

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 909 841**

51 Int. Cl.:

**C12Q 1/68** (2008.01)

**C12P 19/34** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **01.12.2010 E 18201917 (4)**

97 Fecha y número de publicación de la concesión europea: **05.01.2022 EP 3492601**

54 Título: **Nuevo protocolo para preparar 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 45584910 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**10.05.2022**

73 Titular/es:

**VERINATA HEALTH, INC. (100.0%)**  
**5200 Illumina Way**  
**San Diego, California 92122, US**

72 Inventor/es:

**RAVA, RICHARD P.;**  
**CHINNAPPA, MANJULA;**  
**COMSTOCK, DAVID A.;**  
**HEILEK, GABRIELLE y**  
**RHEES, BRIAN KENT**

74 Agente/Representante:

**LEHMANN NOVO, María Isabel**

**ES 2 909 841 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 para preparar bibliotecas de secuenciación

5 **1. Antecedentes de la invención**

La detección y el diagnóstico prenatales son una parte rutinaria de la atención prenatal. Actualmente, el diagnóstico prenatal de afecciones genéticas y cromosómicas implica pruebas invasivas, tales 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 ~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 declarar efectivamente el final de los procedimientos invasivos.

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

El advenimiento de tecnologías de secuenciación de próxima generación (NGS), que permiten la secuenciación de genomas enteros en un tiempo relativamente corto, ha proporcionado la oportunidad de comparar el material genético procedente de un cromosoma que se desea comparar con el de otro sin los riesgos asociados con métodos de muestreo invasivos. No obstante, 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 todas de entre especificidad, sensibilidad y aplicabilidad para diagnosticar de forma fiable aneuploidías fetales en una diversidad de entornos clínicos.

A medida que la secuenciación de ácidos nucleicos ha entrado en el ámbito clínico para pruebas de cáncer, organizaciones tales como el NCCLS (National Council Of Clinical Laboratory Services) y la Association of Clinical Cytogenetics han proporcionado directrices para la normalización de los ensayos basados en secuenciación existentes que utilizan secuenciación basada en PCR, por terminador didesoxi y por extensión de cebador realizada en secuenciadores basados en gel o en capilares (NCCLS: Nucleic Acid Sequencing Methods in Diagnostic Laboratory Medicine MM9-A, Vol. 24 N° 40), secuenciación de Sanger y QF-PCR (Association for Clinical Cytogenetics and Clinical Molecular Genetics Society, Practice Guidelines for Sanger Sequencing Analysis and Interpretation ratificadas por el CMGS Executive Committee del 7 de agosto de 2009, disponible en la dirección de Internet [cmgs.org/BPGs/pdfs%20current%20bpgs/Sequencingv2.pdf](http://cmgs.org/BPGs/pdfs%20current%20bpgs/Sequencingv2.pdf) QF-PCR for the diagnosis of aneuploidy best practice guidelines (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, preservando al mismo tiempo la calidad y la fiabilidad de los ensayos. Dado que los laboratorios clínicos ya están experimentando con NIPD, se desarrollarán procedimientos de calidad para implementar las nuevas tecnologías de secuenciación con el fin de proporcionar sistemas de atención médica seguros y apropiados.

Fan et al. (PNAS 105, 16266-16271, 2008) describen el diagnóstico no invasivo de aneuploidía fetal mediante la secuenciación de escopeta de ADN procedente de sangre materna, y los autores utilizaron un protocolo de preparación de bibliotecas de Solexa/Illumina. Chu et al. (Bioinformatics, 25 (10), mayo de 2009, páginas 1244-1250) describe un "Modelo estadístico para la secuenciación de genoma completo y su aplicación a un diagnóstico mínimamente invasivo de enfermedad genética fetal".

La presente invención se refiere a métodos de secuenciación de próxima generación fiables que se pueden aplicar al menos para la práctica de diagnóstico prenatal no invasivo, y abarca procedimientos que aumentan la rapidez y la calidad de los métodos a la vez que minimizan la pérdida de material y reducen la probabilidad de errores de muestra.

55 **2. Sumario de la invención**

La invención se define en las reivindicaciones adjuntas.

La invención proporciona un método para la secuenciación de ácidos nucleicos que comprende: (a) proporcionar una muestra de ensayo que comprende moléculas de ácido nucleico, en el que dichas moléculas de ácido nucleico son moléculas de ADN genómico humano; (b) realizar la reparación de extremos de las moléculas de ácido nucleico para generar ácidos nucleicos de extremos romos; (c) realizar la adición de colas de dA a los ácidos nucleicos de extremos romos para generar ácidos nucleicos con cola de dA; (d) ligar adaptadores a los ácidos nucleicos con cola de dA para generar una biblioteca de polinucleótidos ligados a adaptadores; (e) opcionalmente amplificar la biblioteca usando

cebadores de amplificación, comprendiendo dichos cebadores de amplificación una porción específica de adaptador; y (f) someter la biblioteca a una secuenciación masivamente paralela; en el que las etapas (b), (c) y (d) son etapas consecutivas.

5 El método de la invención es aplicable a métodos para determinar aneuploidía y/o la fracción fetal en muestras maternas que comprenden ADNcf fetal y materno mediante secuenciación masivamente paralela. El método de la invención comprende un protocolo novedoso para preparar bibliotecas de secuenciación que mejora inesperadamente la calidad del ADN de la biblioteca a la vez que agiliza el proceso de análisis de muestras para diagnósticos prenatales.

10 En un caso, se divulga 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 el que el método comprende: (a) preparar una biblioteca de secuenciación a partir de la mezcla de moléculas de ácidos nucleicos fetales y maternos; en el que la preparación de dicha biblioteca comprende las etapas consecutivas de reparación de extremos, adición de colas de dA y ligación de adaptadores 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 ácidos nucleicos fetales y maternos de una muestra de sangre materna; (c) utilizar 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, e identificar así la presencia o la ausencia de aneuploidía fetal.

20 En otro caso, se divulga 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 el que el método comprende: (a) preparar una biblioteca de secuenciación a partir de la mezcla de moléculas de ácidos nucleicos fetales y maternos; en el que la preparación de dicha biblioteca comprende las etapas consecutivas de reparación de extremos, formación de cola dA y ligadura de adaptadores 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 ácidos nucleicos fetales y maternos de una muestra de sangre materna; (c) utilizar la información de secuencia para obtener una dosis cromosómica para un cromosoma aneuploide; y (d) comparar la dosis cromosómica con al menos un valor umbral, e identificar así la presencia o la ausencia de aneuploidía fetal. El método comprende además utilizar la información de secuencia para identificar un número de etiquetas de secuencia mapeadas para al menos un cromosoma de normalización y para un cromosoma aneuploide; y utilizar el número de etiquetas de secuencia mapeadas identificadas para dicho cromosoma aneuploide y el número de etiquetas de secuencia mapeadas identificadas para el, al menos un, cromosoma de normalización para calcular una dosis de cromosoma para dicho cromosoma aneuploide como una relación del número de etiquetas de secuencia mapeadas identificadas para dicho cromosoma aneuploide y el número de etiquetas de secuencia mapeadas identificadas para el, al menos un, cromosoma de normalización. Opcionalmente, calcular la dosis de cromosoma comprende (i) calcular una relación de densidad de etiquetas de secuencia para el cromosoma aneuploide, relacionando el número de etiquetas de secuencia mapeadas identificadas para el cromosoma aneuploide en la etapa con la longitud de dicho cromosoma aneuploide; (ii) calcular una relación de densidad de etiquetas de secuencia para el, al menos un, cromosoma de normalización, relacionando el número de etiquetas de secuencia mapeadas identificadas para dicho, al menos un, cromosoma de normalización con la longitud del, al menos un, cromosoma de normalización; y (iii) utilizar las relaciones de densidad de etiquetas de secuencia calculadas en las etapas (i) y (ii) para calcular una dosis de cromosoma para el cromosoma aneuploide, calculándose la dosis de cromosoma como la relación entre la relación de densidad de etiquetas de secuencia para el cromosoma aneuploide y la relación de densidad de etiquetas de secuencia para el, al menos un, cromosoma de normalización.

45 En otro caso, se divulga 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 el que el método comprende: (a) preparar una biblioteca de secuenciación a partir de la mezcla de moléculas de ácidos nucleicos fetales y maternos; en el que la preparación de dicha biblioteca comprende las etapas consecutivas de reparación de extremos, formación de cola dA y ligadura de adaptadores 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 ácidos nucleicos fetales y maternos de una muestra de sangre materna; (c) utilizar la información de secuencia para obtener una dosis cromosómica para un cromosoma aneuploide; y (d) comparar la dosis cromosómica con al menos un valor umbral, e identificar así la presencia o la ausencia de aneuploidía fetal. El método comprende además utilizar la información de secuencia para identificar un número de etiquetas de secuencia mapeadas para al menos un cromosoma de normalización y para un cromosoma aneuploide; y utilizar el número de etiquetas de secuencia mapeadas identificadas para dicho cromosoma aneuploide y el número de etiquetas de secuencia mapeadas identificadas para el, al menos un, cromosoma de normalización para calcular una dosis de cromosoma para dicho cromosoma aneuploide como una relación del número de etiquetas de secuencia mapeadas identificadas para dicho cromosoma aneuploide y el número de etiquetas de secuencia mapeadas identificadas para el, al menos un, cromosoma de normalización. El, al menos un, cromosoma de normalización es un cromosoma que posee la variabilidad más reducida y/o la diferenciabilidad más elevada. Opcionalmente, calcular la dosis de cromosoma comprende (i) calcular una relación de densidad de etiquetas de secuencia para el cromosoma aneuploide, relacionando el número de etiquetas de secuencia mapeadas identificadas para el cromosoma aneuploide en la etapa con la longitud de dicho cromosoma aneuploide; (ii) calcular una relación de densidad de etiquetas de secuencia para el, al menos un, cromosoma de normalización, relacionando el número de etiquetas de secuencia mapeadas

identificadas para dicho, al menos un, cromosoma de normalización con la longitud del, al menos un, cromosoma de normalización; y (iii) utilizar las relaciones de densidad de etiquetas de secuencia calculadas en las etapas (i) y (ii) para calcular una dosis de cromosoma para el cromosoma aneuploide, calculándose la dosis de cromosoma como la relación entre la relación de densidad de etiquetas de secuencia para el cromosoma aneuploide y la relación de densidad de etiquetas de secuencia para el, al menos un, cromosoma de normalización.

En otro caso, se divulga 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 el que el método comprende: (a) preparar una biblioteca de secuenciación a partir de la mezcla de moléculas de ácidos nucleicos fetales y maternos; en el que la preparación de dicha biblioteca comprende las etapas consecutivas de reparación de extremos, adición de colas de dA y ligación de adaptadores 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 ácidos nucleicos fetales y maternos de una muestra de sangre materna; (c) utilizar 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, e identificar así la presencia o la ausencia de aneuploidía fetal. El método comprende además utilizar la información de secuencia para identificar un número de etiquetas de secuencia mapeadas para al menos un cromosoma de normalización y para un cromosoma aneuploide; y utilizar el número de etiquetas de secuencia mapeadas identificadas para dicho cromosoma aneuploide y el número de etiquetas de secuencia mapeadas identificadas para el, al menos un, cromosoma de normalización para calcular una dosis de cromosoma para dicho cromosoma aneuploide como una relación del número de etiquetas de secuencia mapeadas identificadas para dicho cromosoma aneuploide y el número de etiquetas de secuencia mapeadas identificadas para el, al menos un, cromosoma de normalización. Opcionalmente, calcular la dosis de cromosoma comprende (i) calcular una relación de densidad de etiquetas de secuencia para el cromosoma aneuploide, relacionando el número de etiquetas de secuencia mapeadas identificadas para el cromosoma aneuploide en la etapa con la longitud de dicho cromosoma aneuploide; (ii) calcular una relación de densidad de etiquetas de secuencia para el, al menos un, cromosoma de normalización, relacionando el número de etiquetas de secuencia mapeadas identificadas para dicho, al menos un, cromosoma de normalización con la longitud del, al menos un, cromosoma de normalización; y (iii) utilizar las relaciones de densidad de etiquetas de secuencia calculadas en las etapas (i) y (ii) para calcular una dosis de cromosoma para el cromosoma aneuploide, calculándose la dosis de cromosoma como la relación entre la relación de densidad de etiquetas de secuencia para el cromosoma aneuploide y la relación de densidad de etiquetas de secuencia para el, al menos un, cromosoma de normalización. En los casos en los que el cromosoma aneuploide es el cromosoma 21, el, al menos un, cromosoma de normalización se selecciona de 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 de normalización para el cromosoma 21 es un grupo de cromosomas seleccionados de entre el cromosoma 9, el cromosoma 1, el cromosoma 2, el cromosoma 11, el cromosoma 12 y el cromosoma 14. En los casos en los que el cromosoma aneuploide es el cromosoma 18, el, al menos un, cromosoma de normalización se selecciona de entre el cromosoma 8, el cromosoma 2, el cromosoma 3, el cromosoma 5, el cromosoma 6, el cromosoma 12 y el cromosoma 14. Alternativamente, el, al menos un, cromosoma de normalización para el cromosoma 18 es un grupo de cromosomas seleccionados de entre el cromosoma 8, el cromosoma 2, el cromosoma 3, el cromosoma 5, el cromosoma 6, el cromosoma 12 y el cromosoma 14. En los casos en los que el cromosoma aneuploide es el cromosoma 13, el, al menos un, cromosoma de normalización se selecciona de 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 de normalización para el cromosoma 13 es un grupo de cromosomas seleccionados de entre el cromosoma 2, el cromosoma 3, el cromosoma 4, el cromosoma 5, el cromosoma 6 y el cromosoma 8. En los casos en los que el cromosoma aneuploide es el cromosoma X, el, al menos un, cromosoma de normalización se selecciona de 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 de normalización para el cromosoma X es un grupo de cromosomas seleccionados de entre el cromosoma 2, el cromosoma 3, el cromosoma 4, el cromosoma 5, el cromosoma 6 y el cromosoma 8.

La muestra materna utilizada en los casos del método para determinar una aneuploidía cromosómica fetal es un fluido biológico seleccionado de entre sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma. En algunos casos, las moléculas de ácido nucleico comprendidas en la muestra materna son moléculas de ADN libre de células. En algunos casos, las etapas consecutivas comprendidas en la preparación de la biblioteca de secuenciación se realizan en menos de una hora. Preferentemente, las etapas consecutivas se realizan en ausencia de polietilenglicol. De forma más preferida, las etapas consecutivas excluyen una purificación. La secuenciación de la biblioteca de secuenciación se realiza mediante métodos de secuenciación de próxima generación (NGS). En algunos casos, la secuenciación comprende una amplificación. En otros casos, la secuenciación es una secuenciación masivamente paralela que utiliza secuenciación por síntesis con terminadores de colorantes reversibles. En otros casos, la secuenciación es una secuenciación masivamente paralela que utiliza secuenciación por ligación. En otros casos más, la secuenciación es secuenciación de una sola molécula.

En otro caso, se divulga en el presente documento un método para determinar la presencia o la ausencia de una aneuploidía en una muestra de sangre materna que comprende una mezcla de moléculas de ácidos nucleicos fetales y maternos, comprendiendo el método: (a) preparar una biblioteca de secuenciación a partir de la mezcla; en el que la preparación de dicha biblioteca comprende las etapas consecutivas de reparación de extremos, adición de colas de dA y ligación de adaptadores a dichos ácidos nucleicos fetales y maternos; (b) secuenciar al menos una porción de la

biblioteca de secuenciación, en el que la secuenciación comprende proporcionar una pluralidad de etiquetas de secuencia; y (c) basándose en la secuenciación, determinar la presencia o la ausencia de aneuploidía en la muestra.

5 En otro caso, se divulga en el presente documento un método para determinar la presencia o la ausencia de una aneuploidía cromosómica o parcial en una muestra de sangre materna que comprende una mezcla de de ácidos nucleicos fetales y maternos, en el que el método comprende: (a) preparar una biblioteca de secuenciación a partir de la mezcla; en el que la preparación de dicha biblioteca comprende las etapas consecutivas de reparación de extremos, adición de colas de dA y ligación de adaptadores a dichos ácidos nucleicos fetales y maternos; (b) secuenciar al menos una porción de la biblioteca de secuenciación, en el que la secuenciación comprende proporcionar una pluralidad de  
10 etiquetas de secuencia; y (c) basándose en la secuenciación, determinar la presencia o la ausencia de la aneuploidía cromosómica o parcial en la muestra.

15 En otro caso, se divulga en el presente documento un método para determinar la presencia o la ausencia de una aneuploidía cromosómica en una muestra de sangre materna que comprende una mezcla de de ácidos nucleicos fetales y maternos, en el que el método comprende: (a) preparar una biblioteca de secuenciación a partir de la mezcla; en el que la preparación de dicha biblioteca comprende las etapas consecutivas de reparación de extremos, adición de colas de dA y ligación de adaptadores a dichos ácidos nucleicos fetales y maternos; (b) secuenciar al menos una porción de la biblioteca de secuenciación, en el que la secuenciación comprende proporcionar una pluralidad de  
20 etiquetas de secuencia; y (c) basándose en la secuenciación, determinar la presencia o la ausencia de la aneuploidía cromosómica en la muestra. Las aneuploidías cromosómicas que pueden determinarse según el método incluyen trisomía 8, trisomía 13, trisomía 15, trisomía 16, trisomía 18, trisomía 21, trisomía 22, monosomía X y XXX.

25 En otro caso, se divulga en el presente documento un método para determinar la presencia o la ausencia de una aneuploidía cromosómica o parcial en una muestra de sangre materna que comprende una mezcla de de ácidos nucleicos fetales y maternos, en el que el método comprende: (a) preparar una biblioteca de secuenciación a partir de la mezcla; en el que la preparación de dicha biblioteca comprende las etapas consecutivas de reparación de extremos, adición de colas de dA y ligación de adaptadores a dichos ácidos nucleicos fetales y maternos; (b) secuenciar al menos una porción de la biblioteca de secuenciación, en el que la secuenciación comprende proporcionar una pluralidad de  
30 etiquetas de secuencia; y (c) basándose en la secuenciación, determinar la presencia o la ausencia de la aneuploidía cromosómica o parcial en la muestra que comprende calcular una dosis de cromosoma basada en el número de dichas etiquetas de secuencia para un cromosoma de interés y para un cromosoma de normalización, y comparar dicha dosis con un valor umbral.

35 En otro caso, se divulga en el presente documento un método para determinar la presencia o la ausencia de una aneuploidía cromosómica en una muestra de sangre materna que comprende una mezcla de de ácidos nucleicos fetales y maternos, en el que el método comprende: (a) preparar una biblioteca de secuenciación a partir de la mezcla; en el que la preparación de dicha biblioteca comprende las etapas consecutivas de reparación de extremos, adición de colas de dA y ligación de adaptadores a dichos ácidos nucleicos fetales y maternos; (b) secuenciar al menos una porción de la biblioteca de secuenciación, en el que la secuenciación comprende proporcionar una pluralidad de  
40 etiquetas de secuencia; y (c) basándose en la secuenciación, determinar la presencia o la ausencia de la aneuploidía cromosómica en la muestra que comprende calcular una dosis de cromosoma basada en el número de dichas etiquetas de secuencia para un cromosoma de interés y para un cromosoma de normalización, y comparar dicha dosis con un valor umbral. Las aneuploidías cromosómicas que pueden determinarse según el método incluyen trisomía 8, trisomía 13, trisomía 15, trisomía 16, trisomía 18, trisomía 21, trisomía 22, monosomía X y XXX

45 La muestra materna utilizada en los casos del método para determinar la presencia o la ausencia de una aneuploidía es un fluido biológico seleccionado de entre sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma. En algunos casos, las moléculas de ácidos nucleicos comprendidas en la muestra materna son moléculas de ADN desprovistas de células. En algunos casos, las etapas consecutivas comprendidas en la preparación de la biblioteca de secuenciación se realizan en menos de una hora. Preferentemente, las etapas consecutivas se realizan en ausencia de polietilenglicol. De forma más preferida, las etapas consecutivas excluyen una purificación. La secuenciación de la biblioteca de secuenciación se realiza mediante métodos de secuenciación de próxima generación (NGS). En algunos casos, la secuenciación comprende una amplificación. En otros casos, la secuenciación es una secuenciación masivamente paralela que utiliza secuenciación por síntesis con terminadores de colorantes reversibles. En otros casos, la secuenciación es una secuenciación masivamente paralela que utiliza  
50 secuenciación por ligación. En otros casos más, la secuenciación es secuenciación de una sola molécula.

55 En otro caso, se divulga en el presente documento un método para determinar la fracción de moléculas de ácido nucleico fetal en una muestra de sangre materna que comprende una mezcla de moléculas de ácidos nucleicos fetales y maternos, en el que 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 el que la preparación de la biblioteca comprende las etapas consecutivas de reparación de extremos, adición de colas de dA y ligación de adaptadores a dichas moléculas de ácidos nucleicos fetales y maternos; (c) secuenciar al menos una porción de la biblioteca de secuenciación; y (d) basándose en dicha secuenciación, determinar la fracción de las moléculas de ácido nucleico fetal. Opcionalmente, el método puede comprender también la determinación de la presencia o la ausencia de aneuploidía en una muestra materna.

En otro caso, se divulga en el presente documento un método para determinar la fracción de moléculas de ácido nucleico fetal en una muestra de sangre materna que comprende una mezcla de moléculas de ácidos nucleicos fetales y maternos, en el que 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 el que la preparación de la biblioteca comprende las etapas consecutivas de reparación de extremos, formación de cola dA y ligación de adaptadores a dichas moléculas de ácidos nucleicos fetales y maternos; (c) secuenciar al menos una porción de la biblioteca de secuenciación; y (d) basándose en dicha secuenciación, determinar la fracción de las moléculas de ácido nucleico fetal. Determinar la fracción comprende determinar el número de etiquetas de secuencia maternas y fetales mapeadas a un genoma diana de referencia que comprende al menos un ácido nucleico polimórfico. Opcionalmente, el método puede comprender también la determinación de la presencia o la ausencia de aneuploidía en una muestra materna.

En otro caso, se divulga en el presente documento un método para determinar la fracción de moléculas de ácido nucleico fetal en una muestra de sangre materna que comprende una mezcla de moléculas de ácidos nucleicos fetales y maternos, en el que el método comprende: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en una porción de la mezcla, en la que cada uno de dicha pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un solo nucleótido (SNP); (b) preparar una biblioteca de secuenciación del producto amplificado obtenido en la etapa (a) en el que la preparación de la biblioteca comprende las etapas consecutivas de reparación de extremos, adición de colas de dA y ligación de adaptadores a dichas moléculas de ácidos nucleicos fetales y maternos; (c) secuenciar al menos una porción de la biblioteca de secuenciación; y (d) basándose en dicha secuenciación, determinar la fracción de las moléculas de ácido nucleico fetal. Opcionalmente, el método puede comprender también la determinación de la presencia o la ausencia de aneuploidía en una muestra materna.

En otro caso, se divulga en el presente documento un método para determinar la fracción de moléculas de ácido nucleico fetal en una muestra de sangre materna que comprende una mezcla de moléculas de ácidos nucleicos fetales y maternos, en el que el método comprende: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en una porción de la mezcla, en el que cada uno de dicha pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un solo nucleótido (SNP); (b) preparar una biblioteca de secuenciación del producto amplificado obtenido en la etapa (a) en el que la preparación de la biblioteca comprende las etapas consecutivas de reparación de extremos, formación de cola dA y ligación de adaptadores a dichas moléculas de ácidos nucleicos fetales y maternos; (c) secuenciar al menos una porción de la biblioteca de secuenciación; y (d) basándose en dicha secuenciación, determinar la fracción de las moléculas de ácido nucleico fetal. Determinar la fracción comprende determinar el número de etiquetas de secuencia maternas y fetales mapeadas a un genoma diana de referencia que comprende al menos un ácido nucleico polimórfico. Opcionalmente, el método puede comprender también la determinación de la presencia o la ausencia de aneuploidía en una muestra materna.

En otro caso, se divulga en el presente documento un método para determinar la fracción de moléculas de ácido nucleico fetal en una muestra de sangre materna que comprende una mezcla de moléculas de ácidos nucleicos fetales y maternos, en el que el método comprende: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en una porción de la mezcla, en la que cada uno de dicha pluralidad de ácidos nucleicos diana polimórficos comprende al menos una repetición en tándem corta (STR); (b) preparar una biblioteca de secuenciación del producto amplificado obtenido en la etapa (a) en el que la preparación de la biblioteca comprende las etapas consecutivas de reparación de extremos, adición de colas de dA y ligación de adaptadores a dichas moléculas de ácidos nucleicos fetales y maternos; (c) secuenciar al menos una porción de la biblioteca de secuenciación; y (d) basándose en dicha secuenciación, determinar la fracción de las moléculas de ácido nucleico fetal. Opcionalmente, el método puede comprender también la determinación de la presencia o la ausencia de aneuploidía en una muestra materna.

En otro caso, se divulga en el presente documento un método para determinar la fracción de moléculas de ácido nucleico fetal en una muestra de sangre materna que comprende una mezcla de moléculas de ácidos nucleicos fetales y maternos, en el que el método comprende: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en una porción de la mezcla, en el que cada uno de dicha pluralidad de ácidos nucleicos diana polimórficos comprende al menos una repetición en tándem corta (STR); (b) preparar una biblioteca de secuenciación del producto amplificado obtenido en la etapa (a) en el que la preparación de la biblioteca comprende las etapas consecutivas de reparación de extremos, formación de cola dA y ligación de adaptadores a dichas moléculas de ácidos nucleicos fetales y maternos; (c) secuenciar al menos una porción de la biblioteca de secuenciación; y (d) basándose en dicha secuenciación, determinar la fracción de las moléculas de ácido nucleico fetal. Determinar la fracción comprende determinar el número de etiquetas de secuencia maternas y fetales mapeadas a un genoma diana de referencia que comprende al menos un ácido nucleico polimórfico. Opcionalmente, el método puede comprender también la determinación de la presencia o la ausencia de aneuploidía en una muestra materna.

En otro caso, se divulga en el presente documento un método para determinar la fracción de moléculas de ácido nucleico fetal en una muestra de sangre materna que comprende una mezcla de moléculas de ácidos nucleicos fetales y maternos, en el que el método comprende: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en una porción de la mezcla, en la que cada uno de la pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un nucleótido (SNP); (b) preparar una biblioteca de secuenciación del producto amplificado

obtenido en la etapa (a) en el que la preparación de la biblioteca comprende las etapas consecutivas de reparación de extremos, adición de colas de dA y ligación de adaptadores a dichas moléculas de ácidos nucleicos fetales y maternos; (c) secuenciar al menos una porción de la biblioteca de secuenciación; y (d) basándose en dicha secuenciación, determinar la fracción de las moléculas de ácido nucleico fetal. En los métodos en los 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 los métodos en los que cada uno de la pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un nucleótido (SNP), el, al menos un, SNP es un SNP en tándem seleccionado de los 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 también la determinación de la presencia o la ausencia de aneuploidía en una muestra materna.

En otro caso, se divulga en el presente documento un método para determinar la fracción de moléculas de ácido nucleico fetal en una muestra de sangre materna que comprende una mezcla de moléculas de ácidos nucleicos fetales y maternos, en el que el método comprende: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en una porción de la mezcla, en el que cada uno de dicha pluralidad de ácidos nucleicos diana polimórficos comprende al menos una repetición en tándem corta (STR); (b) preparar una biblioteca de secuenciación del producto amplificado obtenido en la etapa (a) en el que la preparación de la biblioteca comprende las etapas consecutivas de reparación de extremos, formación de cola dA y ligadura de adaptadores a dichas moléculas de ácidos nucleicos fetales y maternos; (c) secuenciar al menos una porción de la biblioteca de secuenciación; y (d) basándose en dicha secuenciación, determinar la fracción de las moléculas de ácido nucleico fetal. El, al menos un, STR se selecciona de entre CSF1PO, FGA, TH01, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, D2S1338, PentaD, PentaE, 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 también la determinación de la presencia o la ausencia de aneuploidía en una muestra materna.

En otro caso, se divulga en el presente documento un método para determinar la fracción de moléculas de ácido nucleico fetal en una muestra de sangre materna que comprende una mezcla de moléculas de ácidos nucleicos fetales y maternos, en el que el método comprende: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en una porción de la mezcla, en el que cada uno de la pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un nucleótido (SNP); (b) preparar una biblioteca de secuenciación del producto amplificado obtenido en la etapa (a) en el que la preparación de la biblioteca comprende las etapas consecutivas de reparación de extremos, formación de cola dA y ligación de adaptadores a dichas moléculas de ácidos nucleicos fetales y maternos; (c) secuenciar al menos una porción de la biblioteca de secuenciación; y (d) basándose en dicha secuenciación, determinar la fracción de las moléculas de ácido nucleico fetal. Determinar la fracción comprende determinar el número de etiquetas de secuencia maternas y fetales mapeadas a un genoma diana de referencia que comprende al menos un ácido nucleico polimórfico. En los métodos en los 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 los métodos en los que cada uno de la pluralidad de ácidos nucleicos diana polimórficos comprende al menos un polimorfismo de un solo nucleótido (SNP), el, al menos un, SNP es un SNP en tándem seleccionado de los 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 también la determinación de la presencia o la ausencia de aneuploidía en una muestra materna.

5 En otro caso, se divulga en el presente documento un método para determinar la fracción de moléculas de ácido nucleico fetal en una muestra de sangre materna que comprende una mezcla de moléculas de ácidos nucleicos fetales y maternos, en el que el método comprende: (a) amplificar una pluralidad de ácidos nucleicos diana polimórficos en una porción de la mezcla, en el que cada uno de dicha pluralidad de ácidos nucleicos diana polimórficos comprende al menos una repetición en tándem corta (STR); (b) preparar una biblioteca de secuenciación del producto amplificado obtenido en la etapa (a) en el que la preparación de la biblioteca comprende las etapas consecutivas de reparación de extremos, formación de cola dA y ligación de adaptadores a dichas moléculas de ácidos nucleicos fetales y maternos; (c) secuenciar al menos una porción de la biblioteca de secuenciación; y (d) basándose en dicha secuenciación, determinar la fracción de las moléculas de ácido nucleico fetal. Determinar la fracción comprende determinar el número de etiquetas de secuencia maternas y fetales mapeadas a un genoma diana de referencia que comprende al menos un ácido nucleico polimórfico. El, al menos un, STR se selecciona de entre CSF1PO, FGA, TH01, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, D2S1338, PentaD, PentaE, 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 también la determinación de la presencia o la ausencia de aneuploidía en una muestra materna.

La muestra materna utilizada en el método para determinar la fracción de moléculas de ácido nucleico fetal es un fluido biológico seleccionado de entre sangre, plasma, suero, orina y saliva. Preferentemente, la muestra materna es una muestra de plasma. En algunos casos, las moléculas de ácidos nucleicos comprendidas en la muestra materna son moléculas de ADN desprovistas de células. En algunos casos, las etapas consecutivas comprendidas en la preparación de la biblioteca de secuenciación se realizan en menos de una hora. Preferentemente, las etapas consecutivas se realizan en ausencia de polietilenglicol. De forma más preferida, las etapas consecutivas excluyen una purificación. La secuenciación de la biblioteca de secuenciación se realiza mediante métodos de secuenciación de próxima generación (NGS). En algunos casos, la secuenciación comprende una amplificación. En otros casos, la secuenciación es una secuenciación masivamente paralela que utiliza secuenciación por síntesis con terminadores de colorantes reversibles. En otros casos, la secuenciación es una secuenciación masivamente paralela que utiliza secuenciación por ligación. En otros casos más, la secuenciación es secuenciación de una sola molécula.

5 En otro caso, se divulga en el presente documento un medio legible por ordenador que tiene almacenadas instrucciones legibles por ordenador para llevar a cabo el método para determinar la presencia o la 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.

40 En un caso, el medio legible por ordenador tiene almacenadas instrucciones legibles por ordenador para llevar a cabo el método que comprende las etapas siguientes: (a) usar información de secuencia obtenida a partir de una pluralidad de moléculas de ácidos nucleicos fetales y maternos en una muestra de plasma materno para identificar un número de etiquetas de secuencia mapeadas para un cromosoma de interés; (b) usar información de secuencia obtenida a partir de una pluralidad de moléculas de ácidos nucleicos fetales y maternos en una muestra de plasma materno para identificar un número de etiquetas de secuencia mapeadas para al menos un cromosoma de normalización; (c) usar el número de etiquetas de secuencia mapeadas identificadas para un cromosoma de interés en la etapa (a) y el número de etiquetas de secuencia mapeadas identificadas para el, al menos un, cromosoma de normalización en la etapa (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, e identificar así la presencia o la ausencia de aneuploidía fetal. Los cromosomas de interés pueden ser cualquiera de los cromosomas 21, 13, 18 y X.

55 En otro caso, se divulga en el presente documento un sistema de procesamiento informático que está adaptado o configurado para realizar el método para determinar la presencia o la 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.

60 En un caso, el sistema de procesamiento informático está adaptado o configurado para realizar las etapas siguientes: (a) usar información de secuencia obtenida a partir de una pluralidad de moléculas de ácidos nucleicos fetales y maternos en una muestra de plasma materno para identificar un número de etiquetas de secuencia mapeadas para un cromosoma de interés; (b) usar información de secuencia obtenida a partir de una pluralidad de moléculas de ácidos nucleicos fetales y maternos en una muestra de plasma materno para identificar un número de etiquetas de secuencia mapeadas para al menos un cromosoma de normalización; (c) usar el número de etiquetas de secuencia mapeadas identificadas para un cromosoma de interés en la etapa (a) y el número de etiquetas de secuencia mapeadas identificadas para el, al menos un, cromosoma de normalización en la etapa (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,

e identificar así la presencia o la ausencia de aneuploidía fetal. Los cromosomas de interés pueden ser cualquiera de los cromosomas 21, 13, 18 y X.

En otro caso, se divulga en el presente documento un aparato adaptado o configurado para determinar una aneuploidía fetal en una muestra de plasma materno que comprende una mezcla de moléculas de ácidos nucleicos fetales y maternos, y en el que dicho aparato comprende: (a) un dispositivo de secuenciación adaptado o configurado para secuenciar al menos un porción de las moléculas de ácido nucleico en una muestra de plasma materno que comprende moléculas de ácidos nucleicos fetales y maternos, generando así información de secuencia; y (b) un sistema de procesamiento informático configurado para realizar las etapas siguientes: (i) usar 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 información de secuencia generada por el dispositivo de secuenciación para identificar un número de etiquetas de secuencia mapeadas para al menos un cromosoma de normalización; (iii) usar el número de etiquetas de secuencia mapeadas identificadas para un cromosoma de interés en la etapa (i) y el número de etiquetas de secuencia mapeadas identificadas para el, al menos un, cromosoma de normalización en la etapa (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, e identificar así la presencia o la ausencia de aneuploidía fetal. Los cromosomas de interés pueden ser cualquiera de los cromosomas 21, 13, 18 y X.

Aunque los ejemplos del presente documento se refieren a seres humanos y el lenguaje se refiere principalmente a cuestiones humanas, el concepto de la presente invención puede aplicarse a genomas de cualquier planta o animal.

### 3. Breve descripción de los dibujos

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

La figura 3 es un diagrama de flujo de un método **300** para determinar simultáneamente la presencia o la ausencia de aneuploidía fetal y la fracción fetal en una muestra de ensayo de plasma materno enriquecida en ácidos nucleicos polimórficos.

La figura 4 es un diagrama de flujo de un método **400** para determinar simultáneamente la presencia o la ausencia de aneuploidía fetal y la fracción fetal en una muestra de ensayo de ADNcf purificado materno que se ha enriquecido con ácidos nucleicos polimórficos.

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

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

La figura 7 muestra electroferogramas de una biblioteca de secuenciación de ADNcf preparada según 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 relación del número de etiquetas de secuencia mapeadas a cada cromosoma (eje X) y el número total de etiquetas mapeadas a todos los cromosomas (1-22, X e Y) para la muestra M11281 cuando se preparó la biblioteca utilizando el protocolo abreviado del ejemplo, 2a (**◆**) y cuando se preparó según el protocolo de longitud completa del ejemplo, 2b (**■**). También se muestran las relaciones de etiquetas para la muestra M11297 obtenidas a partir de la secuenciación de una biblioteca preparada según el protocolo abreviado del ejemplo, 2a (**▲**) y según 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 de ADNcf extraído de un conjunto de 48 muestras de sangre obtenidas de sujetos humanos, cada uno embarazado de un feto masculino o femenino. Las dosis de cromosoma 21 para muestras de ensayo calificadas, es decir, normales para el cromosoma 21 (**O**), y trisomía 21 (**Δ**) se muestran 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 de ADNcf extraído de un conjunto de 48 muestras de sangre obtenidas de sujetos humanos, cada uno embarazado de un feto masculino o femenino. Las dosis de cromosoma 18 para muestras de ensayo calificadas, es decir, normales para el cromosoma 18 (**O**), y trisomía 18 (**Δ**) se muestran para los cromosomas 1-12 y X (**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 de ADNcf extraído de un conjunto de 48 muestras de sangre obtenidas de sujetos humanos, cada uno embarazado de un feto masculino o femenino. Las dosis de cromosoma 13 para muestras de ensayo calificadas, es decir, normales para el cromosoma 13 (**O**), y trisomía 13 (**Δ**) se muestran para los cromosomas 1-12 y X (**A**), y para los cromosomas 13-21 y X (**B**).

La figura 12 muestra la distribución de la dosis de cromosoma para el cromosoma X determinada a partir de la secuenciación de ADNcf extraído de un conjunto de 48 muestras de sangre de ensayo obtenidas de sujetos humanos, cada uno embarazado de un feto masculino o femenino. Las dosis de cromosoma X para muestras de sujetos masculinos (46,XY; (O)), femeninos (46,XX; ( $\Delta$ )); monosomía X (45,X; (+)) y cariotipos complejos (Cplx (X)) se muestran para cromosomas 1-12 y X (**A**) y para cromosomas 1-22 y X (**B**).

La figura 13 muestra la distribución de la dosis de cromosoma para el cromosoma Y determinada a partir de la secuenciación de ADNcf extraído de un conjunto de 48 muestras de sangre de ensayo obtenidas de sujetos humanos, cada uno embarazado de un feto masculino o femenino. Las dosis de cromosoma Y para muestras de sujetos masculinos (46,XY; ( $\Delta$ )), femeninos (46,XX; (O)); monosomía X (45,X; (+)) y cariotipos complejos (Cplx (X)) se muestran para cromosomas 1-12 (**A**) y para cromosomas 1-22 (**B**).

La figura 14 muestra el coeficiente de variación (CV) para 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 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 las dosis de secuencia (eje Y) para un segmento del cromosoma 11 (81000082-103000103 pb) determinadas a partir de la secuenciación de ADNcf extraído de un conjunto de 7 muestras calificadas (O) obtenidas y 1 muestra de ensayo ( $\blacklozenge$ ) de sujetos humanos embarazados. Se identificó una muestra de un sujeto que portaba un feto con una aneuploidía parcial del cromosoma 11 ( $\blacklozenge$ ).

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

La figura 18 muestra un diagrama de barras que representa la identificación de secuencias polimórficas (SNP) fetales y maternas utilizadas para determinar la fracción fetal en una muestra de ensayo. Se muestra el número total de lecturas de secuencias (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 un caso de uso de la fracción fetal para determinar los umbrales de corte para la detección de aneuploidías.

La figura 20 ilustra la distribución de 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) con respecto a la desviación estándar de la media (eje Y) para la dosis de cromosoma correspondiente en muestras no afectadas.

#### 4. Descripción detallada de la invención

La invención proporciona un método para la secuenciación de ácidos nucleicos que comprende: (a) proporcionar una muestra de ensayo que comprende moléculas de ácido nucleico, en el que dichas moléculas de ácido nucleico son moléculas de ADN genómico humano; (b) realizar la reparación de extremos de las moléculas de ácido nucleico para generar ácidos nucleicos de extremos romos; (c) realizar la adición de colas de dA a los ácidos nucleicos de extremos romos para generar ácidos nucleicos con cola de dA; (d) ligar adaptadores a los ácidos nucleicos con cola de dA para generar una biblioteca de polinucleótidos ligados a adaptadores; (e) opcionalmente amplificar la biblioteca usando cebadores de amplificación, comprendiendo dichos cebadores de amplificación una porción específica de adaptador; y (f) someter la biblioteca a una secuenciación masivamente paralela; en el que las etapas (b), (c) y (d) son etapas consecutivas.

El método de la invención es aplicable a métodos para determinar la presencia o la ausencia de una aneuploidía, por ejemplo, una aneuploidía cromosómica o parcial, y/o la fracción fetal en muestras maternas que comprenden ácidos nucleicos fetales y maternos mediante secuenciación masivamente paralela. El método de la invención comprende un protocolo novedoso para preparar bibliotecas de secuenciación que mejora inesperadamente la calidad del ADN de la biblioteca a la vez que agiliza el proceso de análisis de muestras para diagnósticos prenatales. El método de la invención es aplicable a métodos que permiten determinar variaciones en el número de copias (CNV) de cualquier secuencia de interés en una muestra de ensayo que comprende una mezcla de ácidos nucleicos que se sabe o se sospecha que difieren en la cantidad de una o más secuencias de interés, y/o determinar la fracción de una de al menos dos poblaciones de ácidos nucleicos aportados a la muestra por diferentes genomas. Las secuencias de interés incluyen secuencias genómicas que varían desde cientos de bases a decenas de megabases a cromosomas completos que se sabe o se sospecha que están asociados con una condición genética o patológica. Los 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 tales como el cáncer, por ejemplo, trisomía parcial 8 en leucemia mieloide aguda. Los métodos para determinar CNV divulgados en el presente documento pueden comprender un enfoque estadístico que tenga en cuenta la variabilidad acumulada derivada de la variabilidad relacionada con el proceso, intercromosómica e intersecuenciación. El método de la invención es aplicable a métodos para determinar CNV de cualquier aneuploidía fetal y CNV que se sabe o se sospecha que están asociadas con una diversidad de condiciones médicas.

A menos que se indique lo contrario, la puesta en práctica de la presente invención implica técnicas convencionales utilizadas comúnmente en biología molecular, microbiología, purificación de proteínas, ingeniería de proteínas,

secuenciación de proteínas y ADN y campos de ADN recombinante, que se encuentran dentro de la experiencia en la técnica. Dichas técnicas son conocidas por los expertos en la técnica y se describen en numerosos textos estándar y trabajos de referencia

5 Los intervalos numéricos incluyen los números que definen el intervalo. Se pretende que cada limitación numérica máxima dada a lo largo de la presente memoria descriptiva incluya cada limitación numérica inferior, como si dichas limitaciones numéricas inferiores estuvieran expresamente escritas en el presente documento. Cada limitación numérica mínima dada a lo largo de la presente memoria descriptiva incluirá cada limitación numérica superior, como si dichas limitaciones numéricas superiores estuvieran expresamente escritas en el presente documento. Cada intervalo numérico dado a lo largo de la presente memoria descriptiva incluirá cada intervalo numérico más estrecho que se encuentre dentro de dicho intervalo numérico más amplio, como si dichos intervalos numéricos más estrechos estuvieran todos expresamente escritos en el presente documento

### 5.1 Definiciones

15 Tal como se utilizan en el presente documento, los términos en singular "un", "una", "el" y "la" 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 orientación 5' a 3' y las secuencias de aminoácidos se escriben de izquierda a derecha en orientación amino a carboxi, respectivamente.

20 El término "evaluar" en el presente documento se refiere a caracterizar el estado de una aneuploidía cromosómica mediante uno de los tres tipos de designaciones: "normal", "afectada" y "sin designación". Por ejemplo, en presencia de trisomía la designación "normal" se determina mediante el valor de un parámetro, por ejemplo, una dosis de cromosoma de ensayo que se encuentra por debajo de un umbral de fiabilidad definido por el usuario, la designación "afectada" se determina por un parámetro, por ejemplo, una dosis de cromosoma de ensayo, que se encuentra por encima de un umbral de fiabilidad definido por el usuario, y el resultado de "sin designación" se determina por un parámetro, por ejemplo, una dosis de cromosoma de ensayo, que se encuentra entre los umbrales de fiabilidad definidos por el usuario para realizar una designación "normal" o "afectada".

30 El término "variación del número de copias" en el presente documento se refiere a la variación en el número de copias de una secuencia de ácido nucleico de 1 kb o mayor presente en una muestra de ensayo en comparación con el número de copias de la secuencia de ácido nucleico presente en una muestra calificada. Una "variante del número de copias" se refiere a la secuencia de ácido nucleico de 1 kb o mayor en la que se encuentran diferencias en el número de copias mediante comparación de una secuencia de interés en la muestra de ensayo con la presente en una muestra calificada. Las variantes/variaciones del número de copias incluyen deleciones, incluidas microdeleciones, inserciones, incluidas microinserciones, duplicaciones, multiplicaciones, inversiones, translocaciones y variantes complejas de sitios múltiples. La CNV abarca las aneuploidías cromosómicas y las aneuploidías parciales.

40 El término "aneuploidía" se refiere en el presente documento a un desequilibrio del material genético provocado por una pérdida o una ganancia de un cromosoma completo, o parte de un cromosoma.

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

El término "aneuploidía parcial" se refiere en el presente documento a un desequilibrio del material genético provocado por una pérdida o una ganancia de parte de un cromosoma, por ejemplo, monosomía parcial y trisomía parcial, y abarca los desequilibrios resultantes de translocaciones, deleciones e inserciones.

50 El término "pluralidad" se utiliza en el presente documento con referencia a un número de moléculas de ácido nucleico o etiquetas de secuencia que es suficiente para identificar diferencias significativas en variaciones del número de copias (por ejemplo, dosis de cromosomas) en muestras de ensayo y muestras calificadas utilizando los métodos del presente documento. En algunas formas de realización se obtienen para cada muestra de ensayo al menos aproximadamente  $3 \times 10^6$  etiquetas de secuencia, al menos aproximadamente  $5 \times 10^6$  etiquetas de secuencia, al menos aproximadamente  $8 \times 10^6$  etiquetas de secuencia, al menos aproximadamente  $10 \times 10^6$  etiquetas de secuencia, al menos aproximadamente  $15 \times 10^6$  etiquetas de secuencia, al menos aproximadamente  $20 \times 10^6$  etiquetas de secuencia, al menos aproximadamente  $30 \times 10^6$  etiquetas de secuencia, al menos aproximadamente  $40 \times 10^6$  etiquetas de secuencia, o al menos aproximadamente  $50 \times 10^6$  etiquetas de secuencia que comprenden lecturas de entre 20 y 40 pb.

60 Los términos "polinucleótido", "ácido nucleico" y "moléculas de ácido nucleico" se utilizan indistintamente y se refieren a una secuencia de nucleótidos unida covalentemente (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, e incluyen secuencias de cualquier forma de ácido nucleico, incluidas, pero sin limitación, moléculas de ARN, ADN y ADNcf. El término "polinucleótido" incluye, sin limitación, polinucleótido monocatenario y bicatenario.

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

5 El término "muestra de ensayo" se refiere en el presente documento a una muestra que comprende una mezcla de ácidos nucleicos que comprende al menos una secuencia de ácido nucleico cuyo número de copias se sospecha que ha sufrido una variación. Los ácidos nucleicos presentes en una muestra de ensayo se denominan "ácidos nucleicos de ensayo".

10 El término "muestra calificada" se refiere en el presente documento a una muestra que comprende una mezcla de ácidos nucleicos que están presentes en un número de copias conocido con los que se comparan los ácidos nucleicos de 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 de normalización para el cromosoma 21 es una muestra que no es una muestra de trisomía 21.

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

20 El término "secuencia de interés" se refiere en el presente documento a una secuencia de ácido nucleico que está asociada con una diferencia en la representación de la secuencia en individuos sanos frente a enfermos. Una secuencia de interés puede ser una secuencia en un cromosoma que está tergiversada, es decir, sobrerrepresentada o subrepresentada, 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 subrepresentado en un cáncer. Las secuencias de interés incluyen secuencias que están sobrerrepresentadas o subrepresentadas en la población total, o una subpoblación de células de un sujeto. Una "secuencia calificada de interés" es una secuencia de interés en una muestra calificada. Una "secuencia de ensayo de interés" es una secuencia de interés en una muestra de ensayo.

25 El término "secuencia de normalización" se refiere en el presente documento a una secuencia que muestra una variabilidad en el número de etiquetas de secuencia que se mapean a la misma entre muestras y ejecuciones de secuenciación que mejor se aproxima a la de la secuencia de interés para la que se utiliza como parámetro de normalización, y que mejor puede diferenciar una muestra afectada de una o más muestras no afectadas. Un "cromosoma de normalización" es un ejemplo, de una "secuencia de normalización".

30 El término "diferenciabilidad" se refiere en el presente documento 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.

35 El término "dosis de secuencia" se refiere en el presente documento a un parámetro que relaciona la densidad de etiquetas de secuencia de una secuencia de interés con la densidad de etiquetas de una secuencia de normalización. Una "dosis de secuencia de ensayo" es un parámetro que relaciona la densidad de etiquetas de secuencia de una secuencia de interés, por ejemplo, el cromosoma 21, con la de una secuencia de normalización, por ejemplo, el cromosoma 9, determinada en una muestra de ensayo. De forma similar, una "dosis de secuencia calificada" es un parámetro que relaciona la densidad de etiquetas de secuencia de una secuencia de interés con la de una secuencia de normalización determinada en una muestra calificada.

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

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

50

Los términos "valor umbral" y "valor umbral calificado" se refieren en el presente documento a cualquier número que se calcula utilizando un conjunto de datos calificados y que sirve como límite de diagnóstico de una variación del número de copias, por ejemplo, una aneuploidía, en un organismo. Si los resultados obtenidos en la puesta en práctica de los métodos divulgados en el presente documento superan un umbral, se puede diagnosticar a un sujeto una variación del número de copias, por ejemplo, trisomía 21.

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

El término "etiqueta de secuencia" se utiliza en el presente documento de forma intercambiable con el término "etiqueta de secuencia mapeada" para referirse a una lectura de secuencia que se ha asignado específicamente, es decir, mapeado, a una secuencia más grande, por ejemplo, un genoma de referencia, mediante alineamiento. Las etiquetas de secuencia mapeadas se mapean de forma única a un genoma de referencia, es decir, se asignan a una única ubicación en el genoma de referencia. Las etiquetas que se pueden mapear a más de una ubicación en un genoma de referencia, es decir, las etiquetas que no se mapean de forma única, no se incluyen en el análisis.

Tal como se utilizan en el presente documento, los términos "alineada", "alineamiento" o "alinear" se refieren a una o más secuencias que se identifican como una coincidencia en términos del orden de sus moléculas de ácido nucleico con una secuencia conocida de un genoma de referencia. Dicho alineamiento se puede realizar manualmente o mediante un algoritmo informático, incluyéndose entre los ejemplos el programa informático Efficient Local Alignment of Nucleotide Data (ELAND) distribuido como parte de la línea Illumina Genomics Analysis. La coincidencia de una lectura de secuencia en el alineamiento puede ser una coincidencia de secuencia del 100% o inferior al 100% (coincidencia no perfecta).

Tal como se utiliza en el presente documento, el término "genoma de referencia" se refiere a cualquier secuencia genómica conocida particular, ya sea parcial o completa, de cualquier organismo o virus que pueda utilizarse para referenciar secuencias identificadas de un sujeto. Por ejemplo, en el National Center for Biotechnology Information, en [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov), se encuentra un genoma de referencia utilizado para sujetos humanos, así como para muchos otros organismos. Un "genoma" se refiere a la información genética completa de un organismo o un virus, expresada en secuencias de ácido nucleico.

Los términos "genoma de secuencias diana artificiales" y "genoma de referencia artificial" se refieren en el presente documento 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 abarca alelos de SNP conocidos.

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

El término "derivado" cuando se utiliza en el contexto de un ácido nucleico o una mezcla de ácidos nucleicos, se refiere en el presente documentos a los medios mediante los que se obtienen los ácidos nucleicos de la fuente de la que se originan. Por ejemplo, en una forma de realización, una mezcla de ácidos nucleicos que se deriva de dos genomas diferentes significa que los ácidos nucleicos, por ejemplo, ADNcf, se liberaron de forma natural por las células a través de procesos naturales tales como la necrosis o la apoptosis. En otra forma de realización, una mezcla de ácidos nucleicos que se deriva de dos genomas diferentes significa que los ácidos nucleicos se extrajeron de dos tipos diferentes de células de un sujeto.

El término "muestra mixta" se refiere en el presente documento a una muestra que contiene una mezcla de ácidos nucleicos que se derivan de diferentes genomas.

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

El término "muestra materna original" se refiere en el presente documento a una muestra biológica obtenida de un sujeto embarazado, por ejemplo, una mujer embarazada, que sirve como fuente de la que 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 purificado extraída de una muestra de plasma materno.

El término "fluido biológico" se refiere en el presente documento a un líquido tomado de una fuente biológica e incluye, por ejemplo, sangre, suero, plasma, esputo, líquido de lavado, líquido cefalorraquídeo, orina, semen, sudor, lágrimas, saliva y similares. Tal como se utilizan en el presente documento, los términos "sangre", "plasma" y "suero" abarcan expresamente fracciones o porciones procesadas de los mismos. De forma similar, cuando se tome una muestra de

una biopsia, hisopo, frotis, etc., la "muestra" englobará expresamente una fracción o porción procesada derivada de la biopsia, hisopo, frotis, etc.

5 Los términos "ácidos nucleicos maternos" y "ácidos nucleicos fetales" se refieren en el presente documento a los ácidos nucleicos de un sujeto femenino embarazado y los ácidos nucleicos del feto que porta el sujeto femenino embarazado, respectivamente.

10 Tal como se utiliza en el presente documento, 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 sujetos, y que no necesariamente tiene la misma secuencia en todos los genomas, pero que sirve para proporcionar la identidad en lugar de la información genética de una secuencia de interés, por ejemplo, un gen o un cromosoma.

15 Tal como se utiliza en el presente documento, el término "sustancialmente libre de células" abarca preparaciones de la muestra deseada de la que se eliminan los componentes que normalmente están asociados con la misma. Por ejemplo, una muestra de plasma se vuelve esencialmente libre de células al eliminar las células sanguíneas, por ejemplo, glóbulos rojos, que normalmente se asocian con la misma. En algunas formas de realización, las muestras sustancialmente libres de células se procesan para eliminar células que de otro modo contribuirían al material genético deseado que se va a analizar para evaluar una CNV.

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

25 Tal como se utiliza en el presente documento, el término "cromosoma" se refiere al portador del gen portador de la herencia de una célula viva que se deriva de cromatina y que comprende ADN y componentes proteicos (especialmente histonas). En el presente documento se emplea el sistema de numeración de cromosomas del genoma humano individual internacionalmente reconocido convencional.

30 Tal como se utiliza en el presente documento, el término "longitud de polinucleótido" se refiere al 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 dada en pares de bases, por ejemplo, proporcionada en el ensamblaje NCBI36/hg18 del cromosoma humano que se encuentra en Internet en [genome.ucsc.edu/cgi-bin/hgTracks?hgid=167155613&chromInfoPage=](http://genome.ucsc.edu/cgi-bin/hgTracks?hgid=167155613&chromInfoPage=)

35 El término "sujeto" se refiere en el presente documento 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 del presente documento se refieren a seres humanos y el lenguaje se refiere principalmente a cuestiones humanas, el concepto de la presente invención puede aplicarse a genomas de cualquier planta o animal, y es útil en los campos de la medicina veterinaria, ciencias animal, laboratorios de investigación y similares.

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

45 El término "cromosoma aneuploide" se refiere en el presente documento a un cromosoma que está implicado en una aneuploidía.

50 El término "aneuploidía" se refiere en el presente documento a un desequilibrio del material genético provocado por una pérdida o una ganancia de un cromosoma completo, o parte de un cromosoma.

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

55 Los términos "extremo romo" y "reparación de extremos" se utilizan en el presente documento indistintamente para referirse a un proceso enzimático que da como resultado que ambas cadenas de una molécula de ADN bicatenario terminen en un par de bases, y no incluye la purificación de los productos de extremos romos de la enzima de extremos romos.

60 El término "adición de colas de d-A" en el presente documento 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 con cola de d-A de la enzima de adición de colas de d-A.

65 El término "ligación de adaptador" se refiere en el presente documento a un proceso enzimático que liga una secuencia de adaptador de ADN a fragmentos de ADN, y no incluye la purificación del producto ligado al adaptador de la enzima de ligación.

El término "recipiente de reacción" se refiere en el presente documento a un recipiente de cualquier forma, tamaño, capacidad o material que se puede utilizar para procesar una muestra durante un procedimiento de laboratorio, por ejemplo, de investigación o clínico.

5 El término "etapas consecutivas" se utiliza en el presente documento con referencia a las etapas enzimáticas sucesivas de formación de extremos romos, adición de colas de dA y ligación de adaptadores a ADN que no están interpuestos por etapas de purificación.

10 Tal como se utiliza en el presente documento, el término "purificado" se refiere a material (por ejemplo, un polinucleótido aislado) que se encuentra en un estado relativamente puro, por ejemplo, que es al menos aproximadamente el 80% puro, al menos aproximadamente el 85% puro, al menos aproximadamente el 90% puro, al menos aproximadamente el 95% puro, al menos aproximadamente el 98% puro, o incluso al menos aproximadamente el 99% puro.

15 Los términos "extraído", "recuperado", "aislado" y "separado" se refieren a un compuesto, proteína, célula, ácido nucleico o aminoácido que se retira de al menos un componente con el que está asociado naturalmente y se encuentra en la naturaleza.

20 El término "SNP en tándem" en el presente documento se refiere a dos o más SNP que están presentes dentro de una secuencia de ácido nucleico diana polimórfico.

25 Los términos "ácido nucleico diana polimórfico", "secuencia polimórfica", "secuencia de ácido nucleico diana polimórfico" y "ácido nucleico polimórfico" se utilizan indistintamente en el presente documento para referirse a una secuencia de ácido nucleico, por ejemplo, una secuencia de ADN que comprende uno o más sitios polimórficos.

30 El término "sitio polimórfico" se refiere en el presente documento a un polimorfismo de un solo nucleótido (SNP), una delección o inserción de múltiples bases a pequeña escala, un polimorfismo de múltiples nucleótidos (MNP) o una repetición en tándem corta (STR).

35 El término "pluralidad de ácidos nucleicos diana polimórficos" se refiere en el presente documento a una serie de secuencias de ácidos nucleicos, cada una de las cuales comprende 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.

40 El término "enriquecer" se refiere en el presente documento al proceso de amplificación de ácidos nucleicos diana polimórficos contenidos en una porción de una muestra materna y la combinación del producto amplificado con el resto de la muestra materna de la que se extrajo la porción.

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

50 Tal como se utiliza en el presente documento, el término "amplificación en fase sólida" se refiere, tal como se utiliza en el presente documento, a cualquier reacción de amplificación de ácido nucleico llevada a cabo en un soporte sólido o en asociación con el mismo, de modo que la totalidad o una parte de los productos amplificados se inmovilicen en el soporte sólido a medida que se vayan formando. 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 cebadores de amplificación directo e inverso están inmovilizados en el soporte sólido. La PCR en fase sólida abarca sistemas tales como emulsiones, en los que un cebador está anclado a una perla y el otro está en solución libre, y la formación de colonias en matrices de gel en fase sólida en las que un cebador está anclado a la superficie y el otro está en solución libre. El término fase sólida, o superficie, se utiliza para referirse a 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 celda de flujo similares; perlas, en las que uno o dos cebadores se unen a las perlas y las perlas se amplifican; o una matriz de perlas sobre una superficie después de que las perlas se hayan amplificado.

60 Tal como se utiliza en el presente documento, el término "grupo de cromosomas" se refiere en el presente documento a un grupo de dos o más cromosomas.

65

Un "polimorfismo de un solo nucleótido" (SNP) se produce en un sitio polimórfico ocupado por un solo nucleótido, que es el sitio de variación entre secuencias alélicas. El sitio suele estar 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 pirimidina por otra pirimidina. Una transversión es la sustitución de una purina por una pirimidina o viceversa. Los SNP también pueden surgir de una delección de un nucleótido o una inserción de un nucleótido con respecto a un alelo de referencia. Los polimorfismos de un solo nucleótido (SNP) son posiciones en las que tienen lugar dos bases alternativas con una frecuencia apreciable (> 1%) en la población humana y son el tipo más común de variación genética humana.

Tal como se utiliza en el presente documento, el término "repetición en tándem corta" o "STR", tal como se utiliza en el presente documento, se refiere a una clase de polimorfismos que se producen cuando se repite un patrón de dos o más nucleótidos y las secuencias repetidas son directamente adyacentes entre sí. El patrón puede tener una longitud de 2 a 10 pares de bases (pb) (por ejemplo, (CATG)<sub>n</sub> en una región genómica) y normalmente se encuentra en la región del intrón no codificante. Examinando varios loci de STR y contando cuántas repeticiones de una secuencia de STR específica hay en un locus dado, es posible crear un perfil genético único de un individuo.

Tal como se utiliza en el presente documento, el término "miniSTR" se refiere en el presente documento a la repetición en tándem de cuatro o más pares de bases que abarca 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 aproximadamente 100 pares de bases, menos de aproximadamente 50 pares de bases o menos de aproximadamente 25 pares de bases. Los "miniSTR" son STR que se pueden amplificar a partir de plantillas de ADNcf.

El término "SNP en tándem" se refiere en el presente documento a dos o más SNP que están presentes dentro de una secuencia de ácido nucleico diana polimórfico.

Tal como se utiliza en el presente documento, el término "biblioteca enriquecida" se refiere en el presente documento a una biblioteca de secuenciación que comprende secuencias de ácidos nucleicos diana polimórficos amplificadas. 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" se refiere en el presente documento a una biblioteca de secuenciación que no comprende, es decir, una biblioteca generada a partir de secuencias de ADNcf de origen natural. Una "biblioteca de ácidos nucleicos polimórficos diana" es una biblioteca generada a partir de ácidos nucleicos diana amplificados.

Tal como se utiliza en el presente documento, el término "secuencias de ADNcf de origen natural" se refiere en el presente documento a fragmentos de ADNcf tal como están presentes en una muestra, y en contraste con fragmentos de ADN genómico que se obtienen mediante métodos de fragmentación descritos en el presente documento.

## 5.2 Descripción

La invención proporciona un método para la secuenciación de ácidos nucleicos que comprende: (a) proporcionar una muestra de ensayo que comprende moléculas de ácido nucleico, en el que dichas moléculas de ácido nucleico son moléculas de ADN genómico humano; (b) realizar la reparación de extremos de las moléculas de ácido nucleico para generar ácidos nucleicos de extremos romos; (c) realizar la adición de colas de dA a los ácidos nucleicos de extremos romos para generar ácidos nucleicos con cola de dA; (d) ligar adaptadores a los ácidos nucleicos con cola de dA para generar una biblioteca de polinucleótidos ligados a adaptadores; (e) opcionalmente amplificar la biblioteca usando cebadores de amplificación, comprendiendo dichos cebadores de amplificación una porción específica de adaptador; y (f) someter la biblioteca a una secuenciación masivamente paralela; en el que las etapas (b), (c) y (d) son etapas consecutivas.

El método de la invención es aplicable a métodos para determinar la presencia o la ausencia de una aneuploidía, por ejemplo, una aneuploidía cromosómica o parcial, y/o la fracción fetal en muestras maternas que comprenden ácidos nucleicos fetales y maternos mediante secuenciación masivamente paralela. El método de la invención comprende un protocolo novedoso para preparar bibliotecas de secuenciación que mejora inesperadamente la calidad del ADN de la biblioteca a la vez que agiliza el proceso de análisis de muestras para diagnósticos prenatales. El método de la invención es aplicable a métodos que permiten determinar variaciones en el número de copias (CNV) de cualquier secuencia de interés en una muestra de ensayo que comprende una mezcla de ácidos nucleicos que se sabe o se sospecha que difieren en la cantidad de una o más secuencias de interés, y/o determinar la fracción de una de al menos dos poblaciones de ácidos nucleicos aportados a la muestra por diferentes genomas.

### Métodos de secuenciación

En una forma de realización, el método descrito en el presente documento emplea tecnología de secuenciación de próxima generación (NGS) en la que se secuencian plantillas de ADN amplificadas clonalmente o moléculas de ADN individuales de forma masivamente paralela dentro de una celda de flujo (por ejemplo, tal como se describe por

Volkerding et al. Clin Chem 55:641-658 [2009]; Metzker M Nature Rev 11:31-46 [2010]). Además de la información de secuencias de alto rendimiento, la NGS proporciona información cuantitativa digital, ya que cada lectura de secuencia es una "etiqueta de secuencia" contable que representa una plantilla de ADN clonal individual o una sola molécula de ADN. Esta cuantificación permite a la NGS expandir el concepto de PCR digital de contar moléculas de ADN libre 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 colorantes reversibles, secuenciación por ligación de sonda de oligonucleótidos y secuenciación en tiempo real.

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 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), tal como se describe a continuación. Además de la secuenciación de una sola molécula realizada utilizando la secuenciación por síntesis de Helicos Biosciences, el método de la invención abarca otras tecnologías de secuenciación de una sola molécula e incluyen la tecnología SMRT™ de Pacific Biosciences, la tecnología Ion Torrent™ y la secuenciación por nanoporos que está desarrollando, por ejemplo, Oxford Nanopore Technologies.

Aunque el método de Sanger automatizado se considera una tecnología de "primera generación", la secuenciación de Sanger, incluida la secuenciación de Sanger automatizada, también puede emplearse por el método de la invención. También están abarcados por el método de la invención métodos de secuenciación adicionales que comprenden el uso de tecnologías de formación de imágenes de ácidos nucleicos en desarrollo, por ejemplo, la microscopía de fuerza atómica (AFM) o la microscopía electrónica de transmisión (TEM). A continuación, se describen tecnologías de secuenciación ilustrativas.

En una forma de realización, la tecnología de secuenciación de ADN que se utiliza en el método de la invención es la secuenciación de una sola molécula verdadera de Helicos (tSMS) (por ejemplo, tal como se describe por Harris T.D. et al., Science 320:106-109 [2008]). En la técnica tSMS, una muestra de ADN se escinde en cadenas de aproximadamente 100 a 200 nucleótidos y se añade una secuencia de poliA al extremo 3' de cada cadena de ADN. Cada cadena se marca mediante la adición de un nucleótido de adenosina marcado con fluorescencia. Después, las cadenas de ADN se hibridan a una celda de flujo, que contiene millones de sitios de captura de oligo-T que están inmovilizados en la superficie de la celda de flujo. Las plantillas pueden tener una densidad de aproximadamente 100 millones de plantillas/cm<sup>2</sup>. Después, la celda de flujo se carga en un instrumento, por ejemplo, un secuenciador HeliScope™, y un láser ilumina la superficie de la celda de flujo, revelando la posición de cada plantilla. Una cámara CCD puede mapear la posición de las plantillas en la superficie de la celda de flujo. Después, la etiqueta fluorescente de plantilla se corta y se elimina por lavado. 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 de una forma dirigida por plantilla. La polimerasa y los nucleótidos no incorporados se eliminan. Las plantillas que han dirigido la incorporación del nucleótido marcado con fluorescencia se distinguen mediante la obtención de imágenes de la superficie de la celda 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 la secuencia se recopila con cada etapa de adición de nucleótidos.

En una forma de realización, la tecnología de secuenciación de ADN que se utiliza en el método de la invención es la secuenciación 454 (Roche) (por ejemplo, tal como se describe por Margulies, M. et al. Nature 437:376-380 [2005]). La secuenciación 454 implica dos etapas. En la primera etapa, el ADN se corta en fragmentos de aproximadamente 300 a 800 pares de bases, y los extremos de los fragmentos se vuelven romos. Después, los adaptadores de oligonucleótidos se ligan a los extremos de los fragmentos. Los adaptadores sirven como cebadores para la amplificación y la secuenciación de los fragmentos. Los fragmentos pueden unirse a perlas de captura de ADN, por ejemplo, perlas recubiertas de estreptavidina, utilizando, por ejemplo, el adaptador B, que contiene la etiqueta 5'-biotina. Los fragmentos unidos a las perlas se amplifican por PCR dentro de las gotas de una emulsión de aceite y agua. El resultado son múltiples copias de fragmentos de ADN amplificados clonalmente en cada perla. En la segunda etapa, las perlas se capturan en pocillos (de un tamaño del orden de picolitros). Se realiza una pirosecuenciación en cada fragmento de ADN en paralelo. La adición de uno o más nucleótidos genera una señal de luz que es registrada por una cámara CCD en un instrumento de secuenciación. La intensidad de la señal es proporcional al número de nucleótidos incorporados. La pirosecuenciación utiliza pirofosfato (PPI) que se libera tras la adición de nucleótidos. El PPI se convierte en ATP mediante ATP sulfúrilasa en presencia de adenosina 5' fosfosulfato. La luciferasa utiliza ATP para convertir la luciferina en oxiluciferina, y esta reacción genera luz que se discierne y se analiza.

En una forma de realización, la tecnología de secuenciación de ADN que se utiliza en el método de la invención es tecnología SOLiD™ (Applied Biosystems). En la secuenciación por ligación SOLiD™, el ADN genómico se corta en fragmentos y se unen adaptadores a los extremos 5' y 3' de los fragmentos para generar una biblioteca de fragmentos. Alternativamente, se pueden introducir adaptadores internos ligando adaptadores a los extremos 5' y 3' de los fragmentos, circularizando los fragmentos, digiriendo el fragmento circularizado para generar un adaptador interno y uniendo adaptadores a los extremos 5' y 3' de los fragmentos resultantes para generar una biblioteca de elementos

emparejados. 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 de las perlas seleccionadas se someten a una modificación 3' que permite la unión a un portaobjetos de vidrio. La secuencia se puede determinar mediante hibridación secuencial y ligación de oligonucleótidos parcialmente aleatorios con una base determinada central (o un par de bases) que se identifica mediante un fluoróforo específico. Después de registrar un color, el oligonucleótido ligado se escinde y se elimina y después se repite el proceso.

En una forma de 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 individuales de polimerasa de ADN se unen a la superficie inferior de los identificadores de longitud de onda de modo cero individuales (identificadores ZMW) que obtienen información de secuencia mientras los nucleótidos fosfoenlazados se incorporan a la cadena de cebador en crecimiento. Un ZMW es una estructura de confinamiento que permite la observación de la incorporación de un solo nucleótido por la ADN polimerasa contra el fondo de nucleótidos fluorescentes que se difunden rápidamente dentro y fuera del ZMW (en microsegundos). Se necesitan varios milisegundos para incorporar un nucleótido en una cadena en crecimiento. Durante este periodo de tiempo, la etiqueta fluorescente se excita y produce una señal fluorescente y la etiqueta fluorescente se escinde. La identificación de la fluorescencia correspondiente del colorante indica qué base se ha incorporado. El proceso se repite.

En una forma de realización, la tecnología de secuenciación de ADN que se utiliza en el método de la invención es la secuenciación por nanoporos (por ejemplo, tal como se describe por Soni GV y Meller A. Clin Chem 53: 1996-2001 [2007]). Varias empresas, incluida Oxford Nanopore Technologies (Oxford, Reino Unido), están desarrollando industrialmente técnicas de análisis de ADN de secuenciación por nanoporos. La secuenciación por nanoporos es una tecnología de secuenciación de una sola molécula mediante la cual una sola molécula de ADN se secuencia directamente a medida que pasa a través de un nanoporo. Un nanoporo es un agujero pequeño, del orden de 1 nanómetro de diámetro. La inmersión de un nanoporo en un fluido conductor y la aplicación de un potencial (voltaje) a través del mismo da como resultado 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 la forma del nanoporo. Cuando una molécula de ADN pasa a través de un nanoporo, cada nucleótido de la molécula de ADN obstruye el nanoporo en un grado diferente, cambiando la magnitud de la corriente a través del nanoporo en diferentes grados. Así, este cambio en la corriente a medida que la molécula de ADN pasa a través del nanoporo representa una lectura de la secuencia de ADN.

En una forma de realización, la tecnología de secuenciación de ADN que se utiliza en el método de la invención es la matriz de transistores de efecto de campo sensible a productos químicos (chemFET) (por ejemplo, tal como se describe en la publicación de solicitud de patente de Estados Unidos N° 20090026082). En un ejemplo, de la técnica, las moléculas de ADN pueden disponerse en cámaras de reacción y las moléculas de plantilla pueden hibridarse 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 mediante un chemFET. Una matriz puede tener múltiples sensores chemFET. En otro ejemplo, los ácidos nucleicos individuales pueden unirse a perlas, y los ácidos nucleicos se pueden amplificar en la perla, y las perlas individuales se pueden transferir a cámaras de reacción individuales en una matriz chemFET, en la que cada cámara tiene un sensor chemFET, y los ácidos nucleicos se pueden secuenciar.

En una forma de realización, la tecnología de secuenciación de ADN que se utiliza en el método de la invención es el método de Halcyon Molecular que usa microscopía electrónica de transmisión (TEM). El método, denominado Individual Molecule Placement Rapid Nano Transfer (IMPRNT), comprende la utilización de imágenes de microscopio electrónico de transmisión de resolución de un solo átomo de ADN de alto peso molecular (150 kb o más grande) marcado selectivamente con marcadores de átomos pesados y la disposición de estas moléculas en películas ultrafinas en matrices paralelas ultradensas (3 nm de cadena a cadena) 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 átomos pesados y para extraer información de la secuencia de bases del ADN. El método se describe más detalladamente en la publicación de patente PCT WO 2009/046445. El método permite secuenciar genomas humanos completos en menos de diez minutos.

En una forma de realización, la tecnología de secuenciación de ADN es la secuenciación de una sola molécula por 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 una polimerasa incorpora un nucleótido a una cadena de ADN, se libera un ion de hidrógeno como subproducto. Ion Torrent utiliza una matriz de alta densidad de pocillos micromecanizados para realizar este proceso bioquímico de forma masivamente paralela. Cada pocillo contiene una molécula de ADN diferente. Debajo de los pozos hay una capa sensible a los iones y debajo de la misma un sensor de iones. Cuando se añade un nucleótido, por ejemplo, una C, a una plantilla de ADN y después se incorpora a una cadena 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 del Ion Torrent. El secuenciador, esencialmente el medidor de pH en estado sólido más pequeño del mundo, designa la base, yendo directamente de la información química a la información digital. El secuenciador Ion personal Genome Machine

(PGM™) inunda después secuencialmente el chip con un nucleótido tras otro. Si el siguiente nucleótido que inunda el chip no es una coincidencia, no se registrará ningún cambio de voltaje y no se designará ninguna base. Si hay dos bases idénticas en la cadena de ADN, el voltaje será el doble y el chip registrará dos bases idénticas designadas. La identificación directa permite registrar la incorporación de nucleótidos en segundos.

Otros métodos de secuenciación incluyen PCR digital y secuenciación por hibridación. Puede utilizarse la reacción en cadena de la polimerasa digital (PCR digital o dPCR) 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 de microfluidos, y cada ácido nucleico se amplifica individualmente mediante PCR. Los ácidos nucleicos se pueden separar de forma que haya una media 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 de polinucleótidos con una pluralidad de sondas de polinucleótidos, pudiendo unirse cada una de la pluralidad de sondas de polinucleótidos opcionalmente a un sustrato. El sustrato podría ser una superficie plana que comprende una matriz de secuencias de nucleótidos conocidas. El patrón de hibridación con la matriz se puede utilizar para determinar las secuencias de polinucleótidos presentes en la muestra. En otras formas de realización, cada sonda está unida a una perla, por ejemplo, una perla magnética o similar. La hibridación con las perlas se puede identificar y utilizar para identificar la pluralidad de secuencias de polinucleótidos dentro de la muestra.

En una forma de realización, el método emplea la secuenciación masivamente paralela de millones de fragmentos de ADN utilizando la secuenciación por síntesis de Illumina y la química de secuenciación basada en terminadores reversibles (por ejemplo, tal como se describe por Bentley et al., *Nature* 6:53-59 [2009]). El ADN de plantilla puede ser ADN genómico por ejemplo, ADNcf. En algunas formas de realización, se utiliza ADN genómico de células aisladas como plantilla y se fragmenta en longitudes de varios cientos de pares de bases. En otras formas de realización, se utiliza ADNcf como la plantilla y no se requiere fragmentación ya que el ADNcf existe como fragmentos cortos. Por ejemplo, el ADNcf fetal circula en el torrente sanguíneo como fragmentos de < 300 pb, y se ha estimado 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 en la que se unen los anclajes de oligonucleótidos. El ADN de plantilla se prepara en los extremos para generar extremos romos fosforilados en 5', y la actividad de polimerasa del fragmento Klenow se utiliza para añadir una sola base A al extremo 3' de los fragmentos de ADN fosforilados romos. Esta adición prepara los fragmentos de ADN para la ligación a los adaptadores de oligonucleótidos, que tienen un saliente de una sola base T en su extremo 3' para aumentar la eficacia de la ligación. Los oligonucleótidos adaptadores son complementarios a los anclajes de la celda de flujo. En condiciones de dilución limitantes, se añade ADN de plantilla monocatenario modificado con adaptador a la celda de flujo y se inmoviliza mediante hibridación con los anclajes. Los fragmentos de ADN unidos se extienden y se amplifican en puente para crear una celda de flujo de secuenciación de densidad ultraalta con cientos de millones de clústeres, cada una con ~1000 copias de la misma plantilla. En una forma de realización, el ADN genómico fragmentado al azar, por ejemplo, ADNcf, se amplifica mediante PCR antes de someterse a la amplificación de clústeres. Como alternativa, se utiliza una preparación de biblioteca genómica sin amplificación y el ADN genómico fragmentado aleatoriamente, por ejemplo, ADNcf, se enriquece utilizando solo la amplificación de clústeres (Kozarewa et al., *Nature Methods* 6:291-295 [2009]). Las plantillas se secuencian utilizando una tecnología de secuenciación por síntesis de ADN de cuatro colores robusta que emplea terminadores reversibles con colorantes fluorescentes extraíbles. La identificación por fluorescencia de alta sensibilidad se realiza utilizando excitación láser y óptica de reflexión interna total. Las lecturas de secuencia corta de aproximadamente 20-40 pb, por ejemplo, de 36 pb, se alinean contra un genoma de referencia enmascarado de repetición y las diferencias genéticas se identifican utilizando un programa informático de canalización de análisis de datos especialmente desarrollado. Después de completar la primera lectura, las plantillas se pueden regenerar in situ para permitir una segunda lectura desde el extremo opuesto de los fragmentos. Así, según el método, se utiliza la secuenciación de un solo extremo o de extremos emparejados de los fragmentos de ADN. Se realiza la secuenciación parcial de los fragmentos de ADN presentes en la muestra y se cuentan etiquetas de secuencia que comprenden lecturas de longitud predeterminada, por ejemplo, 36 pb, que se mapean a un genoma de referencia conocido.

La longitud de la secuencia leída está asociada con la tecnología de secuenciación particular. Los métodos NGS proporcionan lecturas de secuencias que varían en tamaño de decenas a cientos de pares de bases. En algunas formas de realización del método descrito en el presente documento, las lecturas de secuencia son 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, 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 permitan lecturas de un solo extremo de más de 500 pb, permitiendo lecturas de más de aproximadamente 1000 pb cuando se generen lecturas de extremos emparejados. En una forma de realización, las lecturas de secuencia son de 36 pb. Otros métodos de secuenciación que pueden emplearse por el método de la

invencción 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 canalización de análisis que transforma la salida de imágenes primarias del secuenciador en cadenas de bases. Un paquete de algoritmos integrados realiza las etapas básicas de transformación de datos primarios: análisis de imágenes, puntuación de intensidad, designación de base y alineamiento.

En una forma de realización, se realiza la secuenciación parcial de los fragmentos de ADN presentes en la muestra y se cuentan etiquetas de secuencia que comprenden lecturas de longitud predeterminada, por ejemplo, 36 pb, que se mapean a un genoma de referencia conocido. Solo las lecturas de secuencia que se alinean de forma única con el genoma de referencia se cuentan como etiquetas de secuencia. En una forma de realización, el genoma de referencia es la secuencia NCBI36/hg18 del genoma de referencia humano, que está disponible en Internet en [genoma.ucsc.edu/cgi-bin/hgGateway?org=Human&db=hg18&hgsid=166260105](http://genoma.ucsc.edu/cgi-bin/hgGateway?org=Human&db=hg18&hgsid=166260105). Otras fuentes de información pública sobre secuencias incluyen GenBank, dbEST, dbSTS, EMBL (el Laboratorio Europeo de Biología Molecular) y DDBJ (el Banco de datos de ADN de Japón). En otra forma de realización, el genoma de referencia comprende la secuencia NCBI36/hg18 del genoma de referencia humano y un genoma de secuencias diana artificiales, que incluye secuencias diana polimórficas, por ejemplo, un genoma de SNP que comprende las SEQ ID NO: 1-56. En otra forma de realización más, el genoma de referencia es un genoma de secuencia diana artificial que comprende secuencias diana polimórficas, por ejemplo, secuencias SNP de las SEQ ID NO: 1-56.

El mapeo de las etiquetas de secuencia se realiza comparando la secuencia de la etiqueta con la secuencia del genoma de referencia para determinar el origen cromosómico de la molécula de ácido nucleico secuenciada (por ejemplo, ADNcf), y no se necesita información de secuencia genética específica. Hay varios algoritmos informáticos disponibles para alinear secuencias, incluidos, sin limitación, BLAST (Altschul et al., 1990), BLITZ (MPsrch) (Sturrock & Collins, 1993), FASTA (Person & Lipman, 1988), BOWTIE (Langmead et al., Genome Biology 10:R25.1-R25.10 [2009]) o ELAND (Illumina, Inc., San Diego, CA, Estados Unidos). En una forma de realización, un extremo de las copias expandidas clonalmente de las moléculas de ADNcf de plasma se secuencia y se procesa mediante análisis de alineamiento bioinformático para el analizador de genoma de Illumina, que utiliza el programa informático Efficient Large-Scale Alignment of Nucleotide Databases (ELAND). El análisis de la información de secuenciación para la determinación de la aneuploidía puede permitir un pequeño grado de desajuste (0-2 desajustes por etiqueta de secuencia) para tener en cuenta los polimorfismos secundarios que pueden existir entre el genoma de referencia y los genomas de la muestra mixta. El análisis de la información de secuenciación para la determinación de la fracción fetal puede permitir un pequeño grado de desajuste dependiendo de la secuencia polimórfica. Por ejemplo, se puede permitir un pequeño grado de desajuste si la secuencia polimórfica es una STR. En los casos en los que la secuencia polimórfica es un SNP, todas las secuencias que coinciden exactamente con cualquiera de los dos alelos en el sitio del SNP se cuentan en primer lugar y se filtran de las lecturas restantes, para lo cual se puede permitir un pequeño grado de desajuste.

