

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 699 895**

51 Int. Cl.:

C12Q 1/68 (2008.01)

G06F 19/18 (2011.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **08.07.2015 E 15175962 (8)**

97 Fecha y número de publicación de la concesión europea: **19.09.2018 EP 2977466**

54 Título: **Detección de la aneuploidía cromosómica**

30 Prioridad:

22.07.2014 US 201462027258 P
03.03.2015 US 201514636193

45 Fecha de publicación y mención en BOPI de la
traducción de la patente:
13.02.2019

73 Titular/es:

YOURGENE BIOSCIENCE (100.0%)
6F.-3, No.178, Nanyang St. Xizhi Dist.
221 New Taipei City, TW

72 Inventor/es:

CHAN, CHIA-HAN

74 Agente/Representante:

IZQUIERDO BLANCO, María Alicia

ES 2 699 895 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

Detección de la aneuploidía cromosómica

5 ANTECEDENTES

Campo de la invención

[0001] La presente divulgación está relacionada con la bioinformática.

10

Descripción de la técnica relacionada

[0002] La aneuploidía es una afección en la que el número de cromosomas no es un múltiplo exacto del número característico de una especie en particular. Un cromosoma de más o de menos es una causa común en los desórdenes genéticos, incluidos los defectos de nacimiento en humanos. La amniocentesis (también denominada 'prueba de fluido amniótico' o 'AFT', por sus siglas en inglés) es un procedimiento médico que se utiliza en el diagnóstico prenatal de anomalías cromosómicas e infecciones fetales. Las anomalías detectadas más a menudo son el síndrome de Down (trisomía 21), el síndrome de Edwards (trisomía 18), el síndrome de Patau (trisomía 13) y el síndrome de Turner (monosomía X). Sin embargo, la amniocentesis entraña diversos riesgos, incluyendo el aborto espontáneo, las heridas debidas a las agujas, la pérdida de fluido amniótico, la sensibilización de Rh, las infecciones y la transmisión de infecciones.

RESUMEN

[0003] El alcance de la invención queda definido o delimitado por las reivindicaciones anexas. La presente invención está relacionada con un método para detectar una aneuploidía relacionada con una zona o región de ácido nucleico diana, e incluye los siguientes pasos o etapas. Se obtiene una base de datos de referencia. La base de datos de referencia se establece secuenciando diversas muestras biológicas de referencia mediante una plataforma de secuenciación. Se determina al menos un factor de normalización basándose en la base de datos de referencia. Se determina un valor límite basándose en la base de datos de referencia. Se secuencia una muestra biológica bajo prueba mediante la plataforma de secuenciación a fin de obtener un número de lecturas de objetivo de la muestra biológica bajo prueba. La muestra biológica bajo prueba se obtiene de una hembra embarazada y tiene moléculas de ácido nucleico de la hembra embarazada y su feto. Las lecturas de objetivo de la muestra biológica bajo prueba proceden de la región de ácido nucleico diana. El número de lecturas de objetivo de la muestra biológica bajo prueba se normaliza mediante el factor de normalización. El número normalizado de las lecturas de objetivo de la muestra biológica bajo prueba se compara con el valor límite. Basándose en esta comparación se determina si la aneuploidía cromosómica relacionada con la región de ácido nucleico diana está presente en el feto.

[0004] En un ejemplo, la base de datos de referencia puede estar basada en el género.

[0005] En un ejemplo, la determinación del factor de normalización también incluye los siguientes pasos. Se determina un número de lecturas de objetivo de cada muestra biológica de referencia, de manera que las lecturas de objetivo de cada muestra biológica de referencia proceden de la región de ácido nucleico diana. Se determina un número de lecturas correlacionadas de cada muestra biológica de referencia, de manera que los números de las lecturas correlacionadas de las muestras biológicas de referencia se correlacionan con los números de las lecturas de objetivo de las muestras biológicas de referencia, y las lecturas correlacionadas de cada muestra biológica de referencia proceden de una región de ácido nucleico correlacionada. El factor de normalización se determina basándose en los números de las lecturas de objetivo de las muestras biológicas de referencia y los números de las lecturas correlacionadas de las muestras biológicas de referencia.

[0006] En un ejemplo, los números de las lecturas correlacionadas de las muestras biológicas de referencia pueden correlacionarse de manera lineal con los números de las lecturas de objetivo de las muestras biológicas de referencia.

[0007] En un ejemplo, los números de las lecturas correlacionadas de las muestras biológicas de referencia y los números de las lecturas de objetivo de las muestras biológicas de referencia pueden tener un coeficiente de correlación en un rango o intervalo de entre alrededor de un 0,7 y alrededor de un 0,99.

[0008] En un ejemplo, la determinación del factor de normalización, basada en los números de las lecturas de objetivo de las muestras biológicas de referencia y los números de las lecturas correlacionadas de las muestras biológicas de referencia, también incluye los siguientes pasos. Se determinan las frecuencias locales diana de referencia, de manera que cada frecuencia local diana de referencia es un ratio o proporción del número de lecturas de objetivo de cada muestra biológica de referencia con respecto a un número de lecturas locales de la mencionada muestra biológica de referencia, y las lecturas locales de cada muestra biológica de referencia proceden del propio cromosoma de la región de ácido nucleico diana. Se determinan las frecuencias globales correlacionadas, de manera que cada frecuencia global correlacionada de referencia es un ratio o proporción del número de lecturas

correlacionadas de cada muestra biológica de referencia con respecto al número de lecturas totales de cada muestra biológica de referencia. Se determina una primera correspondencia entre las frecuencias locales diana de referencia y las frecuencias globales correlacionadas de referencia.

[0009] En un ejemplo, la normalización del número de lecturas de objetivo de la muestra biológica bajo prueba también incluye los siguientes pasos. Se determina una frecuencia global diana de prueba, de manera que la frecuencia global diana de prueba es un ratio o proporción de las lecturas de objetivo de la muestra biológica bajo prueba con respecto al número de lecturas totales de la muestra biológica bajo prueba. Se determina una frecuencia local diana de prueba, de manera que la frecuencia local diana de prueba es un ratio o proporción de las lecturas de objetivo de la muestra biológica bajo prueba con respecto al número de lecturas locales de la muestra biológica bajo prueba, y las lecturas locales de la muestra biológica bajo prueba proceden del propio cromosoma de la región de ácido nucleico diana. Se determina una frecuencia global correlacionada de prueba, de manera que la frecuencia global correlacionada de prueba es un ratio o proporción de un número de lecturas correlacionadas de la muestra biológica bajo prueba con respecto al número de lecturas totales de la muestra biológica bajo prueba, y las lecturas correlacionadas de la muestra biológica bajo prueba proceden de la región de ácido nucleico correlacionada. La frecuencia local diana de prueba se normaliza por medio de la frecuencia global correlacionada de prueba y la primera correspondencia. La frecuencia global diana de prueba se normaliza por medio de la frecuencia local diana de prueba normalizada.

[0010] En un ejemplo, la determinación del factor de normalización, basada en los números de las lecturas de objetivo de las muestras biológicas de referencia y los números de las lecturas correlacionadas de las muestras biológicas de referencia, también incluye los siguientes pasos. Se determinan las frecuencias globales diana de referencia, de manera que cada frecuencia global diana de referencia es un ratio o proporción del número de lecturas de objetivo de cada muestra biológica de referencia con respecto al número de lecturas totales de la mencionada muestra biológica de referencia. Las frecuencias locales diana de referencia se normalizan, respectivamente, mediante las frecuencias globales correlacionadas de referencia y la primera correspondencia. Las frecuencias globales diana de referencia se normalizan, respectivamente, mediante las frecuencias locales diana de referencia normalizadas. Se determina una segunda correspondencia entre las frecuencias globales diana de referencia normalizadas y las frecuencias globales correlacionadas de referencia.

[0011] En un ejemplo, la determinación del valor límite también incluye los siguientes pasos. Los valores estimados de referencia se estiman, respectivamente, basándose en las frecuencias globales correlacionadas de referencia y en la segunda correspondencia. Se determinan, respectivamente, los valores de diferencia de referencia entre las frecuencias globales diana de referencia normalizadas y los valores estimados de referencia. Los valores de diferencia de referencia respecto a los resultados o puntuaciones estándares de referencia se estandarizan, respectivamente, basándose en la base de datos de referencia. El valor límite se determina basándose en las puntuaciones estándares de referencia.

[0012] En un ejemplo, la normalización del número de lecturas de objetivo de la muestra biológica bajo prueba también incluye los siguientes pasos. Se estima o calcula un valor estimado de prueba basándose en la frecuencia global correlacionada de prueba y la segunda correspondencia. Se determina un valor de diferencia de prueba entre la frecuencia global diana de prueba normalizada y el valor estimado de prueba. El valor de diferencia de prueba respecto a una puntuación estándar de prueba se estandariza basándose en la base de datos de referencia, de manera que la comparación incluye comparar la puntuación estándar de prueba y el valor límite.

[0013] De acuerdo con algunas realizaciones de la presente invención, un medio no transitorio legible por una máquina almacena un programa que, al ejecutarse mediante al menos una unidad de procesamiento, detecta una aneuploidía cromosómica en relación con una región de ácido nucleico diana. El programa incluye instrucciones para los siguientes pasos. Se obtiene una base de datos de referencia. La base de datos de referencia se establece secuenciando diversas muestras biológicas de referencia mediante una plataforma de secuenciación. Se determina al menos un factor de normalización basándose en la base de datos de referencia. Se determina un valor límite basándose en la base de datos de referencia. Se secuencian una muestra biológica bajo prueba mediante la plataforma de secuenciación a fin de obtener un número de lecturas de objetivo de la muestra biológica bajo prueba. La muestra biológica bajo prueba se obtiene de una hembra embarazada y tiene moléculas de ácido nucleico de la hembra embarazada y su feto. Las lecturas de objetivo de la muestra biológica bajo prueba proceden de la región de ácido nucleico diana. El número de lecturas de objetivo de la muestra biológica bajo prueba se normaliza mediante el factor de normalización. El número normalizado de las lecturas de objetivo de la muestra biológica bajo prueba se compara con el valor límite. Basándose en esta comparación se determina si la aneuploidía cromosómica relacionada con la región de ácido nucleico diana está presente en el feto.

[0014] En un ejemplo, la base de datos de referencia puede estar basada en el género.

[0015] En un ejemplo, el conjunto de instrucciones para determinar el factor de normalización incluye un conjunto de instrucciones para los siguientes pasos. Se determina un número de lecturas de objetivo de cada muestra biológica de referencia, de manera que las lecturas de objetivo de cada muestra biológica de referencia proceden de la región de ácido nucleico diana. Se determina un número de lecturas correlacionadas de cada muestra biológica de

referencia, de manera que los números de las lecturas correlacionadas de las muestras biológicas de referencia se correlacionan con los números de las lecturas de objetivo de las muestras biológicas de referencia, y las lecturas correlacionadas de cada muestra biológica de referencia proceden de una región de ácido nucleico correlacionada. El factor de normalización se determina basándose en los números de las lecturas de objetivo de las muestras biológicas de referencia y los números de las lecturas correlacionadas de las muestras biológicas de referencia.

[0016] En un ejemplo, los números de las lecturas correlacionadas de las muestras biológicas de referencia pueden correlacionarse de manera lineal con los números de las lecturas de objetivo de las muestras biológicas de referencia.

[0017] En un ejemplo, los números de las lecturas correlacionadas de las muestras biológicas de referencia y los números de las lecturas de objetivo de las muestras biológicas de referencia pueden tener un coeficiente de correlación en un rango o intervalo de entre alrededor de un 0,7 y alrededor de un 0,99.

