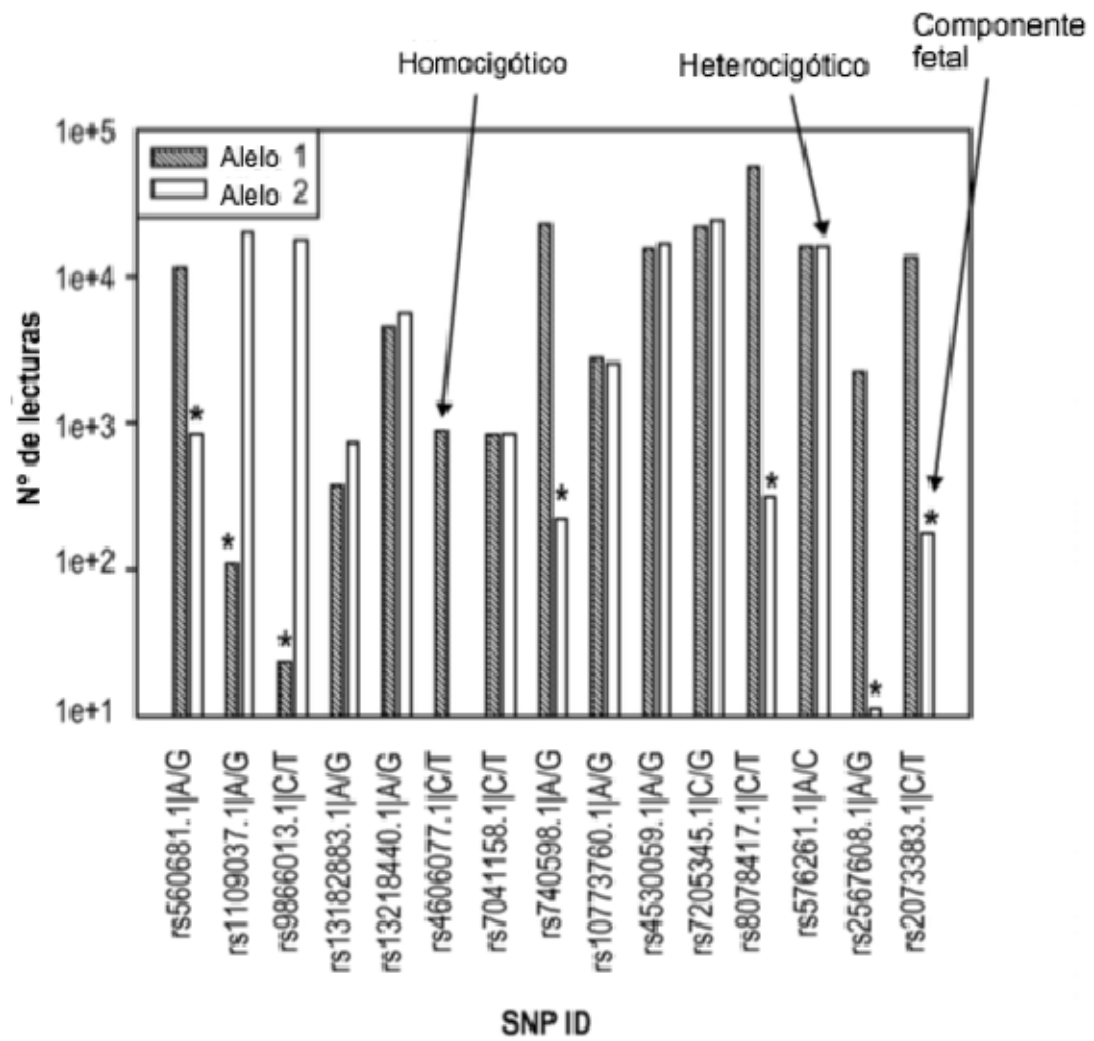


FIG. 18

POBLACIÓN DE PUNTUACIONES
"NORMALES", DISTRIBUCIÓN