#### **Preparación de bibliotecas de secuenciación**

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

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 con adaptador, que están listos para ser secuenciados. Las bibliotecas de secuenciación de polinucleótidos se pueden preparar a partir de ADN o ARN, incluidos equivalentes, análogos de ADN o ADNc, que es ADN complementario o copia producido a partir de una plantilla de ARN, por ejemplo, mediante la acción de la transcriptasa inversa. Los polinucleótidos pueden originarse en forma de ADN bicatenario (ADNbc) (por ejemplo, fragmentos de ADN genómico, productos de PCR y de amplificación) o polinucleótidos que pueden haberse originado en forma monocatenaria, tales como ADN o ARN, y convertirse en forma de ADNbc. A modo de ejemplo, las moléculas de ARNm se pueden copiar en ADNc bicatenarios adecuados para su uso en la preparación de una biblioteca de secuenciación. La secuencia precisa de las moléculas de polinucleótidos primarios generalmente no es importante para el método de preparación de bibliotecas y puede ser conocida o desconocida. En una forma de realización, las moléculas de polinucleótidos son moléculas de ADN. Más particularmente, las moléculas de polinucleótidos representan el complemento genético completo de un organismo y son moléculas de ADN genómico, por ejemplo, moléculas de ADNcf, que incluyen secuencias tanto de intrones como de exones (secuencia codificante), así como secuencias reguladoras no codificantes, tales como secuencias promotoras y potenciadoras. Aún más particularmente, las moléculas de polinucleótidos primarios son moléculas de ADN genómico humano, por ejemplo, moléculas de ADNcf presentes en la sangre periférica de un sujeto embarazado. La preparación de bibliotecas de secuenciación para algunas plataformas de secuenciación NGS requiere que los polinucleótidos se encuentren en un intervalo específico de tamaños de fragmentos, por ejemplo, 0-1200 pb. Por lo tanto, es posible que se requiera la fragmentación de polinucleótidos, por ejemplo, ADN genómico. El 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. La fragmentación de moléculas de polinucleótidos por medios mecánicos, por ejemplo, nebulización,

sonicación e hidro-cizallamiento, da como resultado fragmentos con una mezcla heterogénea de extremos romos y salientes en 3' y 5'. Los polinucleótidos, tanto si se fragmentan a la fuerza o existen de forma natural en forma de fragmentos, se convierten en ADN de extremos romos que tienen 5-fosfatos y 3'-hidroxilo.

5 Por lo general, los extremos de los fragmentos se reparan en los extremos, es decir, los extremos se vuelven 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 formas de realización, un solo desoxinucleótido, por ejemplo, la desoxiadenosina (A), se añade a los extremos 3' de los polinucleótidos, por ejemplo, mediante la actividad de determinados tipos de ADN polimerasa, tales como la polimerasa Taq o la polimerasa Klenow  
10 exo minus. 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 unen en una etapa posterior. La adición de colas de dA evita la autoligación de ambos polinucleótidos de extremos romos, de modo que existe un desplazamiento hacia la formación de las secuencias ligadas al adaptador. Los polinucleótidos con cola de dA se ligan a secuencias de polinucleótidos adaptadores bicatenarios. Se puede utilizar 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 tales como ADN ligasa para unir covalentemente el adaptador al polinucleótido con cola de d-A. El adaptador puede contener un resto 5'-fosfato para facilitar la unión al 3'-OH diana. El polinucleótido con cola de dA contiene un resto fosfato en 5', ya sea residual del proceso de cizallamiento o añadido mediante una etapa de tratamiento enzimático, y se ha reparado en los extremos y opcionalmente se ha extendido por una base o bases salientes, para dar un 3'-OH  
15 adecuado para la ligación. Los productos de la reacción de ligación se purifican para eliminar los adaptadores no ligados, los adaptadores que pueden haberse ligado entre sí, y para seleccionar un intervalo de tamaños de plantillas para la generación de clústeres, lo que puede venir precedido por una amplificación, por ejemplo, una amplificación por PCR. La purificación de los productos de ligación se puede obtener mediante métodos que incluyen electroforesis en gel e inmovilización reversible en fase sólida (SPRI).

25 Los protocolos estándar, por ejemplo, los protocolos para la secuenciación que utilizan, por ejemplo, la plataforma Illumina, instruyen a los usuarios en la purificación de los productos reparados antes de la adición de colas de dA y en la purificación de los productos de adición de colas de dA antes de las etapas de ligación del adaptador de la preparación de la biblioteca. La purificación de los productos reparados en los extremos y los productos con cola de dA elimina enzimas, tampones, sales y similares para proporcionar condiciones de reacción favorables para la etapa enzimática posterior. En la invención reivindicada, las etapas de reparación de extremos, adición de colas de dA y ligación de adaptadores excluyen las etapas de purificación. Así, el método de la invención abarca la preparación de una biblioteca de secuenciación que comprende las etapas consecutivas de reparación de extremos, adición de colas de dA y ligación de adaptador. En casos de preparación de bibliotecas de secuenciación que no requieren la etapa de  
30 adición de colas de dA, por ejemplo, protocolos para secuenciación que utilizan las plataformas Roche 454 y SOLID™ 3, las etapas de reparación de extremos y de ligación de adaptadores excluyen la etapa de purificación de los productos de reparación de extremos antes de la ligación de adaptadores.

40 En la etapa siguiente de una forma de realización del método, se prepara una reacción de amplificación. La etapa de amplificación introduce en las moléculas de plantilla ligadas al adaptador las secuencias de oligonucleótidos requeridas para la hibridación con la celda de flujo. El contenido de una reacción de amplificación es conocido por los expertos en la técnica e incluye sustratos apropiados (tales como dNTP), enzimas (por ejemplo, una ADN polimerasa) y componentes de tampón que se requieren para una reacción de amplificación. Opcionalmente, se puede omitir la amplificación de polinucleótidos ligados a adaptadores. Generalmente, las reacciones de amplificación requieren al menos dos cebadores de amplificación, es decir, oligonucleótidos cebadores, que pueden ser idénticos, e incluyen una 'porción específica de adaptador', capaz de hibridarse con una secuencia de unión al cebador en la molécula polinucleotídica que se va a amplificar (o el complemento de la misma si la plantilla se ve como una sola cadena) durante la etapa de hibridación. Una vez formada, la biblioteca de plantillas preparada según los métodos descritos anteriormente se puede utilizar para la amplificación de ácidos nucleicos en fase sólida. El término "amplificación en fase sólida" se refiere, tal como se utiliza en el presente documento, a cualquier reacción de amplificación de ácido nucleico llevada a cabo en un soporte sólido o en asociación con el mismo, de modo que la totalidad o una parte de los productos amplificados se inmovilicen en el soporte sólido a medida que se vayan formando. 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  
45 cebadores de amplificación directo e inverso están inmovilizados en el soporte sólido. La PCR en fase sólida abarca sistemas tales como emulsiones, en los que un cebador está anclado a una perla y el otro en solución libre, y la formación de colonias en matrices de gel en fase sólida en las que un cebador está anclado a la superficie y el otro en solución libre. Después de la amplificación, las bibliotecas de secuenciación se pueden analizar mediante electroforesis capilar microfluídica para garantizar que la biblioteca esté exenta de dímeros adaptadores o ADN monocatenario. La biblioteca de moléculas de polinucleótido de plantilla es particularmente adecuada para su uso en  
50 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 la biblioteca proporcionan plantillas para la amplificación del genoma completo.

65 En una forma de realización, la biblioteca de polinucleótidos ligados a adaptadores se somete a una secuenciación masivamente paralela, que incluye técnicas para secuenciar millones de fragmentos de ácidos nucleicos, por ejemplo, utilizando la unión de ADN genómico fragmentado aleatoriamente a una superficie plana, ópticamente transparente y

la amplificación en fase sólida para crear una celda de flujo de secuenciación de alta densidad con millones de clústeres. Las matrices agrupadas se pueden preparar o bien utilizando un proceso de termociclado, tal como se describe en la patente WO9844151, o bien un proceso mediante el cual la temperatura se mantiene constante y los ciclos de extensión y desnaturalización se realizan mediante cambios de reactivos. El método Solexa/Illumina al que se hace referencia en el presente documento se basa en la unión de ADN genómico fragmentado aleatoriamente a una superficie plana ópticamente transparente. Los fragmentos de ADN unidos se extienden y se amplifican en puente para crear una celda de flujo de secuenciación de densidad ultraalta con millones de clústeres que contienen cada una, miles de copias de la misma plantilla (documentos WO 00/18957 y WO 98/44151). Las plantillas de clústeres se secuencian utilizando una sólida tecnología de secuenciación por síntesis de ADN de cuatro colores que emplea terminadores reversibles con colorantes fluorescentes extraíbles. Alternativamente, la biblioteca se puede amplificar en perlas en las que cada perla contiene un cebador de amplificación directo e inverso.

La secuenciación de las bibliotecas amplificadas se puede llevar a cabo utilizando cualquier técnica de secuenciación adecuada tal como se describe en el presente documento. En una forma de realización, la secuenciación es una secuenciación masivamente paralela que utiliza secuenciación por síntesis con terminadores de colorantes reversibles. En otras formas de realización, la secuenciación es una secuenciación masivamente paralela que utiliza secuenciación por ligación. En otras formas de realización, la secuenciación es secuenciación de una sola molécula.

### Determinación de aneuploidía

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 al genoma de referencia entre muestras dentro de una ejecución de secuenciación (variabilidad intercromosómica), y la variación del número de etiquetas de secuencia que se mapean al genoma de referencia en diferentes ejecuciones de secuenciación (variabilidad intersecuenciación). Por ejemplo, las variaciones pueden ser particularmente pronunciadas para las etiquetas que se mapean a secuencias de referencia ricas en GC o pobres en GC. En un caso, el método utiliza información de secuenciación para calcular la dosis de cromosoma, que intrínsecamente explica la variabilidad acumulada derivada de la variabilidad intercromosómica, intersecuenciación 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 un caso del método **100** con el que se identifican secuencias de normalización, por ejemplo, cromosomas de normalización, y se determina la presencia o la ausencia de una aneuploidía.

En la etapa **110** se obtiene un conjunto de muestras maternas calificadas para identificar secuencias de normalización calificadas, por ejemplo, cromosomas de normalización, y proporcionar valores de varianza para su uso en la determinación de la identificación estadísticamente significativa de una aneuploidía en muestras de ensayo. En la etapa **110** se obtiene una pluralidad de muestras calificadas biológicas 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 un caso, las muestras calificadas se obtienen de madres embarazadas con un feto que se ha confirmado mediante medios citogenéticos que tiene un número normal de copias de cromosomas con respecto al cromosoma de interés. Las muestras biológicas maternas calificadas pueden ser muestras de fluidos biológicos, por ejemplo, muestras de plasma, o cualquier muestra adecuada tal como se ha descrito anteriormente que contiene una mezcla de moléculas de ADNcf fetal y materno. La muestra es una muestra materna que se obtiene de un sujeto femenino embarazado, por ejemplo, una mujer embarazada. Cualquier muestra biológica materna puede usarse como fuente de ácidos nucleicos fetales y maternos que están contenidos en células o que están "libres de células". En algunos casos, es ventajoso obtener una muestra materna que comprenda ácidos nucleicos libres de células, por ejemplo, ADNcf. Preferentemente, la muestra biológica materna es una muestra de fluido biológico. Un fluido biológico incluye, como ejemplos no limitantes, sangre, plasma, suero, sudor, lágrimas, esputo, orina, flujo de oído, linfa, saliva, líquido cefalorraquídeo, estragos, suspensión de médula ósea, flujo vaginal, lavado transcervical, líquido cerebral, ascitis, leche, secreciones del tracto respiratorio, intestinal y genitourinario, líquido amniótico y muestras de leucoforesis. En algunos casos, 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 algunos casos, la muestra biológica es una muestra de sangre periférica, o el plasma y/o las fracciones de suero del mismo. En otros casos, la muestra es una mezcla de dos o más muestras biológicas, por ejemplo, una muestra biológica puede comprender dos o más muestras de fluido biológico. Tal como se utilizan en el presente documento, los términos "sangre", "plasma" y "suero" abarcan expresamente fracciones o porciones procesadas de los mismos. En algunos casos, la muestra biológica se procesa para obtener una fracción de muestra, por ejemplo, plasma, que contiene la mezcla de ácidos nucleicos fetales y maternos. En algunos casos, la mezcla de ácidos nucleicos fetales y maternos 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 fetales y maternos, por ejemplo, ADNcf. Los ácidos nucleicos libres de células, incluido el ADN libre de células, pueden obtenerse mediante varios métodos conocidos en la técnica a partir de muestras biológicas, incluidas, pero sin limitación, 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 ADNcf de las células, se pueden utilizar métodos de fraccionamiento, centrifugación (por ejemplo, centrifugación por gradiente de densidad), precipitación específica de ADN o clasificación y/o separación de células de alto rendimiento. Están disponibles kits comercialmente disponibles para la separación manual y automatizada de ADNcf (Roche Diagnostics, Indianapolis, IN, Qiagen, Valencia, CA, Macherey-Nagel, Duren, DE). En algunos casos, puede resultar ventajoso fragmentar las moléculas de ácido nucleico en la muestra de ácido nucleico. La fragmentación puede ser aleatoria o puede ser específica, tal como se logra, por ejemplo, usando digestión con endonucleasas de restricción. Los métodos para la fragmentación aleatoria son bien conocidos en la técnica e incluyen, por ejemplo, digestión limitada con ADNasa, tratamiento con álcali y cizallamiento físico. En un caso, los ácidos nucleicos de muestra se obtienen como ADNcf, que no se somete a fragmentación. En otros casos, los ácidos nucleicos de muestra se obtienen como ADN genómico, que se somete a fragmentación en fragmentos de aproximadamente 500 pares de bases o más, y a los que se pueden aplicar fácilmente los métodos NGS. Se prepara una biblioteca de secuenciación a partir de ADN fragmentado de forma natural o fragmentado a la fuerza. En un caso, la preparación de la biblioteca de secuenciación comprende las etapas consecutivas de reparación de extremos, adición de colas de dA y ligación de adaptador a los fragmentos de ADN. En un caso, la preparación de la biblioteca de secuenciación comprende las etapas consecutivas de reparación de extremos y ligación de adaptador a los fragmentos de ADN.

En el paso **120**, se secuencia al menos una porción de cada uno de todos los ácidos nucleicos calificados contenidos en las muestras maternas calificadas. Antes de la secuenciación, la mezcla de ácidos nucleicos fetales y maternos, por ejemplo, ADNcf purificado, se modifica para preparar una biblioteca de secuenciación para generar lecturas de secuencias de entre 20 y 40 pb por ejemplo, 36 bp, que están alineadas con un genoma de referencia, por ejemplo, hg18. En algunos casos, 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 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, 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 permitan lecturas de un solo extremo de más de 500 pb, permitiendo lecturas de más de aproximadamente 1000 pb cuando se generen lecturas de extremos emparejados. En un caso, las lecturas de secuencia comprenden 36 pb. Las lecturas de secuencia se alinean con un genoma de referencia humano, y las lecturas que se mapean de forma única al genoma de referencia humano se cuentan como etiquetas de secuencia. En un caso se obtienen al menos aproximadamente  $3 \times 10^6$  etiquetas de secuencia calificadas, al menos aproximadamente  $5 \times 10^6$  etiquetas de secuencia calificadas, al menos aproximadamente  $8 \times 10^6$  etiquetas de secuencia calificadas, al menos aproximadamente  $10 \times 10^6$  etiquetas de secuencia calificadas, al menos aproximadamente  $15 \times 10^6$  etiquetas de secuencia calificadas, al menos aproximadamente  $20 \times 10^6$  etiquetas de secuencia calificadas, al menos aproximadamente  $30 \times 10^6$  etiquetas de secuencia calificadas, al menos aproximadamente  $40 \times 10^6$  etiquetas de secuencia calificadas, o al menos aproximadamente  $50 \times 10^6$  etiquetas de secuencia calificadas que comprenden lecturas de entre 20 y 40 pb a partir de lecturas que se mapean únicamente a un genoma de referencia.

En la etapa **130**, todas las etiquetas obtenidas de la secuenciación de los ácidos nucleicos en las muestras maternas calificadas se cuentan para determinar una densidad de etiqueta de secuencia calificada. En un caso, la densidad de etiquetas de secuencia se determina como el número de etiquetas de secuencia calificadas mapeadas a la secuencia de interés en el genoma de referencia. En otro caso, la densidad de etiquetas de secuencias calificadas se determina como el número de etiquetas de secuencias calificadas mapeadas a una secuencia de interés normalizada a la longitud de la secuencia de interés calificada a la que están mapeadas. Las densidades de etiquetas de secuencia que se determinan como una relación de la densidad de etiquetas con respecto a la longitud de la secuencia de interés se denominan en el presente documento relaciones de densidad de etiquetas. No se requiere la normalización a la longitud de la secuencia de interés y se puede incluir como una etapa para reducir el número de dígitos en un número para simplificarlo para la interpretación humana. Como todas las etiquetas de secuencia calificadas se mapean y se cuentan en cada una de las muestras calificadas, la densidad de etiquetas de secuencia para una secuencia de interés, por ejemplo, un cromosoma de interés, se determina en las muestras calificadas, al igual que las densidades de etiquetas de secuencias para secuencias adicionales a partir de las cuales se identifican secuencias de normalización, por ejemplo, cromosomas, posteriormente. En un caso, la secuencia de interés es un cromosoma asociado con una aneuploidía cromosómica, por ejemplo, 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 etiquetas de secuencia se aproxima mejor a la del cromosoma 21. Por ejemplo, una secuencia de normalización calificada es una secuencia que tiene la menor variabilidad. En algunos casos, la secuencia de normalización es una secuencia que distingue mejor una o más muestras calificadas 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 calificadas y la(s) dosis de cromosoma(s) en una o más muestras de ensayo. En otro caso, la secuencia de interés es un segmento de un cromosoma asociado con una aneuploidía parcial, por ejemplo, una delección o una inserción cromosómica, o una translocación cromosómica desequilibrada, y la secuencia de normalización es un segmento cromosómico que no

está asociado con la aneuploidía parcial y cuya variación en la densidad de etiquetas de secuencia se aproxima mejor a la del segmento cromosómico asociado con la aneuploidía parcial.

5 En la etapa **140**, en base a las densidades de etiquetas calificadas calculadas, se determina una dosis de secuencia calificada para una secuencia de interés como la relación entre la densidad de etiquetas de secuencia para la secuencia de interés y la densidad de etiquetas de secuencia calificada para secuencias adicionales a partir de las cuales se identifican posteriormente las secuencias de normalización. En un caso, las dosis para el cromosoma de interés, por ejemplo, el cromosoma 21, se determinan como una relación de la densidad de etiquetas de secuencia del cromosoma 21 y la densidad de etiquetas de secuencia para cada uno de los cromosomas restantes, es decir, los cromosomas 1-20, el cromosoma 22, el cromosoma X y el cromosoma Y (véanse los ejemplos 3-5 y las figuras 9-15).

15 En la etapa **145**, una secuencia de normalización, por ejemplo, un cromosoma de normalización, se identifica para una secuencia de interés, por ejemplo, el cromosoma 21, en una muestra calificada basada en las dosis de secuencia calculadas. El método identifica secuencias que inherentemente tienen características similares y que son propensas a variaciones similares entre muestras y ejecuciones de secuenciación, y que son útiles para determinar dosis de secuencia en muestras de ensayo. En algunos casos, la secuencia de normalización es la que mejor diferencia una muestra afectada, es decir, una muestra aneuploide, de una o más muestras calificadas. En otros casos, una secuencia de normalización es una secuencia que muestra una variabilidad en el número de etiquetas de secuencia que se mapean a la misma entre muestras y ejecuciones de secuenciación que mejor se aproxima a la secuencia de interés para la que se utiliza como parámetro de normalización y/o que mejor puede diferenciar una muestra afectada de una o más muestras no afectadas.

20 En algunos casos, 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 menor cuando se utiliza la densidad de etiqueta de secuencia del cromosoma 14. En otros casos, se identifican dos, tres, cuatro, cinco, seis, siete, ocho o más secuencias de normalización para su uso en la determinación de una dosis de secuencia para una secuencia de interés en una muestra de ensayo.

30 En un caso, la secuencia de normalización para el cromosoma 21 se selecciona de entre el cromosoma 9, el cromosoma 1, el cromosoma 2, el cromosoma 3, el cromosoma 4, el cromosoma 5, el cromosoma 6, el cromosoma 7, el cromosoma 8, el cromosoma 10, el cromosoma 11, el cromosoma 12, el cromosoma 13, el cromosoma 14, el cromosoma 15, el cromosoma 16 y el cromosoma 17. Preferentemente, la secuencia de normalización para el cromosoma 21 se selecciona de entre el cromosoma 9, el cromosoma 1, el cromosoma 2, el cromosoma 11, el cromosoma 12 y el cromosoma 14. En un caso, la secuencia de normalización para el cromosoma 21 es un grupo de cromosomas seleccionados de entre el cromosoma 9, el cromosoma 1, el cromosoma 2, el cromosoma 3, el cromosoma 4, el cromosoma 5, el cromosoma 6, el cromosoma 7, el cromosoma 8, el cromosoma 10, el cromosoma 11, el cromosoma 12, el cromosoma 13, el cromosoma 14, el cromosoma 15, el cromosoma 16 y el cromosoma 17. En otros casos, la secuencia de normalización para el cromosoma 21 es un grupo de cromosomas seleccionados de entre el cromosoma 9, el cromosoma 1, el cromosoma 2, el cromosoma 11, el cromosoma 12 y el cromosoma 14.

40 En un caso, la secuencia de normalización para el cromosoma 18 se selecciona de entre el cromosoma 8, el cromosoma 2, el cromosoma 3, el cromosoma 4, el cromosoma 5, el cromosoma 6, el cromosoma 7, el cromosoma 9, el cromosoma 10, el cromosoma 11, el cromosoma 12, el cromosoma 13 y el cromosoma 14. Preferentemente, la secuencia de normalización para el cromosoma 18 se selecciona de entre el cromosoma 8, el cromosoma 2, el cromosoma 3, el cromosoma 5, el cromosoma 6, el cromosoma 12 y el cromosoma 14. Alternativamente, la secuencia de normalización para el cromosoma 18 es un grupo de cromosomas seleccionados de entre el cromosoma 8, el cromosoma 2, el cromosoma 3, el cromosoma 4, el cromosoma 5, el cromosoma 6, el cromosoma 7, el cromosoma 9, el cromosoma 10, el cromosoma 11, el cromosoma 12, el cromosoma 13 y el cromosoma 14. En otros casos, la secuencia de normalización para el cromosoma 18 es un grupo de cromosomas seleccionados de entre el cromosoma 8, el cromosoma 2, el cromosoma 3, el cromosoma 5, el cromosoma 6, el cromosoma 12 y el cromosoma 14.

55 En un caso, la secuencia de normalización para el cromosoma X se selecciona de entre el cromosoma 1, el cromosoma 2, el cromosoma 3, el cromosoma 4, el cromosoma 5, el cromosoma 6, el cromosoma 7, el cromosoma 8, el cromosoma 9, el cromosoma 10, el cromosoma 11, el cromosoma 12, el cromosoma 13, el cromosoma 14, el cromosoma 15 y el cromosoma 16. Preferentemente, la secuencia de normalización para el cromosoma X se selecciona de 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 cromosomas seleccionados de entre el cromosoma 1, el cromosoma 2, el cromosoma 3, el cromosoma 4, el cromosoma 5, el cromosoma 6, el cromosoma 7, el cromosoma 8, el cromosoma 9, el cromosoma 10, el cromosoma 11, el cromosoma 12, el cromosoma 13, el cromosoma 14, el cromosoma 15 y el cromosoma 16. En otros casos, la secuencia de normalización para el cromosoma X es un grupo de cromosomas seleccionados de entre el cromosoma 2, el cromosoma 3, el cromosoma 4, el cromosoma 5, el cromosoma 6 y el cromosoma 8.

65 En un caso, la secuencia de normalización para el cromosoma 13 se selecciona de entre el cromosoma 2, el cromosoma 3, el cromosoma 4, el cromosoma 5, el cromosoma 6, el cromosoma 7, el cromosoma 8, el cromosoma 9, el cromosoma 10, el cromosoma 11, el cromosoma 12, el cromosoma 14, el cromosoma 18 y el cromosoma 21.

Preferentemente, la secuencia de normalización para el cromosoma 13 se selecciona de entre el cromosoma 2, el cromosoma 3, el cromosoma 4, el cromosoma 5, el cromosoma 6 y el cromosoma 8. En otro caso, la secuencia de normalización para el cromosoma 13 es un grupo de cromosomas seleccionados de entre el cromosoma 2, el cromosoma 3, el cromosoma 4, el cromosoma 5, el cromosoma 6, el cromosoma 7, el cromosoma 8, el cromosoma 9, el cromosoma 10, el cromosoma 11, el cromosoma 12, el cromosoma 14, el cromosoma 18 y el cromosoma 21. En otros casos, la secuencia de normalización para el cromosoma 13 es un grupo de cromosomas seleccionados de entre el cromosoma 2, el cromosoma 3, el cromosoma 4, el cromosoma 5, el cromosoma 6 y el cromosoma 8.

La variación en la dosis de cromosoma para el cromosoma Y es superior a 30, independientemente de qué cromosoma de normalización se utilice para determinar la dosis del cromosoma Y. Por lo tanto, cualquier cromosoma, o un grupo de dos o más cromosomas seleccionados de entre los cromosomas 1-22 y el cromosoma X, puede usarse como la secuencia de normalización para el cromosoma Y. En un caso, el, al menos un, cromosoma de normalización es un grupo de cromosomas que consiste en los cromosomas 1-22 y el cromosoma X. En otro caso, el, al menos un, cromosoma de normalización es un grupo de cromosomas seleccionados de entre el cromosoma 2, el cromosoma 3, el cromosoma 4, el cromosoma 5 y el cromosoma 6.

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

En la etapa 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 se necesita determinar la presencia o la ausencia de una aneuploidía fetal.

Una biblioteca de secuenciación se prepara tal como se describe para la etapa 120, y en la etapa **125**, al menos una parte de los ácidos nucleicos de ensayo en la muestra de ensayo se secuencian para generar millones de lecturas de secuencias que comprenden entre 20 y 500 pb, por ejemplo, 36pb. Como en la etapa **120**, las lecturas generadas a partir de la secuenciación de los ácidos nucleicos en la muestra de ensayo se mapean de forma única a un genoma humano de referencia y se cuentan. Tal como se describe en la etapa **120**, se obtienen al menos aproximadamente  $3 \times 10^6$  etiquetas de secuencia calificadas, al menos aproximadamente  $5 \times 10^6$  etiquetas de secuencia calificadas, al menos aproximadamente  $8 \times 10^6$  etiquetas de secuencia calificadas, al menos aproximadamente  $10 \times 10^6$  etiquetas de secuencia calificadas, al menos aproximadamente  $15 \times 10^6$  etiquetas de secuencia calificadas, al menos aproximadamente  $20 \times 10^6$  etiquetas de secuencia calificadas, al menos aproximadamente  $30 \times 10^6$  etiquetas de secuencia calificadas, al menos aproximadamente  $40 \times 10^6$  etiquetas de secuencia calificadas, o al menos aproximadamente  $50 \times 10^6$  etiquetas de secuencia calificadas que comprenden lecturas de entre 20 y 40 pb a partir de lecturas que se mapean únicamente a un genoma de referencia humano.

En la etapa **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 ensayo. En un caso, el número de etiquetas de secuencia de ensayo mapeadas a una secuencia de interés se normaliza a la longitud conocida de una secuencia de interés a la que están mapeadas para proporcionar una densidad de etiqueta de secuencia de ensayo. Tal como se ha descrito para las muestras calificadas, no se requiere la normalización a la longitud conocida de una secuencia de interés y se puede incluir como una etapa para reducir el número de dígitos en un número para simplificarlo para la interpretación humana. A medida que todas las etiquetas de secuencia de ensayo mapeadas se cuentan en la muestra de ensayo, la densidad de etiquetas de secuencia para una secuencia de interés, por ejemplo, una secuencia clínicamente relevante, tal como el cromosoma 21, se determina en las muestras de ensayo, al igual que las densidades de etiquetas de secuencia para secuencias adicionales que corresponden a al menos una secuencia de normalización identificada en las muestras calificadas.

En la etapa **150**, basándose en la identidad de al menos una secuencia de normalización en las muestras calificadas, se determina una dosis de secuencia de ensayo para una secuencia de interés en la muestra de ensayo. La dosis de secuencia, por ejemplo, la dosis de cromosoma, para una secuencia de interés en una muestra de ensayo es una relación entre la densidad de etiquetas de secuencia determinada para la secuencia de interés en la muestra de ensayo y la densidad de etiquetas de secuencia de al menos una secuencia de normalización determinada en la muestra de ensayo, en la que la secuencia de normalización en la muestra de ensayo 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 la secuencia de ensayo para el cromosoma 21 (secuencia de interés) se determina como la relación de la densidad de etiquetas de secuencia para el cromosoma 21 y la densidad de etiqueta de secuencia para el cromosoma 14 determinada en la muestra de ensayo. De forma similar, se determinan las dosis de cromosomas para los cromosomas 13, 18, X, Y y otros cromosomas asociados con aneuploidías cromosómicas. Tal como se ha descrito anteriormente, una secuencia de interés puede ser parte de un cromosoma, por ejemplo, un segmento cromosómico. En consecuencia, la dosis para un segmento cromosómico se puede determinar como la relación entre la densidad de etiquetas de secuencia determinada para el segmento en la muestra de ensayo y la densidad de etiquetas de secuencia para el segmento cromosómico de normalización en la muestra de ensayo,

correspondiendo el segmento de normalización en la muestra de ensayo al segmento de normalización identificado en las muestras calificadas para el segmento particular de interés.

En la etapa **155**, los valores umbrales se derivan de los 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. Preferentemente, los umbrales se eligen a partir de la distribución empírica para cada tipo de aneuploidía, por ejemplo, trisomía 21. Los posibles valores umbrales que se establecieron para clasificar las aneuploidías de trisomía 13, trisomía 18, trisomía 21 y monosomía X son tal como se describen en los ejemplos, que describen el uso del método para determinar las aneuploidías cromosómicas mediante la secuenciación de ADNcf extraído de una muestra materna que comprende una mezcla de ácidos nucleicos fetales y maternos.

En la etapa **160**, la variación del número de copias de la secuencia de interés, por ejemplo, la aneuploidía cromosómica o parcial se determina en la muestra de ensayo comparando la dosis de la secuencia de ensayo para la secuencia de interés con al menos un valor umbral establecido a partir de las dosis de la secuencia calificada.

En la etapa **160**, la dosis calculada para una secuencia de ensayo de interés se compara con la establecida como los valores umbrales que se eligen según un umbral de fiabilidad definido por el usuario para clasificar la muestra como "normal", "afectada" o "sin designación" en la etapa **165**. Las muestras "sin designación" son muestras para las que no se puede realizar un diagnóstico definitivo con fiabilidad.

Otro caso divulgado en el presente documento es un método para proporcionar un diagnóstico prenatal de una aneuploidía cromosómica fetal en una muestra biológica que comprende moléculas de ácidos nucleicos fetales y maternos. El diagnóstico se realiza basándose en la recepción de los datos de secuenciación de al menos una porción de la mezcla de moléculas de ácidos nucleicos fetales y maternos derivadas de una muestra de ensayo biológica, por ejemplo, una muestra de plasma materno, calculando a partir de los datos de secuenciación una dosis de cromosoma de normalización para uno o más cromosomas de interés, determinando una diferencia estadísticamente significativa entre la dosis de cromosoma de normalización para el cromosoma de interés en la muestra de ensayo y un valor umbral establecido en una pluralidad de muestras calificadas (normales), y proporcionando el diagnóstico prenatal basado en la diferencia estadística. Tal como se describe en la etapa **165** del método, se realiza un diagnóstico de normal o afectada. Se proporciona una "sin designación" en caso de que el diagnóstico de normal o afectada no se pueda realizar con seguridad.

#### **Determinación de CNV para diagnósticos prenatales**

El ADN y el ARN fetales libres de células que circulan en la sangre materna se pueden utilizar para el diagnóstico prenatal temprano no invasivo (NIPD) de un número cada vez mayor de afecciones genéticas, tanto para la gestión del embarazo como para ayudar en 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 ha descubierto 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 muertas, y se ha demostrado que el ADN fetal libre de células (ADNcf) consiste en fragmentos cortos, por lo general de menos de 200 pb de longitud Chan et al., *Clin Chem* 50:88-92 [2004], que se puede distinguir ya desde las 4 semanas de gestación (Illanes et al., *Early Human Dev* 83:563-566 [2007]), y se sabe que se elimina de la circulación materna a las pocas horas del parto (Lo et al., *Am J Hum Genet* 64:218-224 [1999]). Además del ADNcf, también se pueden distinguir fragmentos de ARN fetal libre de células (ARNcf) en el torrente sanguíneo materno, que se originan a partir de genes que se transcriben en el feto o la placenta. La extracción y posterior análisis de estos elementos genéticos fetales a partir de una muestra de sangre materna ofrece nuevas oportunidades para el NIPD.

El presente método es un método independiente de polimorfismos que se usa en NIPD y que no requiere que el ADNcf fetal se distinga del ADNcf materno para permitir la determinación de una aneuploidía fetal. En algunos casos, la aneuploidía es una trisomía o monosomía cromosómica completa, o una trisomía o monosomía parcial. Las aneuploidías parciales están provocadas por la pérdida o la ganancia de parte de un cromosoma y abarcan los desequilibrios cromosómicos resultantes de translocaciones desequilibradas, inversiones desequilibradas, deleciones e inserciones. Con mucho, la aneuploidía conocida más común compatible con la vida es la trisomía 21, es decir, el síndrome de Down (DS), que está provocado por la presencia de parte o la totalidad del cromosoma 21. En raras ocasiones, el síndrome de Down puede estar provocado por un defecto hereditario o esporádico por el cual una copia adicional de la totalidad o de parte del cromosoma 21 se une a otro cromosoma (generalmente el cromosoma 14) para formar un solo cromosoma aberrante. El DS está asociado con deterioro intelectual, graves dificultades de aprendizaje y un exceso de mortalidad provocado por problemas de salud a largo plazo, tales como enfermedades cardíacas. Otras aneuploidías con significado clínico conocido incluyen el síndrome de Edward (trisomía 18) y el síndrome de Patau (trisomía 13), que frecuentemente son mortales en los primeros meses de vida. También se conocen anomalías asociadas con el número de cromosomas sexuales e incluyen la monosomía X, por ejemplo, el síndrome de Turner (XO) y el síndrome triple X (XXX) en los nacimientos femeninos y el síndrome de Klinefelter (XXY) y el síndrome XYY en los nacimientos masculinos, todos ellos asociados con varios fenotipos que incluyen esterilidad y reducción de las

habilidades intelectuales. El método de la divulgación se puede usar para diagnosticar estas y otras anomalías cromosómicas prenatalmente.

Según los casos de la presente divulgación, la trisomía determinada por los métodos divulgados en el presente documento se selecciona de entre 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 las trisomías XXY (síndrome de Klinefelter), XYY o XXX. Se apreciará que se pueden determinar otras diversas trisomías y trisomías parciales en ADNcf fetal según las enseñanzas del presente documento. Estas incluyen, pero sin limitación, trisomía parcial 1q32-44, trisomía 9p con trisomía, trisomía 4 mosaicismo, trisomía 17p, trisomía parcial 4q26-qter, trisomía 9, trisomía parcial 2p, trisomía parcial 1q y/o trisomía parcial 6p/monosomía 6q.

El método de la presente divulgación también se puede utilizar para determinar monosomía cromosómica X y monosomías parciales tales como la monosomía 13, la monosomía 15, la monosomía 16, la monosomía 21 y la monosomía 22, que se sabe que están involucradas en el aborto espontáneo. La monosomía parcial de los cromosomas típicamente implicados en la aneuploidía completa también puede determinarse mediante el método de la divulgación. La monosomía 18p es un trastorno cromosómico raro en el que se elimina la totalidad o parte del brazo corto (p) del cromosoma 18 (monosómico). El trastorno se caracteriza típicamente por baja estatura, 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 gravedad de un caso a otro. Las condiciones provocadas por 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 actividad génica en la misma parte del cromosoma 15, la región 15q11-q13. Se apreciará que varias translocaciones y microdeleciones pueden ser asintomáticas en el progenitor portador, pero pueden causar una enfermedad genética importante en la descendencia. Por ejemplo, una madre sana que porta la microdelección 15q11-q13 puede dar a luz a un niño con síndrome de Angelman, un trastorno neurodegenerativo grave. Así la presente divulgación se puede utilizar para identificar dicha deleción 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 neonatos 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 que nacen con este trastorno. Casi todos los casos de monosomía parcial 13q se producen aleatoriamente sin razón aparente (esporádicos). El síndrome de deleción 22q11.2, también conocido como síndrome de DiGeorge, es un síndrome provocado por la deleción de una pequeña porción del cromosoma 22. La deleción (22 q11.2) se produce cerca de la mitad del cromosoma en el brazo largo de un cromosoma del par de cromosomas. Las características de este síndrome varían ampliamente, incluso entre miembros de la misma familia, y afectan a muchas partes del cuerpo. Los signos y síntomas característicos pueden incluir defectos de nacimiento tales como cardiopatías congénitas, defectos en el paladar, más comúnmente relacionados con problemas neuromusculares con cierre (insuficiencia velofaríngea), problemas de aprendizaje, diferencias leves en los rasgos faciales e infecciones recurrentes. Las microdeleciones en la región cromosómica 22q112 se asocian con un riesgo de esquizofrenia de 20 a 30 veces superior. En un caso, el método descrito en el presente documento se utiliza para determinar monosomías parciales que incluyen, pero sin limitación, la monosomía 18p, la monosomía parcial del cromosoma 15 (15q11-q13), la monosomía parcial 13q, y la monosomía parcial del cromosoma 22 también se puede determinar usando el método. El ejemplo, 6 y la figura 16 ilustran el uso del método de la divulgación para determinar la presencia de una deleción parcial del cromosoma 11.

El método de divulgación también se puede utilizar para determinar cualquier aneuploidía si uno de los padres es un portador conocido de dicha anomalía. Estas incluyen, pero sin limitación, mosaico para un pequeño cromosoma marcador supernumerario (SMC); translocación t(11;14)(p15;p13); translocación desequilibrada t(8;11)(p23.2;p15.5); microdelección 11q23; síndrome de Smith-Magenis; deleción 17p11.2; deleción 22q13.3; microdelección Xp22.3; deleción 10p14; microdelección 20p, síndrome de DiGeorge [del(22)(q112q1123)], síndrome de Williams (deleciones 7q1123 y 7q36); deleción 1p36; microdelección 2p; neurofibromatosis tipo 1 (microdelección 17q112), deleción Yq; síndrome de Wolf-Hirschhorn (WHS, microdelección 4p16.3); microdelección 1p36.2; deleción 11q14; microdelección 19q132; Rubinstein-Taybi (microdelección 16 p13.3); microdelección 7p21; síndrome de Miller-Dieker (17p13.3), deleción 17p11.2; y microdelección 2q37.

#### **Determinación de CNV de trastornos clínicos**

Además de la determinación temprana de defectos de nacimiento, los métodos descritos en el presente documento pueden aplicarse 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 inadecuados 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, por lo tanto, puede usarse en el diagnóstico de una condición médica, por ejemplo, cáncer.

Los casos de la divulgación se relacionan con 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 ensayo que comprende una mezcla de ácidos nucleicos derivados de dos genomas diferentes, y que se sabe 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.

5 En un caso, la mezcla de ácidos nucleicos se deriva de células normales y cancerosas derivadas de un sujeto que padece una condición médica, por ejemplo, cáncer.

Se cree que muchos tumores sólidos, tales como el cáncer de mama, progresan desde su 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 de desarrollar resistencia a fármacos rápidamente, y angiogénesis, proteólisis y metástasis potenciadas. Las aberraciones genéticas pueden afectar tanto a los "genes supresores de tumores" recesivos como a los oncogenes de acción dominante. Se cree que las deleciones 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.

Se ha encontrado ADNcf en la circulación de pacientes diagnosticados con neoplasias malignas que incluyen, pero sin limitación, cáncer de pulmón (Pathak et al. *Clin Chem* 52:1833-1842 [2006]), cáncer de prostata (Schwartzbach et al. *Clin Cancer Res* 15:1032-8 [2009]) y cáncer de mama (Schwartzbach et al. disponible en línea en [breast-cancer-research.com/content/11/5/R71](http://breast-cancer-research.com/content/11/5/R71) [2009]). La identificación de inestabilidades genómicas asociadas con cánceres que se pueden determinar en el ADNcf en circulación en pacientes con cáncer es una herramienta potencial de diagnóstico y pronóstico. En un caso, 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 del que se sospecha o se sabe que tiene cáncer, por ejemplo, carcinoma, sarcoma, linfoma, leucemia, tumores de células germinales y blastoma. En un caso, 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 otro caso, 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, pero sin limitación, suero, sudor, lágrimas, esputo, orina, esputo, flujo del oído, linfa, saliva, líquido cefalorraquídeo, estragos, suspensión de médula ósea, flujo vaginal, lavado transcervical, líquido cerebral, ascitis, leche, secreciones de las vías respiratorias, intestinales y genitourinarias y muestras de leucoforesis, o en biopsias de tejido, hisopos o frotis.

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

Los genes de acción dominante asociados con tumores sólidos humanos ejercen típicamente su efecto por sobreexpresión o expresión alterada. La amplificación génica es un mecanismo común que conduce a la regulación al alza 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 protooncogén receptor 2 del factor de crecimiento epidérmico humano (HER2) ubicado en el cromosoma 17 (17(17q21-q22)) da como resultado una sobreexpresión de los receptores HER2 en la superficie celular, lo que da lugar a una señalización excesiva y desregulada en el cáncer de mama y otras neoplasias malignas (Park et al., *Clinical Breast Cancer* 8:392-401 [2008]). Se ha encontrado que una diversidad de oncogenes se amplifica en otras neoplasias malignas humanas. Los ejemplos de amplificación de oncogenes celulares en tumores humanos incluyen amplificaciones de: c-myc en la línea celular de leucemia promielocítica HL60 y en líneas celulares de carcinoma de pulmón de células pequeñas, N-myc en neuroblastomas primarios (estadios III y IV), líneas celulares de neuroblastoma, línea celular de retinoblastoma y tumores primarios, y líneas y tumores de carcinoma de pulmón de células pequeñas, L-myc en líneas celulares y tumores de carcinoma de pulmón de células pequeñas, c-myc en leucemia mieloide aguda y en líneas celulares de carcinoma de colon, c-erbB en células de carcinoma epidermoide y gliomas primarios, cK-ras-2 en carcinomas primarios de pulmón, colon, vejiga y recto, N-ras en la línea celular de carcinoma mamario (Varmus H., *Ann Rev Genética* 18: 553-612 (1984) [citado por Watson et al., *Molecular Biology of the Gene* (4<sup>a</sup> ed.; Benjamin/Cummings Publishing Co. 1987)).

Las deleciones cromosómicas que involucran genes supresores de tumores pueden desempeñar un papel importante en el desarrollo y la progresión de tumores sólidos. El gen supresor de tumores de retinoblastoma (Rb-1), ubicado en el cromosoma 13q14, es el gen supresor de tumores caracterizado más ampliamente. 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 está provocada por la inactivación de los alelos de ambos genes ya sea a través de una mutación puntual o una deleción cromosómica. Se ha descubierto que las alteraciones del gen Rb-1 están presentes no solo en los retinoblastomas, sino también en otras neoplasias malignas tales como osteosarcomas, cáncer de pulmón de células pequeñas (Rygaard et al., *Cancer Res* 50: 5312-5317 [1990]) y cáncer de mama. Los estudios de polimorfismo de longitud de fragmentos de restricción (RFLP) han indicado que dichos tipos de tumores con frecuencia han perdido heterocigosidad en 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 que

incluyen duplicaciones, deleciones y translocaciones desequilibradas que involucran al cromosoma 6 y otros cromosomas asociados indican que las regiones del cromosoma 1, en particular 1q21-1q32 y 1p11-13, podrían albergar oncogenes o genes supresores de tumores que son patogenéticamente relevantes para fases de neoplasias mieloproliferativas tanto crónicas como avanzadas (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 la anomalía cariotípica más común en síndromes mielodisplásicos (SMD). Los pacientes con SMD del(5q)/5q- aislados tienen un pronóstico más favorable que aquellos con defectos cariotípicos adicionales, que tienden a desarrollar neoplasias mieloproliferativas (NMP) y leucemia mieloide aguda. La frecuencia de deleciones 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 (HSC/HPC). El mapeo citogenético de regiones comúnmente eliminadas (CDR) centrado en 5q31 y 5q32 identificó genes candidatos a supresores de tumores, incluida la subunidad ribosómica RPS14, el factor de transcripción Egr1 / Krox20 y la proteína de remodelación del citoesqueleto, alfa-catenina (Eisenmann et al., Oncogen 28:3429-3441 [2009]). Los estudios citogenéticos y de alelotipado de tumores recientes y líneas de células 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 frecuentes involucradas en un amplio espectro de los principales cánceres epiteliales de pulmón, mama, riñón, cabeza y cuello, ovario, cuello uterino, colon, páncreas, esófago, vejiga y otros órganos. Se han mapeado varios genes supresores de tumores 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 de la totalidad del cromosoma 3 en el desarrollo de carcinomas. (Angeloni D., Briefings Functional Genomics 6:19-39 [2007]).

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 localizados en el cromosoma 21 que desempeñan un papel importante en la tumorigénesis. Las aberraciones numéricas somáticas, así como estructurales del cromosoma 21 están asociadas con leucemias, y genes específicos, incluidos RUNX1, Tmprss2 y TFF, que se encuentran en 21q, y desempeñan un papel en la tumorigénesis. (Fonatsch C Gene Chromosomes Cancer 49: 497-508 [2010]).

En un caso, el método se refiere a un medio para evaluar la asociación entre la amplificación génica y el grado de evolución del tumor. La correlación entre la amplificación y/o deleción y el estadio o grado de un cáncer puede ser importante desde el punto de vista del pronóstico porque dicha información puede contribuir a la definición de un grado tumoral con base genética que podría predecir mejor el curso futuro de la enfermedad, teniendo los tumores más avanzados el peor pronóstico. Además, la información sobre eventos tempranos de amplificación y/o deleción puede ser útil para asociar esos eventos como factores de pronóstico de la progresión posterior de la enfermedad. La amplificación y las deleciones de genes identificadas por el método se pueden asociar con otros parámetros conocidos, tales como el grado del tumor, la histología, el índice de marcado Brd/Urđ, el estado hormonal, la afectación de los ganglios, el tamaño del tumor, la duración de la supervivencia y otras propiedades del tumor disponibles a partir de estudios epidemiológicos y bioestadísticos. Por ejemplo, el ADN tumoral que se va a analizar mediante el método podría incluir hiperplasia atípica, carcinoma ductal in situ, cáncer en estadio I-III y ganglios linfáticos metastásicos para permitir la identificación de asociaciones entre amplificaciones y deleciones y estadio. Las asociaciones realizadas pueden posibilitar una intervención terapéutica eficaz. Por ejemplo, las regiones amplificadas sistemáticamente pueden contener un gen sobreexpresado, cuyo producto puede ser atacado terapéuticamente (por ejemplo, la tirosina quinasa del receptor del factor de crecimiento, p185<sup>HER2</sup>).

El método se puede utilizar para identificar eventos de amplificación y/o deleción que están asociados con la resistencia a medicamentos mediante la determinación de la variación del número de copias de ácidos nucleicos de cánceres primarios con respecto a los de células que se han metastatizado en otros sitios. Si la amplificación y/o la deleción de genes es una manifestación de inestabilidad cariotípica que permite el desarrollo rápido de resistencia a medicamentos, se esperaría una mayor amplificación y/o deleción en tumores primarios de pacientes quimiorresistentes que en tumores de pacientes quimiosensibles. Por ejemplo, si la amplificación de genes específicos es responsable del desarrollo de la resistencia a medicamentos, se esperaría que las regiones que rodean esos genes se amplifiquen sistemáticamente en las células tumorales de derrames pleurales de pacientes quimiorresistentes, pero no en los tumores primarios. El descubrimiento de asociaciones entre la amplificación y/o la deleción de genes y el desarrollo de resistencia a medicamentos puede permitir la identificación de pacientes que se beneficiarán o no de un tratamiento adyuvante.

#### **Determinación simultánea de aneuploidía y fracción fetal**

En otro caso, el método permite la determinación simultánea de la fracción del componente secundario de ácido nucleico fetal, es decir, la 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 aportada 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 la fracción fetal se determina mediante la contribución relativa de un alelo polimórfico derivado del genoma

fetal a la contribución del alelo polimórfico correspondiente derivado del genoma materno. Se pueden utilizar secuencias polimórficas junto con ensayos de diagnóstico clínicamente relevantes como control positivo de la presencia de ADNcf con el fin de resaltar resultados falsos negativos o falsos positivos derivados de niveles bajos de ADNcf por debajo del límite de identificación. El método descrito es útil en una diversidad de edades gestacionales.

5 En las **figuras 2-5** se representan casos ilustrativos del método para determinar simultáneamente la fracción fetal y la presencia o ausencia de una aneuploidía de la forma siguiente.

10 La figura 2 proporciona un diagrama de flujo de un caso 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 ensayo que comprende una mezcla de ácidos nucleicos fetales y maternos. Las muestras de ensayo incluyen las muestras descritas en la etapa **110** del caso del método **100**. En algunos casos, la muestra de ensayo es una muestra de sangre periférica obtenida de un sujeto femenino embarazado, por ejemplo, una mujer embarazada. En la etapa **220** la mezcla de ácidos nucleicos presente en la muestra se enriquece en ácidos nucleicos diana polimórficos, cada uno de los cuales comprende un sitio polimórfico. En algunos casos, los ácidos nucleicos que se enriquecen son ADNcf. Los ácidos nucleicos diana son segmentos de material genético que se sabe que comprenden al menos un sitio polimórfico. En algunos casos, los ácidos nucleicos diana comprenden un SNP. En otros casos, el ácido nucleico diana comprende un STR. En otros casos más, los ácidos nucleicos diana comprenden un STR en tándem. El enriquecimiento de una mezcla de ácidos nucleicos fetales y maternos comprende amplificar secuencias diana de una porción de ácidos nucleicos contenidos en la muestra materna original y combinar parte o la totalidad del producto amplificado con el resto de la muestra materna original. En la etapa **230**, se secuencian al menos una parte de la mezcla enriquecida, se identifican las diferencias de secuencia derivadas de la naturaleza polimórfica de las secuencias diana y se determina la contribución relativa de las secuencias polimórficas derivadas del genoma fetal, es decir, la fracción fetal, en la etapa **240**. En algunos casos, la muestra de ensayo materna original es una muestra de fluido biológico, por ejemplo, plasma. En otros casos, la muestra materna original es una fracción procesada de plasma que comprende ADNcf fetal y materno purificado.

### Secuencias polimórficas

30 Los sitios polimórficos que están contenidos en los ácidos nucleicos diana incluyen, sin limitación, polimorfismos de un solo nucleótido (SNP), SNP en tándem, deleciones o inserciones de múltiples bases a pequeña escala, denominadas IN-DELS (también denominadas polimorfismos de deleción-inserción o DIP), polimorfismos de múltiples nucleótidos (MNP) y repeticiones en tándem cortas (STR). Los sitios polimórficos que abarca el método de la divulgación están ubicados en cromosomas autosómicos, lo que permite la determinación de la fracción fetal independientemente del sexo del feto. Cualquier sitio polimórfico que pueda estar abarcado por las lecturas generadas por los métodos de secuenciación descritos en el presente documento puede utilizarse para determinar simultáneamente la fracción fetal y la presencia o la ausencia de una aneuploidía en una muestra materna.

40 En un caso, la mezcla de ácidos nucleicos fetales y maternos en la muestra se enriquece en ácidos nucleicos diana que comprenden al menos un SNP. En algunos casos, los ácidos nucleicos diana comprenden un único SNP, es decir, uno. Las secuencias de ácido nucleico diana que comprenden SNP están disponibles en bases de datos de acceso público, incluidas, pero sin limitación, la base de datos de SNP humanos en la dirección de Internet [wi.mit.edu](http://wi.mit.edu), la página de inicio de NCBI dbSNP en la dirección de Internet [ncbi.nlm.nih.gov](http://ncbi.nlm.nih.gov), la dirección de Internet [lifesciences.perkinelmer.com](http://lifesciences.perkinelmer.com), la base de datos Celera Human SNP en la dirección de Internet [celera.com](http://celera.com), la base de datos SNP del Genome Analysis Group (GAN) en la dirección de Internet [gan.iarc.fr](http://gan.iarc.fr). En un caso, los SNP elegidos para enriquecer el ADNcf fetal y materno se seleccionan del grupo de 92 SNP de identificación individual (IISNP) descrito por Pakstis et al. (Pakstis et al. *Hum Genet* 127:315-324 [2010]), que se ha demostrado que tiene una variación muy pequeña en la frecuencia entre las poblaciones ( $F_{st} < 0,06$ ), y que 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 SNP vinculados y no vinculados. Cada ácido nucleico diana comprende al menos un sitio polimórfico, por ejemplo, un único 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, se puede configurar un panel de SNP para que comprenda al menos un SNP informativo.

60 En un caso, los SNP que son el objetivo de la amplificación se seleccionan 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.

65 En otros casos, cada ácido nucleico diana comprende dos o más SNP, es decir, cada ácido nucleico diana comprende SNP en tándem. Preferentemente, 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 en el presente documento como conjuntos de dos SNP. Para identificar secuencias de SNP en tándem adecuadas, se puede buscar en la base de datos del

International HapMap Consortium (The International HapMap Project, Nature 426:789-796 [2003]). La base de datos está disponible en Internet en hapmap.org. En un caso, los SNP en tándem que son el objetivo de la amplificación se seleccionan de entre los siguientes conjuntos de pares en tándem de SNP rs7277033-rs2110153; rs2822654-rs1882882; rs368657-rs376635; rs2822731-rs2822732; rs1475881-rs7275487; rs1735976-rs2827016; rs447340-rs2824097; rs418989-rs13047336; rs987980-rs987981; rs4143392-rs4143391; rs1691324-rs13050434; rs11909758-rs9980111; rs2826842-rs232414; rs1980969-rs1980970; rs9978999-rs9979175; rs1034346-rs12481852; rs7509629-rs2828358; rs4817013-rs7277036; rs9981121-rs2829696; rs455921-rs2898102; rs2898102-rs458848; rs961301-rs2830208; rs2174536-rs458076; rs11088023-rs11088024; rs1011734-rs1011733; rs2831244-rs9789838; rs8132769-rs2831440; rs8134080-rs2831524; rs4817219-rs4817220; rs2250911-rs2250997; rs2831899-rs2831900; rs2831902-rs2831903; rs11088086-rs2251447; rs2832040-rs11088088; rs2832141-rs2246777; rs2832959-rs9980934; rs2833734-rs2833735; rs933121-rs933122; rs2834140-rs12626953; rs2834485-rs3453; rs9974986-rs2834703; rs2776266-rs2835001; rs1984014-rs1984015; rs7281674-rs2835316; rs13047304-rs13047322; rs2835545-rs4816551; rs2835735-rs2835736; rs13047608-rs2835826; rs2836550-rs2212596; rs2836660-rs2836661; rs465612-rs8131220; rs9980072-rs8130031; rs418359-rs2836926; rs7278447-rs7278858; rs385787-rs367001; rs367001-rs386095; rs2837296-rs2837297; y rs2837381-rs4816672.

En otro caso, la mezcla de ácidos nucleicos fetales y maternos en la muestra se enriquece en ácidos nucleicos diana que comprenden al menos un STR. Los loci STR se encuentran en casi todos los cromosomas del genoma y se pueden amplificar utilizando una diversidad de cebadores de reacción en cadena de la polimerasa (PCR). Los científicos forenses han preferido las repeticiones de tetranucleótidos debido a su fidelidad en la amplificación por PCR, aunque también se utilizan algunas repeticiones de trinucleótidos y pentanucleótidos. En STRBase se compila una lista completa de referencias, hechos e información de secuencias sobre STR, cebadores de PCR publicados, sistemas multiplex comunes y datos de población relacionados, a la que se puede acceder a través de Internet en [ibm4.carb.nist.gov:8800/adn/home.htm](http://ibm4.carb.nist.gov:8800/adn/home.htm). La información de secuencia de GenBank® (<http://www2.ncbi.nlm.nih.gov/cgi-bin/genbank>) para loci STR de uso común también está disponible a través de STRBase. La naturaleza polimórfica de las secuencias de ADN repetidas en tándem que están muy extendidas en todo el genoma humano las ha convertido en marcadores genéticos importantes para estudios de mapeo de genes, análisis de unión y pruebas de identidad humana. Debido al alto polimorfismo de los STR, la mayor parte de los individuos serán heterocigóticos, es decir, la mayor parte de las personas poseerá dos alelos (versiones) de cada uno heredado de cada padre, con un número diferente de repeticiones. Por lo tanto, la secuencia STR fetal no heredada de la madre diferirá en el número de repeticiones de la secuencia materna. La amplificación de estas secuencias STR dará como resultado dos productos de amplificación principales correspondientes a los alelos maternos (y el alelo fetal heredado de la madre) y un producto secundario correspondiente al alelo fetal no heredado de la madre. Esta técnica se notificó por primera vez en 2000 (Pertl et al., Human Genetics 106:45-49 [2002]) y posteriormente se ha desarrollado 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]). Así, la fracción de ácido nucleico fetal en una muestra materna también puede determinarse mediante la secuenciación de ácidos nucleicos diana polimórficos que comprenden STR, que varían entre individuos en el número de unidades repetidas en tándem entre alelos. En un caso, la determinación simultánea de aneuploidía y fracción fetal comprende la secuenciación de al menos una parte de los ácidos nucleicos fetales y maternos presentes en una muestra materna que se ha enriquecido en secuencias polimórficas que comprenden STR. Dado que el tamaño del ADNcf fetal es < 300 pb, las secuencias polimórficas comprenden miniSTR, que se pueden amplificar para generar amplicones cuya longitud es aproximadamente del tamaño de los fragmentos de ADN fetal circulante. El método puede utilizar uno o una combinación de cualquier número de miniSTR informativos para determinar la fracción de ácido nucleico fetal. Por ejemplo, puede utilizarse cualquiera o una combinación de cualquier número de miniSTR, por ejemplo, los miniSTR divulgados en la tabla 22. En un caso, la fracción de ácido nucleico fetal en una muestra materna se realiza mediante un método que incluye la determinación del número de copias del ácido nucleico materno y fetal presentes en la muestra materna mediante la amplificación de al menos un miniSTR autosómico elegido de entre CSF1PO, FGA, TH01, TPOX, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, PentaD, 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 otro caso, el, al menos un, miniSTR autosómico es el grupo de miniSTR CSF1PO, FGA, D13S317, D16S539, D18S51, D2S1338, D21S11 y D7S820.

El enriquecimiento de la muestra en los ácidos nucleicos diana se realiza mediante métodos que comprenden amplificar específicamente las 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, incluidas, pero sin limitación, PCR asimétrica, amplificación dependiente de helicasa, PCR de inicio en caliente, qPCR, PCR en fase sólida y PCR de contacto. Alternativamente, la replicación de las secuencias de ácidos nucleicos diana se puede obtener mediante métodos independientes de enzimas, por ejemplo, síntesis química en fase sólida utilizando las fosforamiditas. La amplificación de las secuencias diana se realiza utilizando pares de cebadores, cada uno de los cuales es 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 amplificadas que comprenden al menos dos, al menos tres, al menos 5, al menos 10, al menos 15, al menos 20, al menos 25, al menos 30, al menos 30, al menos 35, al menos 40 o más sitios polimórficos en la misma reacción de secuenciación. Cualquier panel de conjuntos de cebadores se puede configurar para amplificar al menos una secuencia polimórfica informativa.

5

### Amplificación de secuencias polimórficas

Una serie de cebadores de ácidos nucleicos que están ya disponibles para amplificar fragmentos de ADN que contienen los polimorfismos de SNP y sus secuencias se pueden obtener, por ejemplo, de las bases de datos identificadas anteriormente. También se pueden diseñar cebadores adicionales, por ejemplo, utilizando un método similar al publicado por Vieux, E. F., Kwok, P-Y y Miller, R. D. en *BioTechniques* (junio de 2002) vol. 32. Suplemento: "SNPs: Discovery of Marker Disease, páginas 28-32. En un caso, 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. En un caso, los conjuntos de cebadores comprenden cebadores directos e inversos que abarcan al menos un SNP informativo seleccionado 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 el ejemplo, 7 y en las tablas 10 y 11 se proporcionan ejemplos de conjuntos de cebadores que se utilizan para amplificar los SNP descritos en el presente documento, y se divulgan como SEQ ID NO: 57-112. En otro caso, el grupo de 13 conjuntos de cebadores SEQ ID NO: 57-82 se utiliza para amplificar un ácido nucleico objetivo, cada uno de los cuales comprende al menos un SNP, por ejemplo, un solo SNP, en una porción de una mezcla de ADNcf fetal y materno.

25

En otro caso, se utiliza al menos un conjunto de cebadores para amplificar un ácido nucleico diana, cada uno de los cuales comprende al menos un SNP en tándem, por ejemplo, un conjunto de dos SNP en tándem, en una porción de una mezcla de ADNcf fetal y materno. En un caso, los conjuntos de cebadores comprenden cebadores directos e inversos que abarcan al menos un SNP en tándem informativo seleccionado de entre SNP 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 diana que comprenden los SNP en tándem están diseñados para abarcar ambos sitios SNP. En el ejemplo, 12 se proporcionan conjuntos de cebadores ilustrativos que se utilizan para amplificar los SNP divulgados en el presente documento y se divulgan como SEQ ID NO: 197-310.

45

La amplificación de los ácidos nucleicos diana se realiza utilizando cebadores específicos de secuencia que permiten la amplificación específica de secuencia. Por ejemplo, los cebadores de PCR están diseñados para discriminar la amplificación de genes similares o parálogos que se encuentran en otros cromosomas aprovechando las diferencias de secuencia entre el ácido nucleico diana y cualquier parálogo de otros cromosomas. Los cebadores de PCR directos o inversos están diseñados para hibridarse cerca del sitio SNP y amplificar una secuencia de ácido nucleico de longitud suficiente para incluirla 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 en puente que puede utilizarse opcionalmente antes de la secuenciación. Así, los cebadores de PCR utilizados para amplificar los ácidos nucleicos diana están diseñados para amplificar secuencias que tienen una longitud suficiente para ser amplificadas en puente e identificar los SNP que están abarcados por las lecturas de secuencias. En algunos casos, el primero de dos cebadores del conjunto de cebadores que comprende el cebador directo y el cebador inverso para amplificar el ácido nucleico diana está diseñado para identificar un solo SNP presente dentro de una lectura de secuencia 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, 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 las tecnologías de secuenciación masivamente paralela permitan lecturas de un solo extremo de más

65

de 500 pb. En un caso, uno de los cebadores de PCR está diseñado para amplificar SNP que están incluidos 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 en puente. En un caso, los cebadores de PCR ilustrativos están diseñados para amplificar los ácidos nucleicos diana que contienen un solo SNP seleccionado de 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. En otros casos, los cebadores directo e inverso están diseñados para amplificar los ácidos nucleicos diana, cada uno de los cuales comprende un conjunto de dos SNP en tándem, estando presente cada uno de los mismos en una lectura de secuencia 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, 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. En un caso, 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 en puente.

Los SNP, SNP individuales o en tándem, están contenidos en amplicones de ácido nucleico diana amplificados de al menos 100 pb, al menos 150 pb, al menos 200 pb, al menos 250 pb, al menos 300 pb, al menos 350 pb o al menos 400 pb. En un caso, 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 los 36 pb desde el extremo 3' o 5' del amplicón. En otro caso, 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 los 36 pb desde el extremo 3' del amplicón, y/o el segundo SNP dentro de los 36 pb desde el extremo 5' del amplicón.

En un caso, 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**

Una serie de cebadores de ácidos nucleicos que están ya disponibles para amplificar fragmentos de ADN que contienen los STR 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 fetal y materno circulantes, y se ha demostrado que las moléculas de ADN fetal en el plasma de las mujeres embarazadas son generalmente más cortas que las moléculas de ADN materno. Chan et al., Clin Chem 50:8892 [2004]. El fraccionamiento del tamaño del ADN fetal circulante ha confirmado que la longitud 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 coherentes con los de Fan et al., quien determinó utilizando NGS que el ADNcf fetal rara vez es > 340bp (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, cada uno de los cuales comprende un miniSTR que comprende la cuantificación de al menos un alelo fetal y uno materno en un miniSTR polimórfico, que se puede amplificar para generar amplicones cuya longitud es aproximadamente del tamaño de los fragmentos de ADN fetal circulante.

En un caso, el método comprende determinar el número de copias de al menos un alelo fetal y al menos un alelo materno en al menos un miniSTR polimórfico que se amplifica para generar amplicones que tienen menos de aproximadamente 300 pb, menos de aproximadamente 250 pb, menos de aproximadamente 200 pb, menos de aproximadamente 150 pb, menos de aproximadamente 100 pb o menos de aproximadamente 50 pb. En otro caso, los amplicones que se generan al amplificar los miniSTR tienen menos de aproximadamente 300 pb. En otro caso, los amplicones que se generan al amplificar los miniSTR tienen menos de aproximadamente 250 pb. En otro caso, los amplicones que se generan al amplificar los miniSTR tienen menos de aproximadamente 200 pb. La amplificación del alelo informativo incluye el uso de cebadores miniSTR, que permiten la amplificación de amplicones de tamaño reducido para discernir alelos STR que tienen menos de aproximadamente 500 pb, menos de aproximadamente 450 pb, menos de aproximadamente 400 pb, menos de aproximadamente 350 pb, menos de aproximadamente 300 pares de bases (pb), menos de aproximadamente 250 pb, menos de aproximadamente 200 pb, menos de aproximadamente 150 pb, menos de aproximadamente 100 pb o menos de aproximadamente 50 pb. Los amplicones de tamaño reducido generados por medio de los cebadores miniSTR se conocen como miniSTR y se identifican según el nombre del marcador correspondiente al locus al que se han mapeado. En un caso, los cebadores miniSTR incluyen cebadores

miniSTR que han permitido la reducción máxima del tamaño del amplicón para los 13 loci CODIS STR además del D2S1338, pentaD y 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 tal como se describe por Coble y Butler (Coble y Butler, J Forensic Sci 50:43-53 [2005]) y otros miniSTR que se han caracterizado en el NIST. Se puede acceder a la información sobre los miniSTR caracterizados en el NIST a través de Internet en [cstl.nist.gov/biotech/strbase/newSTRs.htm](http://cstl.nist.gov/biotech/strbase/newSTRs.htm). Puede usarse cualquier par o una combinación de dos o más pares de cebadores miniSTR para amplificar al menos un miniSTR. Por ejemplo, se selecciona al menos un conjunto de cebadores de los conjuntos de cebadores proporcionados en la tabla 22 (ejemplo, 11) y divulgados como SEQ ID NO: 113-196 que se puede usar para amplificar secuencias diana polimórficas que comprenden un STR.

El enriquecimiento de la muestra se obtiene amplificando ácidos nucleicos diana contenidos en una porción de la mezcla de ácidos nucleicos fetales y maternos en la muestra original, y combinando al menos una porción o la totalidad del producto amplificado con el resto de la muestra original no amplificada. El enriquecimiento comprende la amplificación de los ácidos nucleicos diana que están contenidos en una porción de la muestra de fluido biológico. En un caso, la muestra que se enriquece es la fracción de plasma de una muestra de sangre (véase la figura 3). Por ejemplo, se usa una porción de una muestra de plasma materno original para amplificar las secuencias de ácido nucleico diana. Posteriormente, parte o la totalidad del producto amplificado se combina con la muestra de plasma original sin amplificar restante, enriqueciéndola así (véase el ejemplo, 10). En otro caso, la muestra que se enriquece es la muestra de ADNcf purificado que se extrae del plasma (véase la figura 4). Por ejemplo, el enriquecimiento comprende amplificar los ácidos nucleicos diana que están contenidos en una porción de una muestra original de una mezcla purificada de ácidos nucleicos fetales y maternos, por ejemplo, ADNcf que se ha purificado a partir de una muestra de plasma materno, y posteriormente combinar parte o la totalidad del producto amplificado con la muestra purificada original no amplificada restante (véase el ejemplo, 9). En otro caso más, la muestra que se enriquece es una muestra de biblioteca de secuenciación preparada a partir de una mezcla purificada de ácidos nucleicos fetales y maternos (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 una mezcla purificada de ácidos nucleicos fetales y maternos, por ejemplo, ADNcf que se ha purificado a partir de una muestra de plasma materno, preparar una primera biblioteca de secuenciación de secuencias de ácidos nucleicos no amplificados, preparar una segunda biblioteca de secuenciación de ácidos nucleicos diana polimórficos amplificados y, posteriormente, combinar parte o la totalidad de la segunda biblioteca de secuenciación con parte o la totalidad de la primera biblioteca de secuenciación (véase el ejemplo, 8). La cantidad de producto amplificado que se usa para enriquecer la muestra original se selecciona para obtener suficiente información de secuenciación para determinar tanto la presencia o la ausencia de aneuploidía como la fracción fetal del mismo ciclo de secuenciación. 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 mapean para determinar la fracción fetal.

En un caso, la etapa de enriquecer la mezcla de ácidos nucleicos fetales y maternos en ácidos nucleicos diana polimórficos comprende amplificar los ácidos nucleicos diana en una porción de una muestra de ensayo, por ejemplo, una muestra de ensayo de plasma, y combinar la totalidad o una porción del producto amplificado con la muestra de ensayo de plasma restante. El caso del método **300** se representa en el diagrama de flujo proporcionado en la figura 3. En la etapa **310**, una muestra de ensayo, por ejemplo, una muestra de fluido biológico tal como una muestra de sangre, se obtiene de una mujer embarazada, y en la etapa **320** una porción del ADNcf contenido en la fracción de plasma de la muestra de sangre se utiliza para amplificar los ácidos nucleicos diana que comprenden sitios polimórficos, por ejemplo, SNP. En un caso, se usó al menos aproximadamente el 1%, al menos aproximadamente el 1,5%, al menos aproximadamente el 2% y al menos aproximadamente el 10% del plasma materno para amplificar los ácidos nucleicos diana. En la etapa **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 la etapa **340**, y se usan para preparar una biblioteca que se secuenció en la etapa **350**. La biblioteca se preparó a partir de ADNcf purificado y comprendía 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 producto amplificado. En la etapa **360**, se analizan los datos de las ejecuciones de secuenciación y se realiza la determinación simultánea de la fracción fetal y la presencia o la ausencia de aneuploidía.

En un caso, la etapa de enriquecer la mezcla de ácidos nucleicos fetales y maternos en 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 un caso, una porción de una mezcla de ácidos nucleicos fetales y maternos, por ejemplo, ADNcf, purificados a partir de una muestra de plasma materno, se utiliza para amplificar secuencias de ácidos nucleicos polimórficos, y una parte del producto amplificado se combina con la mezcla no amplificada de ácidos nucleicos fetales y maternos purificados, por ejemplo, ADNcf (véase la figura 4). El caso del método **400** se representa en el diagrama de flujo proporcionado en la figura 4. En la etapa **410**, una muestra de ensayo, por ejemplo, una muestra de fluido biológico tal como una muestra de sangre, que comprende una mezcla de ácidos nucleicos fetales y maternos, se obtiene de una mujer embarazada, y la mezcla de ácidos nucleicos fetales y maternos se purifica a partir de la fracción de plasma en la etapa **420**. Tal como se ha

descrito anteriormente, los métodos para la separación del ADN libre de células del plasma son bien conocidos. En la etapa **430**, una porción del ADNcf contenido en la muestra purificada se utiliza para amplificar los 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%, o al menos aproximadamente el 45% o al menos aproximadamente el 50% del ADNcf purificado se usa para amplificar los ácidos nucleicos diana. Preferentemente, la amplificación de las secuencias diana se puede realizar mediante cualquier método que utilice PCR o variaciones del método, incluidas, pero sin limitación, PCR asimétrica, amplificación dependiente de la helicasa, PCR de inicio en caliente, qPCR, PCR en fase sólida y PCR de contacto. En la etapa **440**, una porción, por ejemplo, al menos aproximadamente el 0,01%, del producto amplificado se combina con la muestra de ADNcf purificado sin amplificar, y la mezcla de ácidos nucleicos fetales y maternos amplificados y sin amplificar se secuencian en la etapa **450**. En un caso, la secuenciación se realiza utilizando una cualquiera de las tecnologías NGS. En la etapa **460**, se analizan los datos de los ejecuciones de secuenciación y se realiza la determinación simultánea de la fracción fetal y la presencia o la ausencia de aneuploidía tal como se ha descrito en la etapa **140** del caso representado en la figura 1.

En otro caso, la etapa **220** de enriquecimiento de la mezcla de ácidos nucleicos fetales y maternos en ácidos nucleicos diana polimórficos (figura 2) comprende combinar al menos una porción de una primera biblioteca de secuenciación de moléculas de ácidos nucleicos fetales y maternos no amplificados con al menos una porción de una segunda biblioteca de secuenciación de ácidos nucleicos diana polimórficos amplificados. Así, 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 en los ácidos nucleicos diana se realiza mediante métodos que comprenden amplificar específicamente las secuencias de ácido nucleico diana que comprenden el sitio polimórfico. En la etapa **510**, una muestra de ensayo, por ejemplo, una muestra de fluido biológico tal como una muestra de sangre, que comprende una mezcla de ácidos nucleicos fetales y maternos, se obtiene de una mujer embarazada, y la mezcla de ácidos nucleicos fetales y maternos se purifica a partir de la fracción de plasma en la etapa **520**. En la etapa **530**, una porción del ADNcf contenido en la muestra purificada se utiliza para amplificar los 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 las secuencias de ácido nucleico diana. Preferentemente, la amplificación de las secuencias diana se puede realizar mediante cualquier método que utilice PCR o variaciones del método, incluidas, pero sin limitación, PCR asimétrica, amplificación dependiente de la helicasa, PCR de inicio en caliente, qPCR, PCR en fase sólida y PCR de contacto. En la etapa **540**, los ácidos nucleicos diana amplificados que comprenden los sitios polimórficos, por ejemplo, SNP, se utilizan para preparar una biblioteca de secuenciación de ácidos nucleicos diana. De forma similar, la porción de ADNcf no amplificado purificado se usa para preparar una biblioteca de secuenciación primaria en la etapa **550**. En la etapa **560**, una porción de la biblioteca diana se combina con la biblioteca primaria generada a partir de la mezcla de ácidos nucleicos sin amplificar, y la mezcla de ácidos nucleicos fetales y maternos comprendida en las dos bibliotecas se secuencian en la etapa **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 la etapa **580**, se analizan los datos de las ejecuciones de secuenciación y se realiza la determinación simultánea de la fracción fetal y la presencia o la ausencia de aneuploidía tal como se ha descrito en la etapa **140** del caso representado en la figura 1.

#### **Determinación de aneuploidía a partir de bibliotecas enriquecidas de secuenciación**

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

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

La determinación de la fracción fetal en las etapas **240** (figura 2), **360** (figura 3), **480** (figura 4) y **580** (figura 5) se basa en el número total de etiquetas que se mapean al primer alelo y el número total de etiquetas que se mapean 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 NCBI36/hg18 del genoma de referencia humano, o el genoma de referencia comprende la secuencia NCBI36/hg18 del genoma de referencia humano y un genoma de secuencias diana artificiales, que incluye las secuencias polimórficas diana. En un caso, el genoma diana artificial abarca secuencias polimórficas que comprenden SNP 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 un caso, el genoma artificial incluye las secuencias diana polimórficas de las SEQ ID NO: 1-56. En otro caso, el genoma artificial incluye las secuencias diana polimórficas de las SEQ ID NO: 1-26 (véase el ejemplo, 7). En otro caso, el genoma diana artificial abarca secuencias polimórficas que comprenden STR seleccionados de entre CSF1PO, FGA, TH01, TPOX, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, PentaD, 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 otro caso 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 utilicen para determinar la fracción fetal. En consecuencia, un genoma de secuencias diana artificiales no se limita a las secuencias SNP o STR ejemplificadas en el presente documento.

El sitio polimórfico informativo, 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 < 10% del ADNcf materno. Así, la presencia de una contribución secundaria de un alelo a la mezcla de ácidos nucleicos fetales y maternos con respecto a la contribución principal del alelo materno se puede asignar al feto. Los alelos que se derivan del genoma materno se denominan en el presente documento alelos principales, y los alelos que se derivan del genoma fetal se denominan en el presente documento alelos secundarios. Los alelos que están representados por niveles similares de etiquetas de secuencia mapeadas representan alelos maternos. Los resultados de una amplificación multiplex ilustrativa 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 por el cambio de un solo nucleótido en un sitio polimórfico predeterminado, y los alelos fetales se distinguen por su contribución relativamente secundaria a la mezcla de ácidos nucleicos fetales y maternos en la muestra en comparación con la contribución principal a la mezcla por parte de los ácidos nucleicos maternos, es decir, las secuencias 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 de la madre y el alelo heredado del padre para calcular la fracción fetal. En consecuencia, la abundancia relativa de ADNcf fetal en la muestra materna se determina como un parámetro del número total de etiquetas de secuencia única 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 un caso, 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<sub>x</sub>) de la forma siguiente:

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

la fracción fetal para la muestra se calcula como el promedio de la fracción fetal de todos los alelos informativos.

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<sub>x</sub>) de la forma siguiente:

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

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

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

De forma similar, la fracción fetal se puede calcular a partir del número de etiquetas mapeadas a alelos de SNP en tándem tal como se realiza para los SNP individuales, pero teniendo en cuenta las etiquetas mapeadas a los dos alelos x e y de SNP en tándem presentes en cada una de las secuencias polimórficas de ácido nucleico diana amplificadas que se amplifican para enriquecer las muestras, es decir, % de fracción fetal alelo<sub>x+y</sub> = (( $\sum$ Etiquetas de secuencia fetal para alelo<sub>x+y</sub>) / ( $\sum$ Etiquetas de secuencia materna para alelo<sub>x+y</sub>))  $\times$  100

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<sub>x+y</sub>) de la forma siguiente:

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

para compensar la presencia de 2 conjuntos de alelos fetales en tándem, estando uno enmascarado por el fondo materno. Las secuencias de 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 de la madre y el haplotipo heredado del padre para calcular la fracción fetal.

La fracción fetal se puede determinar a partir de bibliotecas de secuenciación que comprenden secuencias diana polimórficas amplificadas que comprenden STR contando el número de etiquetas mapeadas a un alelo principal (materno) y uno secundario (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 heredado de la madre) y una secuencia de etiqueta secundaria correspondiente al alelo fetal no heredado de la madre. La fracción fetal se calcula como una relación del número de etiquetas mapeadas a los alelos fetales y maternos.

**Determinación de la fracción fetal por secuenciación masivamente paralela**

Además de utilizar el presente método para determinar simultáneamente la fracción fetal y la aneuploidía, la fracción fetal puede determinarse independientemente de la determinación de la aneuploidía tal como se describe en el presente documento, pero puede determinarse de forma independiente y/o junto con otros métodos utilizados para la determinación de la aneuploidía tales como los métodos descritos en las publicaciones de solicitud de patente de Estados Unidos N° US 2007/0202525A1; US2010/0112575A1, US 2009/0087847A1; US2009/0029377A1; US 2008/0220422A1; US2008/0138809A1, US2008/0153090A1 y la patente de Estados Unidos N° 7.645.576. El método para determinar la fracción fetal también se puede combinar con ensayos para determinar otras condiciones prenatales asociadas con la madre y/o el feto. Por ejemplo, el método se puede utilizar junto con análisis prenatales, por ejemplo, tal como se describe en las publicaciones de solicitud de patente de Estados Unidos N° US2010/0112590A1, US2009/0162842A1, US2007/0207466A1 y US2001/0051341A1.

La figura 6 proporciona un diagrama de flujo de un caso del método de la divulgación para determinar la fracción de ácidos nucleicos fetales en una muestra biológica materna mediante secuenciación masivamente paralela de ácidos nucleicos diana polimórficos amplificados por PCR independientemente de la determinación simultánea de aneuploidía. El método comprende la secuenciación de una biblioteca de secuenciación de ácidos nucleicos diana polimórficos de la forma siguiente. En la etapa **610** se obtiene de un sujeto una muestra de ensayo materna que comprende una mezcla de ácidos nucleicos fetales y maternos. La muestra es una muestra materna que se obtiene de un sujeto femenino embarazado, por ejemplo, una mujer embarazada. Otras muestras maternas pueden ser de mamíferos, por ejemplo, vaca, yegua, perra o gata. Si el sujeto es un ser humano, la muestra se puede tomar en el primer o segundo trimestre de embarazo. Los ejemplos de muestras biológicas maternas son los descritos anteriormente. En la etapa **620**, la mezcla de ácidos nucleicos fetales y maternos 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 fetales y maternos, por ejemplo, ADNcf, tal como se ha descrito para el caso **100**. En la etapa **630**, una porción de la mezcla purificada de ADNcf fetal y materno se utiliza 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 un solo nucleótido (SNP), SNP en tándem, deleciones o inserciones de múltiples bases a pequeña escala, denominadas IN-DELS (también denominadas polimorfismos de deleción-inserción o DIP), polimorfismos de múltiples nucleótidos (MNP), repeticiones en tándem cortas (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 ilustrativas y los métodos para amplificarlas son tal como se describen para los casos que se muestran en las figuras 2-5. En algunos casos, los sitios polimórficos están ubicados 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 pueden utilizarse en los métodos descritos en el presente documento.

En la etapa **640**, una porción o la totalidad de las secuencias polimórficas amplificadas se utilizan para preparar una biblioteca de secuenciación para la secuenciación en una forma paralela tal como se describe. En un caso, la biblioteca se prepara para la secuenciación por síntesis utilizando la química de secuenciación basada en terminadores reversibles de Illumina, tal como se describe en el ejemplo, 13. En la etapa **640**, la información de secuencia que se necesita para determinar la fracción fetal se obtiene utilizando un método NGS. En la etapa **650**, la fracción fetal se determina en base al número total de etiquetas que se mapean al primer alelo y el número total de etiquetas que se mapean al segundo alelo en un sitio polimórfico informativo, por ejemplo, un SNP, contenido en un genoma de referencia artificial, por ejemplo, un genoma de referencia SNP. Los genomas diana artificiales son tal como se describen en el presente documento. Se identifican los sitios polimórficos informativos y se calcula la fracción fetal tal como se describe.

La determinación de la fracción fetal según el presente documento puede utilizarse junto con ensayos de diagnóstico clínicamente relevantes como control positivo de la presencia de ADNcf con el fin de resaltar resultados falsos negativos o falsos positivos derivados de niveles bajos de ADNcf por debajo del límite de identificación. En un caso, la información de la fracción fetal se puede utilizar para establecer umbrales y estimar el tamaño mínimo de la muestra en la detección de aneuploidía. Dicho uso se describe en el ejemplo, 16, más adelante. 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 de plasma o suero materno, pueden utilizarse para enumerar secuencias en una muestra. Las secuencias pueden enumerarse utilizando cualquiera de las técnicas de secuenciación descritas anteriormente. El conocimiento de la fracción fetal se puede utilizar para establecer umbrales de "corte" para designar a los estados "aneuploidía", "normal" o "marginal/sin designación" (incierto). Después se pueden realizar cálculos para estimar el número mínimo de secuencias necesarias para lograr la sensibilidad adecuada (es decir, la probabilidad de identificar correctamente un estado de aneuploidía).