[0018] En un ejemplo, el conjunto de instrucciones para determinar el factor de normalización basándose en los números de las lecturas de objetivo de las muestras biológicas de referencia y los números de las lecturas correlacionadas de las muestras biológicas de referencia incluye un conjunto de instrucciones para los siguientes pasos. Se determinan las frecuencias locales diana de referencia, de manera que cada frecuencia local diana de referencia es un ratio o proporción del número de lecturas de objetivo de cada muestra biológica de referencia con respecto a un número de lecturas locales de la mencionada muestra biológica de referencia, y las lecturas locales de cada muestra biológica de referencia proceden del propio cromosoma de la región de ácido nucleico diana. Se determinan las frecuencias globales correlacionadas de referencia, de manera que cada frecuencia global correlacionada de referencia es un ratio o proporción del número de lecturas correlacionadas de cada muestra biológica de referencia con respecto al número de lecturas totales de cada muestra biológica de referencia. Se determina una primera correspondencia entre las frecuencias locales diana de referencia y las frecuencias globales correlacionadas de referencia.

[0019] En un ejemplo, el conjunto de instrucciones para normalizar el número de lecturas de objetivo de la muestra biológica bajo prueba también incluye un conjunto de instrucciones para los siguientes pasos. Se determina una frecuencia global diana de prueba, de manera que la frecuencia global diana de prueba es un ratio o proporción del número de lecturas de objetivo de la muestra biológica bajo prueba con respecto al número de lecturas totales de la muestra biológica bajo prueba. Se determina una frecuencia local diana de prueba, de manera que la frecuencia local diana de prueba es un ratio o proporción del número de lecturas de objetivo de la muestra biológica bajo prueba con respecto al número de lecturas locales de la muestra biológica bajo prueba, y las lecturas locales de la muestra biológica bajo prueba proceden del propio cromosoma de la región de ácido nucleico diana. Se determina una frecuencia global correlacionada de prueba, de manera que la frecuencia global correlacionada de prueba es un ratio o proporción de un número de lecturas correlacionadas de la muestra biológica bajo prueba con respecto al número de lecturas totales de la muestra biológica bajo prueba, y las lecturas correlacionadas de la muestra biológica bajo prueba proceden de la región de ácido nucleico correlacionada. La frecuencia local diana de prueba se normaliza por medio de la frecuencia global correlacionada de prueba y la primera correspondencia. La frecuencia global diana de prueba se normaliza por medio de la frecuencia local diana de prueba normalizada.

[0020] En un ejemplo, el conjunto de instrucciones para determinar el factor de normalización basándose en los números de las lecturas de objetivo de las muestras biológicas de referencia y los números de las lecturas correlacionadas de las muestras biológicas de referencia incluye un conjunto de instrucciones para los siguientes pasos. Se determinan las frecuencias globales diana de referencia, de manera que cada frecuencia global diana de referencia es un ratio o proporción del número de lecturas de objetivo de cada muestra biológica de referencia con respecto al número de lecturas totales de la mencionada muestra biológica de referencia. Las frecuencias locales diana de referencia se normalizan, respectivamente, mediante las frecuencias globales correlacionadas de referencia y la primera correspondencia. Las frecuencias globales diana de referencia se normalizan, respectivamente, mediante las frecuencias locales diana de referencia normalizadas. Se determina una segunda correspondencia entre las frecuencias globales diana de referencia normalizadas y las frecuencias globales correlacionadas de referencia.

[0021] En un ejemplo, el conjunto de instrucciones para determinar el valor límite incluye un conjunto de instrucciones para los siguientes pasos. Los valores estimados de referencia se estiman, respectivamente, basándose en las frecuencias globales correlacionadas de referencia y en la segunda correspondencia. Se determinan, respectivamente, los valores de diferencia de referencia entre las frecuencias globales diana de referencia normalizadas y los valores estimados de referencia. Los valores de diferencia de referencia respecto a los resultados o puntuaciones estándares de referencia se estandarizan, respectivamente, basándose en la base de datos de referencia. El valor límite se determina basándose en las puntuaciones estándares de referencia.

[0022] En un ejemplo, el conjunto de instrucciones para normalizar el número de lecturas de objetivo de la muestra biológica bajo prueba incluye un conjunto de instrucciones para los siguientes pasos. Se estima o calcula un valor estimado de prueba basándose en la frecuencia global correlacionada de prueba y la segunda correspondencia. Se determina un valor de diferencia de prueba entre la frecuencia global diana de prueba normalizada y el valor

estimado de prueba. El valor de diferencia de prueba respecto a una puntuación estándar de prueba se estandariza basándose en la base de datos de referencia, de manera que el conjunto de instrucciones para la comparación incluye un conjunto de instrucciones para comparar la puntuación estándar de prueba y el valor límite.

5 BREVE DESCRIPCIÓN DE LAS ILUSTRACIONES

[0023]

La Figura 1 (Fig. 1) es un diagrama de flujo de un método para detectar una aneuploidía cromosómica de acuerdo con algunas realizaciones de la presente invención.

Las Figuras 2A y 2B muestran gráficos de lecturas cromosómicas de diversas muestras biológicas de referencia secuenciadas mediante la misma plataforma de secuenciación de acuerdo con diversos ejemplos prácticos de la presente invención.

La Figura 3 es un diagrama de flujo del paso 120 de la Figura 1.

La Figura 4 es un diagrama de flujo del paso 123 de la Figura 3.

La Figura 5 muestra la distribución de las frecuencias globales del cromosoma 21 de las muestras biológicas de referencia de acuerdo con un ejemplo práctico de la presente invención.

La Figura 6 muestra una correspondencia entre las frecuencias globales del cromosoma 22 de las muestras biológicas de referencia y las frecuencias locales de 43-45 Mb del cromosoma 21 de las muestras biológicas de referencia.

La Figura 7 muestra una correspondencia entre las frecuencias globales del cromosoma 22 de las muestras biológicas de referencia y las frecuencias globales de 43-45 Mb del cromosoma 21 de las muestras biológicas de referencia.

La Figura 8 muestra una correspondencia entre las frecuencias globales del cromosoma 22 de las muestras biológicas de referencia y las frecuencias globales normalizadas de 43-45 Mb del cromosoma 21 de las muestras biológicas de referencia.

La Figura 9 es un diagrama de flujo del paso 130 de la Figura 1.

La Figura 10 muestra la distribución de las puntuaciones estándares de referencia de acuerdo con el ejemplo práctico que se muestra en las Figuras 5-8.

La Figura 11 es un diagrama de flujo del paso 150 de la Figura 1.

La Figura 12 muestra la distribución de las puntuaciones estándares de prueba y de referencia de acuerdo con el ejemplo práctico que se muestra en las Figuras 5-8 y 10.

45 DESCRIPCIÓN DETALLADA

[0024] En la descripción detallada que se ofrece a continuación, con fines explicativos se ofrecen numerosos detalles específicos a fin de proporcionar una comprensión rigurosa de las realizaciones desveladas. Sin embargo, resulta evidente que pueden ponerse en práctica una o más realizaciones sin estos detalles específicos. En otros casos, se representan esquemáticamente algunas estructuras y dispositivos muy conocidos a fin de simplificar las ilustraciones.

[0025] Tal y como se utiliza en el presente documento, el término 'muestra biológica' hace referencia a cualquier muestra biológica que se toma o extrae de un sujeto (por ejemplo, un humano, como una mujer embarazada) y que contiene una o más moléculas de interés de ácido nucleico.

[0026] El término 'ácido nucleico' hace referencia al ácido desoxirribonucleico (ADN), al ácido ribonucleico (ARN) o a un polímero de estos, tanto en su forma monocatenaria como en su forma bicatenaria. A menos que se especifique lo contrario, el término abarca los ácidos nucleicos que contienen análogos conocidos de nucleótidos naturales que tienen propiedades de unión similares al ácido nucleico de referencia y se metabolizan de una forma similar a los nucleótidos que existen de forma natural. A menos que se especifique lo contrario, una secuencia de ácido nucleico particular también abarca o incluye implícitamente las variantes modificadas de estos compuestos (por ejemplo, sustituciones de codones degenerados), los alelos, los ortólogos, los polimorfismos de nucleótido simple o único (SNPs, por sus siglas en inglés) y las secuencias complementarias, así como la secuencia indicada explícitamente. El término 'ácido nucleico' se utiliza indistintamente junto con 'gen', 'ADNc', 'ARNm', 'ARN pequeño no codificante', 'micro ARN' ('mi-ARN' o 'miRNA', por sus siglas en inglés), 'ARN asociado a Piwi' y 'ARN horquillado corto' ('shRNA',

por sus siglas en inglés) codificado por un gen o un locus.

[0027] Tal y como se utiliza en el presente documento, el término 'basándose en' o 'basado/a en' quiere decir 'basándose al menos en parte en' o 'basado/a al menos en parte en' y hace referencia a un valor o resultado que se utiliza para determinar otro valor o resultado, tal y como sucede en la relación entre la introducción de un método y los resultados de dicho método.

[0028] Tal y como se utiliza en el presente documento, el término 'aneuploidía cromosómica' hace referencia a una variación en el número cuantitativo de un cromosoma respecto al de un genoma diploide. La variación puede ser una ganancia o una pérdida. Puede afectar a un cromosoma al completo o a una región de un cromosoma.

[0029] Tal y como se utiliza en el presente documento, el término 'valor límite' ('o valor tope') hace referencia a un valor numérico cuyo valor se utiliza para decidirse o mediar entre dos o más estados (por ejemplo, anormal y normal) de clasificación para una muestra biológica. Por ejemplo, si un parámetro es mayor que el valor límite, se realiza una primera clasificación de los datos cuantitativos (por ejemplo, anormal); o, si el parámetro es menor que el valor límite, se realiza una clasificación diferente de los datos cuantitativos (por ejemplo, normal).

[0030] La Figura 1 es un diagrama de flujo de un método para detectar una aneuploidía cromosómica de acuerdo con algunas realizaciones de la presente invención. Tal y como se muestra en la Figura 1, el método para detectar una aneuploidía cromosómica relacionada con una región de ácido nucleico diana incluye los siguientes pasos o etapas. En el paso 110, se obtiene una base de datos de referencia. La base de datos de referencia se establece secuenciando diversas muestras biológicas de referencia mediante una plataforma de secuenciación. Las muestras biológicas de referencia pueden ser plasma, orina, suero o cualesquiera otras muestras adecuadas. Las muestras biológicas de referencia se obtienen, respectivamente, de hembras embarazadas. Por consiguiente, cada muestra biológica de referencia contiene moléculas de ácido nucleico de una hembra embarazada y de su feto. Las moléculas de ácido nucleico pueden ser, por ejemplo, fragmentos de cromosomas. La plataforma de secuenciación puede ser, por ejemplo, una plataforma de secuenciación de última generación, como 454 Platform (Roche), HiSeq, NextSeq y MiSeq System (Illumina Platform), Ion Torrent PGM y Proton System (Life Technologies) o Minion (Oxford Nanopore).

[0031] Las Figuras 2A y 2B muestran gráficos de lecturas cromosómicas de diversas muestras biológicas de referencia secuenciadas mediante la misma plataforma de secuenciación de acuerdo con diversos ejemplos prácticos de la presente invención. En las Figuras 2A y 2B, el eje X es el número cromosómico o número de cromosomas, y el eje Y es la variación respecto a la media en varias lecturas de cada cromosoma. Tal y como se muestra en las Figuras 2A y 2B, hay una fluctuación en el número de lecturas de cada cromosoma. Esta fluctuación puede estar provocada por las propiedades de la muestra, las propiedades del sistema o los factores ambientales. Por consiguiente, debido a la fluctuación no se puede determinar si la aneuploidía cromosómica está presente en el feto basándose únicamente en el número de lecturas de cada cromosoma.