ESTIMADA f1

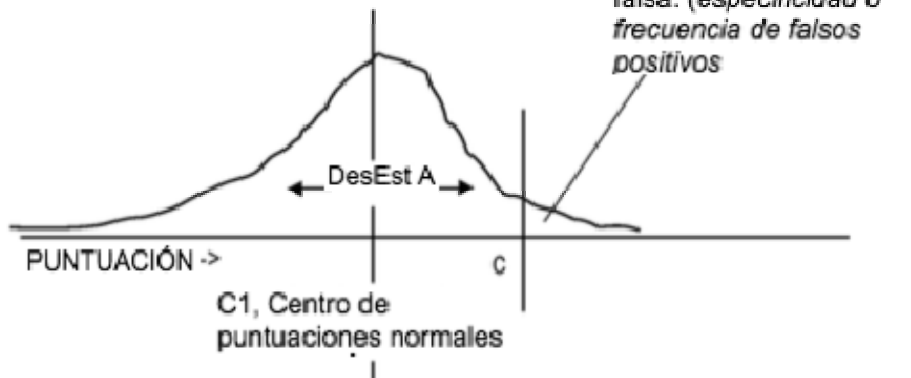


FIG. 19A

← d →

POBLACIÓN DE PUNTUACIONES
"ANEUPLOIDÍA", DISTRIBUCIÓN
ESTIMADA f2

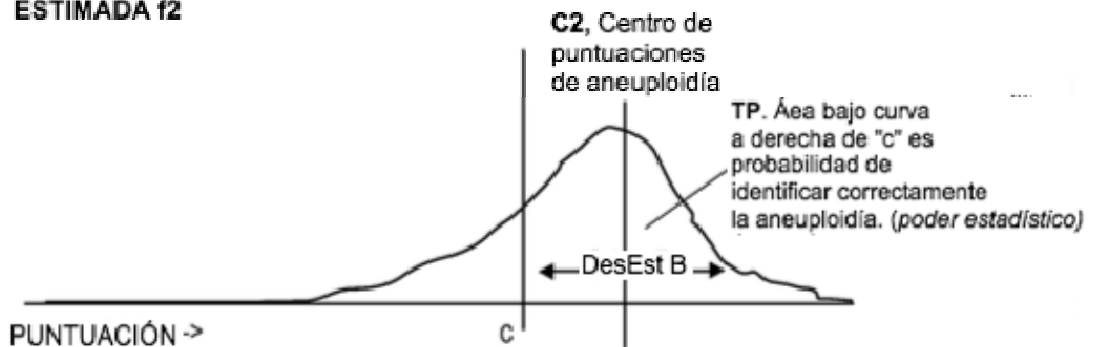