Los presentes métodos pueden aplicarse 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 aportada a una muestra por dos individuos, por ejemplo, los diferentes genomas son aportados por el feto y la madre que porta al feto, los métodos pueden utilizarse para determinar la fracción de un genoma en una mezcla derivada de dos células

diferentes de un individuo, por ejemplo, los genomas son aportados a la muestra por células cancerosas aneuploides y células euploides normales del mismo sujeto.

### Composiciones y kits

5 También se divulgan en el presente documento composiciones y kits o sistemas de reactivos útiles para poner en práctica los métodos descritos en el presente documento.

10 Las composiciones divulgadas en el presente documento se pueden incluir en kits para mezclas de secuenciación masivamente paralela de moléculas de ácidos nucleicos fetales y maternos, por ejemplo, ADNcf, presentes en una muestra materna, por ejemplo, una muestra de plasma. Los kits comprenden una composición que comprende al menos un conjunto de cebadores para amplificar al menos un ácido nucleico diana polimórfico en dichas moléculas de ácidos nucleicos fetales y maternos. 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, denominadas IN-DELS (también denominadas polimorfismos de deleción-inserción o DIP), polimorfismos de múltiples nucleótidos (MNP), repeticiones en tándem cortas (STR), polimorfismo de longitud de fragmentos 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 individuales o moléculas de ácido nucleico amplificadas clonalmente tal 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 colorantes reversibles, secuenciación en tiempo real, secuenciación por ligación de sonda de oligonucleótidos o secuenciación de una sola molécula.

25 En un caso, 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 de los 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 las SEQ ID NO: 57-112.

30 En otro caso, la composición comprende cebadores para amplificar ácidos nucleicos diana polimórficos que comprenden cada uno al menos un SNP en tándem. Los SNP en tándem ilustrativos 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 un caso, la composición incluye cebadores para amplificar los SNP en tándem ilustrativos divulgados en el presente documento, y la composición comprende los cebadores ilustrativos correspondientes de SEQ ID NO: 197-310.

50 En otro caso, la composición comprende cebadores para amplificar ácidos nucleicos diana polimórficos que comprenden cada uno al menos un STR. Los STR ilustrativos incluyen CSF1PO, FGA, TH01, TPOX, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, D2S1338, PentaD, PentaE, 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 un caso, la composición incluye cebadores para amplificar los STR en tándem ilustrativos divulgados en el presente documento, y la composición comprende los cebadores ilustrativos correspondientes de SEQ ID NO: 113-196.

60 Los kits pueden contener una combinación de reactivos que incluye los elementos necesarios para realizar un ensayo según los métodos divulgados en el presente documento. El sistema de reactivos se presenta en forma de paquete comercial, como una composición o mezcla cuando la compatibilidad de los reactivos lo permita, en una configuración de dispositivo de ensayo, o más típicamente como un kit de ensayo, es decir, una combinación envasada de uno o más recipientes, dispositivos o similares que contenga los reactivos necesarios, y que preferentemente incluya instrucciones escritas para la realización de los ensayos. El kit divulgado en el presente documento puede adaptarse a 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 tal como se describe y, cuando corresponda, reactivos para purificar ADNcf, que están dentro del alcance de la

65

divulgación. En un caso, 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, incluyen al menos un conjunto de oligonucleótidos específicos de alelo específicos para un SNP seleccionado y/o una región de repeticiones en tándem. Preferentemente, el kit incluye una pluralidad de conjuntos de cebadores para amplificar un panel de secuencias polimórficas. Un kit puede comprender otros reactivos y/o información para genotipar o cuantificar alelos en una muestra (por ejemplo, tampones, nucleótidos, instrucciones). Los kits también incluyen una pluralidad de recipientes de tampones y reactivos apropiados.

### Productos informáticos

La determinación de aneuploidía y/o la determinación de la fracción fetal se deriva informáticamente de la gran cantidad de información de secuenciación que se obtiene según los métodos descritos en el presente documento. En un caso, se divulga en el presente documento un medio legible informáticamente que tiene almacenadas instrucciones legibles informáticamente para determinar la presencia o la ausencia de aneuploidía a partir de la información obtenida de la secuenciación de ácidos nucleicos fetales y maternos en una muestra materna. En un caso, el medio legible informáticamente utiliza información de secuencia obtenida de una pluralidad de moléculas de ácidos nucleicos fetales y maternos en una muestra de plasma materno para identificar un número 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 de normalización, el medio legible informáticamente 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 la ausencia de aneuploidía fetal. Los ejemplos de cromosomas de interés incluyen, sin limitación, los cromosomas 21, 13, 18 y X.

En otro caso, se divulga en el presente documento un sistema de procesamiento que está adaptado o configurado para determinar la presencia o la ausencia de aneuploidía a partir de la información obtenida de la secuenciación de ácidos nucleicos fetales y maternos en una muestra materna. El sistema de procesamiento informático está adaptado o configurado para (a) usar información de secuencia obtenida a partir de una pluralidad de moléculas de ácidos nucleicos fetales y maternos en una muestra de plasma materno para identificar un número de etiquetas de secuencia mapeadas para un cromosoma de interés; (b) usar información de secuencia obtenida a partir de una pluralidad de moléculas de ácidos nucleicos fetales y maternos en una muestra de plasma materno para identificar un número de etiquetas de secuencia mapeadas para al menos un cromosoma de normalización; (c) usar el número de etiquetas de secuencia mapeadas identificadas para un cromosoma de interés en la etapa (a) y el número de etiquetas de secuencia mapeadas identificadas para el, al menos un, cromosoma de normalización en la etapa (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, e identificar así la presencia o la ausencia de aneuploidía fetal. Los ejemplos de cromosomas de interés incluyen, sin limitación, los cromosomas 21, 13, 18 y X.

En otro caso, se divulga en el presente documento un aparato adaptado o configurado para determinar la presencia o la ausencia de aneuploidía a partir 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 un porción de las moléculas de ácido nucleico en una muestra de plasma materno que comprende moléculas de ácidos nucleicos fetales y maternos, generando así información de secuencia; y (b) un sistema de procesamiento informático configurado para realizar las etapas siguientes: (i) usar 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 información de secuencia generada por el dispositivo de secuenciación para identificar un número de etiquetas de secuencia mapeadas para al menos un cromosoma de normalización; (iii) usar el número de etiquetas de secuencia mapeadas identificadas para un cromosoma de interés en la etapa (i) y el número de etiquetas de secuencia mapeadas identificadas para el, al menos un, cromosoma de normalización en la etapa (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, e identificar así la presencia o la 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

Se recogieron muestras de sangre periférica de mujeres embarazadas en su primer o segundo trimestre de embarazo y que se consideraron en riesgo de aneuploidía fetal. Se obtuvo el consentimiento informado de cada participante antes de la extracción de sangre. La sangre se recogió antes de la amniocentesis o del muestreo de vellosidades coriónicas. El análisis de cariotipos se realizó utilizando muestras de vellosidades coriónicas o de amniocentesis para confirmar el cariotipo fetal.

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

Para la extracción de plasma libre de células, la capa de plasma superior se transfirió a un tubo de centrifugadora de alta velocidad de 15 ml y se centrifugó a  $16.000 \times g$ ,  $4^{\circ}\text{C}$  durante 10 min utilizando una centrifugadora Beckman Coulter Avanti J-E y un rotor JA-14. Las dos etapas 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  $-80^{\circ}\text{C}$  y se descongeló solo una vez antes de la amplificación del plasma ADNcf o para la purificación de ADNcf.

El ADN libre de células purificado (ADNcf) se extrajo del plasma libre de células usando el kit QIAamp Blood DNA Mini (Qiagen) esencialmente según las instrucciones del fabricante. Se añadió un mililitro de tampón AL y  $100 \mu\text{l}$  de solución de proteasa a 1 ml de plasma. La mezcla se incubó durante 15 minutos a  $56^{\circ}\text{C}$ . Se añadió un mililitro de etanol al 100% a la digestión de plasma. La mezcla resultante se transfirió a minicolumnas QIAamp que se ensamblaron con VacValves y VacConnectors proporcionados en el ensamblaje de columna QIAvac 24 Plus (Qiagen) Se aplicó vacío a las muestras y el ADNcf retenido en los filtros de la columna se lavó al vacío con  $750 \mu\text{l}$  de tampón AW1, seguido de un segundo lavado con  $750 \mu\text{l}$  de tampón AW24. La columna se centrifugó a  $14.000 \text{ rpm}$  durante 5 minutos para eliminar cualquier tampón residual del filtro. El ADNcf se eluyó con tampón AE mediante centrifugación a  $14.000 \text{ rpm}$  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

Todas las bibliotecas de secuenciación, es decir, bibliotecas primarias y enriquecidas, se prepararon a partir de aproximadamente 2 ng de ADNcf purificado que se extrajo de plasma materno. La preparación de la biblioteca se realizó utilizando reactivos del conjunto 1 de reactivos de ADN para la preparación de muestras de ADN NEBNext™ (Nº de parte E6000L; New England Biolabs, Ipswich, MA), para Illumina® de la forma siguiente. Debido a que el ADN plasmático libre de células está fragmentado en la naturaleza, no se realizó ninguna 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  $40 \mu\text{l}$  se convirtieron en extremos romos fosforilados según el módulo de reparación de extremos NEBNext® mediante incubación en un tubo de microcentrifugadora de 1,5 ml del ADNcf con  $5 \mu\text{l}$  de tampón de fosforilación 10X,  $2 \mu\text{l}$  de mezcla de solución de desoxinucleótidos (10 mM cada dNTP),  $1 \mu\text{l}$  de una dilución 1:5 de ADN polimerasa I,  $1 \mu\text{l}$  de ADN polimerasa T4 y  $1 \mu\text{l}$  de polinucleótido quinasa T4 proporcionada en el conjunto 1 de reactivos de ADN para la preparación de muestras de ADN NEBNext™ durante 15 minutos a  $20^{\circ}\text{C}$ . Después, las enzimas se inactivaron con calor incubando la mezcla de reacción a  $75^{\circ}\text{C}$  durante 5 minutos. La mezcla se enfrió a  $4^{\circ}\text{C}$  y la adición de colas de dA al ADN de extremos romos se realizó usando  $10 \mu\text{l}$  de la mezcla maestra de adición de colas de dA que contenía el fragmento Klenow (3' a 5' exo minus) (conjunto 1 de reactivos de ADN para preparación de muestras de ADN NEBNext™) e incubando durante 15 minutos a  $37^{\circ}\text{C}$ . Posteriormente, el fragmento de Klenow se inactivó con calor incubando la mezcla de reacción a  $75^{\circ}\text{C}$  durante 5 minutos. Después de la inactivación del fragmento Klenow, se usó  $1 \mu\text{l}$  de una dilución 1:5 de mezcla Illumina Genomic Adapter Oligo Mix (Nº de parte 1000521; Illumina Inc., Hayward, CA) para ligar los adaptadores de Illumina (adaptadores Y sin índice) al ADN con cola de dA utilizando  $4 \mu\text{l}$  de la ADN ligasa T4 proporcionada en el conjunto 1 de reactivos de ADN para preparación de muestras de ADN NEBNext™, incubando la mezcla de reacción durante 15 minutos a  $25^{\circ}\text{C}$ . La mezcla se enfrió a  $4^{\circ}\text{C}$  y el ADNcf ligado al adaptador se purificó de adaptadores no ligados, dímeros de adaptadores y otros reactivos utilizando perlas magnéticas provistas en el sistema de purificación de PCR Agencourt AMPure XP (N.º de parte A63881; Beckman Coulter Genomics, Danvers, MA). Se realizaron dieciocho ciclos de PCR para enriquecer selectivamente el ADNcf ligado con adaptador ( $25 \mu\text{l}$ ) usando la mezcla maestra Phusion® High-Fidelity Master Mix ( $25 \mu\text{l}$ ; Finnzymes, Woburn, MA) y cebadores de PCR de Illumina ( $0,5 \mu\text{M}$  cada uno) complementarios a los adaptadores (Nº de parte 1000537 y 1000537). El ADN ligado al adaptador se sometió a PCR ( $98^{\circ}\text{C}$  durante 30 segundos; 18 ciclos de  $98^{\circ}\text{C}$  durante 10 segundos,  $65^{\circ}\text{C}$  durante 30 segundos y  $72^{\circ}\text{C}$  durante 30; extensión final a  $72^{\circ}\text{C}$  durante 5 minutos) y se mantuvo a  $4^{\circ}\text{C}$ ) usando cebadores de PCR genómicos de Illumina (Nº de parte 100537 y 1000538) y la mezcla maestra Phusion HF PCR Master Mix proporcionada en el conjunto 1 de reactivos de ADN para preparación de muestras de ADN NEBNext™, según las instrucciones del fabricante. El producto amplificado se purificó utilizando el sistema de purificación por PCR Agencourt AMPure XP (Agencourt Bioscience Corporation, Beverly, MA) según las instrucciones del fabricante disponibles en [www.beckmangenomics.com/products/AMPureXPProtocol\\_000387v001.pdf](http://www.beckmangenomics.com/products/AMPureXPProtocol_000387v001.pdf). El producto amplificado purificado se eluyó en  $40 \mu\text{l}$  de tampón EB de Qiagen, y la concentración y la distribución de tamaño de las bibliotecas amplificadas se analizó utilizando el kit Agilent DNA 1000 para el bioanalizador 2100 (Agilent technologies Inc., Santa Clara, CA).

#### b. Preparación de bibliotecas de secuenciación -- protocolo de duración completa

El protocolo de duración completa descrito es esencialmente el protocolo estándar proporcionado por Illumina, y solo difiere del protocolo de Illumina en la purificación de la biblioteca amplificada: el protocolo de Illumina indica que la biblioteca amplificada se purificará mediante electroforesis en gel, mientras que el protocolo descrito en el presente documento utiliza perlas magnéticas para la misma etapa de purificación. Se usaron aproximadamente 2 ng de ADNcf purificado que se había extraído de plasma materno para preparar una biblioteca de secuenciación primaria usando

el conjunto 1 de reactivos de ADN para la preparación de muestras de ADN NEBNext™ (Nº de parte E6000L; New England Biolabs, Ipswich, MA) para Illumina® esencialmente según las instrucciones del fabricante. Todas las etapas, excepto la purificación final de los productos ligados con adaptador, que se realizó con perlas magnéticas y reactivos de Agencourt en lugar de la columna de purificación, se realizaron según el protocolo que acompaña a los reactivos para la preparación de muestras NEBNext™ para una biblioteca de ADN genómico que se secuenciará utilizando Illumina® GAI. El protocolo NEBNext™ sigue esencialmente el proporcionado por Illumina, que está disponible en [grcf.jhmi.edu/hts/protocols/11257047\\_ChIP\\_SamplePrep.pdf](http://grcf.jhmi.edu/hts/protocols/11257047_ChIP_SamplePrep.pdf).

Los salientes de aproximadamente 2 ng de fragmentos de ADNcf purificados contenidos en 40 µl se convirtieron en extremos romos fosforilados según el módulo de reparación de extremos NEBNext® mediante incubación en 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 polinucleótido quinasa T4 proporcionados en el conjunto 1 de reactivos de ADN para la preparación de muestras de ADN NEBNext™ en un tubo de microcentrifugadora de 200 µl en un ciclador térmico durante 30 minutos a 20°C. La muestra se enfrió a 4°C y se purificó usando una columna QIAQuick proporcionada en el kit de purificación QIAQuick PCR (QIAGEN Inc., Valencia, CA) de la forma siguiente. La reacción de 50 µl se transfirió a un tubo de microcentrifugadora de 1,5 ml y se añadieron 250 µl de tampón PB de Qiagen. Los 300 µl resultantes se transfirieron a una columna QIAquick, que se centrifugó a 13.000 rpm durante 1 minuto en una microcentrifugadora. La columna se lavó con 750 µl de tampón PE de Qiagen 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 tampón EB de Qiagen mediante centrifugación. La adición de colas de dA a 34 µl del ADN de extremos romos se realizó usando 16 µl de la mezcla maestra de adición de colas de dA que contenía el fragmento Klenow (3' a 5' exo minus) (conjunto 1 de reactivos de ADN para la preparación de muestras de ADN NEBNext™) e incubando durante 30 minutos a 37°C según el módulo de adición de colas de dA NEBNext® del fabricante. La muestra se enfrió a 4°C y se purificó usando una columna proporcionada en el kit de purificación por PCR MinElute (QIAGEN Inc., Valencia, CA) de la forma siguiente. La reacción de 50 µl se transfirió a un tubo de microcentrifugadora de 1,5 ml y se añadieron 250 µl de tampón PB de Qiagen. Los 300 µl se transfirieron a una columna MinElute, que se centrifugó a 13.000 rpm durante 1 minuto en una microcentrifugadora. La columna se lavó con 750 µl de tampón PE de Qiagen 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 con 15 µl de tampón EB de Qiagen y se volvió a centrifugar. Diez microlitros del eluido de ADN se incubaron con 1 µl de una dilución 1:5 de mezcla Illumina Genomic Adapter Oligo Mix (Nº de parte 1000521), 15 µl de 2X de tampón de reacción de ligación rápida y 4 µl de ADN ligasa T4 rápida, durante 15 minutos a 25°C según el módulo de ligación rápida NEBNext®. La muestra se enfrió a 4°C y se purificó usando una columna MinElute de la forma siguiente. Se añadieron ciento cincuenta microlitros de tampón PE de Qiagen a los 30 µl de reacción, y el volumen completo se transfirió a una columna MinElute, que se centrifugó a 13 000 rpm durante 1 minuto en una microcentrifugadora. La columna se lavó con 750 µl de tampón PE de Qiagen 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 con 28 µl de tampón EB de Qiagen y se volvió a centrifugar. Veintitres microlitros del eluido 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 se mantuvo a 4°C) con cebadores de PCR genómicos de Illumina (Nº de parte 100537 y 1000538) y la mezcla maestra Phusion HF PCR Master Mix proporcionada en el conjunto 1 de reactivos de ADN para preparación de muestras de ADN NEBNext™, según las instrucciones del fabricante. El producto amplificado se purificó utilizando el sistema de purificación por PCR Agencourt AMPure XP (Agencourt Bioscience Corporation, Beverly, MA) según las instrucciones del fabricante disponibles en [www.beckmangenomics.com/products/AMPureXPProtocol\\_000387v001.pdf](http://www.beckmangenomics.com/products/AMPureXPProtocol_000387v001.pdf). El sistema de purificación por PCR AMPure XP de Agencourt elimina los dNTP, los cebadores, los dímeros de cebadores, las sales y otros contaminantes no incorporados, y recupera los amplicones de más de 100 pb. El producto amplificado purificado se eluyó a partir de perlas Agencourt en 40 µl de tampón EB de Qiagen, y la distribución de tamaño de las bibliotecas amplificadas se analizó utilizando el kit Agilent DNA 1000 para el bioanalizador 2100 (Agilent technologies Inc., Santa Clara, CA).

### c. Análisis de bibliotecas de secuenciación preparadas según los protocolos abreviado (a) y de longitud completa (b)

Los electroferogramas generados por el bioanalizador se muestran en la figura 7. La figura 7 (A) muestra el electroferograma del ADN de la biblioteca preparado a partir de ADNcf purificado a partir de la 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 a partir 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 1500, 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 de la figura 7 (A) muestra un pico secundario de fragmentos de 187 pb y un pico principal de fragmentos de 263 pb, mientras que el electroferograma de 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 de la figura 7 (A), una concentración de 7,34 ng/µl para el ADN del pico de 263 pb de la figura 7 (A), y una concentración de 14,72 ng/µl para el ADN del pico de 265 pb de la figura 7 (B). Se sabe que los adaptadores de Illumina que se ligaron al ADNcf tienen 92 pb, que cuando se restan de los 265 pb, indican que el tamaño máximo del

ADNcf es de 173 pb. Es posible que el pico secundario de 187 pb represente fragmentos de dos cebadores que se ligaron extremo a extremo. Los fragmentos de dos cebadores lineales se eliminan del producto final de la biblioteca cuando se utiliza 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 a adaptador purificado es el doble que la del ADNcf ligado a adaptador producido utilizando el protocolo de longitud completa. Se ha observado que la concentración de los fragmentos de ADNcf ligados al adaptador es siempre mayor que la obtenida utilizando el protocolo de longitud completa (datos no mostrados).

Así, una ventaja de preparar la biblioteca de secuenciación utilizando el protocolo abreviado es que la biblioteca obtenida comprende de forma coherente un solo pico principal en el intervalo de 262 a 267 pb, mientras que la calidad de la biblioteca preparada usando el protocolo de longitud completa varía según se refleja por el número y la movilidad de picos distintos de los que representan el ADNcf. Los productos que no son ADNcf ocuparían espacio en la celda de flujo y disminuirían la calidad de la amplificación de clústeres y la obtención de imágenes posterior de las reacciones de secuenciación, lo que es la base de la asignación general del estado de aneuploidía. Se ha demostrado que el protocolo abreviado no afecta a la secuenciación de la biblioteca (véase la figura 8).

Otra ventaja de preparar la biblioteca de secuenciación con el protocolo abreviado es que las tres etapas enzimáticas de formación de extremos romos, adición de colas de d-A y ligación de adaptador tardan menos de una hora en completarse para respaldar la validación e implementación de un servicio de diagnóstico de aneuploides rápida.

Otra ventaja es que las tres etapas enzimáticas de formación de extremos romos, adición de colas de d-A y ligación de adaptador se realizan en el mismo tubo de reacción, lo que evita múltiples transferencias de muestras que podrían conducir a la pérdida de material y, lo que es más importante, a una posible mezcla de muestras y contaminación de la muestra.

### Ejemplo 3

#### Secuenciación masivamente paralela y determinación de aneuploidía

Se obtuvieron muestras de sangre periférica de sujetos embarazados y se purificó ADNcf de la fracción de plasma tal como se describe en el ejemplo, 1. Todas las bibliotecas de secuenciación se prepararon utilizando el protocolo abreviado de preparación de bibliotecas descrito en el ejemplo, 2. El ADN amplificado se secuenció utilizando el analizador de genoma II de Illumina 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 dianas más particulares. En el presente caso se obtuvo un gran número de lecturas de 36 pb, cubriendo aproximadamente el 10% del genoma. La secuenciación del ADN de la biblioteca se realizó con el analizador de genoma II (Illumina Inc., San Diego, CA, Estados Unidos) según los protocolos estándar del fabricante. Se pueden encontrar copias del protocolo para la secuenciación del genoma completo utilizando la tecnología Illumina/Solexa en BioTechniques.RTM. Protocol Guide 2007 Publicada en diciembre de 2006: p 29, y en Internet en [biotechniques.com/default.asp?pagina=protocolo&subsection=article\\_display&id=112378](http://biotechniques.com/default.asp?pagina=protocolo&subsection=article_display&id=112378). La biblioteca de ADN se diluyó a 1 nM y se desnaturalizó. El ADN de la biblioteca (5 pM) se sometió a la amplificación de clústeres según el procedimiento descrito en la Cluster Station User Guide and Cluster Station Operations Guide de Illumina, disponible en Internet en [illumina.com/systems/genome\\_analyzer/cluster\\_station.ilmn](http://illumina.com/systems/genome_analyzer/cluster_station.ilmn). Una vez completada la secuenciación de la muestra, el programa informático "Sequencer Control Software" de Illumina transfirió los archivos de imagen y de designación de bases a un servidor Unix que ejecutaba la versión 1.51 del programa informático "Genome Analyzer Pipeline" de Illumina. Se ejecutó el programa "Gerald" de Illumina para alinear secuencias con el genoma humano de referencia que se deriva del genoma hg18 proporcionado por el National Center for Biotechnology Information (NCBI36/hg18, disponible en Internet en <http://genome.ucsc.edu/cgi-bin/hgGateway?org=Human&db=hg18&hgsid=166260105>). Los datos de secuencia generados a partir del procedimiento anterior que se alinearon de forma única con el genoma se leyeron desde la salida de Gerald (archivos export.txt) mediante un programa (c2c.pl) que se ejecutaba en un ordenador que ejecutaba el sistema operativo Linux. Se permitieron alineamientos de secuencias con desajustes de bases y se incluyeron en los recuentos de alineamientos solo si se alineaban únicamente con el genoma. Se excluyeron los alineamientos de secuencias con coordenadas de inicio y final idénticas (duplicados).

Entre aproximadamente 5 y 15 millones de etiquetas de 36 pb con 2 o menos desajustes se mapearon de forma única al genoma humano. Todas las etiquetas mapeadas se contaron y se incluyeron en el cálculo de las dosis de cromosomas tanto en las muestras de ensayo como en las de calificación. Las regiones que se extienden desde la base 0 hasta la base 2 x 10<sup>6</sup>, desde la base 10 x 10<sup>6</sup> hasta la base 13 x 10<sup>6</sup> y desde la base 23 x 10<sup>6</sup> hasta el final del cromosoma Y, se excluyeron específicamente del análisis porque las etiquetas derivadas de fetos tanto masculinos como femeninos se mapean a estas regiones del cromosoma Y.

Se observó que hubo cierta variación en el número total de etiquetas de secuencia mapeadas a cromosomas individuales entre muestras secuenciadas en la misma ejecución (variación intercromosómica), pero se observó una variación sustancialmente mayor entre diferentes ejecuciones de secuenciación (variación de ejecución intersecuenciación).

**Ejemplo 4**

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

Para examinar el alcance de la variación inter cromosómica e intersecuenciación en el número de etiquetas de secuencia mapeadas para todos los cromosomas, se extrajo y se secuenció ADNcf de plasma obtenido de sangre periférica de 48 sujetos embarazados voluntarios tal como se describe en el ejemplo, 1, y se analizó de la forma siguiente.

Se determinó el número total de etiquetas de secuencia que se mapearon a cada cromosoma (densidad de etiquetas de secuencia). Alternativamente, el número de etiquetas de secuencia mapeadas puede normalizarse a la longitud del cromosoma para generar una relación de densidad de etiquetas de secuencia. La normalización de la longitud del cromosoma no es una etapa necesaria y se puede realizar únicamente para reducir el número de dígitos en un número para simplificarlo para la interpretación humana. Las longitudes de los cromosomas que se pueden usar para normalizar los recuentos de etiquetas de secuencia pueden ser las longitudes proporcionadas en Internet en [genoma.ucsc.edu/goldenPath/stats.html#hg18](http://genoma.ucsc.edu/goldenPath/stats.html#hg18).

La densidad de etiquetas de secuencia resultante para cada cromosoma se relacionó con la densidad de etiquetas de secuencia de cada uno de los cromosomas restantes para derivar una dosis de cromosoma calificado, que se calculó como la relación de la densidad de etiquetas de secuencia para el cromosoma de interés, por ejemplo, el cromosoma 21 y la densidad de etiquetas 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 calificado 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 muestran en las figuras 9-13. Las figuras 9-13 también representan las dosis de cromosomas para las muestras de ensayo. Las dosis de cromosomas para cada uno de los cromosomas de interés en las muestras calificadas proporcionan una medida de la variación en el número total de etiquetas de secuencia mapeadas para cada cromosoma de interés con respecto a la de cada uno de los cromosomas restantes. Así, las dosis de cromosomas calificados pueden identificar el cromosoma o un grupo de cromosomas, es decir, el cromosoma de normalización, que tiene una variación entre las muestras que es 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 de cromosomas promedio calculadas determinadas en una población de muestras calificadas para los cromosomas 13, 18 y 21, y los cromosomas X e Y.

En algunos casos, el mejor cromosoma de normalización puede no tener la menor variación, pero puede tener una distribución de dosis calificadas que distingue mejor una muestra o muestras de ensayo de las muestras calificadas, es decir, el mejor cromosoma de normalización puede no tener la variación más baja, pero puede tener la mayor diferenciabilidad. Así, la diferenciabilidad cuenta para la variación en la dosis de cromosoma y la distribución de las dosis en las muestras calificadas.

Las tablas 2 y 3 proporcionan el coeficiente de variación como medida de variabilidad y los valores de la prueba t de Student como medida de diferenciabilidad para los cromosomas 18, 21, X e Y, en los que cuanto menor es el valor de la prueba T, mayor es la diferenciabilidad. La diferenciabilidad para el cromosoma 13 se determinó como el cociente de la diferencia entre la dosis de cromosoma media en las muestras calificadas y la dosis para el cromosoma 13 en la única muestra de ensayo T13, y la desviación estándar de la media de la dosis calificada.

Las dosis calificadas de cromosomas también sirven como base para determinar los valores umbrales al identificar aneuploidías en las muestras de ensayo, tal como se describe a continuación.

**Tabla 1**

<b>Dosis de cromosoma calificado para los cromosomas 13, 18, 21, X e Y (n=1; muestra #11342, 46 XY)</b>					
<b>Cromosoma</b>	<b>cr 21</b>	<b>cr 18</b>	<b>cr 13</b>	<b>cr X</b>	<b>cr Y</b>
cr1	0,149901	0,306798	0,341832	0,490969	0,003958
cr2	0,15413	0,315452	0,351475	0,504819	0,004069
cr3	0,193331	0,395685	0,44087	0,633214	0,005104
cr4	0,233056	0,476988	0,531457	0,763324	0,006153
cr5	0,219209	0,448649	0,499882	0,717973	0,005787

ES 2 909 841 T3

<b>Dosis de cromosoma calificado para los cromosomas 13, 18, 21, X e Y (n=1; muestra #11342, 46 XY)</b>					
<b>Cromosoma</b>	<b>cr 21</b>	<b>cr 18</b>	<b>cr 13</b>	<b>cr X</b>	<b>cr Y</b>
cr6	0,228548	0,467763	0,521179	0,748561	0,006034
cr7	0,245124	0,501688	0,558978	0,802851	0,006472
cr8	0,256279	0,524519	0,584416	0,839388	0,006766
cr9	0,309871	0,634203	0,706625	1,014915	0,008181
cr10	0,25122	0,514164	0,572879	0,822817	0,006633
cr11	0,257168	0,526338	0,586443	0,8423	0,00679
cr12	0,275192	0,563227	0,627544	0,901332	0,007265
cr13	0,438522	0,897509	1	1,436285	0,011578
cr14	0,405957	0,830858	0,925738	1,329624	0,010718
cr15	0,406855	0,832697	0,927786	1,332566	0,010742
cr16	0,376148	0,769849	0,857762	1,231991	0,009931
cr17	0,383027	0,783928	0,873448	1,254521	0,010112
cr18	0,488599	1	1,114194	1,600301	0,0129
cr19	0,535867	1,096742	1,221984	1,755118	0,014148
cr20	0,467308	0,956424	1,065642	1,530566	0,012338
cr21	1	2,046668	2,280386	3,275285	0,026401
cr22	0,756263	1,547819	1,724572	2,476977	0,019966
crX	0,305317	0,624882	0,696241	1	0,008061
cr Y	37,87675	77,52114	86,37362	124,0572	1

**Tabla 2**

<b>Dosis de cromosoma calificado, varianza y diferenciabilidad para los cromosomas 21, 18 y 13</b>								
	<b>21 (n=35)</b>				<b>18 (n=40)</b>			
	Promedio	Desv Est	CV	Prueba t	Promedio	Desv Est	CV	Prueba t
cr1	0,15335	0,001997	1,30	3,18E-10	0,31941	0,008384	2,62	0,001675
cr2	0,15267	0,001966	1,29	9,87E-07	0,31807	0,001756	0,55	4,39E-05
cr3	0,18936	0,004233	2,24	1,04E-05	0,39475	0,002406	0,61	3,39E-05
cr4	0,21998	0,010668	4,85	0,000501	0,45873	0,014292	3,12	0,001349
cr5	0,21383	0,005058	2,37	1,43E-05	0,44582	0,003288	0,74	3,09E-05
cr6	0,22435	0,005258	2,34	1,48E-05	0,46761	0,003481	0,74	2,32E-05
cr7	0,24348	0,002298	0,94	2,05E-07	0,50765	0,004669	0,92	9,07E-05
cr8	0,25269	0,003497	1,38	1,52E-06	0,52677	0,002046	0,39	4,89E-05
cr9	0,31276	0,003095	0,99	3,83E-09	0,65165	0,013851	2,13	0,000559
cr10	0,25618	0,003112	1,21	2,28E-10	0,53354	0,013431	2,52	0,002137
cr11	0,26075	0,00247	0,95	1,08E-09	0,54324	0,012859	2,37	0,000998
cr12	0,27563	0,002316	0,84	2,04E-07	0,57445	0,006495	1,13	0,000125
cr13	0,41828	0,016782	4,01	0,000123	0,87245	0,020942	2,40	0,000164
cr14	0,40671	0,002994	0,74	7,33E-08	0,84731	0,010864	1,28	0,000149
cr15	0,41861	0,007686	1,84	1,85E-10	0,87164	0,027373	3,14	0,003862

ES 2 909 841 T3

Dosis de cromosoma calificado, varianza y diferenciabilidad para los cromosomas 21, 18 y 13								
	21 (n=35)				18 (n=40)			
	Promedio	Desv Est	CV	Prueba t	Promedio	Desv Est	CV	Prueba t
cr16	0,39977	0,018882	4,72	7,33E-06	0,83313	0,050781	6,10	0,075458
cr17	0,41394	0,02313	5,59	0,000248	0,86165	0,060048	6,97	0,088579
cr18	0,47236	0,016627	3,52	1,3E-07				
cr19	0,59435	0,05064	8,52	0,01494	1,23932	0,12315	9,94	0,231139
cr20	0,49464	0,021839	4,42	2,16E-06	1,03023	0,058995	5,73	0,061101
cr21					2,03419	0,08841	4,35	2,81E-05
cr22	0,84824	0,070613	8,32	0,02209	1,76258	0,169864	9,64	0,181808
crX	0,27846	0,015546	5,58	0,000213	0,58691	0,026637	4,54	0,064883

Tabla 3

Dosis de cromosoma calificado, varianza y diferenciabilidad para los cromosomas 13, X e Y								
	13 (n=47)				X (n=19)			
	Promedio	Desv Est	CV	Dif	Promedio	Desv Est	CV	Prueba t
cr1	0,36536	0,01775	4,86	1,904	0,56717	0,025988	4,58	0,001013
cr2	0,36400	0,009817	2,70	2,704	0,56753	0,014871	2,62	9,6E-08
cr3	0,45168	0,007809	1,73	3,592	0,70524	0,011932	1,69	6,13E-11
cr4	0,52541	0,005264	1,00	3,083	0,82491	0,010537	1,28	1,75E-15
cr5	0,51010	0,007922	1,55	3,944	0,79690	0,012227	1,53	1,29E-11
cr6	0,53516	0,008575	1,60	3,758	0,83594	0,013719	1,64	2,79E-11
cr7	0,58081	0,017692	3,05	2,445	0,90507	0,026437	2,92	7,41E-07
cr8	0,60261	0,015434	2,56	2,917	0,93990	0,022506	2,39	2,11E-08
cr9	0,74559	0,032065	4,30	2,102	1,15822	0,047092	4,07	0,000228
cr10	0,61018	0,029139	4,78	2,060	0,94713	0,042866	4,53	0,000964
cr11	0,62133	0,028323	4,56	2,081	0,96544	0,041782	4,33	0,000419
cr12	0,65712	0,021853	3,33	2,380	1,02296	0,032276	3,16	3,95E-06
cr13					1,56771	0,014258	0,91	2,47E-15
cr14	0,96966	0,034017	3,51	2,233	1,50951	0,05009	3,32	8,24E-06
cr15	0,99673	0,053512	5,37	1,888	1,54618	0,077547	5,02	0,002925
cr16	0,95169	0,080007	8,41	1,613	1,46673	0,117073	7,98	0,114232
cr17	0,98547	0,091918	9,33	1,484	1,51571	0,132775	8,76	0,188271
cr18	1,13124	0,040032	3,54	2,312	1,74146	0,072447	4,16	0,001674
cr19	1,41624	0,174476	12,32	1,306	2,16586	0,252888	11,68	0,460752
cr20	1,17705	0,094807	8,05	1,695	1,81576	0,137494	7,57	0,08801
cr21	2,33660	0,131317	5,62	1,927	3,63243	0,235392	6,48	0,00675
cr22	2,01678	0,243883	12,09	1,364	3,08943	0,34981	11,32	0,409449
crX	0,66679	0,028788	4,32	1,114				
cr2-6	0,46751	0,006762	1,45	4,066				
cr3-6	0,50332	0,005161	1,03	5,260				

Dosis de cromosoma calificado, varianza y diferenciabilidad para los cromosomas 13, X e Y								
	13 (n=47)				X (n=19)			
	Promedio	Desv Est	CV	Dif	Promedio	Desv Est	CV	Prueba t
cr_tot					1,13209	0,038485	3,40	2,7E-05
	Y (n=26)							
	Promedio	Desv Est	CV	Prueba t				
Cr 1-22, X	0,00734	0,002611	30,81	1,8E-12				

En el ejemplo, 3 se describen ejemplos de diagnósticos de T21, T13, T18 y un caso de síndrome de Turner obtenidos utilizando los cromosomas de normalización, las dosis de cromosomas y la diferenciabilidad para cada uno de los cromosomas de interés.

5

**Ejemplo 5**

**Diagnóstico de aneuploidía fetal utilizando cromosomas de normalización**

10 Para aplicar el uso de dosis de cromosomas para evaluar la aneuploidía en una muestra de ensayo biológica, se obtuvieron muestras de análisis de sangre materna de voluntarias embarazadas y se preparó ADNcf, y se secuenció y se analizó una biblioteca de secuenciación preparada según el protocolo abreviado descrito en el ejemplo, 2.

**Trisomía 21**

15

La tabla 4 proporciona la dosis calculada para el cromosoma 21 en una muestra de ensayo ilustrativa (#11403). El umbral calculado para el diagnóstico positivo de aneuploidía T21 se estableció en > 2 desviaciones estándar de la media de las muestras calificadas (normales). Se dio un diagnóstico para T21 basado en que la dosis de cromosoma en la muestra de ensayo era superior al umbral establecido. Los cromosomas 14 y 15 se usaron como cromosomas de normalización en cálculos separados para mostrar que se puede utilizar para identificar la aneuploidía o bien un cromosoma que tiene la variabilidad más baja, por ejemplo, el cromosoma 14, o bien un cromosoma que tiene la mayor diferenciabilidad, por ejemplo, el cromosoma 15. Se identificaron trece muestras T21 usando las dosis de cromosomas calculadas, y se confirmó que las muestras con aneuploidía eran T21 por cariotipo.

20

25

**Tabla 4**

Dosis de cromosoma para una aneuploidía T21 (muestra #11403, 47 XY +21)			
Cromosoma	Densidad de etiquetas de secuencia	Dosis de cromosoma para el cr 21	Umbral
Cr21	333,660	0,419672	0,412696
Cr14	795,050		
Cr21	333,660	0,441038	0,433978
Cr15	756,533		

**Trisomía 18**

30

La tabla 5 proporciona la dosis calculada para el cromosoma 18 en una muestra de ensayo (#11390). El umbral calculado para el diagnóstico positivo de aneuploidía T18 se estableció en 2 desviaciones estándar de la media de las muestras calificadas (normales). Se dio un diagnóstico para T18 basado en que la dosis de cromosoma en la muestra de ensayo era superior al umbral establecido. El cromosoma 8 se usó como el cromosoma de normalización. En este caso, el cromosoma 8 tenía la menor variabilidad y la mayor diferenciabilidad. Se identificaron ocho muestras T18 usando dosis de cromosomas y se confirmó que eran T18 por cariotipo.

35

Estos datos muestran que un cromosoma de normalización puede tener tanto la menor variabilidad como la mayor diferenciabilidad.

40

Tabla 5

Dosis de cromosoma para una aneuploidía T18 (muestra #11390, 47 XY +18)			
Cromosoma	Densidad de etiquetas de secuencia	Dosis de cromosoma para el cr 18	Umbral
Cr18	602,506	0,585069	0,530867
Cr8	1,029,803		

**Trisomía 13**

5 La tabla 6 proporciona la dosis calculada para el cromosoma 13 en una muestra de ensayo (#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 dio un diagnóstico para T13 basado en que la dosis de cromosoma en la muestra de ensayo era superior al umbral establecido. La dosis de cromosoma para el cromosoma 13 se calculó utilizando o bien el cromosoma 5 o bien el grupo de cromosomas 3, 4, 5 y 6 como cromosoma de normalización. Se identificó una muestra T13.

Tabla 6

Dosis de cromosoma para una aneuploidía T13 (muestra #51236, 47 XY +13)			
Cromosoma	Densidad de etiquetas de secuencia	Dosis de cromosoma para el cr 13	Umbral
Cr13	692,242	0,541343	0,52594
Cr5	1,278,749		
Cr13	692,242	0,530472	0,513647
Cr3-6 [promedio]	1,304,954		

15 La densidad de etiquetas de secuencia para los cromosomas 3-6 es el recuento promedio de etiquetas para los cromosomas 3-6.

20 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.

25 Así, se puede utilizar un grupo de cromosomas como cromosoma de normalización para determinar las dosis de cromosomas e identificar las aneuploidías.

**Síndrome de Turner (monosomía X)**

30 La tabla 7 proporciona la dosis calculada para el cromosoma X e Y en una muestra de ensayo (#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 Turner (XO) (muestra #51238, 45 X)			
Cromosoma	Densidad de etiquetas de secuencia	Dosis de cromosoma para el cr X y el cr Y	Umbral
CrX	873,631	0,786642	0,803832
Cr4	1,110,582		
CrY	1,321	0,001542101	0,00211208
Cr_total	856,623,6		
(1-22, X) (Promedio)			

Una muestra que tenía una dosis de cromosoma X menor que la del umbral establecido se identificó como que tenía menos de un cromosoma X. Se determinó que la misma muestra tenía una dosis de cromosoma Y inferior al umbral establecido, lo que indica que la muestra no tenía un cromosoma Y. Así, la combinación de dosis de cromosomas para X e Y se utilizó para identificar las muestras con síndrome de Turner (monosomía X).

Así, el método descrito en el presente documento permite la determinación de la CNV de los cromosomas. En particular, el método permite la determinación de aneuploidías cromosómicas con sobrerrepresentación y subrepresentación mediante la secuenciación masivamente paralela de ADNcf de plasma materno y la identificación de cromosomas de normalización para el análisis estadístico de los datos de secuenciación. La sensibilidad y la fiabilidad del método permiten pruebas precisas de aneuploidía en el primer y segundo trimestre.

### Ejemplo 6

#### Determinación de aneuploidía parcial

El uso de dosis de secuencia se aplicó para evaluar la aneuploidía parcial en una muestra de ensayo biológica de ADNcf que se preparó a partir de plasma sanguíneo y se secuenció tal como se describe en el ejemplo, 1. Se confirmó mediante cariotipado que la muestra procedía de un sujeto con una delección parcial del cromosoma 11.

El análisis de los datos de secuenciación para la aneuploidía parcial (delección parcial del cromosoma 11, es decir, q21-q23) se realizó tal como se describe para las aneuploidías cromosómicas en los ejemplos anteriores. El mapeo de las etiquetas de secuencia en el cromosoma 11 en una muestra de ensayo reveló una pérdida significativa de recuentos de etiquetas entre los pares de bases 81000082-103000103 en el brazo q del cromosoma con respecto a los recuentos de etiquetas obtenidos para la secuencia correspondiente en el cromosoma 11 en las muestras calificadas (datos no mostrados). Se utilizaron etiquetas de secuencia mapeadas a la secuencia de interés en el cromosoma 11 (81000082-103000103 pb) en cada una de las muestras calificadas, y etiquetas de secuencia mapeadas a los 20 segmentos de megabase en el genoma completo en las muestras calificadas, es decir, densidades de etiquetas de secuencias calificadas, para determinar las dosis de secuencias calificadas como relaciones de densidades de etiquetas en todas las muestras calificadas. Se calcularon la dosis de secuencia promedio, la desviación estándar y el coeficiente de variación para todos los 20 segmentos de megabase en el genoma completo, y la secuencia de megabase de 20 que tenía la menor variabilidad fue la secuencia de normalización identificada en el cromosoma 5 (13000014-33000033 pb) (véase la tabla 8), que se utilizó para calcular la dosis para la secuencia de interés en la muestra de ensayo (véase la tabla 9). La tabla 8 proporciona la dosis de secuencia para la secuencia de interés en el cromosoma 11 (81000082-103000103 pb) en la muestra de ensayo que se calculó como la relación de etiquetas de secuencia mapeadas a la secuencia de interés y las etiquetas de secuencia mapeadas 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 ensayo (◇). La línea continua muestra la media, y la línea discontinua muestra el umbral calculado para el diagnóstico positivo de aneuploidía parcial que se estableció en 5 desviaciones estándar de la media. El diagnóstico de aneuploidía parcial se basó en que la dosis de secuencia en la muestra de ensayo era inferior al umbral establecido. Se verificó mediante cariotipado que la muestra de ensayo tenía la delección q21-q23 en el cromosoma 11.

Por lo tanto, además de identificar aneuploidías cromosómicas, el método de la divulgación puede usarse para identificar aneuploidías parciales.

Tabla 8

<b>Secuencia de normalización calificada, dosis y varianza para la secuencia Cr11: 81000082-103000103 (muestras calificadas n=7)</b>			
	<b>Cr11: 81000082-103000103</b>		
	<b>Promedio</b>	<b>Desv Est</b>	<b>CV</b>
Cr5 13000014-33000033	1,164702	0,004914	0,42

Tabla 9

<b>Dosis de secuencia para la secuencia de interés (81000082-103000103) en el cromosoma 11 (muestra de ensayo 11206)</b>				
Segmento de cromosoma	de	Densidad de etiquetas de secuencia	Dosis de segmento de cromosoma para Cr 11 (q21-q23)	Umbral
Cr11: 81000082-103000103		27,052	1,0434313	1,1401347
Cr5: 13000014-33000033		25,926		

**Ejemplo 7**

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

10 Se seleccionó un conjunto de 28 SNP autosómicos de una lista de 92 SNP (Pakstis et al., Hum Genet 127:315-324 [2010]), y de secuencias SNP disponibles en Applied Biosystems en la dirección de Internet [applybiosystems.com](http://applybiosystems.com), y se validaron para su uso en la amplificación por PCR multiplexada y para secuenciación masivamente paralela para determinar la fracción fetal con o sin la determinación simultánea de la presencia o la ausencia de aneuploidía. Se diseñaron cebadores para hibridarlos con una secuencia cercana al sitio SNP en el ADNcf para garantizar que se incluyera 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 someterlos a amplificación en puente durante la formación de clústeres. Así, se diseñaron cebadores para generar amplicones de al menos 110 pb, que cuando se combinaron con los adaptadores universales (Illumina Inc., San Diego, CA) utilizados para la amplificación de clústeres, dieron como resultado moléculas de ADNcf de al menos 200 pb. Se identificaron secuencias de cebadores y se sintetizaron conjuntos de cebadores, es decir, cebadores directo e inverso, por Integrated DNA Technologies (San Diego, CA), y se almacenaron como una solución 1  $\mu$ M para su uso para amplificar secuencias diana polimórficas tal 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 ADNcf diana y las secuencias de los amplicones que comprenden los posibles alelos de SNP que se generarían utilizando los cebadores. Los SNP proporcionados en la tabla 10 se utilizaron para la amplificación simultánea de 13 secuencias diana en un ensayo multiplexado para determinar simultáneamente la fracción fetal y la presencia o la ausencia de una aneuploidía en muestras de ADNcf derivadas de mujeres embarazadas. El panel proporcionado en la tabla 10 es un panel SNP ilustrativo. Se pueden emplear menos o más SNP para enriquecer el ADN fetal y materno en ácidos nucleicos diana polimórficos. Los SNP adicionales que se pueden usar incluyen los SNP que se proporcionan en la tabla 11. Los SNP de la tabla 11 se validaron en amplificaciones de PCR multiplex y se secuenciaron con el analizador Genomell A como se ha descrito anteriormente.

30 Los alelos SNP de las tablas 10 y 11 se muestran en negrita y subrayados.

Tabla 10

Panel de SNP para la determinación de la fracción fetal					
SNP ID	Cr	Amplificón: Alelo 1	Amplificón: Alelo 2	Secuencia de cebador directo, nombre y SEQ ID NO:	Secuencia de cebador inverso, nombre y SEQ ID NO:
rs560681	1	CACATGCACAGCCA GCAACCTGTACGC AGGAGTCCACCA GTTTCTTCTGAGAA CATCTGTTACAGGTTT CTCCATCTCTATT TACTCAGGTCACAG GACCTTGGGG (SEQ ID NO:1)	CACATGCACAGCCA GCAACCTGTACGC AGGAGTCCACCA GTTTCTTCTGAGAA CATCTGTTACAGGTTT CTCCATCTCTATT TACTCAGGTCACAG GACCTTGGGG (SEQ ID NO:2)	CACATGCA CAGCCAGC AAGCC (rs560681_C 1_1_F; SEQ ID NO:57)	CCCCAAGGTC CTGTGACCTGA GT (rs560681_C1_1_ R; SEQ ID NO:58)
rs1109037	2	TGAGGAAGTGAGGC TCAGAGGTAAGAA ACTTTGTCACAGGC TGGTGGTGAGGGTG GAGATTTTACACTCC CTGCCCTCCACACCA GTTTCTCCGGAGTGG AAAGACTTTCATCTC GCACTGGCA (SEQ ID NO:3)	TGAGGAAG TGAGGCTC AGAGGCTC AGAGGCTC GAGATTTTACACTCC 2_1_F; SEQ ID NO:59)	TGCCAGTGGC AGATGAAAAGT CTTT (rs110937_C2_1_ R; SEQ ID NO:60)	
rs9866013	3	GTGCCCTCAGAACCT TTGAGATCTGATCT ATTTTAAAGCTTCT TAGAAGAGAGATTG CAAAGTGGGTTGTTT CTTAGCCAGACAG GGCAGCAAAATAGG GGTGGCTGGTGGGA TGGGA (SEQ ID NO:5)	GTGCCCTCAGAACCT TTGAGATCTGATCT ATTTTAAAGCTTCT TAGAAGAGAGATTG CAAAGTGGGTTGTTT CTTAGCCAGACAG GGCAGGTAATAGG GGTGGCTGGTGGGA TGGGA (SEQ ID NO:6)	GTGCCCTC AGAACCTT TGAGATCT GAT (rs9866013_C3_1_F; SEQ ID NO:61)	TCCCATCCAC CAGCCACCC (rs9866013_C3_1_ R; SEQ ID NO:62)

Panel de SNP para la determinación de la fracción fetal					
SNP ID	Cr	Amplificación: Alelo 1	Amplificación: Alelo 2	Secuencia de cebador directo, nombre y SEQ ID NO:	Secuencia de cebador inverso, nombre y SEQ ID NO:
rs13182883	5	AGGTGTGTCCTCTT TTGTGAGGGAGGG GTCCCTTCTGGCCTA GTAGAGGGCCTGGC CTGCAGTGACATTC AAATCCTCAAGGAA CAGGTGGGAGGT GGGACAAAGG (SEQ ID NO:7)	AGGTGTGTCCTCTT TTGTGAGGGAGGG GTCCCTTCTGGCCTA GTAGAGGGCCTGGC CTGCAGTGACATTC AAATCCTCGAGGAA CAGGTGGGAGGT GGGACAAAGG (SEQ ID NO:8)	AGGTGTGT CTCTCTTT GTGAGGGG (rs13182883_C5_1_F; SEQ ID NO:63)	CCTTTGTCCCA CCTCCCAACC (rs13182883_C5_1_R; SEQ ID NO:64)
rs13218440	6	CCTCGCCTACTGTGC TGTTCTAACCATCA TGCTTTCCCTGAAT CTCTGAGCTTTTT CTGCTGTGGACTGA AACTTGATCCTGAG ATTACCTCTAGTCC CTTGAGCAGCCTCC TGGAACTACAGCT GGGATGG (SEQ ID NO:9)	CCTCGCCTACTGTGC TGTTCTAACCATCA TGCTTTCCCTGAAT CTCTGAGCTTTTT CTGCTGTGGACTGA AACTTGATCCTGAG ATTACCTCTAGTCC CTCTGGCAGCCTCC TGGAACTACAGCT GGGATGG (SEQ ID NO:10)	CCTGGCCT ACTGTGCT GTTTCTAA CC (rs13218440_C6_1_F; SEQ ID NO:65)	CCATCCAGCT GAGTATTCCA GGAG (rs13218440_C6_1_R; SEQ ID NO:66)
rs7041158	9	AATTGCAATGGTGA	AATTGCAATGGTGA	AATTGCAA	CCAGTGAGAA
		GAGGTTGATGGTAA AATCAAACGGAAT TGTTATTTGTCAIT CTGATGGACTGGAA CTGAGGATTTCAAT TTCTCTCCAACCA AGACACTTCTCACTG G (SEQ ID NO:11)	GAGGTTGATGGTAA AATCAAACGGAAT TGTTATTTGTCAIT CTGATGGACTGGAA CTGAGGATTTCAAT TTCTCTCCAACCA AGACACTTCTCACTG G (SEQ ID NO:12)	TGGTGAGA GGTTGATG GT (SEQ ID NO:67)	GTGTCTGGGT TGG (SEQ ID NO:68)

Panel de SNP para la determinación de la fracción fetal					
SNP ID	Cr	Amplificón: Alelo 1	Amplificón: Alelo 2	Secuencia de cebador directo, nombre y SEQ ID NO:	Secuencia de cebador inverso, nombre y SEQ ID NO:
rs740598	10	<p>GAAATGCCTTCTCAG GTAATGGAAGGTTA TCCAAATATTTTCG TAAAGTATTTCAAATA GCAATGGCTCGTCTA TGGTTAGTCTCACAG CCACATTTCTCAGAAC TGCTCAAACC (SEQ ID NO:13)</p>	<p>GAAATGCCTTCTCAG GTAATGGAAGGTTA TCCAAATATTTTCG TAAAGTATTTCAAATA GCAATGGCTCGTCTA TGGTTAGTCTCACAG CCACATTTCTCAGAAC TGCTCAAACC (SEQ ID NO:14)</p>	<p>GAAATGCC TTCTCAGG TAAATGGAA GGT (SEQ ID NO:69)</p>	<p>GGTTTGAGCA GTTCGAGAAAT GTGGCT (SEQ ID NO:70)</p>
rs10773760	12	<p>ACCCAAAACACTGG AGGGGCTCTTCTCA TTTTCCGTAAGACTGC AAGTGTAGCCGTC GGACCCAGCTTCTGT CTGGAAGTTCGTCA AATTGCAGTTAAGTC CAAAGTGGCCACAT AGCAGATAAGGG (SEQ ID NO:15)</p>	<p>ACCCAAAACACTGG AGGGGCTCTTCTCA TTTTCCGTAAGACTGC AAGTGTAGCCGTC GGACCCAGCTTCTGT CTGGAAGTTCGTCA AATTGCAGTTAAGTC CAAAGTGGCCACAT AGCAGATAAGGG (SEQ ID NO:16)</p>	<p>ACCCAAAA CACTGGAG GGGCTT (SEQ ID NO:71)</p>	<p>CCCTTATCTGC TATGTGGCATA CTTGG (SEQ ID NO:72)</p>
rs4530059	14	<p>GCACCCAGAAITTTAA ACAACGCTGACAAT AAATATGCACTCGA TGATGACTTCCCAGA GCTCCAGAAGCAAC TCCAGCACACAGAG AGGCGCTGATGTGC CTGTCAGGTGC (SEQ ID NO:17)</p>	<p>GCACCCAGAAITTTAA ACAACGCTGACAAT AAATATGCACTCGA TGATGACTTCCCAGA GCTCCAGAAGCAAC TCCAGCACACAGAG AGGCGCTGATGTGC CTGTCAGGTGC (SEQ ID NO:18)</p>	<p>GCACCAGA ATTTAAAC AACGCTGA CAA (SEQ ID NO:73)</p>	<p>GCACCTGACA GGCACATCAG CG (SEQ ID NO:74)</p>

Panel de SNP para la determinación de la fracción fetal					
SNP ID	Cr	Amplificación: Alelo 1	Amplificación: Alelo 2	Secuencia de cebador directo, nombre y SEQ ID NO:	Secuencia de cebador inverso, nombre y SEQ ID NO:
rs7205345	16	TGACTGTATACCCCA GGTGACCCCTTGGGT CATCTATCATAGA ACTTATCTCACAGAG TATAAGAGCTGATTT CTGTGCTGCCTCTC ACACTAGACTTCCAC ATCCTTAGTGC (SEQ ID NO:19)	TGACTGTATACCCCA GGTGACCCCTTGGGT CATCTATCATAGA ACTTATCTCACAGAG TATAAGAGCTGATTT CTGTGCTGCCTCTC ACACTAGACTTCCAC ATCCTTAGTGC (SEQ ID NO:20)	TGACTGTA TACCCAG GTGCACCC (SEQ ID NO:75)	GCACTAAGGA TGTGGAAATCT AGTGTG (SEQ ID NO:76)
rs8078417	17	TGTACGTGGTCACCA GGGACCGCTGGCG CTGGAGGGAGGCC CCGAGCTCGTGCC CCGTGAAGCTTCAG CTCCCTCCCAGGCT GTCCCTGAGGCTCT CTCACACT (SEQ ID NO:21)	TGTACGTGGTCACCA GGGACCGCTGGCG CTGGAGGGAGGCC CCGAGCTCGTGCC CCGTGAAGCTTCAG CTCCCTCCCAGGCT GTCCCTGAGGCTCT CTCACACT (SEQ ID NO:22)	TGTACGTG GTCACCA GGGACG (SEQ ID NO:77)	AGTGTGAGAA GAGCCTCAAG GACAGC (SEQ ID NO:78)
rs576261	19	CAGTGGACCCCTGCT GCACCTTCTCCCC TCCCATCAACTCTT TTGTGCCCTCCCTC CGTGAACACTTCT CTGTACCAACCCCTG GCCTCACAACTCTCT	CAGTGGACCCCTGCT GCACCTTCTCCCC TCCCATCAACTCTT TTGTGCCCTCCCTC CGTGAACACTTCT CTGTACCAACCCCTG GCCTCACAACTCTCT	CAGTGGAC CCTGTGC ACCTT (SEQ ID NO:79)	GTGGCAAAGG AGAGAGTTGT GAGG (SEQ ID NO:80)
		CCTTTGCCAC (SEQ ID NO:23)	CCTTTGCCAC (SEQ ID NO:24)		

Panel de SNP para la determinación de la fracción fetal					
SNP ID	Cr	Amplicón: Alelo 1	Amplicón: Alelo 2	Secuencia de cebador directo, nombre y SEQ ID NO:	Secuencia de cebador inverso, nombre y SEQ ID NO:
rs2567608	20	CAGTGGCATAGTAG TCCAGGGGCTCTCC TCAGCACTCCAGC ACCTCCAGGAGGC AGCAGGCAGGCAG AGAACCCGCTGGAA GAATCGGCGGAAGT TGTCGGAGAGG (SEQ ID NO:25)	CAGTGGCATAGTAG TCCAGGGGCTCTCC TCAGCACTCCAGC ACCTCCAGGAGGC AGCAGGCAGGCAG AGAACCCGCTGGAA GAATCGGCGGAAGT TGTCGGAGAGG (SEQ ID NO:26)	CAGTGGCA TAGTAGTC CAGGGGCT (SEQ ID NO:81)	CCTCTCCGACA ACTTCCGCCG (SEQ ID NO:82)

Tabla 11

SNP adicionales para la determinación de la fracción fetal					
SNP ID	Cr	Amplificación: Alelo 1	Amplificación: Alelo 2	Secuencia de cebador directo, nombre y SEQ ID NO:	Secuencia de cebador inverso, nombre y SEQ ID NO:
rs430046	16	AGGTCTGGGGCC GCTGAATGCCAAGC TGGGAATCTTAAAT GTTAAGGAACAAG GTCATACAATGAAT GGGTGATGTAAAA GCTTGGGAGGTGAT TTCGAGGGTAGGT GCTGGTTTAAATGG GAGGA (SEQ ID NO:27)	AGGTCTGGGGCCGC TGAATGCCAAGCTGG GAATCTTAAATGTTA AGGAACAAGTCAITA CAATGAATGGTGTGA TGTAAGCTTGGGA GGTATTTTGAGGG TAGGTGTGGGTTTA ATGGGAGGA (SEQ ID NO:28)	AGGTCTGG GGCCCGCT GAAT (rs430046_C I_1_F; SEQ ID NO:83)	TCCTCCCAITTA AACCCAGCAC CT (rs430046_C1_1_ R; SEQ ID NO:84)
rs9951171	18	ACGGTTCTGTCTCG TAGGGGAGAAAAG TCCTCGTTGTCTCT CTGGATGCAACAT GAGAGAGCAGCAC ACTGAGGCTTTATG GATTGCCCTGCCAC AAGTGAAACAGG (SEQ ID NO:29)	ACGGTTCTGTCTCTGT AGGGGAGAAAAGTCC TCCTCGTTCTCTGGG ATGCCAATGAGAGA GCAGCACACTGAGGC TTTTATGGTTGCCCTG CCACAAGTGAAACAGG (SEQ ID NO:30)	ACGGTTCT GTCTGTGA GGGAGAGA (rs9951171_ C1_1_F; SEQ ID NO:85)	CCTGTTCACTT GTGGCAGGGC A (rs9951171_C1_1_ R; SEQ ID NO:86)
rs338882	5	GCCAGTCAGATG GGCGTGTGGGTC TGCTCTCTCTCTC CTGCTCTGGCTT CAITTTCTCTCTCT CTGCTCACCTCTCT TTCGTGTGCTCTG CACACACACCTTTG GGACAAGGG CTGGA (SEQ ID NO:31)	GCCAGTCAGATGGG CGTGTGGCGTCTGT CTCTCTCTCTCTCTG TCTCTGGCTTCAITTT TCTCTCTCTCTCTCT ACCTCTCTCTCTCTG CTGTGCAACACACG TTTGGGACAAAGG CTGGA (SEQ ID NO:32)	GCCAGTC AGATGGGC GTCC (rs338882_C I_1_F; SEQ ID NO:87)	TCCAGCCCTTG TCCCAAACGT GT (rs338882_C1_1_ R; SEQ ID NO:88)

SNP adicionales para la determinación de la fracción fetal					
SNP ID	Cr	Amplificación: Alelo 1	Amplificación: Alelo 2	Secuencia de cebador directo, nombre y SEQ ID NO:	Secuencia de cebador inverso, nombre y SEQ ID NO:
rs10776839	9	GCCGGACCTGCGA AATCCAAAATGCC AAACATTCCCGCT CACATGATCCAGA GAGAGGGGACCCA GTGTTCCAGCTTG CAGCTGAGGAGCC CGAGGTTGCCGTCA	GCCGGACCTCGAAA TCCCAAAATGCCAAA CATTCGCGCTCACA TGATCCAGAGAGAG GGGACCCAGTGTTC CAGCTTGCAGCTGAG GAGCCGAGTTGCC GTCAGATCAGAGCCC	GCCGGACC TGCGAAAT CCCAA (rs10776839 Cl. I. F. SEQ ID NO:89)	CGGGCAACTG GGGCTCTGATC (rs10776839_C1_1 _R; SEQ ID NO:90)
		GATCAGAGCCCCA GTTGCCCG (SEQ ID NO:33)	CAGTTGCCCG (SEQ ID NO:34)		
rs9905977	17	AGCAGCCTCCCTCG ACTAGCTCACACTA CGATAAGGAAAAT TCATGAGCTGGTGT CCAAGGAGGGCTG GGTGACTCGTGGCT CAGTCAGCATCAAG ATTCCCTTCGTCTT CCCCCTGCC (SEQ ID NO:35)	AGCAGCCTCCCTCGA CTAGCTCACACTACG ATAAGGAAAATTCAT GAGCTGGTGTCCAAG GAGGGCTGGGTGACT CGTGGCTCAGTCAGC GTC AAGATTCCTTTC GTCCTTCCTCTGCCC (SEQ ID NO:36)	AGCAGCCT CCCTCGAC TAGCT (rs9905977_1 Cl. I. F. SEQ ID NO:91)	GGCAGAGGGG AAAGACGAAA GGA (rs9905977_C1_1 _R; SEQ ID NO:92)

SNP adicionales para la determinación de la fracción fetal					
SNP ID	Cr	Amplificón: Alelo 1	Amplificón: Alelo 2		
rs1277284	4	TGGCAITGCGCTGTA ATATACATAGCCAT GGTTTTTATAGGC AATTTAAGATGAAT AGCTTCTAAACTAT AGATAAGTTTCATT ACCCCAAGGAAGCT GAACTATAGCTACT TTACCCAAAATCAT TAGAATGGTGCTT (SEQ ID NO:37)	TGGCAITGCGCTGTA TATACATAGCCATGG TTTTTATAGGCAATT TAAAGATGAATAGCTT CTAAACTATAGATAA GTTTCATTACCCAG GAAAGCTGAACTATAG CTACTTTCGCCAAA TCATTAGAAITGGTGC TT (SEQ ID NO:38)	TGGCAITG CCTGTAAT ATACATAG (rs1277284_C4_1_F; SEQ ID NO:93)	AAGCACCAATT CTAATGATTTTT GG (rs1277284_C4_1_R; SEQ ID NO:94)
rs258684	7	ATGAAGCCITCCAC CAACTGCGCTGTATG ACTCATCTGGGGAC TTCTGCTCTATACT CAAAGTGGCTTGTAGT CACTGCCAATGTAT TTCCATATGAGGGA CGATGATTAATAAG GAAATATAGAAAC AACAACTGATC (SEQ ID NO:39)	ATGAAGCCCTCCACC AACTGCGCTGTATGAC TCAATCTGGGGACTTC TGCTCTATACTCAA GTGGCTTAGTCACTG CCAATGTATTTCCAT ATGAGGGACGGTGTAT TACTAAGGAAATATA GAAACAACAACCTGAT C (SEQ ID NO:40)	ATGAAGCC TTCCACCA ACTG (rs258684_C7_1_F; SEQ ID NO:95)	GATCAGTTGTT GTTTCTATATT TCCTT (rs258684_C7_1_R; SEQ ID NO:96)
rs1347696	8	ACAACAGAATCAG GTGATTTGGAGAAA AGATCACAGGCCTA GGCACCAAGGCCTT GAAAGGATGAAAGA ATGAAAAGATGGAC GGAACAAAATTAG GACCTTAAATTCITT GTTCAGITTCAG (SEQ ID NO:41)	ACAACAGAAATCAGGT GATTTGGAGAAAAGAT CACAGGCCTAGGCAC CCAAGGCTTGAAGGA TGAAAGAAATGAAAGA TGGACGGAAAGAAAAT TAGGACCTTAATTCITT TGTTCAAGITTCAG (SEQ ID NO:42)	ACAACAGA ATCAGGTG ATTTGGA (rs1347696_C8_4_F; SEQ ID NO:97)	CTGAACCTGAA CAAAGAAATTA AGGTC (rs1347696_C8_4_F; SEQ ID NO:98)

SNP adicionales para la determinación de la fracción fetal					
SNP ID	Cr	Amplión: Alelo 1	Amplión: Alelo 2		
rs508485	11	TTGGGGTAAATTTT CATTGTCAATGTG GAATTTAAATATAC CATCATCTACAAAG AATTCACAGAGTT AAATATCTTAAAGTT AAACACTTAAATA AGTGTTCGCTGAT ATTTGTGACAGA TAAACAGAGTCTAA TTCCCAACCCC (SEQ ID NO:43)	TTGGGGTAAATTTT AATTGTCATATGTGGA AITTTAAATATACCAT CATCTACAAAGAAAT CCACAGAGTTAAATA TCCTTAAGTTAAACAC TTAAAATAAGTGTTT GCGTGATATTTTGT GATAGATAAACAGAG TCTAAATCCCACCCC (SEQ ID NO:44)	TTGGGGTA AAATTTCA TTGTCA (rs508485_C 11_1.F; SEQ ID NO:99)	GGGGTGGGAA TTAGACTCTG (rs508485_C11_1 R; SEQ ID _NO100)
rs9788670	15	TGCAATTCAAATCA GGAAGTATGACCA AAAGACAGAGATC	TGCAATTCAAATCAG GAAAGTATGACCAAAA GACAGAGATCTTTT	TGCAATTC AAATCAGG AAGTATG	GCAACATCGA GGTTTGTCTAG (rs9788670_c15_ _
		TTTTTGGATGATC CCTAGCCTAGCAAT GCCTGGCAGCCATG CAGGTGCAATGTCA ACCTTAAATAATGT ATTGCAAACTCAGA GCTGACAAAACCTCG ATGTTGC (SEQ ID NO:45)	TGGATGATCCCTAGC CTAGCAATGCCCTGGC AGCCATGCGAGTGA ATGTCAAACCTTAAAT AATGATTTGCAAAIT CAGAGCTGACAAACC TCGATGTTGC (SEQ ID NO:46)	(rs9788670_c 15_2.F; SEQ ID NO:101)	2_R; SEQ ID NO:102)

SNP adicionales para la determinación de la fracción fetal					
SNP ID	Cr	Amplificación: Alelo 1	Amplificación: Alelo 2	Secuencia de cebador directo, nombre y SEQ ID NO:	Secuencia de cebador inverso, nombre y SEQ ID NO:
rs8137254	22	CTGTGCTCTGGGAA TAGCTGCAGAAAGTA ACTTGGGACCCAA AATAAAGCAGAAT GCTAATGTCAGATC CTGAGAACCAAGC CTGGGACTCTGGT GCCAATTCGGATTC TCCATGAGCATGGT (SEQ ID NO:47)	CTGTGCTCTGGAAAT AGCTGCAGAAAGTAAC TTGGGGACCCAAAAT AAAGCAGAATGCTAA TGTC AAGTCTGAGA ACCAAAGCCCTGGGAC TCTGGTGCCATTJGG 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 AAATTTCTTAATAT AGTTAATTATCGAG GGGAGGGAAAGCA AAGGAGCACAGGT AGTCCACAGAATA AGACACAAAGAAA CTCAAGCTGTG (SEQ ID NO:49)	TTTTTCCAGCCAACCTC AAGGCCAAAAAAAAT TTCTTAATATAGTTAT TATGGGAGGGAGGG GAAGCAAAAGGAGCA CAGGTAGTCCACAGA ATAGGCACAAAGAAA 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	TCTTCTGTCCTCCCT AAGCAAAACAACAT CCGCTTGTCTCTGT CTGTGTAACACAG TGAATGGGTGTGCA CGCTTGATGGCCCT CTGAGCCCTGTGTG CACAAACCAAAA (SEQ ID NO:51)	TCTTCTGTCCTCCCTAA GCAAAACAACATCCGC TTGTCTCTGTCTGTGT AACACAGTGAATGG GTGTGACACGCTTGTG GGCCTCTGAGCCCT TGTTGCACAAACCCAG AAA (SEQ ID NO:52)	TCTTCTCG TCCCCTAA GCAA (rs2182957_c13_1_F: SEQ ID NO:107)	TTTCTGGTTTG TGCAACAGG (rs2182957_c13_1_R: SEQ ID NO:108)

SNP adicionales para la determinación de la fracción fetal					
SNP ID	Cr	Amplión: Alelo 1	Amplión: Alelo 2	Secuencia de cebador directo, nombre y SEQ ID NO:	Secuencia de cebador inverso, nombre y SEQ ID NO:
rs3739005	2	CACATGGGGCATT AAGAAATGCCAG GGAGGAGAGGGA GAACGGTGCITTT CACATTTGCATTTG AATTTCCGAGTTCC CAGGATGTGTTTTT GTGTCATCGATGT (SEQ ID NO:53)	CACATGGGGCATT AGAAATGCCAGGGA GGAGGAGGAGAAC GGTGTCTTTTACAT TGCAITTTGAAATTTG AGITCCCAGGATGTG TTTTTGTCTCATCGA TCT (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 GTGAAATAAAAAAC AAATGCCATGTCT GTCTTTTATGGCA TTTTGGGACITTTAC ATTTCAAACATTTCC AGACATGTATCACA ACACGAAGGAATA ACAGTTCCAGGGAT ATCT (SEQ ID NO:55)	GGGCTCTGAGGTGT TGAAATAAAAAACAA TGTCCATGTCTGTCT TTTTATGGCAITTTGGG ACITTTACATTTCAAA CATTTCCAGCATGTGA TCACAACACGAGGGA ATAACAGTTCCAGGG ATATCT (SEQ ID NO:56)	GGGCTCTG AGGTGTGT GAAA (rs530022_c1 _2_F_SEQ ID NO:111)	AGATATCCCCTG GAACTGTATT CC (rs530022_c1_2_ _R_SEQ ID NO:112)
		TTTTTTGGATGATC CCTAGCCTAGCAAT GCCTGGCAGCCATG CAGGTGCAATGTCA ACCTTAAATAATGT ATTGCAAACTCAGA GCTGACAAAACCTCG ATGTTGC (SEQ ID NO:45)	TGGATGATCCCTAGC CTAGCAATGCCCTGGC AGCCATGCAGGTGCA ATGTCAAACCTTAAAT AATGTATTGCAAAIT CAGAGCTGACAAAAC TCGATGTTGC (SEQ ID NO:46)	(rs9788670_c 15_2_F_SEQ ID NO:101)	2_R_SEQ ID NO:102)

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

Para determinar simultáneamente la fracción fetal y la presencia o la ausencia de una aneuploidía en una muestra materna, se enriqueció una biblioteca de secuenciación primaria de ácidos nucleicos fetales y maternos en secuencias polimórficas de ácidos nucleicos diana y se secuenció de la forma siguiente.

Se preparó ADNcf purificado a partir de una muestra de plasma materno tal como se describe en el ejemplo, 1. Se usó una primera porción de la muestra de ADNcf purificado para preparar una biblioteca de secuenciación primaria usando el protocolo abreviado descrito en el ejemplo, 2. Se usó una segunda porción de la muestra de ADNcf purificado para amplificar secuencias polimórficas de ácidos nucleicos diana, es decir, SNP y preparar una biblioteca de secuenciación diana de la forma siguiente. Se amplificó ADNcf contenido en 5 µl de ADNcf purificado en un volumen de reacción de 50 µl que contenía 7,5 µl de una mezcla de cebadores 1 µM (tabla 5), 10 µl de mezcla maestra NEB 5X y 27 µl de agua. Se realizó un termociclado con el Gene Amp9700 (Applied Biosystems). Se utilizaron las condiciones de ciclado siguientes: incubación a 95°C durante 1 minuto, seguida de 30 ciclos a 95°C durante 20 segundos, 68°C durante 1 minuto y 68°C durante 30 s, seguidos de una incubación final a 68 °C durante 5 minutos. Se añadió un mantenimiento final a 4°C hasta que se retiraron las muestras para combinarlas con la porción no amplificada de la muestra de ADNcf purificado. El producto amplificado se purificó utilizando el sistema de purificación por PCR Agencourt AMPure XP (Nº de parte A63881; Beckman Coulter Genomics, Danvers, MA). Se añadió un mantenimiento final a 4°C hasta que se retiraron las muestras 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 tal como se describe en el ejemplo, 2. Las bibliotecas de secuenciación primaria y 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 tal como se describe en el ejemplo. El análisis de los datos de secuenciación para determinar aneuploidía se realizó tal como se describe en el ejemplo, 3 usando el genoma humano hg18 como genoma de referencia. El análisis de los datos de secuenciación para determinar la fracción fetal se realizó de la forma siguiente. De forma concomitante al análisis para determinar la aneuploidía, se analizaron los datos de secuenciación para determinar la fracción fetal. Después de la transferencia de la imagen y los archivos de designación de bases al servidor Unix que ejecuta el programa informático de Illumina "Genome Analyzer Pipeline" versión 1.51 tal como se describe en el ejemplo, 2c, las lecturas de 36 pb se alinearon con un "genoma SNP" utilizando el programa BOWTIE. El genoma SNP se identificó como la agrupación de las secuencias de ADN polimórfico, es decir, SEQ ID NO: 1-56, que abarcan los alelos de los 13 SNP divulgados en la tabla 10 en el ejemplo, 7. Para el análisis de la fracción fetal, solo se usaron las lecturas que se mapearon únicamente al genoma del SNP. Las lecturas que coincidían perfectamente con el genoma SNP se contaron como etiquetas y se filtraron. De las lecturas restantes, solo las lecturas que tenían uno o dos desajustes se contaron como etiquetas y se incluyeron en el análisis. Se contaron las etiquetas mapeadas a cada uno de los alelos SNP y se determinó la fracción fetal. Aproximadamente un millón del número total de etiquetas de secuencia obtenidas a partir de la secuenciación de la biblioteca enriquecida correspondía a etiquetas mapeadas al genoma de referencia SNP. La figura 17 muestra un gráfico de la relación del número de etiquetas de secuencia mapeadas a cada cromosoma y el número total de etiquetas mapeadas a todos los cromosomas (1-22, X e Y) obtenidas a partir de la secuenciación de una biblioteca de ADNcf no enriquecida (●) y una biblioteca de ADNcf enriquecida con el 5% (■) o el 10% (◆) de biblioteca de 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 a la información de secuenciación utilizada para determinar aneuploidía. En las tablas 12, 13 y 14 siguientes se proporcionan ejemplos de determinación de la fracción fetal para muestras obtenidas de sujetos portadores de un feto con aneuploidía cromosómica.

**a. Determinación de la fracción fetal**

La fracción fetal se calculó como:

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

en la que alelo<sub>x</sub> es un alelo informativo.

Tabla 12

<b>Determinación simultánea de aneuploidía y fracción fetal: determinación de la fracción fetal</b>				
<b>Muestra (cariotipo)</b>	<b>ID</b>	<b>SNP</b>	<b>RECuentos de Etiquetas de SNP</b>	<b>Fracción Fetal (%)</b>
11409 (47, XY+21)		rs13182883.1 Cr.5 longitud=111 alelo=A	261	4,41
		rs13182883.2 Cr.5 longitud=111 alelo=G	5918	
		rs740598.1 Cr.10 longitud=114 alelo=A	5545	7,30
		rs740598.2 Cr.10 longitud=114 alelo=G	405	
		rs8078417.1 Cr.17 longitud=110 alelo=C	8189	6,74
		rs8078417.2 Cr.17 longitud=110 alelo=T	121470	
		rs576261.1 Cr.19 longitud=114 alelo=A	58342	7,62
		rs576261.2 Cr.19 longitud=114 alelo=C	4443	
<b>Fracción fetal (Media±S.D.) = 6,53±1,45</b>				
<b>Muestra ID</b>				
95133 (47, XX+18)		rs1109037.1 Cr.2 longitud=126 alelo=A	12229	2,15
		rs1109037.2 Cr.2 longitud=126 alelo=G	263	
		rs13218440.1 Cr.6 longitud=139 alelo=A	55949	3,09
		rs13218440.2 Cr.6 longitud=139 alelo=G	1729	
		rs7041158.1 Cr.9 longitud=117 alelo=C	7281	4,12
		rs7041158.2 Cr.9 longitud=117 alelo=T	300	
		rs7205345.1 Cr.16 longitud=116 alelo=C	53999	2,14
		rs7205345.2 Cr.16 longitud=116 alelo=G	1154	
<b>Fracción fetal (Media±S.D.) = 2,9± 0,9</b>				
<b>Muestra ID</b>				
51236 (46, XY+13)		rs13218440.1 Cr.6 longitud=139 alelo=A	1119	1,65
		rs13218440.2 Cr.6 longitud=139 alelo=G	67756	
		rs560681.1 Cr.1 longitud=111 alelo=A	14123	5,18
		rs560681.2 Cr.1 longitud=111 alelo=G	732	
		rs7205345.1 Cr.16 longitud=116 alelo=C	18176	1,63
		rs7205345.2 Cr.16 longitud=116 alelo=G	296	
		rs9866013.1 Cr.3 longitud=121 alelo=C	117	2,33
		rs9866013.2 Cr.3 longitud=121 alelo=T	5024	

<b>Fración fetal (Media±S.D.) = 2,7± 1,7</b>			
<b>Muestra ID</b>			
<b>54430 (45, XO)</b>	rs1109037.1 Cr.2 longitud=126 alelo=A	19841	1,80
	rs1109037.2 Cr.2 longitud=126 alelo=G	357	
	rs9866013.1 Cr.3 longitud=121 alelo=C	12931	3,81
	rs9866013.2 Cr.3 longitud=121 alelo=T	493	
	rs7041158.1 Cr.9 longitud=117 alelo=C	2800	4,25
	rs7041158.2 Cr.9 longitud=117 alelo=T	119	
	rs740598.1 Cr.10 longitud=114 alelo=A	12903	4,87
	rs740598.2 Cr.10 longitud=114 alelo=G	628	
	rs10773760.1 Cr.12 longitud=128 alelo=A	46324	4,65
	rs10773760.2 Cr.12 longitud=128 alelo=G	2154	
<b>Fración fetal (Media±S.D.) = 3,9± 1,2</b>			

**b. Determinación de aneuploidía**

5 La determinación de aneuploidía de los cromosomas 21, 13, 18 y X se realizó utilizando dosis de cromosomas tal como se describe en el ejemplo, 4. La dosis de cromosoma calificado, la varianza y la diferenciabilidad para los cromosomas 21, 18, 13, X e Y se proporcionan en las tablas X e Y. La clasificación de los cromosomas de normalización identificados por las dosis de cromosomas determinadas a partir de la secuenciación de la biblioteca enriquecida fue la misma que la determinada a partir de la secuenciación de una biblioteca primaria (no enriquecida) del ejemplo, 4. La figura 17 muestra que la secuenciación de una biblioteca que se ha enriquecido en secuencias diana polimórficas, por ejemplo, SNP, no se ve afectada por la inclusión de los productos SNP amplificados.