[0032] En referencia a la Figura 1, en el paso 120 se determina al menos un factor de normalización basándose en la base de datos de referencia. Tal y como se muestra en las Figuras 2A y 2B, los números de las lecturas de los cromosomas se correlacionan entre sí. Es decir, un aumento en el número de lecturas de un cromosoma disminuiría el número de lecturas de otro cromosoma. Puesto que las muestras biológicas de referencia de las Figuras 2A y 2B se obtuvieron básicamente del mismo modo y se secuenciaron mediante la misma plataforma de secuenciación, las fluctuaciones para las diferentes muestras biológicas de referencia son básicamente iguales. Por lo tanto, el cromosoma correlacionado puede usarse para normalizar el cromosoma de interés y eliminar la fluctuación.

[0033] Puesto que un feto femenino no tiene cromosoma Y, una muestra biológica de un feto femenino tendría un perfil diferente de lecturas cromosómicas cuando se comparara con el de una muestra de un feto masculino. Por lo tanto, la base de datos de referencia del paso 110 puede estar basada en el género. Es decir, si una muestra biológica bajo prueba es de un feto masculino, se utiliza una base de datos de referencia de fetos masculinos. Si la muestra biológica bajo prueba es de un feto femenino, se utiliza una base de datos de referencia de fetos femeninos.

[0034] La Figura 3 es un diagrama de flujo del paso 120 de la Figura 1. El paso 120 incluye los siguientes pasos o etapas. En el paso 121 se determina un número de lecturas de objetivo de cada muestra biológica de referencia. Las lecturas de objetivo de cada muestra biológica de referencia proceden de la región de ácido nucleico diana. En el paso 122, se determina un número de lecturas correlacionadas de cada muestra biológica de referencia. Los números de las lecturas correlacionadas de las muestras biológicas de referencia se correlacionan con los números de las lecturas de objetivo de las muestras biológicas de referencia. Por ejemplo, los números de las lecturas correlacionadas de las muestras biológicas de referencia pueden correlacionarse de manera lineal con los números de las lecturas de objetivo de las muestras biológicas de referencia. Los números de las lecturas correlacionadas de las muestras biológicas de referencia y los números de las lecturas de objetivo de las muestras biológicas de referencia pueden tener un coeficiente de correlación en un rango o intervalo de entre alrededor de un 0,7 y alrededor de un 0,99. Las lecturas correlacionadas de cada muestra biológica de referencia proceden de una región de ácido nucleico correlacionada. En el paso 123, el factor de normalización se determina basándose en los números de las lecturas de objetivo de las muestras biológicas de referencia y los números de las lecturas

correlacionadas de las muestras biológicas de referencia.

[0035] La Figura 4 es un diagrama de flujo del paso 123 de la Figura 3. Más específicamente, el paso 123 incluye los siguientes pasos o etapas. En el paso 124, se determinan las frecuencias locales diana de referencia. Cada frecuencia local diana de referencia es un ratio o proporción del número de lecturas de objetivo de cada muestra biológica de referencia con respecto a un número de lecturas locales de la mencionada muestra biológica de referencia. Las lecturas locales de cada muestra biológica de referencia proceden del propio cromosoma de la región de ácido nucleico diana. En el paso 125, se determinan las frecuencias globales correlacionadas de referencia. Cada frecuencia global correlacionada de referencia es un ratio o proporción del número de lecturas correlacionadas de cada muestra biológica de referencia con respecto al número de lecturas totales de cada muestra biológica de referencia. En el paso 126, se determina una primera correspondencia entre las frecuencias locales diana de referencia y las frecuencias globales correlacionadas de referencia. En el paso 127, se determinan las frecuencias globales diana de referencia. Cada frecuencia global diana de referencia es un ratio o proporción del número de lecturas de objetivo de cada muestra biológica de referencia con respecto al número de lecturas totales de la mencionada muestra biológica de referencia. En el paso 128a, las frecuencias locales diana de referencia se normalizan, respectivamente, mediante las frecuencias globales correlacionadas de referencia y la primera correspondencia. En el paso 128b, las frecuencias globales diana de referencia se normalizan, respectivamente, mediante las frecuencias locales diana de referencia normalizadas. En el paso 129, se determina una segunda correspondencia entre las frecuencias globales diana de referencia normalizadas y las frecuencias globales correlacionadas de referencia.

[0036] La descripción que se ofrece a continuación tomará como ejemplo la detección de la trisomía 21 en varones para ilustrar cómo llevar a cabo el paso 120. La Figura 5 muestra la distribución de las frecuencias globales del cromosoma 21 para las muestras biológicas de referencia de acuerdo con un ejemplo práctico de la presente invención. En la Figura 5, el eje Y es la frecuencia global del cromosoma 21. La frecuencia global del cromosoma 21, $r(21)/r(all)$, es un ratio o proporción de un número de lecturas de cada muestra biológica de referencia que procede del cromosoma 21, $r(21)$, con respecto al número de lecturas totales de cada -mencionada- muestra biológica de referencia, $r(all)$. Tal y como se muestra en la Figura 5, algunas muestras anómalas o anormales están mezcladas con las muestras normales y, por lo tanto, no se puede determinar un valor límite. En este ejemplo práctico de la presente invención, las muestras biológicas de referencia son 1004, y las muestras biológicas de referencia son plasma obtenido de hembras o mujeres embarazadas. La plataforma de secuenciación es Illumina HiSeq.

[0037] En este ejemplo práctico de la presente invención, se seleccionó la región de 43-45 Mb del cromosoma 21 para que fuera la región de ácido nucleico diana. Tras la inspección, se descubrió que los números de las lecturas de las muestras biológicas de referencia procedentes del cromosoma 22 y los números de las lecturas de las muestras biológicas de referencia procedentes de la región de 43-45 Mb del cromosoma 21 tienen un coeficiente de correlación de 0,94. Por lo tanto, se seleccionó el cromosoma 22 para que fuera la región de ácido nucleico correlacionada. De acuerdo con el paso 124, se determinaron las frecuencias locales de 43-45 Mb del cromosoma 21 de las muestras biológicas de referencia. La frecuencia local de 43-45 Mb del cromosoma 21 de cada muestra biológica de referencia es un ratio o proporción del número de lecturas de cada muestra biológica de referencia procedentes de la región de 43-45 Mb del cromosoma 21 con respecto a un número de lecturas de cada muestra biológica de referencia procedentes del cromosoma 21. De acuerdo con el paso 125, se determinaron las frecuencias globales del cromosoma 22 de las muestras biológicas de referencia. La frecuencia global del cromosoma 22 de cada muestra biológica de referencia es un ratio o proporción del número de lecturas de cada muestra biológica de referencia procedentes del cromosoma 22 con respecto al número de lecturas totales de cada muestra biológica de referencia.

[0038] La Figura 6 muestra una correspondencia entre las frecuencias globales del cromosoma 22 de las muestras biológicas de referencia y las frecuencias locales de 43-45 Mb del cromosoma 21 de las muestras biológicas de referencia. En la Figura 6, el eje X es la frecuencia global del cromosoma 22, y el eje Y es la frecuencia local de 43-45 Mb del cromosoma 21. Se observó que existe una relación lineal entre las frecuencias globales del cromosoma 22 de las muestras biológicas de referencia que son normales y las frecuencias locales de 43-45 Mb del cromosoma 21 de las muestras biológicas de referencia que son normales. Las frecuencias globales del cromosoma 22 de las muestras biológicas de referencia que son normales y las frecuencias locales de 43-45 Mb del cromosoma 21 de las muestras biológicas de referencia que son normales tienen un coeficiente de determinación de 0,9475 y una línea de regresión ($Y=8,1892X-0,0341$). La línea de regresión ($Y=8,1892X-0,0341$) puede considerarse la primera correspondencia del paso 126.

[0039] De acuerdo con el paso 127, se determinaron las frecuencias globales de 43-45 Mb del cromosoma 21 de las muestras biológicas de referencia. La frecuencia global de 43-45 Mb del cromosoma 21 de cada muestra biológica de referencia es un ratio o proporción del número de lecturas de cada muestra biológica de referencia procedentes de la región de 43-45 Mb del cromosoma 21 con respecto al número de lecturas totales de cada muestra biológica de referencia.

[0040] La Figura 7 muestra una correspondencia entre las frecuencias globales del cromosoma 22 de las muestras biológicas de referencia y las frecuencias globales de 43-45 Mb del cromosoma 21 de las muestras biológicas de

referencia. En la Figura 7, el eje X es la frecuencia global del cromosoma 22, y el eje Y es la frecuencia global de 43-45 Mb del cromosoma 21. Se observó que existe una relación lineal entre las frecuencias globales del cromosoma 22 de las muestras biológicas de referencia que son normales y las frecuencias globales de 43-45 Mb del cromosoma 21 de las muestras biológicas de referencia que son normales. Las frecuencias globales del cromosoma 22 de las muestras biológicas de referencia que son normales y las frecuencias globales de 43-45 Mb del cromosoma 21 de las muestras biológicas de referencia que son normales tienen un coeficiente de determinación de 0,9374 y una línea de regresión ($Y=0,0851X-0,0003$).

[0041] De acuerdo con el paso 128a, las frecuencias locales de 43-45 Mb del cromosoma 21 de las muestras biológicas de referencia se normalizaron, respectivamente, mediante las frecuencias globales del cromosoma 22 de las muestras biológicas de referencia y la primera correspondencia. De acuerdo con el paso 128b, las frecuencias globales de 43-45 Mb del cromosoma 21 de las muestras biológicas de referencia se normalizaron, respectivamente, mediante las frecuencias locales normalizadas de 43-45 Mb del cromosoma 21 de las muestras biológicas de referencia. Específicamente, la frecuencia global normalizada de 43-45 Mb del cromosoma 21 de cada muestra biológica de referencia se calculó mediante la siguiente Fórmula I:

$$y_1 = x_3 / (x_2 / (8.1892 * x_1 - 0.0341)) \quad \text{.....fórmula I}$$

donde y_1 es la frecuencia global normalizada de 43-45 Mb del cromosoma 21 de cada muestra biológica de referencia, x_1 es la frecuencia global del cromosoma 22 de cada muestra biológica de referencia, x_2 es la frecuencia local de 43-45 Mb del cromosoma 21 de cada muestra biológica de referencia y x_3 es la frecuencia global de 43-45 Mb del cromosoma 21 de cada muestra biológica de referencia.

[0042] La Figura 8 muestra una correspondencia entre las frecuencias globales del cromosoma 22 de las muestras biológicas de referencia y las frecuencias globales normalizadas de 43-45 Mb del cromosoma 21 de las muestras biológicas de referencia. En la Figura 8, el eje X es la frecuencia global del cromosoma 22, y el eje Y es la frecuencia global normalizada de 43-45 Mb del cromosoma 21. Se observó que existe una relación lineal entre las frecuencias globales del cromosoma 22 de las muestras biológicas de referencia que son normales y las frecuencias globales normalizadas de 43-45 Mb del cromosoma 21 de las muestras biológicas de referencia que son normales. Las frecuencias globales del cromosoma 22 de las muestras biológicas de referencia que son normales y las frecuencias globales normalizadas de 43-45 Mb del cromosoma 21 de las muestras biológicas de referencia que son normales tienen un coeficiente de determinación de 0,9971 y una línea de regresión ($Y=0,0852X-0,0003$). La línea de regresión ($Y=0,0852X-0,0003$) puede considerarse la segunda correspondencia del paso 129.