FIG. 19B

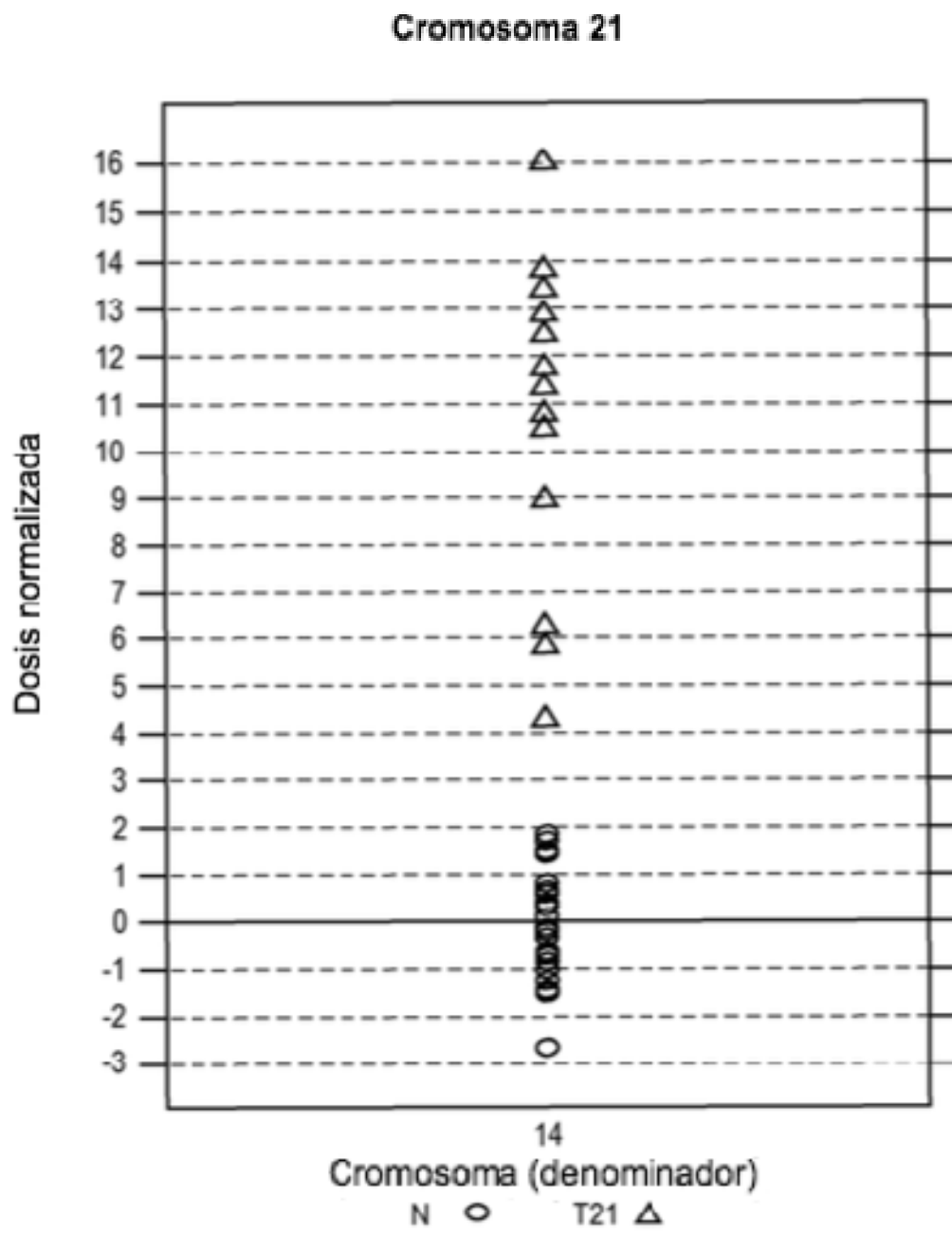


FIG. 20A

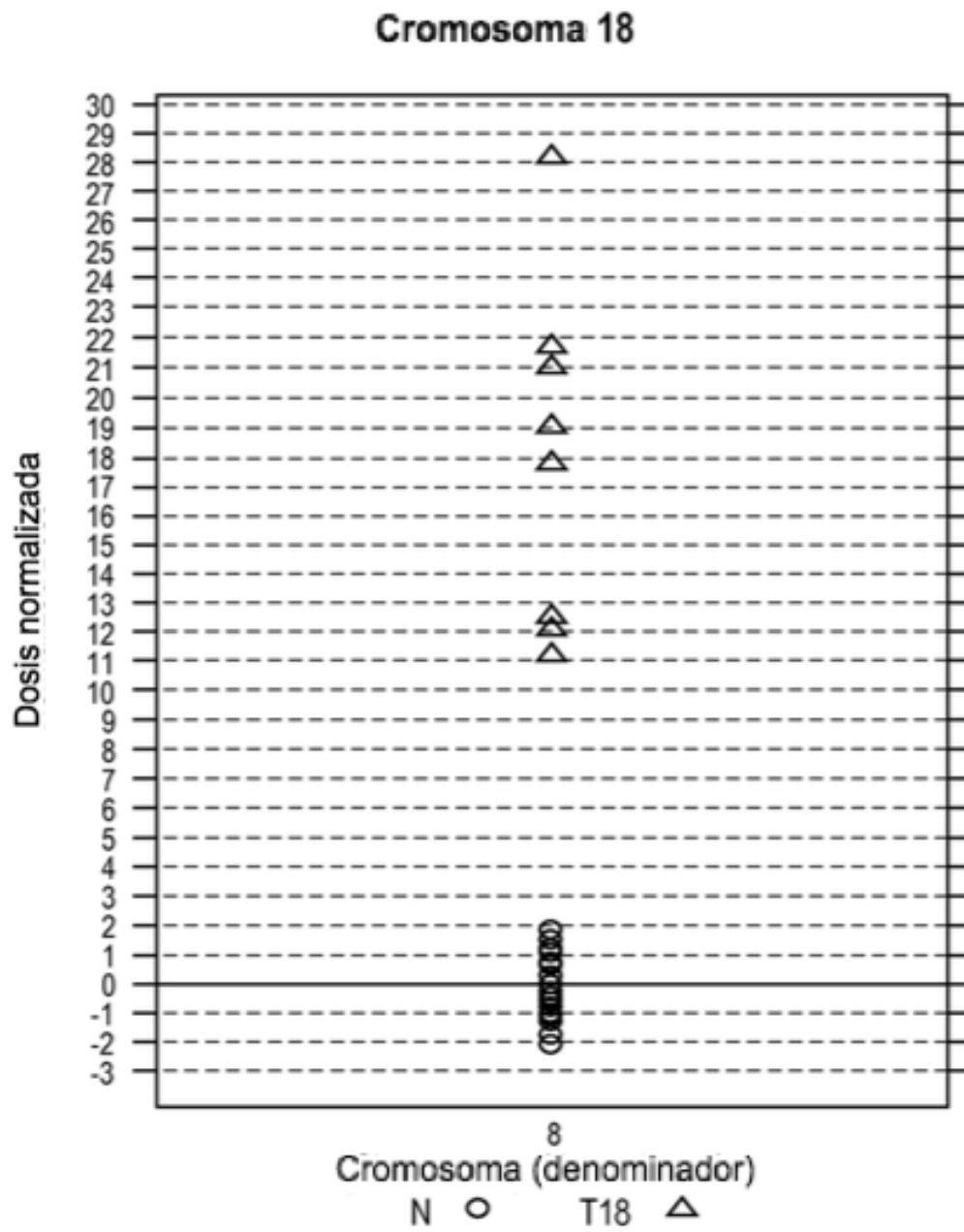