10

**Tabla 13**

<b>Dosis de cromosoma calificado, Varianza y Diferenciabilidad para los cromosomas 21 y 18</b>								
	<b>21 (n=35)</b>				<b>18 (n=40)</b>			
	Promedio	Desv Est	CV	Prueba t	Promedio	Desv Est	CV	Prueba t
cr1	0,15332	0,002129	1,39	1,06E-10	0,32451	0,008954	2,76	2,74E-03
cr2	0,15106	0,002053	1,36	8,52E-08	0,31984	0,001783	0,56	5,32E-05
cr3	0,18654	0,004402	2,36	8,07E-07	0,39511	0,002364	0,60	1,93E-05
cr4	0,21578	0,011174	5,18	1,47E-04	0,45714	0,014794	3,24	1,37E-03
cr5	0,21068	0,005332	2,53	1,08E-06	0,44626	0,003250	0,73	3,18E-05
cr6	0,22112	0,005453	2,47	1,74E-06	0,46818	0,003434	0,73	2,24E-05
cr7	0,24233	0,002314	0,96	2,39E-08	0,51341	0,005289	1,03	1,24E-04
cr8	0,24975	0,003772	1,51	1,06E-07	0,52898	0,002161	0,41	6,32E-05
cr9	0,31217	0,003050	0,98	1,60E-09	0,66100	0,014413	2,18	8,17E-04
cr10	0,25550	0,003164	1,24	2,42E-11	0,54091	0,013953	2,58	2,26E-03
cr11	0,26053	0,002596	1,00	1,32E-10	0,55158	0,013283	2,41	1,29E-03
cr12	0,27401	0,002061	0,75	1,40E-08	0,58032	0,007198	1,24	1,57E-04
cr13	0,41039	0,017637	4,30	3,09E-05	0,86961	0,021614	2,49	2,36E-04
cr14	0,40482	0,002908	0,72	1,10E-08	0,85732	0,011748	1,37	2,16E-04
cr15	0,41821	0,008238	1,97	1,24E-10	0,88503	0,029199	3,30	5,72E-03
cr16	0,40668	0,021232	5,22	2,91E-05	0,86145	0,056245	6,53	1,04E-01
cr17	0,42591	0,027001	6,34	5,85E-04	0,90135	0,068151	7,56	1,24E-01
cr18	0,46529	0,016239	3,49	8,02E-09				

ES 2 909 841 T3

<b>Dosis de cromosoma calificado, Varianza y Diferenciabilidad para los cromosomas 21 y 18</b>								
	<b>21 (n=35)</b>				<b>18 (n=40)</b>			
	Promedio	Desv Est	CV	Prueba t	Promedio	Desv Est	CV	Prueba t
cr19	0,63003	0,063272	10,04	3,30E-02	1,33522	0,150794	11,29	3,04E-01
cr20	0,49925	0,023907	4,79	1,65E-05	1,05648	0,064440	6,10	7,98E-02
cr21					2,06768	0,087175	4,22	5,10E-05
cr22	0,88726	0,083330	9,39	3,43E-02	1,87509	0,198316	10,58	2,43E-01
crX	0,27398	0,016109	5,88	1,16E-04	0,58665	0,027280	4,65	7,50E-02

Tabla 14

<b>Dosis de cromosoma calificado, varianza y diferenciabilidad para los cromosomas 13, X e Y</b>								
	<b>13 (n=47)</b>				<b>X (n=20)</b>			
	Promedio	Desv Est	CV	Dif	Promedio	Desv Est	CV	Prueba t
cr1	0,37213	0,018589	5,00	2,41	0,58035	0,02706	4,66	5,68E-05
cr2	0,36707	0,010067	2,74	3,03	0,57260	0,01432	2,50	1,53E-09
cr3	0,45354	0,008121	1,79	3,67	0,70741	0,01126	1,59	9,04E-13
cr4	0,52543	0,005306	1,01	2,39	0,82144	0,01192	1,45	5,86E-16
cr5	0,51228	0,008273	1,61	3,95	0,79921	0,01100	1,38	2,32E-13
cr6	0,53756	0,008901	1,66	3,91	0,83880	0,01261	1,50	3,64E-13
cr7	0,58908	0,018508	3,14	2,83	0,91927	0,02700	2,94	1,86E-08
cr8	0,60695	0,015797	2,60	3,05	0,94675	0,02173	2,30	3,40E-10
cr9	0,75816	0,033107	4,37	2,59	1,18180	0,04827	4,08	9,63E-06
cr10	0,62018	0,029891	4,82	2,56	0,96642	0,04257	4,40	4,55E-05
cr11	0,63248	0,029204	4,62	2,55	0,98643	0,04222	4,28	1,82E-05
cr12	0,66574	0,023047	3,46	2,76	1,03840	0,03301	3,18	1,26E-07
cr13					1,56355	0,01370	0,88	6,33E-17
cr14	0,98358	0,035331	3,59	2,67	1,58114	0,08076	5,11	2,29E-04
cr15	1,01432	0,055806	5,50	2,39	1,53464	0,12719	8,29	2,01E-02
cr16	0,98577	0,085933	8,72	2,17	1,61094	0,14829	9,21	2,68E-02
cr17	1,03217	0,100389	9,73	2,13	1,74904	0,07290	4,17	1,62E-04
cr18	1,13489	0,040058	3,53	2,62	2,38397	0,30515	12,80	1,07E-01
cr19	1,52678	0,203732	13,34	1,98	1,88186	0,14674	7,80	1,56E-02
cr20	1,20919	0,100371	8,30	2,27	3,71853	0,22406	6,03	4,21E-04
cr21	2,38087	0,132418	5,56	2,29	3,35158	0,40246	12,01	8,66E-02
cr22	2,14557	0,271281	12,64	2,13	0,58035	0,02706	4,66	5,68E-05
crX	0,66883	0,029157	4,36	1,04				
cr2-6	0,46965	0,006987	1,49	4,17				
cr3-6	0,50496	0,005373	1,06	5,16				
	<b>Y (n=25)</b>							
	<b>Promedio</b>	<b>Desv Est</b>	<b>CV</b>	<b>Prueba t</b>				
Cr 1-22, X	0,00728	0,00227	31,19	1,30E-13				

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

La tabla 12 muestra los datos para la determinación de la fracción fetal en muestras ilustrativas. Los valores de dosis de cromosoma calculados para los cromosomas 21, 18, 13, X e Y en muestras de ensayo ilustrativas correspondientes se proporcionan en las tablas 15, 16, 17 y 18, respectivamente.

**Trisomía 21**

La tabla 8 proporciona la dosis calculada para el cromosoma 21 en la muestra de ensayo (11409). Se utilizó el cromosoma 14 como cromosoma de normalización. El umbral calculado para el diagnóstico positivo de aneuploidía T21 se estableció en 2 desviaciones estándar de la media de las muestras calificadas (normales). Se dio un diagnóstico para T21 basado en que la dosis de cromosoma en la muestra de ensayo era superior al umbral establecido. Las doce muestras de T21 que se confirmó que eran T21 por cariotipo se identificaron en una población de 48 muestras de sangre.

**Tabla 15**

<b>Dosis de cromosoma para una aneuploidía T21</b>			
<b>Cromosoma</b>	<b>Densidad de etiquetas de secuencia</b>	<b>Dosis de cromosoma para el cr 21</b>	<b>Umbral</b>
Cr21	264,404	0,439498	0,410634
Cr14	601,605		

**Trisomía 18**

La tabla 9 proporciona la dosis calculada para el cromosoma 18 en una muestra de ensayo (95133). El cromosoma 8 se usó como el cromosoma de normalización. En este caso, el cromosoma 8 tenía la menor variabilidad y la mayor diferenciabilidad. El umbral calculado para el diagnóstico positivo de aneuploidía T18 se estableció en > 2 desviaciones estándar de la media de las muestras calificadas (no T18). Se dio un diagnóstico para T18 basado en que la dosis de cromosoma en la muestra de ensayo era superior al umbral establecido. Se identificaron ocho muestras T18 usando dosis de cromosomas y se confirmó que eran T18 por cariotipado.

**Tabla 16**

<b>Dosis de cromosoma para una aneuploidía T18</b>				
<b>Cromosoma</b>	<b>Densidad de etiquetas de secuencia</b>	<b>Dosis de cromosoma para el cr 18</b>	<b>Umbral</b>	
Cr18	604,291	0,550731	0,533297	
Cr8	1,097,253			

**Trisomía 13**

Las tablas 10 y 11 proporcionan la dosis calculada para el cromosoma 13 en una muestra de ensayo (51236). El 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 etiquetas de secuencia para el cromosoma 4 como el cromosoma de normalización, mientras que la dosis proporcionada en la tabla 11 se determinó utilizando el promedio de las relaciones de densidades de etiquetas de secuencia para el grupo de cromosomas 3, 4, 5 y 6 como el cromosoma de normalización. Se dio un diagnóstico para T13 basado en que la dosis de cromosoma en la muestra de ensayo era superior al umbral establecido. Se identificó una muestra T13 usando dosis de cromosomas y se confirmó que eran T13 por cariotipado.

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

Tabla 17

Dosis de cromosoma para una aneuploidía T13				
Cromosoma	Densidad de etiquetas de secuencia		Dosis de cromosoma para el cr 13	Umbral
Cr13	669,872		0,538140	0,536044
Cr4	1,244,791			

5

Tabla 18

Dosis de cromosoma para una aneuploidía T13				
Cromosoma	Densidad de etiquetas de secuencia		Dosis de cromosoma para el cr 13	Umbral
Cr13	669,872		0,532674	0,515706
Cr3	1,385,881			
Cr4	1,244,791			
Cr5	1,229,257			
Cr6	1,170,331			

**Síndrome de Turner (monosomía X)**

10

Tres muestras que tenían una dosis de cromosoma inferior a la del umbral establecido se identificaron como que tenían menos de un cromosoma X. Se determinó que las mismas muestras tenían una dosis de cromosoma Y inferior al umbral establecido, lo que indica que las muestras no tenían cromosoma Y.

15

Las dosis calculadas para los cromosomas X e Y en la muestra de ensayo de monosomía X ilustrativa (54430) se proporcionan en la tabla 12. Se seleccionó el cromosoma 4 como el cromosoma de normalización para calcular la dosis para el cromosoma X; y todos los cromosomas, es decir, 1-22 e Y se usaron como cromosomas de normalización. 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 (sin monosomía X).

20

Tabla 19

Dosis de cromosoma para un síndrome de Turner (monosomía X)			
Cromosoma	Densidad de etiquetas de secuencia	Dosis de cromosoma para el Cr X	Umbral
CrX	904,049	0,777990	0,797603
Cr4	1,162,031		
CrY	390	0,0004462	0,002737754
Cr (1-22, X) (Promedio)	874,108,1		

25

Así, el método permite la determinación simultánea de aneuploidías cromosómicas y la fracción fetal mediante la secuenciación masivamente paralela de una muestra materna que comprende una mezcla de ADNcf fetal y materno que se ha enriquecido en una pluralidad de secuencias polimórficas, cada una de las cuales comprende un SNP. En este ejemplo, la mezcla de ácidos nucleicos fetales y maternos se enriqueció mediante la combinación de una porción de una biblioteca de secuenciación que se construyó a partir de secuencias polimórficas fetales y maternas amplificadas con una biblioteca de secuenciación que se construyó a partir de la mezcla de ADNcf fetal y materna original no amplificada restante.

30

**Ejemplo 9****Determinación simultánea de aneuploidía y fracción fetal:****enriquecimiento de ácidos nucleicos fetales y maternos en una muestra de ADNcf purificado**

Para enriquecer el ADNcf fetal y materno contenido en una muestra purificada de ADNcf extraído de una muestra de plasma materno, se usó una porción del ADNcf purificado para amplificar secuencias polimórficas de ácido nucleico diana, comprendiendo cada una de las mismas un SNP elegido del panel de SNP que se proporciona en la tabla 6.

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

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 cebadores 1 uM (tabla 5), 10 μl de mezcla maestra NEB 5X y 27 μl de agua. Se realizó un termociclado con el Gene Amp9700 (Applied Biosystems). Se utilizaron las condiciones de ciclado siguientes: incubación a 95°C durante 1 minuto, seguida de 30 ciclos a 95°C durante 20 segundos, 68°C durante 1 minuto y 68°C durante 30 s, seguidos de una incubación final a 68 °C durante 5 minutos. Se añadió un mantenimiento final a 4°C hasta que se retiraron las muestras para combinarlas con la porción no amplificada de la muestra de ADNcf purificado. El producto amplificado se purificó con el sistema de purificación por PCR Agencourt AMPure XP (Nº de parte A63881; Beckman Coulter Genomics, Danvers, MA) y la concentración se cuantificó usando Nanodrop 2000 (Thermo Scientific, Wilmington, DE). El producto de amplificación purificado se diluyó 1:10 en agua y se añadieron 0,9 μl (371 μg) a 40 μl de muestra de ADNcf purificado para obtener un pico del 10%. El ADNcf fetal y materno enriquecido presente en la muestra de ADNcf purificado se usó para preparar una biblioteca de secuenciación y se secuenció tal como se describe en el ejemplo, 2.

La tabla 13 proporciona los recuentos de etiquetas obtenidos para cada uno de los cromosomas 21, 18, 13, X e Y, es decir, densidad de etiquetas de secuencia y recuentos de etiquetas obtenidos para las secuencias polimórficas informativas contenidas en el genoma de referencia de SNP, es decir, la densidad de etiqueta de SNP. Los datos muestran que la información de secuenciación se puede obtener mediante la secuenciación de una sola biblioteca construida a partir de una muestra de ADNcf materno purificado que se ha enriquecido en secuencias que comprenden SNP para determinar simultáneamente la presencia o la 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 era cuantificable a partir de los resultados de secuenciación de cinco SNP informativos y se determinó que era del 3,84%. Las densidades de etiquetas de secuencia se proporcionan para los cromosomas 21, 13, 18, X e Y. La muestra AFR105 fue la única muestra que se sometió al protocolo de enriquecimiento de ADNcf purificado para secuencias polimórficas amplificadas. Así, no se proporcionaron coeficientes de variación ni ensayos 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 único proceso de secuenciación.

**Tabla 20**

<b>Determinación simultánea de aneuploidía y fracción fetal: enriquecimiento de ácidos nucleicos fetales y maternos en una muestra de ADNcf purificado</b>					
<b>Aneuploidía</b>					
	Cromosoma 21	Cromosoma 18	Cromosoma 13	Cromosoma X	Cromosoma Y
Densidad de etiquetas de secuencia de	178763	359529	388204	572330	2219
Cariotipo	No afectado	No afectado	No afectado	No afectado	No afectado
<b>Fracción fetal</b>					
SNP	DENSIDAD DE ETIQUETAS DE SNP		FRACCIÓN FETAL (%)		
rs10773760.1 Cr.12 longitud=128 alelo=A	18903		2,81		
rs10773760.2 Cr.12 longitud=128 alelo=G	532				
rs1109037.1 Cr.2 longitud=126 alelo=A	347		5,43		
rs1109037.2 Cr.2 longitud=126 alelo=G	6394				

<b>Determinación simultánea de aneuploidía y fracción fetal: enriquecimiento de ácidos nucleicos fetales y maternos en una muestra de ADNcf purificado</b>		
<b>Aneuploidía</b>		
rs2567608.1 Cr.20 longitud=110 alelo=A	94503	1,74
rs2567608.2 Cr.20 longitud=110 alelo=G	1649	
rs7041158.1 Cr.9 longitud=117 alelo=C	107	5,61
rs7041158.2 Cr.9 longitud=117 alelo=T	6	
rs8078417.1 Cr.17 longitud=110 alelo=C	162668	3,61
rs8078417.2 Cr.17 longitud=110 alelo=T	5877	
<b>Fracción fetal (Media±S.D.) = 3,8± 1,6</b>		

**Ejemplo 10****5 Determinación simultánea de aneuploidía y fracción fetal: enriquecimiento de ácidos nucleicos fetales y maternos en una muestra de plasma**

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

Se amplificó ADNcf contenido en 15 µl de plasma libre de células en un volumen de reacción de 50 µl que contenía 9 µl de una mezcla de cebadores 1 µM (15 plex tabla 5), 1 µl de ADN polimerasa de sangre Phusion, 25 µl del tampón de PCR en sangre 2X Phusion que contenía trifosfatos de desoxinucleótidos (dNTP: dATP, dCTP, dGTP y dTTP). Se realizó un termociclado con Gene Amp9700 (Applied Biosystems) utilizando las condiciones de ciclado siguientes: incubación a 95°C durante 3 minutos, seguida de 35 ciclos a 95°C durante 20 segundos, 55°C durante 30 s y 70°C durante 1 minuto, seguidos de una incubación final a 68°C durante 5 minutos. Se añadió un mantenimiento final a 4°C hasta que se retiraron las muestras para combinarlas con la porción no amplificada de la muestra libre de células. El producto amplificado se diluyó 1:2 con agua y se analizó con 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 tal 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 tal como se describe en los ejemplos 2 y 3.

Los resultados se proporcionan en la tabla 21. En el ejemplo, dado, los datos muestran que la fracción de ADN fetal en la muestra de plasma SAC2517 era cuantificable a partir de los resultados de secuenciación de un SNP informativo y se determinó que era del 9,5%. En el ejemplo, dado, el cariotipado demostró que la muestra SAC2517 no se vio 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 en plasma para secuencias polimórficas amplificadas. Así, no se pudieron determinar coeficientes de variación ni ensayos de diferenciabilidad. El ejemplo, demuestra que el enriquecimiento de la mezcla de ADNcf fetal y materno presente en una muestra de plasma para secuencias de ácido nucleico que comprenden al menos un SNP informativo se puede usar para proporcionar la secuencia requerida y los recuentos de etiquetas de SNP para determinar la aneuploidía y la fracción fetal a partir de un único proceso de secuenciación.

Tabla 21

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

<b>Aneuploidía</b>					
	Cromosoma 21	Cromosoma 18	Cromosoma 13	Cromosoma X	Cromosoma Y
Densidad de etiquetas de secuencia	183851	400582	470526	714055	2449
Cariotipo	No afectado	No afectado	No afectado	No afectado	No afectado
<b>Fracción fetal</b>					
SNP	RECUENTOS DE ETIQUETAS			FRACCIÓN FETAL (%)	
rs10773760.1 Cr12 longitud=128 alelo=A	8536			9,49	
rs10773760.2 Cr12 longitud=128 alelo=G	89924				

**Ejemplo 11**

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

Para determinar simultáneamente la presencia o la ausencia de una aneuploidía y la fracción fetal en una mezcla de ADNcf fetal y materno obtenido de una muestra materna, la mezcla se enriquece en secuencias polimórficas que comprenden STR, se secuencian y se analizan los datos. El enriquecimiento puede ser de una biblioteca de secuenciación tal como se describe en el ejemplo, 8, de una muestra de ADNcf purificada tal como se describe en el ejemplo, 9, o de una muestra de plasma tal como se describe en el ejemplo, 10. En cada caso, la información de secuenciación se obtiene secuenciando una sola biblioteca, lo que permite determinar simultáneamente la presencia o la ausencia de una aneuploidía y la fracción fetal. Preferentemente, la biblioteca de secuenciación se prepara utilizando el protocolo abreviado proporcionado en el ejemplo, 2.

Los STR que se amplifican se eligen de los STR codis y no codis divulgados en la tabla 22, y la amplificación de las secuencias polimórficas de STR se obtiene utilizando los correspondientes conjuntos de cebadores proporcionados. Algunos de los STR que se han divulgado y/o analizado previamente para determinar la fracción fetal se enumeran en la tabla 22 y se divulgan en las solicitudes provisionales de Estados Unidos 61/296.358 y 61/360.837.

Tabla 22

<b>miniSTR CODIS y NO CODIS</b>				
<b>Locus STR (Nombre del marcador)</b>	<b>Ubicación del cromosoma</b>	<b>Intervalo de tamaño (pb)</b>	<b>Acceso de GenBank</b>	<b>Secuencias de cebadores (directo/inverso)</b>
<b>Loci de miniSTR codis*</b>				
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	CCTGTTCCCTTATTTCCC(TH01_F; SEQ ID NO:117)
				GGGAACACAGACTCCATGGTG(TH01_R; SEQ ID NO:118)

ES 2 909 841 T3

miniSTR CODIS y NO CODIS				
Locus STR (Nombre del marcador)	Ubicación del cromosoma	Intervalo de tamaño (pb)	Acceso de GenBank	Secuencias de cebadores (directo/inverso)
<b>Loci de miniSTR codis*</b>				
TPOX	2p25.3	65-101	M68651	CTTAGGGAACCCTCACTGAATG(TPOX_F ; SEQ ID NO:119)
				GTCCTTGTCAGCGTTTATTTGC(TPOX _R; SEQ ID NO:120)
vWA	12p13.31	88-148	M25858	AATAATCAGTATGTGACTTGGATTGA(v WA_F; SEQ ID NO:121)
				ATAGGATGGATGGATAGATGGA(vWA_R ; SEQ ID NO:122)
D3S1358	3p21.31	72-120	NT_005997	CAGAGCAAGACCCTGTCTCAT(D3 S 1358_ F; SEQ ID NO:123)
				TCAACAGAGGCTTGCATGTAT(D3 S 1358_ R; SEQ ID NO:124)
D5S818	5q23.2	81-117	AC008512	GGGTGATTTTCCTCTTTGGT(D5S818 F; SEQ ID NO:125)
				AACATTTGTATCTTTATCTGTATCCTTAT TTAT(D5S818_ R; SEQ ID NO:126)
D7S820	7q21.11	136-176	AC004848	GAACACTTGTCATAGTTTAGAACGAAC( D7S820 F; SEQ ID NO:127)
				TCA TTGACAGAA TTGCACCA(D7S820 _R; SEQ ID NO:128)
D8S1179	8q24.13	86-134	AF216671	TTTGTATTTTCATGTGTACATTCGTATC(D 7S820 F; SEQ ID NO:129)
				ACCTATCCTGTAGATTATTTTCACTGTG(D7S820_R; SEQ ID NO:130)
D13S317	13q31.1	88-132	AL353628	TCTGACCCATCTAACGCCTA(D13 S317 _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(D16S53 9_R; SEQ ID NO:134)
D18S51	18q21.33	113-193	AP001534	TGAGTGACAAATTGAGACCTT(D18S51_F ; SEQ ID NO:135)
				GTCTTACAATAACAGTTGCTACTATT(D 1 8S51_R; SEQ ID NO:136)
D21S11	21q21.1	153-221	AP000433	ATTCCCCAAGTGAATTGC(D21S11_F; SEQ ID NO:137)
				GGTAGATAGACTGGATAGATAGACGA(D 21S11_R; SEQ ID NO:138)
D2S1338	2q35	90-142	AC01036	TGAAACAGAAATGGCTTGG(D2S1338_F ; SEQ ID NO:139)
				GATTGCAGGAGGGAAGGAAG(D2S1338_ R; SEQ ID NO:140)

ES 2 909 841 T3

miniSTR CODIS y NO CODIS				
Locus STR (Nombre del marcador)	Ubicación del cromosoma	Intervalo de tamaño (pb)	Acceso de GenBank	Secuencias de cebadores (directo/inverso)
<b>Loci de miniSTR codis*</b>				
Penta D	21q22.3	94-167	AP001752	GAGCAAGACACCATCTCAAGAA(Penta D_F; SEQ ID NO:141)
				GAAATTTTACATTTATGTTTATGATTCTC T(Penta D_R; SEQ ID NO:142)
Penta E	15q26.2	80-175	AC027004	GGCGACTGAGCAAGACTC(Penta E_F; SEQ ID NO:143)
				GGTTATTAATTGAGAAAACCTTACA(Penta E_R; SEQ ID NO:144)
<b>Loci de miniSTR no codis*</b>				
D22S1045	22q12.3	82 - 115	AL022314 (17)	ATTTTCCCGATGATAGTAGTCT(D22S1045_F; SEQ ID NO:145)
				GCGAATGTATGATTGGCAATATTTT(D22S1045_R; SEQ ID NO:146)
D20S1082	20q13.2	73 - 101	AL158015	ACATGTATCCCAGAACTTAAAGTAAAC(D20S1082 F; SEQ ID NO:147)
				GCAGAAGGGAAAATTGAAGCTG(D20S1082_R; SEQ ID NO:148)
D20S482	20p13	85 - 126	AL121781 (14)	CAGAGACACCGAACCAATAAGA(D20S482_F; SEQ ID NO:149)
				GCCACATGAATCAATTCCTATAATAAA(D20S482_R; SEQ ID NO:150)
D18S853	18p11.31	82 - 104	AP005130 (11)	GCACATGTACCCTAAAACCTTAAAAT(D18S853_F; SEQ ID NO:151)
				GTCAACCAAACCTCAACAAGTAGTAA(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)	GCACCCAAAACCTGAATGTCATA(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(D12ATA63 SEQ ID NO:159) F;
				GGAAAAGACATAGGATAGCAATTT(D12ATA63_R; SEQ ID NO:160)

ES 2 909 841 T3

miniSTR CODIS y NO CODIS					
Locus STR (Nombre del marcador)	Ubicación del cromosoma	Intervalo de tamaño (pb)	Acceso de GenBank	Secuencias de cebadores (directo/inverso)	
<b>Loci de miniSTR codis*</b>					
D11S4463	11q25	88 - 116	AP002806 (14)	TCTGGATTGATCTGTCTGTCC(D 11 S4463 F; SEQ ID NO:161)	
				GAATTAAATACCATCTGAGCACTGAA(D11S 4463_R; SEQ ID NO:162)	
D10S1435	10p15.3	82-139	AL354747 (11)	TGTTATAATGCATTGAGTTTTATTCTG(D10S 1435_F; SEQ ID NO:163)	
				GCCTGTCTCAAAAATAAAGAGATAGACA(D 10S1435 R; SEQ ID NO:164)	
D10S1248	10q26.3	79 - 123	AL391869 (13)	TTAATGAATTGAACAAATGAGTGAG(D10S 248_F; SEQ ID NO:165) 1	
				GCAACTCTGGTTGTATTGTCTTCAT(D10S12 48_R; SEQ ID NO:166)	
D9S2157	9q34.2	71 - 107	AL162417 (10)	CAAAGCGAGACTCTGTCTCAA(D9S2157 _F; SEQ ID NO:167)	
				GAAAATGCTATCCTCTTTGGTATAAAT(D9S 2157_R; SEQ ID NO:168)	
D9S1122	9q21.2	93 - 125	AL161789 (12)	GGGTATTTCAAGATAACTGTAGATAGG(D9 S1122 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)	
				GTGAAAAGTAGATATAATGGTTGGTG(D6 S1017_R; SEQ ID NO:174)	
D6S474	6q21	107 - 136	AL357514 (17)	GGTTTTCCAAGAGATAGACCAATTA(D6S47 4_F; SEQ ID NO:175)	
				GTCCTCTCATAAATCCCTACTCATATC(D6S 474_R; SEQ ID NO:176)	
D5S2500	5q11.2	85 - 126	AC008791 (17)	CTGTTGGTACATAATAGGTAGGTAGGT(D5S 2500_F; SEQ ID NO:177)	
				GTCGTGGGCCCCATAAATC(D5S2500_R; SEQ ID NO:178)	
D4S2408	4p15.1	85 - 109	AC110763 (9)	AAGGTACATAACAGTTCAATAGAAAGC(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)	

miniSTR CODIS y NO CODIS					
Locus (Nombre del marcador)	STR del	Ubicación del cromosoma	Intervalo de tamaño (pb)	Acceso de GenBank	Secuencias de cebadores (directo/inverso)
<b>Loci de miniSTR codis*</b>					
D3S4529	3p12.1	111 - 139	AC117452 (13)	CCCAAAATTACTTGAGCCAAT(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(D2S 1776_ R; SEQ ID NO:188)	
D2S441	2p14	78 - 110	AC079112 (12)	CTGTGGCTCATCTATGAAACTT(D2S441_F; SEQ ID NO:189)	
				GAAGTGGCTGTGGTGTATGAT(D2S441_R; SEQ ID NO:190)	
D1S1677	1q23.3	81 - 117	AL513307 (15)	TTCTGTTGGTATAGAGCAGTGTTT(D1S 1677 _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)	
				GTTTTAATTTTCTCCAAATCTCCA(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)					

Los miniSTR proporcionados en la tabla 22 se han utilizado con éxito para determinar la fracción fetal en muestras de ADN en plasma obtenidas de mujeres embarazadas con fetos o masculinos o femeninos, mediante electroforesis capilar (véase

- 5 la tabla 24 en el ejemplo, 15) para identificar y cuantificar los alelos fetales y maternos. Por lo tanto, se espera que las secuencias polimórficas que comprenden otros STR, por ejemplo, los STR restantes de la tabla 22, puedan usarse para determinar la fracción fetal mediante métodos de secuenciación masivamente paralelos.
- 10 La secuenciación de la biblioteca enriquecida en secuencias STR polimórficas se realiza mediante tecnología NGS, por ejemplo, secuenciación masivamente paralela por síntesis. Las lecturas de secuencias de longitudes de al menos 100 pb se alinean con un genoma de referencia, por ejemplo, la secuencia NCBI36/hg18 del genoma de referencia humano, y con un genoma STR, y el número de etiquetas de secuencia mapeadas al genoma humano de referencia y el genoma de referencia STR obtenido para los alelos informativos se utiliza para determinar la presencia o la
- 15 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 proporcionados.

## Ejemplo 12

**Determinación simultánea de aneuploidía y fracción fetal mediante secuenciación masivamente paralela de muestras maternas enriquecidas en secuencias polimórficas que comprenden SNP en tándem**

5 Para determinar simultáneamente la aneuploidía y la fracción fetal en muestras maternas que comprenden ácidos nucleicos fetales y maternos, las muestras de plasma, las muestras de ADNcf purificado y las muestras de la biblioteca de secuenciación se enriquecen en secuencias polimórficas de ácido nucleico diana, cada una de las cuales comprende un par de SNP en tándem seleccionado de entre rs7277033-rs2110153; rs2822654-rs1882882; rs368657-  
 10 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 diana que comprenden los SNP en tándem están diseñados para abarcar ambos sitios SNP. Por ejemplo, el cebador directo está diseñado para abarcar el primer SNP y el cebador inverso está diseñado para abarcar el segundo del par de SNP en tándem, es decir, cada uno de los sitios SNP en el par en tándem está abarcado dentro de los 36 pb generados por el método de secuenciación. La secuenciación de extremos emparejados se puede utilizar para identificar todas las secuencias que abarcan los sitios SNP en tándem. Conjuntos ilustrativos de cebadores que se utilizan para amplificar SNP en tándem divulgados en el presente documento son rs7277033-rs2110153\_F: TCCTGGAACAAAAGTATT (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: TGCTGTGGAATACTAAAAGG (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:GTATAATCCATGAATCTGTTT (SEQ ID NO:203) y rs2822731-rs2822732\_R:TTCAAATTGTATATAAGAGAGT (SEQ ID NO:204), rs1475881-rs1475881\_F:GCAGGAAAGTTATTTTTAAT (SEQ ID NO:205) y rs1475881-rs1475881\_R:TGCTTGAGAAAGCTAACACTT (SEQ ID NO:206), rs1735976-rs1735976\_F:CAGTGTGGAAATTGCTG (SEQ ID NO:207) y rs1735976-rs1735976\_R:GGCACTGGGAGATTATTGTA (SEQ ID NO:208), rs447349-rs2824097\_F:TCCTGTTGTTAAGTACACAT (SEQ ID NO:209) y rs447349-rs2824097\_R:GGGCCGTAATTACTTTTTG (SEQ ID NO:210), rs418989-rs13047336\_F:ACTCAGTAGGCACTTTGTGTC (SEQ ID NO:211) y rs418989-rs13047336\_R:TCTTCCACCACCAATC (SEQ ID NO:212), rs987980-rs987981\_F:TGGCTTTTCAAAGGTAAAA (SEQ ID NO:213) y rs987980-rs987981\_R: GCAACGTTAATCATCTGAATTT (SEQ ID NO:214), rs4143392-rs4143391\_F: rs4143392-rs4143391 (SEQ ID NO:215) y rs4143392-rs4143391\_R:ATTTTATATGTCTATGATCTAAG (SEQ ID NO:216), rs1691324-rs13050434\_F: AGAGATTACAGGTGTGAGC (SEQ ID NO:217) y rs1691324-rs13050434\_R: ATGATCCTCAACTGCCTCT (SEQ ID NO:218), rs11909758-rs9980111\_F: TGAAACTCAAAGAGAAAAAG (SEQ ID NO:219) y rs11909758-rs9980111\_R: ACAGATTTCTACTTAAAATT (SEQ ID NO:220), rs2826842-rs232414\_F: TGAAACTCAAAGAGAAAAAG (SEQ ID NO:221) y rs2826842-rs232414\_R: ACAGATTTCTACTTAAAATT (SEQ ID NO:222), rs2826842-rs232414\_F: GCAAAGGGGTACTCTATGTA (SEQ ID NO:223) y rs2826842-rs232414\_R: TATCGGGTCATCTTGTAAA (SEQ ID NO:224), rs1980969-rs1980970\_F: TCTAACAAAGCTCTGTCCAAAA (SEQ ID NO:225) y rs1980969-rs1980970\_R: CCACACTGAATAACTGGAACA (SEQ ID NO:226), rs9978999-rs9979175\_F: GCAAGCAAGCTCTACCTTC (SEQ ID NO:227) y rs9978999-rs9979175\_R: TGTTCTTCCAAAATTCACATGC (SEQ ID NO:228), rs1034346-rs12481852\_F: ATTTCACTATTCCTTCATTTT (SEQ ID NO:229) y rs1034346-rs12481852\_R: TAATTGTTGCACACTAAATTAC (SEQ ID NO:230), rs4817013-rs7277036\_F: AAAAAGCCACAGAAATCAGTC (SEQ ID NO:231) y rs4817013-rs7277036\_R: TTCTTATATCTCACTGGGCATT (SEQ ID NO:232), rs9981121-rs2829696\_F: GGATGGTAGAAGAGAAGAAAGG (SEQ ID NO:233) y rs9981121-rs2829696\_R: GGATGGTAGAAGAGAAGAAAGG (SEQ ID NO:234), rs455921-rs2898102\_F: TGCAAAGATGCAGAACCAAC (SEQ ID NO:235) y rs455921-rs2898102\_R: TTTTGTTCCTTGTCTGGCTGA (SEQ ID NO:236), rs2898102-rs458848\_F: TGCAAAGATGCAGAACCAAC (SEQ ID NO:237) y rs2898102-rs458848\_R: GCCTCCAGCTCTATCCAAGTT (SEQ ID NO:238), rs961301-rs2830208\_F: CCTTAATATCTTCCCATGTCCA (SEQ ID NO:239) y rs961301-rs2830208\_R: ATTGTTAGTGCCCTTCTGCTT (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: CAACTTGTCTTTTATTGATGT (SEQ ID NO:244), rs1011734-rs1011733\_F: CTATGTTGATAAAACATTGAAA (SEQ ID NO:245) y rs1011734-rs1011733\_R: GCCTGCTGGAATATAGTTT (SEQ ID NO:246), rs2831244-rs9789838\_F: CAGGGCATAATACTAAGCTGT (SEQ ID NO:247) y rs2831244-rs9789838\_R: CAATGACTCTGAGTTGAGCAC (SEQ ID NO:248), rs8132769-rs2831440\_F: ACTCTCTCCCTCCCTCT (SEQ ID NO:249) y rs8132769-rs2831440\_R: TATGGCCCCAAAATCTTCT (SEQ ID NO:250), rs8134080-rs2831524\_F:

ACAAGTACTGGGCAGATTGA (SEQ ID NO:251) y rs8134080-rs2831524\_R: GCCAGGTTTAGCTTTCAAGT (SEQ ID NO:252), rs4817219-rs4817220\_F: TTTTATATCAGGAGAAACACTG (SEQ ID NO:253) y rs4817219-rs4817220\_R: CCAGAATTTTGGAGGTTTAAT (SEQ ID NO:254), rs2250911-rs2250997\_F: TGTCACTCCTCCTTTATCTCCA (SEQ ID NO:255) y rs2250911-rs2250997\_R: TTCTTTTGCCTCTCCCAAAG (SEQ ID NO:256), rs2831899-rs2831900\_F: ACCCTGGCACAGTGTGACT (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: TTTTCCCATTTCCAACCTCT (SEQ ID NO:260), rs11088086-rs2251447\_F: AAAAGATGAGACAGGCAGT (SEQ ID NO:261) y rs11088086-rs2251447\_R: ACCCCTGTGAATCTCAAAT (SEQ ID NO:262), rs2832040-rs11088088\_F: GCACTTGCTTCTATTGTTTGT (SEQ ID NO:263) y rs2832040-rs11088088\_R: CCCTTCTCTCTTCCATTCT (SEQ ID NO:264), rs2832141-rs2246777\_F: AGCACTGCAGGTA (SEQ ID NO:265) y rs2832141-rs2246777\_R: ACAGATACCAAAGAACTGCAA (SEQ ID NO:266), rs2832959-rs9980934\_F: TGGACACCTTTCAACTTAGA (SEQ ID NO:267) y rs2832959-rs9980934\_R: GAACAGTAATGTTGAACTTTTT (SEQ ID NO:268), rs2833734-rs2833735\_F: TCTTGCAAAAAGCTTAGCACA (SEQ ID NO:269) y rs2833734-rs2833735\_R: AAAAAGATCTCAAAGGGTCCA (SEQ ID NO:270), rs933121-rs933122\_F: GCTTTTGTGAACATCAAGT (SEQ ID NO:271) y rs933121-rs933122\_R: CCTTCCAGCAGCATAGTCT (SEQ ID NO:272), rs2834140-rs12626953\_F: AAATCCAGGATGTGCAGT (SEQ ID NO:273) y rs2834140-rs12626953\_R: ATGATGAGGTCAGTGGTGT (SEQ ID NO:274), rs2834485-rs3453\_F: CATCACAGATCATAGTAAATGG (SEQ ID NO:275) y rs2834485-rs3453\_R: AATTATTATTTTGCAGGCAAT (SEQ ID NO:276), rs9974986-rs2834703\_F: CATGAGGCAAACACCTTCC (SEQ ID NO:277) y rs9974986-rs2834703\_R: GCTGGACTCAGGATAAAGAACA (SEQ ID NO:278), rs2776266-rs2835001\_F: TGGAAAGCCTGAGCTGACTAA (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: GAATGACCTTGGCACTTTTATCA (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: CCCAGGAAGAGTGGAAAGATT (SEQ ID NO:291) y rs2836550-rs2212596\_R: TTAGCTTGCATGTACCTGTGT (SEQ ID NO:292), rs2836660-rs2836661\_F: AGCTAGATGGGGTGAATTTT (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: AATGAATAAGGTCTCAGAG (SEQ ID NO:296), rs9980072-rs8130031\_F: TTTAATCTGATCATTGCCCTA (SEQ ID NO:297) y rs9980072-rs8130031\_R: AGCTGTGGGTGACCTTGA (SEQ ID NO:298), rs418359-rs2836926\_F: TGTCCCACCATTTGTGTATTA (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: CCATCATGGAAGCATGG (SEQ ID NO:303) y rs385787-rs367001\_R: TCATCTCCATGACTGCACTA (SEQ ID NO:304), rs367001-rs386095\_F: GAGATGACGGAGTAGCTCAT (SEQ ID NO:305) y rs367001-rs386095\_R: CCCAGCTGCACTGTCTAC (SEQ ID NO:306), rs2837296-rs2837297\_F: TCTTGTTCGAATCACAGGAC (SEQ ID NO:307) y rs2837296-rs2837297\_R: 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 se puede incluir en 5' del cebador directo o en 3' del cebador inverso es GCCGCCTGCAGCCCGCGCCCCCGTGCCTCCGCCCCGCGCCGCGCCGGGCGCC (SEQ ID NO:311).

La preparación de muestras y el enriquecimiento de la biblioteca de secuenciación de ADNcf, una muestra de ADNcf purificada y una muestra de plasma se realizan según el método descrito en los ejemplos 8, 9 y 10, respectivamente. Todas las bibliotecas de secuenciación se preparan tal como se describe en el ejemplo, 2a y la secuenciación se realiza tal como se describe en el ejemplo, 2b e incluye la secuenciación de extremos emparejados. El análisis de los datos de secuenciación para la determinación de la aneuploidía fetal se realiza tal como se describe en los ejemplos 4 y 5. De forma concomitante al análisis para determinar la aneuploidía, se analizaron los datos de secuenciación para determinar la fracción fetal de la forma siguiente. Después de la transferencia de los archivos de imagen y de designación de bases al servidor Unix que ejecuta el programa informático de Illumina "Genome Analyzer Pipeline" versión 1.51 tal como se describe, las lecturas de 36 pb se alinearon con un "genoma SNP en tándem" utilizando 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 divulgados anteriormente. Para el análisis de la fracción fetal, solo se usaron las lecturas que se mapean únicamente al genoma del SNP en tándem. Las lecturas que coinciden perfectamente con el genoma SNP se cuentan como etiquetas y se filtran. De las lecturas restantes, solo las lecturas que tenían uno o dos desajustes se cuentan como etiquetas y se incluyen en el análisis. Se cuentan las etiquetas mapeadas a cada uno de los alelos de SNP en tándem, y la fracción fetal se determina esencialmente tal como se ha descrito en el ejemplo, 6 anterior, pero teniendo en cuenta las etiquetas mapeadas a los dos alelos x e y de SNP en tándem presentes en cada una de las secuencias polimórficas de ácido nucleico diana amplificadas, que se amplifican para enriquecer las muestras, es decir,

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

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<sub>x+y</sub>) de la forma siguiente:

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

para compensar la presencia de 2 conjuntos de alelos fetales en tándem, estando uno enmascarado por el fondo materno. Las secuencias de 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 de la madre y el haplotipo heredado del padre para calcular la fracción fetal mediante 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 de alelos en tándem. En un caso, la fracción fetal es la fracción fetal promedio determinada para al menos 3 conjuntos de alelos en tándem informativos.

### Ejemplo 13

#### Determinación de la fracción fetal mediante secuenciación masivamente paralela de una biblioteca diana que comprende ácidos nucleicos polimórficos que comprenden SNP

Para determinar la fracción de ADNcf fetal en una muestra materna, se amplificaron las secuencias de ácido nucleico polimórfico diana, cada una de las cuales comprende un SNP, y se usaron para preparar una biblioteca diana para la secuenciación de forma masivamente paralela.

Se extrajo ADNcf tal como se describe en el ejemplo, 1. Se preparó una biblioteca de secuenciación diana de la forma siguiente: se amplificó ADNcf contenido en 5 µl de ADNcf purificado en un volumen de reacción de 50 µl que contenía 7,5 µl de una mezcla de cebadores 1 µM (tabla 10), 10 µl de mezcla maestra NEB 5X y 27 µl de agua. Se realizó un termociclado con Gene Amp9700 (Applied Biosystems) utilizando las condiciones de ciclado siguientes: incubación a 95°C durante 1 minuto, seguida de 20-30 ciclos a 95°C durante 20 segundos, 68°C durante 1 minuto y 68°C durante 30 s, seguidos de una incubación final a 68°C durante 5 minutos. Se añadió un mantenimiento final a 4°C hasta que se retiraron las muestras para combinarlas con la porción no amplificada de la muestra de ADNcf purificado. El producto amplificado se purificó con el sistema de purificación por PCR Agencourt AMPure XP (Nº de parte A63881; Beckman Coulter Genomics, Danvers, MA) y la concentración se cuantificó usando Nanodrop 2000 (Thermo Scientific, Wilmington, DE). Se añadió un mantenimiento final a 4°C hasta que se retiraron las muestras 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 usando el protocolo abreviado descrito en el ejemplo, 2, y se secuenció de forma masivamente paralela usando secuenciación por síntesis con terminadores de colorantes reversibles y según el protocolo de Illumina. El análisis y el recuento de etiquetas mapeadas a un genoma de referencia que consiste en 26 secuencias (13 pares, cada una de las cuales representa dos alelos) que comprenden un SNP, es decir, SEQ ID NO: 1-56 se realizó tal como se describe.

La tabla 23 proporciona los recuentos de etiquetas obtenidos 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 etiquetas de SNP	Fracción fetal (%)
rs10773760.1 Cr.12 longitud=128 alelo=A	236590	1,98
rs10773760.2 Cr.12 longitud=128 alelo=G	4680	
rs13182883.1 Cr.5 longitud=111 alelo=A	3607	4,99
rs13182883.2 Cr.5 longitud=111 alelo=G	72347	
rs4530059.1 Cr.14 longitud=110 alelo=A	3698	1,54
rs4530059.1 Cr.14 longitud=110 alelo=G	239801	

<b>Determinación de la fracción fetal mediante secuenciación masivamente paralela de una biblioteca de ácidos nucleicos polimórficos que comprenden SNP</b>		
<b>SNP</b>	<b>RECuentos de etiquetas de SNP</b>	<b>Fracción fetal (%)</b>
rs8078417.1 Cr. 17 longitud= 110 alelo=C	1E+06	3,66
rs8078417.2 Cr.17 longitud= 110 alelo=T	50565	
<b>Fracción fetal (Media±S.D.) = 12,4± 6,6</b>		

Los resultados muestran que las secuencias de ácidos nucleicos polimórficos, cada una de las cuales comprende al menos un SNP, se pueden amplificar a partir de ADNcf derivado de una muestra de plasma materno para construir una biblioteca que se puede secuenciar de forma masivamente paralela para determinar la fracción de ácidos nucleicos fetales en la muestra materna. Los métodos de secuenciación masivamente paralela para determinar la fracción fetal pueden usarse en combinación con otros métodos para proporcionar diagnóstico de aneuploidía fetal y otros ensayos prenatales.

#### **Ejemplo 14**

#### **Determinación de la fracción fetal mediante secuenciación masivamente paralela de una biblioteca diana que comprende ácidos nucleicos polimórficos que comprenden STR o SNP en tándem**

La fracción fetal se puede determinar independientemente de la determinación de la aneuploidía utilizando una biblioteca diana que comprende SNP en tándem o STR tal como se describe para la biblioteca diana de SNP del ejemplo, 13. Para preparar una biblioteca diana de SNP en tándem, se utiliza una porción de una biblioteca de ADNcf purificado que comprende ácidos nucleicos fetales y maternos para amplificar secuencias diana utilizando una mezcla de cebadores, por ejemplo, las tablas 10 y 11. Para preparar una biblioteca diana de STR, se utiliza una porción de una biblioteca de ADNcf purificado que comprende ácidos nucleicos fetales y maternos para amplificar secuencias diana utilizando una mezcla de cebadores, por ejemplo, la tabla 22. La biblioteca diana de SNP en tándem se secuencia tal como se describe en el ejemplo, 12.

Las bibliotecas diana se secuencian tal como se describe, y la fracción fetal se determina a partir del número de etiquetas de secuencia mapeadas al genoma de referencia de STR o SNP en tándem, respectivamente, que comprende todos los alelos posibles de STR o SNP en tándem abarcados por los cebadores. Se identifican los alelos informativos y se determina la fracción fetal 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**

Para determinar la fracción fetal en muestras maternas que comprendían ADNcf fetal y materno, se recolectaron muestras de sangre periférica de mujeres embarazadas voluntarias que portaban fetos o masculinos o femeninos. Se obtuvieron muestras de sangre periférica y se procesaron para proporcionar ADNcf purificado tal como se describe en el ejemplo, 1.

Se analizaron diez microlitros de muestras de ADNcf usando el kit de amplificación por PCR AmpFISTR® MiniFiler™ (Applied Biosystems, Foster City, CA) según las instrucciones del fabricante. Brevemente, el ADNcf contenido en 10 µl se amplificó en un volumen de reacción de 25 µl que contenía 5 µl de cebadores marcados con fluorescencia (conjunto de cebadores AmpF/STR® MiniFiler™), y la mezcla maestra AmpF/STR® MiniFiler™, que incluye ADN polimerasa AmpliTaq Gold® y tampón asociado, sal (MgCl<sub>2</sub> 1,5 mM) y trifosfatos de desoxinucleótidos 200 µM (dNTP: dATP, dCTP, dGTP y dTTP). Los cebadores marcados con fluorescencia son cebadores directos que están marcados con colorantes 6FAM™, VIC™, NED™ y PET™. Se realizó un termociclado con Gene Amp9700 (Applied Biosystems) utilizando las condiciones de ciclado siguientes: incubación a 95°C durante 10 minutos, seguida de 30 ciclos a 94°C durante 20 segundos, 59°C durante 2 minutos y 72°C durante 1 minuto, seguidos de una incubación final a 60°C durante 45 minutos. Se añadió un mantenimiento final a 4°C hasta que se retiraron las muestras para el análisis. El producto amplificado se preparó diluyendo 1 ul de producto amplificado en 8,7 ul de formamida Hi-Di™ (Applied Biosystems) y 0,3 µl de patrón de tamaño interno GeneScan™-500 LIZ\_ (Applied Biosystems), y se analizó con un analizador genético ABI PRISM3130xl (Applied Biosystems) utilizando recopilación de datos HID\_G5\_POP4 (Applied Biosystems) y matriz de capilares de 36 cm. Todo el genotipado se realizó con el programa informático GeneMapper\_ID v3.2 (Applied Biosystems) utilizando escaleras alélicas e intervalos y paneles proporcionados por el fabricante.

Todas las mediciones de genotipado se realizaron en el analizador genético Applied Biosystems 3130xl, usando una "ventana" de ±0,5 nt alrededor del tamaño obtenido para cada alelo para permitir la detección y la asignación correcta

de alelos. Se determinó que cualquier alelo de la muestra cuyo tamaño estuviera fuera de la ventana de  $\pm 0,5$  nt era OL, es decir, "Off Ladder". Los alelos OL son alelos de un tamaño que no está representada 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 mínimo de altura del pico de  $> 50$  RFU se estableció en base a los experimentos de validación realizados para evitar el tipado 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 intervalos preestablecidos para los STR que se analizan.

Los cálculos de la fracción fetal se realizaron utilizando la altura máxima promedio para los alelos principales y secundarios en cada locus STR determinado a partir de inyecciones por triplicado. Las reglas que se aplican al cálculo son:

1. los datos de alelos fuera de escalera (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 relativas de fluorescencia)
3. si solo hay un intervalo, el marcador se considera no informativo; y
4. si se designa un segundo intervalo, pero los picos del primer y segundo intervalo están dentro del 50-70% de sus unidades relativas de fluorescencia (RFU) en la altura del pico, la fracción minoritaria no se mide y el marcador no se considera informativo.

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

$$\text{fracción fetal} = \left( \frac{\sum \text{altura de pico de alelo secundario}}{\sum \text{altura de pico de alelo(s) principal(es)}} \right) \times 100,$$

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

La tabla 8 proporciona los datos obtenidos del análisis de ADNcf de una mujer embarazada de un feto masculino.

**Tabla 24**

Fracción fetal determinada en ADNcf de un sujeto embarazado por análisis de STR										
STR	Alelo 1	Alelo 2	Alelo 3	Alelo 1 Altura	Alelo 2 Altura	Alelo 3 Altura	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					

Fracción fetal determinada en ADNcf de un sujeto embarazado por análisis de STR										
STR	Alelo 1	Alelo 2	Alelo 3	Alelo 1 Altura	Alelo 2 Altura	Alelo 3 Altura	Fracción fetal	Fracción fetal (Media/STR)		
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		
<b>Fracción fetal = 3,5</b>										

Los resultados muestran que se puede utilizar ADNcf para determinar la presencia o la ausencia de ADN fetal según se indica por la detección de un componente secundario en uno o más alelos STR, para determinar el porcentaje de fracción fetal y para determinar el sexo del feto tal como se indica por la presencia o la ausencia del alelo amelogenina.

5

**Ejemplo 16**

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

10

Los recuentos de secuencias coincidentes con diferentes cromosomas se manipulan para generar una puntuación que variará con el número de copias cromosómicas que puede interpretarse para identificar la amplificación o la delección cromosómica. Por ejemplo, dicha puntuación podría generarse comparando la cantidad relativa de etiquetas de secuencia en un cromosoma que sufre cambios en el número de copias con un cromosoma que se sabe que es un euploide. Los ejemplos de puntuaciones que se pueden usar para identificar la amplificación o la delección incluyen, pero sin limitación: recuentos del cromosoma de interés divididos por los recuentos de otro cromosoma de la misma ejecución experimental, recuentos del cromosoma de interés divididos por el número total de recuentos de la serie experimental, comparación de recuentos de la muestra de interés frente a 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 utilizar para establecer umbrales de "corte" para designar los estados "aneuploidía", "normal" o "marginal" (incierto). Después se realizan cálculos para estimar el número mínimo de secuencias necesarias para lograr la sensibilidad adecuada (es decir, la probabilidad de identificar correctamente un estado de aneuploidía).

15

20

25

La figura 19 es un gráfico 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 univariante donde la condición de aneuploidía representa una amplificación del número de copias. Los casos multivariante y/o las anomalías de reducción/delección son simples extensiones o reordenamientos de las descripciones dadas y se pretende que se encuentren dentro del alcance de esta técnica.

30

35

La cantidad de "superposición" entre las poblaciones puede determinar cómo de bien se pueden discriminar los casos normales y los de aneuploidía. En general, el aumento de la fracción fetal, ff, aumenta la potencia 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, teniendo cuatro copias del cromosoma en lugar de una trisomía) de la amplificación también aumentará la separación de los centros de población, lo que conducirá a una mayor potencia (es decir, una mayor probabilidad de identificar correctamente los estados de aneuploidía).

40

El aumento del 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.

**Establecimiento de umbrales y estimación del tamaño de la muestra**

5 El siguiente procedimiento se puede utilizar para establecer "c", el valor crítico para designar los estados "aneuploidía", "normal" o "marginal" (incierto). Sin pérdida de generalidad, a continuación, se utilizan pruebas estadísticas unilaterales.

10 En primer lugar, se decide una tasa aceptable de falsos positivos, FP (a veces también denominada "error de tipo I" o "especificidad"), que es la probabilidad de un falso positivo o falsa designación 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, 0,008, 0,009, 0,01, 0,02, 0,03, 0,04, 0,05, 0,06, 0,07, 0,08, 0,09 o 0,1.

15 En segundo lugar, el valor de "c" se puede determinar resolviendo la ecuación:  $FP = \int_c^{\infty} f1(x)dx$ .

$$FP = \int_c^{\infty} f1(x)dx \quad \text{(Ecuación 1)}$$

20 Una vez que se ha determinado un valor crítico, c, se puede estimar el número mínimo de secuencias requeridas para lograr una determinado TP = tasa de verdaderos positivos. La tasa de verdaderos positivos puede ser, por ejemplo, de aproximadamente 0,5, 0,6, 0,7, 0,8 o 0,9. En un caso, la tasa de verdaderos positivos puede ser 0,8. En otras palabras, N es el número mínimo de secuencias necesarias para identificar la aneuploidía 100\*TP por ciento del tiempo. N = número mínimo tal que  $TP = \int_c^{\infty} f2(x,ff)dx > 0,8$ . N se determina resolviendo

$$\min_N s.t. \{TB \geq \int_c^{\infty} f2(x, N)dx\} \quad \text{(Ecuación 2)}$$

25 En las pruebas estadísticas clásicas, f1 y f2 son a menudo F, distribuciones F no centrales (un caso especial de distribuciones t y t no centrales) aunque esa no es una condición necesaria para esta solicitud.

**Establecimiento de "niveles" de umbrales para proporcionar más control de errores**

Los umbrales también se pueden establecer en etapas utilizando los métodos anteriores. Por ejemplo, se puede establecer un umbral para la designación 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 la puntuación, S:

35 (S > ca) entonces designar "Trisomía"

40 (cb > S <= ca) entonces designar "Marginal"

45 (S < cb) entonces designar "Normal"

**Algunas generalizaciones triviales que se encuentran dentro del alcance de esta técnica**

Se pueden utilizar diferentes valores para umbrales tales como TP, FP, etc. Los procedimientos se pueden ejecutar en cualquier orden. Por ejemplo, se puede comenzar con N y resolver c, etc. Las distribuciones pueden depender de ff de modo que f1(x,N,ff), f2(x,N,ff) y/u otras variables. Las ecuaciones integrales anteriores se pueden resolver con referencia a tablas o mediante métodos informáticos iterativos. Se puede estimar un parámetro de no centralidad y se puede leer la potencia a partir de tablas estadísticas estándar. La potencia estadística y los tamaños de muestra pueden derivarse del cálculo o la estimación de las medias cuadráticas esperadas. Se pueden utilizar distribuciones teóricas de forma cerrada tales como f, t, no central, normal, etc. o estimaciones (kernel u otras) para modelar las distribuciones f1, f2. El establecimiento de un umbral empírico y la selección de parámetros mediante curvas características del operador del receptor (ROC) se pueden utilizar y cotejar con la fracción fetal. Se pueden utilizar varias estimaciones de la dispersión de la distribución (varianza, desviación absoluta media, rango intercuartílico, etc.). Se pueden utilizar varias estimaciones del centro de distribución (media, mediana, etc.). Se pueden utilizar pruebas estadísticas bilaterales en lugar de unilaterales. La prueba de hipótesis simple se puede reformular como regresión

lineal o no lineal. Se pueden utilizar métodos combinatorios, simulación (por ejemplo, monte carlo), maximización (por ejemplo, maximización esperada), iterativos u otros métodos de forma independiente o junto con los anteriores para establecer la potencia estadística o los umbrales.

5 **Ejemplo 17**

**Demostración de detección de aneuploidía**

10 Los datos de secuenciación obtenidos para las muestras descritas en los ejemplos 4 y 5, y que se muestran 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).

15 La figura 20 (A) muestra la distribución de dosis de cromosomas relativas a la desviación estándar de la media de la dosis del cromosoma 21 en las muestras no afectadas (normales) (o) y las muestras con trisomía 21 (T21; Δ) cuando se utiliza el cromosoma 14 como cromosoma de normalización para el cromosoma 21. La figura 20 (B) muestra la distribución de dosis de cromosomas relativas a la desviación estándar de la media de la dosis del cromosoma 18 en las muestras no afectada (o) y las muestras con trisomía 18 (T18; Δ) cuando se utiliza el cromosoma 8 como cromosoma de normalización para el cromosoma 18. La figura 20 (C) muestra la distribución de las dosis de cromosomas relativas a la desviación estándar de la media de la dosis del cromosoma 13 en las muestras no afectadas (o) y las muestras con trisomía 13 (T13; Δ) usando la densidad de etiqueta de secuencia promedio del grupo de cromosomas 3, 4, 5 y 6 como el cromosoma de normalización para determinar la dosis de cromosoma para el cromosoma 13. La figura 20 (D) muestra la distribución de las dosis de cromosomas relativas a la desviación estándar de la media de 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 de normalización para el cromosoma X. La figura 20 (E) muestra la distribución de las dosis de cromosomas relativas a 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 con monosomía X (+), cuando se utiliza la densidad de etiqueta de secuencia promedio del grupo de cromosomas 1-22 y X como el cromosoma de normalización para determinar la dosis de cromosoma para el cromosoma Y.

Los datos muestran que la trisomía 21, la trisomía 18 y la trisomía 13 se distinguían claramente de las muestras no afectadas (normales). Las muestras con monosomía X fueron fácilmente identificables por tener dosis de cromosoma X que era claramente inferior a la de las muestras femeninas no afectadas (figura 20 (D)), y por tener dosis de cromosoma Y que eran claramente inferiores a las de las muestras masculinas no afectadas (figura 20 (E)).

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

LISTA DE SECUENCIAS

45 <110> VERINATA HEALTH, INC.  
 <120> Nuevo protocolo para preparar bibliotecas de secuenciación  
 <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  
 50 <150> 61/296,358 <151> 2010-01-19  
 <160> 427  
 <170> PatentIn versión 3.5  
 <210> 1  
 <211> 111  
 55 <212> ADN  
 <213> Homo sapiens  
 <400> 1  
 cacatgcaca gccagcaacc ctgtcagcag gagttccac cagtttcttt ctgagaacat 60  
 ctgttcaggt ttctctccat ctctatttac tcaggtcaca ggaccttggg g 111  
 <210> 2  
 60 <211> 111

ES 2 909 841 T3

<212> ADN  
 <213> Homo sapiens  
 <400> 2  
 cacatgcaca gccagcaacc ctgtcagcag gagttcccac cagtttcttt ctgagaacat 60  
  
 ctgttcaggt ttctctccat ctctgtttac tcaggtcaca ggaccttggg g 111  
 5 <210> 3  
 <211> 126  
 <212> ADN  
 <213> Homo sapiens  
 <400> 3  
 tgaggaagtg aggctcagag ggtaagaaac tttgtcacag agctgggtgt gaggggtggag 60  
  
 attttacct ccctgctcc cacaccagtt tctccagagt ggaaagactt tcatctcgca 120  
  
 ctggca 126  
 10 <210> 4  
 <211> 126  
 <212> ADN  
 <213> Homo sapiens  
 15 <400> 4  
 tgaggaagtg aggctcagag ggtaagaaac tttgtcacag agctgggtgt gaggggtggag 60  
  
 attttacct ccctgctcc cacaccagtt tctccgaggt ggaaagactt tcatctcgca 120  
  
 ctggca 126  
 <210> 5  
 <211> 121  
 <212> ADN  
 <213> Homo sapiens  
 20 <400> 5  
 gtgccttcag aacctttgag atctgattct atttttaag cttcttagaa gagagattgc 60  
  
 aaagtgggtt gtttctctag ccagacaggg caggcaata ggggtggctg gtgggatggg 120  
  
 a 121  
 <210> 6  
 <211> 121  
 <212> ADN  
 <213> Homo sapiens  
 25 <400> 6  
 gtgccttcag aacctttgag atctgattct atttttaag cttcttagaa gagagattgc 60  
  
 aaagtgggtt gtttctctag ccagacaggg caggtaata ggggtggctg gtgggatggg 120  
  
 a 121  
 <210> 7  
 <211> 111  
 <212> ADN  
 <213> Homo sapiens  
 30 <400> 7  
 aggtgtgtct ctcttttgtg aggggagggg tcccttctgg cctagtagag ggcctggcct 60  
  
 gcagtgagca ttcaaactct caaggaacag ggtggggagg tgggacaaag g 111  
 35 <210> 8  
 <211> 111  
 <212> ADN  
 <213> Homo sapiens  
 <400> 8  
 aggtgtgtct ctcttttgtg aggggagggg tcccttctgg cctagtagag ggcctggcct 60  
  
 gcagtgagca ttcaaactct cgaggaacag ggtggggagg tgggacaaag g 111  
 40 <210> 9

ES 2 909 841 T3

	<211> 139		
	<212> ADN		
	<213> Homo sapiens		
	<400> 9		
	cctcgcctac tgtgctggtt ctaaccatca tgcttttccc tgaatctctt gagtcttttt	60	
	ctgctgtgga ctgaaacttg atcctgagat tcacctctag tccctctgag cagcctcctg	120	
5	gaatactcag ctgggatgg	139	
	<210> 10		
	<211> 139		
	<212> ADN		
	<213> Homo sapiens		
10	<400> 10		
	cctcgcctac tgtgctggtt ctaaccatca tgcttttccc tgaatctctt gagtcttttt	60	
	ctgctgtgga ctgaaacttg atcctgagat tcacctctag tccctctggg cagcctcctg	120	
	gaatactcag ctgggatgg	139	
	<210> 11		
	<211> 117		
	<212> ADN		
	<213> Homo sapiens		
15	<400> 11		
	aattgcaatg gtgagagggt gatggtaaaa tcaaacggaa cttgttattt tgtcattctg	60	
	atggactgga actgaggatt ttcaatttcc tctccaacc aagacacttc tcaactgg	117	
	<210> 12		
	<211> 117		
	<212> ADN		
	<213> Homo sapiens		
20	<400> 12		
	aattgcaatg gtgagagggt gatggtaaaa tcaaacggaa cttgttattt tgtcattctg	60	
	atggactgga actgaggatt ttcaatttcc tttccaacc aagacacttc tcaactgg	117	
	<210> 13		
	<211> 114		
	<212> ADN		
	<213> Homo sapiens		
25	<400> 13		
	gaaatgcctt ctcaggtaat ggaaggttat ccaaataattt ttcgtaagta tttcaaatag	60	
	caatggctcg tctatgggta gtctcacagc cacattctca gaactgctca aacc	114	
	<210> 14		
	<211> 114		
	<212> ADN		
	<213> Homo sapiens		
30	<400> 14		
	gaaatgcctt ctcaggtaat ggaaggttat ccaaataattt ttcgtaagta tttcaaatag	60	
	caatggctcg tctatgggta gtctcgagc cacattctca gaactgctca aacc	114	
	<210> 15		
	<211> 128		
	<212> ADN		
	<213> Homo sapiens		
40	<400> 15		
	accctaaaaca ctggaggggc ctcttctcat tttcggtaga ctgcaagtgt tagccgctcg	60	
	gaccagcttc tgtctggaag ttogtcaaat tgcagttaag tccaagtatg ccacatagca	120	
	gataaggg	128	
	<210> 16		

ES 2 909 841 T3

<211> 128  
 <212> ADN  
 <213> Homo sapiens  
 <400> 16  
 acccaaaaca ctggaggggc ctcttctcat tttcggtaga ctgcaagtgt tagccgtcgg 60  
 gaccagcttc tgtctggaag ttcgtcaaat tgcagttagg tccaagtatg ccacatagca 120  
 5 gataaggg 128  
 <210> 17  
 <211> 110  
 <212> ADN  
 <213> Homo sapiens  
 10 <400> 17  
 gcaccagaat ttaaacaacg ctgacaataa atatgcagtc gatgatgact tcccagagct 60  
 ccagaagcaa ctccagcaca cagagaggcg ctgatgtgcc tgtcaggtgc 110  
 <210> 18  
 <211> 110  
 <212> ADN  
 15 <213> Homo sapiens  
 <400> 18  
 gcaccagaat ttaaacaacg ctgacaataa atatgcagtc gatgatgact tcccagagct 60  
 ccagaagcaa ctccagcaca cggagaggcg ctgatgtgcc tgtcaggtgc 110  
 <210> 19  
 <211> 116  
 <212> ADN  
 20 <213> Homo sapiens  
 <400> 19  
 tgactgtata cccaggtgc acccttgggt catctctatc atagaactta tctcacagag 60  
 tataagagct gatttctgtg tctgcctctc aactagact tccacatcct tagtgc 116  
 <210> 20  
 <211> 116  
 <212> ADN  
 25 <213> Homo sapiens  
 <400> 20  
 tgactgtata cccaggtgc acccttgggt catctctatc atagaactta tctcacagag 60  
 tataagagct gatttctgtg tctgcctctc aactagact tccacatcct tagtgc 116  
 <210> 21  
 <211> 110  
 <212> ADN  
 <213> Homo sapiens  
 30 <400> 21  
 tgtactggt caccagggga cgcctggcgc tgcgaggag gccccgagcc tcgtgcccc 60  
 gtgaagcttc agctcccctc cccggctgtc cttgaggctc ttctcacact 110  
 <210> 22  
 <211> 110  
 <212> ADN  
 <213> Homo sapiens  
 40 <400> 22  
 tgtactggt caccagggga cgcctggcgc tgcgaggag gccccgagcc tcgtgcccc 60  
 gtgaagcttc agctcccctc cctggctgtc cttgaggctc ttctcacact 110  
 <210> 23  
 <211> 114  
 <212> ADN  
 45 <213> Homo sapiens

ES 2 909 841 T3

<400> 23  
 cagtggacc tgetgcacct ttctcccct cccatcaacc tcttttgtgc ctccccctcc 60  
  
 gtgtaccacc ttctctgtca ccaaccctgg cctcacaact ctctcctttg ccac 114  
 <210> 24  
 <211> 114  
 <212> ADN  
 <213> Homo sapiens  
 <400> 24  
 cagtggacc tgetgcacct ttctcccct cccatcaacc tcttttgtgc ctccccctcc 60  
  
 gtgtaccacc ttctctgtca ccaccctgg cctcacaact ctctcctttg ccac 114  
 <210> 25  
 <211> 110  
 <212> ADN  
 <213> Homo sapiens  
 <400> 25  
 cagtggcata gtagtccagg ggctcctcct cagcacctcc agcaccttcc aggaggcagc 60  
  
 agcgcaggca gagaaccgc tggaagaatc ggcggaagtt gtcggagagg 110  
 <210> 26  
 <211> 110  
 <212> ADN  
 <213> Homo sapiens  
 <400> 26  
 cagtggcata gtagtccagg ggctcctcct cagcacctcc agcaccttcc aggaggcagc 60  
  
 agcgcaggca gagaaccgc tggaaggatc ggcggaagtt gtcggagagg 110  
 <210> 27  
 <211> 129  
 <212> ADN  
 <213> Homo sapiens  
 <400> 27  
 aggtctgggg gccgctgaat gccaaactgg gaatcttaa tgtaaggaa caaggtcata 60  
  
 caatgaatgg tgtgatgtaa aagcttggga ggtgatttct gagggtaggt gctgggttta 120  
  
 atgggagga 129  
 <210> 28  
 <211> 129  
 <212> ADN  
 <213> Homo sapiens  
 <400> 28  
 aggtctgggg gccgctgaat gccaaactgg gaatcttaa tgtaaggaa caaggtcata 60  
  
 caatgaatgg tgtgatgtaa aagcttggga ggtgattttt gagggtaggt gctgggttta 120  
  
 atgggagga 129  
 <210> 29  
 <211> 107  
 <212> ADN  
 <213> Homo sapiens  
 <400> 29  
 acggttctgt cctgtagggg agaaaagtcc tcgttgttcc tctgggatgc aacatgagag 60  
  
 agcagcacac taggcttta tggattgcc tgccacaagt gaacagg 107  
 <210> 30  
 <211> 107  
 <212> ADN  
 <213> Homo sapiens  
 <400> 30

ES 2 909 841 T3

acggttctgt cctgtagggg agaaaagtcc tcgttgttcc tctgggatgc aacatgagag 60  
 agcagcacac tgaggcttta tggggtgccc tgccacaagt gaacagg 107  
 <210> 31  
 <211> 127  
 <212> ADN  
 5 <213> Homo sapiens  
 <400> 31  
 gcgcagtcag atggggcgtgc tggcgtctgt cttctctctc tcctgctctc tggcttcatt 60  
 tttctctctc tctgtctcac cttctttcgt gtgcctgtgc acacacacgt ttgggacaag 120  
 ggctgga 127  
 <210> 32  
 <211> 127  
 10 <212> ADN  
 <213> Homo sapiens  
 <400> 32  
 gcgcagtcag atggggcgtgc tggcgtctgt cttctctctc tcctgctctc tggcttcatt 60  
 tttctctctc tctgtctcac cttctttcgt gtgcctgtgc atacacacgt ttgggacaag 120  
 ggctgga 127  
 <210> 33  
 15 <211> 130  
 <212> ADN  
 <213> Homo sapiens  
 <400> 33  
 gcccgacctg cgaaatccca aaatgccaaa cattcccgcc tcacatgatc ccagagagag 60  
 gggaccacgt gttcccagct tgcagctgag gagcccagag ttgccgtcag atcagagccc 120  
 cagttgcccg 130  
 20 <210> 34  
 <211> 130  
 <212> ADN  
 <213> Homo sapiens  
 <400> 34  
 gcccgacctg cgaaatccca aaatgccaaa cattcccgcc tcacatgatc ccagagagag 60  
 gggaccacgt gttcccagct tgcagctgag gagcccagag ttgccgtcag atcagagccc 120  
 cagttgcccg 130  
 25 <210> 35  
 <211> 121  
 <212> ADN  
 <213> Homo sapiens  
 30 <400> 35  
 agcagcctcc ctcgactagc tcacactacg ataaggaaaa ttcatgagct ggtgtccaag 60  
 gagggtctgg tgactcgtgg ctcagtcagc atcaagattc ctttctctt tcccctctgc 120  
 c 121  
 <210> 36  
 <211> 121  
 <212> ADN  
 35 <213> Homo sapiens  
 <400> 36  
 agcagcctcc ctcgactagc tcacactacg ataaggaaaa ttcatgagct ggtgtccaag 60  
 gagggtctgg tgactcgtgg ctcagtcagc gtcaagattc ctttctctt tcccctctgc 120  
 c 121

ES 2 909 841 T3

	<210> 37		
	<211> 138		
	<212> ADN		
	<213> Homo sapiens		
5	<400> 37		
	tggcattgcc tgtaatatac atagccatgg ttttttatag gcaatttaag atgaatagct	60	
	tctaaactat agataagttt cattacccca ggaagctgaa ctatagctac tttacccaaa	120	
	atcattagaa tggtgctt	138	
	<210> 38		
	<211> 138		
	<212> ADN		
10	<213> Homo sapiens		
	<400> 38		
	tggcattgcc tgtaatatac atagccatgg ttttttatag gcaatttaag atgaatagct	60	
	tctaaactat agataagttt cattacccca ggaagctgaa ctatagctac tttcccaaaa	120	
	atcattagaa tggtgctt	138	
	<210> 39		
	<211> 136		
15	<212> ADN		
	<213> Homo sapiens		
	<400> 39		
	atgaagcctt ccaccaactg cctgtatgac tcatctgggg acttctgctc tatactcaaa	60	
	gtggcttagt cactgccaat gtatttccat atgagggacg atgattacta aggaaatata	120	
	gaaacaacaa ctgatc	138	
	<210> 40		
20	<211> 136		
	<212> ADN		
	<213> Homo sapiens		
	<400> 40		
	atgaagcctt ccaccaactg cctgtatgac tcatctgggg acttctgctc tatactcaaa	60	
	gtggcttagt cactgccaat gtatttccat atgagggacg gtgattacta aggaaatata	120	
	gaaacaacaa ctgatc	138	
25	<210> 41		
	<211> 118		
	<212> ADN		
	<213> Homo sapiens		
	<400> 41		
	acaacagaat caggtgattg gagaaaagat cacaggccta ggcaccaag gcttgaagga	60	
30	tgaaagaatg aaagatggac ggaacaaaat taggacctta attctttggt cagttcag	118	
	<210> 42		
	<211> 118		
	<212> ADN		
35	<213> Homo sapiens		
	<400> 42		
	acaacagaat caggtgattg gagaaaagat cacaggccta ggcaccaag gcttgaagga	60	
	tgaaagaatg aaagatggac ggaagaaaat taggacctta attctttggt cagttcag	118	
	<210> 43		
	<211> 150		
	<212> ADN		
40	<213> Homo sapiens		
	<400> 43		

ES 2 909 841 T3

	ttggggtaaa ttttcattgt catatgtgga atttaaatat accatcatct acaaagaatt	60
	ccacagagtt aaatatctta agttaaacac ttaaaataag tgtttgcgtg atattttgat	120
	gacagataaa cagagtctaa ttcccacccc	150
	<210> 44	
	<211> 150	
	<212> ADN	
5	<213> Homo sapiens	
	<400> 44	
	ttggggtaaa ttttcattgt catatgtgga atttaaatat accatcatct acaaagaatt	60
	ccacagagtt aaatatctta agttaaacac ttaaaataag tgtttgcgtg atattttgat	120
	gatagataaa cagagtctaa ttcccacccc	150
	<210> 45	
	<211> 145	
10	<212> ADN	
	<213> Homo sapiens	
	<400> 45	
	tgcaattcaa atcaggaagt atgaccacaaa gacagagatc ttttttggat gatccctagc	60
	ctagcaatgc ctggcagcca tgcaggtgca atgtcaacct taaataatgt attgcaaact	120
	cagagctgac aaacctcgat gttgc	145
	<210> 46	
	<211> 145	
15	<212> ADN	
	<213> Homo sapiens	
	<400> 46	
	tgcaattcaa atcaggaagt atgaccacaaa gacagagatc ttttttggat gatccctagc	60
	ctagcaatgc ctggcagcca tgcaggtgca atgtcaacct taaataatgt attgcaaact	120
	cagagctgac aaacctcgat gttgc	145
	<210> 47	
	<211> 124	
20	<212> ADN	
	<213> Homo sapiens	
	<400> 47	
	ctgtgctctg cgaatagctg cagaagtaac ttggggaccc aaaataaagc agaatgctaa	60
	tgtcaagtcc tgagaaccaa gccctgggac tctggtgcca tttcggattc tccatgagca	120
	tggt	124
25	<210> 48	
	<211> 124	
	<212> ADN	
	<213> Homo sapiens	
30	<400> 48	
	ctgtgctctg cgaatagctg cagaagtaac ttggggaccc aaaataaagc agaatgctaa	60
	tgtcaagtcc tgagaaccaa gccctgggac tctggtgcca tttcggattc tccatgagca	120
	tggt	124
	<210> 49	
	<211> 118	
	<212> ADN	
35	<213> Homo sapiens	
	<400> 49	
	tttttccagc caactcaagg ccaaaaaaaaa tttcttaata tagttattat gcgaggggag	60
	gggaagcaaa ggagcacagg tagtccacag aataagacac aagaaacctc aagctgtg	118

ES 2 909 841 T3

<210> 50  
 <211> 118  
 <212> ADN  
 <213> Homo sapiens  
 5 <400> 50  
 tttttccagc caactcaagg ccaaaaaaaaa tttcttaata tagttattat gcgaggggag 60  
  
 gggaagcaaa ggagcacagg tagtccacag aataggacac aagaaacctc aagctgtg 118  
 <210> 51  
 <211> 110  
 <212> ADN  
 10 <213> Homo sapiens  
 <400> 51  
 tcttctcgtc ccctaagcaa acaacatccg cttgcttctg tctgtgtaac cacagtgaat 60  
  
 ggggtgtgcac gcttgatggg cctctgagcc cctggtgcac aaaccagaaa 110  
 <210> 52  
 <211> 110  
 <212> ADN  
 15 <213> Homo sapiens  
 <400> 52  
 tcttctcgtc ccctaagcaa acaacatccg cttgcttctg tctgtgtaac cacagtgaat 60  
  
 ggggtgtgcac gcttggtggg cctctgagcc cctggtgcac aaaccagaaa 110  
 <210> 53  
 <211> 110  
 <212> ADN  
 20 <213> Homo sapiens  
 <400> 53  
 cacatggggg cattaagaat cgcccagggg ggaggagggg gaacgcgtgc ttttcacatt 60  
  
 tgcatttgaa ttttcagatt cccaggatgt gtttttgtgc tcatcgatgt 110  
 <210> 54  
 <211> 110  
 <212> ADN  
 25 <213> Homo sapiens  
 <400> 54  
 cacatggggg cattaagaat cgcccagggg ggaggagggg gaacgcgtgc ttttcacatt 60  
  
 tgcatttgaa tttttgagtt cccaggatgt gtttttgtgc tcatcgatgt 110  
 <210> 55  
 <211> 128  
 <212> ADN  
 30 <213> Homo sapiens  
 <400> 55  
 gggctctgag gtgtgtgaaa taaaaacaaa tgtccatgtc tgtcctttta tggcattttg 60  
  
 ggactttaca tttcaaacat ttcagacatg tatcacaaca cgaaggaata acagttccag 120  
 ggatatct 128  
 <210> 56  
 <211> 128  
 <212> ADN  
 40 <213> Homo sapiens  
 <400> 56  
 gggctctgag gtgtgtgaaa taaaaacaaa tgtccatgtc tgtcctttta tggcattttg 60  
  
 ggactttaca tttcaaacat ttcagacatg tatcacaaca cgaggaata acagttccag 120  
 ggatatct 128  
 <210> 57  
 45

<211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 5 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 57  
 cacatgcaca gccagcaacc c 21  
 <210> 58  
 <211> 23  
 10 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 58  
 15 ccccaaggctc ctgtgacctg agt 23  
 <210> 59  
 <211> 23  
 <212> ADN  
 <213> Secuencia artificial  
 20 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 59  
 tgaggaagtg aggctcagag ggt 23  
 <210> 60  
 25 <211> 24  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 30 <400> 60  
 tgccagtgcg agatgaaagt cttt 24  
 <210> 61  
 <211> 27  
 <212> ADN  
 35 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 61  
 gtgccttcag aacctttgag atctgat 27  
 40 <210> 62  
 <211> 20  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 45 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 62  
 tcccatccca ccagccaccc 20  
 <210> 63  
 <211> 25  
 50 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 63  
 55 aggtgtgtct ctcttttgtg agggg 25  
 <210> 64  
 <211> 21  
 <212> ADN

## ES 2 909 841 T3

<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 64  
5 cctttgtccc acctccccac c 21  
<210> 65  
<211> 26  
<212> ADN  
<213> Secuencia artificial  
10 <220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 65  
cctcgcctac tgtgctgttt ctaacc 26  
<210> 66  
15 <211> 25  
<212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
20 <400> 66  
ccatcccagc tgagtattcc aggag 25  
<210> 67  
<211> 26  
<212> ADN  
25 <213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 67  
aattgcaatg gtgagaggtt gatggt 26  
30 <210> 68  
<211> 24  
<212> ADN  
<213> Secuencia artificial  
<220>  
35 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 68  
ccagtgagaa gtgtcttggg ttgg 24  
<210> 69  
<211> 27  
40 <212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 69  
45 gaaatgcctt ctcaggtaat ggaaggt 27  
<210> 70  
<211> 27  
<212> ADN  
<213> Secuencia artificial  
50 <220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 70  
ggtttgagca gttctgagaa tgtggct 27  
<210> 71  
55 <211> 22  
<212> ADN  
<213> Secuencia artificial  
<220>

ES 2 909 841 T3

5 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 71  
acccaaaaca ctggaggggc ct 22  
<210> 72  
<211> 27  
<212> ADN  
<213> Secuencia artificial  
<220>  
10 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 72  
cccttatctg ctatgtggca tacttgg 27  
<210> 73  
<211> 27  
<212> ADN  
15 <213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 73  
gcaccagaat ttaaacaacg ctgacaa 27  
20 <210> 74  
<211> 22  
<212> ADN  
<213> Secuencia artificial  
<220>  
25 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 74  
gcacctgaca ggcacatcag cg 22  
<210> 75  
<211> 24  
30 <212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 75  
35 tgactgtata ccccaggtgc accc 24  
<210> 76  
<211> 27  
<212> ADN  
<213> Secuencia artificial  
40 <220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 76  
gcactaagga tgtggaagtc tagtgtg 27  
<210> 77  
45 <211> 22  
<212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
50 <400> 77  
tgtactgtggt caccagggga cg 22  
<210> 78  
<211> 26  
<212> ADN  
55 <213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 78

# ES 2 909 841 T3

5 agtgtgagaa gagcctcaag gacagc 26  
<210> 79  
<211> 21  
<212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 79  
10 cagtggaccc tgctgcacct t 21  
<210> 80  
<211> 24  
<212> ADN  
<213> Secuencia artificial  
<220>  
15 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 80  
gtggcaaagg agagagttgt gagg 24  
<210> 81  
<211> 24  
20 <212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 81  
25 cagtggcata gtagtccagg ggct 24  
<210> 82  
<211> 21  
<212> ADN  
<213> Secuencia artificial  
30 <220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 82  
cctctccgac aacttccgcc g 21  
<210> 83  
35 <211> 20  
<212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
40 <400> 83  
aggtctgggg gccgctgaat 20  
<210> 84  
<211> 23  
<212> ADN  
45 <213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 84  
tcctcccatt aaaccagca cct 23  
50 <210> 85  
<211> 23  
<212> ADN  
<213> Secuencia artificial  
<220>  
55 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 85  
acggttctgt cctgtagggg aga 23  
<210> 86

# ES 2 909 841 T3

<211> 22  
<212> ADN  
<213> Secuencia artificial  
<220>  
5 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 86  
cctgttcact tgtggcaggg ca 22  
<210> 87  
<211> 20  
10 <212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 87  
15 ggcagtcag atgggcgtgc 20  
<210> 88  
<211> 23  
<212> ADN  
<213> Secuencia artificial  
20 <220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 88  
tccagccctt gtcccaaacg tgt 23  
<210> 89  
25 <211> 21  
<212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
30 <400> 89  
gccggacctg cgaaatccca a 21  
<210> 90  
<211> 21  
<212> ADN  
35 <213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 90  
cgggcaactg gggctctgat c 21  
40 <210> 91  
<211> 21  
<212> ADN  
<213> Secuencia artificial  
<220>  
45 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 91  
agcagcctcc ctgactagc t 21  
<210> 92  
<211> 23  
50 <212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 92  
55 ggcagagggg aaagacgaaa gga 23  
<210> 93  
<211> 24  
<212> ADN