[0043] En referencia a la Figura 1, en el paso 130 se determina un valor límite basándose en la base de datos de referencia. La Figura 9 es un diagrama de flujo del paso 130 de la Figura 1. En el paso 131, los valores estimados de referencia se estiman, respectivamente, basándose en las frecuencias globales correlacionadas de referencia y en la segunda correspondencia. En el paso 132, se determinan, respectivamente, los valores de diferencia de referencia entre las frecuencias globales diana de referencia normalizadas y los valores estimados de referencia. En el paso 133, los valores de diferencia de referencia se estandarizan, respectivamente, a los resultados o puntuaciones estándares de referencia basándose en la base de datos de referencia. En el paso 134, el valor límite se determina basándose en las puntuaciones estándares de referencia.

[0044] La descripción que se ofrece a continuación continuará con la detección de la trisomía 21 en varones que se muestra en las Figuras 5-8 para ilustrar cómo llevar a cabo el paso 130. En este ejemplo práctico, de acuerdo con el paso 131, los valores estimados de referencia se estimaron, respectivamente, basándose en las frecuencias globales del cromosoma 22 de las muestras biológicas de referencia y la segunda correspondencia. De acuerdo con el paso 132, se determinaron los valores de diferencia de referencia. Cada valor de diferencia de referencia se encuentra entre la frecuencia global normalizada de 43-45 Mb del cromosoma 21 de cada muestra biológica de referencia y el correspondiente valor estimado de referencia. Específicamente, cada valor de diferencia de referencia se calculó mediante la siguiente Fórmula II:

$$y_2 = y_1 - (0.0852 * x_1 - 0.0003) \quad \text{.....fórmula II}$$

donde y_1 es el ratio global normalizado de 43-45 Mb del cromosoma 21 de cada muestra biológica de referencia, y_2 es el valor de diferencia de referencia y x_1 es la frecuencia global del cromosoma 22 de cada muestra biológica de referencia.

[0045] De acuerdo con el paso 133, los valores de diferencia de referencia se estandarizaron, respectivamente, a las puntuaciones estándares de referencia. Específicamente, cada puntuación estándar de referencia se calculó mediante la siguiente Fórmula III:

$$Z = \frac{y_2 - 0.000028}{0.0000094} \dots\dots\dots \text{fórmula III}$$

donde Z es la puntuación estándar de referencia, y_2 es el valor de diferencia de referencia, 0,000028 es el promedio de los valores de diferencia de referencia y 0,0000094 es la desviación estándar de los valores de diferencia de referencia.

[0046] La Figura 10 muestra la distribución de las puntuaciones estándares de referencia de acuerdo con el ejemplo práctico que se muestra en las Figuras 5-8. En la Figura 10, el eje Y es la puntuación estándar. Tal y como se muestra en la Figura 10, las muestras anómalas o anormales están separadas de las muestras normales y, por lo tanto, se podría determinar un valor límite. En el presente ejemplo práctico, se determinó que el valor límite era de 1,75.

[0047] En referencia a la Figura 1, en el paso 140, una muestra biológica bajo prueba se secuenciará mediante la plataforma de secuenciación a fin de obtener un número de lecturas de objetivo de la muestra biológica bajo prueba. La muestra biológica bajo prueba puede ser plasma, orina, suero o cualesquiera otras muestras adecuadas. La muestra biológica bajo prueba se obtiene de una hembra embarazada y contiene moléculas de ácido nucleico de la hembra embarazada y de su feto. Las moléculas de ácido nucleico pueden ser, por ejemplo, fragmentos de cromosomas. Las lecturas de objetivo de la muestra biológica bajo prueba proceden de la región de ácido nucleico diana.

[0048] En el paso 150, el número de lecturas de objetivo de la muestra biológica bajo prueba se normaliza mediante el factor de normalización. La Figura 11 es un diagrama de flujo del paso 150 de la Figura 1. En el paso 151, se determina una frecuencia global diana de prueba. La frecuencia global diana de prueba es un ratio o proporción del número de lecturas de objetivo de la muestra biológica bajo prueba con respecto al número de lecturas totales de la muestra biológica bajo prueba. En el paso 152, se determina una frecuencia local diana de prueba. La frecuencia local diana de prueba es un ratio o proporción del número de lecturas de objetivo de la muestra biológica bajo prueba con respecto al número de lecturas locales de la muestra biológica bajo prueba. Las lecturas locales de la muestra biológica bajo prueba proceden del propio cromosoma de la región de ácido nucleico diana. En el paso 153, se determina una frecuencia global correlacionada de prueba. La frecuencia global correlacionada de prueba es un ratio o proporción de un número de lecturas correlacionadas de la muestra biológica bajo prueba con respecto al número de lecturas totales de la muestra biológica bajo prueba. Las lecturas correlacionadas de la muestra biológica bajo prueba proceden de la región de ácido nucleico correlacionada. En el paso 154a, la frecuencia local diana de prueba se normaliza por medio de la frecuencia global correlacionada de prueba y la primera correspondencia. En el paso 154b, la frecuencia global diana de prueba se normaliza por medio de la frecuencia local diana de prueba normalizada.

[0049] En el paso 155, se estima o calcula un valor estimado de prueba basándose en la frecuencia global correlacionada de prueba y la segunda correspondencia. En el paso 156, se determina un valor de diferencia de prueba entre la frecuencia global diana de prueba normalizada y el valor estimado de prueba. En el paso 157, el valor de diferencia de prueba se estandariza a una puntuación estándar de prueba basándose en la base de datos de referencia.

[0050] En referencia a la Figura 1, en el paso 160, el número normalizado de las lecturas de objetivo de la muestra biológica bajo prueba se compara con el valor límite. En el paso 170, basándose en esta comparación se determina si la aneuploidía cromosómica relacionada con la región de ácido nucleico diana está presente en el feto. Específicamente, en el paso 160, la puntuación estándar de prueba calculada mediante el paso 157 se compara con el valor límite determinado mediante el paso 130. Después, en el paso 170, si la puntuación estándar de prueba calculada mediante el paso 157 es mayor que el valor límite determinado mediante el paso 130, se determina que la aneuploidía cromosómica relacionada con la región de ácido nucleico diana está presente en el feto, es decir, 'anormal'. Si la puntuación estándar de prueba calculada mediante el paso 157 es menor que el valor límite determinado mediante el paso 130, se determina que la aneuploidía cromosómica relacionada con la región de ácido nucleico diana está ausente del feto, es decir, 'normal'.

[0051] La descripción que se ofrece a continuación continuará con la detección de la trisomía 21 en varones que se muestra en las Figuras 5-8 y 10 para ilustrar cómo llevar a cabo los pasos 140-170. De acuerdo con el paso 140, se obtuvo una muestra biológica bajo prueba (plasma) de una hembra o mujer embarazada. Después, la muestra biológica bajo prueba se secuenció mediante Illumina HiSeq.

[0052] De acuerdo con el paso 151, se determinó una frecuencia global de 43-45 Mb del cromosoma 21 de la muestra biológica bajo prueba. La frecuencia global de 43-45 Mb del cromosoma 21 de la muestra biológica bajo prueba es un ratio o proporción de un número de lecturas de la muestra biológica bajo prueba procedentes de la región de 43-45 Mb del cromosoma 21 con respecto a un número de lecturas totales de la muestra biológica bajo prueba. De acuerdo con el paso 152, se determinó una frecuencia local de 43-45 Mb del cromosoma 21 de la muestra biológica bajo prueba. La frecuencia local de 43-45 Mb del cromosoma 21 de la muestra biológica bajo prueba es un ratio o proporción del número de lecturas de la muestra biológica bajo prueba procedentes de la región

de 43-45 Mb del cromosoma 21 con respecto a un número de lecturas de la muestra biológica bajo prueba procedentes del cromosoma 21. De acuerdo con el paso 153, se determinó una frecuencia global del cromosoma 22 de la muestra biológica bajo prueba. La frecuencia global del cromosoma 22 de la muestra biológica bajo prueba es un ratio o proporción de las lecturas de la muestra biológica bajo prueba procedentes del cromosoma 22 con respecto al número de lecturas totales de la muestra biológica bajo prueba.

[0053] De acuerdo con el paso 154a, la frecuencia local de 43-45 Mb del cromosoma 21 de la muestra biológica bajo prueba se normalizó mediante la frecuencia global del cromosoma 22 de la muestra biológica bajo prueba y la primera correspondencia. En el paso 154b, la frecuencia global de 43-45 Mb del cromosoma 21 de la muestra biológica bajo prueba se normalizó mediante la frecuencia local normalizada de 43-45 Mb del cromosoma 21 de la muestra biológica bajo prueba. Específicamente, la frecuencia global normalizada de 43-45 Mb del cromosoma 21 de la muestra biológica bajo prueba se calculó mediante la siguiente Fórmula IV:

$$y1 = x3 / (x2 / (8.1892 * x1 - 0.0341)) \dots\dots\dots \text{fórmula IV}$$

donde y1 es la frecuencia global normalizada de 43-45 Mb del cromosoma 21 de la muestra biológica bajo prueba, x1 es la frecuencia global del cromosoma 22 de la muestra biológica bajo prueba, x2 es la frecuencia local de 43-45 Mb del cromosoma 21 de la muestra biológica bajo prueba y x3 es la frecuencia global de 43-45 Mb del cromosoma 21 de la muestra biológica bajo prueba.

[0054] De acuerdo con el paso 155, se estimó un valor estimado de prueba basándose en la frecuencia global del cromosoma 22 de la muestra biológica bajo prueba y la segunda correspondencia. De acuerdo con el paso 156, se determinó un valor de diferencia de prueba entre la frecuencia global normalizada de 43-45 Mb del cromosoma 21 de la muestra biológica bajo prueba y el valor estimado de prueba. Específicamente, el valor de diferencia de prueba se calculó mediante la siguiente Fórmula V:

$$y2 = y1 - (0.0852 * x1 - 0.0003) \dots\dots\dots \text{fórmula V}$$

donde y1 es la frecuencia global normalizada de 43-45 Mb del cromosoma 21 de la muestra biológica bajo prueba, y2 es el valor de diferencia de prueba y x1 es la frecuencia global del cromosoma 22 de la muestra biológica bajo prueba.

[0055] De acuerdo con el paso 157, el valor de diferencia de prueba se estandarizó a una puntuación estándar de prueba basándose en la base de datos de referencia. Específicamente, la puntuación estándar de prueba se calculó mediante la siguiente Fórmula VI:

$$Z = y2 - 0.000028 / 0.0000094 \dots\dots\dots \text{fórmula VI}$$

donde Z es la puntuación estándar de prueba, y2 es el valor de diferencia de prueba, 0,000028 es el promedio de los valores de diferencia de referencia y 0,0000094 es la desviación estándar de los valores de diferencia de referencia.

[0056] La Figura 12 muestra la distribución de las puntuaciones estándares de prueba y de referencia de acuerdo con el ejemplo práctico que se muestra en las Figuras 5-8 y 10. En la Figura 12, el eje Y es la puntuación estándar. Tal y como se muestra en la Figura 12, la puntuación estándar es -0,16, que es menos que el valor límite, esto es, 1,75. Por lo tanto, se determinó que la muestra biológica bajo prueba era normal. Es decir, se determinó que la trisomía 21 no estaba presente en el feto (o, en otras palabras, estaba ausente del feto).