FIG. 20B

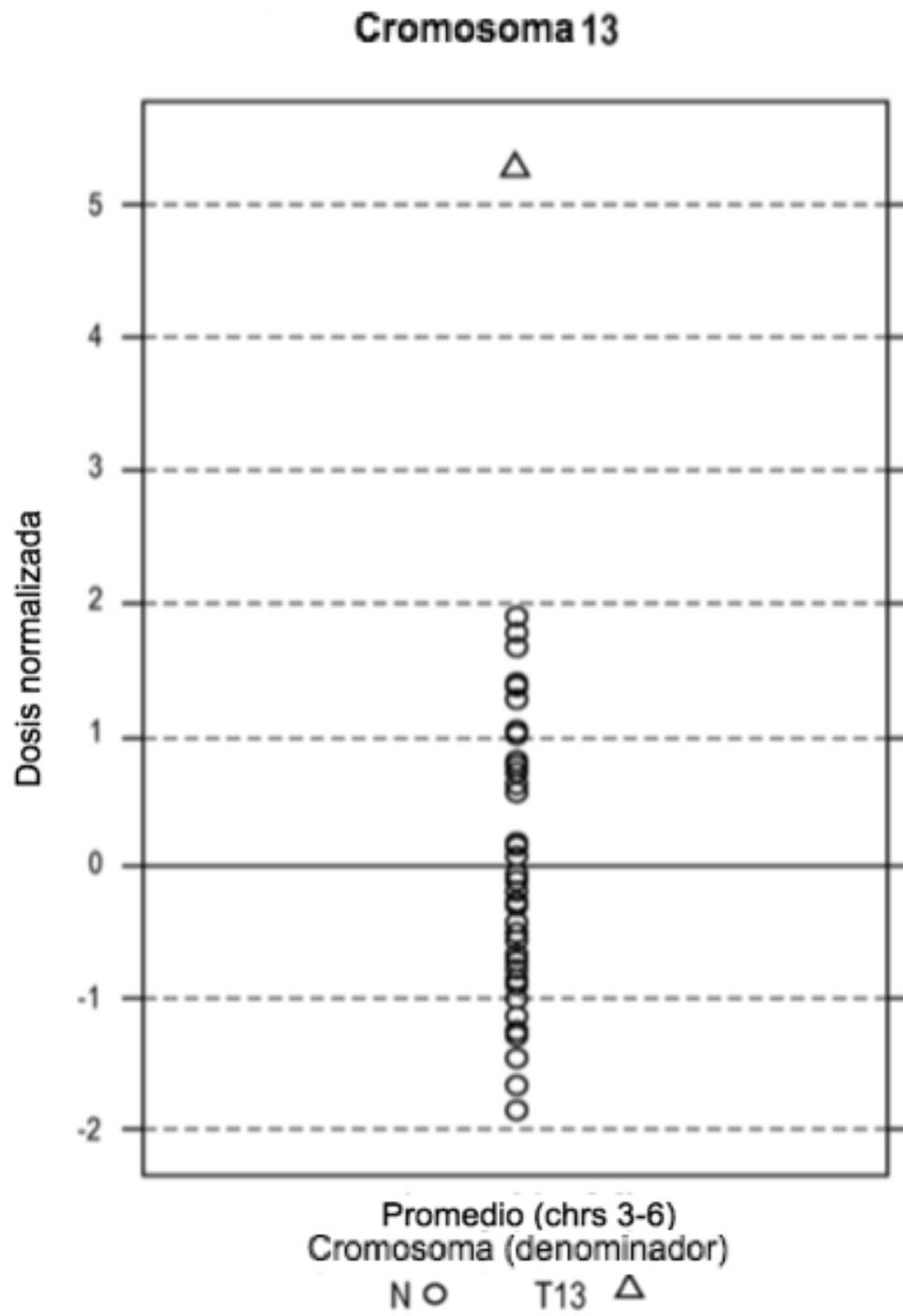


FIG. 20C