## ES 2 909 841 T3

<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 93  
5 tggcattgcc tgtaatatac atag 24  
<210> 94  
<211> 23  
<212> ADN  
<213> Secuencia artificial  
10 <220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 94  
aagcaccatt ctaatgattt tgg 23  
<210> 95  
15 <211> 20  
<212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
20 <400> 95  
atgaagcctt ccaccaactg 20  
<210> 96  
<211> 27  
<212> ADN  
25 <213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 96  
gatcagttgt tgtttctata tttcctt 27  
30 <210> 97  
<211> 22  
<212> ADN  
<213> Secuencia artificial  
<220>  
35 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 97  
acaacagaat caggtgattg ga 22  
<210> 98  
<211> 25  
40 <212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 98  
45 ctgaactgaa caaagaatta aggtc 25  
<210> 99  
<211> 22  
<212> ADN  
<213> Secuencia artificial  
50 <220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 99  
ttggggtaaa ttttcattgt ca 22  
<210> 100  
55 <211> 20  
<212> ADN  
<213> Secuencia artificial  
<220>

## ES 2 909 841 T3

5 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 100  
ggggtgggaa ttagactctg 20  
<210> 101  
<211> 23  
<212> ADN  
<213> Secuencia artificial  
<220>

10 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 101  
tgcaattcaa atcaggaagt atg 23  
<210> 102  
<211> 20  
<212> ADN  
<213> Secuencia artificial  
<220>

15 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 102  
gcaacatcga ggtttgcag 20  
<210> 103  
<211> 20  
<212> ADN  
<213> Secuencia artificial  
<220>

20 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 103  
ctgtgctctg cgaatagctg 20  
<210> 104  
<211> 20  
<212> ADN  
<213> Secuencia artificial  
<220>

25 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 104  
accatgctca tggagaatcc 20  
<210> 105  
<211> 20  
<212> ADN  
<213> Secuencia artificial  
<220>

30 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 105  
ttttccagc caactcaagg 20  
<210> 106  
<211> 21  
<212> ADN  
<213> Secuencia artificial  
<220>

35 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 106  
cacagcttga ggtttcttgt g 21  
<210> 107  
<211> 20  
<212> ADN  
<213> Secuencia artificial  
<220>

40 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 107

45

50

55

ES 2 909 841 T3

5 tcttctcgtc ccctaagcaa 20  
<210> 108  
<211> 20  
<212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 108  
10 tttctggttt gtgcaacagg 20  
<210> 109  
<211> 20  
<212> ADN  
<213> Secuencia artificial  
<220>  
15 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 109  
cacatggggg cattaagaat 20  
<210> 110  
<211> 22  
20 <212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 110  
25 acatcgatga gcacaaaaac ac 22  
<210> 111  
<211> 20  
<212> ADN  
<213> Secuencia artificial  
30 <220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 111  
gggctctgag gtgtgtgaaa 20  
<210> 112  
35 <211> 24  
<212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
40 <400> 112  
agatatccct ggaactgta ttcc 24  
<210> 113  
<211> 23  
<212> ADN  
45 <213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 113  
50 acagtaactg cttcataga tag 23  
<210> 114  
<211> 22  
<212> ADN  
<213> Secuencia artificial  
<220>  
55 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 114  
gtgtcagacc ctggtctaag ta 22  
<210> 115

<211> 26  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 5 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 115  
 aaataaaatt aggcataattt acaagc 26  
 <210> 116  
 <211> 23  
 10 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 116  
 15 gctgagtgat ttgtctgtaa ttg 23  
 <210> 117  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 20 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 117  
 cctggtcctc ccttatttcc c 21  
 <210> 118  
 25 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 30 <400> 118  
 gggaacacag actccatggt g 21  
 <210> 119  
 <211> 22  
 <212> ADN  
 35 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 119  
 cttaggaac cctcactgaa tg 22  
 40 <210> 120  
 <211> 22  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 45 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 120  
 gtccttgta gcgtttattt gc 22  
 <210> 121  
 <211> 26  
 50 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 121  
 55 aataatcagt atgtgacttg gattga 26  
 <210> 122  
 <211> 22  
 <212> ADN

<213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 122  
 5 ataggatgga tggatagatg ga 22  
 <210> 123  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 10 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 123  
 cagagcaaga ccctgtctca t 21  
 <210> 124  
 15 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 20 <400> 124  
 tcaacagagg cttgcatgta t 21  
 <210> 125  
 <211> 20  
 <212> ADN  
 25 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 125  
 30 gggatgatttt cctctttggt 20  
 <210> 126  
 <211> 33  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 35 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 126  
 aacattgta tctttatctg tatccttatt tat 33  
 <210> 127  
 <211> 27  
 40 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 127  
 45 gaacacttgt catagtttag aacgaac 27  
 <210> 128  
 <211> 20  
 <212> ADN  
 <213> Secuencia artificial  
 50 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 128  
 tcattgacag aattgcacca 20  
 <210> 129  
 55 <211> 27  
 <212> ADN  
 <213> Secuencia artificial  
 <220>

ES 2 909 841 T3

<223> Descripción de secuencia artificial: Cebador sintético  
<400> 129  
tttgatttc atgtgtacat tcgtatc 27  
5 <210> 130  
<211> 28  
<212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
10 <400> 130  
acctatcctg tagattatct tcactgtg 28  
<210> 131  
<211> 20  
<212> ADN  
15 <213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 131  
tctgaccat ctaacgccta 20  
20 <210> 132  
<211> 27  
<212> ADN  
<213> Secuencia artificial  
<220>  
25 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 132  
cagacagaaa gatagataga tgattga 27  
<210> 133  
<211> 21  
30 <212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 133  
35 atacagacag acagacaggt g 21  
<210> 134  
<211> 23  
<212> ADN  
<213> Secuencia artificial  
40 <220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 134  
gcatgtatct atcatccatc tct 23  
<210> 135  
45 <211> 21  
<212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
50 <400> 135  
tgagtgacaa attgagacct t 21  
<210> 136  
<211> 26  
<212> ADN  
55 <213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 136

gtcttacaat aacagttgct actatt 26  
 <210> 137  
 <211> 18  
 <212> ADN  
 5 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 137  
 attccccaag tgaattgc 18  
 10 <210> 138  
 <211> 26  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 15 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 138  
 ggtagataga ctggatagat agacga 26  
 <210> 139  
 <211> 20  
 20 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 139  
 25 tggaaacaga aatggcttgg 20  
 <210> 140  
 <211> 20  
 <212> ADN  
 <213> Secuencia artificial  
 30 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 140  
 gattgcagga gggaaggaag 20  
 <210> 141  
 35 <211> 22  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 40 <400> 141  
 gagcaagaca ccatctcaag aa 22  
 <210> 142  
 <211> 30  
 <212> ADN  
 45 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 142  
 gaaattttac atttatgttt atgattctct 30  
 50 <210> 143  
 <211> 18  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 55 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 143  
 ggcgactgag caagactc 18  
 <210> 144

<211> 27  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 5 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 144  
 ggttattaat tgagaaaact ccttaca 27  
 <210> 145  
 <211> 23  
 10 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 145  
 15 attttcccg atgatagtag tct 23  
 <210> 146  
 <211> 26  
 <212> ADN  
 <213> Secuencia artificial  
 20 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 146  
 gcgaatgtat gattggcaat attttt 26  
 <210> 147  
 25 <211> 27  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 30 <400> 147  
 acatgtatcc cagaacttaa agtaaac 27  
 <210> 148  
 <211> 22  
 <212> ADN  
 35 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 148  
 40 gcagaaggga aaattgaagc tg 22  
 <210> 149  
 <211> 22  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 45 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 149  
 cagagacacc gaaccaataa ga 22  
 <210> 150  
 <211> 27  
 50 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 150  
 55 gccacatgaa tcaattccta taataaa 27  
 <210> 151  
 <211> 25  
 <212> ADN

ES 2 909 841 T3

<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 151  
5 gcacatgtac cctaaaactt aaaat 25  
<210> 152  
<211> 26  
<212> ADN  
<213> Secuencia artificial  
10 <220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 152  
gtcaaccaaa actcaacaag tagtaa 26  
<210> 153  
15 <211> 25  
<212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
20 <400> 153  
aagatgaaat tgccatgtaa aaata 25  
<210> 154  
<211> 27  
<212> ADN  
25 <213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 154  
gtgtgtataa caaaattcct atgatgg 27  
30 <210> 155  
<211> 22  
<212> ADN  
<213> Secuencia artificial  
<220>  
35 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 155  
gcacccaaaa ctgaatgtca ta 22  
<210> 156  
<211> 20  
40 <212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 156  
45 ggtgagagtg agaccctgtc 20  
<210> 157  
<211> 27  
<212> ADN  
<213> Secuencia artificial  
50 <220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 157  
tgaataact ctacgactgt ctgtctg 27  
<210> 158  
55 <211> 21  
<212> ADN  
<213> Secuencia artificial  
<220>

ES 2 909 841 T3

5 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 158  
gaataggagg tggatggatg g 21  
<210> 159  
<211> 20  
<212> ADN  
<213> Secuencia artificial  
<220>  
10 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 159  
gagcgagacc ctgtctcaag 20  
<210> 160  
<211> 24  
<212> ADN  
15 <213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 160  
ggaaaagaca taggatagca attt 24  
20 <210> 161  
<211> 21  
<212> ADN  
<213> Secuencia artificial  
<220>  
25 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 161  
tctggattga tctgtctgtc c 21  
<210> 162  
<211> 26  
30 <212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 162  
35 gaattaaata ccatctgagc actgaa 26  
<210> 163  
<211> 27  
<212> ADN  
<213> Secuencia artificial  
40 <220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 163  
tgttataatg cattgagttt tattctg 27  
<210> 164  
45 <211> 28  
<212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
50 <400> 164  
gcctgtctca aaaataaaga gatagaca 28  
<210> 165  
<211> 25  
<212> ADN  
55 <213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 165

ES 2 909 841 T3

5 ttaatgaatt gaacaaatga gtgag 25  
<210> 166  
<211> 25  
<212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 166  
10 gcaactctgg ttgtattgtc ttcac 25  
<210> 167  
<211> 21  
<212> ADN  
<213> Secuencia artificial  
<220>  
15 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 167  
caaagcgaga ctctgtctca a 21  
<210> 168  
<211> 27  
20 <212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 168  
25 gaaaatgcta tcctctttgg tataaat 27  
<210> 169  
<211> 27  
<212> ADN  
<213> Secuencia artificial  
30 <220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 169  
gggtatttca agataactgt agatagg 27  
<210> 170  
35 <211> 24  
<212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
40 <400> 170  
gcttctgaaa gcttctagtt tacc 24  
<210> 171  
<211> 19  
<212> ADN  
45 <213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 171  
50 tccacatcct caccaacac 19  
<210> 172  
<211> 21  
<212> ADN  
<213> Secuencia artificial  
<220>  
55 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 172  
gcctaggaag gctactgtca a 21  
<210> 173

<211> 18  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 5 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 173  
 ccacccgtcc atttaggc 18  
 <210> 174  
 <211> 27  
 10 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 174  
 15 gtgaaaaagt agatataatg gttggtg 27  
 <210> 175  
 <211> 25  
 <212> ADN  
 <213> Secuencia artificial  
 20 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 175  
 ggttttccaa gagatagacc aatta 25  
 <210> 176  
 25 <211> 27  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 30 <400> 176  
 gtcctctcat aaatccctac tcatatc 27  
 <210> 177  
 <211> 27  
 <212> ADN  
 35 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 177  
 ctgttgttac ataataggta ggtaggt 27  
 40 <210> 178  
 <211> 19  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 45 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 178  
 gtcgtgggcc ccataaatc 19  
 <210> 179  
 <211> 27  
 50 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 179  
 55 aaggtacata acagttcaat agaaagc 27  
 <210> 180  
 <211> 26  
 <212> ADN

ES 2 909 841 T3

<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 180  
5 gtgaaatgac tgaaaaatag taacca 26  
<210> 181  
<211> 24  
<212> ADN  
<213> Secuencia artificial  
10 <220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 181  
ctaggagatc atgtgggtat gatt 24  
<210> 182  
15 <211> 24  
<212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
20 <400> 182  
gcagtgaata aatgaacgaa tgga 24  
<210> 183  
<211> 21  
<212> ADN  
25 <213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 183  
30 cccaaaatta cttgagccaa t 21  
<210> 184  
<211> 24  
<212> ADN  
<213> Secuencia artificial  
<220>  
35 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 184  
gagacaaaat gaagaacag acag 24  
<210> 185  
<211> 25  
40 <212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 185  
45 tctttgctct catgaataga tcagt 25  
<210> 186  
<211> 24  
<212> ADN  
<213> Secuencia artificial  
50 <220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 186  
gttttgata atgaaccac tcag 24  
<210> 187  
55 <211> 27  
<212> ADN  
<213> Secuencia artificial  
<220>

ES 2 909 841 T3

<223> Descripción de secuencia artificial: Cebador sintético  
<400> 187  
tgaacacaga tgттаagtgt gtatatg 27  
5 <210> 188  
<211> 23  
<212> ADN  
<213> Secuencia artificial  
<220>  
10 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 188  
gtctgagggtg gacagttatg aaa 23  
<210> 189  
<211> 23  
<212> ADN  
15 <213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 189  
ctgtggctca tctatgaaa ctt 23  
20 <210> 190  
<211> 22  
<212> ADN  
<213> Secuencia artificial  
<220>  
25 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 190  
gaagtggctg tgggttatg at 22  
<210> 191  
<211> 24  
30 <212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 191  
35 ttctgttggt atagagcagt gttt 24  
<210> 192  
<211> 20  
<212> ADN  
<213> Secuencia artificial  
40 <220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 192  
gtgacaggaa ggacggaatg 20  
<210> 193  
45 <211> 27  
<212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
50 <400> 193  
catgaggttt gcaaatacta tcttaac 27  
<210> 194  
<211> 24  
<212> ADN  
55 <213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 194

gttttaattt tctccaaatc tcca 24  
 <210> 195  
 <211> 27  
 <212> ADN  
 5 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 195  
 tcttagccta gatagatact tgcttcc 27  
 10 <210> 196  
 <211> 23  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 15 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 196  
 gtcaaccttt gaggctatag gaa 23  
 <210> 197  
 <211> 19  
 20 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 197  
 25 tcctggaaac aaaagtatt 19  
 <210> 198  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 30 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 198  
 aaccttaaa caaagctaga a 21  
 <210> 199  
 35 <211> 20  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 40 <400> 199  
 actaagcctt ggggatccag 20  
 <210> 200  
 <211> 21  
 <212> ADN  
 45 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 200  
 tgctgtggaa atactaaaag g 21  
 50 <210> 201  
 <211> 20  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 55 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 201  
 ctccagaggt aatcctgtga 20  
 <210> 202

ES 2 909 841 T3

<211> 21  
<212> ADN  
<213> Secuencia artificial  
<220>  
5 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 202  
tgggtgaga tggatatctag g 21  
<210> 203  
<211> 22  
10 <212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 203  
15 gtataatcca tgaatcttgt tt 22  
<210> 204  
<211> 22  
<212> ADN  
<213> Secuencia artificial  
20 <220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 204  
ttcaaattgt atataagaga gt 22  
<210> 205  
25 <211> 20  
<212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
30 <400> 205  
gcaggaaagt tatttttaat 20  
<210> 206  
<211> 21  
<212> ADN  
35 <213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 206  
tgcttgagaa agctaacact t 21  
40 <210> 207  
<211> 20  
<212> ADN  
<213> Secuencia artificial  
<220>  
45 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 207  
cagtgtttgg aaattgtctg 20  
<210> 208  
<211> 20  
50 <212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 208  
55 ggcactggga gattattgta 20  
<210> 209  
<211> 20  
<212> ADN

ES 2 909 841 T3

<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 209  
5 tcctgttggtt aagtacacat 20  
<210> 210  
<211> 18  
<212> ADN  
<213> Secuencia artificial  
10 <220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 210  
gggccgtaat tacttttg 18  
<210> 211  
15 <211> 21  
<212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
20 <400> 211  
actcagtagg cactttgtgt c 21  
<210> 212  
<211> 18  
<212> ADN  
25 <213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 212  
tcttcacca caccaatc 18  
30 <210> 213  
<211> 19  
<212> ADN  
<213> Secuencia artificial  
<220>  
35 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 213  
tggcttttca aaggtaaaa 19  
<210> 214  
<211> 21  
40 <212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 214  
45 gcaacgttaa catctgaatt t 21  
<210> 215  
<400> 215  
000  
<210> 216  
50 <211> 22  
<212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
55 <400> 216  
atattatag tcatgatcta ag 22  
<210> 217  
<211> 19

<212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 5 <400> 217  
 agagattaca ggtgtgagc 19  
 <210> 218  
 <211> 19  
 <212> ADN  
 10 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 218  
 atgacctca actgcctct 19  
 15 <210> 219  
 <211> 20  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 20 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 219  
 tgaaactcaa aagagaaaag 20  
 <210> 220  
 <211> 20  
 25 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 220  
 30 acagatttct acttaaaatt 20  
 <210> 221  
 <211> 20  
 <212> ADN  
 <213> Secuencia artificial  
 35 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 221  
 tgaaactcaa aagagaaaag 20  
 <210> 222  
 40 <211> 20  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 45 <400> 222  
 acagatttct acttaaaatt 20  
 <210> 223  
 <211> 20  
 <212> ADN  
 50 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 223  
 gcaaaggggt actctatgta 20  
 55 <210> 224  
 <211> 20  
 <212> ADN  
 <213> Secuencia artificial

ES 2 909 841 T3

<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 224  
5 tatcgggtca tcttgtaaa 20  
<210> 225  
<211> 22  
<212> ADN  
<213> Secuencia artificial  
<220>  
10 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 225  
tctaacaaag ctctgtccaa aa 22  
<210> 226  
<211> 21  
15 <212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 226  
20 ccacactgaa taactggaac a 21  
<210> 227  
<211> 21  
<212> ADN  
<213> Secuencia artificial  
25 <220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 227  
gcaagcaagc tctctacctt c 21  
<210> 228  
30 <211> 22  
<212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
35 <400> 228  
tgttcttcca aaattcacat gc 22  
<210> 229  
<211> 21  
<212> ADN  
40 <213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 229  
atttactat tccttcattt t 21  
45 <210> 230  
<211> 22  
<212> ADN  
<213> Secuencia artificial  
<220>  
50 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 230  
taattgtgac aactaaatt ac 22  
<210> 231  
<211> 21  
55 <212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético

<400> 231  
 aaaaagccac agaaatcagt c 21  
 <210> 232  
 <211> 22  
 5 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 232  
 10 ttcttatatc tcaactgggca tt 22  
 <210> 233  
 <211> 22  
 <212> ADN  
 <213> Secuencia artificial  
 15 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 233  
 ggatggtaga agagaagaaa gg 22  
 <210> 234  
 20 <211> 22  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 25 <400> 234  
 ggatggtaga agagaagaaa gg 22  
 <210> 235  
 <211> 20  
 <212> ADN  
 30 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 235  
 tgcaaagatg cagaaccaac 20  
 35 <210> 236  
 <211> 22  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 40 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 236  
 ttttgttcct tgtcctggct ga 22  
 <210> 237  
 <211> 20  
 45 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 237  
 50 tgcaaagatg cagaaccaac 20  
 <210> 238  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 55 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 238  
 gcctccagct ctatccaagt t 21

ES 2 909 841 T3

5 <210> 239  
<211> 22  
<212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 239  
ccttaatatc ttcccatgtc ca 22  
10 <210> 240  
<211> 22  
<212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
15 <400> 240  
attgtagtg cctcttctgc tt 22  
<210> 241  
<211> 20  
<212> ADN  
20 <213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 241  
gagaagtgag gtcagcagct 20  
25 <210> 242  
<211> 22  
<212> ADN  
<213> Secuencia artificial  
<220>  
30 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 242  
tttctaaatt tccattgaac ag 22  
<210> 243  
<211> 20  
35 <212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 243  
40 gaaattggca atctgattct 20  
<210> 244  
<211> 21  
<212> ADN  
<213> Secuencia artificial  
45 <220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 244  
caacttgccc tttattgatg t 21  
50 <210> 245  
<211> 22  
<212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
55 <400> 245  
ctatgtgat aaaacattga aa 22  
<210> 246  
<211> 20

ES 2 909 841 T3

<212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
5 <400> 246  
gcctgtctgg aatatagttt 20  
<210> 247  
<211> 22  
<212> ADN  
10 <213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 247  
cagggcatat aatctaagct gt 22  
15 <210> 248  
<211> 21  
<212> ADN  
<213> Secuencia artificial  
<220>  
20 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 248  
caatgactct gagttgagca c 21  
<210> 249  
<211> 18  
25 <212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 249  
30 actctctccc tcccctct 18  
<210> 250  
<211> 20  
<212> ADN  
<213> Secuencia artificial  
35 <220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 250  
tatggcccca aaactattct 20  
<210> 251  
40 <211> 20  
<212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
45 <400> 251  
acaagtactg ggcagattga 20  
<210> 252  
<211> 20  
<212> ADN  
50 <213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 252  
gccaggttta gctttcaagt 20  
55 <210> 253  
<211> 22  
<212> ADN  
<213> Secuencia artificial

<220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 253  
 ttttatatca ggagaaacac tg 22  
 5 <210> 254  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 10 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 254  
 ccagaatttt ggaggtttaa t 21  
 <210> 255  
 <211> 22  
 15 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 255  
 20 tgtcatcct ctttatctc ca 22  
 <210> 256  
 <211> 20  
 <212> ADN  
 <213> Secuencia artificial  
 25 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 256  
 ttcttttgcc tctcccaaag 20  
 <210> 257  
 30 <211> 20  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 35 <400> 257  
 accctggcac agtggtgact 20  
 <210> 258  
 <211> 20  
 <212> ADN  
 40 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 258  
 tgggcctgag ttgagaagat 20  
 45 <210> 259  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 50 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 259  
 aatttgtaag tatgtgcaac g 21  
 <210> 260  
 <211> 20  
 55 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético

<400> 260  
 tttttcccat ttccaactct 20  
 <210> 261  
 <211> 20  
 5 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 261  
 10 aaaagatgag acaggcaggt 20  
 <210> 262  
 <211> 20  
 <212> ADN  
 <213> Secuencia artificial  
 15 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 262  
 acccctgtga atctcaaaat 20  
 <210> 263  
 20 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 25 <400> 263  
 gcacttgctt ctattgttg t 21  
 <210> 264  
 <211> 20  
 <212> ADN  
 30 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 264  
 cccttctct cttcattct 20  
 35 <210> 265  
 <211> 13  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 40 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 265  
 agcactgcag gta 13  
 <210> 266  
 <211> 21  
 45 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 266  
 50 acagatacca aagaactgca a 21  
 <210> 267  
 <211> 20  
 <212> ADN  
 <213> Secuencia artificial  
 55 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 267  
 tggacacctt tcaacttaga 20

<210> 268  
 <211> 22  
 <212> ADN  
 <213> Secuencia artificial  
 5 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 268  
 gaacagtaat gttgaacttt tt 22  
 10 <210> 269  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 15 <400> 269  
 tcttgcaaaa agcttagcac a 21  
 <210> 270  
 <211> 21  
 <212> ADN  
 20 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 270  
 aaaaagatct caaagggtcc a 21  
 25 <210> 271  
 <211> 20  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 30 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 271  
 gcttttgctg aacatcaagt 20  
 <210> 272  
 <211> 19  
 35 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 272  
 40 ccttcagca gcatagtct 19  
 <210> 273  
 <211> 18  
 <212> ADN  
 <213> Secuencia artificial  
 45 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 273  
 aaatccagga tgtgcagt 18  
 <210> 274  
 50 <211> 19  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 55 <400> 274  
 atgatgaggt cagtgggtg 19  
 <210> 275  
 <211> 22

<212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 5 <400> 275  
 catcacagat catagtaaatt gg 22  
 <210> 276  
 <211> 21  
 <212> ADN  
 10 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 276  
 aattattatt ttgcaggcaa t 21  
 15 <210> 277  
 <211> 20  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 20 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 277  
 catgaggcaa acacctttcc 20  
 <210> 278  
 <211> 22  
 25 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 278  
 30 gctggactca ggataaagaa ca 22  
 <210> 279  
 <211> 20  
 <212> ADN  
 <213> Secuencia artificial  
 35 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 279  
 tggaagcctg agctgactaa 20  
 <210> 280  
 40 <211> 20  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 45 <400> 280  
 ccttcttttc ccccagaatc 20  
 <210> 281  
 <211> 21  
 <212> ADN  
 50 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 281  
 taggagaaca gaagatcaga g 21  
 55 <210> 282  
 <211> 22  
 <212> ADN  
 <213> Secuencia artificial

ES 2 909 841 T3

<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 282  
5 aaagactatt gctaaatgct tg 22  
<210> 283  
<211> 20  
<212> ADN  
<213> Secuencia artificial  
<220>  
10 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 283  
taagcgtagg gctgtgtgtg 20  
<210> 284  
<211> 21  
15 <212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 284  
20 ggacggatag actccagaag g 21  
<210> 285  
<211> 23  
<212> ADN  
<213> Secuencia artificial  
25 <220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 285  
gaatgacctt ggcactttta tca 23  
<210> 286  
30 <211> 27  
<212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
35 <400> 286  
aaggatagag atatacagat gaatgga 27  
<210> 287  
<211> 18  
<212> ADN  
40 <213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 287  
45 catgcaccgc gcaaatac 18  
<210> 288  
<211> 19  
<212> ADN  
<213> Secuencia artificial  
<220>  
50 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 288  
atgcctcacc cacaacac 19  
<210> 289  
<211> 20  
55 <212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético

<400> 289  
 tccaagccct tctcactcac 20  
 <210> 290  
 <211> 20  
 5 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 290  
 10 ctgggacggt gacattttct 20  
 <210> 291  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 15 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 291  
 cccaggaaga gtggaagat t 21  
 <210> 292  
 20 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 25 <400> 292  
 ttagcttgca tgtacctgtg t 21  
 <210> 293  
 <211> 20  
 <212> ADN  
 30 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 293  
 35 agctagatgg ggtgaatttt 20  
 <210> 294  
 <211> 18  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 40 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 294  
 tgggctgagg ggagattc 18  
 <210> 295  
 <211> 22  
 45 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 295  
 50 atcaagctaa ttaatgttat ct 22  
 <210> 296  
 <211> 20  
 <212> ADN  
 <213> Secuencia artificial  
 55 <220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 296  
 aatgaataag gtcctcagag 20

ES 2 909 841 T3

5 <210> 297  
<211> 21  
<212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 297  
tttaatctga tcattgccct a 21  
10 <210> 298  
<211> 18  
<212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
15 <400> 298  
agctgtgggt gaccttga 18  
<210> 299  
<211> 20  
<212> ADN  
20 <213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 299  
tgtcccacca ttgtgtatta 20  
25 <210> 300  
<211> 20  
<212> ADN  
<213> Secuencia artificial  
<220>  
30 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 300  
tcagacttga agtccaggat 20  
<210> 301  
<211> 20  
35 <212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 301  
40 gcttcagggg tgtagtttt 20  
<210> 302  
<211> 20  
<212> ADN  
<213> Secuencia artificial  
45 <220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 302  
ctttgtgaaa agtcgtccag 20  
50 <210> 303  
<211> 18  
<212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
55 <400> 303  
ccatcatgga aagcatgg 18  
<210> 304  
<211> 20

ES 2 909 841 T3

<212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
5 <400> 304  
tcattctccat gactgcacta 20  
<210> 305  
<211> 20  
<212> ADN  
10 <213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 305  
gagatgacgg agtagctcat 20  
15 <210> 306  
<211> 18  
<212> ADN  
<213> Secuencia artificial  
<220>  
20 <223> Descripción de secuencia artificial: Cebador sintético  
<400> 306  
cccagctgca ctgtctac 18  
<210> 307  
<211> 20  
25 <212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 307  
30 tcttgttcca atcacaggac 20  
<210> 308  
<211> 20  
<212> ADN  
<213> Secuencia artificial  
35 <220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 308  
atgctgttag ctgaagctct 20  
<210> 309  
40 <211> 20  
<212> ADN  
<213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
45 <400> 309  
tgaaagctcc taaagcagag 20  
<210> 310  
<211> 20  
<212> ADN  
50 <213> Secuencia artificial  
<220>  
<223> Descripción de secuencia artificial: Cebador sintético  
<400> 310  
ttgaagagat gtgctatcat 20  
55 <210> 311  
<211> 54  
<212> ADN  
<213> Secuencia artificial

ES 2 909 841 T3

<220>  
 <223> Descripción de secuencia artificial: Cebador sintético  
 <400> 311  
 5 gccgcctgca gcccgcgcc cccgtgcccc cgccccgccg ccggcccggg cgcc 54  
 <210> 312  
 <211> 601  
 <212> ADN  
 <213> Homo sapiens  
 <400> 312  
 catagtgaca ggtatatgcc caactaactg tggaaaacag ttctttcttt caaccttact 60  
 catcaccctc acggtctggt tatgaggctc tcctccacca gccagaaagg atgacgtgcc 120  
 atacctgcaa aacttataca gcatcaacag aatgaatcct tccaacaagc cgaaacattg 180  
 agtattgtgg cacagaatat gcccaccca ttactcaatc tagatatcct tttattccac 240  
 cgtctcatga ttttcttttt cctggaaaac aaaagtatth ctttcatagc ccagctagca 300  
 ygataaatca gcgagtcaga attctagctt tgttgtaagg ttttgcgaaat atctgatcct 360  
 cttatthtgt actthtctat thcctaggca aatctgagta thtcaccag thttccttaa 420  
 ctaggcattg aaaactcagt thttthctta caaaccttca tgtcttctg ctcatthtga 480  
 cagtcttatc thgcacctcc tataaaatgg agaaaactga cattaaaacg taaththttht 540  
 tacaththtga gggattccca gagaaththt cccaatctc cttaggtagg gactthttht 600  
 10 c 601  
 <210> 313  
 <211> 601  
 <212> ADN  
 <213> Homo sapiens  
 15 <400> 313  
 gtgggaacta tagtaagaa gtccctacct aaggagattg gggaaaaatt ctctgggaat 60  
 ccctcaaaat gtaataaaaa ttacgtthta atgtcaagtt tctccaththt ataggaggtg 120  
 caagataaga ctgtgcaaat gagcaggaag acatgaaggt ttgtaagaaa aaaactgagt 180  
 thtcaatgcc tagttaagga aaactgggtg aaatactcag atthtgcctag gaaatagaaa 240  
 agtacaaaat aaggagatca gatathtgcga aaaccttaca acaaagctag aathtctgact 300  
 ygctgaththt tcgtgctagc tgggctatga aagaaatact thtgththtcc aggaaaaaga 360  
 aatcatgag acggtggaat aaaaggatat ctgathtgag taatgggtgg ggcatathtct 420  
 gtgccacaat actcaatgtht tcggctthttht ggaaagathc atthtctgthtga tgctgtataa 480  
 gththtgcagg tatggcacgt catcctthtct ggctggthtga ggagagctc ataaacagac 540  
 cgtgagggtg atgagthaag thgaaagaaa gaactgththt ccacagthttag thgggcatat 600  
 a 601  
 <210> 314  
 <211> 650  
 <212> ADN  
 20 <213> Homo sapiens  
 <400> 314

ES 2 909 841 T3

tttattggtc ctgactggta caaatactga taaaaaggat tttaagatca tattcatact 60  
 tttggggaat gagagccaca attaattaac aatgtctgcc atgagattgg atgcaagagt 120  
 atggcactca tactattcct acttctgtct aattacacta tttgtttctg tgtgcaaaaa 180  
 tctttggtag gtgggtggatg tgcccaagac acaggaaga aaaagaagta aacagggaga 240  
 tacaacacag actctgaaat ggggcatcat ggaagacgga gctttgtcgt cttggctctt 300  
 gctgtatatt cacttctac aacagtgcta aataccttgt ggatgcttaa atataataa 360  
 tgaatgcata aatgaaaaga gtaaataaag agtgtatatg aaagtatgta gataaaattc 420  
 ttcactaagc cttggggatc cagctgcttm aggactaaga ccgtatctag ctccttttag 480  
 tatttcaca gcatgccatg gagatacatg tttctgatta tatatgatac atggaaatta 540  
 tatgttggtg aatgagtgat tgagtaaagtg tgtactaggg cagctaatca taaatatttc 600  
 tactattgct aaaatgactg gatttatcca ttccttctga gagtttatac 650  
 <210> 315  
 <211> 626  
 <212> ADN  
 <213> Homo sapiens  
 <400> 315  
 ctgcttaagg actaagaccr tatctagctc cttttagtat ttccacagca tgccatggag 60  
 atacatgttt ctgattatat atgatacatg gaaattatat gttggtgaat gagtgattga 120  
 gtaaagtgtg actagggcag ctaatcataa atatttctac tattgctaaa atgactggat 180  
 ttatccattc cttctgagag tttatactga ttgcttatat tgtatcaaat accgtaactg 240  
 agggcaatgt ttaactcaaac taatagcacc attcaaattt atgcaaaca taacactata 300  
 tctttaaaat gttttcacta aaagctgcat aaagagtgta ttcaacaaca atagaataat 360  
 tttacaatct tttttcttgc ttaatggcca tttgtgcctt ctgacatgct gctagccatt 420  
 caaaggtcac actaccttga agttgaagat caagacaaat gattagactc ataaaagaca 480  
 aatcacgtct ttctggacag gtgattatta ataattaatt agcatttaa catgtattat 540  
 ttaagttctt ttttaagttat aaagtctttg atttgctaaa cagtttaaat aatgaataaa 600  
 acataaaata ataatagtta ccattt 626  
 <210> 316  
 <211> 1113  
 <212> ADN  
 <213> Homo sapiens  
 <400> 316

5

10

ES 2 909 841 T3

caagagctgc atctcactcc aatTTTTtctt ctccctataa ccttatctag attcccagtt	60
gagggaaaccg atgacctaat tcctctcagt ttaaATGcaa cacaggagca aattccaaat	120
atctatgctg gtcttgctgg gattgcagaa ccccagggTg gttatcctcc tccagaggta	180
atcctgtgat cagcactaac rccacatacc agccctttca tcagcttgTt ggagaagcat	240
ctttacttcc caccaagcag tgacctagat accatctcac accagttaga atcaggatca	300
ttaaaaaagtc aagaaaaaac agatgctgaa gaggatgtgg agaaatagga atgcttttac	360
actgttagtg ggaatgtaaa ttagttcaac cattgtcaaa gacagtgtgg cgatccctca	420
cagatctaga accagaaata ccatttgacc cagcaatccc attactgggt ctatacccaa	480
aggattataa attactctac tataaagaca catgcacaca tatgtttatt gcagcaccat	540
tcacaatagc aaagaattgc aaccaaccct aatgcccctc aatgacagac tggataaaga	600
aaatctggca catatacacc atggaatact acgcagccat aaaaaggat 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 aaaaaaaaa agtgtaatcc agtttacatt ttcaaggTca aagtgggtac	1020
aatgctatct atcttgggct aagaagagaa aaggaaaaat tcttgcttta aatcttagaa	1080
gtctggtttt tttccctgTt ttgtacccca tcc	1113
<210> 317	
<211> 1113	
<212> ADN	
<213> Homo sapiens	
<400> 317	

5

ES 2 909 841 T3

ttcccagttg agggaaccga tgacctaatt cctctcagtt taaatgcaac acaggagcaa	60
attccaaata tctatgctgg tcttgctggg attgcagaac cccaggggtg ttatcctcct	120
ccagaggtaa tcctgtgatc agcactaacg ccacatacca gccctttcat cagcttgttg	180
gagaagcatc tttacttccc rccaagcagt gacctagata ccatctcaca ccagttagaa	240
tcaggatcat taaaaagtca agaaaaaaca gatgctgaag aggatgtgga gaaataggaa	300
tgcttttaca ctgttagtgg gaatgtaaat tagttcaacc attgtcaaag acagtgtggc	360
gatccctcac agatctagaa ccagaaatac catttgacc agcaatcca ttactgggtc	420
tatacccaa ggattataa ttactctact ataaagacac atgcacacat atgtttattg	480
cagcaccatt cacaatagca aagaattgca accaacccta atgccatca atgacagact	540
ggataaagaa aatctggcac atatacacca tggaatacta cgcagccata aaaaaggatg	600
agtttatgtc ctttacaggg acatggatga agctggaaac catcattctc agcaaactaa	660
cacaggaaca gaaaaccaa cacatgttct cactcacaag tgggagttga acaatgagaa	720
cacatggaca cagggagggg aacatcacac accactgctt gtcaggggggt ggggggctag	780
gggaaggata gcattaggag aaatacctaa tgtagatgaa gggttgatgg gtgcagcaaa	840
ccaccatggc atgtgtatac ctgtgtaaca aacctccatg ttctgcacgt gtatccaga	900
acttaaagta caatacaaaa aaaaaaaaaa gtgtaatcca gtttacattt tcaaggtcaa	960
agtgggtaca atgctatcta tcttgggcta agaagagaaa aggaaaaatt cttgctttaa	1020
atcttagaag tctgggtttt ttccctgttt tgtaccccat cctcttggtc tctctagata	1080
tatttaagac tcacatagga cttgtctttt cta	1113
<210> 318	
<211> 1001	
<212> ADN	
<213> Homo sapiens	
<400> 318	
tcatcaacta aatagttgat gaggggaaat tgttctgtat atgttcatac ttcagctaat	60
caattaaana tgatgaaata ataagattac cattttgcaa acccctaata caatgttga	120
tccaggcaat gatcatcaat ggccactaaa atcacacaaa aggagataac cagaatatgt	180

5

ES 2 909 841 T3

	gctttgtgat ggaagcatta aatacaacta atgagatatt gtttataaga aagaaaggaa	240
	gcaagaaagc aatcacacca agctctgtat ctagctacca catttaagga aaaaaagaga	300
	cagaagagca tgttaaatgt taccaagaag atacagtcag tcggaaaaaa tacagacaag	360
	aaaatacaga gcaaaacaac ccagcttctt cagcaaatca atataaaaaa attttaagaa	420
	agagttaaag tataaactga gagacttcag aaacatatta tccaagtata atccatgaat	480
	cttgtttaa tatagatcaa rtaaaccact ataccaaaaa catcaaaaga caactgggta	540
	aatthtttaa atgactagct atttgatgtt aaggaagtaa tgttactctc ttatatacaa	600
	tttgaataa tctagcgagg agcagcaaat gtgcggctat gaggaagaaa cacaattggc	660
	cattcttgaa tcattagctg gatgggtgct atatgggggt agattttact actctctaat	720
	tttacaataa tttaaaatgt tccataataa attggtgagt tatcaaaaga aatatttcta	780
	tataatagct aaaattatth ataaaagta gtggtctcat aactttatth atthatttac	840
	ttattttgag accgagtctc cctctgttat gcaggctgga gtgcagtggc tccatctcgg	900
	ctcactgcaa acttcacctc ctggattgaa gcgattctcc tgcctcagcc cccccgagta	960
	gctgggatta caggcttgca cccccagcc cagctaathh t	1001
	<210> 319	
	<211> 601	
	<212> ADN	
5	<213> Homo sapiens	
	<400> 319	
	agctaccaca ttttaaggaaa aaaagagaca gaagagcatg ttaaatgtta ccaagaagat	60
	acagtcagtc ggaaaaaata cagacaagaa aatacagagc aaaacaacc agcttcttca	120
	gcaaatcaat ataaaaaaat ttttaagaaag agttaaagta taaactgaga gacttcagaa	180
	acataattatc caagtataat ccatgaatct tgtttaaata tagatcaaat aaaccactat	240
	acaaaaaaca tcaaaagaca actgggtaaa ttttttaaat gactagctat ttgatgttaa	300
	rgaagtaatg ttactctctt atatacaatt tgaaataatc tagcgaggag cagcaaatgt	360
	gcggtatga ggaagaaaca caattggcca ttcttgaatc attagctgga tgggtgctat	420
	atggggtag atthttactac tctctaathh tacatatatt taaaatgttc cataataaat	480
	tgttgagtta tcaaaagaaa tathhctata taatagctaa aathhthhtht aaaagthhtht	540
	ggtctcataa cthhthhthhtht thhthhthhtht atthhthhthhtht cgagthhthhthhtht	600
	a	601
	<210> 320	
	<211> 601	
	<212> ADN	
10	<213> Homo sapiens	
	<400> 320	

ES 2 909 841 T3

ccaactgata taattagata aacttagtca atatatttga atccacatt ccagcagcta 60  
 ttttctccat ttgcttttat tgctgtttgt ggtgagtttg atatataatt ttaaggtggt 120  
 aacatcccta acttatgtat ggttacagct cataaatacg aacctgtgtc atgcaactca 180  
 tatatgactg tgttcaaaat aatgtgtatt agactgtaaa acgattttaa tattttaaat 240  
 aactttcctg catttgtcgg tttcagcagg aaagttattt ttaataactt ccctgtattt 300  
 sttggtttca gtattaatta atctcattaa tgctaaactt tgtgaccta ggttaaaaaa 360  
 catattcaag atagcttcag aatgtttggg atacaaatag gtctggctaa atataagtgt 420  
 tagctttctc aagcatctaa atgctggcgg gcttttaaaa aaccagggct ttaaggagaa 480  
 aacacctgct ctgtggtttt gtagcagata tgaagtattc aaatttctta ataaatagaa 540  
 aaagaaatat ataacagaaa caggttgcac ttgtctttct cattaagcag gtggttagta 600  
 c 601  
 <210> 321  
 <211> 501  
 <212> ADN  
 5 <213> Homo sapiens  
 <400> 321  
 agctcataaa tacgaacctg tgtcatgcaa ctcatatatg actgtgttca aaataatgtg 60  
 tattagactg taaaaagatt ttaatatttt aaataacttt cctgcatttg toggtttcag 120  
 caggaaagt atttttaata acttccctgt atttgttggg ttcagtatta attaactca 180  
 ttaatgctaa actttgtgat cctaggttaa aaaacatatt caagatagct tcagaatggt 240  
 tggatatacaa rtaggctctgg ctaaataataa gtgtagctt tctcaagcat ctaaagtctg 300  
 gcgggctttt aaaaaaccag ggctttaagg agaaaacacc tgctctgtgg tttttagca 360  
 gatatgaagt attcaaatth ctttaataaat agaaaaagaa atatataaca gaaacaggtt 420  
 gcacttgtct ttctcattaa gcagggtggt agtaccatta tttgcattct catagcctta 480  
 atatacattt tccttctcta g 501  
 <210> 322  
 <211> 601  
 10 <212> ADN  
 <213> Homo sapiens  
 <400> 322  
 ttttagtth ctactttagt gtcttagtgc tttctcgata tgggagaatt catgtcctcc 60  
 attcagaagt atgcactaag taagaggtat catgtctggt tcttgattag gtactaatct 120  
 tgaaatatta tcctacaata ggtagagca cgtatatctc ctgataatat attgaatatg 180  
 atagatttaa ataattggtt aactaaatac taaagcaaat tgctgcacgt atcatttatt 240  
 attcattgtg tagaaagtgc ctgactcagt gtttggaat tgtctgactt ttcctcatat 300  
 rtagtgtggt ttcatgttat tgtatataag acctgacatg aactctgtht acaataatct 360  
 ccagtgcca taaagacat aataaataat ataaccaatt ggtttcttht tgctgtcatt 420  
 tattagggca tatggcatta gtggaggatt accttgtatt acctatagtg cttagagtat 480  
 gaatcacaca tgcacctga aggaaaagag gtgcaatgta ataagaaacc agatattgaa 540  
 aatgcaagtt ttgttatgtht attctgggta tgttaacctt tattctgtcc ctccatatgc 600  
 15 a 601

ES 2 909 841 T3

<210> 323  
 <211> 501  
 <212> ADN  
 <213> Homo sapiens  
 5 <400> 323  
 aagaggtatc atgtctgggtt cttgattagg tactaatctt gaaatactat cctacagtag 60  
 gttagagcac gtatatctcc tgataatata ttgaatatga tagatttaa taattggta 120  
 actaaatact aaagcaaatt gctgcacgta tcatttatta ttattgtgt agaaagtgcc 180  
 tgactcagtg tttggaaatt gtctgacttt tcctcatata tagtgtgggtt tcatgttatt 240  
 gtatataaga mctgacatga accctgttta caataatctc ccagtgccat aaagaccata 300  
 ataaataata taaccaattg gtttctttat gctgtcattt attagggcat atggcattag 360  
 tggaggatta ccttgtatta cccatagtgc ttagagtatg aatcacacat gcacctgaa 420  
 ggaaaagagg tgcaatgtaa taagaaacca gatattgaaa atgcaagttt tgttatgta 480  
 ttctgggtat gttaaccttt a 501  
 <210> 324  
 <211> 854  
 <212> ADN  
 10 <213> Homo sapiens  
 <400> 324  
 tttcagcact gagagccaga gtggaattgt ctccttcatt gccactgcct tcacgttttg 60  
 tgtgtcgtat ctgttttggt atcactgaga cccaagaacc cccgacttgc cgacatacta 120  
 tgtggccccg agagaggact tgagctctct gggtttcac attaccatca attaaataaa 180  
 caggacagta gcttcttctc tggattgta atttaaggct ctggataata catgtaaccg 240  
 ccttatgata gagcagaatt gtaagtaggc tcatggtaga atcgttcaat gacatttccc 300  
 tttcctttgg gagaacaga aattcacagc tctaattctt ttcctattaa tagttcctgr 360  
 ccattattcc agaactgtcc taaaggaatt ctttctcctt aaggacacca cctcccagga 420  
 gggatattaa agatttgcac aggccgggca cgggtggctca tgcttgtaat ccagcagtt 480  
 tgggaggcca agcggggtgg atcacttgtg ctgaggggtt caagaccggc ctggccaaca 540  
 tgggtgaaacc ctatctctac taaaaacaca aaagttagct gggcctggct atgcatgcct 600  
 gtaattccag ctactogga ggctgaggct ggagaatagc ttgaaccagg gaggtggaga 660  
 taacagtgag ctgagatgcc actatgacac tccagcctgg gtgacagagc aagactctct 720  
 ctcaaaaaaa aaaaaagatt tttatagtcc agtattcaac gttcatagta caccttctt 780  
 atcctagtaa atcttctttt atcaaggtat atgatcccat atagtagtta actcttactc 840  
 ttactttatg acaa 854  
 15 <210> 325  
 <211> 501  
 <212> ADN  
 <213> Homo sapiens  
 <400> 325

ES 2 909 841 T3

aaatacttac tattaatat gagaaactgt ggtgtttatac ggtaagatcc acgaaggaag 60  
 aagttttaa gaaaaaact ttaaccgtgg aaaaaaaaaa ctttaatgtc tattatcgaa 120  
 taggggccgt aattactttt gcaaaataaa aaaacaaaca agactagcta tagtgtaaat 180  
 gtaatctgta tgctttttaa tgaacaatt aagtaggttg cccatttaca attagcctga 240  
 ttttctcctg ygtggtatta tgtgtactta acaacaggac ccagtggaaa ttcactcatt 300  
 taacaaagtc tgcctacatg gtttcaaata tgggcctaac ttgaaaattc agtcataatt 360  
 aaatctaagg actaaaaca atctgtataa aaagattctg ctaaataagg gaaaattcaa 420  
 gtctagggct acattctgaa agatattgaa gtagaacctc tgcagcaaga ctaggcttgg 480  
 aaagtgcggg gaggagggaa a 501  
 <210> 326  
 <211> 601  
 <212> ADN  
 5 <213> Homo sapiens  
 <400> 326  
 ccacatcaga aacatgagga aattctacat ggtaaaaaca gcaacaacca aaaaatactt 60  
 aaagtcaaca aaccaggaaa agacatctct gaatatagga atgccaacc tttaacacaa 120  
 taaaacacag attatatttc agaaggctat attatatgtg tataccaaca tcaatatgtc 180  
 cagagtagct gcacagagtt ccatatttta gtctttataa gttcccctcc tcaccctact 240  
 cagtaggcac tttgtgtcta gaaacttctg tgtcaacagt tttccctctc tctggaattc 300  
 mtcaggacag aagtgattgg tgtggtgaa gagggttgtg ctaagagtga agttatatga 360  
 aagtaggatg gaggttagca agtagttaa gtccagaaag gcaataaggt gttaaggaag 420  
 aacttttcca ttttacaggt ctgagcaagc aggaaatcaa ctctacaaac tttgaaactt 480  
 ggtaaatatg aaaacattct caataccatt tgtcatttaa taaatacaaa ttatactatt 540  
 ttactgcttg catctagaag tttgtcaaag atctcgtctt aattattcat tgtgtcggcg 600  
 a 601  
 10 <210> 327  
 <211> 601  
 <212> ADN  
 <213> Homo sapiens  
 <400> 327  
 gacgagatct ttgacaaact tctagatgca agcagtaaaa tagtataatt tgtatttatt 60  
 aaatgacaaa tggatttgag aatgttttca tatttaccac gtttcaaagt ttgtagagtt 120  
 gatttcctgc ttgctcagac ctgtaaaatg gaaaagtctc tccttaacac cttattgcct 180  
 ttctggactt taactacttg ctaacctcca tcctactttc atataacttc actcttagca 240  
 caaccctctt ccaccacacc aatcacttct gtcctgatga attccagaga gagggaaaac 300  
 ygttgacaca gaagtttcta gacacaaagt gcctactgag tagggtgagg aggggaactt 360  
 ataaagacta aaatatggaa ctctgtgcag ctactctgga catattgatg ttggtataca 420  
 catataatat agccttctga aatataatct gtgttttatt gtgttaaagg tttggcattc 480  
 ctatattcag agatgtcttt tcctggtttg ttgactttaa gtatTTTTTg gttgttctg 540  
 tttttaccat gtagaatttc ctcatgttcc tgatgtggaa agtataagaa tatcagccag 600  
 15 a 601

ES 2 909 841 T3

<210> 328  
 <211> 811  
 <212> ADN  
 <213> Homo sapiens  
 5 <400> 328  
 taaataatct ctaattagta taatgggtgt tcttagtgca gtgggtactt ttaaagtgct 60  
 ttgtggcttt tgatgaaaat tgtcttagta tttaaaactt tttcttacc aattttttgt 120  
 tcccatcgaa ttagcaatgc tgtaaagaaa ggcatcttat tccatttttt gttgctataa 180  
 aggaataactt gaggctgggt 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 gatacaaaaa 480  
 cctcccatta ggccccacct ccaaattgta tcccattggg gatcaaattt caaaaagaga 540  
 tttggaagga acaaacaaac catatctaag ccatagtaaa aggaatggct tttcaaaggt 600  
 aaaatttact ragtgtatta atattttacc aatttccagc caggagagta tgaatgttgc 660  
 attattacat tgctttgaaa caaagcatta gtcttaattc agaagtttaa attcagatgt 720  
 taacgttgca tatttaataa tgcacaacca gtactaaaat cctcattgaa atgacaaata 780  
 attttatttc gaatccctta tagaggttca c 811  
 10 <210> 329  
 <211> 811  
 <212> ADN  
 <213> Homo sapiens  
 <400> 329  
 tgtcttagta tttaaaactt tttcttacc aattttttgt tcccatcgaa ttagcaatgc 60  
 tgtaaagaaa ggcatcttat tccatttttt gttgctataa aggaataactt gaggctgggt 120  
 aatttataaa gatgaaaagt ttatttggct cgcaattctg gatggctgga aggttaagta 180  
 ctgggccaca gcatctggtg ggggcctcga gctgcttcta gtcataatgg aaggtgaagg 240  
 gtgtaaagat catgtgacaa gggaggaaag aagagaagga aggaggtgct ggttctttct 300  
 atcaaccaat tcgcaagaga actaatagag aaagaactca cttagccctg tgggaacaca 360  
 ttaatctatt cataagggat ctggctgtat gatacaaaaa cctcccatta ggccccacct 420  
 ccaaattgta tcccattggg gatcaaattt caaaaagaga tttggaagga acaaacaaac 480  
 catatctaag ccatagtaaa aggaatggct tttcaaaggt aaaatttact aagtgtatta 540  
 atattttacc aatttccagc caggagagta tgaatgttgc attattacat tgctttgaaa 600  
 caaagcatta ktcttaattc agaagtttaa attcagatgt taacgttgca tatttaataa 660  
 tgcacaacca gtactaaaat cctcattgaa atgacaaata attttatttc gaatccctta 720  
 tagaggttca caatgtttta acaatgtagt tttgactaaa tagaagtagt caaacctgt 780  
 cagattggaa atagtattta taaaacataa a 811  
 15 <210> 330  
 <211> 601  
 <212> ADN  
 <213> Homo sapiens

ES 2 909 841 T3

	<400> 330		
	gctcatcaat tttgacttaa gaaaattcta gcaacattta tagatthttgc caaaattcag	60	
	cttcttccca aatcaatcta taagaaggct cttccttaaa cataatthttt atatctatga	120	
	actgcactag catttactat atatthtttat cactctcacc attactggat aataaataaa	180	
	agctcattaa aagagttaac aaaacatatt tathtttaggc atcctgaaaa aaagattcaa	240	
	thtttattatc atthttctaaa taagtattga agaaaaggaga atthtaatta cttcatatac	300	
	stgataaagg aaaacatatg caaggcaaat aaacatctta gatcatgaca tataaaaataa	360	
	tagattatta ctaaagatta aaatactthtc ttaagaatta aagcaattct aaaagcaata	420	
	gtaaataaca thttthttctag tgatcagaca ctggatacta tghthtgagat agacagtgaa	480	
	thgggaatgt tghthttacag aagctcctac cthgcaagga caggcaagth taaatgtcag	540	
	ctagaaaact atctthgagth thcagtaatg taagatththc ctattcaatt tcacactthta	600	
	a	601	
5	<210> 331		
	<211> 601		
	<212> ADN		
	<213> Homo sapiens		
	<400> 331		
	agaaaattct agcaacatth atagatththg ccaaaattca gctthttccc aatcaatct	60	
	ataagaaggc thttccttaa acataatthtt tatatctatg aactgcacta gcattthacta	120	
	tataththta thactctcac cactactgga taataaataa aagctcatta aaagagthaa	180	
	caaaacatat thaththtagg catcctgaaa aaaagattca atthttattat cattthctaca	240	
	ataagtattg aagaaaggag aaththaaatt actthcatata cctgataaag gaaaacatat	300	
	rcaaggcaaa taaacatctt agatcatgac atataaaaata atagattatt actaaagatt	360	
	aaaatactth cthtaagaatt aaagcaatth taaaagcaat agthaaataac atthctthcta	420	
	gtgatcagac actggatact atghthtgaga tagacagtga atthgggaatg thghththaca	480	
	gaagctccta cctthgcaagg acaggcaagth thaaatgtca gctagaaaac thctthgagth	540	
	ththcagthaat gthaaagatthth cctattthcaat thcacactthth aaathththata thathathaaa	600	
10	a	601	
	<210> 332		
	<211> 1110		
	<212> ADN		
	<213> Homo sapiens		
15	<400> 332		

ES 2 909 841 T3

tgtagaagtt cttatcactt cctggccttt tggctaagat caagtgtgaa atgtagaagt 60  
 tcctctaagc tttacttccc tcaaaaacta gttttatcct gtcagcagga ttcacttaaa 120  
 aagacaaatt cagattatga attttttct tttttacagg gtctgctctg ttgcccaggc 180  
 tggagtgcag aggacaatc tgggctcact gcagcctccg cctcctgggt tcaagcaatt 240  
 ctcttgctc agcctcccga gtaactggga ttacaggcat gtgccaccac ccagctaatt 300  
 tttgtatfff tagtagagat ggggtttcac cacattggtc aggctgggtc cgaactgctg 360  
 gcctcaagtg atccacttgc ctgggctcc caaagtgcag agattacagg tgtgagccac 420  
 cgtgcccagc ctcataaccg tttcaactac tttttcactt gacaagcaga tgtgaagtta 480  
 acaaagtac ccatatftga aataaagata gtatattcct ggggyaggca gaggcagttg 540  
 aggatcatga aataactatg ttggcatagt tatttaggtg ttgatactgt tattatgcca 600  
 ttgaaagtta aacagagaac cctctgggta catgtttttat accaatgcac actatcttat 660  
 tagtccctct cataatgtgc agtcatcatt actgttacgg gttgaggtgt ccccatcctc 720  
 tatgggacac ctctatggtg aagtctcaga ttccctagaa tctcagaatg tgaccttggt 780  
 tggaaacaga tttgctacag acgcaattag ttgagatgcg cttatatggg taggtcctaa 840  
 ttcagtgact ggtgtcctta aaaaaatgga aatgtacaca cggtggtaga catgcataga 900  
 ggggaagagag atggagaaaa tggtcaccta caagccaaag acaggggtct ggagcagatc 960  
 cttccctcac agccctcaga aggaaccaat cttgccaata ccttgatfff ggacttccac 1020  
 ctccagaact ataacacatt tctgttcttc aagcaatttg tagccatttg ttacagctaa 1080  
 tacaatcaca catagaaatg acttgtaaat 1110  
 <210> 333  
 <211> 691  
 <212> ADN  
 <213> Homo sapiens  
 <400> 333  
 taaaacatgt acccagaggg ttctctgttt aactttcaat ggcataataa cagtatcaac 60  
 acctaaataa ctatgccaac atagttatft catgatcctc aactgcctct gcctacccca 120  
 ggaatatact atctftatft caaatatggg tgactttggt aacttcacat ctgcttgcca 180  
 agtgaaaaag tagttgaaac rgttatgagg ctgggcacgg tggctcacac ctgtaatctc 240  
 tgcactttgg gaggccgagg caagtggatc acttgaggcc agcagttcga gaccagcctg 300  
 accaatgtgg tgaaacccca tctctactaa aaatacaaaa attagctggg tgggtggcaca 360  
 tgcctgtaat cccagttact cgggaggctg aggcaagaga attgcttgaa cccaggaggc 420  
 ggaggctgca gtgagccgag attgtgcctc tgcactccag cctgggcaac agagcagacc 480  
 ctgtaaaaaa gaaaaaatt cataatctga atftgtctft ttaagtgaat cctgctgaca 540  
 agataaaact agtftttgag ggaagtaaag cttagaggaa cttctacatt tcacacttga 600  
 tcttagccaa aaggccagga agtgataaga acttctacat ttaagttat tcacaagata 660  
 actattaatg aacctgaaat agtttgtaaa g 691  
 <210> 334  
 <211> 640  
 <212> ADN  
 <213> Homo sapiens

5

10

ES 2 909 841 T3

	<400> 334					
	aaaccttttt	cctgttttac	tattactaaa	ggtggcacia	cagcaacctc	aacaactttg 60
	caccatgccca	acactgatgt	ttacaccag	cacagcattt	ttggtctcta	tttttattct 120
	cctctgaatg	taatgaggat	tcctagatgg	ctagccaatt	cgaatattta	aggcaactga 180
	aagttagaat	gtttctgaaa	catagtgttg	ttgccagaga	gtacgaaagt	tttcaagaat 240
	atcgggcaat	tctgaaagta	caaagaagcc	agattaaatg	aaataacact	ggcgaagttt 300
	tagcaagggtg	actctcatat	aatgatcatt	atcattacca	cagttaaaag	aaaagagttg 360
	tttatgaaag	gccatgtgtc	tgcaatgaaa	ctcaaaagag	aaaagttaac	aggtgcaara 420
	ggtagtttta	ttataaaag	agggtaggca	acaagaatat	gtttaatttt	tcttcctttt 480
	catgagtaag	gacaagagtt	tcatatatgt	gaatattttt	atttaatttt	aagtagaaat 540
	ctgtttttta	aatatgggta	tatgcttatt	tgtgtaagtg	taagaaacag	aagtaagtac 600
	agcaaaccag	aataggcca	aacactcctg	agcataattt		640
5	<210> 335					
	<211> 919					
	<212> ADN					
	<213> Homo sapiens					
	<400> 335					
	tacaccagc	acagcatttt	tggtctctat	ttttattctc	ctctgaatgt	aatgaggatt 60
	cctagatggc	tagccaattc	gaatatttaa	ggcaactgaa	agttagaatg	tttctgaaac 120
	atagtgttgt	tgccagagag	tacgaaagtt	ttcaagaata	tcgggcaatt	ctgaaagtac 180
	aaagaagcca	gattaaatga	aataacactg	gcgaagtttt	agcaagggtg	ctctcatata 240
	atgatcatta	tcattaccac	agttaaaaga	aaagagttgt	ttatgaaagg	ccatgtgtct 300
	gcaatgaaac	tcaaaagaga	aaagttaaca	ggtgcaaaag	gtagttttat	tataaaagga 360
	gggtaggcaa	caagaatatg	tttaattttt	cttccttttc	atgagtaagg	acaagagtkt 420
	catatatgtg	aatattttta	tttaatttta	agtagaatc	tgttttttaa	atatgggtat 480
	atgcttattt	gtgtaagtgt	aagaaacaga	agtaagtaca	gcaaaccaga	aataggccaa 540
	acactcctga	gcataatttt	acttggtaga	ttattcctga	aacttaagga	atcatctttg 600
	aactcttttc	ctcacttgac	ttccaggatt	caccatgcac	ttgtgatttt	cctttcattt 660
	cactctccgt	tcctcctcag	tctttttttc	tccccaggt	cttttttggt	catcttaaac 720
	tctaaatttt	agaatatccc	aggggtctgc	cttcggcctt	ctcttttata	tctacactgg 780
	cctcatacat	aatcttaacc	aagtcattat	tttaaatacc	tacaatatac	tgaaaacttc 840
	taaatttgta	ttttaattct	tgacttcttc	catacagtct	agatttgtat	gtccataggc 900
10	tgacatcatt	ggctgatac				919
	<210> 336					
	<211> 1001					
	<212> ADN					
	<213> Homo sapiens					
15	<400> 336					
	ttactaaata	ttctccaaca	aatatatact	tagtatatac	tattagtgat	gcatgctttc 60

ES 2 909 841 T3

aaatatttgg actatatcaa tgaatgaaac aaaaaattat ttgcccttaa ggagcttaga 120  
 ttctaacaga tggattcaga tgatTTTTat gccttatttc gtaggtttaa aagagcaatg 180  
 gggaaaaggg aagaagagag ggattgaaaa tattgagaag gttgggagac ttagcaatTT 240  
 taagtaaggt agtgagggta ggttttattg gcaaagtgat ttttcagcag agactgggaa 300  
 agatgaacgt ggtatcctgg aggaaagcct cccaggcaga gttaagctgc taacaaaagt 360  
 gcccttaggc tggagtgggc ttgtttgatt aaggaacaaa gaggtcagca tggttgcaact 420  
 agagagaaaa aatcagatgg cgtaaggaga tgaaatcaga aagatacgag gctaggcaaa 480  
 ggggtactct atgtaatgaa yatgacctgg cagtactgac atctcctgag ggactgttag 540  
 aagtgcagac tcttgtatct tttctcaagt ctatgaaatc tagacttcat ttaacaaga 600  
 tgacccgata ttacataca cattaagtt ccagaagcac tgatataca cattgtaaga 660  
 tgcacacagga cttcaattct ttttctggtt tttagaggca gtcctttggg gtgttttggg 720  
 tagagtataa tgacctgaaa tatctaggat cactctagct actatcttga ggaaagagtg 780  
 caataaggcg gaacagtcca gaggcaatgg tggcttcta aatgaaagac acacagcaact 840  
 caaacaggc agttgaggag ggatgggaaag aagttgtcaa attctagaca tattttaaag 900  
 gtagtgtcca gagaatttcc ttagatgcgt aggaacatgg aggataggac atagggtgga 960  
 aataaacgaa ataagaaac tgaagctgat tctgacattt t 1001

<210> 337

<211> 1576

<212> ADN

<213> Homo sapiens

<400> 337

atacctttta agtgacatcc tagtgaatct ccatttgtca cgagacctca agctttccag 60  
 ttctggcaca aagtgattac tcataccatc acttcaaaat gatgattatc ttcatttatt 120  
 ttagttatat tgaacaaaat atacatttaa aaaatctaata tactaaatat tctccaacaa 180  
 atatatactt agtatatact attagtgatg catgctttca aatatttggga ctatatcaat 240  
 gaatgaaaca aaaaattatt tgccttaag gagcttagat tctaacagat ggattcagat 300  
 gatttttatg ccttatttctg taggtttaaa agagcaatgg ggaaaaggga agaagagagg 360  
 gattgaaaat attgagaagg ttgggagact tagcaatTTT aagtaaggta gtgagggtag 420  
 gttttattgg caaagtgatt tttcagcaga gactgggaaa gatgaacgtg gtatcctgga 480  
 ggaaagcctc ccaggcagag ttaagctgct aacaaaagtg cccttaggct ggagtgggct 540  
 tgtttgatta aggaacaaag aggtcagcat ggttgcaacta gagagaaaa atcagatggc 600  
 gtaaggagat gaaatcagaa agatacgagg ctaggcaaag ggttactcta tgtaatgaac 660  
 atgacctggc agtactgaca tctcctgagg gactgttaga agtgcagact cttgtatctt 720

5

ES 2 909 841 T3

tctcaartc tatgaaatct agacttcatt ttaacaagat gacccgatat ttacatacac	780
attaaagttc cagaagcact gatataacac attgtaagat cgcacaggac ttcaattcctt	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
tagatgctga ggaacatgga ggataggaca tagggtggaa ataaacgaaa taaagaaact	1140
gaagctgatt ctgacatttt agacctaaaa tctcaactaa aagttgcaa gatgggaaaa	1200
actaggtgca tcttgtttg tgagtggaaa tcagccttgt gaattaagac ttaaactgat	1260
gtctttaatc ccgtagaaat accatgaagg cagtagaaga tggctaaaga gaggtctaga	1320
ctgtaggtagc aaatttaaaa gtcacttgca tttggatgct taaagtcagg atattgtgaa	1380
gtcaacagag gaataaataa atgcagagag gggaaagaaa aggcccatag actgagccat	1440
tgtctggttt attcacatat tagtatatat tttcttaaag atgtttgcta tataataatg	1500
agttacctaa agtgtgactt ttctaaatth atggggaatt ttctacattg tgttatggca	1560
ctactaaaa taataa	1576
<210> 338	
<211> 1275	
<212> ADN	
<213> Homo sapiens	
<400> 338	
gtaaaactaa ttataattaa aatcaaaata tttactgaac ctacttactc ctataatttg	60
cgttgctggt taaaaccag ctataaaaat tttgatcaaa aatttttatt ttgtaaataga	120
tctgacacag cataaatggt aatcacattt ctttatttta tttgcagatt aatttgagta	180
atttgaaaaa ttattaatgt tacttaatta ctctcaacac cttacagtgt ctctgtgtaag	240
cactattggt gatactgaat ttaagttaca ttttaacaact atcagaaaat agtttttaa	300
gtaaaaatta tgatttgag tttaccaact aaatcttggt agctttcact gcctctattg	360
agaagagcag cagttcttat ctctctcctt tttcttcttt aattaacaag agattatttg	420
tatcatagcc ataaaatcag ttcaggtatt acatgaacga caccctgac tgcaatggtg	480
tagtttattg tattagtcca tttcatgct gctgataaag acatacataa gactgggtaa	540
tttataaaga aatagaagtt taacggactc acagttccat gtggctgggg aagcctcaca	600
atcatgatcg aaggcaaaag gcacatctta catggcaaca ggcaagagag aatgagagcc	660
aagtgaaagg agaaaccct tataaaacct tcagacctca tgagacttat tcactaccac	720
aagaacagta tgtgagaac agtcccatga tccagttatc tcccactggg tccctcccac	780

5

ES 2 909 841 T3

	cacacaagg	aattatggga	actgcaattc	aagatgaaat	gtgggtggaa	gcacaacgga	840
	actatatcat	gatcaaagca	ttattgtttt	ctctgataag	ctgatctaga	aagtgctgct	900
	tgtgatcagc	tttggtgacc	atgatcagtg	aaatggttaa	ggaaatctac	agattttgta	960
	ggtttgcc	ttgacagacg	accggtatct	gtttctcttt	tcatgatgaa	gtatctaaca	1020
	aagctctgtc	caaaattttg	aatttctcgt	taaawgcatc	atgattatag	aacagaggtt	1080
	acaatcaatt	atcagtcac	acaatcactc	tcatcagtc	ttaaggtgca	tacctggtgt	1140
	tccagttatt	cagtgtggta	taacaaacta	cctggaactt	aatggcttga	aatagtcacc	1200
	attacattat	gattgtccat	tctctgcatc	aataattagg	atttggcaaa	gagggaatgg	1260
	tttgtttaca	gacag					1275
	<210>	339					
	<211>	1275					
5	<212>	ADN					
	<213>	Homo sapiens					
	<400>	339					
	gtaaaactaa	ttataattaa	aatcaaaata	tttactgaac	ctacttactc	ctataatttg	60
	cgttgctggt	taaaaccag	ctataaaaat	ttgatcaaa	aatttttatt	ttgtaaatga	120
	tctgacacag	cataaatggt	aatcacattt	ctttatttta	tttgagatt	aatttgagta	180
	atgtgaaaa	ttattaatgt	tacttaatta	ctctcaacac	cttacagtg	ctcctgtaag	240
	cactattggt	gatactgaat	ttaagttaca	tttaacaact	atcagaaaat	agtttttaa	300
	gtaaaaatta	tgatttgag	tttaccact	aaatcttgtt	agctttcact	gcctctattg	360
	agaagagcag	cagttcttat	cttctcctt	tttctcttt	aattaacaag	agattatttg	420
	tatcatagcc	ataaaatcag	ttcaggtatt	acatgaacga	caccctgac	tgcaatggtg	480
	tagtttattg	tattagtcca	ttttcatgct	gctgataaag	acatacataa	gactgggtaa	540
	tttataaaga	aatagaagtt	taacggactc	acagttccat	gtggctggg	aagcctcaca	600
	atcatgatcg	aaggcaaaag	gcacatctta	catggcaaca	ggcaagagag	aatgagagcc	660
	aagtgaaagg	agaaaccct	tataaaacct	tcagacctca	tgagacttat	tcactaccac	720
	aagaacagta	tgtgagaaac	agtcccatga	tccagttatc	tcccactggg	tccctcccac	780
	cacacaagg	aattatggga	actgcaattc	aagatgaaat	gtgggtggaa	gcacaacgga	840
	actatatcat	gatcaaagca	ttattgtttt	ctctgataag	ctgatctaga	aagtgctgct	900
	tgtgatcagc	tttggtgacc	atgatcagtg	aaatggttaa	ggaaatctac	agattttgta	960
	ggtttgcc	ttgacagacg	accggtatct	gtttctcttt	tcatgatgaa	gtatctaaca	1020
	aagctctgtc	caaaattttg	aatttctcgt	taaatgcatc	atgattatag	aacagaggtt	1080
	acaatcaatt	atcagtcac	acaatcactc	tcatcagtc	ttaaggtgcr	tacctggtgt	1140
	tccagttatt	cagtgtggta	taacaaacta	cctggaactt	aatggcttga	aatagtcacc	1200
	attacattat	gattgtccat	tctctgcatc	aataattagg	atttggcaaa	gagggaatgg	1260
	tttgtttaca	gacag					1275
10	<210>	340					
	<211>	1001					
	<212>	ADN					
	<213>	Homo sapiens					

ES 2 909 841 T3

<400> 340  
gaaacaaaa attgcttttt atatattgat atttttgcac ggattttctta ggattttcta 60  
tgtacatgac catgtcatct gcaaatgaaa tagttttatt tctttatcaa tccggatgaa 120  
tttattaanaa ttatcttgcc taatttccca aatagggcct ccatgttgaa cataagtggg 180  
ggcaagggg atctgttgct aatctcagtg gatgatattc agtgttttac aatgatcttc 240  
gacagctctg gctgttaaat tatcatagtc tgtatggcct aaacaaacaa aatacttatg 300  
attatggggg aggctgggat atccaagatc aagttgctgg caggtctagc aacctgccac 360  
tgggaagccc tgcttcccag ttttcagatg gccaccttct tatagtatct tcaccaaga 420  
tagggcagag agagcaagca agctctctac cttctcatat aagggcacta atcccaccat 480  
gaaggcgcca ctgtcatgac stgattatgt cacaaagacc ccggggcaaa tattaccact 540  
gtgaggagta cagttttagc atgtgaattt tggaagaaca caaacattta gtacagagtg 600  
actattaagt atgttattaa ctatggagtt tttgtaggca ttttttaaca cattgagaaa 660  
gtttcctcta ttcctacttt tgttgagaag tttttatgat gacaaggcat tacattttat 720  
ccaatgactt ttctgtgtgt attgagatga ctgatttggc ctgccaatth aatccattg 780  
ttgattctct ctaggatttt ttttatttca gttattaaat ttttcaacag gagaattact 840  
gtcttgttct tttttttgta atttctgtcc cttactggg attccatatt taataaggca 900  
tcataatagt actcttcttt agtttcttaa agatggtttt ctttagtttt taacatattt 960  
atgtctattt agaagtcttt gttaagtctg acatctgagc t 1001

<210> 341  
<211> 1001  
<212> ADN  
<213> Homo sapiens

5

<400> 341  
ggattttctta ggattttcta tgtacatgac catgtcatct gcaaatgaaa tagttttatt 60  
tctttatcaa tccggatgaa tttattaanaa ttatcttgcc taatttccca aatagggcct 120  
ccatgttgaa cataagtggg ggcaagggg atctgttgct aatctcagtg gatgatattc 180  
agtgttttac aatgatcttc gacagctctg gctgttaaat tatcatagtc tgtatggcct 240  
aaacaaacaa aatacttatg attatggggg aggctgggat atccaagatc aagttgctgg 300

ES 2 909 841 T3

caggcttagc aacctgccac tgggaagccc tgcttcccag ttttcagatg gccaccttct 360  
 tatagtatct tcaccaaaaga tagggcagag agagcaagca agctctctac cttctcatat 420  
 aagggcacta atcccacat gaaggcgcca ctgtcatgac ctgattatgt cacaaagacc 480  
 cgggggcaaa tattaccact stgaggagta cagttttagc atgtgaattt tggagaaca 540  
 caaacattta gtacagagtg actattaagt atgttattaa ctatggagt tttgtaggca 600  
 ttttttaaca cattgagaaa gtttcctcta ttcctacttt tgttgagaag tttttatgat 660  
 gacaaggcat tacatthtat ccaatgactt ttctgtgtgt attgagatga ctgatttgtt 720  
 ctgccaatth aaatccattg ttgattctct ctaggattht ttttatttca gttattaaat 780  
 ttttcaacag gagaattact gtcttgttct tttttttgta atttctgtcc ccttactggt 840  
 attccatatt taataaggca tcataatagt actcttcttt agtttcttaa agatggtttt 900  
 ctttagtht taacatatht atgtctatht agaagtcttt gttaatgtctg acatctgagc 960  
 tctctcaaag tttctgctga ttttttttt cctatgtht g 1001  
 <210> 342  
 <211> 701  
 <212> ADN  
 <213> Homo sapiens  
 <400> 342  
 ggaaaccctg gcctcttgat cacactttcc tggagthtag tcccctctgc aatatgtacc 60  
 tgggagthcat aagaaatgcc agttacaaaa acttctctgta cagatathcct agcactcaac 120  
 tggaaaccgg ggagagthcac aattctgtct ttccagccat atgtaactga aatggagatc 180  
 ttttaccct gagccagggg tgatgggaaa gggagctggt catggctcaa tgtthtagct 240  
 tttcttggtc ttcaagatth catagacatt cttaaataca tgtthctthc aatgaagtht 300  
 gcccttagga caatthcacag ctacathagg tactthttaa ataathctth tgaccatccg 360  
 tggthattthc attgaagaaa atctatagag cacctcagcc atcathccag aagthgactat 420  
 cctcctcagth aatggtthctt attctathth taaathatcat tgatgthgaa cathctathth 480  
 cactathctc tcatththatt rthtatgggaa attathataca gthctccaga tththaaagc 540  
 cthgthaaaca tgtththaat cacacaaata thctthctgtg ggaaathgac agthaththtag 600  
 tgtgcaacaa thathathgaa ctathththca aactthataaa cgaagthgaaa thcthaataa 660  
 aathcaththt caaacacaaa aaththgagcc agaathagga a 701  
 <210> 343  
 <211> 701  
 <212> ADN  
 <213> Homo sapiens  
 <400> 343

5

10

ES 2 909 841 T3

aatgccagtt acaaaaactt cctgtacaga tctcctagca ctcaactgga aaccggggag 60  
 agtcacaatt ctgtctttcc agccatatgt aactgaaatg gagatctttt caccctgagc 120  
 caggggtgat gggaaagga gctgggtcatg gctcaatggt tagccttttc ttggcttca 180  
 agatttcata gacattctta aatacatggt tctttcaatg aagtttgccc ttaggacaat 240  
 tcacagctac attaggtact ttttaataa tacttttgac catccgtggt tatttcattg 300  
 aagaaaatct atagagcacc tcagccatca ttccagaagt gactatcctc cttagtaatg 360  
 gttcttattc taatttttaa tatcattgat gtagaacatt ctatttcact attccttcat 420  
 tttattatta tgggaaatta tatacagttc tccagatfff taaagccttg ctaacatggt 480  
 ttaagtcaca caaatattct yctgtggaa aatgacagta atttagtgtg caacaattat 540  
 atagaactat ttttcaaact tataaacgaa gtgaaattct aaataaaatc atttatcaaa 600  
 cacaaaaatt tgagccagaa taaggaatgt aaattacaat ttaaacacag attataaact 660  
 atcttacttt taaaatgtta aaattcctaa cttgtttgaa a 701

<210> 344  
 <211> 768  
 <212> ADN

5

<213> Homo sapiens

<400> 344  
 ctaaaatcta ccattatatg atatccttcc caatacataa attaaaaaaaa aaaactgt 60  
 agaggaaaaa gcaatatttt gaaatgatat gcttttcttt gtttgtcttc aaacaattac 120  
 atcttcatca taatggttgt attagtctgt ttttactctg ctataaagaa ttgcctgaga 180  
 ctgagtaaca tataaagaaa aaagttttaa ttgaccacag tttcacaggc ttaataggaa 240  
 gcatgactgg gaaacttaga atcatggcag aagaggaagg ggaagcaagg atcttcttca 300  
 catggtagca ggagagagag cacaaagggg gacacgctac acactttcaa acaacgagat 360  
 ctctgagaa ctctatcggg agaacagcaa gagggaagt caccctatg attcaatcag 420  
 ctcccaccgg gcttctcccc tgacacatga ggaattaaa ttggatgaga gatttgggtg 480  
 gggacacaca gacaaacct atcaactgtc atggacttaa acaattgtct ttgaattgtc 540  
 tttttcata cttttatttg catcttlyca ctaaaagat gacacaaagt aatcctagtt 600  
 tacatttttt accatgtaat tccatattac ttttctctga aagtactta tttttaaact 660  
 tcaaagctct tcatacttat ggtttgatct gcacttacia ctggatctca gaaagattga 720  
 atttcccat cataccaagt tcatgtctct cactcttaat atttgctc 768

<210> 345  
 <211> 701  
 <212> ADN

10

<213> Homo sapiens

<400> 345

ES 2 909 841 T3

aaatgatatg cttttctttg tttgtcttca aacaattaca tcttcatcat aatggttgta 60  
 ttagtctggtt tttacactgc tataaagaat tgcctgagac tgagtaacat ataaagaaaa 120  
 aagttttaat tgaccacagt ttcacaggct taataggaag catgactggg aaacttagaa 180  
 tcatggcaga agaggaaggg gaagcaagga tcttcttcac atggtagcag gagagagagc 240  
 acaaaggggg acacgctaca cactttcaaa caacgagatc tcctgagaac tctatcggga 300  
 gaacagcaag agggaagttc acccctatga ttcaatcagc tcccaccggg cttctcccct 360  
 gacacatgag gaattacaat tggatgagag atttgggtgg ggacacacag acaaaccata 420  
 tcaactgtca tggacttaaa caattgtctt tgaattgtct tttttcatac ttttatttgc 480  
 atcttttcac taaaagatg rcacaaagta atcctagttt acatttttta ccatgtaatt 540  
 ccatattact ttttctgaa agttacttat ttttaaatct caaagctctt catacttatg 600  
 gtttgatctg cacttacaac tggatctcag aaagattgaa ttctcccatc ataccaagtt 660  
 catgtctctc actcttaata tttgttccca agacaacaat t 701  
 <210> 346  
 <211> 6758  
 <212> ADN  
 <213> Homo sapiens  
 <400> 346  
 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 agtatataca caagcccaaa 240  
 taagatactc caaagatggt tgataggttc tagatctcta gatatactgc tcaatgaaag 300  
 tgtccccctg aacaaagcca gtctgcaaag actggytgag atgatttttt ttaaattgca 360  
 agtctcagca acaacaaaaa tgacaagaca tgcacagaag caagaaaata taacacaatc 420  
 aaagaaaaaa aagccacaga aatcagtcct agagaaaacy gatctatgag ctgcctgaca 480  
 ataattataa aataactatc ataaaaatgc ccagtgagat ataagaaaac acagacaact 540  
 aatgaatca ggaaaatgat gcatgaacaa aatgggcata tcaacagaga tggaaatgac 600  
 aaagataaac aaacagaaat tttggagctt aaaaatacag taagtaaagt gaataattca 660  
 ctaaaaaatat 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

5

ES 2 909 841 T3

atctaaagag actcacacta agacatatta tcatcaaact gtcaaatca aagacaaaga 1020  
 attgtgaaat ctgccaagga aaagtgactc atcacacata agagatataa cataagattg 1080  
 tcacaggatt tctgaacaga cactttgcag gtcagagggga agtaggggtga catattccag 1140  
 gtgctgaaag aagaaaacac cctgccaacc aagaatatgg catccagaaa aactttccta 1200  
 gaagaatgaa ggagaaatth agactttccc aaataaacia aagctgaggg agttcattac 1260  
 taccagacct gctctgcaaa atgctaaga gaaacctca ggtgaaacia aaagatgcta 1320  
 gacagtaaca caaacact cataaataac ttcttcagta aaaataatac atcgaciaat 1380  
 atggtaacct gtattaatac tgggtgcacia attcactttc aaatthtata aataagaatt 1440  
 taaaggatga aaacatctaa aactaactat aaatctatat aatgaatata caatatataa 1500  
 aaaaatttgt gatcacata acataaaatg ggggaggtag agctgtatag gggtagagct 1560  
 tttgtatgca attgaaatta ccatcagttt aaactgaact gttataacat taagatgttt 1620  
 tatgtaattg caatggtaac tatattctat agaatatatt aaaaagaaaa agaaaatagg 1680  
 aagggaatca aagcatgtcc ttgtaaaaaa gtcaatgaaa gcaaaagaaa ggcagaaaaa 1740  
 gtgaaaagga ggaataaaaa gttataagac ataaaaaaa tgaaaatagt aatagtctctg 1800  
 ccatatcagt aattacatta aatataaatg gattaaactc cctaatcaaa tcatagattg 1860  
 gtttgaaga actaacttta caattaaaga cacacagctg acggtgaagg gagaaaaaa 1920  
 acttccatgc agtgaccaa atagaggagg gtggctgtat tactgtcaga caaataaaa 1980  
 ttttaagtcaa aaactgttac aagagtaaaa gaagggcatt atacagttaa aaaagtaaat 2040  
 tcgccaggca gacacaacia ttataaatat caatacacia aaataagagc tctaaatat 2100  
 atgcagcaaa cagacataat tgaagaaaga aataaatagc taaaatggta gaagacttta 2160  
 atacccccac ttacaataat gtataaaata acaagacaga atgtaaaaa aaatgtagag 2220  
 aatttgagca aactgtaga ccaattggac ctaataaata tactcagaat aatccatcca 2280  
 accaaagcag aaacagaata tacattcttt tcaagtacac atttgacatt ctctgggatt 2340  
 aactacatgt tatgcaacia acaagtctca acaatgttta aaagtctgat attacaciaa 2400  
 gtattgtttc tgatgacgat ggaaagaacc tagaagcaa tagcaaaaag aaaatagaaa 2460  
 atccacacat atgtggaaat taaactacat gcaattaagc aaagggcaa agaagaagaa 2520  
 gaaaaagaa aacaccgtga acaaaaaaa acaaaaaata cagcatatga aaatgcatgg 2580  
 gatgcagcaa aagtgatggt aagagaaatg tttatagtta taaatgcaaa ccttaaaaa 2640  
 gaagaaagaa acaaaaaata ctcaaattaa caactttaca agtcaagaag gtagagaaaa 2700  
 aagaaciaac tataccaaaa gctaacacag aaagaaaaa ataaagatta aaaaciaaaa 2760  
 caatttaaaa aatagcagaa ctaaaagttg gttctttgaa aagatcaaca gaattgacia 2820