[0057] En algunas realizaciones, el método que se ha descrito previamente se implementa como un programa almacenado en un medio no transitorio legible por una máquina. Cuando se ejecuta el programa al menos mediante una unidad de procesamiento, se lleva a cabo el método previamente descrito. El medio no transitorio legible por una máquina puede incluir -pero no se limita a- disquetes o discos flexibles, discos ópticos, discos compactos (CDs), discos de vídeo digitales (DVDs), discos magneto-ópticos, memorias solo de lectura (ROMs), memorias de acceso aleatorio (RAMs), memorias EPROM, memorias EEPROM, tarjetas ópticas o magnéticas o cualquier otro medio o medio legible por una máquina que sea adecuado para guardar o almacenar instrucciones.

REIVINDICACIONES

1. Un método para detectar una aneuploidía cromosómica relacionada con una región de ácido nucleico diana (100), de manera que el método comprende:

obtener una base de datos de referencia (110), de manera que la base de datos de referencia se establece secuenciando diversas muestras biológicas de referencia mediante una plataforma de secuenciación; determinar los factores de normalización basándose en la base de datos de referencia, de manera que la determinación de los factores de normalización (120) comprende:

determinar un número de lecturas de objetivo (o 'target reads', en inglés) de cada muestra biológica de referencia (121), de manera que las lecturas de objetivo de cada muestra biológica de referencia proceden de la región de ácido nucleico diana;

determinar un número de lecturas correlacionadas de cada muestra biológica de referencia (122), de manera que el número de lecturas correlacionadas de las muestras biológicas de referencia se correlaciona de manera lineal con el número de lecturas de objetivo de las muestras biológicas de referencia, y de manera que tienen un coeficiente de correlación en un rango o intervalo de entre alrededor de un 0,7 y alrededor de un 0,99, que se calcula mediante un modelo de regresión lineal, y de manera que las lecturas correlacionadas de cada muestra biológica de referencia proceden de una región de ácido nucleico correlacionada; y

determinar los factores de normalización basándose en el número de lecturas de objetivo de las muestras biológicas de referencia y el número de lecturas correlacionadas de las muestras biológicas de referencia (123), de manera que los factores de normalización comprenden:

una primera correspondencia entre las frecuencias locales diana de referencia y las frecuencias globales correlacionadas de referencia,

y una frecuencia global correlacionada de prueba para normalizar una frecuencia local diana de prueba, de manera que la frecuencia global correlacionada de prueba es un ratio o proporción de un número de lecturas correlacionadas de la muestra biológica bajo prueba con respecto al número de lecturas totales de la muestra biológica bajo prueba, y las lecturas correlacionadas de la muestra biológica bajo prueba proceden de la región de ácido nucleico correlacionada, y la frecuencia local diana de prueba es un ratio o proporción del número de lecturas de objetivo de la muestra biológica bajo prueba con respecto a un número de lecturas locales de la muestra biológica bajo prueba, y las lecturas locales de la muestra biológica bajo prueba proceden del propio cromosoma de la región de ácido nucleico diana; y

una frecuencia local diana de prueba normalizada, de manera que la frecuencia global diana de prueba es un ratio o proporción del número de lecturas de objetivo de la muestra biológica bajo prueba con respecto al número de lecturas totales de la muestra biológica bajo prueba;

determinar un valor límite basándose en la base de datos de referencia (130);

secuenciar una muestra biológica bajo prueba mediante la plataforma de secuenciación a fin de obtener un número de lecturas de objetivo de la muestra biológica bajo prueba (140), de manera que la muestra biológica bajo prueba se obtiene de una hembra o mujer embarazada y tiene moléculas de ácido nucleico de la hembra embarazada y su feto, y las lecturas de objetivo de la muestra biológica bajo prueba proceden de la región de ácido nucleico diana;

normalizar el número de lecturas de objetivo de la muestra biológica bajo prueba mediante los factores de normalización (150);

comparar el número normalizado de las lecturas de objetivo de la muestra biológica bajo prueba con el valor límite (160); y

basándose en esta comparación, determinar si la aneuploidía cromosómica relacionada con la región de ácido nucleico diana está presente en el feto (170).

2. El método de la reivindicación 1, de manera que la base de datos de referencia está basada en el género.

3. El método de la reivindicación 1, de manera que la determinación de los factores de normalización basada en los números de las lecturas de objetivo de las muestras biológicas de referencia y los números de las lecturas correlacionadas de las muestras biológicas de referencia (123) comprende:

determinar las frecuencias locales diana de referencia (124), de manera que cada frecuencia local diana de referencia es un ratio o proporción del número de lecturas de objetivo de cada muestra biológica de referencia con respecto a un número de lecturas locales de cada muestra biológica de referencia, y las lecturas locales de cada muestra biológica de referencia proceden del propio cromosoma de la región de ácido nucleico diana;

determinar las frecuencias globales correlacionadas de referencia (125), de manera que cada frecuencia

global correlacionada de referencia es un ratio o proporción del número de lecturas correlacionadas de cada muestra biológica de referencia con respecto al número de lecturas totales de cada muestra biológica de referencia; y

5 determinar la primera correspondencia entre las frecuencias locales diana de referencia y las frecuencias globales correlacionadas de referencia (126).

4. El método de la reivindicación 3, de manera que la normalización del número de lecturas de objetivo de la muestra biológica bajo prueba (150) comprende:

10 determinar la frecuencia global diana de prueba (127);
determinar la frecuencia local diana de prueba (124);
determinar la frecuencia global correlacionada de prueba (125);
normalizar la frecuencia local diana de prueba mediante la frecuencia global correlacionada de prueba y la primera correspondencia (128a); y
15 normalizar la frecuencia global diana de prueba mediante la frecuencia local diana de prueba normalizada (128b).

5. El método de la reivindicación 4, de manera que la determinación de los factores de normalización basada en los números de las lecturas de objetivo de las muestras biológicas de referencia y los números de las lecturas correlacionadas de las muestras biológicas de referencia (123) comprende:

20 determinar las frecuencias globales diana de referencia (127), de manera que cada frecuencia global diana de referencia es un ratio o proporción del número de lecturas de objetivo de cada muestra biológica de referencia con respecto al número de lecturas totales de cada muestra biológica de referencia;
25 normalizar, respectivamente, las frecuencias locales diana de referencia mediante las frecuencias globales correlacionadas de referencia y la primera correspondencia (128a);
normalizar, respectivamente, las frecuencias globales diana de referencia mediante las frecuencias locales diana de referencia normalizadas (128b); y
30 determinar una segunda correspondencia entre las frecuencias globales diana de referencia normalizadas y las frecuencias globales correlacionadas de referencia (129).

6. El método de la reivindicación 5, de manera que la determinación del valor límite (130) comprende:

35 estimar, respectivamente, los valores estimados de referencia basándose en las frecuencias globales correlacionadas de referencia y la segunda correspondencia (131);
determinar, respectivamente, los valores de diferencia de referencia entre las frecuencias globales diana de referencia normalizadas y los valores estimados de referencia (132);
estandarizar, respectivamente, los valores de diferencia de referencia a los resultados o puntuaciones estándares de referencia basándose en la base de datos de referencia (133); y
40 determinar el valor límite basándose en las puntuaciones estándares de referencia (134).

7. El método de la reivindicación 6, de manera que la normalización del número de lecturas de objetivo de la muestra biológica bajo prueba (150) comprende:

45 estimar un valor estimado de prueba basándose en la frecuencia global correlacionada de prueba y la segunda correspondencia (155);
determinar un valor de diferencia de prueba entre la frecuencia global diana de prueba normalizada y el valor estimado de prueba (156); y
estandarizar el valor de diferencia de prueba a una puntuación estándar de prueba basándose en la base de datos de referencia (157); y
50 de manera que la comparación comprende:
comparar la puntuación estándar de prueba con el valor límite.

8. Un medio no transitorio legible por una máquina que almacena un programa que, al ejecutarse mediante al menos una unidad de procesamiento, detecta una aneuploidía cromosómica relacionada con una región de ácido nucleico diana, de manera que el programa comprende un conjunto de instrucciones para:

60 obtener una base de datos de referencia (110), de manera que la base de datos de referencia se establece secuenciando diversas muestras biológicas de referencia mediante una plataforma de secuenciación;
determinar los factores de normalización basándose en la base de datos de referencia (120), de manera que el conjunto de instrucciones para la determinación de los factores de normalización (120) comprende un conjunto de instrucciones para:

65 determinar un número de lecturas de objetivo de cada muestra biológica de referencia (121), de manera que las lecturas de objetivo de cada muestra biológica de referencia proceden de la región de ácido nucleico diana;

- 5 determinar un número de lecturas correlacionadas de cada muestra biológica de referencia (122), de manera que el número de lecturas correlacionadas de las muestras biológicas de referencia se correlaciona de manera lineal con el número de lecturas de objetivo de las muestras biológicas de referencia, y de manera que tienen un coeficiente de correlación en un rango o intervalo de entre alrededor de un 0,7 y alrededor de un 0,99, que se calcula mediante un modelo de regresión lineal, y de manera que las lecturas correlacionadas de cada muestra biológica de referencia proceden de una región de ácido nucleico correlacionada; y
- 10 determinar los factores de normalización basándose en el número de lecturas de objetivo de las muestras biológicas de referencia y el número de lecturas correlacionadas de las muestras biológicas de referencia (123), de manera que los factores de normalización comprenden:
- 15 una primera correspondencia entre las frecuencias locales diana de referencia y las frecuencias globales correlacionadas de referencia,
- 20 y una frecuencia global correlacionada de prueba para normalizar una frecuencia local diana de prueba, de manera que la frecuencia global correlacionada de prueba es un ratio o proporción de un número de lecturas correlacionadas de la muestra biológica bajo prueba con respecto al número de lecturas totales de la muestra biológica bajo prueba, y las lecturas correlacionadas de la muestra biológica bajo prueba proceden de la región de ácido nucleico correlacionada, y la frecuencia local diana de prueba es un ratio o proporción del número de lecturas de objetivo de la muestra biológica bajo prueba con respecto a un número de lecturas locales de la muestra biológica bajo prueba, y las lecturas locales de la muestra biológica bajo prueba proceden del propio cromosoma de la región de ácido nucleico diana; y
- 25 una frecuencia local diana de prueba normalizada para normalizar una frecuencia global diana de prueba, de manera que la frecuencia global diana de prueba es un ratio o proporción del número de lecturas de objetivo de la muestra biológica bajo prueba con respecto a un número de lecturas totales de la muestra biológica bajo prueba;
- 30 determinar un valor límite basándose en la base de datos de referencia (130);
- 35 secuenciar una muestra biológica bajo prueba mediante la plataforma de secuenciación a fin de obtener un número de lecturas de objetivo de la muestra biológica bajo prueba (140), de manera que la muestra biológica bajo prueba se obtiene de una hembra o mujer embarazada y tiene moléculas de ácido nucleico de la hembra embarazada y su feto, y las lecturas de objetivo de la muestra biológica bajo prueba proceden de la región de ácido nucleico diana;
- 40 normalizar el número de lecturas de objetivo de la muestra biológica bajo prueba mediante los factores de normalización (150);
- comparar el número normalizado de las lecturas de objetivo de la muestra biológica bajo prueba con el valor límite (160); y
- basándose en esta comparación, determinar si la aneuploidía cromosómica relacionada con la región de ácido nucleico diana está presente en el feto (170).
- 9.** El medio no transitorio legible por una máquina de la reivindicación 8, de manera que la base de datos de referencia está basada en el género.
- 45 **10.** El medio no transitorio legible por una máquina de la reivindicación 8, de manera que el conjunto de instrucciones para la determinación de los factores de normalización basada en los números de las lecturas de objetivo de las muestras biológicas de referencia y los números de las lecturas correlacionadas de las muestras biológicas de referencia (123) comprende un conjunto de instrucciones para:
- 50 determinar las frecuencias locales diana de referencia (124), de manera que cada frecuencia local diana de referencia es un ratio o proporción del número de lecturas de objetivo de cada muestra biológica de referencia con respecto a un número de lecturas locales de cada muestra biológica de referencia, y las lecturas locales de cada muestra biológica de referencia proceden del propio cromosoma de la región de ácido nucleico diana;
- 55 determinar las frecuencias globales correlacionadas de referencia (125), de manera que cada frecuencia global correlacionada de referencia es un ratio o proporción del número de lecturas correlacionadas de cada muestra biológica de referencia con respecto al número de lecturas totales de cada muestra biológica de referencia; y
- 60 determinar la primera correspondencia entre las frecuencias locales diana de referencia y las frecuencias globales correlacionadas de referencia (126).
- 11.** El medio no transitorio legible por una máquina de la reivindicación 10, de manera que el conjunto de instrucciones para normalizar el número de lecturas de objetivo de la muestra biológica bajo prueba (150) comprende un conjunto de instrucciones para:
- 65 determinar la frecuencia global diana de prueba (127);