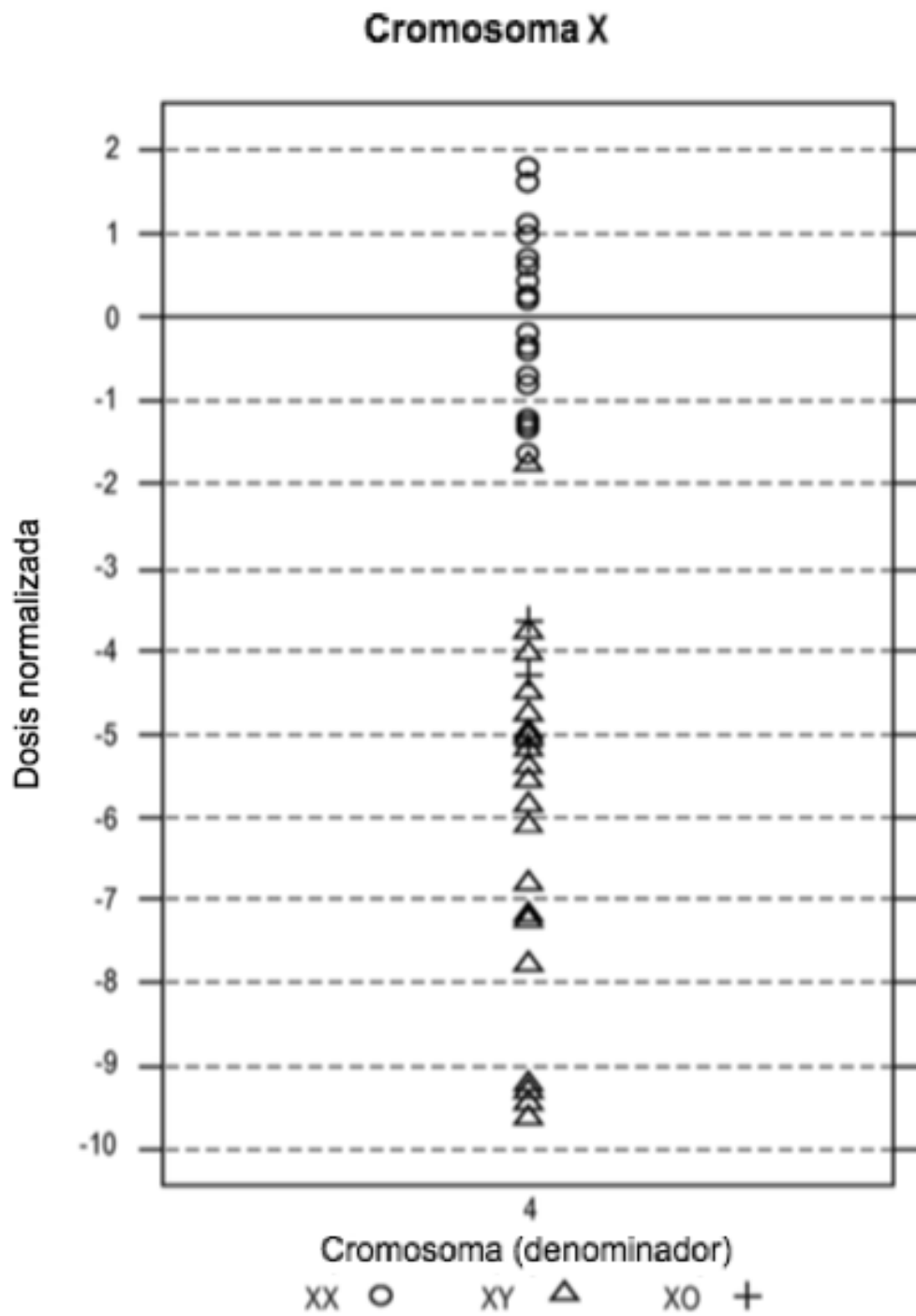


FIG. 20D

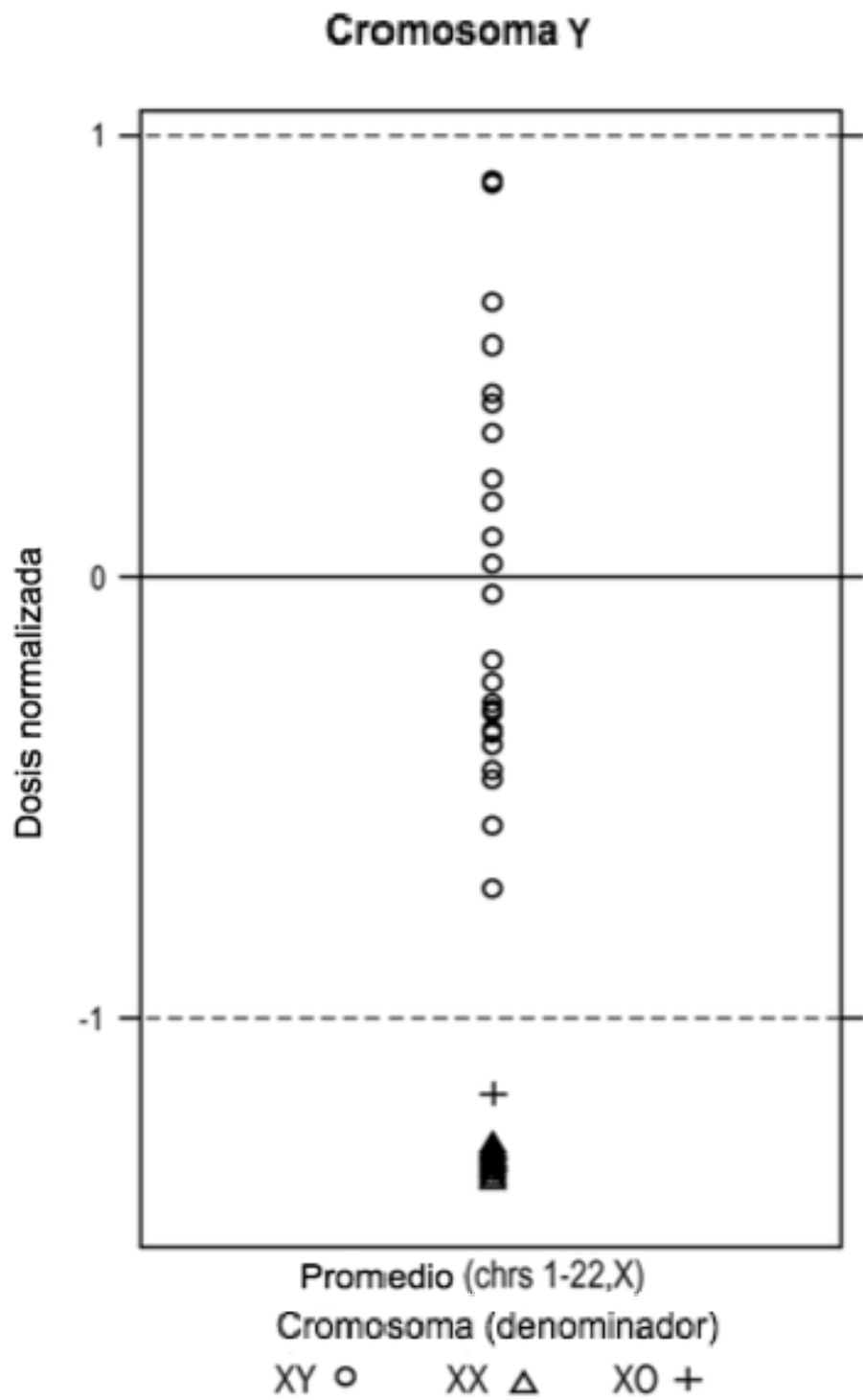


FIG. 20E