ES 2 909 841 T3

tttcttagct acattaagaa aaatacaaga ctcaaataac acaaatcagt ggtgaaaggg 2880  
 ggtattataa ctgatgccac agaaatacaa aaggatcata agggactact acaaatgtga 2940  
 tgacaacaaa ttgagtaacc taggatacct tgataaatc caaaaaatgc acaatatact 3000  
 gaatcatgaa tacatgaccc ttataaatca agactaaatc ataaagaaat agaaaatatac 3060  
 aacagaccaa taattagtaa ggagaataaa ctagtaatca gaaacctccc acaaaagaaa 3120  
 agcttaggac caaatggcctt tactggagaa ttctaccaac cattaaaagg ataattaaga 3180  
 ccaatcttcc tcaaatctttt aaaacaaatg ttaaagagga ggaaactctt tcaatctcat 3240  
 tcataaggtc agcattatcc ttataccaaa accagacaaa gacactatta aaaaaactta 3300  
 gaccaatatac cctgatgaat ttcgatgcaa gaatcctcag caaaatacta tcaacaatt 3360  
 caacagcata cttaaagat tatatgctgt aatcaagatg catttattct ttgaatgcaa 3420  
 gtgtaattca acacataaaa ttcaatcaat gtaatacacc acattaacag aatgagagac 3480  
 aaaaaccaca taattatatac aactgatgca gaaaaaatc tgacacagtt caacaccttt 3540  
 tgtgataaaa acactcaaca aactaggaag agaaggaaac aactttaaca catcatatgc 3600  
 tcactgatga aaatctacaa gttctttata aaagatcagg aacaagacaa taatctgcat 3660  
 tgttaccact tctattatac gtagtattgg aagttctaatac 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 ctgcaggaat gtaaaaattt atcctaaaat tcatatagaa 4200  
 tctctaggta ccctgagggc caaacaattt tgagaaaaaa aaaagaacaa aattggagga 4260  
 ctcacacttc cagattacaa gaatatttac aaattacata tttacaaaaa aaattacaaa 4320  
 gccacaataa tcaaaacaac gtgggatttg cataaaggca gatatataga ccagtggaaat 4380  
 agtattgaga gtccagaaat aaaccttag 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 tcatgctaatac ttggcaacat tttgtgatgt gacacaaaaa gcagagtcaa 4680  
 taaaagcaaa aattagacag atggaaatcc atcatagttt ataacttttg gtcattaag 4740

ES 2 909 841 T3

aacagtcaac agagtgaaaa ggcaatctat aaaatggggg aaaaacagaa aatatgtgca 4800  
 aatcacagat atctgataggg ggattcatat ccagaataaa taaagaactc ctatatctca 4860  
 acaacaaaa atctaatacca atcaaaaaat gggccaaggg agtgaagata catttctcca 4920  
 aagatgttat acaaatggcc aggaagcata tgaaaagatg ttcaatgtca ctaatcatca 4980  
 gagaaatgca aatcaaaacc acagtgcaat atcacttcac attcattaga atggcttctg 5040  
 tcatgaacaa cagaaaataa caagtgttga tgagtgtgta gagaaattga gacctttata 5100  
 taatthttggc agaaattcaa aatgggtgcaa ccactataaa aatgatatg gaggtcctca 5160  
 aaaaattaaa aatagaacta ccatatgatc cacaatccca cctctgggta catattcaaa 5220  
 agaattgaaa gcaggggtgtt gaagatatat ttgcacactc tttatagcag cactgttcac 5280  
 aatagccaag agatgaaagt aaccctaaagg ttcatgaagc aatgaataaa caaaatatat 5340  
 tatgtacata gagtaaaata ctgtgcagct ttaaagagaa aggaaatcctt atactatgct 5400  
 acaacatgaa tggaacttta gggcattata gtaagtaaaa taagccagtt ttttttaag 5460  
 gacaaataaa cactatacga ttctacttaa gtatttaatg ttgtcaaatt tataaatata 5520  
 gaatgtagaa tagtgggttac cctgagctgg gggaaagggg caaaggggaa ttgttatthtt 5580  
 aatgggtata gtttcagttc tgcaaaatga aaaggttctg gaaatctggtt tcacaatggt 5640  
 gtaaataataa ttactctgaa attgtacact taaaatggt taagatgaca aatagagttg 5700  
 tgatgtcttc tthttgttatt atatagaaaa actthttcat atgataatag tctthttttt 5760  
 taagctgact ttgctgatat taatataatc cttccatttht tctthtaaat gctatatgct 5820  
 ttcacataat tthgtctthac gttgatgtat ttatacataa ggtgggtthc ttatagatac 5880  
 cacgttgtgt gtctthttta tctaagttga tagacttgcc tthttgttagg gtattthaaat 5940  
 aatthtatatt taatgtaatt attgatatag ttgagtgtgt tgattthttgt thtctatttg 6000  
 ctccatctgt tgttggthct cattattcct ctgthttctac cttctthttgt actaattatt 6060  
 atattthtatt atthttctac tcaactgttg gctthattagc cacattgctt thaaaattht 6120  
 taatgattgc tctagggtht ataataaaca aatgthtagc atthttctacc atcaaatatt 6180  
 thtacctat tcatgtatac thcaatthct thctthcccat cctthtgaact atatcttcat 6240  
 acattthtact ctacattthgt tataactcag tgctthtgaaa gtcaattatt thtgtcttht 6300  
 acagtcaatg atthtttaag agthtaacag tgaaaaaaa tggctthtcat cthttthccat 6360  
 tagattthcat actctthctg cctgaagaat thctthttaat agactthtga ctgctgggtct 6420  
 caggcaagaa atthctctcag cctthgtthg thtgaaaaac tgctthattac acctthgtth 6480  
 thgaaagata thttcactag gtatagaagt ctgggtthgac agthctcatt gthttgtcaca 6540  
 gcattthttaa gatgcccatt caattgtctt gtctthgtata atthttggatt agtctggtgt 6600  
 atthcttacc thtthtctc tctgtgcaat gctthcaacca tcccactthca ggctgcctth 6660  
 aagatgthtt cthttthcctt aatctthtagt thtttagctgg thgacagtga cgcatctaa 6720  
 tgtagtgtat gaggtthgctt thattgtcac tgtthgtg 6758  
 <210> 347  
 <211> 6758  
 <212> ADN  
 <213> Homo sapiens

ES 2 909 841 T3

<400> 347  
agagtgggcc attgttctga ctagtctggg gctcccaaaa gaactggtat ctgtctcacc 60  
tgactcagaa caatgataag gctgtagatc tttttggaag tctatgaaaa caggcacaat 120  
gaaggcagca tgttagagat ataattccac aggaagatgc caggtaaaac aaaagagaaa 180  
aagcaggaac aagctgatta ggaaatttgt gatgactaaa agtatataca caagcccaaa 240  
taagatactc caaagatggt tgataggttc tagatctcta gatatactgc tcaatgaaag 300  
tgtccccctg aacaaagcca gtctgcaaag actgggtgag atgatttttt ttaaagtca 360  
agtctcagca acaacaaaaa tgacaagaca tgcacagaag caagaaaata taacacaatc 420  
aaagaaaaaa aagccacaga aatcagtcct agagaaaact gatctatgag ctgcctgama 480  
ataattataa aataactatc ataaaaatgc ccagtggagat ataagaaaac acagacaact 540  
aatgaatca ggaaaatgat 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  
atctaaagag actcacacta agacatatta tcatcaaact gtcaaatca aagacaaaga 1020  
attgtgaaat ctgccaaagga aaagtgactc atcacacata agagatataa cataagattg 1080  
tcacaggatt tctgaacaga cactttgcag gtcagaggga agtagggtga catattccag 1140  
gtgctgaaag aagaaaacac cctgccaaac aagaatatgg catccagaaa aactttccta 1200  
gaagaatgaa ggagaaatth agactttccc aaataaacia aagctgaggg agttcattac 1260  
taccagacct gctctgcaa atgctaaaga gaaaccttca ggtgaaacia aaagatgcta 1320  
gacagtaaca caaaaccact cataaataac ttcttcagta aaaataatac atcgacaaat 1380  
atggtaacct gtattaatac tggtgcaaaa attcactttc aaattttata aataagaatt 1440  
taaaggatga aaacatctaa aactaactat aatctatat aatgaatata caatatataa 1500

ES 2 909 841 T3

aaaaatthgt 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 ggcagaaaaga 1740  
gtgaaaagga ggaataaaaa gttataagac ataaaaaaa tgaaaatagt aatagtcctg 1800  
ccatatcagt aattacatta aatataaatg gattaaactc cctaatacaa tcatagattg 1860  
gtttgcaaga actaacttta caattaaaga cacacagctg acggtgaagg gagaaaaaaa 1920  
acttccatgc agtgacccaa atagaggagg gtggctgtat tactgtcaga caaaataaaa 1980  
tttaagtcaa aaactgttac aagagtaaaa gaagggcatt atacagttaa aaaagtaaat 2040  
tcgccaggca gacacaacaa ttataaatat caatacataa aaataagagc tcctaaatat 2100  
atgcagcaaa cagacataat tgaagaaaga aataaatagc taaaatggtg gaagacttta 2160  
atacccccac ttacaataat gtataaaata acaagacaga atgtaataaa aaatgtagag 2220  
aatttgagca acactgtaga ccaattggac ctaataaata tactcagaat aatccatcca 2280  
accaaagcag aaacagaata 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 acaaaaaata cagcatatga aaatgcatgg 2580  
gatgcagcaa aagtgatggg aagagaaatg tttatagtta taaatgcaaa ccttaaaaaa 2640  
gaagaagaa aacaaaaata ctcaaataa caactttaca agtcaagaag gtagagaaaa 2700  
aagaacaaac tataccaaaa gctaacacag aaagaaaaga ataaagatta aaaacaaaaa 2760  
caatttaaaa aatagcagaa ctaaaagttg gttctttgaa aagatcaaca gaattgacaa 2820  
tttcttagct acattaagaa aaatacaaga ctcaaataac acaaatcagt ggtgaaaggg 2880  
ggattataa ctgatgccac agaaatacaa aaggatcata agggactact acaaatgtgta 2940  
tgacaacaaa ttgagtaacc taggatacct tgataaatc caaaaaatgc acaatatact 3000  
gaatcatgaa tacatgacct ttataaatca agactaaatc ataaagaaat agaaaatctc 3060  
aacagaccaa taattagtaa ggagaataaa ctagtaatca gaaacctccc aacaaagaaa 3120  
agcttaggac caaatggctt tactggagaa ttotaccaac cattaaaagg ataattaaga 3180  
ccaatcttcc tcaactttt aaaacaaatg ttaaagagga ggaaactctt tcaatctcat 3240  
tcataaggtc agcattatcc ttataccaaa accagacaaa gacactatta aaaaaactta 3300  
gaccaatctc cctgatgaat ttcgatgcaa gaatcctcag caaaatacta tcaacaatt 3360  
caacagcata cttaaatgat tatatgctgt aatcaagatg catttattct ttgaatgcaa 3420

ES 2 909 841 T3

gtgtaattca acacataaaa ttcaatcaat gtaatacacc acattaacag aatgagagac 3480  
 aaaaaccaca taattatatac aactgatgca gaaaaaaatc tgacacagtt caacaccttt 3540  
 tgtgataaaa aactcaaca 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 ctacaatat tgccaaaaga aaataaagga gacacaaata 4080  
 agtgatatgt ttttaatatg tccacccaaa gtgatcttca gattcaatga aatccctatc 4140  
 aaagttataa tggcattttt ctgcaggaat gtaaaaaattt atcctaaaat tcatatagaa 4200  
 tctctaggtg ccctgagggc caaacaattt tgagaaaaaa aaaagaacaa aattggagga 4260  
 ctcacacttc cagattacaa gaatatttac aaattacata tttacaaaaa aaattacaaa 4320  
 gccacaataa tcaaaacaac gtgggatttg cataaaggca gatatataga ccagtggaaat 4380  
 agtattgaga gtccagaaat aaacccttag gtatatcatc aaatgacatt tgacaaagtg 4440  
 ctggtaccac tcaatgggaa tgggacaatt tgttcaacaa atagagcaaa gaaaactaaa 4500  
 catccatgtg caaagaata aatctggacc cttatattac actatagaca aaattaattc 4560  
 aaaatggatt aaagatctaa atgaaagatc taaaactata aaactcctag gagaaaacag 4620  
 aggaaaaatt tcatgctaata ttggcaacat tttgtgatgt gacacccaaa gcagagtcaa 4680  
 taaaagcaaa aattagacag atggaatcc atcatagttt ataacttttg gtcattaag 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 cagaaaataa caagtgttga tgagtgtgta gagaaattga gacctttata 5100  
 taatthttggc agaaattcaa aatggtgcaa ccactataaa aatgatatg gaggtcctca 5160  
 aaaaattaaa aatagaacta ccatatgatc cacaatccca cctctgggta catattcaaa 5220  
 agaattgaaa gcaggggtgtt gaagatatat ttgcacactc tttatagcag cactgttcac 5280

ES 2 909 841 T3

aatagccaag agatgaaagt aacccaaagg ttcatgaagc aatgaataaa caaaatatat 5340  
tatgtacata gagtaaaata ctgtgcagct ttaaagagaa aggaaatctt atactatgct 5400  
acaacatgaa tggaaacttta gggcattata gtaagtaaaa taagccagtt ttttttaaag 5460  
gacaaataaa cactatacga ttctacttaa gtatttaatg ttgtcaaatt tataaatata 5520  
gaatgtagaa tagtggttac cctgagctgg gggaaagggg caaaggggaa ttgttatttt 5580  
aatgggtata gtttcagttc tgcaaaatga aaaggttctg gaaatctgtt tcacaatggt 5640  
gtaaataata 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 ttttgtagg gtatttfaat 5940  
aatttatatt taatgtaatt attgatatag ttgagtgtgt tgatttttgt tttctatttg 6000  
ctccatctgt tgttggttct cattattcct ctgtttctac cttcttttgt actaattatt 6060  
atattttatt atttttcatc tcaactgttg gcttattagc cacattgctt ttaaaatfff 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 tgaaaaaaaa tggctttcat ctttttccat 6360  
tagatttcat actccttctg cctgaagaat ttcttttaat agaccttgta ctgcggtct 6420  
caggcaagaa attctctcag cctttgttgg tttgaaaaac tgcttattac acctttgttt 6480  
ttgaaagata ttttcaactag gtatagaagt ctgggttgac agttctcatt gtttgcaca 6540  
gcatttttaa gatgccatt caattgtctt gtcttgata attttgatt agtctggtgt 6600  
atcttacc tttgttctc tctgtgcaat gcttcaacca tcccacttca ggctgccttt 6660  
aagatgtttt cttttocctt aatctttagt ttttagctgg ttgacagtga cgcactaag 6720  
tgtagtgtat gaggttgctt ttattgtcac tgttgttg 6758  
<210> 348  
<211> 501  
<212> ADN  
<213> Homo sapiens  
<400> 348  
gaccatgta tgacatttta gtgcttgcta agcagtaaat actgacttac tttcctgcta 60  
cactctcag agcagaaaga gaaatctaca aaaagggcaa thtagttggg atccaccaca 120  
gccttgagac tgggcatgt ttctacagct taccacatt ttacccccac tttctctgag 180

5

ES 2 909 841 T3

	aaacaatgca aactggagaa caaggtcaga gaagttatct tggatggtag aagagaagaa	240
	aggagaagaa rggataagca gaaaatcaaa aagggcataa aaaaattact ggggaaaata	300
	attcttagtc actcaccatt tcttatgttt gtgaaaacag aaacgaggag caagtgtgt	360
	tgtaagaatt gttcttgccc ctccccctcc accaccaca tctgtcaagc tatccctgtt	420
	tcactgtttc ctctgcactc tctattaact tctttgtcct cctcttttct tttcctacag	480
	caaagacttt ttgtcatggt t	501
	<210> 349	
	<211> 501	
	<212> ADN	
5	<213> Homo sapiens	
	<400> 349	
	tgacttactt tcctgctaca ctcttcagag cagaaagaga aatctacaaa aagggcaatg	60
	tagttgggat ccaccacagc cttgagactg ggccatgttt ctacagctta cccacathtt	120
	accccactt tctctgagaa acaatgcaaa ctggagaaca aggtcagaga agttatcttg	180
	gatggtagaa gagaagaaag gagaagaaag gataagcaga aaatcaaaaa gggcataaaa	240
	aaattactgg rgaaaataat tcttagtcac tcaccathtt ttatgtttgt gaaaacagaa	300
	acgaggagca agtgttggtg taagaattgt tcttgcccct ccccctccac caccacatc	360
	tgtcaagcta tcctgtttc actgtttcct ctgcactctc tattaacttc tttgtcctcc	420
	tcttttcttt tcctacagca aagacttttt gtcatgtttt gtttcttttt ctattgtttc	480
	tttcctttt ctaatccttg a	501
	<210> 350	
	<211> 1148	
	<212> ADN	
10	<213> Homo sapiens	
	<400> 350	
	tatgagatth aatgttaaga aataaaatgt aggatctaaa acgtaatcta tagcataatc	60
	tcaaaaatgg tttagaaatg acataataat acagacatth gtgggtggta ggattatgca	120
	tatttttata tattttttaa tatatttttc aaaagcttcc tataaagaat gtaattcttt	180
	cccaattcca aatctagctt aaacataatt ttacaaaaat tattctctca gaatgtaaac	240
	tagtaccacc tctatggaaa acattatgga gatttcctaa agagttaaaa gtagatctac	300
	catttgatcc agcaatctta atactgggta tctaccgga ggaaaagaag tcattgtatg	360
	aaaaagacac ttgtacacat atgtttacag gaccacaatt cacaaatgca aagatgcaga	420
	accaacctaa gtggccastg actaatgaga ggataaagaa gatgtggcat atatatatca	480
	gggactacta ctcagccatt acaaggaaca aaataatgtc ttttgcaaca acttgatag	540
	agctggaggc cattattcta agtaaagtaa ttcaggaatt ggaaaaccaa aaaccgtatg	600

ES 2 909 841 T3

	ttctctctta taagtgggaa ctaagttag	aataagcaaa	ggcacacaga	gggacatatt	660
	ggactttaga gactcacgag gaggaggta	ataggggact	agggattaaa	agaaaaacta	720
	gacattaggt acaaggtacc ctacttaagt	gcactaaaat	ctcagaattc	accactacgt	780
	aattcaacta agtaacaaga aaccacttgt	accccaaaag	ctactgaaat	aaaaattatt	840
	ctctcaaaaa ttttaagccc taaacttcag	ttcctattgt	ttatatttac	taagaaaaac	900
	aacagaaaaac actgttttaa aaatgggtga	tttttttaag	gttaaaggta	tataagacag	960
	ctgcctaagg aaacgcagat acccctgtac	cttgtgtgtg	ttgttgtttt	tcactttttt	1020
	aaaaaacata gagatgggat ctccctatgc	tgcccaggct	tgtctcaaac	tcctgagctc	1080
	aagcaatcct ctgacctcag actctcaaag	ttttgggact	acaggcgaca	gtcaccatgc	1140
	cagccaat				1148
	<210>	351			
	<211>	1148			
	<212>	ADN			
5	<213>	Homo sapiens			
	<400>	351			
	tatgagattt aatgttaaga aataaaatgt	aggatctaaa	acgtaatcta	tagcataatc	60
	tcaaaaatgg tttagaaatg acataataat	acagacattt	gtgggtggta	ggattatgca	120
	tatttttata tattttttaa tatatttttc	aaaagcttcc	tataaagaat	gtaattcttt	180
	ccaattcca aatctagctt aaacataatt	ttacaaaaat	tattctctca	gaatgtaaac	240
	tagtaccacc tctatggaaa acattatgga	gatttcctaa	agagttaaaa	gtagatctac	300
	catttgatcc agcaatctta atactgggta	tctaccggga	ggaaaagaag	tcattgtatg	360
	aaaaagacac ttgtacacat atgtttacag	gaccacaatt	cacaaatgca	aagatgcaga	420
	accaacctaa gtggccactg actaatgaga	ggataaagaa	gatgtggcat	atatayatca	480
	gggactacta ctcagccatt acaaggaaca	aaataatgtc	ttttgcaaca	acttggatag	540
	agctggaggc cattattcta agtaaagtaa	ttcaggaatt	ggaaaaccaa	aaaccgtatg	600
	ttctctctta taagtgggaa ctaagttag	aataagcaaa	ggcacacaga	gggacatatt	660
	ggactttaga gactcacgag gaggaggta	ataggggact	agggattaaa	agaaaaacta	720
	gacattaggt acaaggtacc ctacttaagt	gcactaaaat	ctcagaattc	accactacgt	780
	aattcaacta agtaacaaga aaccacttgt	accccaaaag	ctactgaaat	aaaaattatt	840
	ctctcaaaaa ttttaagccc taaacttcag	ttcctattgt	ttatatttac	taagaaaaac	900
	aacagaaaaac actgttttaa aaatgggtga	tttttttaag	gttaaaggta	tataagacag	960
	ctgcctaagg aaacgcagat acccctgtac	cttgtgtgtg	ttgttgtttt	tcactttttt	1020
	aaaaaacata gagatgggat ctccctatgc	tgcccaggct	tgtctcaaac	tcctgagctc	1080
	aagcaatcct ctgacctcag actctcaaag	ttttgggact	acaggcgaca	gtcaccatgc	1140
	cagccaat				1148
10	<210>	352			
	<211>	1148			
	<212>	ADN			
	<213>	Homo sapiens			
	<400>	352			

ES 2 909 841 T3

tatgagat	ttt	aatg	ttaaga	aataaa	aatgt	aggat	ctaaa	acg	taat	cta	tag	cataat	c	60										
tcaaaa	aatgg	tttag	aaatg	acata	ataat	acag	acatt	tg	gggt	ggta	ggatt	atg	ca	120										
tatttt	tata	tatttt	tata	tatttt	tc	aaa	agct	tc	tata	aga	at	gtaatt	cttt	180										
ccaatt	cca	aatct	agctt	aa	cataat	ttaca	aaaat	tatt	ctct	ca	gaat	gtaaac		240										
tagtacc	acc	tctat	ggaaa	acatt	atgga	gattt	cctaa	agag	ttaaaa	gtag	atct	ac		300										
catttg	atcc	agcaat	ctta	atact	gggta	tctacc	cgga	ggaa	agaag	tcatt	g	atg		360										
aaaa	agac	ac	ac	atg	ttac	ag	gacc	aca	att	cacaa	atg	ca	aagat	gcaga	420									
accaac	ctaa	gtg	gccact	g	actaat	gaga	ggata	aaaga	gat	gtg	gcat	ata	tay	atca	480									
gggact	acta	ctc	agcc	att	aca	agga	aca	aaata	atg	tc	tttg	ca	aca	actt	g	atag	540							
agctg	gag	gc	catt	att	cta	ag	taa	ag	taa	ttc	agga	att	gg	aaa	ac	ca	aaacc	g	atg	600				
ttctc	ctta	taag	tgg	aa	cta	ag	ttag	aa	ta	ag	ca	aaa	gg	ca	ca	gaga	ggg	ca	tatt	660				
ggact	tt	aga	gact	ca	cg	ag	gag	gg	gta	atag	ggg	act	agg	gatt	aaa	ag	aaaa	acta		720				
gacatt	ag	gt	tacc	ctact	t	aa	gt	gcact	aaa	at	ctc	aga	att	c	acc	act	ac	g		780				
aattc	a	acta	ag	ta	aca	aga	aacc	act	gt	ac	ccca	aaa	ag	ctact	g	aa	att	tatt		840				
ctctc	aaaa	ttt	ta	ag	ccc	t	aa	act	t	ca	g	ttc	ct	att	gt	ttat	att	ta	ag	aaa	ac	900		
aacag	aaa	ac	act	gtt	tt	aa	aat	gg	t	gga	tttt	tt	ta	ag	g	ta	ta	ag	ac	ag		960		
ctgc	ct	aa	gg	aa	ac	g	agat	ac	ccc	ct	gt	ac	ctt	gt	gt	tt	g	ttt	t	ca	ctt	ttt	1020	
aaaa	a	acata	gag	at	ggg	at	ctc	ct	at	g	tc	ccc	ag	g	ct	g	ca	aac	tc	ct	g	ag	ctc	1080
aagca	at	cct	ctg	ac	ct	ca	ag	ttt	t	ggg	act	ac	ag	g	ca	g	ca	g	tc	ac	cat	g	1140	
cag	cca	at																					1148	
<210>	353																							
<211>	1148																							
<212>	ADN																							
<213>	Homo sapiens																							
<400>	353																							
tatgagat	ttt	aatg	ttaaga	aataaa	aatgt	aggat	ctaaa	acg	taat	cta	tag	cataat	c	60										
tcaaaa	aatgg	tttag	aaatg	acata	ataat	acag	acatt	tg	gggt	ggta	ggatt	atg	ca	120										

5

ES 2 909 841 T3

	tatattttata tatatttttaa tatatttttc aaaagcttcc tataaagaat gtaattcttt	180
	cccaattcca aatctagctt aaacataatt ttacaaaaat tattctctca gaatgtaaac	240
	tagtaccacc tctatggaaa acattatgga gatttcctaa agagttaaaa gtagatctac	300
	catttgatcc agcaatctta atactgggta tctaccggga ggaaagaag tcattgtatg	360
	aaaaagacac ttgtacacat atgtttacag gaccacaatt cacaaatgca aagatgcaga	420
	accaacctaa gtggccactg actaatgaga ggataaagaa gatgtggcat atatatatca	480
	gggactactr ctcagccatt acaaggaaca aaataatgtc ttttgaaca acttgatag	540
	agctggaggc cattattcta agtaaagtaa ttcaggaatt ggaaaacca aaaccgatg	600
	ttctctctta taagtgggaa ctaagttagg aataagcaaa ggcacacaga gggacatatt	660
	ggactttaga gactcacgag gaggagggta ataggggact agggattaaa agaaaaacta	720
	gacattaggt acaaggtacc ctacttaagt gcactaaaat ctcagaattc accactacgt	780
	aattcaacta agtaacaaga aaccacttgt accccaaaag ctactgaaat aaaaattatt	840
	ctctcaaaaa ttttaagccc taaacttcag ttcctattgt ttatatttac taagaaaaac	900
	aacagaaaac actgttttaa aaatgggtga ttttttaag gttaaaggta tataagacag	960
	ctgcctaagg aaacgcagat acccctgtac cttgtgtgtg ttgtgtttt tcactttttt	1020
	aaaaaacata gagatgggat ctccctatgc tgcccaggct tgtctcaaac tcctgagctc	1080
	aagcaatcct ctgacctcag actctcaaag ttttgggact acaggcgaca gtcaccatgc	1140
	cagccaat	1148
	<210> 354	
	<211> 611	
	<212> ADN	
5	<213> Homo sapiens	
	<400> 354	
	caaaacctca accttcaga taagtctaag ggtgagaact tcacacaaga tgaataagaa	60
	ccaatttctt ccaggcgat gttgaacctg gaaatgaaag ccaatctctc ttggaagcc	120
	tggtttgtag aaatgtcagt ctttgtttca agctgtggga gaatgagaag caagacttta	180
	gggaaagagg aataaaatag atgtgcagaa ataacagagt gagaaagtct tcagggtgct	240
	gctagcccta attgcaggca tcctgaatc ctagacctg gattgcaaga gactccttaa	300
	tatcttccca tgtccacatt tgcttcacat agtttgaatg tggcttctat tatatacaga	360
	tacaagattc aaatccaacc tctaygatga ctggtcttgt gaataagcag aagaggcact	420
	aacaatatga cgtgagggat tcagggaaga gcactttctt gagcacatat cttccctggg	480
	ctgccagctg tagtttatga aattccacaa tgaggatgaa atggaatcac catttacaga	540
	gtactctcca gatgtctaac cctaagctag gtacctcaa aatattatct agtttagata	600
	atcaaccctt t	611
10	<210> 355	
	<211> 601	
	<212> ADN	
	<213> Homo sapiens	
	<400> 355	

ES 2 909 841 T3

	t	601
	<210> 356	
	<211> 527	
	<212> ADN	
5	<213> Homo sapiens	
	<400> 356	
	gctctagaat atggcattcc agaagtggga tgctacaaat agtctcattg agagtcaact	60
	tgcaacaatgt atcgtctctac ccttacatca atttctgaaa caacttctct ttgcacttcc	120
	cctatagtta catgcataat aaattctgac aactcttatg aagtcattga ataactttct	180
	tcttatgttt cctatcaatg tcattagccc tttatcttgt ttgagtttcc atcagcaatg	240
	ttttcaagtc ccaagatcat tcatgtatcc acaagcaatg atacgccaga tttggacaaa	300
	taatactgaa tactatctta ttttactgac catgatcaag gcagtgtgga ttgctgccaa	360
	gtccaagaga agtgagggtca gcagctgcaa gccacctccg tcatttagaa aagcttcatg	420
	atgtagtggtg tcgtttcgat gtgacactgt ctcacagagt taaaatgatg tgmaaggaac	480
	tgttcaatgg aaatttagaa atttctcttt ttctcaattt tagtgta	527
	<210> 357	
	<211> 601	
10	<212> ADN	
	<213> Homo sapiens	
	<400> 357	
	gaacaagatt ttcttgcttt taaaataact acattaaagc tgaaaattta ggccaaaatt	60
	ttcaagtggg aatagttaca ggcaattcat ctttctggtc agaaaagggt gttactgcag	120
	ctatttctgc ctgaaactgg gtggcactac tacttttttt tttttttttt taactgagca	180
	gacattttcc ttacactaaa attgagaaaa agagaaattt ctaaatttcc attgaacagt	240
	tccttgcaca tcattttaac tctgtgagac agtgtcacat cgaaacgaca cactacatca	300
	ygaagctttt ctaaagacg gaggtggctt gcagctgctg acctcacttc tcttgactt	360
	ggcagcaatc cacactgcct tgatcatggc agtgaaaata agatagtatt cagtattatt	420
	tgtccaaatc tggcgtatca ttgcttggg ataatgaat gatcttggga cttgaaaaca	480
	ttgctgatgg aaactcaaac aagataaagg gctaatagaca ttgataggaa acataagaag	540
	aaagttattc catgacttca taagagttgt cagaatttat tatgcatgta actacagggg	600
	a	601

ES 2 909 841 T3

<210> 358  
 <211> 1001  
 <212> ADN  
 <213> Homo sapiens  
 5 <400> 358  
 gcttaataacc tgagtgatgg aatattctgt tcaacaaacc cctctgacat aggtttgcct 60  
 atataataaa cctgttcatg tactcctgaa cctaaaagtt taaaaaagat tatgtagaaa 120  
 acccaaagga atctataaaa agtctactag agctagagtg attttaacaa gatttcaata 180  
 cacaaattca aatgtctttc tatatattaa tgacaatcaa caataaaatt ttaaacatt 240  
 attaaagtat aatgaaaata tcaactgttt aggagaaat gtaacaagaa tggatgaagga 300  
 cctatacact aaaaagcttc aatatgttgt tgagattaac tgaagaaggt ctaaatagat 360  
 ttttttttca tgtctcggaa gacttaatat gtgaagatac caattcttcc ccaaatgatc 420  
 aacaggtgaa atgcaatccc aatcaaaaac ccagcaatta ttttaagggg gaaattggca 480  
 atctgattct aaaattcata yggaaaaaaa caatggagtt agaataacta aaacaagtcc 540  
 gaaaaagaaa aagaaatgga ggactaatgc tacctgattt caagtcttat cgtataaatc 600  
 tacatcaata aaggacaagt tggatttggg ttaaagatag ataaatacat cagtgaata 660  
 gaatattgaa tccagaataa atccacacat atatggataa aaataccaga caattcagtg 720  
 gagatggttt tgtttttaca acaaatgtta ctggaacaaa ttgatatatg tattagtcag 780  
 atatggctgc cataacaaag aaccacaaac aggtggttta aataatggaa ataaatttcc 840  
 tcagaattct ggagtatgga agcccaagat caagttgctg ggaggattcg tttcttctga 900  
 gtgtctcttt ttttgatgac agatgactat cttttaccaa tgtcttctact tggttttccc 960  
 tctgtgtgtg cctaggtcct attctccaat tcctataagg a 1001  
 <210> 359  
 <211> 1001  
 <212> ADN  
 10 <213> Homo sapiens  
 <400> 359

ES 2 909 841 T3

ctaaaagttt aaaaaagatt atgtagaaaa cccaaaggaa tctataaaaa gtctactaga	60
gctagagtga ttttaacaag atttcaatac acaaattcaa atgtctttct atatattaat	120
gacaatcaac aataaaatth taaaacatta ttaaagtata atgaaaatat caactgttta	180
gggagaaatg taacaagaat ggtgaaggac ctatacacta aaaagcttca atatgttggt	240
gagattaact gaagaaggtc taaatagatt tttttttcat gtctcggaag acttaatatg	300
tgaagatacc aattcttccc caaatgatca acaggtgaaa tgcaatccca atcaaaatcc	360
cagcaattat ttttaagggg aaattggcaa tctgattcta aaattcatat ggaaaaaac	420
aatggagtta gaataactaa aacaagtccg aaaaagaaaa agaaatggag gactaatgct	480
acctgatttc aagtcttctc rtataaatct acatcaataa aggacaagtt ggtattgggt	540
taaagataga taaatacatc agtggatag aatattgaat ccagaataaa tccacacata	600
tatggataaa aataccagac aattcagtgg agatggtttt gtttttacia caaatgttac	660
tggacaacaa tgatatatgt attagtcaga tatggctgcc ataacaaaga accacaaaca	720
ggtggtttaa ataatggaaa taaatttcct cagaattctg gagtatggaa gcccaagatc	780
aagttgctgg gaggattcgt ttcttctgag tgtctctttt tttgatgaca gatgactatc	840
ttttaccaat gtcttcactt ggttttccct ctgtgtgtgc ctaggctcta ttctccaatt	900
cctataagga aaccagtcac attggattag ggcccactct aatggcccca ttttacttgc	960
attatctctt taaagacact atctccagat gtagccacac t	1001
<210> 360	
<211> 1058	
<212> ADN	
<213> Homo sapiens	
<400> 360	
catgattagc tatgctactt tccactgctc ttagtatact gagaggcagc ataagtaaaa	60
ctaaaatata tgaagatagc aatagactat ttaaagtaga agaagtatgc tatttttggt	120
ttgttttcat ttcgaaggaa atatgcaaag gtttattgag tatttcagct tctcttacag	180
taggtttttt ttggattcct tctgtgtttg tctatgttga taaaacattg aaatgccata	240
tagctcaaag gtcattcact taagaaatct aagtaactgat aacatcttag ccccgattct	300
tcataggcat tgtaagcct attataatth ttggtwcagag agaaggtaaa ctatattcca	360
gacaggcata taaagcaatt tctcctataa ttggagttca cgaaaaattc acatatttct	420
ttttaatagt aactctcaca gcaagaacat atgtttgtaa ataatacatc acagaatctt	480

5

ES 2 909 841 T3

attggcagac aaggaaattc ctaaaatatt ttttactgcc acatcaatta agatatataa 540  
 aataccttat atagaagatg tttgcaccca ggccaaacaa atcaaacaag aatagaagca 600  
 ctgacagtct tatttcaaaa ttggtttaac ttgtatttac aggatattgt agtaccttat 660  
 aaagttgatt gctgattggc cgtcttttac agaattctgt cagattgta ttatttcttg 720  
 taaagattga ttcaaacaaa taaaaattgt caggattgga tatgtcctat agtgaggtgt 780  
 agttatgtca catgagattt ttaattacaa agaaatggaa aataaaatga gaatagaatt 840  
 gagactcccc tgtcacctca caaatatggt gaaatacaat gaaatttcca aagatgttaa 900  
 agcatataaa gttgaataat tcttattatg tattaaactt acagaaattt aatttcttta 960  
 ctttataaga ggtagtgaag atataaaatt aattatgaag acagagtagt cttagtcaga 1020  
 catggcccta taaagcatat tcccattcgt tacatcaa 1058  
 <210> 361  
 <211> 1058  
 <212> ADN  
 5 <213> Homo sapiens  
 <400> 361  
 catgattagc tatgtactt tccactgctc ttagtatact gagaggcagc ataagtaaaa 60  
 ctaaaatatac tgaagatagc aatagactat ttaaagtaga agaagtatgc tatttttgtt 120  
 ttgttttcat ttcgaaggaa atatgcaaag gtttattgag tatttcagct tctcttacag 180  
 taggtttttt ttggattcct tctgtgtttg tctatgttga taaaacattg aatgccaya 240  
 tagctcaaag gtcattcact taagaaatct aagtactgat aacatcttag ccccgattct 300  
 tcataggcat tgtaagcct attataattt tggtagagag agaaggtaaa ctatattcca 360  
 gacaggcata taaagcaatt tctcctataa ttggagttca cgaaaaattc acatatttct 420  
 ttttaatagt aactctcaca gcaagaacat atgtttgtaa ataatacatc acagaatctt 480  
 attggcagac aaggaaattc ctaaaatatt ttttactgcc acatcaatta agatatataa 540  
 aataccttat atagaagatg tttgcaccca ggccaaacaa atcaaacaag aatagaagca 600  
 ctgacagtct tatttcaaaa ttggtttaac ttgtatttac aggatattgt agtaccttat 660  
 aaagttgatt gctgattggc cgtcttttac agaattctgt cagattgta ttatttcttg 720  
 taaagattga ttcaaacaaa taaaaattgt caggattgga tatgtcctat agtgaggtgt 780  
 agttatgtca catgagattt ttaattacaa agaaatggaa aataaaatga gaatagaatt 840  
 gagactcccc tgtcacctca caaatatggt gaaatacaat gaaatttcca aagatgttaa 900  
 agcatataaa gttgaataat tcttattatg tattaaactt acagaaattt aatttcttta 960  
 ctttataaga ggtagtgaag atataaaatt aattatgaag acagagtagt cttagtcaga 1020  
 catggcccta taaagcatat tcccattcgt tacatcaa 1058  
 <210> 362  
 <211> 956  
 10 <212> ADN  
 <213> Homo sapiens  
 <400> 362

ES 2 909 841 T3

aaaacaagga acaaaacaac aaaaatgtta caaccgaaca acagactttt gagtcatggt 60  
 tcaggccaag aggtgatgag ttactgtagt tgcttgagct ggttggtgaa atattacctg 120  
 gcaacaaaac tgaatataga ggtggcttag taaaatgcag attcagaatg agtgccttaa 180  
 ggtaaggca tataagacca aactgatttt ctttttcacg aggtcttcag gtaaggccat 240  
 tgtagaagat accttgtttg cgaacttcag taaattactt cacttgtctc atattttcat 300  
 tttcaggatg gaggcttgag attgaattgt agtgcaatta ggtaaatfff taccattfff 360  
 aatataata ttaaaatatt aattataaat taccttattt gaatctggaa taatatttat 420  
 tgcagggcat ataatctaag ctgtaaactg cctgtyagaa gacaacatat tcatcttgct 480  
 aaggataag ctatatgact ggcactgtgc tcaactcaga gtcattgaat gaacagtatt 540  
 tattaatct atgaatgaga gcacttcaag tatacagaaa gatatctcaa aagattcagc 600  
 cttacattgc tcataacttc aatgacttag atgaaaact cctgaacatt tttatcagtt 660  
 gtataggtag cccaaatcat aagggaatgt ttatcaatta gatgatgaaa tggggatgca 720  
 actacatcat ggcaggctaa agcaatagaa tgactttgac aagaggaaat tacatagagg 780  
 cacctgagtc tcctaaacca atttcaaagg tatgagaggg gggatgata aataaatagt 840  
 tgatagatga aaaaactcag aagttatagt tgacagcaat ttaatatata tatgaaaaat 900  
 gtggttgagc ttttagggaa aaaaactaa taaaatctaa tggaaattag tgggcc 956  
 <210> 363  
 <211> 956  
 <212> ADN  
 <213> Homo sapiens  
 <400> 363  
 caaccgaaca acagactttt gagtcatggt tcaggccaag aggtgatgag ttactgtagt 60  
 tgcttgagct ggttggtgaa atattacctg gcaacaaaac tgaatataga ggtggcttag 120  
 taaaatgcag attcagaatg agtgccttaa ggtaaggca tataagacca aactgatttt 180  
 ctttttcacg aggtcttcag gtaaggccat tgtagaagat accttgtttg cgaacttcag 240  
 taaattactt cacttgtctc atattttcat tttcaggatg gaggcttgag attgaattgt 300  
 agtgcaatta ggtaaatfff taccattfff aatataata ttaaaatatt aattataaat 360  
 taccttattt gaatctggaa taatatttat tgcagggcat ataatctaag ctgtaaactg 420  
 cctgtcagaa gacaacatat tcatcttgct aaggtrtaag ctatatgact ggcactgtgc 480  
 tcaactcaga gtcattgaat gaacagtatt tattaatct atgaatgaga gcacttcaag 540  
 tatacagaaa gatatctcaa aagattcagc cttacattgc tcataacttc aatgacttag 600  
 atgaaaact cctgaacatt tttatcagtt gtataggtag cccaaatcat aagggaatgt 660  
 ttatcaatta gatgatgaaa tggggatgca actacatcat ggcaggctaa agcaatagaa 720  
 tgactttgac aagaggaaat tacatagagg cacctgagtc tcctaaacca atttcaaagg 780  
 tatgagaggg gggatgata aataaatagt tgatagatga aaaaactcag aagttatagt 840  
 tgacagcaat ttaatatata tatgaaaaat gtggttgagc ttttagggaa aaaaactaa 900  
 taaaatctaa tggaaattag tgggccactc atttctccac ctaggatggt aaaaat 956  
 <210> 364  
 <211> 601

5

10

ES 2 909 841 T3

<212> ADN  
 <213> Homo sapiens  
 <400> 364  
 gtaaaacaca tagatcgctg tatccttggt cagtaagcta caacatactc gtatctcctg 60  
 aaatcctggg cttaaactcga ggtctcaaag gctttgttt gttttgttgt atggttgtat 120  
 ggtgagtgtg tgtgtgtgtg tgtgtgtgtg tgtttattct cctgaaattc tcctcctcac 180  
 ttgacttaag ctaaaagata aacgtcctct tcctttcagc cacagatggt gatggataaa 240  
 ttgaatgtca ttcacattat tcccttaaaa taaactctct ccctcccctc tcccgtctca 300  
 wccttgctcc tttctttata taatgggtaa tgcgttaatg tcagcagaat agttttgggg 360  
 ccataatggc aagtatcacg tggatggttt agcattgttt ttagaatgct gtgaatttgg 420  
 gtatatgtga gttttgggga aagttttgca actatatgtt tgttaattaa atgaggacta 480  
 taaagtaata taaaattatg tttctggaac atattttgga agctataaag tcatctgtat 540  
 ttattatcca cagacataat gtcattgttc aggtcctgca accttcttat aatcaacata 600  
 c 601  
 5 <210> 365  
 <211> 601  
 <212> ADN  
 <213> Homo sapiens  
 <400> 365  
 agtaagctac aacatactcg tatctcctga aatcctgggc ttaaactcgag gtctcaaagg 60  
 ctttgttttg ttttgttgta tggttgtatg gtgagtgtgt gtgtgtgtgt gtgtgtgtgt 120  
 gtttattctc ctgaaattct cctcctcact tgacttaagc taaaagataa acgtcctctt 180  
 cctttcagcc acagatgggtg atggataaat tgaatgtcat tcacattatt cccttaaaat 240  
 aaactctctc cctcccctct cccgtctcat ccttgcctt tttttatat atgggtaat 300  
 10 kcgттаатgt cagcagaata gttttggggc cataatggca agtatcacgt ggatggттта 360  
 gcattgtttt tagaatgctg tgaatttggg tatatgtgag ttttggggaa agttttgcaa 420  
 ctatatgttt gttaattaaa tgaggactat aaagtaatat aaaattatgt ttctggaaca 480  
 tattttggaa gctataaagt catctgtatt tattatccac agacataatg tcattgttca 540  
 ggtcctgcaa ccttcttata atcaacatac gtggggcccag ggattttatg tatcttcgcc 600  
 t 601  
 15 <210> 366  
 <211> 1079  
 <212> ADN  
 <213> Homo sapiens  
 <400> 366

ES 2 909 841 T3

gaatttatgg tctgatggag aagggaatca ttaaagttct atgtagtgag atatccccaa 60  
 ggggtgtatt aggcttacca ccaactggaat ctggatagat gaagacagag tggcagggaa 120  
 gtcgtattaa ggttctgttt ctgctgggag ccacaggtcc tcaggaagca acaagtactg 180  
 ggcagattga tactgtagct rggctctagc tctatacctc tagaataaag gttacaaact 240  
 agcaacttga aagctaaacc tggcccacag atatgtttta tttggctctt aactgtttt 300  
 aaaaaatatt accaacattt aaaactggga agttttatga aaaaaccag acttctggat 360  
 tctgttgaaa aaaaaatca gaagatctgg caatactgag ctgacattcc tatatgacaa 420  
 caattggctg gatctatgca gcttctctcc aaaaagcaaa gaatgtgttc ttgcttaaca 480  
 cagtccccac cactccctca tattctccaa tcctggacct gagcgtcatt tgctatgtat 540  
 cgccatttgc catgaagttt tacactctac agaaatataa tttttttgta gaagactatg 600  
 ctttaatcaa gatcaggata atataaagtg agatctgaaa gtggaaaaaa gataaatgtc 660  
 caacaatgat agactggatt aagaaaatgt ggcacatata caccgtggag tactatgcag 720  
 ccaaaaaaaaa cgatgagttc atgtcctttg tagggacatg gatgaagctg gaaaccacca 780  
 ttctcagcaa actatcgcaa ggacaaaaaa ccaaacgccg catgttctca ctcataggtg 840  
 ggaattgaac aatgagaaca cttgggcaca ggaaggggaa catcacacac cgggccctgt 900  
 tgtggggtgg ggggaggagg gagggatagc atttggagat atacctaag ttaaagact 960  
 agtttctggg tgcagcacac catcatggca catgtataca tatgtaacta acctgcacat 1020  
 tgtgcacatg taccctaaaa cttaaagat aatttttaaa aaaagatatt ttcttatct 1079  
 <210> 367  
 <211> 501  
 <212> ADN  
 <213> Homo sapiens  
 <400> 367  
 ataaattttc tcttccctca agaatttatg gtctgatgga gaagggaatc attaaagttc 60  
 tatgtagtga gatatcccc aaggggtgtat taggcttacc accactggaa tctggataga 120  
 tgaagacaga gtggcagggg agtcgtatta aggttctggt tctgctggga gccacaggtc 180  
 ctcaggaagc aacaagtact gggcagattg atactgtagc tgggctctag ctctatacct 240  
 ctagaataaa kgttacaaac tagcaacttg aaagctaaac ctggcccaca gatatgtttt 300  
 atttggctct tacactgttt taaaaaatat taccaacatt taaaactggg aagttttatg 360  
 aaaaaaccca gacttctgga ttctgttgaa aaaaaaatc agaagatctg gcaactactga 420  
 gctgacattc ctatatgaca acaattggct ggatctatgc agcttctctc caaaaagcaa 480  
 agaatgtggt cttgcttaac a 501  
 <210> 368  
 <211> 1001  
 <212> ADN  
 <213> Homo sapiens  
 <400> 368

5

10

ES 2 909 841 T3

tgaagaagcc gcctggcttc ttgtttcttc tcatagcaaa atgcaatgag aaagagataa	60
tttgagaaaa gaaccgttta aacaaaaaga aaccaagaca taatgatttt ggaaattctc	120
agtttattca gactgcaaaa gatattaataa taaagaaact cagtaacagg gatagataat	180
ctaaagaaaa agcctaggac acggctgtag taaccttctg tttttatacc tcagcaattt	240
gctaatagcct caaaaagatc aaaagtactc aaatataaag ggctctttga agagattaga	300
tttcctcaat caaaccaaag agcatcgagg aagcttaagg ttactgtccc tcacatatct	360
cagcagaagg caaaaataga agactgatta tctaagaaag atctctgaaa gagtctcata	420
ttatggagtg aaccctgtg gcatacatgg gagaccact tggttcttga gaattttata	480
tcaggagaaa cactgtcagt ytgtattgaa aggaacagag aaaatacгаа attaaagaag	540
actattaaac ctccaaaatt ctggcaggaa agaagcttac acagctactc agttgcaaag	600
atctgccact tttcatatac atgaaaggac tcagaggagg aagccacagg tttagaagga	660
aaagctaaaa gcaacatcgt attagtcttg gatctaggaa cctaatttct ctagcagaat	720
ctagaaatgg cttgggacaa gtgattgttt ttttacctag gattttctcc ctcttgaaaa	780
caggactgtc tgtaactatt atcctatgcc tgccctacca tcatatttca gaaacaggta	840
acttatgttt tcactttcaa agattcaciaa taaagagaaa ttgtacctca gaatgatta	900
taccagagct ttcctcatgc ataaattaa taatttaggt tatgtgattt gaagcttttg	960
agtgggtgag gtgacatttt ggatgctgag ttggtgccgt a	1001
<210> 369	
<211> 1001	
<212> ADN	
<213> Homo sapiens	
<400> 369	

5

ES 2 909 841 T3

tcttctcata gcaaaatgca atgagaaaga gataatthga gaaaagaacc gtttaacaa 60  
 aaagaaacca agacataatg atthttggaaa ttctcagtht attcagactg caaaagatat 120  
 taaaataaag aaactcagta acagggatag ataatctaaa gaaaagcct aggcacagc 180  
 tgtagtaacc ttctgtthtt atacctcagc aatthtgctaa tgcctcaaaa agatcaaaa 240  
 tactcaata taaagggctc thtgaagaga ttagatthcc tcaatcaaac caaagagcat 300  
 cgaggaagct taaggttact gtccctcaca tatctcagca gaaggcaaaa atagaagact 360  
 gattatctaa gaaagatctc tgaaagagtc tcatattatg gagtgaacc ctgtggcata 420  
 catgggagac ccacttggtt ctthgagaatt thtatcagc agaaactg tcagthctgta 480  
 thgaaaggaa cagagaaaat rcgaaattaa agaagactat taaacctca aatthctggc 540  
 aggaaagaag cttacacagc tactcagthg caaagatctg ccactthtca tatacatgaa 600  
 aggactcaga ggaggaagcc acaggtthtag aaggaaaagc taaaagcaac atcgtattag 660  
 tctthgatct aggaacctaa thtctctagc agaatctaga aatggctthg gacaagthg 720  
 ththththta cctaggttht thtccctctt gaaaacagga ctgtctgtaa ctattacct 780  
 atgcctgccc taccatcata thtcagaaac aggthactta thththtact thcaagatt 840  
 cacaataaag agaaattgta cctcagaatg gattatacca gagctthcct catgcataaa 900  
 thaaataatt taggtthtgt gattthgagc thththgagthg thgaggtgac atththgagth 960  
 ctgagthggt gccgthgta gtccagaatt ctgcggaact t 1001

<210> 370

<211> 601

<212> ADN

5 <213> Homo sapiens

<400> 370

ctctagactc ctctgtatt thaatthagc cactththta ggcctacaa ththtagatct 60  
 ccacagggct cthgaaactt cthgacctc atcagthaca thtccattag thgcatgacc 120  
 caagagthct agaacatcta thcagcaagt thgtatctgg thagtgaata thcctthctat 180  
 thgtthccctt thgcatcaaa ctacacactg thctthcctc ththctthca aagctthgaa 240  
 aatthctcac thgtatctca thctthctct cthgaaaaac thgatcacctc thgatgaatta 300  
 raacggaatg accaagcttht gggagagthca aaagaatctc gthgtthaaag actcagagtht 360  
 thagaagcaa caaaaagatt atacagatgt gaatthgthga cctthcctca ccagggcatg 420  
 thgcctthgga thaaagataat cthagcacac actthcatagc thgagaacaa thththgagth 480  
 cththctthta thgaththta cataaagcaa atthgthgath thacctaaag gctthgaccaa 540  
 ggcctaatthc ctctagagcc cctthgatcat gaacaccatt cctthctatga thctthagth 600  
 c 601

<210> 371

<211> 601

<212> ADN

10 <213> Homo sapiens

<400> 371

ES 2 909 841 T3

acaagctcca gccatggacg caattccttc tagaagcaaa atttatctct agactcctcc 60  
 tgtattttaa tttagccact tttttagggc ctacaatttt agatctccac agggctcttg 120  
 aaacttcttg aacctcatca gtaacatgtc cattagtggc atgaccaag agttctagaa 180  
 catctattca gcaagtgtgt atctggtaag tgaatattcc ttctatgtgt tcccttttgc 240  
 atcaaaactac aactgtcat tctccttta tctccaaaag cttgaaaatt cctcacttgt 300  
 rtctcattct ttctctctta gaaaactgat cacctctgat gaattagaac ggaatgacca 360  
 agctttggga gaggcaaaag aatctcggtg ttaaagactc agagttaaag aagcaacaaa 420  
 aagattatac agatgtgaat atgtgacctt cctccaccag ggcatgttgc cttggagtaa 480  
 gataatctaa gcacacactt catagcctga gaacaatttt ggaagtcttt gctttatgga 540  
 tatttacata aagcaaatat ggatatttac ctaaaggctg gaccaaggcc taattcctct 600  
 a 601

5

<210> 372  
 <211> 701  
 <212> ADN  
 <213> Homo sapiens  
 <400> 372  
 gaagatgcac tctaattgttt tttcccagaa gctctgtagg tttagctttt acctttctgg 60  
 gtttgttttg ttttgttttt tgagatggag tcccactcgt gtcaccacag ctggagtaca 120  
 atggtgcaat ctcggttcac tgcaacctcc acctcccggg ttcaagcaat tcccctgtct 180  
 ccacctctcg agtagctggg atgggaggcg cctgccacca tacctggcta attttcatat 240  
 ttttagtaaa gataggggtt caccatgta gccaggetgg tctcgaactc ctgacctcaa 300  
 gtgatccacc cgcctcagct tcccaaagtg ctgggattac aggcgtgagc cactgcgccc 360  
 agccctagct ttttgggtcta tgattcctcc caaattaatt tctgtgaacc attaccttaa 420  
 gatgttgaga ttaaatgtcc agaatctcat ttgttcacct ttgaaaatta agaaacctg 480  
 gcacagtgtt gactggagcc wcttacctta atagaaaata aagctcacat atatccataa 540  
 tgaaaagcag agaccagcac aacctagtc acctgacagt tttaaaatcc aaggccagga 600  
 tcttctcaac tcaggccac tcacttactc cacaacatac ttcttctttc ctcagcatct 660  
 actacttgtg ctgggacctt ggtcttccca ttgttcatgt c 701

10

<210> 373  
 <211> 701  
 <212> ADN  
 <213> Homo sapiens  
 <400> 373

ES 2 909 841 T3

agatggagtc cactcgtgt caccaggct ggagtacaat ggtgcaatct cggttcactg 60  
 caacctccac ctcccgggtt caagcaattc ccctgtctcc acctctcgag tagctgggat 120  
 gggagcgcc tgccaccata cctggctaata tttcatatth ttagtaaaga tagggtttca 180  
 ccattgttagc caggctggtc tcgaactcct gacctcaagt gatccaccgc cctcagcttc 240  
 ccaaagtgtt gggattacag gcgtgagcca ctgcgccag ccctagcttt ttggtctatg 300  
 attcctcca aattaatttc tgtgaacct taccttaaga tgttgagatt taatgtccag 360  
 aatctcattt gttcaccttt gaaaattaag aaacctggc acagtgttga ctggagccac 420  
 ttacctaat agaaaataaa gctcacatat atccataatg aaaagcagag accagcacia 480  
 ccatagtcac ctgacagttt waaaatcaa gccaggatc ttctcaactc aggccactc 540  
 acttactcca caacatactt cttctttcct cagcatctac tacttgtgct gggacctgg 600  
 tcttccatt gttcatgtca ttctttcct cacagttccc attcttttct ccctgaaata 660  
 aagaaatttc aaaatatacc atgtttcatg aaaaagacia a 701  
 <210> 374  
 <211> 701  
 <212> ADN  
 <213> Homo sapiens  
 <400> 374  
 gatttccacc ctcaggatgat ggggatggtt gaacatcaa cacctgaaac aggacagacg 60  
 atattgacag tacttgttag ttgcatataa tcacagacca gtggaaacag atgaaccaca 120  
 cagggccaca gcggggtttc actggggaac agagtgaaca atcaggaggt gtgggaggca 180  
 ggttttagtag tttaaagagg ttgaggtgtc ccctggatc ccatgggagg atcacattgg 240  
 ctcatgtgaa ttatcatagc gactggcagg gaactgaaat cttctactca gggataagca 300  
 gaaactgtcc ctggtttcct tgataaaaag ggttgttga taggggacct tatccatggg 360  
 aggaaagtga ggagggaaat ttgtggctaa gccattcaag gccctcccag ttttactaga 420  
 tgtcaaggca gcacacgtaa tattgggact taattttagc cacataacta ataaatttgt 480  
 aagtatgtgc aacggctcac rcttgcttcc agaatggcac ctaaaaaca gatttacctc 540  
 tccccaaatt cagatatgga attaaatgta atgtcaggaa aattgtctaa gagttgaaa 600  
 tgggaaaaaa atgttctttt ggtggagtta tggactccag aggttatcag attctattga 660  
 ataactgact tttgattgta tttgtaacia ttaggtatt t 701  
 <210> 375  
 <211> 1001  
 <212> ADN  
 <213> Homo sapiens  
 <400> 375

5

10

ES 2 909 841 T3

gcatataatc acagaccagt ggaacagat gaaccacaca gggccacagc ggggtttcac 60  
 tggggaacag agtgaacaat caggaggtgt gggaggcagg tttagtagtt taaagaggtt 120  
 gaggtgtccc cctggatccc atgggaggat cacattggct catttgaatt atcatacgga 180  
 ctggcagggg actgaaatct tctactcagg gataagcaga aactgtccct ggtttccttg 240  
 ataaaaaggg ttgtttgata ggggacctta tccatgggag gaaagtgagg agggaaatth 300  
 gtggctaagc cattcaaggc cctcccagtt ttactagatg tcaaggcagc acacgtaata 360  
 ttgggactta attttagcca cataactaat aaatttghaa gtatgtgcaa cggctcacac 420  
 ttgcttccag aatggcacct aaaaaacaga tttacctctc cccaaattca gatatggaat 480  
 taaatgtaat gtcaggaaaa ytgtctaaga gttggaatg ggaaaaaat gttcttttgg 540  
 tggagttatg gactccagag gttatcagat tctattgaat aacgtacttt tgattgtatt 600  
 tghaacaatt aggctatttg tgaactcggg aggggtagaa atcgagttgt agaaaatgga 660  
 tggtaatgca agtgatthtt gaccatatca atgcaaatga attctgttgg tagaaatatt 720  
 catttccaca ctgtagatga ccctaaacat atgtcattac attatathtt attgccttat 780  
 agactattaa ccaathttga atcatacagt agcaaattht tttcagcatt cttgtgtgta 840  
 tgtgtttata tatacacgtg catatgtatt taagatataa aattgtataa tcttcaaatt 900  
 cttctttgaa caggtttgaa cctcttatta gtttctcat taaggaattt aataagacct 960  
 ttaatgcatg tttgtathtt catgagagtc attathttac c 1001  
 <210> 376  
 <211> 695  
 <212> ADN  
 <213> Homo sapiens  
 <400> 376  
 tgctccttca ttagtgcaat ggaacagcaa atcaggatac tttcacagtt ctcttaagtg 60  
 agcctagaag tggggagctg cttgttcaca aacttgaagc ctgaatatgt taatattctt 120  
 tcagtggccg gacgcggtgg ctcatgcctg taatcccaac actttgggag gccgaggtag 180  
 gcagatcaac ctgaagtcag gagttcgagg ccagcctggc caacatggtg aaacccacc 240  
 tgttggctct tactaaaaat agaaaaatta gctgggcatg gtggcgcagc cctgtaatcc 300  
 cagctactca ggaggctgtg gcagaagaat cgcctgcacc tgggaggcag aggttgcttt 360  
 gagttgatat cgtgtcactg cactccagcc tgggcaacag agtgagatcc tttcagaaac 420  
 ctgctgtctg tathttgata caatthaaaa aaaaaaaaag atgagacagg caggtgcgaa 480  
 agaaataaaa gtcamaactg atccagttgg gaaactcaga attgacagtt acgtgtcctt 540  
 tcatttattg atathttgag attcacaggg gththaaactt taththttcca agactgaata 600  
 gttcccacct ccttccata tataaathtt gagtagctgg ggagatttht aagaggctcc 660  
 ccataaactc agaagtthaa agagacaagg gtccc 695  
 <210> 377  
 <211> 601  
 <212> ADN  
 <213> Homo sapiens  
 <400> 377

5

10

ES 2 909 841 T3

aacccacact gttggtctgt actaaaaata gaaaaattag ctgggcatgg tggcgcatgc 60  
 ctgtaatccc agctactcag gaggctgtgg cagaagaatc gcctgcacct gggaggcaga 120  
 ggttgctttg agttgatatc gtgtcactgc actccagcct gggcaacaga gtgagatcct 180  
 ttcagaaacc tgctgtctgt atttggatac aattaaanaa aaaaaaaga tgagacaggc 240  
 aggtgcgaaa gaaataaaag tcacaactga tccagttggg aaactcagaa ttgacagtta 300  
 sgtgtccttt cattttattga ttttttgaga ttcacagggg tttaaacttt attcttccaa 360  
 gactgaatag tccccacctc ccttccatat ataaaatttg agtagctggg gagatttaaa 420  
 agaggctccc cataaaactca gaagttaaaa gagacaaggg tcccagtaaa tacaanaatga 480  
 ttggggttga ggaggcagat tttctgtcct cagtgaagtt tgttggttg ttggttggtt 540  
 ggttggttaa ttggttggtt tttgagtcag ggtctcactt tgtcacccaa gctggagtgc 600  
 a 601

5

<210> 378  
 <211> 663  
 <212> ADN  
 <213> Homo sapiens  
 <400> 378  
 tgtagcaaca ggagggatga gacccaaagg tctgaaaagc cagtatttta agaagtcttg 60  
 gaaaatgtgg aggttgaaaa atctaacagg agtgcttgct tcagcagcaa tttagagtag 120  
 attagcatgg cctctgcgcc aggatgacat gcacattcct aaaagtgttc cgtgttttaa 180  
 aaaaaagaga gagacagaat ctaaggggat gtgtacattt gctagagcta ctataacaaa 240  
 gtaccagagg cagggtcact tcaacaacag aaatttattt ctcacagttc tggaggctag 300  
 acgtccaaga ttaaggtggt gactgggttg aattcagccc ataacaggaa ataaggagtt 360  
 aaataaagca cttgcttcta ttgtttgtac ctaaacttaa cagaayacag taagtaacaa 420  
 gtcattggga tgcagaaaag aaaaaagaga gtgaaggaag gagagaaggt gaaggagaa 480  
 tgggaagagag gaaggagggg aggaaagaaa agtttgatga atgattgcag tctaaactgg 540  
 ttcaacaag agatcttggt taattaagga attcatocca tctctgccta ttaggaggag 600  
 gaaaaagtct aaaatagaag atggtgaaag ttgatgacc ccaggcatta aggccattca 660

10

tct 663  
 <210> 379  
 <211> 662  
 <212> ADN  
 <213> Homo sapiens  
 <400> 379

ES 2 909 841 T3

ttaagaagtc ttgaaaaatg tggagggtga aaaatctaac aggagtgcct gcttcagcag 60  
 caatttagag tagattagca tggcctctgc gccaggatga catgcacatt cctaaaaagtg 120  
 ttccgtgttt taaaaaaaaag agagagacag aatctaaggg gatgtgtaca tttgctagag 180  
 ctactataac aaagtaccag aggcagggtc acttcaacaa cagaaattta tttctcacag 240  
 ttctggaggc tagacgtcca agattaaggt gttgactggg ttgaattcag cccataacag 300  
 gaaataagga gttaataaaa gcacttgctt ctattgtttg tacctaaact taacagaaca 360  
 cagtaagtaa caagtcattg ggatgcagaa aagaaaaaag agagtgaagg aaggagaraa 420  
 ggtgaaggga gaatggaaga gaggaaggga gggaggaaag aaaagtttga tgaatgattg 480  
 cagtctaac tggttcaaac aagagatcct gtttaattaa ggaattcatc ccatctctgc 540  
 ctattaggag gaggaaaaag tctaaaatag aagatggtga aagttggatg accccaggca 600  
 ttaaggccat tcatctttaa ctgttatgct tggatcatgc aaatgtgtct ggtagctaca 660

ag 662  
 <210> 380  
 <211> 615  
 <212> ADN  
 <213> Homo sapiens  
 <400> 380

5

ttccatacat tccttcaca ccattgcctt taacctttca aattcctgct taaaactaat 60  
 cccattttta tggctgacct caccctgtat caaaaactcc gacatccctt tacgacagag 120  
 agcacaaaact agtgggtccaa aatgtcatgg gggctctctc agagttgttt tttcaatcag 180  
 gaaatttcac ataaaaatat ggatttctga tttctctttt aaaaacagaa aaacgagcca 240  
 ccagtgggag cactgcaggt atctgtgtga gaccygtact tcacaactcc tgctttccct 300  
 ccataaagta gcttgcattt tccacattga ctttgcagtt ctttggatc tgtattggtt 360  
 ttaagataat ttctactata tcacatatct cctcacagta caaagatc attttctttc 420  
 ccttttcttt ttaaaaaatt tgtattttta attttgtgg gtacacagta gatatttatg 480  
 gggcatatga ggtattttat aggcataata tatgtactag ggtaagtggg gtattcatca 540  
 cctcaagcat ttatcctttc tttgtgtaaa atatagcatt ttctgaacac tatgaatact 600  
 taagtacaag gatca 615

10

<210> 381  
 <211> 994  
 <212> ADN  
 <213> Homo sapiens  
 <400> 381

ES 2 909 841 T3

tcaaagtgta	acaaatattcc	tttctcata	aactagcaga	cattctatcc	cctcattatt	60
gtaacacatt	tctaatatct	ttctcaaatt	gtcttctgt	attacaatgc	actcaccttg	120
gcttagaatg	tctgagacaa	gaaaatctat	tcaccattcc	cacagatgac	tcctcactc	180
tcctcccaag	tctccatac	attccttcca	caccattgcc	ctaaccttt	caaattcctg	240
cttaaaaacta	atcccatttt	tatggctgac	ctcaccctgt	atcaaaaact	cggacatccc	300
tttacgacag	agagcacaaa	ctagtgtcc	aaaatgcat	gggggtcttc	tcagagttgt	360
tttttcaatc	aggaaatttc	acataaaaat	atggatttct	gattttctct	ttaaaaacag	420
aaaaacgagc	caccagtggg	agcactgcag	gtatctgtgt	gagacctgta	cttcacaact	480
cctgctttcc	ctccataaag	yagcttgcag	tttccacatt	gactttgcag	ttctttggta	540
tctgtattgg	ttttaagata	atcttacta	tatcacatat	ctcctcacag	tacaaagata	600
tcattttctt	tccttttct	ttttaaaaa	tttgtatttt	taatttttgt	gggtacacag	660
tagatattta	tggggcatat	gaggtatttt	ataggcatat	aatatgtact	agggtaagtg	720
gggtattcat	cacctcaagc	atctatcctt	tctttgtgta	aatatagca	ttttctgaac	780
actatgaata	cttaagtaca	aggatcaagt	cataggattt	ggaattgatt	tttaaaatat	840
gttgaccaa	gtgctcttat	catcaactt	aacatcacta	atgaaggatg	aacatcccaa	900
atctgaaaat	ccaaaatcca	aaatgctcca	taatctaaaa	cttggtgagc	accaacatga	960
tgcttaaagg	aaatgctcct	ggagcatttc	agat			994
<210>	382					
<211>	1001					
<212>	ADN					
<213>	Homo sapiens					
<400>	382					
ctatgagaaa	tatttttaaa	gtggttagga	acaattcata	gcactgacat	gttatcagta	60
aaaatagaag	aaaataaatt	aatattatga	aatattaatt	atatttcatt	aattatgtaa	120
tatgaattat	gttttagctc	aaatatttcc	caagggacaa	ttaagtaa	gaaaaataca	180
cacagattaa	aataataaat	agagaaggag	atattaatga	ggtacaaaa	gaaaaaatac	240
atgtaatcac	atgaaatgct	attatttgaa	agattaacaa	aacttgtaa	ctacctgcta	300
acttgatcaa	agaaaaaat	cgagaaacca	tatgcgcaat	taatagtaag	agggaaataa	360
acattgaaac	agaagacatt	tgaaatacca	tataagactg	ggtttcagag	ctctatgtac	420
gtaaattgat	aatgtcctgg	agaagtgcag	atgacaaaa	tggacacctt	tcaacttaga	480

5

ES 2 909 841 T3

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 aaaaagaat gcttttatac 720  
acttgggtggg aatgtaaatt agttcagtca ttgtggaaga ctttgatgat tcctagaaga 780  
cctaaataca gaactactat ttgacccaac aatcccatta ctgggtatat actcaaatga 840  
ctataaatca ttctattata aagacacatg catggatatg ttcattacag cactatgcac 900  
aatagcaaaag acttgggaatc aacatgaatg tccatcaatg atagactaga taaagaaaaat 960  
gtggtacaca tataccatgg aatactatgc agccataaaa a 1001  
<210> 383  
<211> 1001  
<212> ADN  
5 <213> Homo sapiens  
<400> 383  
tcagtaaaaa tagaagaaaa taaattaata ttatgaaata ttaattatat ttcattaatt 60  
atgtaatatg aattatgttt tagctcaaat atttcccaag ggacaattaa gtaaagtaaa 120  
aatacacaca gattaaaata ataaatagag aaggagatat taatgaggta caaaaagaaa 180  
aaatacatgt aatcacatga aatgctatta tttgaaagat taacaaaact tgtaaaactac 240  
ctgctaactt gatcaaagaa aaaaatcgag aaacatgatg cgcaattaat agtaagaggg 300  
aaataaacat tgaacagaa gacatttgaa ataccatata agactgggtt tcagagctct 360  
atgtacgtaa attgataatg tcctggagaa gtgcagatga ccaaagtga 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 tactatttga cccaacaatc ccattactgg gtatatactc 780  
aatgactat aatcattct attataaaga cacatgcatg gatatgttca ttacagcaat 840  
atgcacaata gcaaagactt ggaatcaaca tgaatgtcca tcaatgatag actagataaa 900  
gaaaatgtgg tacacatata ccatggaata ctatgcagcc ataaaaatga aggagatcat 960  
gccctttgca gggacacgaa tagaggtgga ggccattatc c 1001  
<210> 384  
10 <211> 501  
<212> ADN  
<213> Homo sapiens  
<400> 384

ES 2 909 841 T3

agttgcttga aagcaaagtt ctgcgagtag ctctctatct agaaggaggc attttattta 60  
 tgtaaggaag tcacctaaaa gaaaattcat ttgttatggg gtggcttta gagttactta 120  
 cttttaatgg aatccccag ataataataa attctgaaaa aaaaaaatca gaatcatggc 180  
 atgttaaac tggatacatt cctagaaata gatggaact gctcttgcaa aaagcttagc 240  
 acatgttaaa rcattttaga aacaatttgc caaagtttat ttagtctagt gatttcgaca 300  
 ggttaaatgg accctttgag atcttttttc ctcaagtaca aaggctcact tgcttaatga 360  
 acacagtccc agaaaagcag ggggctgaac cttggctcta ccatcttacc taagattcta 420  
 gagttagcaa agggtttcca caagcccaaa ttattatggt taatcttttc aattatctgt 480  
 gaagcattag gttggtgcaa a 501  
 <210> 385  
 <211> 501  
 <212> ADN  
 <213> Homo sapiens  
 <400> 385  
 gaggcatttt atttatgtaa ggaagtcacc taaaagaaaa ttcatttggt atgggtgaggc 60  
 ttttaagagtt acttactttt aatggaatcc cccagataat aataaattct gaaaaaaaaa 120  
 aatcagaatc atggcatggt aaaactggat acattcctag aaatagatgg aaactgctct 180  
 tgcaaaaagc ttagcacatg ttaaagcatt ttagaaacaa tttgccaag tttatattag 240  
 ctagtgattt ygacaggtta aatggaccct ttgagatctt ttttctcaa gtacaaaaggc 300  
 tcacttgctt aatgaacaca gtcccagaaa agcagggggc tgaaccttgg ctctaccatc 360  
 ttacctaaga ttctagagtt agcaaagggt ttccacaagc ccaaattatt atgtttaatc 420  
 ttttcaatta tctgtgaagc attaggttgg tgcaaaagta actgcagggt ttgacattaa 480  
 aactggcaaa aactgcaata a 501  
 <210> 386  
 <211> 703  
 <212> ADN  
 <213> Homo sapiens  
 <400> 386  
 gacaccagtt agcatattgt cgcgggggag aggggtggga aaggcgagag aacagcatgt 60  
 ggtccagagg ccatacccag atggaggctg cagtcagctc cccagtcaa ggcaaagccc 120  
 aagtcaaagc catgcttccc tcttgcccac ctgctccaat gccaccaca gagagtgcgc 180  
 cacagctcac aggatgcagg tctggttgaa tcttaacaat aactttgtaa gggaggtgtc 240  
 attagctcca ttctcctggc aggaggatga ggctcaaggc agctaaaggc ttttgctgaa 300  
 catcaagtgg tgagccagga ctcaawgcca gatcttcttg tttccctggt aggtgtatgt 360  
 agcacaactg gtatctgcag actatgctgc tggaagggct agccgtcact gttatcacag 420  
 cgactgctgc ctgagatatg ccaggctact ctgcaagaag tttacaaata taagctcact 480  
 tgatcttcat aacatactac ctaggtacaa tcattatatt tatttgacag atacagagac 540  
 agaggggaca cagaaaggat tagtaacttg ccccaaacca cacagccagc aagggtgtaag 600  
 tgagcacctg cagtctagat gagacaccac tcaaaacgtc atttttctgg cagccccgtg 660  
 cagttaccac agtggtcacc ccagtggtca gctaaaggcc aag 703  
 <210> 387

ES 2 909 841 T3

<211> 704  
 <212> ADN  
 <213> Homo sapiens  
 <400> 387  
 gcatattgtc gcgggggaga ggggtgggaa aggcgagaga acagcatgtg gtccagaggg 60  
 catacccaga tggaggctgc agtcagctcc ccagtcaaag gcaaagccca agtcaaagcc 120  
 atgcttccct cttgccacc tgctccaatg ccaccacag agagtgcgcc acagctcaca 180  
 ggatgcaggt ctggttgaat cttacaata actttgtaag ggaggtgtca ttagctccat 240  
 tctcctggca ggaggatgag gctcaaggca gctaaaggct tttgctgaac atcaagtgg 300  
 gagccaggac tcaatgccag atcttcttgt ttccctgta ggtgtwtgta gcacaactgg 360  
 tatctgcaga ctatgctgct ggaagggcta gccgtcactg ttatcacagc gactgctgcc 420  
 tgagatatgc caggtactgc tgcaagaagt ttacaaatat aagctcactt gatcttcata 480  
 acatactacc taggtacaat cattatattt atttgacaga tacagagaca gaggggacac 540  
 agaaaggatt agtaacttgc cccaaaccac acagccagca aggtgtaagt gagcacctgc 600  
 agtctagatg agacaccact caaacgtca ttttctggc agcccctgc agttaccaca 660  
 5 gtggtcacc cagtggctcag ctaaaggcca agcccacogt ttct 704  
 <210> 388  
 <211> 975  
 <212> ADN  
 <213> Homo sapiens  
 10 <400> 388  
 gacttaagac aaggggtct taatttgatt attttttct gttttatag atttctatga 60  
 aaactacaac aaaataaagt taattctatt taagtactt ttaaatgaat tgcctttgtt 120  
 agaaaaaaaa ttaagtgttt ttgtctcact ctgtcaccga ggctggagca cagtgggtgtg 180  
 atcatggctt actgcagcca tgacctcccg ggctcagggtg atcctccac ctcagcttcc 240  
 caaatagatg ggactacagt tgtgtgccac aacgcctggc taatttttgt atttttttgt 300  
 agagacaggg tctcaccagg ttgccaggc tgatcttga ctcttggct caagcgatcc 360  
 acccacctca gcctccctga gtgctgggat tacaggcatg agccagcgca ccagccaga 420  
 attacatfff tttaaatggt actgtcctag aaaatccagg atgtgcagtg atcaygtatg 480  
 aatgcatgga cctgcacaca caggagtga caaaagacc acccctgcca ggtcaccact 540  
 catatctcac ccagcccac gctagctcac actcctccc acacaccact gacctcatca 600  
 ttgctaggta ccacttgac ttctcaacag gttcaagaca attggccttc ctctctctt 660  
 ctagaaacac cctcttttct gggctttgtg taacacctgg tctttctccc ctctctggcc 720  
 acttctcagc tttcttttt ctttctttct ttttttttt ttttttttg cacttcctc 780  
 ttcctctaca tcaagcttgt ccaaccaca gcccaggaca gctttgaatg cagcetaaca 840  
 caaatcgtg agctttctta aaacattatg agatgtgtgt gtgtgtgtgt gtgtgtgtgt 900  
 gtgtgtgtgt gtgtgtgtgt gtttagctca tcagctatcg ttattgtag tgtattttat 960  
 gtgtggccca agaca 975  
 <210> 389  
 <211> 976  
 <212> ADN  
 15

ES 2 909 841 T3

<213> Homo sapiens  
 <400> 389  
 gacttttttaa tgaattgcct ttgtagaaa aaaaattaag tgtttttgtc tcaactctgtc 60  
 acccaggctg gagcacagtg gtgtgatcat ggcttactgc agccatgacc tcccgggctc 120  
 aggtgatcct cccacctcag cttcccaaat agatgggact acagttgtgt gccacaacgc 180  
 ctggctaatt tttgtatfff tttgtagaga cagggtctca ccaggttgcc caggctgatc 240  
 ttgaactcct tggctcaagc gatccacca cctcagcctc cctgagtgct gggattacag 300  
 gcatgagcca gcgcacccag ccagaattac atttttttaa atggactgt cctagaaaat 360  
 ccaggatgtg cagtgatcac gtatgaatgc atggacctgc acacacagga gtgaacaaaa 420  
 gaccacccc tgccaggta ccaactcatat ctcaccccag cccacgctag ctcacrcctc 480  
 tccccacaca ccaactgacct catcattgct aggtacccac ttgacttctc aacaggttca 540  
 agacaattgg ccttcctcgt ctcttctaga aacaccctct tttctgggct ttgtgtaaca 600  
 cctggtcttt ctcccctctc tggccacttc tcagcttttc tttttctttc tttctttttt 660  
 tttttttttt ttttgccact tcctcttcct ctacatcaag cttgtccaac ccacagccca 720  
 ggacagcttt gaatgcagcc taacacaaat tcgtaagctt tcttaaaaca ttatgagatg 780  
 tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgttta gctcatcagc 840  
 tatcgttatt gttagtgtat tttatgtgtg gcccaagaca tttcttcttc cagtgtggcc 900  
 cagggaaagcc aaaagattgg acaccctgc tctacaacat ctcaatatag gcctttttca 960  
 5 tgtttcattc tagatt 976  
 <210> 390  
 <211> 801  
 <212> ADN  
 10 <213> Homo sapiens  
 <400> 390  
 atccagacgg tgcacatact cctgctctg tctagatggg gtccacattc cctgctccgt 60  
 ctagactgtg cccatattcg ctgctggctg caaatgcgag gagttgacag cagcctcccc 120  
 tttacaaggc aggaggtgcc actgctcgc attgtctcca cctagggctt cacttgcttt 180  
 ctatctgcag acatcagagg gaccacatc tctctgttct gacacgctgt gtgttgatgg 240  
 cagagtttaa ttatccacat gcaatcttac tttccttatt cccaagtccg tggggctgcc 300  
 tcatcaaagc attgtaagaa ctgataacca tcttctagaa gtatcatagt gatattaaga 360  
 acacacatca cagatcatag taaatggctt taatttttta rcgaaatctc actactgcaa 420  
 atgcattggt gtcctagcta atgaatgcat agagtattgc ctgcaaaata ataattgaga 480  
 ttctatffff aagaagctta gaacagtaca tgggtgcatag caaagactct gtgtatgtga 540  
 agccagatff taaaatatgg taacaagtgt ctgaaaatat gtggctcaat ttgtctcccc 600  
 gttacttttc cctctcccc tttaaaatgt agaggaagga gaagaagaga taagaggttt 660  
 gtgagtgaag acaagggccc ttttaagcct gggaaagacta acgcataggt gatctccctc 720  
 tgccttaaaa ggcacaggaa tcttagtggg gaaaaagaag tggtgataaa tagccagtcc 780  
 gtgtgcctgg aatatcaaag t 801  
 <210> 391  
 <211> 801