determinar la frecuencia local diana de prueba (124);
 determinar la frecuencia global correlacionada de prueba (125);
 normalizar la frecuencia local diana de prueba mediante la frecuencia global correlacionada de prueba y la
 primera correspondencia (128a); y
 normalizar la frecuencia global diana de prueba mediante la frecuencia local diana de prueba normalizada
 (128b).

12. El medio no transitorio legible por una máquina de la reivindicación 11, de manera que el conjunto de instrucciones para la determinación de los factores de normalización basada en los números de las lecturas de objetivo de las muestras biológicas de referencia y los números de las lecturas correlacionadas de las muestras biológicas de referencia (123) comprende un conjunto de instrucciones para:

determinar las frecuencias globales diana de referencia (127), de manera que cada frecuencia global diana de referencia es un ratio o proporción del número de lecturas de objetivo de cada muestra biológica de referencia con respecto al número de lecturas totales de cada muestra biológica de referencia;
 normalizar, respectivamente, las frecuencias locales diana de referencia mediante las frecuencias globales correlacionadas de referencia y la primera correspondencia (128a);
 normalizar, respectivamente, las frecuencias globales diana de referencia mediante las frecuencias locales diana de referencia normalizadas (128b); y
 determinar una segunda correspondencia entre las frecuencias globales diana de referencia normalizadas y las frecuencias globales correlacionadas de referencia (129).

13. El medio no transitorio legible por una máquina de la reivindicación 12, de manera que el conjunto de instrucciones para determinar el valor límite (130) comprende un conjunto de instrucciones para:

estimar, respectivamente, los valores estimados de referencia basándose en las frecuencias globales correlacionadas de referencia y la segunda correspondencia (131);
 determinar, respectivamente, los valores de diferencia de referencia entre las frecuencias globales diana de referencia normalizadas y los valores estimados de referencia (132);
 estandarizar, respectivamente, los valores de diferencia de referencia a las puntuaciones estándares de referencia basándose en la base de datos de referencia (133); y
 determinar el valor límite basándose en las puntuaciones estándares de referencia (134).

14. El medio no transitorio legible por una máquina de la reivindicación 13, de manera que el conjunto de instrucciones para la normalización del número de lecturas de objetivo de la muestra biológica bajo prueba (150) comprende un conjunto de instrucciones para:

estimar un valor estimado de prueba basándose en la frecuencia global correlacionada de prueba y la segunda correspondencia (155);
 determinar un valor de diferencia de prueba entre la frecuencia global diana de prueba normalizada y el valor estimado de prueba (156); y
 estandarizar el valor de diferencia de prueba a una puntuación estándar de prueba basándose en la base de datos de referencia (157);
 de manera que el conjunto de instrucciones para la comparación comprende un conjunto de instrucciones para:
 comparar la puntuación estándar de prueba con el valor límite.

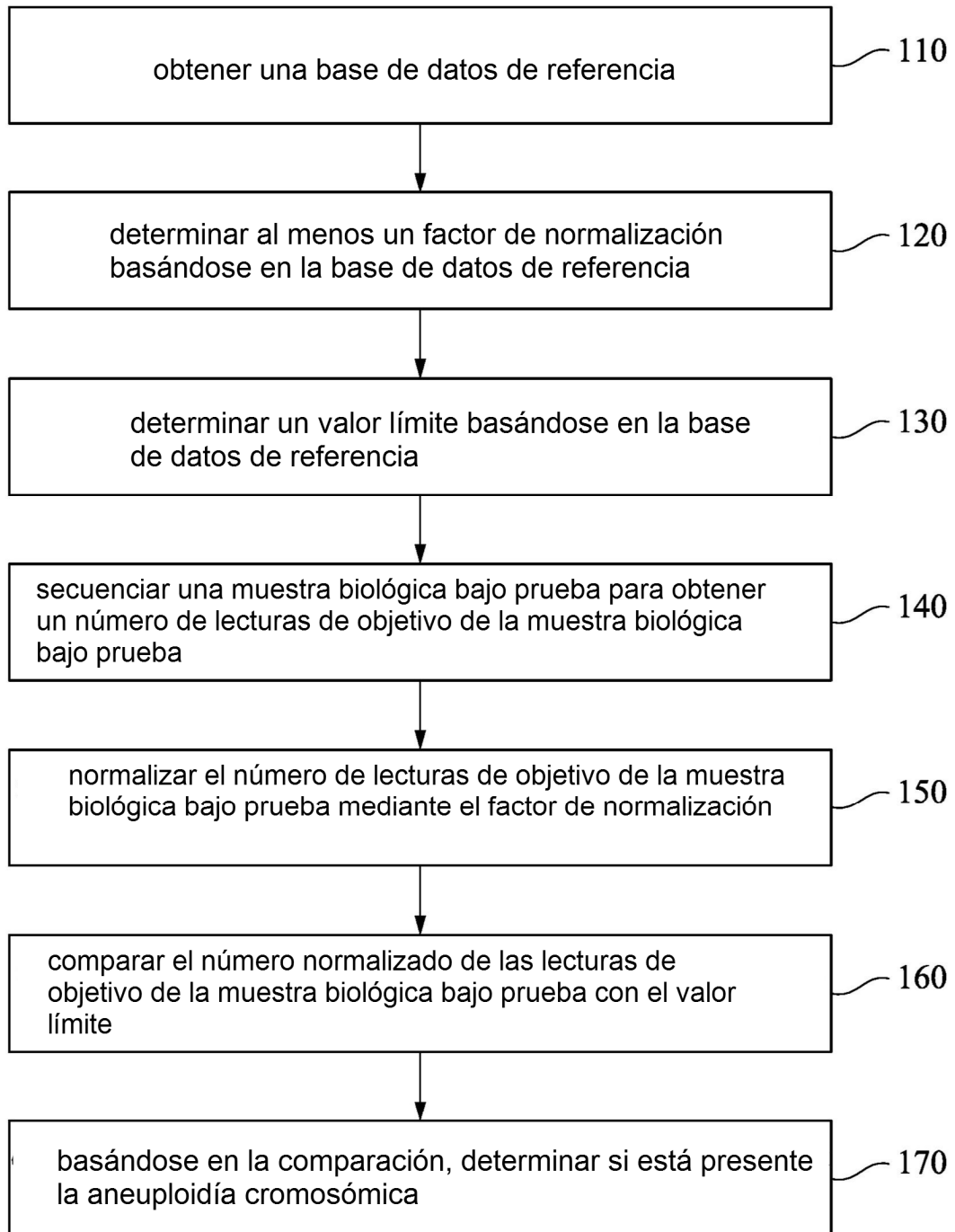
100

Fig. 1

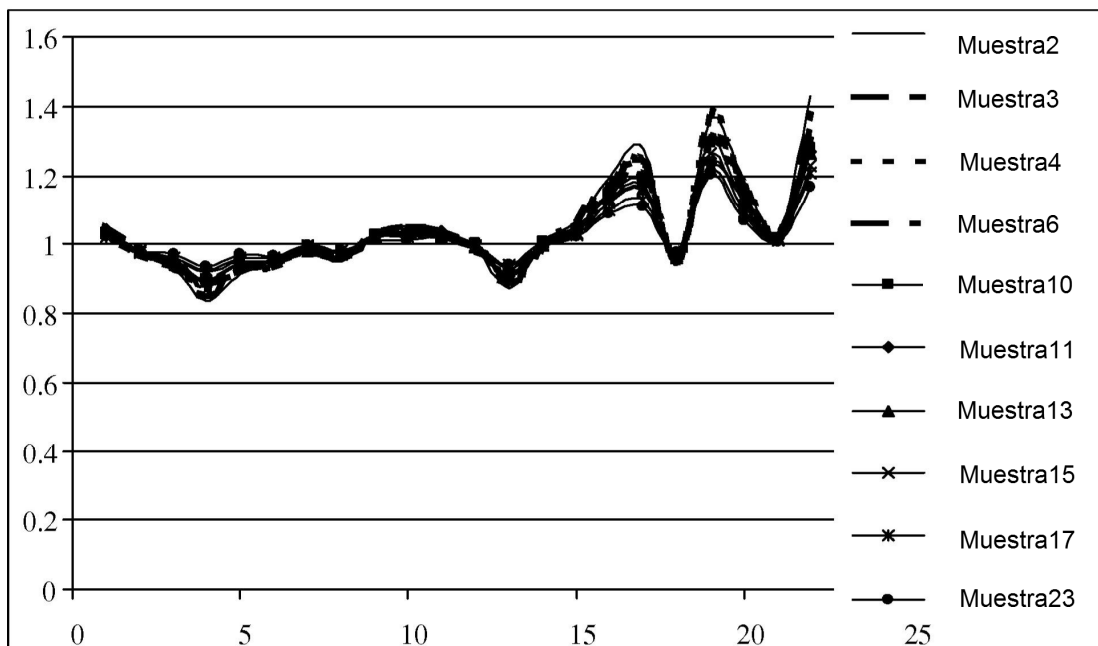


Fig. 2A

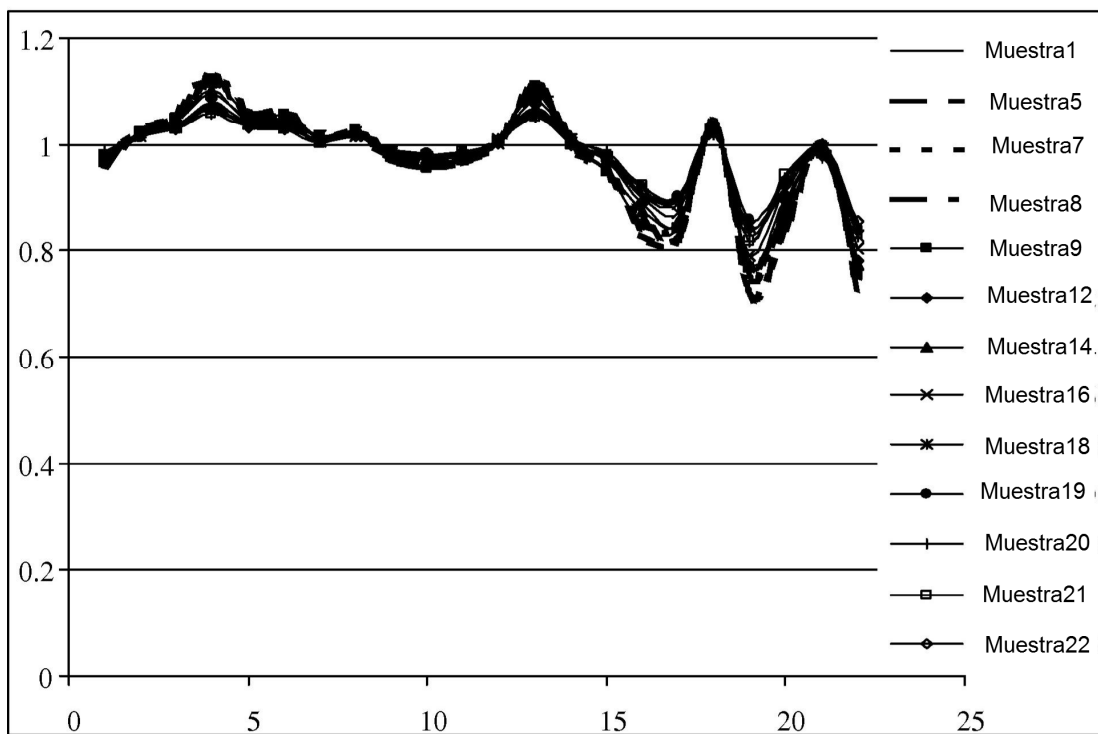


Fig. 2B

120

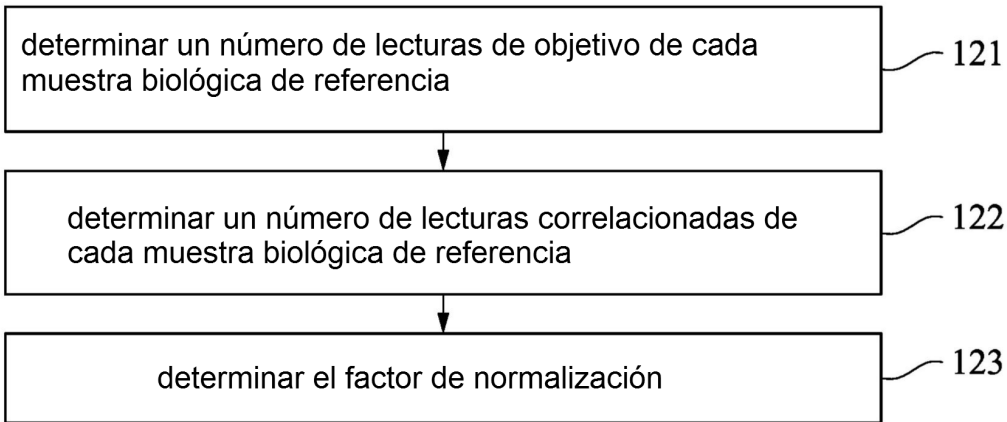


Fig. 3

123

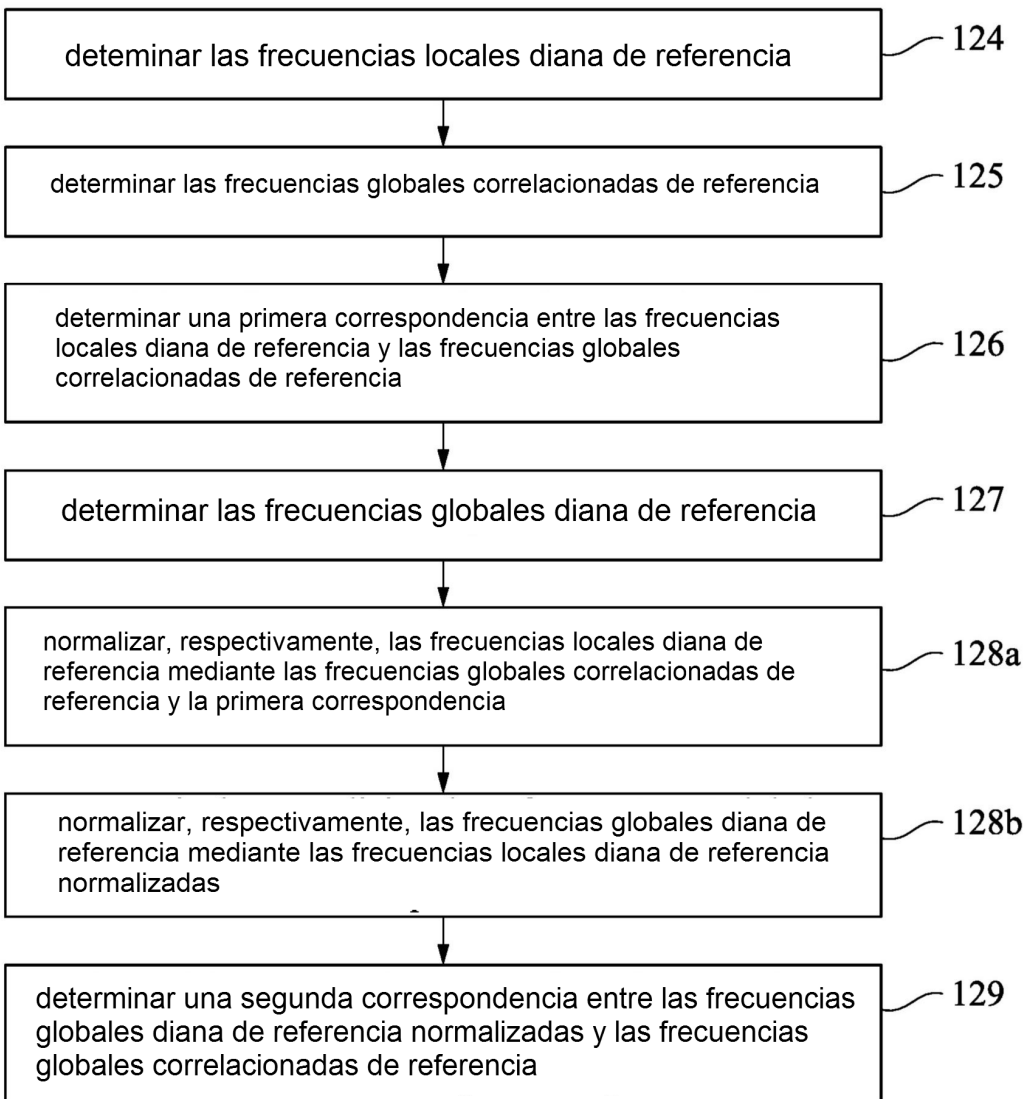


Fig. 4

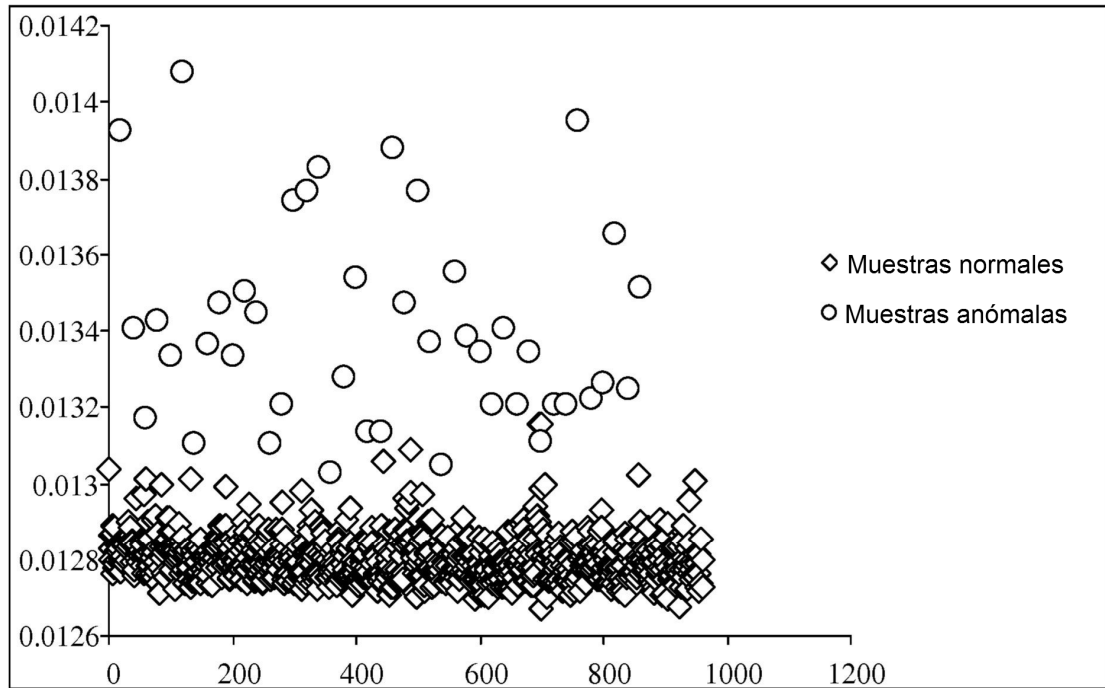