ES 2 909 841 T3

<212> ADN  
 <213> Homo sapiens  
 <400> 391  
 ccctgctccg tctagactgt gcccatattc gctgctggct gcaaatgcga ggagttgaca 60  
 gcagcctccc ctttacaagg caggaggtgc cactgttcgc cattgtctcc acctagggct 120  
 tcacttgctt tctatctgca gacatcagag ggacccacat ctctctgttc tgacacgctg 180  
 tgtgttgatg gcagagttta attatccaca tgcaatctta ctttctctat tcccaagtcc 240  
 gtggggctgc ctcatcaaag cattgtaaga actgataacc atcttctaga agtatcatag 300  
 tgatattaag aacacacatc acagatcata gtaaatggct ttaatttttt agcgaatct 360  
 cactactgca aatgcattgt tgtcctagct aatgaatgca yagagtattg cctgcaaaat 420  
 aataattgag attctatctt taagaagctt agaacagtac atggtgcata gcaaagactc 480  
 tgtgtatgtg aagccagatt ttaaaatag gtaacaagtg tctgaaaata tgtggctcaa 540  
 tttgtctccc ggttactttt ccctctcccc ctttaaaatg tagaggaagg agaagaagag 600  
 5 ataagaggtt tgtgagtga gacaagggcc ctttaaggcc tgggaagact aacgccatag 660  
 ggatctccct ctgccttaa aggcacagga atcttagtgg ggaaaaagaa gtggtgataa 720  
 atagccagtc cgtgtgcctg gaatatcaaa gtcagtgcgt gccagggatc aactgicggg 780  
 tcacgtgcac tctgggtctc t 801  
 <210> 392  
 10 <211> 601  
 <212> ADN  
 <213> Homo sapiens  
 <400> 392  
 ttggcctggg gctgattcct ccaaagcaat gtgtctcttc gcagagtctc ttagagctgc 60  
 aaggcagtat gggatcatca gagaggatgc taggaagctt cagaaatgga ggtcctggta 120  
 gaaagggctc tttggcgtgg cctctgaaga gtccaaatgt gggacaagac cctccgaaag 180  
 cgggtggcctg gggagccaca ggtggggcag ccagcacgga agaggggtggc tttgctacca 240  
 ttgggaaaac ttatcctcca catcctcatg aggcaaacac ctttctacc ttaccgctcc 300  
 ycagtggcct ccctgttgcc ttcttattca agactaagac cctctagaat gttctttatc 360  
 ctgagtccag ctgattgtct atactaatat cagtacgggg ttagatgag gacaaccagt 420  
 gtgcctggct gccaggcacc ccctcccaa acccaggag tttctggaac attccaactc 480  
 tgcttgaggg tatccatgca gcatctacta ctgtgagcag gtggtctgat ctgtggaaaa 540  
 cttctatgat tcacctgagg gtaactgccc tttgtgattt gaaagaatga tgctaacaga 600  
 a 601  
 15 <210> 393  
 <211> 601  
 <212> ADN  
 <213> Homo sapiens  
 <400> 393

ES 2 909 841 T3

gcagagtctc ttagagctgc aaggcagtat gggatcatca gagaggatgc taggaagctt 60  
cagaaatgga ggtcctggta gaaagggctc tttggcgtgg cctctgaaga gtccaaatgt 120  
gggacaagac cctccgaaag cgggtggcctg gggagccaca ggtggggcag ccagcacgga 180  
agagggtggc tttgctacca ttgggaaaac ttatcctcca catcctcatg aggcaaacac 240  
ctttcctacc ttaccgctcc tcagtggcct ccctgttgcc ttcttattca agactaagac 300  
yctctagaat gttctttatc ctgagtcacg ctgattgtct atactaatat cagtacgggg 360  
tgtagatgag gacaaccagt gtgcctggct gccaggcacc ccctcccaa accccaggag 420  
tttctggaac attccaactc tgcttgaggg tatccatgca gcatctacta ctgtgagcag 480  
gtggtctgat ctgtggaaaa cttctatgat tcacctgagg gtaactgcc tttgtgattt 540  
gaaagaatga tgctaacaga aagtgttgtc atttctgaac ttttctgaac tctgcagcga 600  
g 601  
<210> 394  
5 <211> 1001  
<212> ADN  
<213> Homo sapiens  
<400> 394  
agatttggat ggggacacaa aaccaaacca tatcataggt taaattgtgt ctcccacccc 60  
aaaaatgtgt atgttgaagt cctaaccttc agtactcaga atgtgacatt atttggaaat 120  
agggtcattg cagatggagt tagttaagat gaggtcatta ggatgagtcc ctaatccaat 180  
atgactggtg ctcttaca aaaggggaagt ttggacacag agccatgcac atgggtggga 240  
agaatcccaa atgaacggat aggcagaggg ttggagagat gcatcaacaa ggaacaccaa 300  
agattgccag caacccccag aagctggggg agaggcctgg aacagattct ccctcacagc 360  
ctgagaggaa ccaagctggc tgacaccttg atctcaggtt accggccttg agaactgaga 420  
gaccctgggt ttctgttgtt taagcctctc aggggtgcagc actttattat ggaagcctga 480  
gctgactaat acaggtgtct ytatatctca ctgagggaaa gtgacaggaa agtaagaacc 540  
atztatgtcc aagagtccag aggagtcaac cagattctgg gggaaaagaa ggtacaatgc 600  
tggcctctcc atgcagccta gtccccaaca cttgtagggc ccagggcaag atctaagca 660  
ctctctcacc tatgcatcta tatgctgtaa ctgagataaa caaactatta aataatata 720  
gtgtcttgcc tctcaatctg acaattacac ctttataata gcaacatagg aaaataacta 780  
aaactatggt ttttaggcaa ccaaatacca gcaaaatgta ataattccta ttattagata 840  
tgtttaagtg ttctgctggt gggtcagcat ctttggtaga gtcataaaat taaaatgtac 900  
ataattaatt aaatattata tgtttattcc ctaacattta tttctgtcat ttcttttttc 960  
tttttttcag acagtctcac tcttttgccc aggccggagt g 1001  
10 <210> 395  
<211> 1001  
<212> ADN  
<213> Homo sapiens  
<400> 395

ES 2 909 841 T3

ttgtgtctcc caccceaaaa atgtgtatgt tgaagtcccta accttcagta ctcagaatgt	60
gacattatatt ggaaataggg tcattgcaga tggagttagt taagatgagg tcattaggat	120
gagtcctaa tccaatatga ctggtgctct tacaanaagg ggaagtttg acacagagcc	180
atgcacatgg gtggaagaa tcccaaatga acggataggg agagggttgg agagatgcat	240
caacaaggaa caccaaagat tgccagcaac ccccaagaagc tgggggagag gcctggaaca	300
gattctccct cacagcctga gaggaaccaa gctggtgac accttgatct caggttaccg	360
gccttgagaa ctgagagacc ctgggtttct gttgtttaag cctctcaggg tgcagcactt	420
tattatggaa gcctgagctg actaatacag gtgtctctat atctcactga gggaaagtga	480
caggaaagta agaaccattt rtgtccaaga gtccagagga gtcaaccaga ttctggggga	540
aaagaaggta caatgctggc ctctccatgc agcctagtcc ccaacacttg tagggcccag	600
ggcaagatct aaagcactct ctcacctatg catctatatg ctgtaactca gataaanaaa	660
ctattaanaa atatatgtgt cttgcctctc aatctganaa ttacaccttt ataatagcaa	720
catagganaa taactanaac tatggttttt aggcaaccaa ataccagcaa aatgtaanaa	780
ttcctattat tagatatggt taagtgttct gctggtgggt cagcatcttt ggtagagtca	840
tanaaattana atgtacanaa ttaattanaa attatatggt tattccctaa catttatttc	900
tgtcatttct tttttctttt tttcagacag tctcactctt ttgccagggc cggagtgcag	960
tggcgtgatc tcagctcact gcaacctccg cctcccaggt t	1001
<210> 396	
<211> 1218	
<212> ADN	
<213> Homo sapiens	
<400> 396	

5

ES 2 909 841 T3

gataaagaaa ggtcatcctc aatttcaatt tactttatat attctttgag aggtaaccgt	60
gtcttatctc cccccaaaat tccttttaaa aggaaatttc caaagatgct ctattctgtg	120
aataaagcat tgtgccacag ccgagaggat ccagcaatga acatgagatt gcccttgatt	180
cataaggctc acaagctagt aaggatagag aacactttaa aataaaaaaa aatagttttt	240
ggtatattta tattgtgtat ttggtataat tgagttttct acattctcat atatgtattt	300
catattttga agaataatgca gaaaataatc aagcttccaa ataaacattt ttttttaaga	360
actgcacaag tgagaattta ggagaacaga agatcagagg gctgcacrgg ctaaactaga	420
caatgagccc atgcaagtaa gttaagagga gaagcgggta agtatgcacc tgctttgtct	480
aggtgaccag caagcattta gcaatagtct tttcaaaaca acagctcctt atattgtcaa	540
atctcaagaa gtaatattta tggttaaaaa aatctcagac ccaacagaaa atccatgagg	600
gagatggttt tggaaacgca gaattttcag ctatgatatc cttttataaa caagcagata	660
ctttcccca atataattca atgcctcagt ctacctctg ctgaaaccac taacaccacc	720
actaaagctc gactatatgg gaaaatttag gtgtcacttt caaaatatgt cctagcataa	780
aggcaattaa aaaatgtaaa gcaccaaaga tgcaagagag acataaatga ataaaatac	840
tggcacgaaa gttttcaaaa gcttgggaat ctgattcaaa aaaaaataaa atcagccaag	900
cagtgttagt aagtagcca atcaggtttc aagaaggcag aaagacaaaa tcaacatcac	960
cagcatttga caccgctact gggggaaaaa agggggatgg agttcgttta tggccttttt	1020
aaaaatgcca ttacttgac aagagtcata acagagaagc actgcttatt tcagttctgt	1080
taactgtaaa tatcagagcc aacaccaga aaaagttcac cattagccaa ttggttttgc	1140
ctggccaatt ggagatggta ataggcctgc tatggatgac attctttctg atataagttg	1200
tttcttgctt tttctccc	1218
<210> 397	
<211> 1218	
<212> ADN	
<213> Homo sapiens	
<400> 397	

5

ES 2 909 841 T3

gataaagaaa ggtcatcctc aatttcaatt tactttatat attctttgag aggtaaccgt	60
gtcttatctc cccccaaat tccttttaaa aggaaatttc caaagatgct ctattctgtg	120
aataaagcat tgtgccacag ccgagaggat ccagcaatga acatgagatt gcccttgatt	180
cataaggtct acaagctagt aaggatagag aacactttaa aataaaaaaa aatagttttt	240
ggtatattta tattgtgtat ttggtataat tgagttttct acattctcat atatgtattt	300
catattttga agaatatgca gaaaataatc aagcttccaa ataaacattt ttttttaaga	360
actgcacaag tgagaattta ggagaacaga agatcagagg gctgcacggg ctaaactaga	420
caatgagccc atgcaagtaa gttaagagga gaagcgggta agtatgcacc tgctttgtct	480
aggwgaccag caagcattta gcaatagtct tttcaaaaaca acagctcctt atattgtcaa	540
atctcaagaa gtaatattta tggttaaaaa aatctcagac ccaacagaaa atccatgagg	600
gagatggttt tggaaacgca gaattttcag ctatgatatc cttttataaa caagcagata	660
ctttcccaa atataattca atgcctcagt ctacctctg ctgaaaccac taacaccacc	720
actaaagctc gactatatgg gaaaatttag gtgtcacttt caaaatatgt cctagcataa	780
aggcaattaa aaaatgtaa gcaccaaaga tgcaagagag acataaatga ataaaatatac	840
tggcacgaaa gttttcaaaa gcttgggaat ctgattcaaa aaaaaataaa atcagccaag	900
cagtgttagt aagttagcca atcaggtttc aagaaggcag aaagacaaaa tcaacatcac	960
cagcatttga caccgctact gggggaaaaa agggggatgg agttcgttta tggccttttt	1020
aaaaatgcca ttacttgac aagagtcata acagagaagc actgcttatt tcagttctgt	1080
taactgtaaa tatcagagcc aacaccaga aaaagttcac cattagccaa ttggttttgc	1140
ctggccaatt ggagatggta ataggcctgc tatggatgac attctttctg atataagttg	1200
tttcttgctt tttctccc	1218
<210> 398	
<211> 1072	
<212> ADN	
<213> Homo sapiens	
<400> 398	

ES 2 909 841 T3

cacttaaaag ctctggaac ctacgagatt atctttaaaa tcgtggggac caaatggctg	60
gccaaaggact tgtttctgta caggtgcat tgcttctctg ctgtgttcct ttttattacc	120
caagtaaccg gtatttcagc tcacaagatg agaaaatgac aaacaggcaa aataagcgtg	180
gggctgtgtg tgcaacagtt watcataaag ccatcaccag gagacgtcac tgggvcgctt	240
ctggagtcta tccgtcctaa ctttgctttc tttctttttt tttttaaatt taagttctag	300
ggtacatatg cacaacgtgc aggtttgtca cacatgtata catgtgccat gttgggtgtg	360
tgcaccatt aactcgtcat ttacattagg tgtatctcct agtgctatcc ctccccactc	420
ccccgacccc atgacaggcc ccagtggtg atgttcccct tcctgtgtcc aagtgttctc	480
attgttcaat cccacacct gagtgagaac atgccatggt tggttttttg tccttgcat	540
agtttgctga gaatgatggt ttccagcttc atccatgtcc ctacaaagga catgaactca	600
tcctttttta tggctacata gtattccatg gtgtatatgt gccacatttt cttaatccag	660
tctatcatog atggacattt gggttggtc caagtctttg ctattgtgac tagtggtgca	720
ataaatatac gtgtggatgt gtctttatag cagtttgatt tataatcctt tgggtatata	780
cccagtaacg ggatggctgg gtcaaagtgt atttctagtt ctagatcctt gaggaatcgc	840
cacactgact tcacaatgg ttgaactagt taacagtccc accaacagtg tgaaagtgtt	900
cctatttctc cacatcctct ccagcacccc attttgactt tgctataagg gaactttagc	960
atctgaacgt gcggacagct tcattgctgg cttgttacgt aacagtgttt tgtgaccatc	1020
tcatgtcata cccacacatc gaaaccagca gtttaaatgg ccagctgttt gc	1072
<210> 399	
<211> 1072	
<212> ADN	
<213> Homo sapiens	
<400> 399	
agattatctt taaaatcgtg gggaccaa at ggctggccaa ggacttgttt ctgtacaggt	60
gcgattgctt ctctgctgtg ttccctttta ttaccaagt aaccggatt tcagctcaca	120
agatgagaaa atgacaaaca ggcaaaataa gcgtagggct gtgtgtgcaa cagtttatca	180
taaagccatc accaggagac rtcactgggc gccttctgga gtctatccgt cctaactttg	240
ctttctttct tttttttttt aaatttaagt tctagggtac atatgcacaa cgtgcaggtt	300
tgtcacacat gtatacatgt gccatggtgg tgtgctgcac ccattaactc gtcatttaca	360
ttaggtgtat ctctagtgc tatccctccc cactcccccg accccatgac aggccccagt	420
gtgtgatgtt ccccttctg tgtccaagt ttctcattgt tcaatcccca cctatgagtg	480
agaacatgcc atgtttggtt ttttgcctt gcgatagttt gctgagaatg atggtttcca	540

5

ES 2 909 841 T3

gcttcatcca tgtccctaca aaggacatga actcatcctt ttttatggct acatagtatt 600  
 ccatggtgta tatgtgccac attttcttaa tccagtctat catogatgga catttggggt 660  
 ggttccaagt ctttgctatt gtgactagtg ttgcaataaa tatacgtgtg gatgtgtctt 720  
 tatagcagtt tgatttataa tcctttgggt atatacccag taacgggatg gctgggtcaa 780  
 atggtatttc tagttctaga tccttgagga atcgccacac tgacttccac aatggttgaa 840  
 ctagttaaca gtcccaccaa cagtgtgaaa gtgttcctat ttctccacat cctctccage 900  
 accccatfff gactttgcta taagggaaact ttagcatctg aacgtgcgga cagcttcatt 960  
 gctggcttgt tacgtaacag tgttttgtga ccatctcatg tcataccac acatcgaaac 1020  
 cagcagttta aatggccagc tgtttgcttg tgaaaactcc cctcggctgg ct 1072

<210> 400

<211> 948

<212> ADN

5 <213> Homo sapiens

<400> 400

aaattttctt tgctgaagtg tcttttcaaa tttttgcctt ttaaaaaaat tgagttgtct 60  
 taatattgag tcgtaagggt ctttatatat tctggctata tgcctttgt cagatatatg 120  
 tcttgcaaat attttctccc agtctgtggc ttacctttc cattttttaa ctgtgtttta 180  
 taaaaaaaaag aagttttttt agatcaaagt ccattttaat cattttttct tttatagttc 240  
 atgctttttg tgtctcattt aagaaatctt tccctactcc aatgtcacia atatattctc 300  
 tgagaagctt aacagttttt gcaactaaat ttaggtctat gatccgtttt gacttaattt 360  
 ttccatatgg tgcgatgtaa cagttgagat ttttttcta tgcaggcaga tattcaatgg 420  
 ttcaagtacc atttattgaa atggctatct tttctccact gaatgacctt ggcactttta 480  
 tcaaacatca actggccaca yacaggtgag tctacttctg gacacttacc ctgttcatt 540  
 catctgtata tctctatcct tacaccaaca cgcatagtct tgaatactag ggcaagttaa 600  
 ttttaagatg tctcctggat atgtaaaaat tatatctgag ttgaactaca gtttatttat 660  
 atatccaggc agcaaataaa tgtgagaatc tggagggtgag ggaagagatc agagatacca 720  
 ccttgaaac catcaattta gagatgattc ttaaggcagg ggactaaggg aactctgta 780  
 ggacacagac atagagaagg gaaggggctg cggcctgaac accccacctg catgctcact 840  
 cacatacttt cgtcggcctg tgtaaacgaa gtgctgggtc tccccagcct ctctcatctg 900  
 taagcagtgc caacaacgtc caacacagtt ccatccaatt tggatctg 948

<210> 401

<211> 920

<212> ADN

10 <213> Homo sapiens

<400> 401

ES 2 909 841 T3

aatttttgcc ttttaaaaa attgagttgt cttaaatattg agtcgtaagg ttctttatat 60  
attctggcta tatgtccttt gtcagatata tgtcctgcaa atattttctc ccagtctgtg 120  
gcttaccttt tccattttta aactgtgttt tataaaaaaa agaagttttt ttagatcaaa 180  
gtccatttta atcatttttt cttttatagt tcatgctttt tgtgtctcat ttaagaaatc 240  
tttcctact ccaatgtcac aaatatattc tctgagaagc ttaacagttt ttgcaactaa 300  
atthaggtct atgatccggt ttgacttaat ttttccatat ggtgtcatgt aacagttgag 360  
atttttttcc tatgcaggca gatattcaat ggttcaagta ccatttattg aaatggctat 420  
cttttctcca ctgaatgacc ttggcacttt tatcaaacat caactggcca cacacaggtg 480  
agtctacttc tggacactta ycctgttcca ttcatctgta tatctctatc cttacaccaa 540  
cacgcatagt cttgaatact agggcaagtt aattttaaga tgtctcctgg atatgtaaaa 600  
attatatctg agttgaacta cagtttattt atatatccag gcagcaaata aatgtgagaa 660  
tctggaggtg agggaagaga tcagagatac caccttgga accatcaatt tagagatgat 720  
tcttaaggca ggggactaag ggacactctg taggacacag acatagagaa ggggaagggc 780  
tgcggcctga acaccccacc tgcattgctca ctcacatact ttcgtcggcc tgtgttaacg 840  
aagtgtggtg tctcccagc ctctctcatc tgtaagcagt gccaacaacg tccaacacag 900  
ttccatcaa tttgatctg 920  
<210> 402  
<211> 701  
<212> ADN  
5 <213> Homo sapiens  
<400> 402  
tgtgtgctt ccattccata ggcacctgat cctaagtggt aaccaatccc agaactctcc 60  
ccttatttct tgcgtcatgt tttgaattga tgtgataaac aatgtgattc gagcgtctta 120  
actcagccta tgagcctctc tattctgtga ctgctggaat aggctgcttg gccatgttct 180  
tggaagctac caccatatca rggtaatttc ccacacaaca ttccagcccc tgctttcccc 240  
tctggcctta tctagggcca ttcccact caggtgaatg cagactcaa atgtactgag 300  
ctgtgtgcag gggccaggtg caaatgcttt ctgtgcatct gcacatgctg ttctacctgg 360  
gaagtcttt cctcctttca cctattttta ccttaaacct cagacatcat ctaccctgga 420  
aagtccttc tgacctcacg catctaagta ggtccccccc ataatcccta tccatgcctt 480  
ctatagtact taacatggtg acctttaatt gttcatttac ttagctctct gctctcccac 540  
actgtgaact ccttacaac agggaatgct atctctgaat gaatctttca tctccatgta 600  
acacatgcct ccaaccctac ctagcacaca atctggcata taacaggcac tcaataaacc 660  
ttcaatgaat gccttgatca agtacaagga acataagcaa a 701  
<210> 403  
10 <211> 701  
<212> ADN  
<213> Homo sapiens  
<400> 403

ES 2 909 841 T3

ttaaccaatc ccagaactct ccccttattt cttgctgcat gttttgaatt gatgtgataa 60  
 acaatgtgat tcgagcgtct taactcagcc tatgagcctc tctattctgt gactgctgga 120  
 ataggctgct tggccatggt cttggaagct accaccatat cagggtaatt tcccacacaa 180  
 cattccagcc cctgctttcc yctctggcct tatctagggc cattcccacaa ctcaggtgaa 240  
 tgcgactcc aaatgtactg agctgtgtgc aggggccagg tgcaaatgct ttctgtgcat 300  
 ctgcacatgc tgttctacct gggaagtccct ttctctcttt cacctatttt taccttaaac 360  
 ctcagacatc atctaccctg gaaagtcctt cctgacctca cgcatactag taggtcccc 420  
 ccataatccc tatccatgcc ttctatagta cttaacatgg tgacctttaa ttgttcattt 480  
 acttagctct ctgctctccc aactgtgaa ctccttacia acagggaatg tcatctctga 540  
 atgaatcttt catctccatg taacacatgc ctccaacct acctagcaca caatctggca 600  
 tataacaggc actcaataaa ccttcaatga atgccttgat caagtacaag gaacataagc 660  
 aaatttcctg tggaaaaaaa gaattgtatt aagttctttg g 701  
 <210> 404  
 <211> 601  
 <212> ADN  
 <213> Homo sapiens  
 <400> 404  
 atgttcactt acacatcttt ctttcactta attgaatcct ttatttttgt cttagaatct 60  
 tctgaatatt gaaaacagag aactatactg gaagaacata gtgtattaag actcatggag 120  
 agggagatgt gatactgtgt cactgaggtc gttccagtca taggagaaat gttaccactg 180  
 gattgaggtc tggtagcttt taaaagatga tttaattcta tgatatgtgt tcaacttgca 240  
 ctaggatagt ttttactttc acctttgttc catgcaccgc gcaaatacct gggaaccctt 300  
 rttgcccaac tcaagagcca gagtcctctg tcatcatttt gcctctctcc taagtgcacag 360  
 gactgagtgc agacttgggtg tttgtgggtg aggcattgtg actgacaggc aggcttcagt 420  
 ttatttagcg agtgtgagcc ctggcaggaa gattctcttt ctctgcttgc caggttgagg 480  
 aggcctcatt aagcagtttg aacttgtggt tttggcgtgt ctagtctctg tgcaggtggc 540  
 ttggtatcct cacaggcatt tctttggcct cacccttggg gtgactgttc acttgtgttt 600  
 g 601  
 <210> 405  
 <211> 601  
 <212> ADN  
 <213> Homo sapiens  
 <400> 405

5

10

ES 2 909 841 T3

tcttctgaat attgaaaaca gagaactata ctggaagaac atagtgtatt aagactcatg 60  
gagagggaga tgtgatactg tgtcactgag gtcggtccag tcataggaga aatgttacca 120  
ctggattgag gtctggtaca ttttaaaaga tgatttaatt ctatgatatg tgttcaactt 180  
gcactaggat agtttttact ttcacctttg ttccatgcac cgcgcaaata cctgggaacc 240  
cttggtgccc aactcaagag ccagagtcct ctgtcatcat tttgocctctc toctaagtga 300  
saggactgag tgcagacttg gtgtttgtgg gtgaggcatg tggactgaca ggcaggcttc 360  
agtttattta gcgagtgtga gccctggcag gaagattctc tttctctgct tgccaggttg 420  
aggaggcctc attaagcagt ttgaacttgt ggttttgccg tgtctagtcc tgggtgcagg 480  
ggcttggtat cctcacaggc atttctttgg cctcaccctt ggggtgactg ttcacttgtg 540  
tttgagcggc tgggactcag taggttcact ggagtaggta tttctttaga gccactggcg 600  
g 601  
<210> 406  
<211> 701  
<212> ADN  
5 <213> Homo sapiens  
<400> 406  
cagctccttg gcaagcctgc tccttcccca gcaaatggaa acaccattct gaacacctgg 60  
gcattgtctc tgatgtccct tttcatctcc ctactctcac acaatccagc tgcctctctg 120  
ccttccacgg atattaagaa cgtccacat ctctgagtc caagcccttc tcaactcacct 180  
ctttctgaa ctaatttctt yctgtttttt tccagtctc cttctgttc atgtctctcc 240  
tctgcacact tccattttct ggttcagaaa atgtcacogt cccagtcaca cttgccttat 300  
ggctgttggtg tcataaatac agttgacact tgaacaacat gggtttgaac tgcattggatt 360  
cacttataca catatTTTTT caatacaaat atatttaaaa attttgaga tttgcaacaa 420  
tttgaaaaaa cttgcagatg aacagcatag catagaaata ttgaaaaatt aagaaaaagg 480  
tatgtcatga atgcataaaa catatgcaga tactagtcta ttttaacctt tactgccata 540  
aaatatacac aaatctatta taaaaggta aagtttatca aagcttatgc acacaaacac 600  
ttatagacca tatagggagc cattcagtag agagaaatgt aagcgaacgt aaaggtgtgc 660  
tatttaatca caactgcata cacactgtac cactgcacta a 701  
<210> 407  
10 <211> 701  
<212> ADN  
<213> Homo sapiens  
<400> 407

ES 2 909 841 T3

gggcattgtc tctgatgtcc cttttcatct ccctactctc acacaatcca gctgcctctc 60  
 tgccttccac ggatattaag aacgtccacc atctcctgag tccaagccct tctcactcac 120  
 ctctttcttg aactaatttc tttctgtttt tttccagtcc tcccttctgt tcatgtctct 180  
 cctctgcaca cttccatttt stggttcaga aaatgtcacc gtcccagtca cacttgcctt 240  
 atggctgttg tgtcataaat acagttgaca cttgaacaac atggggttga actgcatgga 300  
 ttcacttata cacatatttt ttcaatacaa atatatttaa aaattttga gatttgcAAC 360  
 aatttgaaaa aacttgcaga tgaacagcat agcatagaaa tattgaaaaa ttaagaaaaa 420  
 ggtatgtcat gaatgcataa aacatatgca gatactagtc tattttaacc tttactgcca 480  
 taaaatatac acaaatctat tataaaaggt taaagtttat caaagcttat gcacacaaac 540  
 acttatagac catatagggg gccattcagt agagagaaat gtaagcgaac gtaaaggtgt 600  
 gctatttaat cacaactgca tacacactgt accactgcac taatttcaga gccacctcct 660  
 gttgtgattg tggtagagccc aagtgttgtg aggatctgct t 701  
 <210> 408  
 <211> 501  
 <212> ADN  
 <213> Homo sapiens  
 <400> 408  
 caggtagggt aagcaaatga acacaaattc aaactcggaa ttcaaaacca gcctctgtgt 60  
 attcctgagg accatactgt ctgctaagtg tagagaaagg cacatcctgg ttcaacagca 120  
 gagaaagcaa acaggaggca ctttctgtga gtcactctca ccacagggcc ctctcttttg 180  
 tgatccagcg atacttgttc acagtcaaag cccaggaaga gtggaaagat taacctttgt 240  
 gagccaaacc rtgtgacact tgattacttg acagaactaa tccttctgtc ctgatgacag 300  
 aaattcaact acacaggtac atgcaagcta atatctgttg taatgcctcc cagtttctct 360  
 ggagaattcc ttagtttcct ggacatctct gaaatgcaaa gttttggcaa cgagtctctg 420  
 aattaacctc tgaaaatctc acccagccaa gatggccttc ttgagaagac tgaagaacat 480  
 ggttggtttc aggctgagct g 501  
 <210> 409  
 <211> 604  
 <212> ADN  
 <213> Homo sapiens  
 <400> 409  
 cactttctgt gagtcatctc caccacaggg ccctctcttt tgtgatccag cgatacttgt 60  
 tcacagtcaa agcccaggaa gagtggaaag attaaccttt gtgagccaaa ccgtgtgaca 120  
 cttgattact tgacagaact aatccttctg tcctgatgac agaamttaa ctacacaggt 180

ES 2 909 841 T3

	acatgcaagc taatatctgt tgtaatgcct cccagtttct ctggagaatt ccttagtttc	240
	ctggacatct ctgaaatgca aagttttggc aacgagtctc tgaattaacc tctgaaaatc	300
	tcaccagcc aagatggcct tcttgagaag actgaagaac atggttggtt tcaggctgag	360
	ctggaagtgg tttacctccc aggagaggtt ccccacagtg gtgtttaagg catggggtgg	420
	accaacacca ggaagactca gacatcacac caccacactt caactcagtc acatccacct	480
	acattttctg aaaacaaaag gcagtctccc caaaaagcac tgagactctt gtgtaggtaa	540
	tctgagcaga caccaacttc ccagggcttc cttttatcca ggagagcttg gctgttcttt	600
	ttaa	604
	<210> 410	
	<211> 701	
	<212> ADN	
5	<213> Homo sapiens	
	<400> 410	
	ctccttccgc catggttgta agtttcctga cgcctcccag tcatgcttcc cgtacagcct	60
	gcagaactgc gagtcaatga aatccctttt ctccacaaat taccagctct caggtagttc	120
	cttacagcag cgtgggaaca gactcaagag ctgaagcaag caaggccgtt agcaaggagc	180
	gggctgggga gagcactcca ggcagagggg 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 cacaccatg gcctggcaag ctagatgggg	480
	tgaattttca cctgccacag ycgcaagtca aagccaccgg cttctctctt ctccctccca	540
	ttgctcctga cagccagggg taatatattg cctcatgtaa acagggaggc atccaccgga	600
	gaatctcccc tcagcccaca taagctctgc agagagggct gtgttgctcc agttcccacc	660
	tggacatgag cactttgaag ggcagcttcc ctcccggggt c	701
	<210> 411	
	<211> 612	
	<212> ADN	
10	<213> Homo sapiens	
	<400> 411	
	gggctgggga gagcactcca ggcagagggg 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 cacaccatg gcctggcaag ctagatgggg	300
	tgaattttca cctgccacag tgcgaagtca aagccaccgg cttctctctt ctccctccca	360
	ttgctcctga cagccagggg taatatattg cctcatgtaa acagggaggc aycaccgga	420
	gaatctcccc tcagcccaca taagctctgc agagagggct gtgttgctcc agttcccacc	480
	tggacatgag cactttgaag ggcagcttcc ctcccggggt ctggctgagc tcagggtagg	540
	cgtcagtctg catggattgg atggaggaag gctgtgcgtg gcaggagatg aactgcctct	600
15	tgggctgtgt gg	612

ES 2 909 841 T3

<210> 412  
 <211> 1001  
 <212> ADN  
 <213> Homo sapiens  
 5 <400> 412  
 ttggggaagg aagcactggg ggggaaggaag cactgggctt gggacagggc tgggcgctgc 60  
 ctcttactg gaccatgaca aggttgttac ctcaccaagg agaggtgcaa aaagcttagg 120  
 ggcttgatt tctagatttc agtgccaact atgccactta ctggctttat ccttggggaa 180  
 tttatctact ctgtgaccct cagttttttt atcttaatta ttaatacata cctcataatg 240  
 tgactgtgag gattcactta ataatatatg gaaaaccata gaatagtgcc cagcatctag 300  
 gaagtgccac agcccccttc agaagctagt gaaacctgca gaccactttt cagagtgata 360  
 ttattatfff tttctaggtt tactgagtta taattgaaaa aataaaaatg gaatatagat 420  
 gtacaacatg aagctctgat gcatatatcc attgtgaaat gatgaccaca atcaagctaa 480  
 ttaatgttat ctatcacttc wcatagttca accttttttt gtggtgagag tactgaagat 540  
 ctactctctt agcaatfff aaatctaaaa tacattatta ttaacacagt cactgtgccg 600  
 tacgttagct ctgaggacct tattcatttt atacctaaaa gtctgtatcc tttaaccaac 660  
 ctctcctaat tcccactgt catccctact gccacctctg gtaaccagcc ttctgctctg 720  
 tttctgagtc caaccttctt agattccaca tatgagtgag atcatgctgt gcagtgtttg 780  
 tttttctgtg tctggcttgc tttcacttag cataatgtcc tccaggtcca cccatgttgt 840  
 tgcaaatggc agaatcttct tcttgtaaa gactgaataa tatccctgtg tgtgcgtgca 900  
 tgtgtgtgtg tgtttgtgtg tgtgtgtgta tcacattttc ttcattccatt catccatcaa 960  
 tggacactaa gcactaaggt tgattccgta tcttggctat t 1001  
 <210> 413  
 <211> 2480  
 <212> ADN  
 10 <213> Homo sapiens  
 <400> 413  
 aacattactt ggggaaggaa gcaactggggg gaaggaagca ctgggcttgg gacagggctg 60

ES 2 909 841 T3

ggcgctgcct cttcactgga ccatgacaag gttgttacct caccaaggag aggtgcaaaa 120  
 agcttagggg cttggatttc tagatttcag tgccaactat gccacttact ggctttatcc 180  
 ttggggaatt tatctactct gtgaccctca gtttttttat cttaattatt aatacatacc 240  
 tcataatgtg actgtgagga ttcacttaat aatatatgga aaaccataga atagtgccca 300  
 gcatctagga agtgccacag ccccttcag aagctagtga aacctgcaga ccacttttca 360  
 gagtgatatt attatTTTTT tctaggttta 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 cgtagactct gaggacctta ttcattttat acctaaaagt ctgtatcctt 660  
 taaccaacct ctccataatt cccactgtca tccctactgc cacctctggt aaccagcctt 720  
 ctgctctggt tctgagtcca acctcttag attccacata tgagtgagat catgctgtgc 780  
 agtgtttgtt tttctgtgtc tggcttgctt tcacttagca taatgtcctc caggtccacc 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 atttgcattt tctgtatgat 1320  
 tagtggtgta gagcatcttt tcatattccc attggaatt cgtatatctt cctttgagaa 1380  
 atatttattc agatcttttg cccattgtta gctgagttat atgtgagttg gttttggttt 1440  
 gttgttggtt tttgtttttg ctattgagct gagttccttg tataattttg atattaaatc 1500  
 cttctcagct gtatggttga cagatacatt cttgcattct gtaagttgca tctgtaggtt 1560  
 gcaacagagt ctctttactc tgttgattgc ttgctttact gtgtgaaagc ttttttagct 1620  
 tgatgtaatt gtgtttgtct atttttgctt ttgttgcttg tacttttagt gtcatatcca 1680  
 aaaagtatt gccagacca gtgtcatccc ctatgttttc ttctagtaat tttaaagttt 1740  
 caggtcttat gtctatgtct ttaatccatt ttgagttaat ttttgtgtag ggtttaagat 1800  
 aagaatccaa ttttattttt attttttgta tatggatata caatttccc aacaccattt 1860  
 attgaaaatt ctatcctttc tttgttgtgt attaacatca gaataatatt tttaaataca 1920

ES 2 909 841 T3

	taaaattcag aagatgacaa aggaaaccaa ttacattgaa atgcatacag agttataatt	1980
	ctgaaagagc aatatatgtg cctctttgta aacacatcat atatcaaact gcagtgaccg	2040
	ttctaacaac tattgcaatt tcaaagtcatt gttgagtagg aggagtactt tgagattctg	2100
	aaacaacggt cttgtgctat gaaatatcca tgattttgat tggatgatggt atcccaggtc	2160
	ttgttaatgc tgctgtaatc tgttgcttcc attccatagt tgaataaaat gcttgatatac	2220
	tgttggaat tagtaaaaat aaaaacgtat ttttttccat ccaagttcat tctcagacct	2280
	tgaagagtca cttctctgga ttctgcagca aagttcccag ctggggcagc aagatttagg	2340
	caattgaaaa gaacatacac cttgttctca gtggcaaac acatggaaag ctttaaatgt	2400
	cagagaagaa ttctgccatt ttgctgactt tttttagtct ctcctaataa acaagtgtta	2460
	agtgacaagc ttttcagagg	2480
	<210> 414	
	<211> 601	
	<212> ADN	
5	<213> Homo sapiens	
	<400> 414	
	ccccccccc ccccgcatat ctcaggtggg cattttttaa cttaactaga taacaaaaca	60
	cagctaagac aagtcctttt ctccagcaaa gatggcaatg ctctaataac tctgagcata	120
	ttaaagattc tccaagactc tagcctctgc tgcaaaaaca catacaaata cctactacta	180
	ctgctgctgt gatgatgatg atgacagcaa tagtgagaat attttaaata tgccaggcac	240
	ggtggcaact gctttccaaa tattatcata tttaatctga tcattgccct atgaggtagg	300
	ragtattctg attcccattt tataaataag gaacccgagg cttagagagc atcgggtgact	360
	tgttcaaggt caccacagc tgtcaagtga cagaacttcg ataaaaatcc agactccttt	420
	aatggagtat ggaggggagt cagaaaacat aggaagtaag ggattgtgat tgacaatgtg	480
	tccttgcaaa gggacagggt aagagacaca agggcagctg tctgaggtgt gccattcacc	540
	agcttcagga gagaagtggc aggctacctc cagctatcca gccctatcca gccaaggaag	600
	c	601
	<210> 415	
	<211> 601	
	<212> ADN	
10	<213> Homo sapiens	
	<400> 415	
	caaacacag 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

ES 2 909 841 T3

aggtaggag tattctgatt cccattttat aaataaggaa cccgaggctt agagagcatc 300  
 rgtgacttgt tcaaggtcac ccacagctgt caagtgacag aacttcgata aaaatccaga 360  
 ctcccttaat ggagtatgga gggaggtcag aaaacatagg aagtaaggga ttgtgattga 420  
 caatgtgtcc ttgcaaaggg acagggttaag agacacaagg gcagctgtct gaggtgtgcc 480  
 attcaccagc ttcaggagag aagtggcagg ctacctccag ctatccagcc ctatccagcc 540  
 aaggaagctt gggagacatg ttagttcccg ccttcatttc catcagcaac ctcaaagcca 600  
 c 601  
 <210> 416  
 <211> 5823  
 <212> ADN  
 <213> Homo sapiens  
 <400> 416  
 5 tatttcaggc tttcttcttt ctatggataa gaaagctcct caggtggcaa caaaggccat 60  
 ttctttggaa gcaggcatgg catgtgacga aaaaaagaca tctcagaaaa gagccaagaa 120  
 taagactgga gagccactgt cagagaacag aaactgggct taatcaagga acatctcttg 180  
 ttcccagagt aggaggctgg caatattttc tcaactgaaat ttcagaattg ttatggacca 240  
 gtgactgctc tatgtgttca atttgttccc ttttcaaatg gaagcattta ttgcagacga 300  
 cctgcctctg tcccaccatt gtgtattaggt tttgtagagy gtagacaact tgccttttta 360  
 gttttagagt ttctgtatca agagaagatg tgtgtgggccc taacctagat tacaggatcc 420  
 tggacttcaa gtctgatata atgactggat gagactttga ctgtcctaga attgggatga 480  
 acatattttg ccggtgggag ggcgtgagta attgcgggta gagggcagac tgtccctcac 540  
 acctattcct tttcatggtg ccttcccaaa ctgcctctgg aggtggccac acaaatggct 600  
 ttggccattg tgaccatggg aaacttgatg cagaggctgg aaaaagcact tgcattgttc 660  
 tgtctcctct cttgttctc tacaatcaca agaaatgtct aggcaggctct gagcaggccc 720  
 aggctcatct gccatggaag aagaatggca catggaagag ggtcacattg tccaaccaa 780  
 gacgatccta gaccagccag gcccagttc atggttcaag acacatgaac atagttgcac 840  
 gaaccaagat tagttgtgta tggcccagac tagcagcagc acccatccaa cctacagact 900  
 ctgagaaata aatactagtt gtcttaagct tccaagtttc agtgtgagca ttaggtagta 960  
 acagttaatg aataagacag ataactcattt tatctgtctg gatacttata caatgatttc 1020  
 tattttttat tgatacataa tattttacat attgctgggg tacatgtgac attttgctac 1080  
 atacatagaa tgtgtaatga tccagtcagg atatctgagg tgtccatcac tttgagaatt 1140  
 tctcacttct gtgtgttggg aacaattcaa gtcgtctctt ctagttattt taaaatatac 1200  
 aatacattgt taactgtagt cttttttatt gaatgacagg acttgtacct tttatctaac 1260

ES 2 909 841 T3

tgtatgtttg tatctattaa gctagttctc tttatccctg cccctccta cccactcact	1320
cttccaacc tctaacatgt atcatcctat tctatatctc catgagatca acttctttag	1380
ctcccacata tgagcaaaaa catatgatgt ttgtctttct gtgcccggtt tatttcactt	1440
atgacctoca tttocatcca tgttactata aatgacagga tttcattcctt tttgtggcca	1500
aacagtatth catttgtgtat atatactaca ttttctttat ccattcatcc attgatgaac	1560
acttacgttg attccatatac tttgtctattg tgaatggtgc tgcaataaac atgcacgtgc	1620
agttatccct ttgatacact gatttatttt cctttggata aatacccagt agtgagattg	1680
ctggatcata cggtagttct acttttagtt tttgagacat 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 gctttttaa agatttattt tttgttgttg ttgagtttag tgcatacct	2040
ggatattagt ctcttatctg atgaagagtt tgccaatatt ttctccatt caacaggttg	2100
tctcttcatt ctgttgactg tttcctttgc tgtgcagaag cactttatat acagtcccat	2160
ttgtctatth tttagtagtc tatgcattta aggtctcagc cacaaaatct ttgcctagac	2220
cagtcctaaa gtgtttccc tatattttct tctagtagtt ttattgtttc atgtcttata	2280
tttaagtcta taatccattg tgagttgatt tttgtatatg gtgagatagg ggccttgttt	2340
cattcctctg catatagata tttaattttc tcagcaccat ttattgaagg tgccttccc	2400
tattgtatgt tcttggtgcc tttgtcaaaa ttcagttggc tataaatatg tgaatttatt	2460
tctgggttct ctatgtggtt ccattagtct atgtgtctat tttatacca atatcatgct	2520
gttttgatta ccatagcctt gtaatatatt ttgaagtcag gtagtgtgat gcctccagct	2580
ttgttctttt tgctcaggat tgctgtgcat actctggctt tttggttaca taaaaattc	2640
aggatttttg tatttctgtg aaaaatggca ttagtatttt gataggaatt gcactggatc	2700
tgtatattgt cctggacaac atggtcattt taacaatatt aattcttcta atctatgagt	2760
atgagacgtc tcccacttg tttgtgtcct cttcaatttc tttcattggt gtttcataat	2820
ctcccttcta caggcctttc acctccttgg ttaaattaat tcctaggat tttttgtag	2880
ctactgtaaa tgggactgcc ttctttctca gctagttcat ttttggtgca tagaaacct	2940
atthttgtat gttcattttc tatcctgcaa cattaccaa tttgcttate agctttaagt	3000
gtgtatthtg ctttgcttgt agagtcttct ggthtctcta aatgtaagac gatgtcatct	3060
gcaaacgggg acaatttgac ttctcttaa aaatctgtat gccttttatt cctttctctt	3120

ES 2 909 841 T3

gcctgattgc tctggctcta cctccagtac tatactgaat aaaagtggta aaagtgagca 3180  
tccttccttg tcttgctcta gttcttagag gaaatacttt cagtttttcc ccaactcagta 3240  
tgatgtagc tgtgggtcat atatagcctt tattatgtta agatatgttt cttctgtacc 3300  
tggtttggtg acagcttttt atcataaaaag gatgtagaat tttatcaaat gttttttctg 3360  
catctgttga gataatcata tggtttttgt cattccttct actgttgtga tgtatcatgt 3420  
ttattgattt gtgtatgtta aaccatcctt gtgtccttgg tataaattat acttggatcat 3480  
gggtgattat ctttttggca tctctgcgaa ttgtttgeta gctttttgtt ttgttctttt 3540  
tgagaatfff tatgtctagg ttccttagaa acaactggcct gtagttctct ttttgtgtgt 3600  
gtgtccttgt ctagtttggg gtcagggaaa tgggtgtcct gtagaatgag ttgttttttc 3660  
tttgatffff ttgcaagagt ttgaggagaa tgggtattag ttcttcttta tgtggttggt 3720  
caaattggca gtgaattcat tcagtcatga gcttttcttt ttttgggagg gttctcatta 3780  
ctgagttaat cacactgctc attactgacg tgttcagatt ttctatttct tctggaatct 3840  
cagtagttgt atgtttccag caatttatcc atttcctcta ggttttctag tttggtagta 3900  
tatagctatt cataatagtc tctgatgacg ttttgtatff ctgtgatata agttgtaatg 3960  
tctttttcat ttcctatfff atttgggtct tttcttgttt agtctagcaa ggggtttatc 4020  
tattttatct ttttgaagaa ccaacttttt gtttcattga ccctttctac gtcttttagtc 4080  
tttatttcat ttagatttgc tctgaaacttt actatgtcct tccttctaaf tttgggttgg 4140  
gtttgttctt ttctagttcc ttgaggtgca tcattgaatt gtttctttga tatctatcta 4200  
ctcatttgat gtaggtgttt attgctatac actctcccct cctagagctc cttttgttgt 4260  
gtcccatagg tcttggtagt ttgtttctat tttcatttgt ttcaaacatt ttatttccat 4320  
attaatffff atcattcagg aggagcatat tatttaattc ccatgtatff gtatagtttc 4380  
caaagtccct cttatffcta tttttactcc attgtgtctg gagaagatac ttcatatgat 4440  
ttcaatffff aaaaatttgt caagacttgt tttttgtcct aacatatggg ctatcctgga 4500  
gaatgttcca tgtgctgatg agaaaaatgt gtactcagca gttgttgagt aacatgttct 4560  
acaaatatct gttagatcca tttggcttaa agtctagttt aaatccaatg agtttttgtt 4620  
aattttgtct agatgatcat gatctgagac tgagggtgaag tccccaacaa ttatcgtgtt 4680  
ggagtctacc tctcttttta aatctagaaa tatttgcttt ataaatccgg gtggcttagt 4740  
gttgggtgca tatatttagt tgttatttcc tcattagatt gatctcttta ctattatata 4800  
ataactgttt actgcttctg gcataaagtc gtttttatgt aagtacagcc attcctgctt 4860  
gagtttagta ccatgttgac aaagggatgc atagagagtt ggtaaagcat gatttctggg 4920  
tgtctgtgtg aagggttttt gagaagagtt tagcatgagt ctgtggagtg agtgggaaga 4980  
ttctccctca atgtcagcag gaaccatcca tccactgggg gccaggtag aaaaagatga 5040

ES 2 909 841 T3

agaaatggtg aattctctct ctctcctgga gctgggtcac ccttcttctg cccttgaaca 5100  
 ggacatcaca actccaggct ctccagcctt tggactccaa gactgacacc agtgcccctc 5160  
 cccaattacc ccaggccctc aggcctttgg cctaggattg agacttacac catcagcttc 5220  
 cctggttctg aggcttctgg acttgactg ggccatacta ccagcatccc agggctctcca 5280  
 gcttgacagag agcctgttgt gggacttttc agcctccata atcaagtaag ccaatttccc 5340  
 tggatatctat atagatatac aatcatgttt tgcttaccag cctgaaaaat gtatcgctag 5400  
 atgagtctgt cattgcataa acatcatagt gtacttacac aaacctagat tctatagcct 5460  
 actacacacc tagtctataa acatgtacag catgttactg tactgaatat tgtaggcaac 5520  
 tgtaacacaa tggatgaatat ttgcatatth aaacatatct tatcattaaa aagatacagt 5580  
 aaacataagg tataaaagac aaaaaccggc acacctatat agggcactta ccataaatgc 5640  
 agcttgacag actagaagtc actcagggtg agtcagtgag cgaacgtgaa ggcctaggtt 5700  
 attactgtcc actacggtag actttatcaa cactgtacac aggctacact aaatttattt 5760  
 tttaaaaatt tgctctccaa taataaatta atcttcgcat cctttttttg ttgttcactg 5820  
 tgg 5823  
 <210> 417  
 <211> 5823  
 <212> ADN  
 <213> Homo sapiens  
 <400> 417  
 tatttcaggc tttcttcttt ctatggataa gaaagctcct caggtggcaa caaaggccat 60  
 ttctttggaa gcaggcatgg catgtgacga aaaaaagaca tctcagaaaa gagccaagaa 120  
 taagactgga gagccactgt cagagaacag aaactgggct taatcaagga acatctcttg 180  
 ttcccagagt aggaggctgg caatattttc tcaactgaaat ttcagaattg ttatggacca 240  
 gtgactgctc tatgtgttca atttgttccc ttttcaaatg gaagcattta ttgcagacga 300  
 cctgcctctg tcccaccatt gtgtattagg tttgtagagt gtagacaact tgccttttta 360  
 gttttagagt ttctgtatca agagaagatg tgtgtrggcc taacctagat tacaggatcc 420  
 tggacttcaa gtctgatata atgactggat gagactttga ctgtcctaga attgggatga 480  
 acatattttg ccggtgggag ggcgtgagta attgcgggta gagggcagac tgtccctcac 540  
 acctattcct tttcatggtg ccttcccaaa ctgcctctgg aggtggccac acaaatggct 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

5

ES 2 909 841 T3

gaaccaagat tagttgtgta tggcccagac tagcagcagc acccatccaa cctacagact 900  
 ctgagaaata aatactagtt gtcttaagct tccaagtffc agtgtgagca ttaggtagta 960  
 acagttaatg aataagacag ataatcattt tatctgtctg gatacttata caatgatttc 1020  
 tattttttat tgatacataa tattttacat attgctgggg tacatgtgac attttgctac 1080  
 atacatagaa tgtgtaatga tccagtcagg atatctgagg tgtccatcac tttgagaatt 1140  
 tctcacttct gtgtgttggg aacaattcaa gtcgtctctt ctagtatttt taaaatatac 1200  
 aatacattgt taactgtagt cttttttatt gaatgacagg acttgtacct tttatctaac 1260  
 tgtatgtttg tatctattaa gctagtcttc tttatccctg cccctccta cccactcact 1320  
 cttcccaacc tctaacatgt atcatcctat tctatatctc catgagatca acttctttag 1380  
 ctcccacata tgagcaaaaa catatgatgt ttgtctttct gtgcccggtt tatttcaett 1440  
 atgacctcca tttccatcca tgttactata aatgacagga tttcattctt tttgtggcca 1500  
 aacagtattt cattgtgtat atatactaca ttttctttat ccattcatcc attgatgaac 1560  
 acttacgttg attccatatac tttgctattg tgaatggtgc tgcaataaac atgcacgtgc 1620  
 agttatccct ttgatacact gatttatttt cctttggata aataccaggt agtgagattg 1680  
 ctggatcata cggtagttct acttttagtt tttgagacat ttccatactt ttccagtgtt 1740  
 tgtattaatt tacattccca tcaacaatgt ataagatttc cctttcctcc acatcctcac 1800  
 cagcatctgc tattttttgt ctttttaata atagtcattc taactggggg gagaggatat 1860  
 ctcgctatgg ttttgatttg catttccctg atatttaatg atattgagca tttcttcata 1920  
 taacctattg gccatttgtg tgtctttttt tttttttttt ttttttttga gaattgtcta 1980  
 ctcatttttg gctttttaaa agatttattt tttgttgttg ttgagtttag tgcatatcct 2040  
 ggatattagt ctcttatctg atgaagagtt tgccaatatt ttctccatt caacaggtg 2100  
 tctcttcatt ctgttgactg tttcctttgc tgtgcagaag cactttatat acagtccat 2160  
 ttgtctattt tttagtagtc tatgcattta aggtctcagc cacaaaatct ttgcctagac 2220  
 cagtcctaaa gtgtttcccc tatattttct tctagtagtt ttattgtttc atgtcttata 2280  
 ttttaagtcta taatccattg tgagttgatt tttgtatatg gtgagatagg ggccttgttt 2340  
 cattcctctg catatagata ttttaattttc tcagcacat ttattgaagg tgccttccc 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 taaaaatttc 2640  
 aggatttttg tatttctgtg aaaaatggca ttagtatttt gataggaatt gcactggatc 2700

ES 2 909 841 T3

tgtatattgt cctggacaac atggtcattt taacaatatt aattcttcta atctatgagt	2760
atgagacgtc ttcccacttg tttgtgtcct cttcaatttc tttcattggt gtttcataat	2820
ctccccttcta caggcctttc acctccttgg ttaaattaat tcctaggtat tttttttag	2880
ctactgtaaa tgggactgcc ttctttctca gctagttcat ttttggtgca tagaaacct	2940
atTTTTgtat gttcattttc tAtcctgcaa cattaccaa tttgcttAtc agctttaaagt	3000
gtgtatTTTT gtttgcctgt agagtcctct ggtttctcta aatgtaagac gatgtcatct	3060
gcaaaacgggg acaatttgac ttctcttaa aaatctgtat gccttttatt cctttctctt	3120
gcctgattgc tctggctcta cctccagtac tatactgaat aaaagtggta aaagtgagca	3180
tccttctctg tcttgcctca gttcttagag gaaatacttt cagtttttcc cactcagta	3240
tgatgttagc tgtgggtcat atatagcctt tattatgtta agatatgttt cttctgtacc	3300
tggtttgttg acagcttttt atcataaaa gatgtagaat tttatcaaat gttttttctg	3360
catctgttga gataatcata tggtttttgt cattccttct actgttgtga tgtatcatgt	3420
ttattgattt gtgtatgtta aaccatcctt gtgtccttgg tataaattat acttgggtcat	3480
ggtgtattat ctttttggca tcctgtcgaa ttgtttgcta gccttttgtt ttgttctttt	3540
tgagaatttt tatgtctagg ttcttagaa aactggcct gtagttctct ttttgtgtgt	3600
gtgtccttgt ctagtttggt gtcaggaaa tgggtgctt gtagaatgag ttgtttttc	3660
tttgattttt ttgcaagagt ttgaggagaa tgggtattag ttcttcttta tgtggttgg	3720
caaattggca gtgaattcat tcagtcatga gcttttcttt ttttgggagg gttctcatta	3780
ctgagttaat cacactgctc attactgac tgttcagatt ttctatttct tctggaatct	3840
cagtagtgt atgtttccag caatttatcc atttctcta ggttttctag tttggtagta	3900
tatagctatt cataatagtc tctgatgac ttttgtatt ctgtgatatc agttgtaatg	3960
tctttttcat ttctatttt atttgggtct tttcttgtt agtctagcaa ggggtttatc	4020
tattttatct ttttgaagaa ccaactttt gtttcattga ccctttctac gtcttttagtc	4080
tttatttcat ttagatttgc tctgaacttt actatgtct tccttcta at tttgggttg	4140
gtttgttctt ttctagttcc ttgaggtgca tcattgaatt gtttcttga tatctatcta	4200
ctcatttgat gtaggtgtt attgctatac actctccct cctagagctc ctttgttgt	4260
gtcccatag tcttggtag ttgtttctat tttcatttgt ttcaaactt ttatttccat	4320
attaattttt atcattcagg aggagcatat tatttaattc ccatgtatt gtatagtttc	4380
caaagtctct cttatttcta ttttactcc attgtgtct gagaagatac ttcatatgat	4440
ttcaattttt aaaaatttgt caagacttgt tttttgtcct aacatatggt ctatcctgga	4500
gaatgttcca tgtgctgatg agaaaaatgt gtactcagca gttgttgagt aacatgttct	4560
acaaatatct gttagatcca tttggtctaa agtctagttt aaatccaatg agtttttgtt	4620

ES 2 909 841 T3

aattttgtct agatgatcat gatctgagac tgaggtgaag tccccaaaca ttatcgtgtt 4680  
ggagctacc tctcttttta aatctagaaa tatttgcttt ataaatccgg gtggctagt 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 agtgggaaga 4980  
ttctccctca atgtcagcag gaaccatcca tccactgggg gccaggtag aaaaagatga 5040  
agaaatggtg aattctctct ctctcctgga gctgggtcac cttcttctg cccttgaaca 5100  
ggacatcaca actccaggct ctccagcctt tggactccaa gactgacacc agtggccctc 5160  
ccaattacc ccaggccctc aggcctttgg cctaggattg agacttacac catcagcttc 5220  
cctggttctg aggccttctg acttgactg gcccatacta ccagcatccc agggctccta 5280  
gcttgacagag agcctgttgt gggacttttc agcctccata atcaagtaag ccaatttccc 5340  
tggtatctat atagatatac aatcatgttt tgcttaccag cctgaaaaat gtatcgttag 5400  
atgagtctgt cattgcataa acatcatagt gtacttacac aaacctagat tctatagcct 5460  
actacacacc tagtctataa acatgtacag catgttactg tactgaatat tgtaggcaac 5520  
tgtaacacaa tggtgaatat ttgcatattt aaacatatct tatcattaaa aagatacagt 5580  
aaacataagg tataaaagac aaaaaccggc acacctatat agggcactta ccataaatgc 5640  
agcttgacag actagaagtc actcaggtg agtcagtgag cgaacgtgaa ggcctaggtt 5700  
attactgtcc actacggtag actttatcaa cactgtacac aggctacact aaatttattt 5760  
tttaaaaatt tgctctccaa taataaatta atcttcgcat cttttttttg ttgttactg 5820  
tgg 5823  
<210> 418  
<211> 707  
<212> ADN  
5 <213> Homo sapiens  
<400> 418  
aacggtgtca gctggagtga actcctgtgt gtgcaaggcc tgggtctcct ggtcagacta 60  
ctttctatgg gaaaggcata gtgtatagtc tatatactat acataggggt gctgggagga 120  
actggggttt tcacagccag ctttggtttt cattaggttt gtttagtttc cattgcttca 180  
ggggtgttag ttttgtgttc mcaactagat tataaactcc tcttgattc ctgatggcag 240  
tgacttgaag gcatttattt gaagaataat agacatacag aaaggggcgc atgtcataaa 300  
ggtacagctg gacgactttt cacaaagtga gcacatttgt atgatcgatg ttgagaccaa 360  
gagcattcag tggacaactc ctttccagtt actccacccc actcccagtg accatcattc 420  
tgacttctaa ctgtgtagac atgttttgct tgttttgtac tttacaaaca tatctactct 480  
atthtaggtg gctagacaat gtgttttaca atgctggcca tgacagtgtt tgaaagaata 540  
aatggaatc aaatagaatg ggcagtatca gagtgtgttg cctgcctaag aatgttttg 600  
tgacattttg gctttgggtc tatttacaca ttaaatctaa gagcaccaga atgtggtgtc 660  
aaaatgtgtt tggggatgaa gatattctaa agtcctgtag taagcaa 707  
<210> 419

10

ES 2 909 841 T3

<211> 712  
 <212> ADN  
 <213> Homo sapiens  
 <400> 419  
 cagactactt tctatgggaa aggcatagtg tatagtctat atactataca taggggtgct 60  
 gggaggaact ggggttttca cagccagctt tggttttcat taggtttggt tagtttccat 120  
 tgcttcaggg gtgttagttt tgtgttccca actagattat aaactcctct tgcattcctg 180  
 atggcagtga cttgaaggca tttatattgaa gaataataga catacagaaa ggggcrctatg 240  
 tcataaaggc acagctggac gacttttcac aaagtgagca catttgatg atcgatgttg 300  
 agaccaagag cattcagtgg acaactcctt tccagttact ccacccact cccagtgacc 360  
 atcattctga cttctaactg tgtagacatg ttttgcttgt tttgtacttt acaaacatat 420  
 ctactctatt tttagtggct agacaatgtg ttttacaatg ctggccatga cagtgtttga 480  
 aagaataaaa tggaatcaaa tagaatgggc agtatcagag tgtgttgctt gcctaagaaa 540  
 tgttttgtga cattttggct ttgggtctat ttacacatta aatctaagag caccagaatg 600  
 tgggtgcaaa atgtgtttgg ggatgaagat attctaagc cctgtagtaa gcaatgcaaa 660  
 5 acgttctgga ggtgtttatt aaacatttgt ttgtagaatg gagaggaaga ca 712  
 <210> 420  
 <211> 1210  
 <212> ADN  
 <213> Homo sapiens  
 10 <400> 420  
 aaagcagcac tgctctgcat tcagccttgc tacgtctcct tcagatgggc gcactagata 60  
 ctgagtgatg atcatgcctt gtctaggatc tcaccaagac agttcatgaa agagacagtg 120  
 cagctcatgg aggagatggg gcagctcaca gagaggatgg tgccatcatg gaaagcatgg 180  
 ggcagtcatg gagatgacgg rgtagctcat ggagaatata atgccatcat ggaaggcata 240  
 gtgcagtcac ggagatgatg gtgcagctca tggagaagat ggtgccatca tgggaaggcat 300  
 ggtgcaatca tggagtagac agtgcagctg ggccaagatt ctccctgact aagctcttct 360  
 caggcacctc tgagccgtcg tcttaactag gcctccagct tggcttgtga aaactgcaga 420

ES 2 909 841 T3

ctctcagcac aaatgatttg cctcctacat taagagactt aaataaacac ttgcatggct 480  
 gtgtttatatt aaacagctca aggctgtgtc cctgggatga caatgactcc agcccctaaa 540  
 attcctgctt gtgaaagctc attgctgaca gaaggatcta ccatttgctc cagccaacac 600  
 ctgggtggcag gcagataggg cctgagcccc atttaagagc agttccttta gaaagcttgc 660  
 aattgtaaat cttttctctg cccatttgag atgtaaatct tctaccacct agaactgtct 720  
 tctcaaggac ctgtgagctg actcactgaa atgcaaacat tcagggagat aactccactc 780  
 ctgtccccat gcgacggcga ggccctgact ttgggtggga ccttgctctt atttgcacca 840  
 ccacctcctg tcctaaagac atgagacggt tgtctctcct ctggataagt gcctattaac 900  
 caaccaggt gtccctggta catgaaccag tccagcctag cacctggcac tgcctttccc 960  
 tcagcacact ccagtctgta aaagtctcct tatggttgtt ttggcaaagt tgagcttagt 1020  
 taatgctaga ccccttctct actgcaatag ttactgctga ataaagtcta tccttaccac 1080  
 tttacttagt gttgggcttt gtttctcttt cataagctca tggagaagac aatgcagttc 1140  
 catcaagttt ctggctctta cactgctaac agtcagctct ggggtccctg agagggacag 1200  
 actcacacca 1210  
 <210> 421  
 <211> 1194  
 <212> ADN  
 <213> Homo sapiens  
 <400> 421  
 gcattcagcc ttgctacgtc tccttcagat gggcgacta gatactgagt gatgatcatg 60  
 ccttgtctag gatctcacca agacagttca tgaagagac agtgcagctc atggaggaga 120  
 tgggtgcagct cacagagagg atggtgccat catggaagc atggggcagt catggagatg 180  
 acggagtagc tcatggagaa kataatgccca tcatggaagg catagtgcag tcatggagat 240  
 gatggtgcag ctcatggaga agatggtgcc atcatggaag gcatggtgca atcatggagt 300  
 agacagtgca gctgggcca gattctccct gactaagctc ttctcaggca cctctgagcc 360  
 gtcgtcttaa ctaggcctcc agcttggctt gtgaaaactg cagactctca gcacaaatga 420  
 tttgcctcct acattaagag acttaataa acacttgcag ggctgtgttt atttaaacag 480  
 ctcaaggctg tgtccctggg atgacaatga ctccagcccc taaaattcct gcttgtgaaa 540  
 gctcattgct gacagaagga tctaccattt gttccagcca acacctggtg gcaggcagat 600  
 aggccctgag ccccatntaa gagcagttcc tttagaaagc ttgcaattgt aaatcttttc 660  
 tctgcccatt tgagatgtaa atcttctacc acctagaact gtcttctcaa ggacctgtga 720  
 gctgactcac tgaatgcaa acattcaggg agataactcc actcctgtcc ccatgcgacg 780  
 gcgagggcct gactttggtg ggcaccttgc tcttatttgc accaccacct cctgtcctaa 840

5

ES 2 909 841 T3

agacatgaga cgtttgtctc tcctctggat aagtgcctat taaccaaccc aggtgtcctg 900  
 gtcacatgaa ccagtcocagc ctagcacctg gcactgcctt tccctcagca cactccagtc 960  
 tgtaaaagtc tccttatggg tgttttggca aagttgagct tagttaatgc tagaccctt 1020  
 ctctactgca atagttagct ctgaataaag tctatcctta ccactttaac tagtggtggg 1080  
 ctttgtttct ctttcataag ctcatggaga agacaatgca gttccatcaa gtttctggct 1140  
 cttacactgc taacagtcag ctctggggtc cctgagaggg acagactcac acca 1194  
 <210> 422  
 <211> 1194  
 <212> ADN  
 <213> Homo sapiens  
 <400> 422  
 5 gcatcagcc ttgctacgtc tccttcagat gggcgcaacta gatactgagt gatgatcatg 60  
 ctttgtctag gatctcacca agacagttca tgaaagagac agtgcagctc atggaggaga 120  
 tgggtgcagct cacagagagg atgggtccat catggaagc atggggcagt catggagatg 180  
 acggagtagc tcatggagaa kataatgcca tcatggaagg catagtgcag tcatggagat 240  
 gatggtgcag ctcatggaga agatggtgcc atcatggaag gcatggtgca atcatggagt 300  
 agacagtgca gctgggcaa gattctccct gactaagctc ttctcaggca cctctgagcc 360  
 gtcgtcttaa ctaggcctcc agcttggctt gtgaaaactg cagactctca gcacaaatga 420  
 tttgcctcct acattaagag acttaataa acacttgcac ggctgtgttt atttaaacag 480  
 ctcaaggctg tgtccctggg atgacaatga ctccagcccc taaaattcct gcttgtgaaa 540  
 gctcattgct gacagaagga tctaccattt gttccagcca acacctggtg gcaggcagat 600  
 aggcctgag ccccatttaa gagcagttcc tttagaaagc ttgcaattgt aaatcttttc 660  
 tctgccatt tgagatgtaa atcttctacc acctagaact gtcttctcaa ggacctgtga 720  
 gctgactcac tgaatgcaa acattcaggg agataactcc actcctgtcc ccatgcgacg 780  
 gcgaggccct gactttggtg ggcaccttgc tcttatttgc accaccacct cctgtcctaa 840  
 agacatgaga cgtttgtctc tcctctggat aagtgcctat taaccaaccc aggtgtcctg 900  
 gtcacatgaa ccagtcocagc ctagcacctg gcactgcctt tccctcagca cactccagtc 960  
 tgtaaaagtc tccttatggg tgttttggca aagttgagct tagttaatgc tagaccctt 1020  
 ctctactgca atagttagct ctgaataaag tctatcctta ccactttaac tagtggtggg 1080  
 ctttgtttct ctttcataag ctcatggaga agacaatgca gttccatcaa gtttctggct 1140  
 cttacactgc taacagtcag ctctggggtc cctgagaggg acagactcac acca 1194  
 <210> 423  
 <211> 1118  
 <212> ADN  
 <213> Homo sapiens  
 <400> 423  
 10

ES 2 909 841 T3

accaagacag ttcatgaaag agacagtgca gctcatggag gagatggtgc agctcacaga 60  
 gaggatggtg ccatcatgga aagcatgggg cagtcatgga gatgacggag tagctcatgg 120  
 agaagataat gccatcatgg aaggcatagt gcagtcatgg agatgatggt gcagctcatg 180  
 gagaagatgg tgccatcatg raaggcatgg tgcaatcatg gagtagacag tgcagctggg 240  
 ccaagattct ccctgactaa gctcttctca ggcacctctg agccgtcgtc ttaactaggc 300  
 ctccagcttg gcttgtgaaa actgcagact ctgagcacia atgatttgcc toctacatta 360  
 agagacttaa ataaacactt gcatggctgt gtttatttaa acagctcaag gctgtgtccc 420  
 tgggatgaca atgactccag cccctaaaat tcctgcttgt gaaagctcat tgctgacaga 480  
 aggatctacc atttgttcca gccaacacct ggtggcaggc agataggccc tgagcccat 540  
 ttaagagcag ttcctttaga aagcttgcaa ttgtaaatct tttctctgcc catttgagat 600  
 gtaaactctt taccacctag aactgtcttc tcaaggacct gtgagctgac tcaactgaaat 660  
 gcaaacattc agggagataa ctccactcct gtccccatgc gacggcgagg ccctgacttt 720  
 ggtgggcacc ttgctcttat ttgaccacc acctcctgtc ctaaagacat gagacgtttg 780  
 tctctcctct ggataagtgc ctattaacca acccaggtgt cctggtcaca tgaaccagtc 840  
 cagcctagca cctggcactg cctttccctc agcacactcc agtctgtaaa agtctcctta 900  
 tggttgtttt ggcaaagtgt agcttagtta atgctagacc cttctctac tgcaatagtt 960  
 actgctgaat aaagtctatc cttaccactt taactagtgt tgggctttgt ttctctttca 1020  
 taagctcatg gagaagacaa tgcagttcca tcaagtttct ggctcttaca ctgctaacag 1080  
 tcagctctgg ggtccctgag agggacagac tcacacca 1118  
 <210> 424  
 <211> 601  
 <212> ADN  
 <213> Homo sapiens  
 <400> 424  
 gtaggggcac tgtctatact ggctgcactc tggccagtgc tgtcccaacg ctgaccctc 60  
 tggaagctaa tctggcttat aatgaggatg ctttctttag aggggactct ccatgcacag 120  
 cagaaaatcc caatggagtg gttcttccct atgtcccaa gggactggga atattctttc 180  
 agtaacaatg gccattggg ggaagaagga tgaaagtggg gtgagagacg tgaaatttgg 240  
 agaggtccct caaagattgt gatgtgcctc tcttgttcca atcacaggac aggggtataa 300  
 yggctttcct ttgaaacacg gggatgaatt taactattca ctcccagggt agattcatca 360  
 gggctctagag cttcagctaa cagcatgagg aagattccaa atgtgcccc atcagcatag 420  
 gaactgggta tgttgagtct atggtctcat aaaaccagaa gaaggacaag ggattgtggc 480  
 tccaggcttg ggagcacctt ttccttacca tgggctacag tatttattta gggtaaagga 540  
 aggaaactcc tgaggtgcta tggggtgcca gcaatttggg gcatcagtaa ttcaatgtcc 600  
 c 601  
 <210> 425  
 <211> 601  
 <212> ADN  
 <213> Homo sapiens  
 <400> 425

5

10

ES 2 909 841 T3

acgctgacct ctctggaagc taatctggct tataatgagg atgctttctt tagaggggac 60  
 tctccatgca cagcagaaaa tccaatgga gtggttcttc cctatgtccc caagggactg 120  
 ggaatattct ttcagtaaca atggccatt gggggaagaa ggatgaaagt ggggtgagag 180  
 acgtgaaatt tggagaggtc cctcaaagat tgtgatgtgc ctctcttggt ccaatcacag 240  
 gacaggggta taacggcttt cctttgaaac acggggatga atttaactat tcaactccca 300  
 rgtagattca tcagggtcta gagcttcagc taacagcatg aggaagattc caaatgtgcc 360  
 cccatcagca taggaactgg gtatgttgag tctatggtct cataaaacca gaagaaggac 420  
 aagggttgt ggctccaggc ttgggagcac ctttctcta ccatgggcta cagtatttat 480  
 ttagggtaaa ggaaggaaac tctgagggtg ctatgggtg ccagcaattt ggagcatcag 540  
 taattcaatg tcccttcagc catgtgtatt caactcctgc tgtgggtgtg gacttggtgc 600  
 a 601

5

<210> 426  
 <211> 701  
 <212> ADN  
 <213> Homo sapiens  
 <400> 426  
 ttcttgggca tcgtcatatt ctgtaaaaca aggaagctca gccagtggtg ttctaactg 60  
 acctcctttc tacatcctta ggtgttgta tgcgtgaatc acgtccccc aaaagacatg 120  
 ttcattgtcct aaccccagc acctcagaat gtgtgatctg gtttggaat aaggatcatca 180  
 cagatgaaat tagctaagac aaggatcatat tggaataggg ttggccctta atccactgtg 240  
 actggtgtcc ttttaagaag aggacacaga cacaggaggg gagagggcca tgggatgatg 300  
 cagggtgaga ctggagtgtc acagctgcaa gcaatacat ttctgtgctg tgaagccacc 360  
 catttggtgg tactacgtta aaacagctct aggaaattaa tacagatgtt gcctgtattt 420  
 ttgtttctca tattactact cattgtttta atgatgactg ttttattcat taagttgaaa 480  
 gctcctaaag cagagggacc rtatttttat gtccaactc tccttaaggc cttgcctatg 540  
 atagcacatc tcttcaatag aattgtccta actttaacag agacaacttg ggattattaa 600  
 tatggagaac aaagggttaa gctggtgcca gatgggtttc attttctcta aatctggaac 660  
 caaaggcagc aagtctatgg ggtggacgga gttcttagct c 701

10

<210> 427  
 <211> 701  
 <212> ADN  
 <213> Homo sapiens  
 <400> 427

ES 2 909 841 T3

caaggaagct cagcccagtg tgttctaaca tgacctcct tctacatcct taggtgttgt	60
tatgctgaa tcacgtcccc ccaaaagaca tgttcatgtc ctaaccccca ggacctcaga	120
atgtgtgatc tggtttgaa ataaggcat cacagatgaa attagctaag acaaggcat	180
attggaatag ggttggcct taatccactg tgactggtgt ccttttaaga agaggacaca	240
gacacaggag gggagagggc catgggatga tgcaggtgga gactggagtg ctacagctgc	300
aagcaaatac atttctgtgc tgtgaagcca cccatttggg ggtactacgt taaaacagct	360
ctaggaaatt aatacagatg ttgcctgtat ttttgttct catattacta ctcttgttt	420
taatgatgac tgttttattc attaagttga aagctcctaa agcagagga ccatatttt	480
atgtccaac tctccttaag sccttgccta tgatagcaca tctcttcaat agaattgtcc	540
taactttaac agagacaact tgggttattt aatatggaga acaaagggtt aagctggtgc	600
cagatgggtt tcattttctc taaatctgga accaaaggca gcaagtctat ggggtggacg	660
gagttcttag ctcaaccctt tggtgaggta agaagaagga t	701

**REIVINDICACIONES**

1. Un método para la secuenciación de ácidos nucleicos que comprende:
  - (a) proporcionar una muestra de ensayo que comprende moléculas de ácido nucleico, en el que dichas moléculas de ácido nucleico son moléculas de ADN genómico humano;
  - (b) realizar la reparación de extremos de las moléculas de ácido nucleico para generar ácidos nucleicos de extremos romos;
  - (c) realizar la adición de colas de dA a los ácidos nucleicos de extremos romos para generar ácidos nucleicos con cola de dA;
  - (d) ligar adaptadores a los ácidos nucleicos con cola de dA para generar una biblioteca de polinucleótidos ligados a adaptadores;
  - (e) opcionalmente amplificar la biblioteca usando cebadores de amplificación, comprendiendo dichos cebadores de amplificación una porción específica de adaptador; y
  - (f) someter la biblioteca a una secuenciación masivamente paralela;
 en el que las etapas (b), (c), y (d) son etapas consecutivas.
2. El método de la reivindicación 1, en el que dichas etapas consecutivas se realizan en ausencia de polietilenglicol.
3. El método de la reivindicación 1 o la reivindicación 2, en el que dichas etapas consecutivas se realizan en menos de 1 hora.
4. El método de cualquier reivindicación anterior, en el que dichas etapas consecutivas se realizan en el mismo tubo de reacción.
5. El método de cualquier reivindicación anterior, en el que dichos ácidos nucleicos son moléculas de ADN libre de células (ADNcf).
6. El método de la reivindicación 1, en el que dicho ADN genómico se somete a fragmentación antes de las etapas consecutivas de reparación de extremos, adición de colas de dA y ligación de adaptador a dichos ácidos nucleicos.
7. El método de cualquiera de las reivindicaciones 1-5, en el que dichos ácidos nucleicos son moléculas de ADN libre de células (ADNcf) y no se somete a fragmentación antes de las etapas consecutivas de reparación de extremos, adición de colas de dA y ligación de adaptador a dichos ácidos nucleicos.
8. El método de la reivindicación 1, en el que dicha secuenciación masivamente paralela comprende la amplificación en fase sólida para crear una celda de flujo de secuenciación de alta densidad con millones de clústeres.
9. El método de la reivindicación 1, en el que la biblioteca se amplifica en perlas y en el que cada perla comprende un cebador de amplificación directo e inverso.
10. El método de la reivindicación 1, en el que:
  - (i) dicha secuenciación es una secuenciación masivamente paralela que utiliza secuenciación por síntesis con terminadores de colorantes reversibles; o
  - (ii) dicha secuenciación es una secuenciación masivamente paralela que utiliza secuenciación por ligación.
11. El método de la reivindicación 1, en el que dicha secuenciación comprende una amplificación.
12. El método de la reivindicación 1, en el que dicha secuenciación es secuenciación de una sola molécula.
13. El método de cualquiera de las reivindicaciones 1-12, en el que la muestra:
  - (a) es una muestra de sangre periférica, o el plasma y/o las fracciones de suero del mismo;
  - (b) es una muestra de plasma derivada de sangre periférica;
  - (c) es una muestra de plasma derivada de sangre periférica que comprende una mezcla de ADNcf derivado de células normales y cancerosas;
  - (d) se deriva de una mezcla de células cancerosas y no cancerosas de un fluido biológico seleccionado de entre suero, sudor, lágrimas, esputo, orina, flujo del oído, linfa, saliva, líquido cefalorraquídeo, lavados, suspensión de médula ósea, flujo vaginal, lavado transcervical, fluido cerebral, ascitis, leche, secreciones de las vías respiratorias, intestinales y genitourinarias, y muestras de leucoforesis; o
  - (e) se selecciona de entre biopsias de tejido, hisopos o frotis.

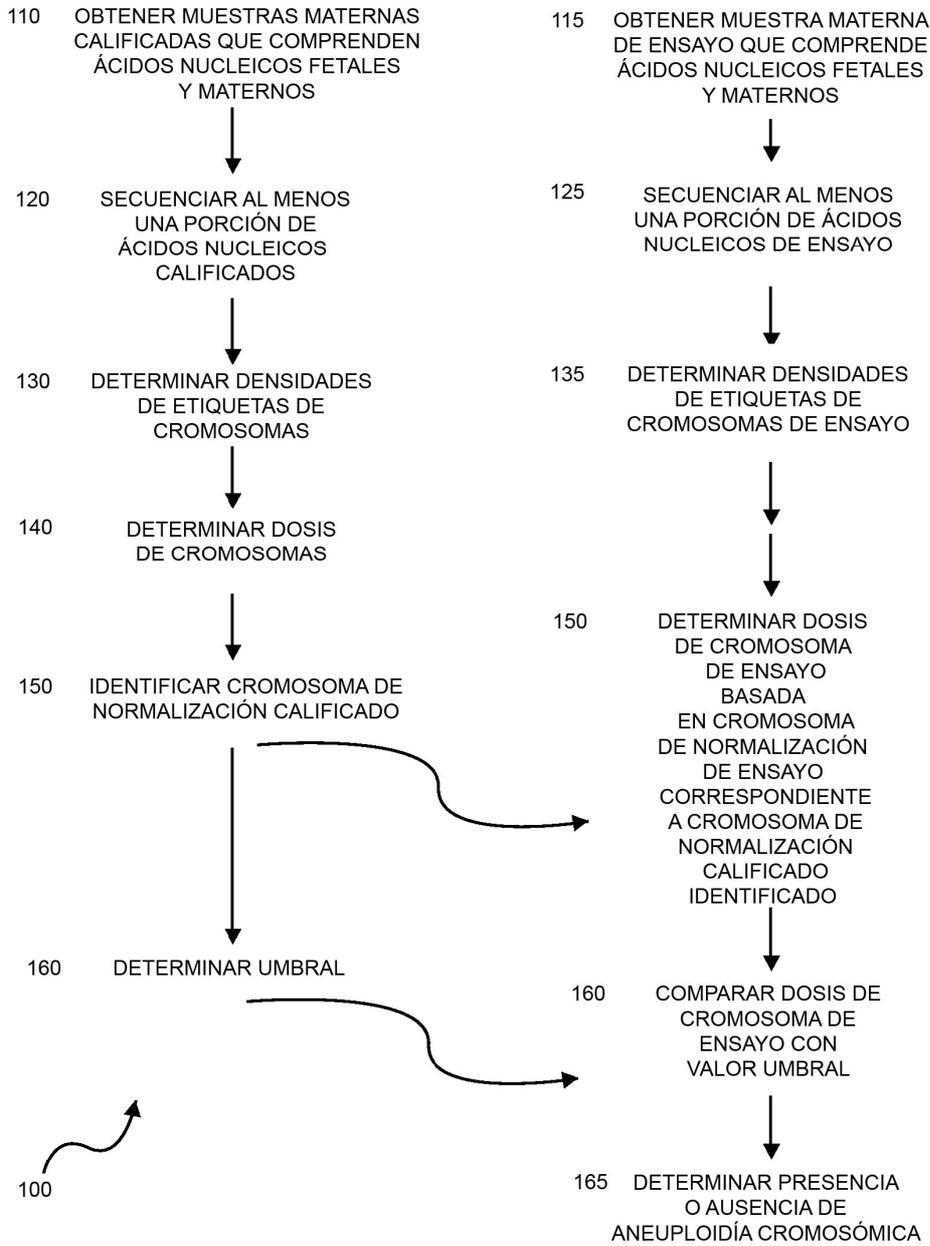
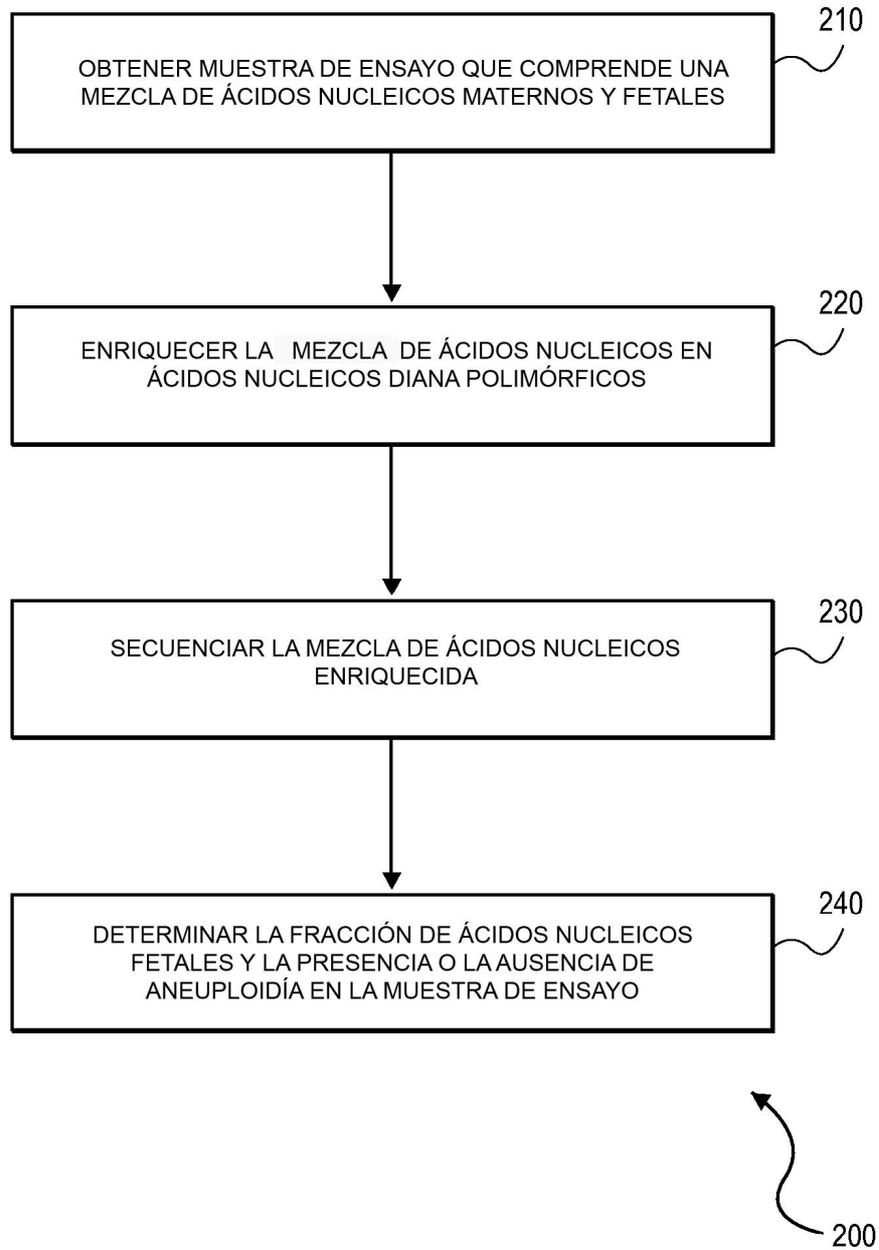
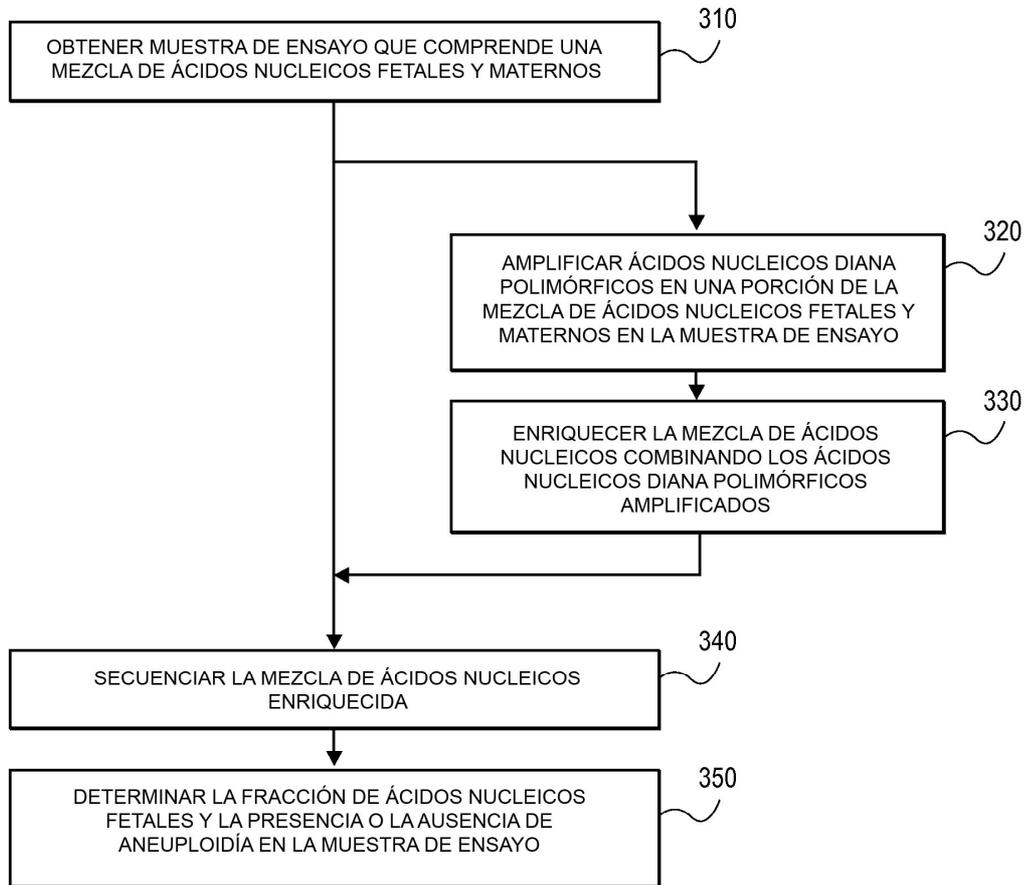


FIG. 1

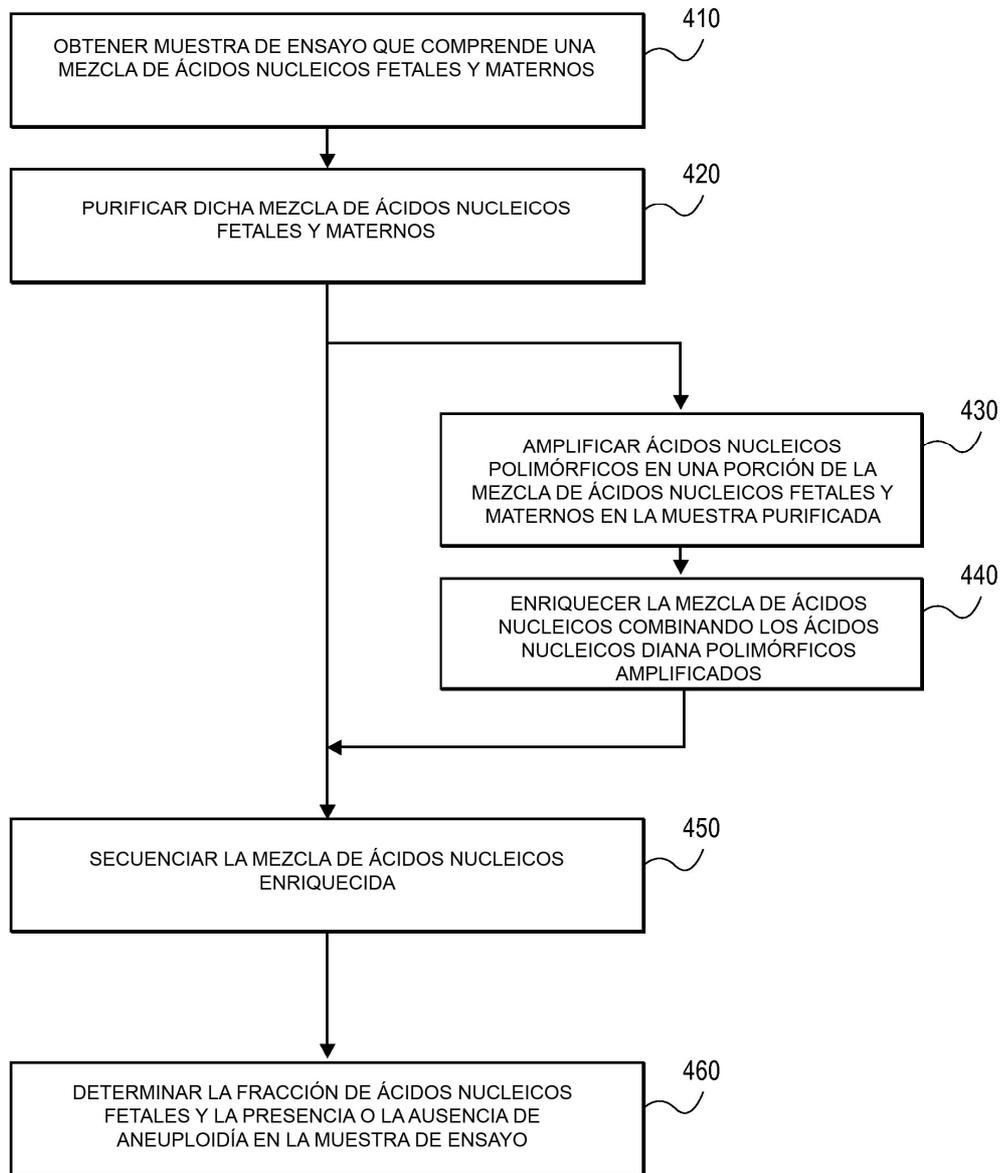


**FIG. 2**



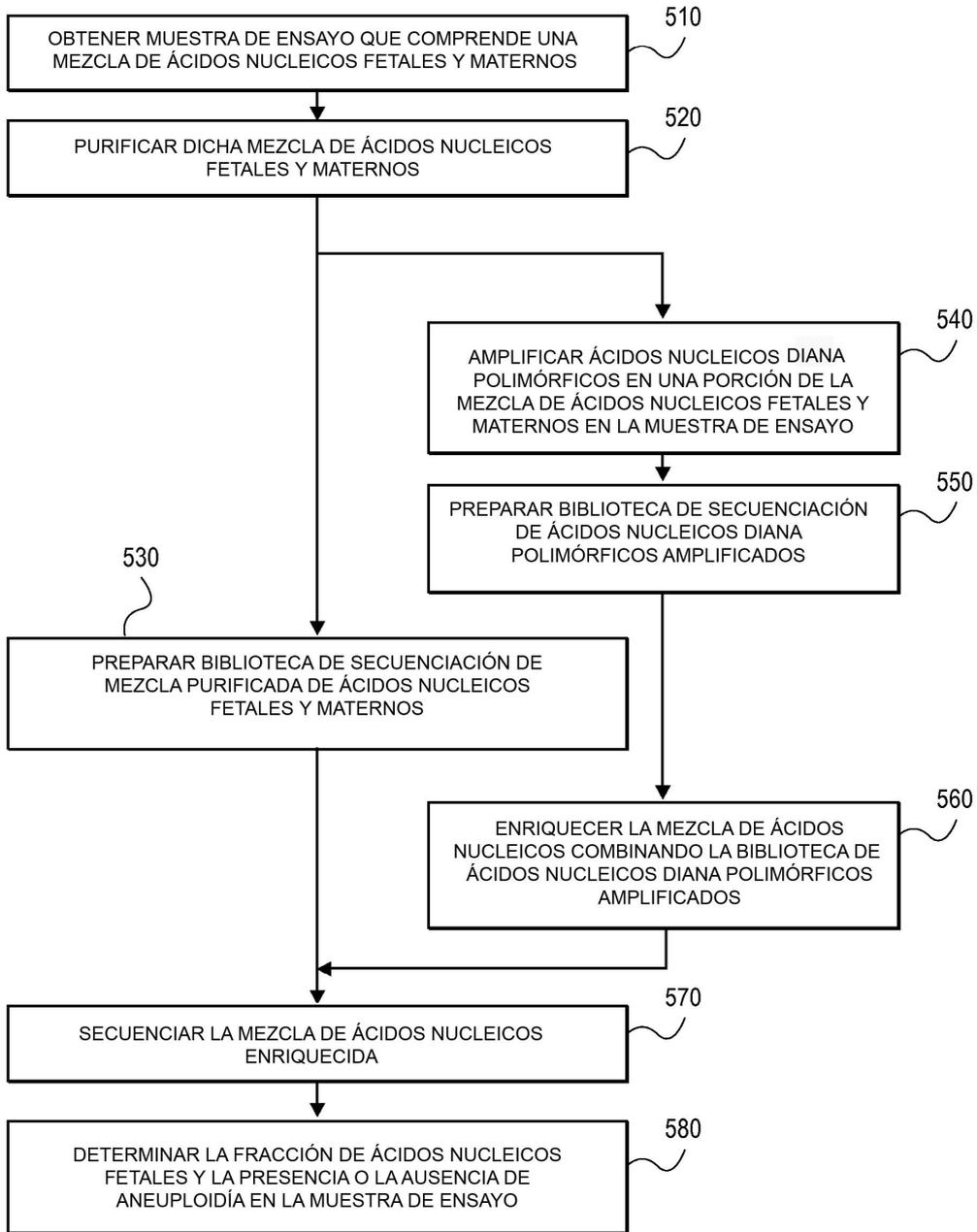
300

**FIG. 3**



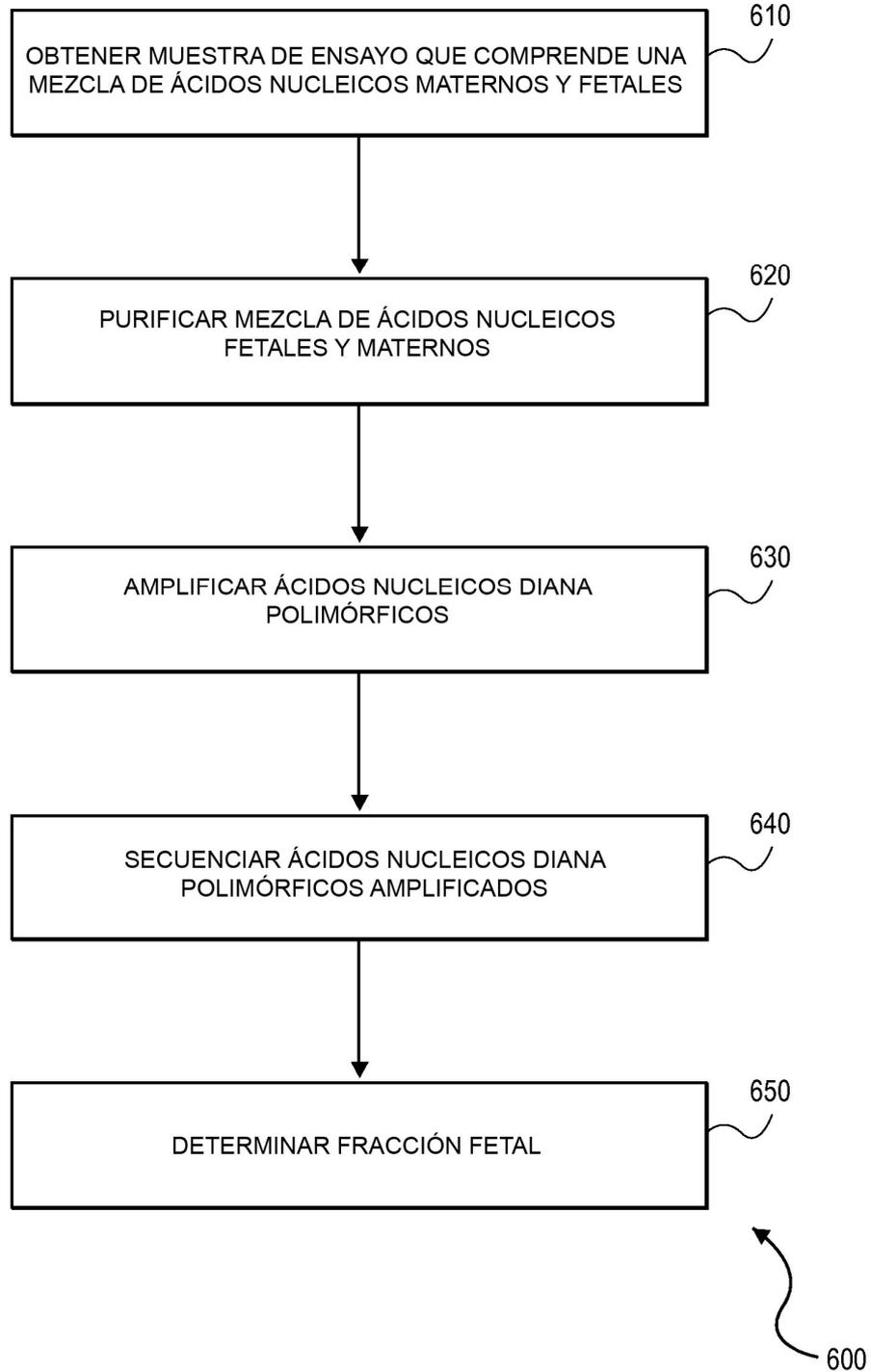
400 ↗

**FIG. 4**

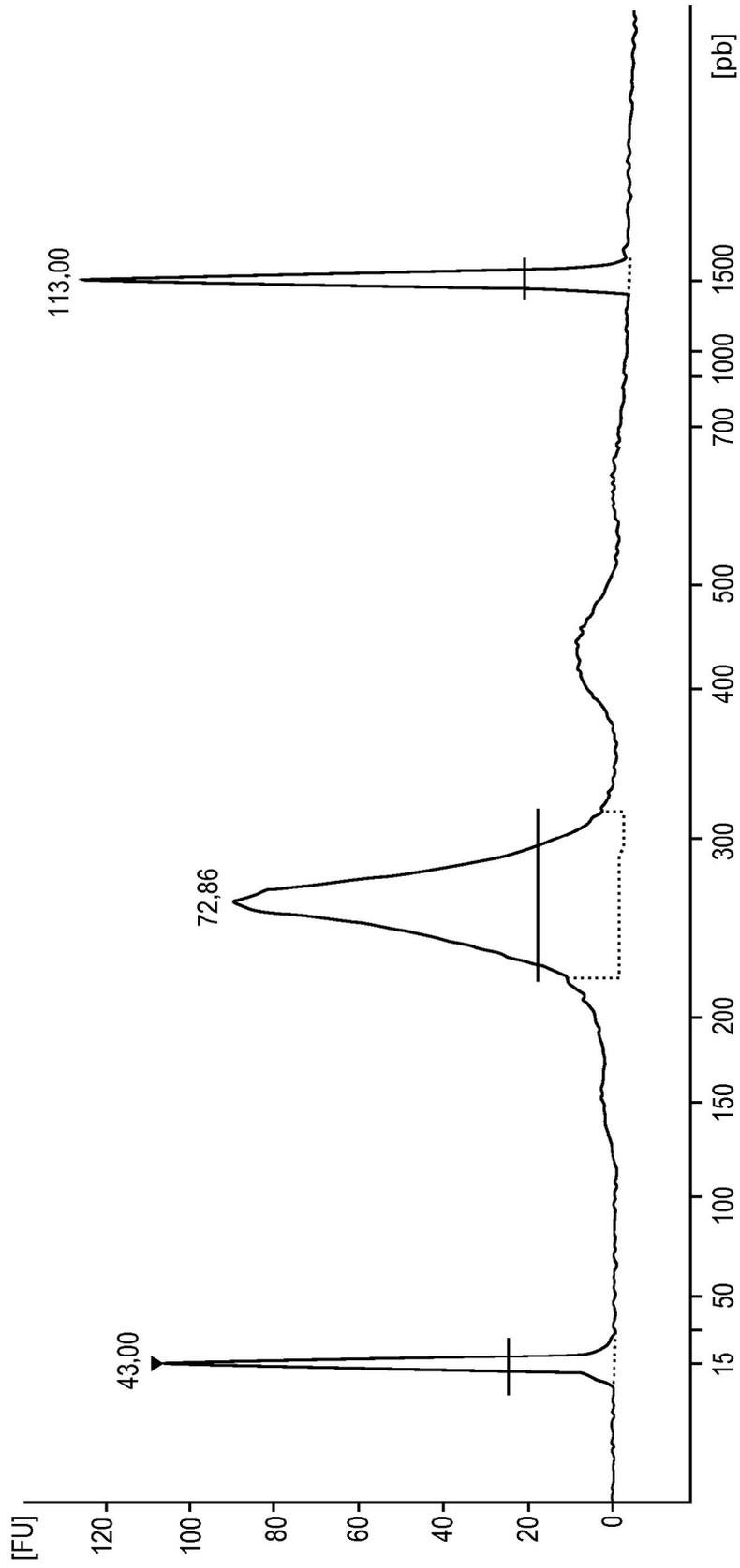


500

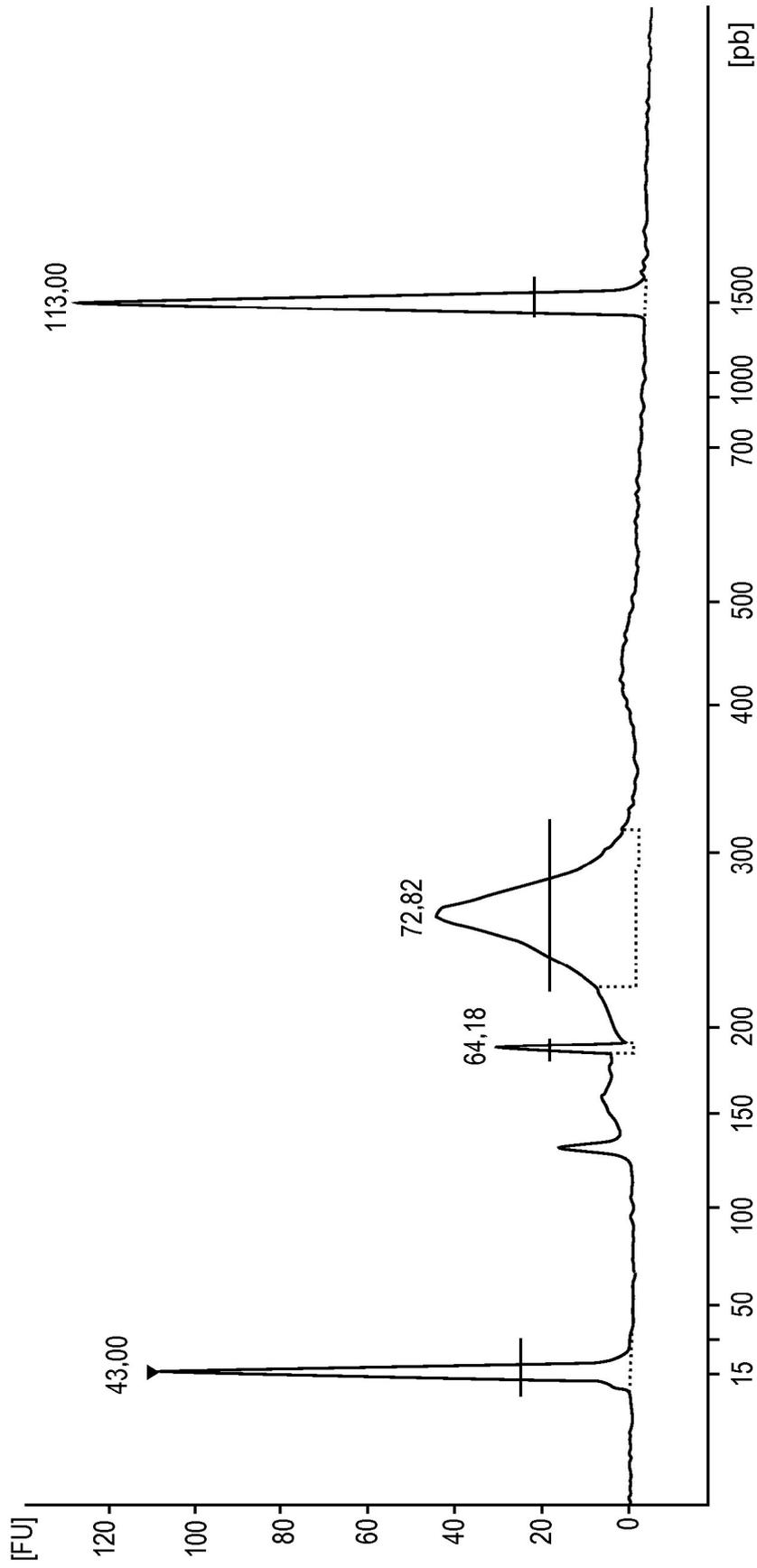
FIG. 5



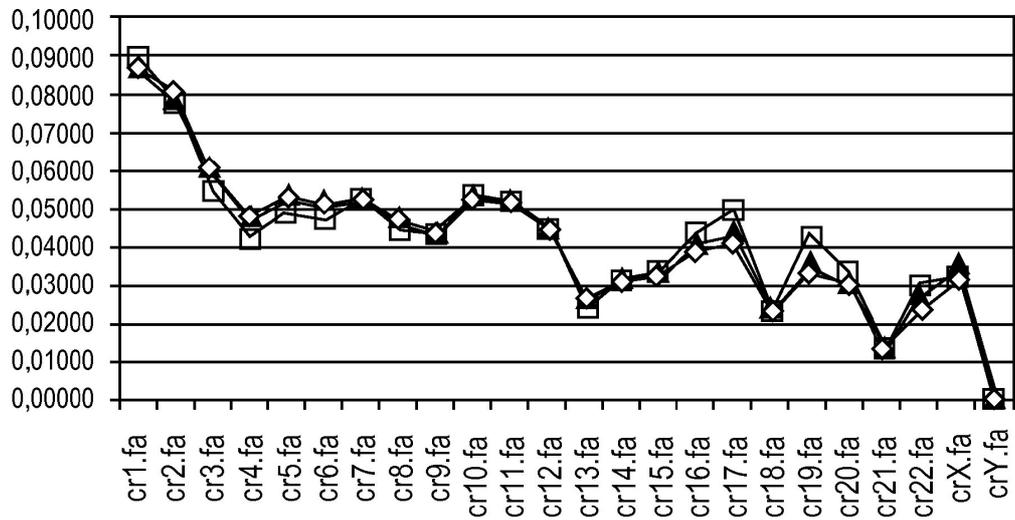
**FIG. 6**



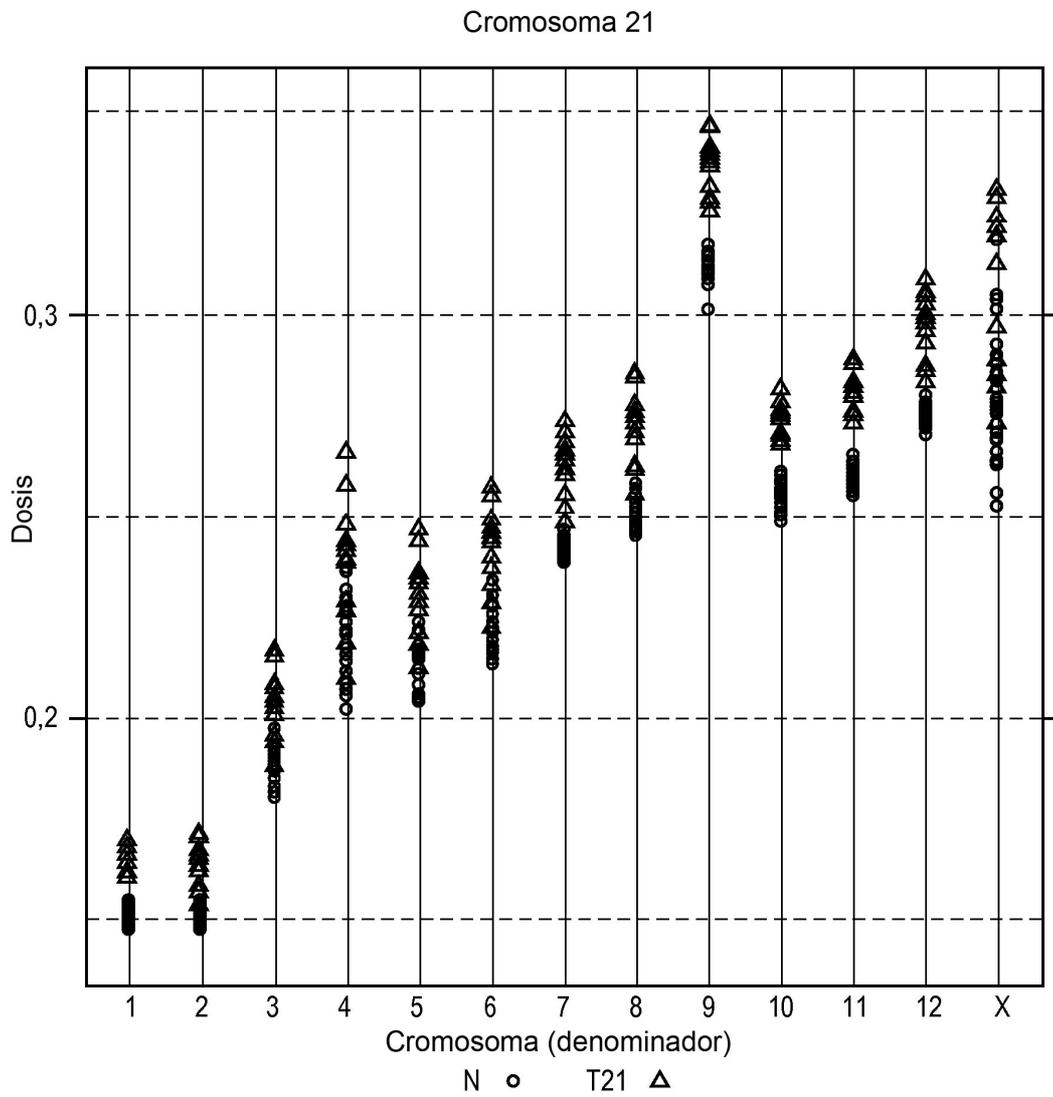
**FIG. 7A**



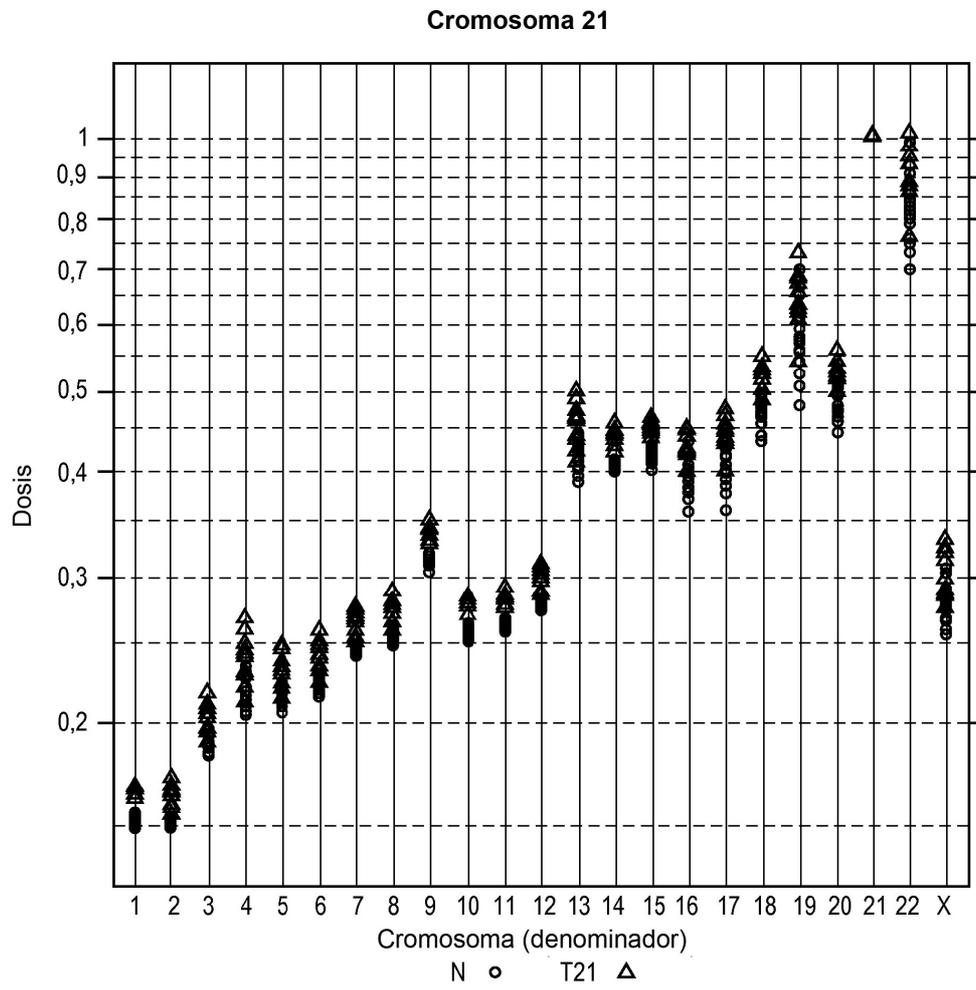
**FIG. 7B**



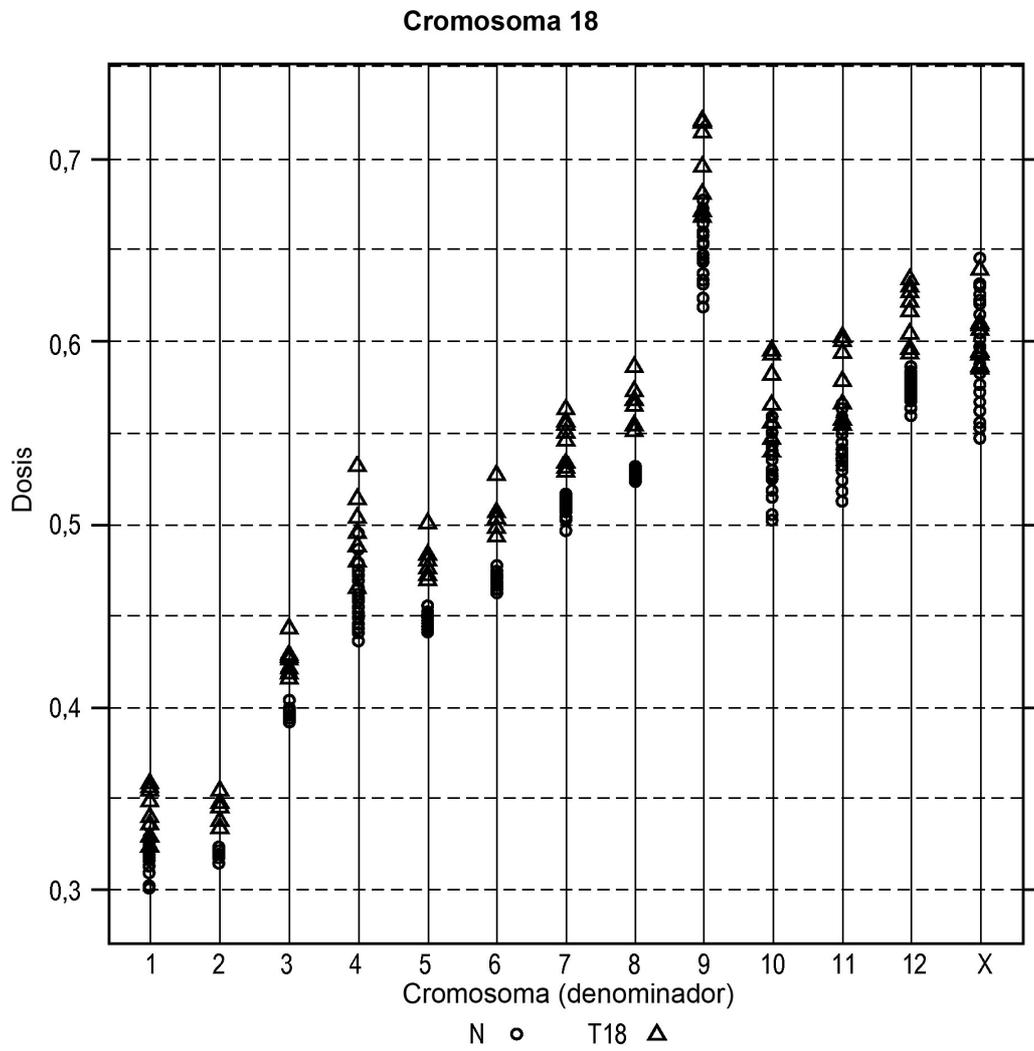
**FIG. 8**



**FIG. 9A**



**FIG. 9B**



**FIG. 10A**

Cromosoma 18

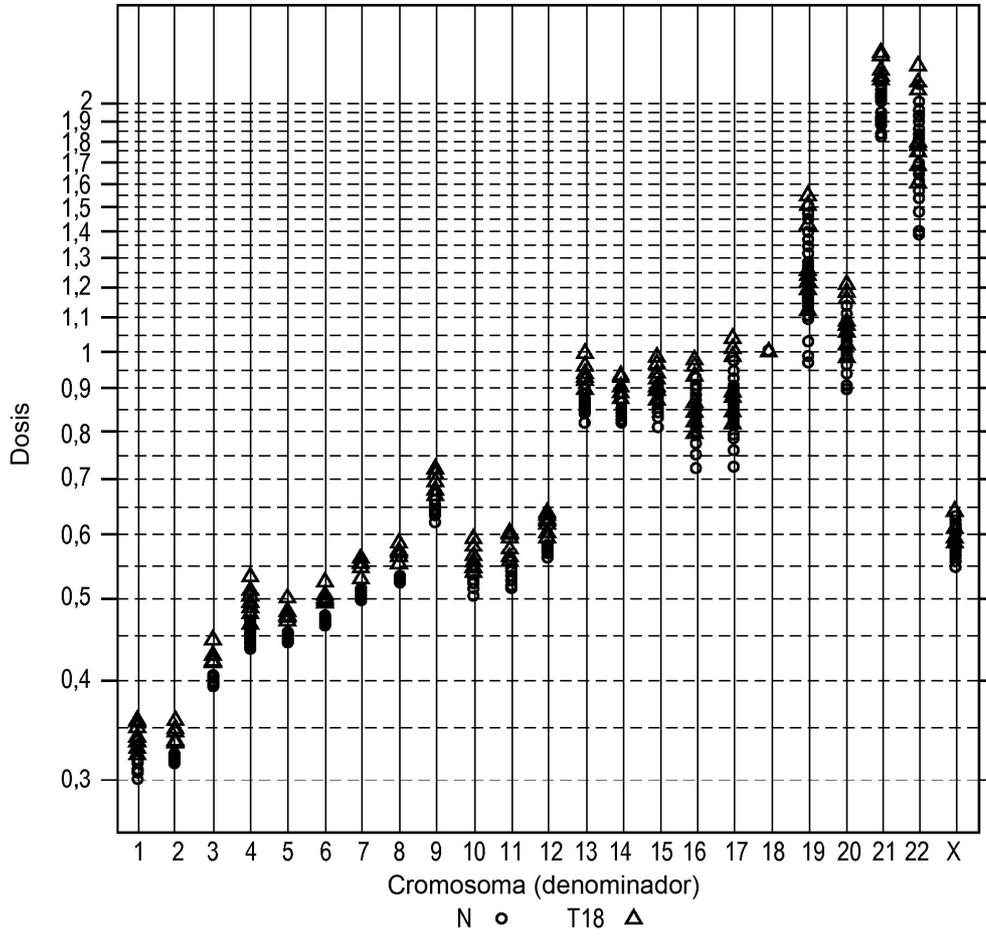
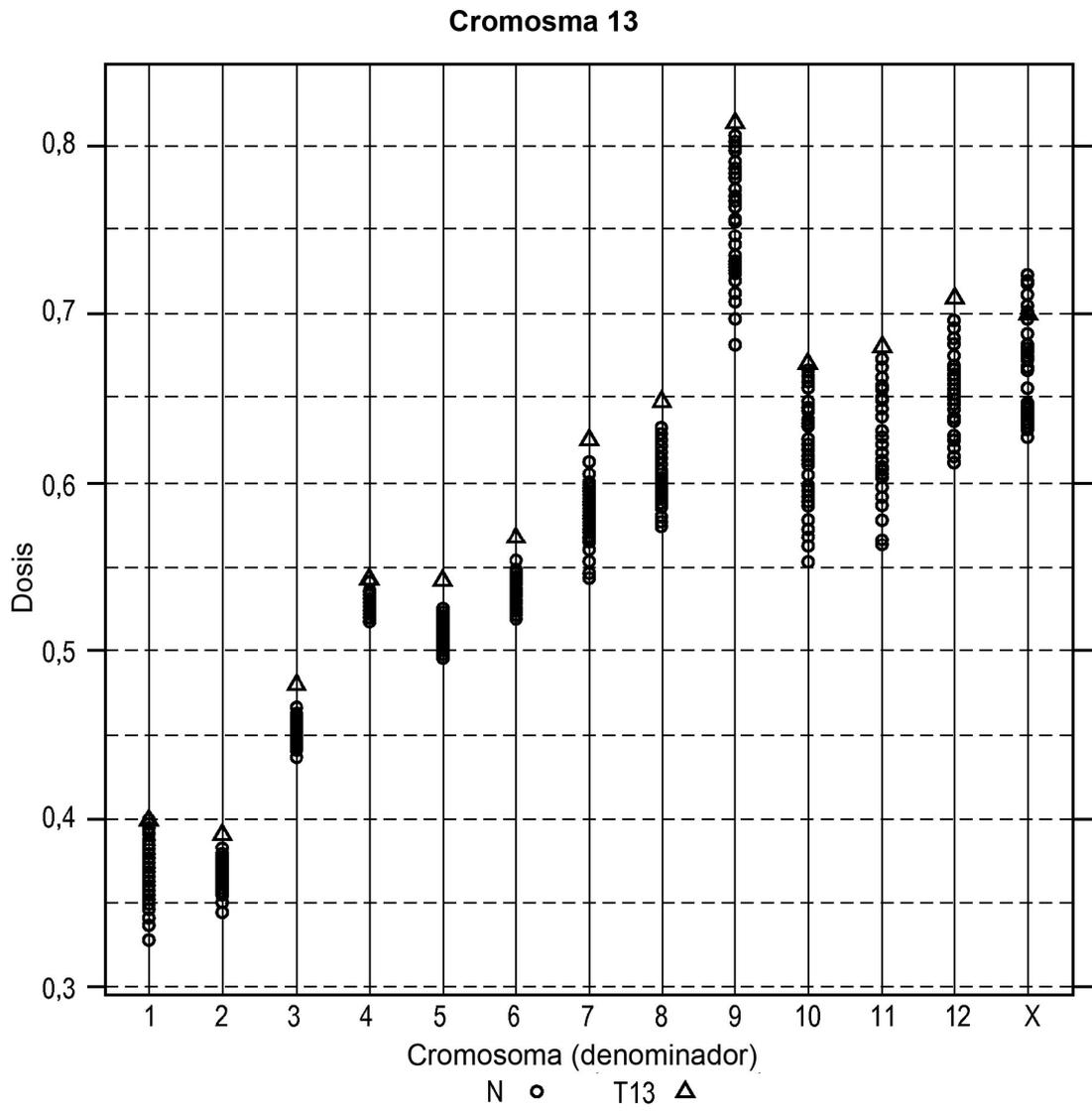


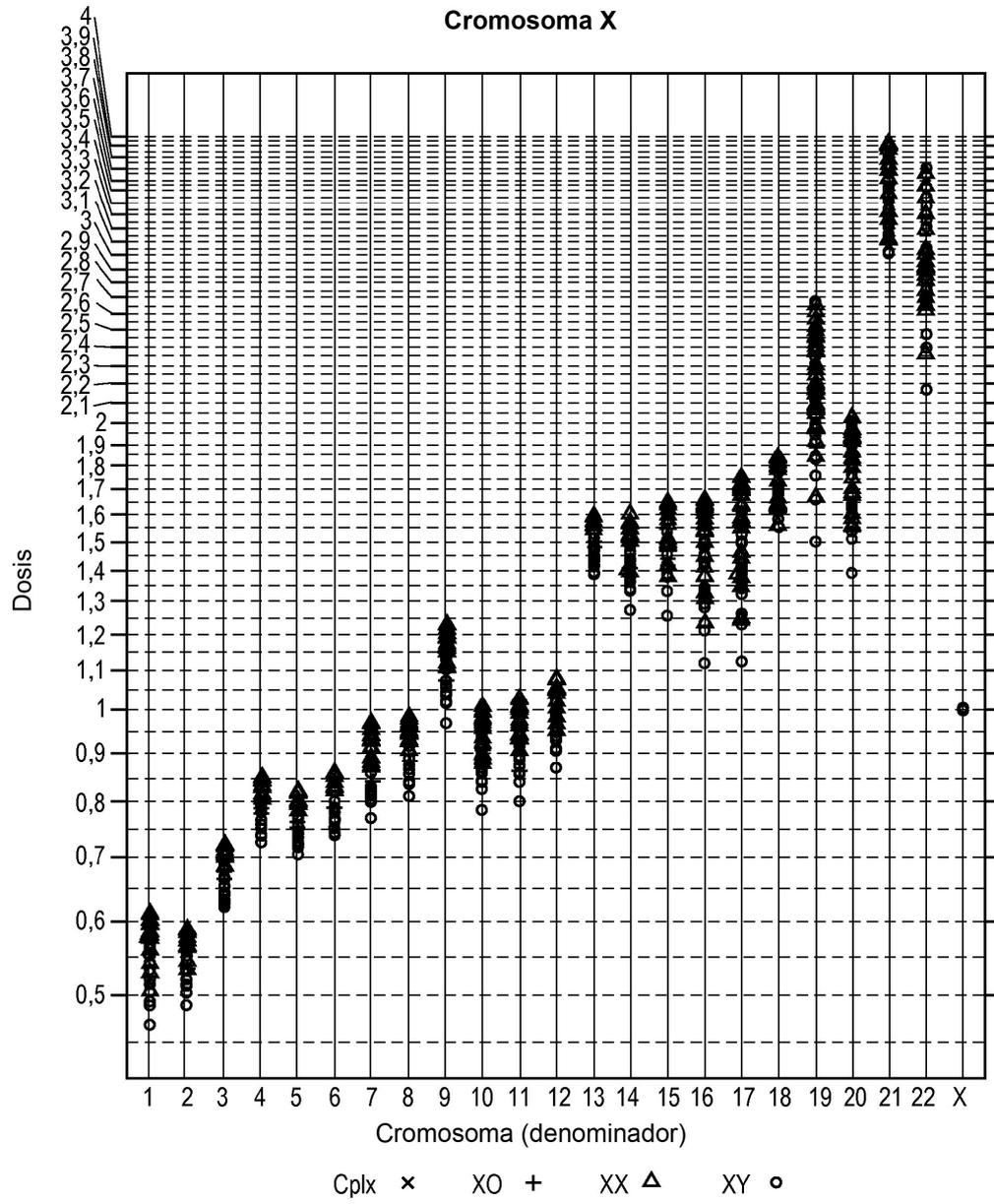
FIG. 10B



**FIG. 11A**

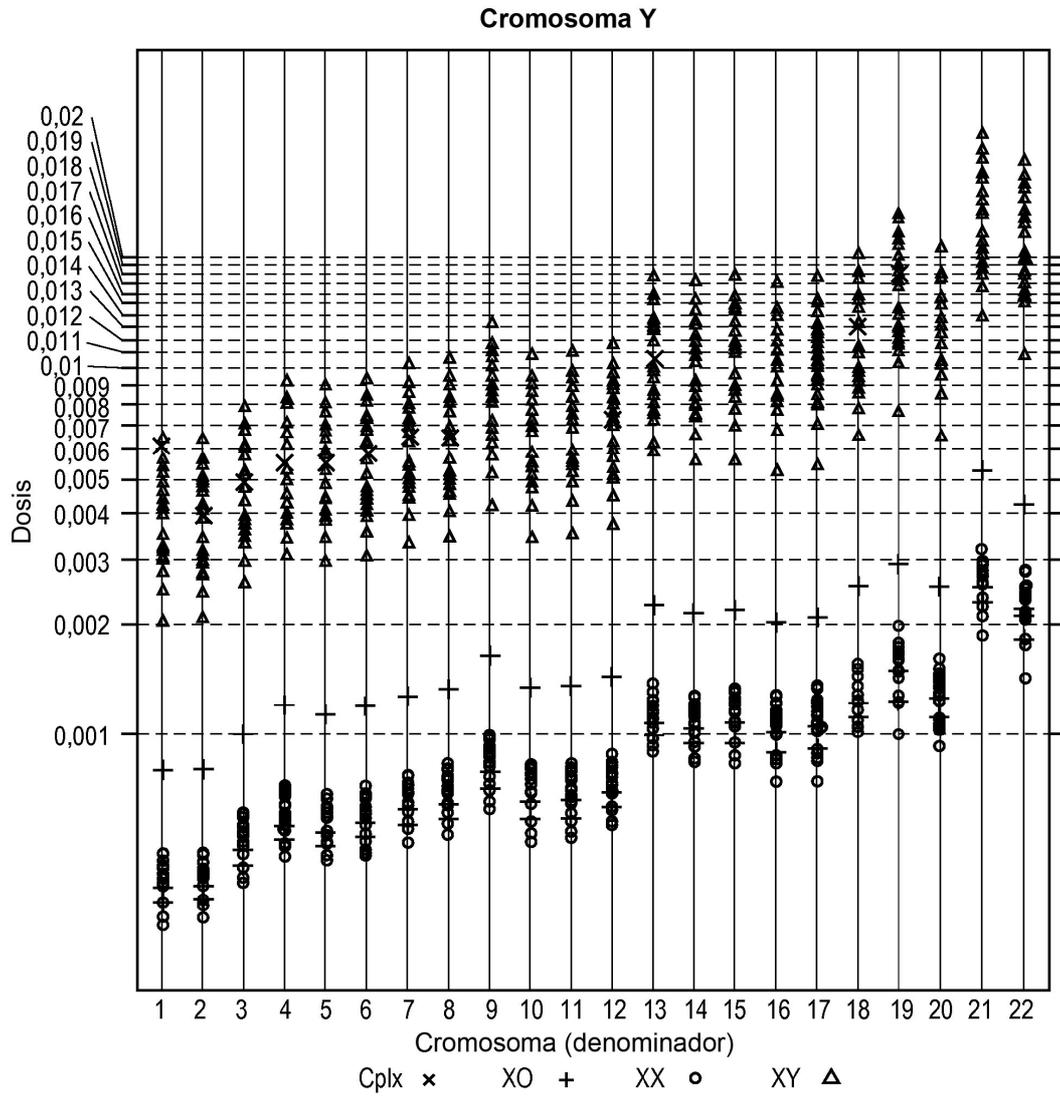




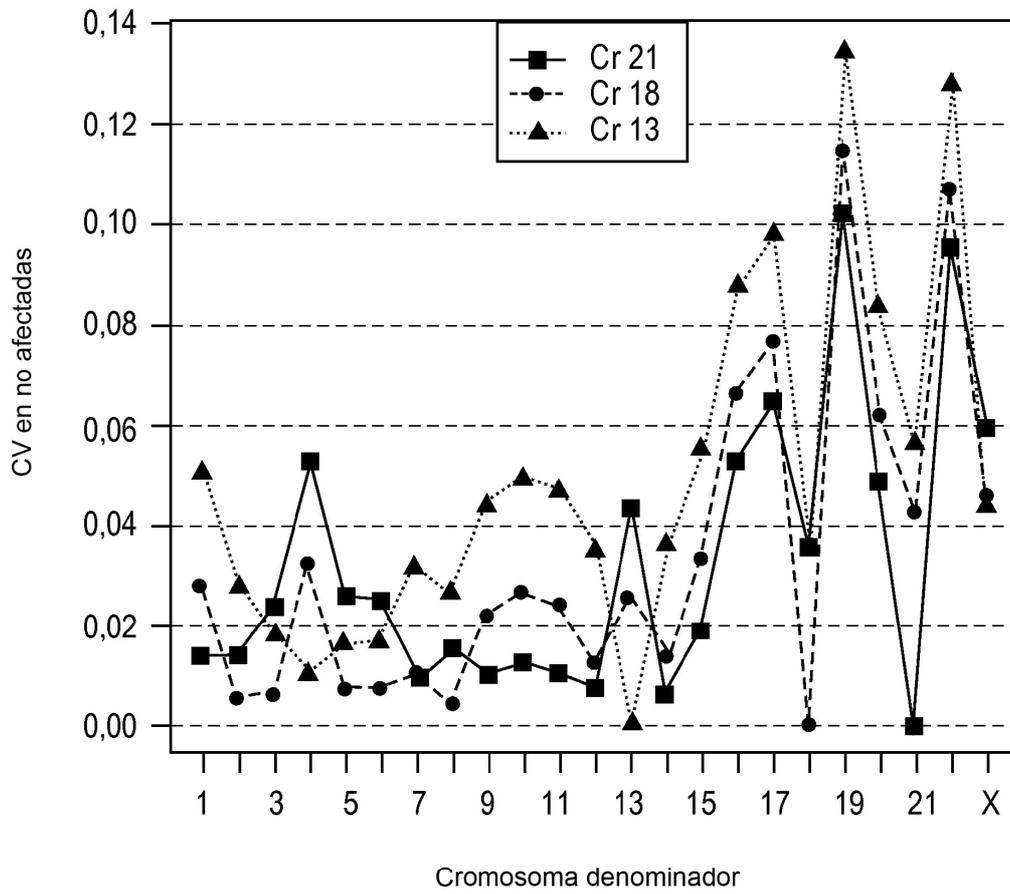


**FIG. 12B**

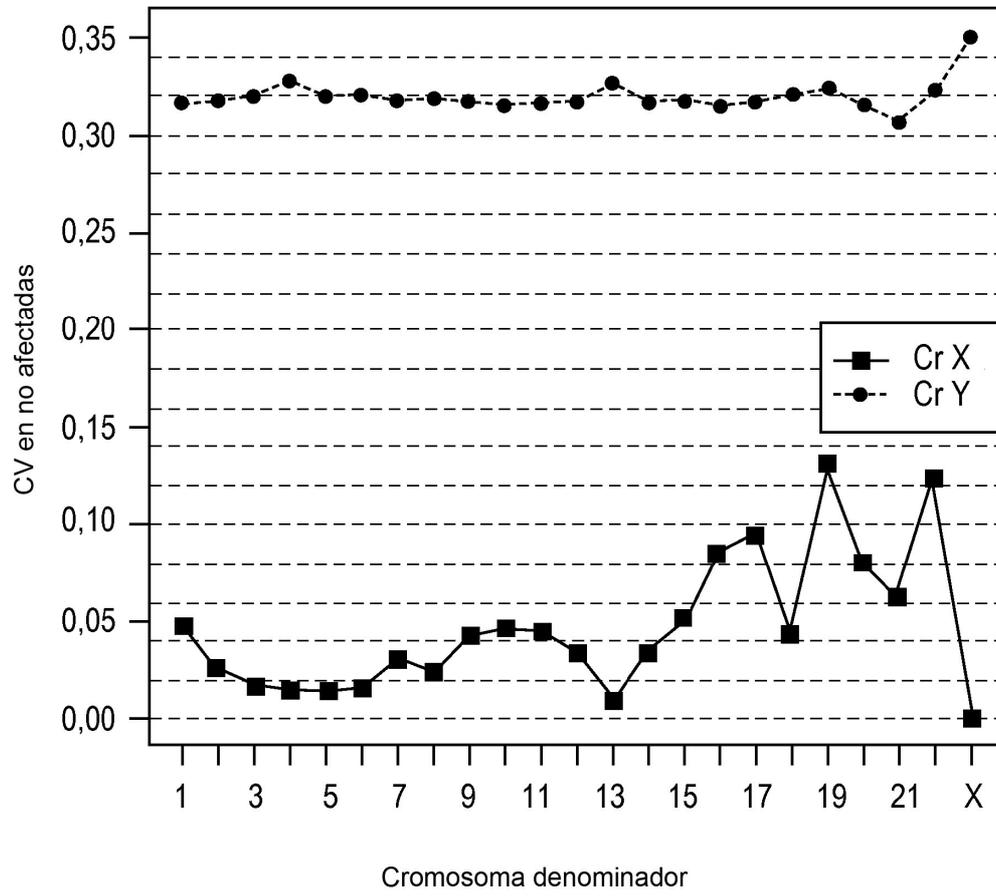




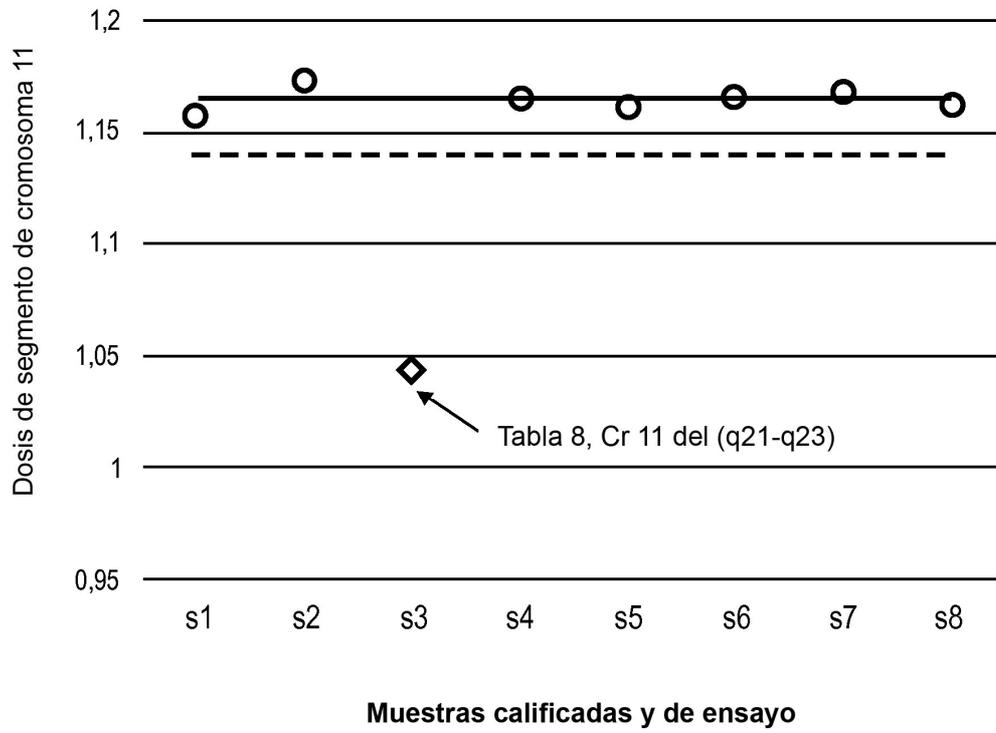
**FIG. 13B**



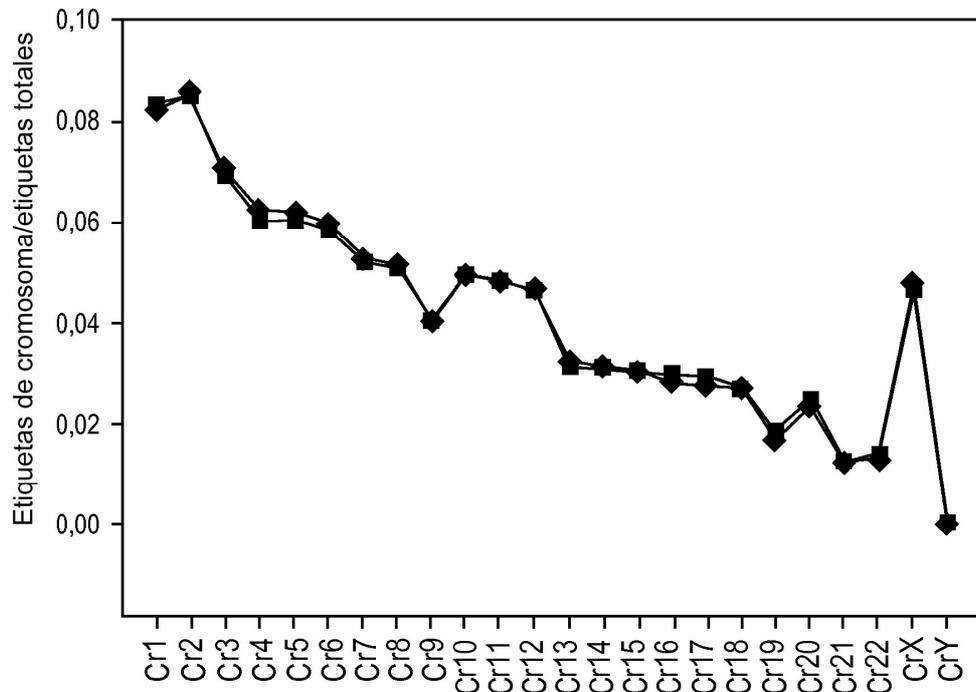
**FIG. 14**



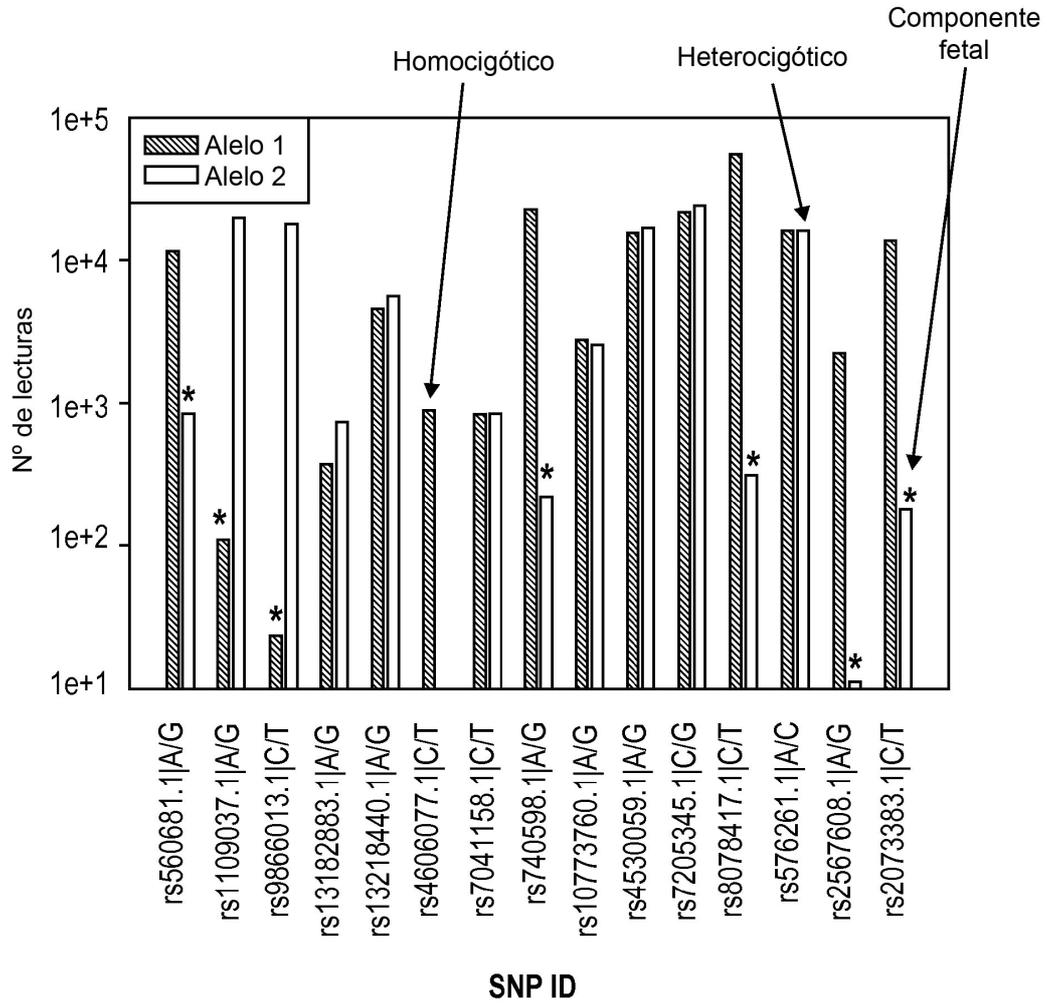
**FIG. 15**



**FIG. 16**

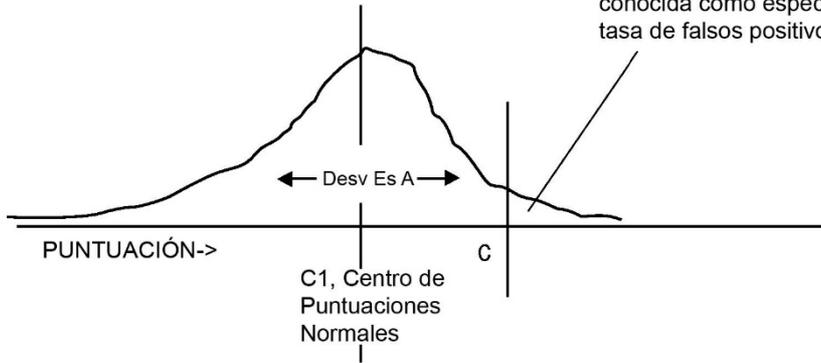


**FIG. 17**



**FIG. 18**

**POBLACIÓN DE PUNTUACIONES "NORMALES",  
ESTIMACIÓN DE DISTRIBUCIÓN f1**

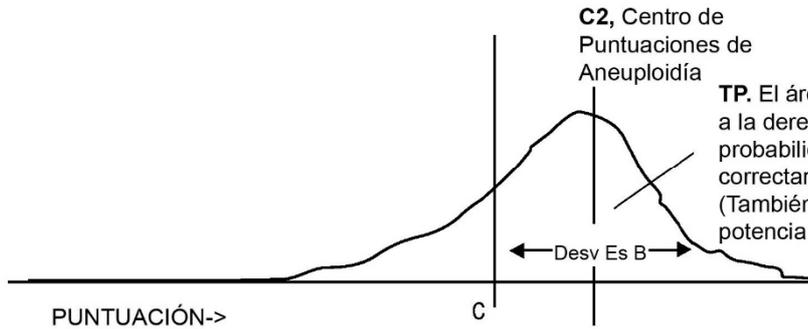


**FP.** El área bajo la curva a la derecha de "c" es la probabilidad de designación falsa. (También conocida como especificidad o tasa de falsos positivos)

**FIG. 19A**

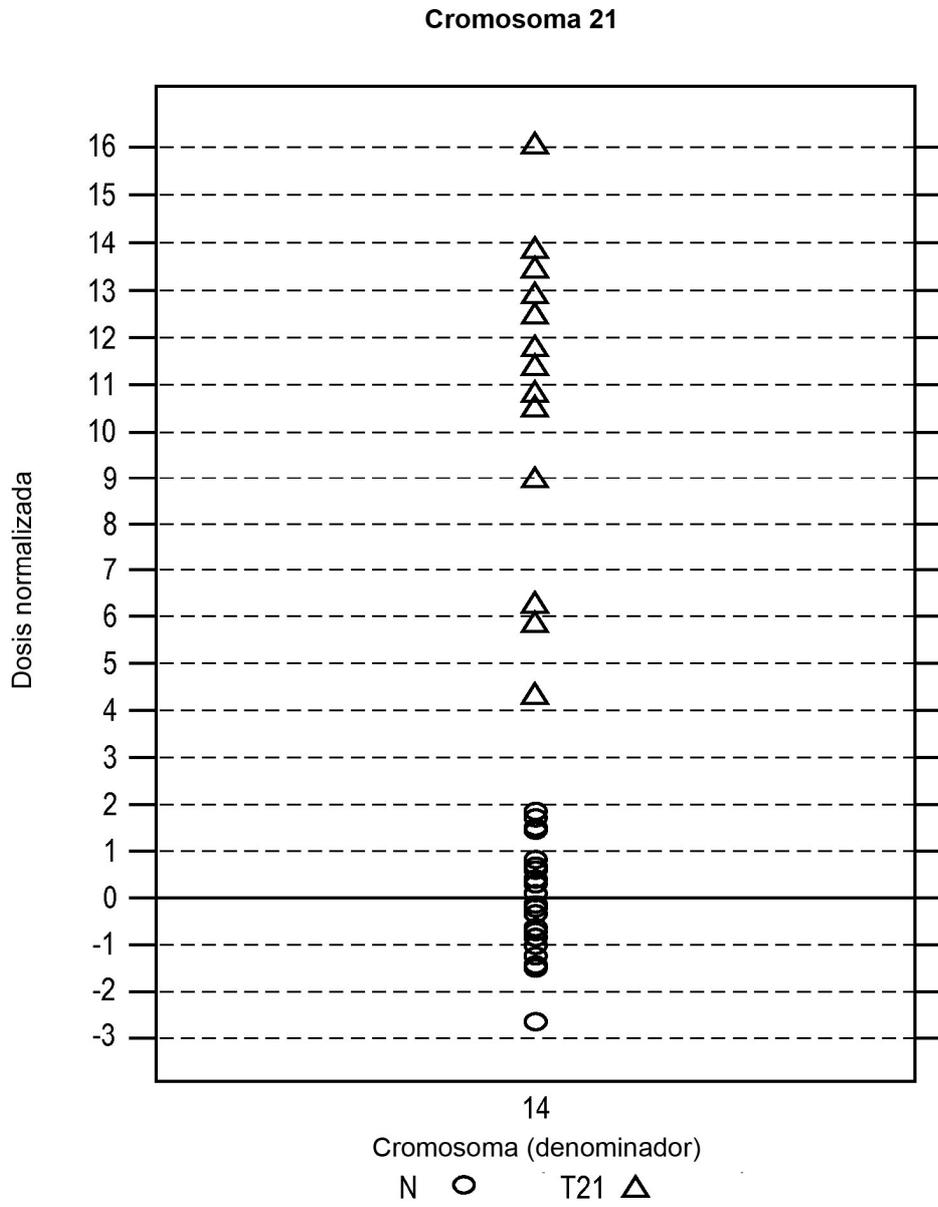
← d →

**POBLACIÓN DE PUNTUACIONES DE "ANEUPLOIDÍA",  
ESTIMACIÓN DE DISTRIBUCIÓN f2**

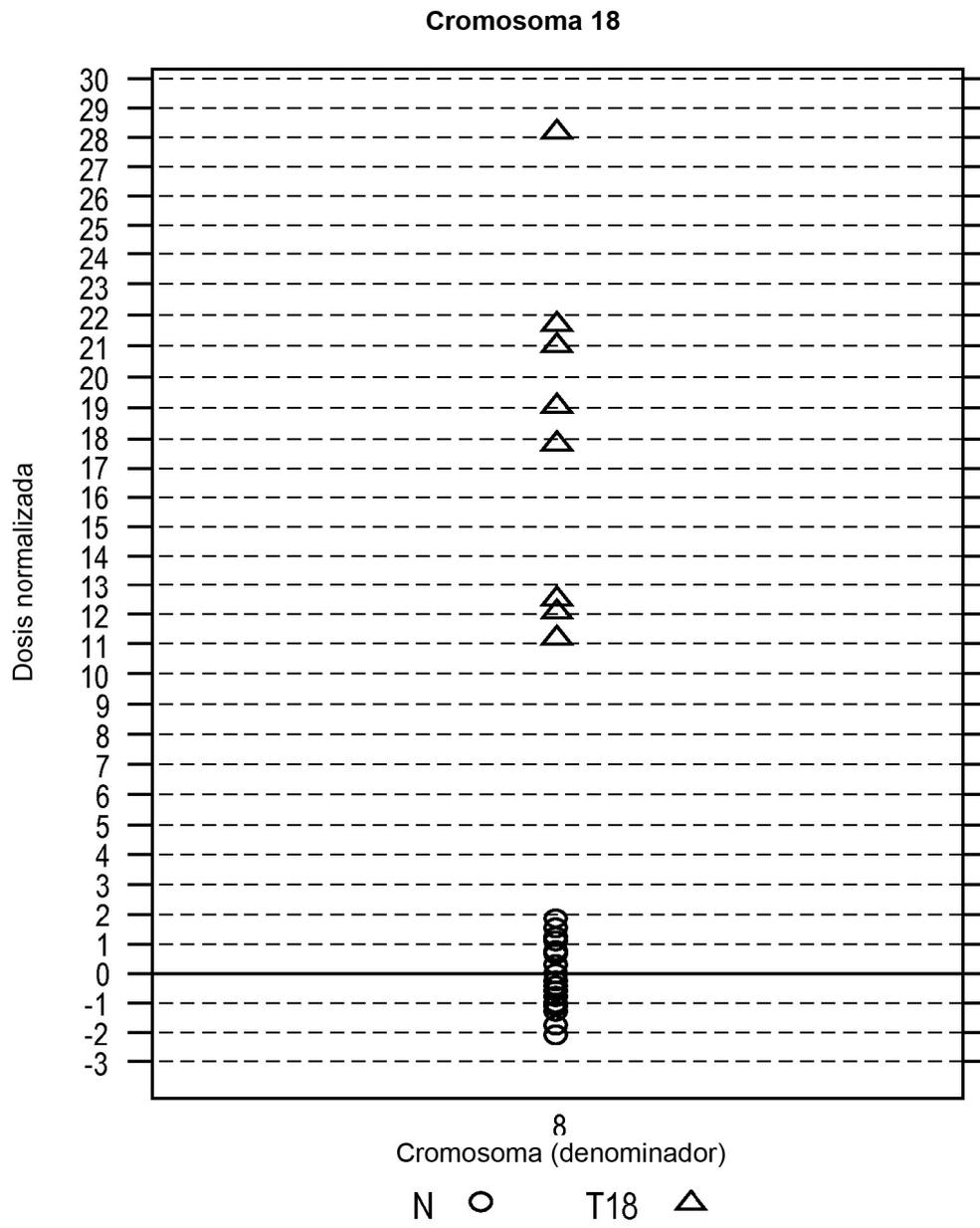


**TP.** El área bajo la curva a la derecha de "c" es la probabilidad de identificar correctamente la aneuploidía. (También conocida como potencia estadística).

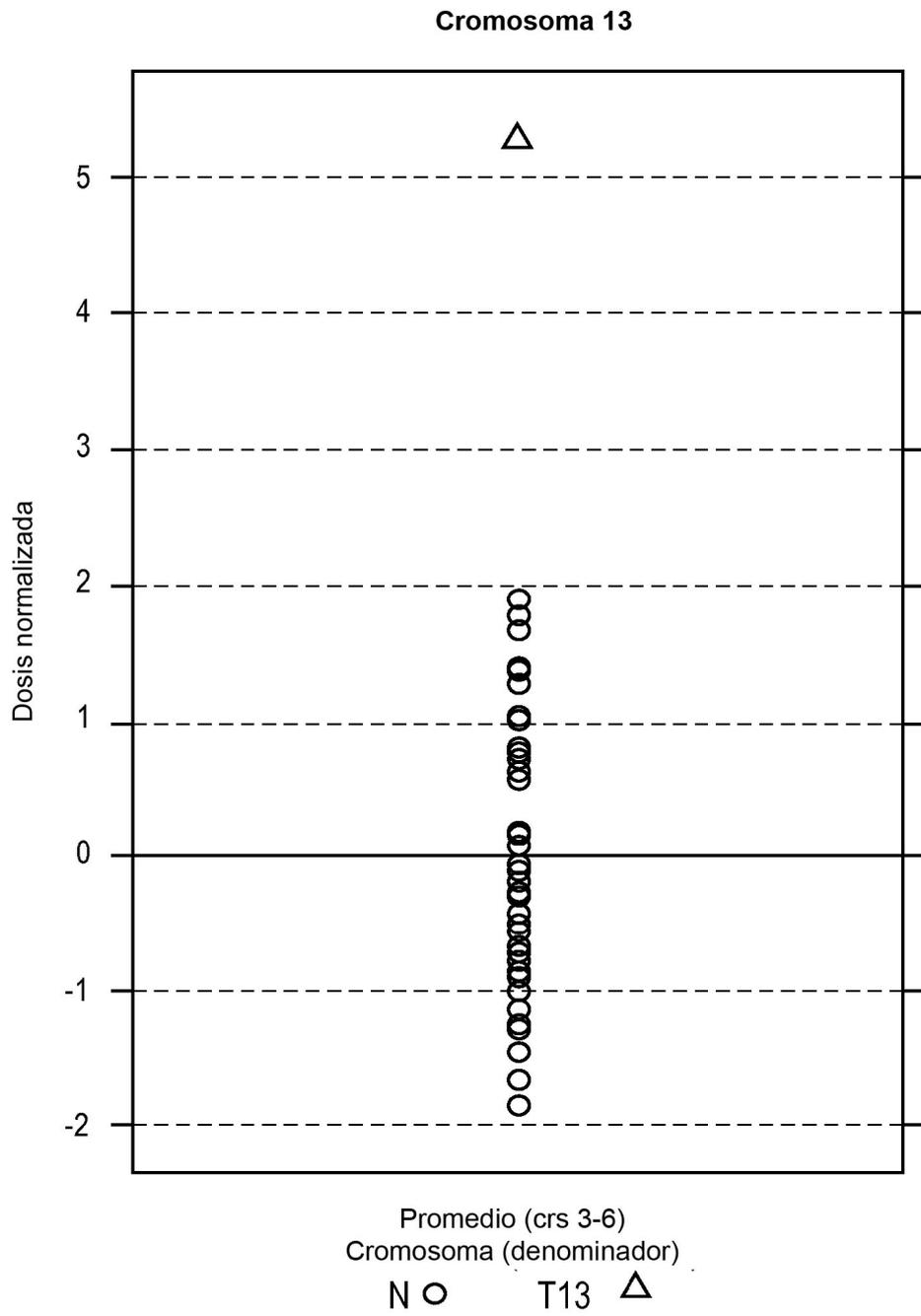
**FIG. 19B**



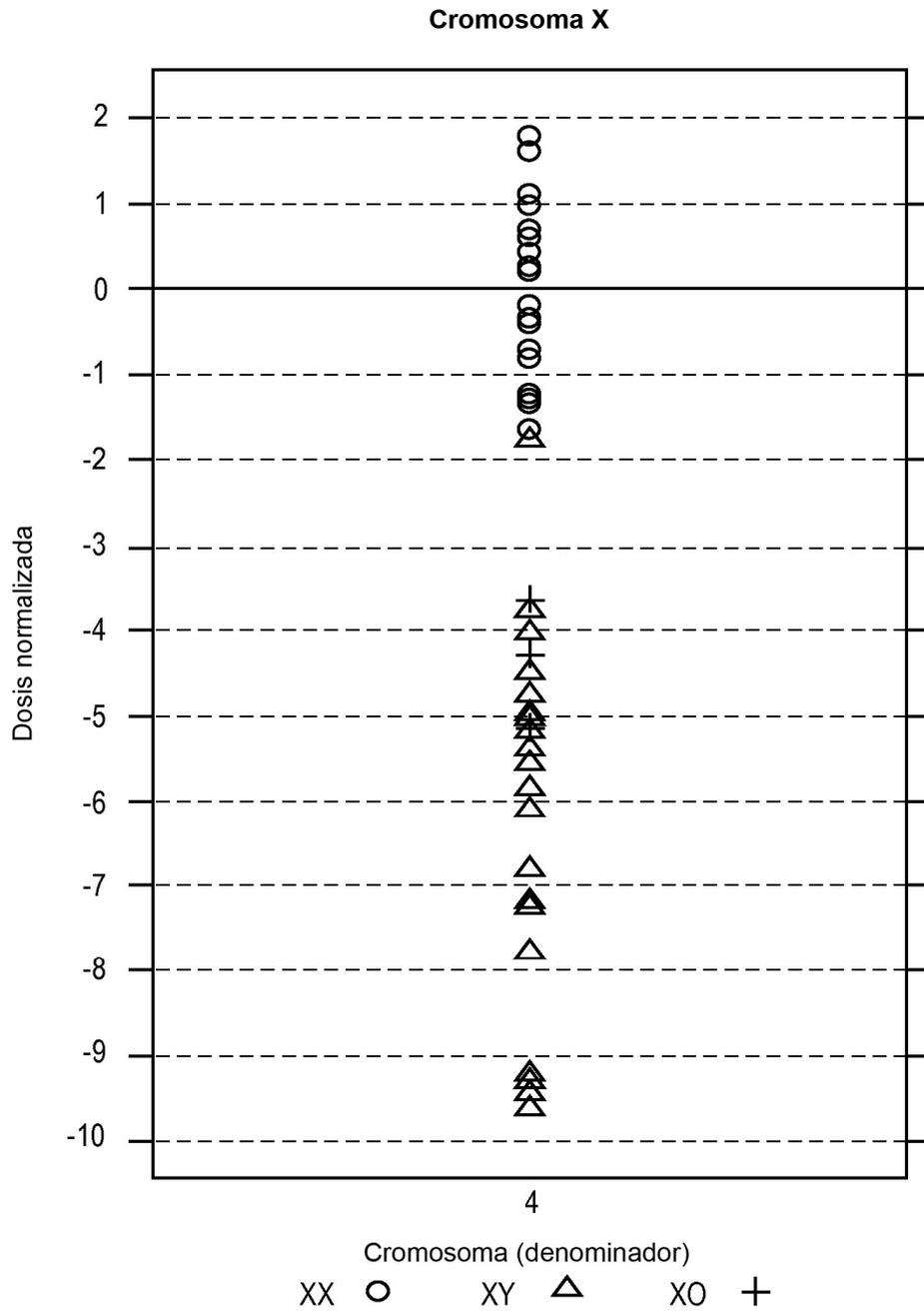
**FIG. 20A**



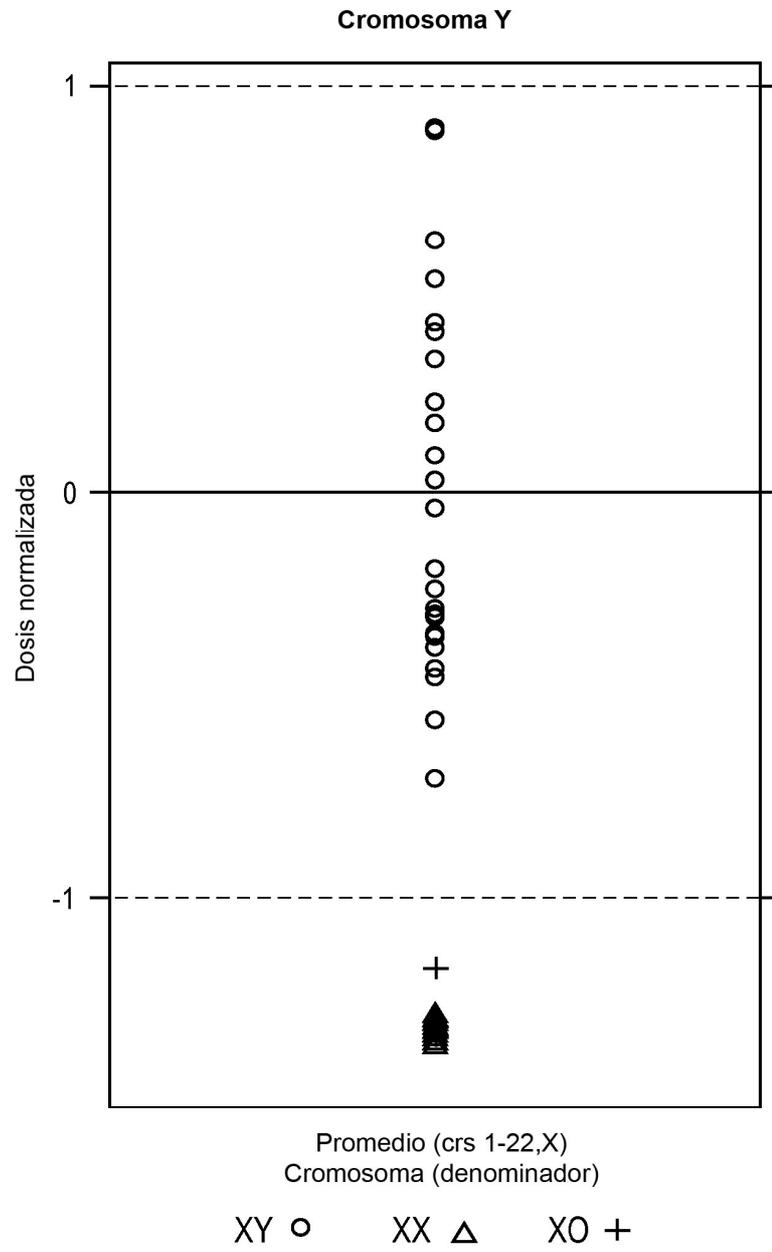
**FIG. 20B**



**FIG. 20C**



**FIG. 20D**



**FIG. 20E**