Fig. 5

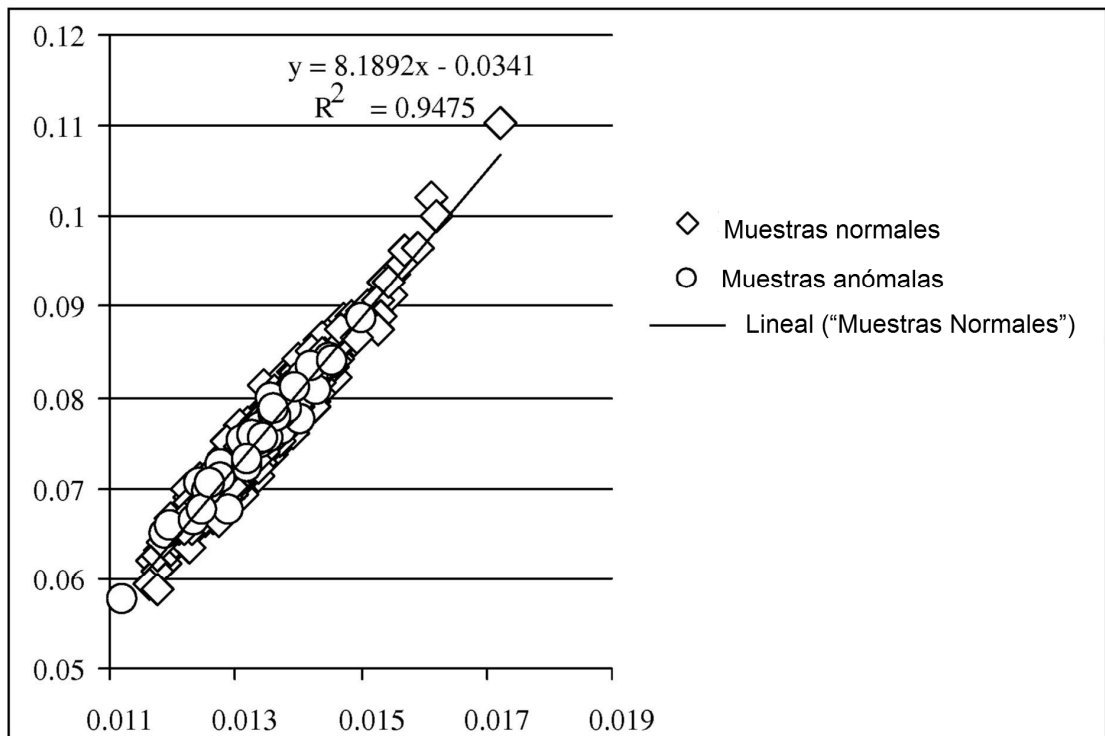


Fig. 6

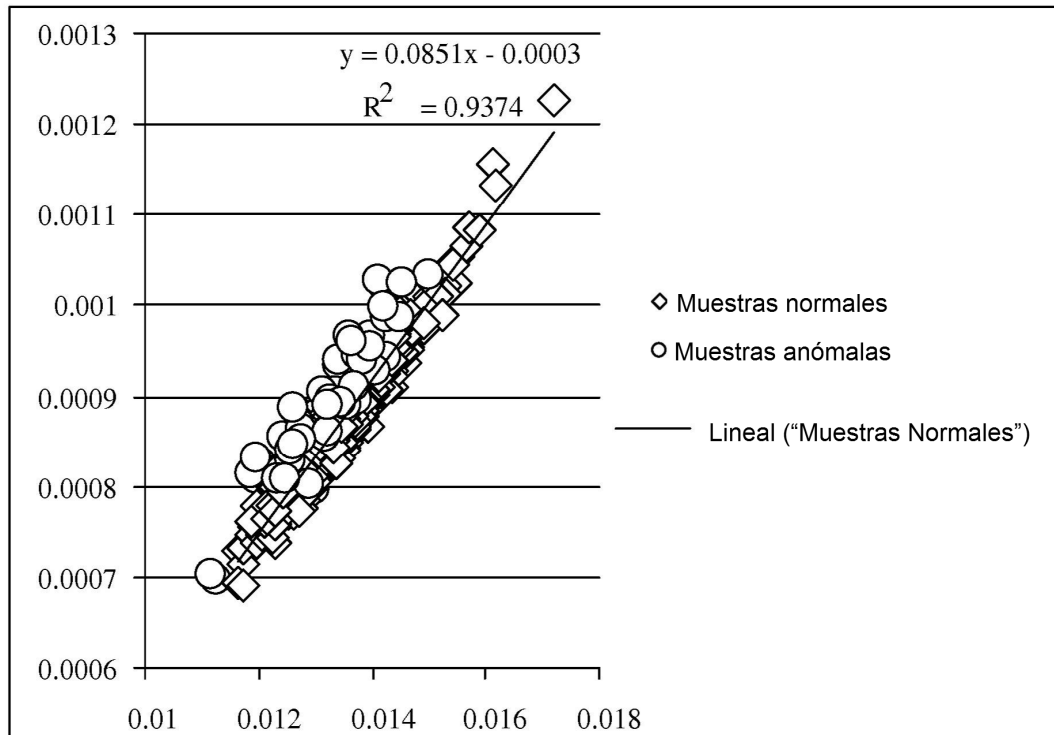


Fig. 7

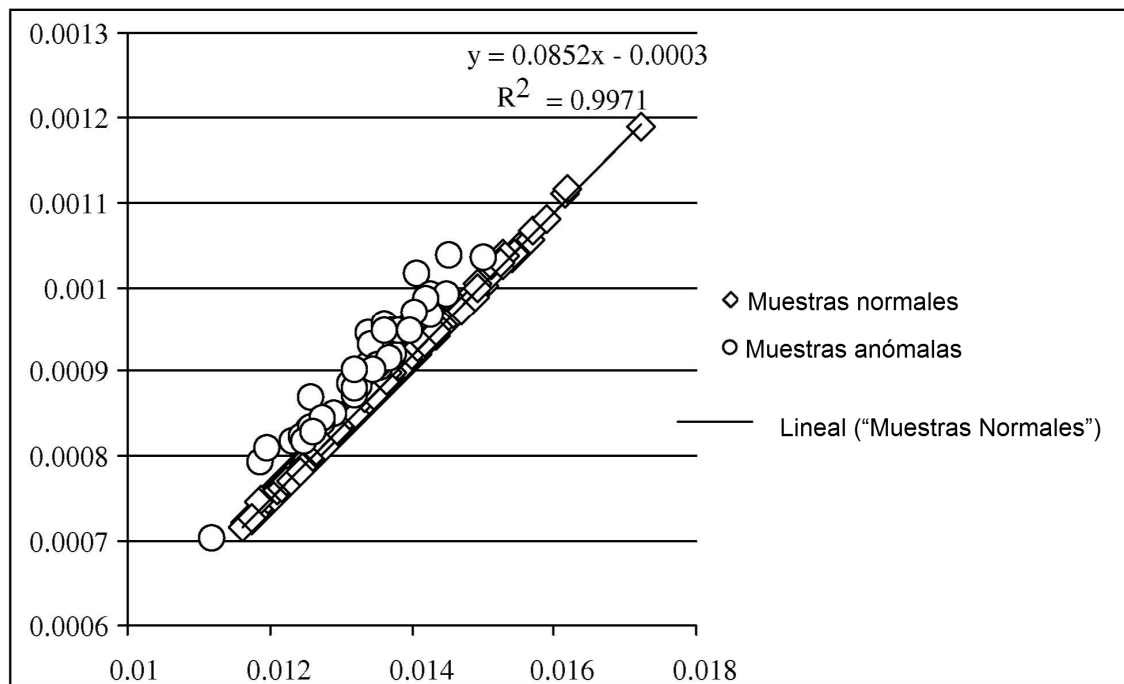


Fig. 8

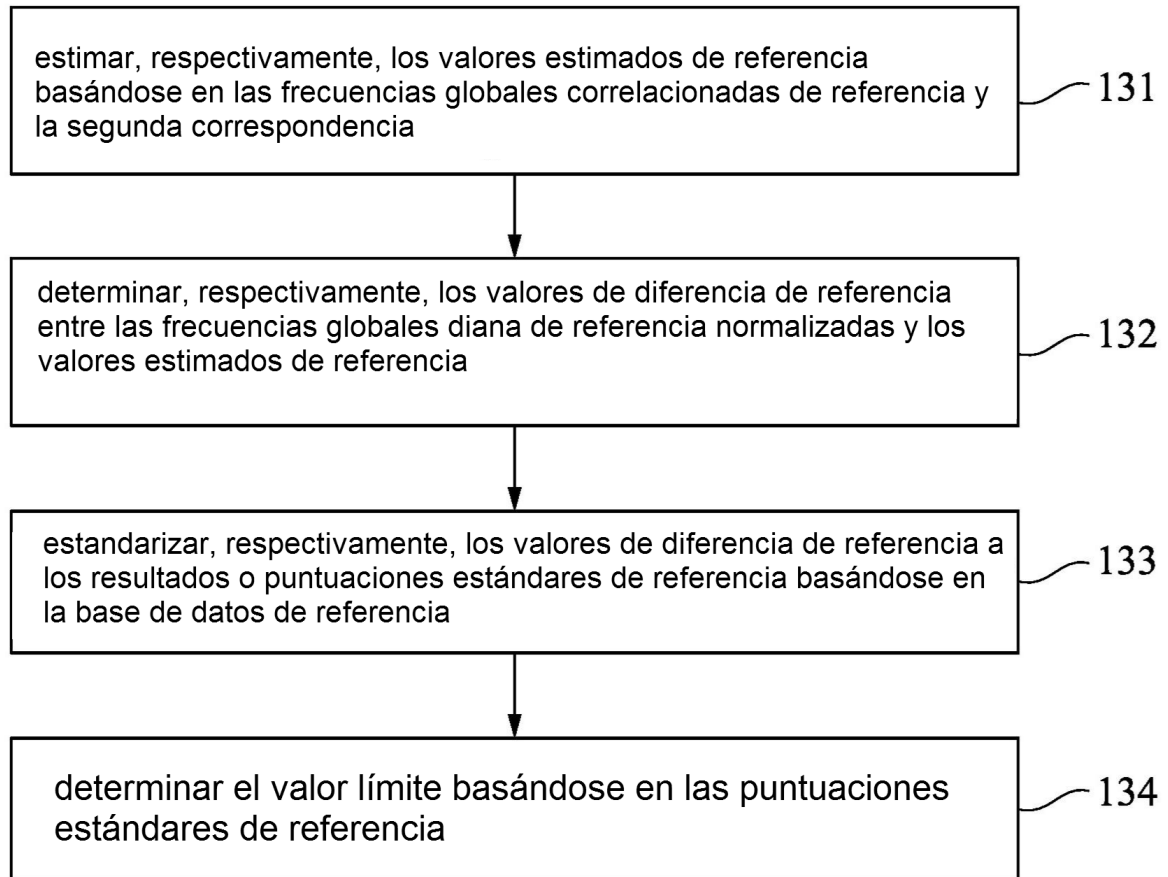
130

Fig. 9

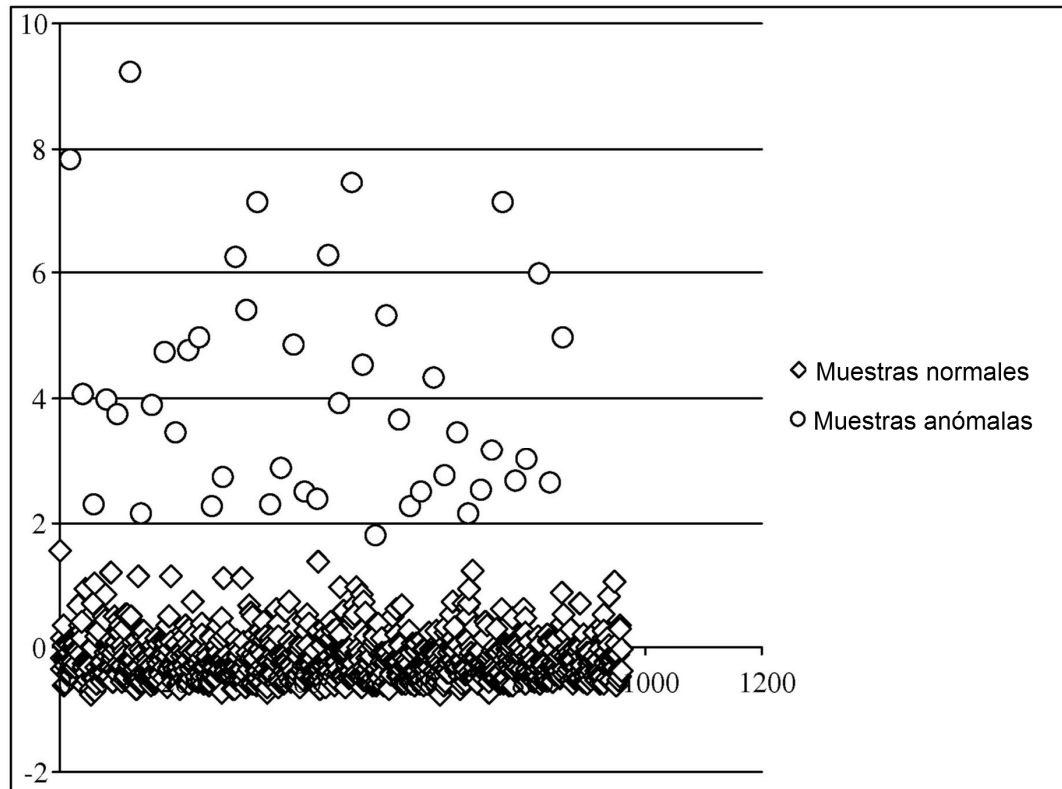


Fig. 10

150

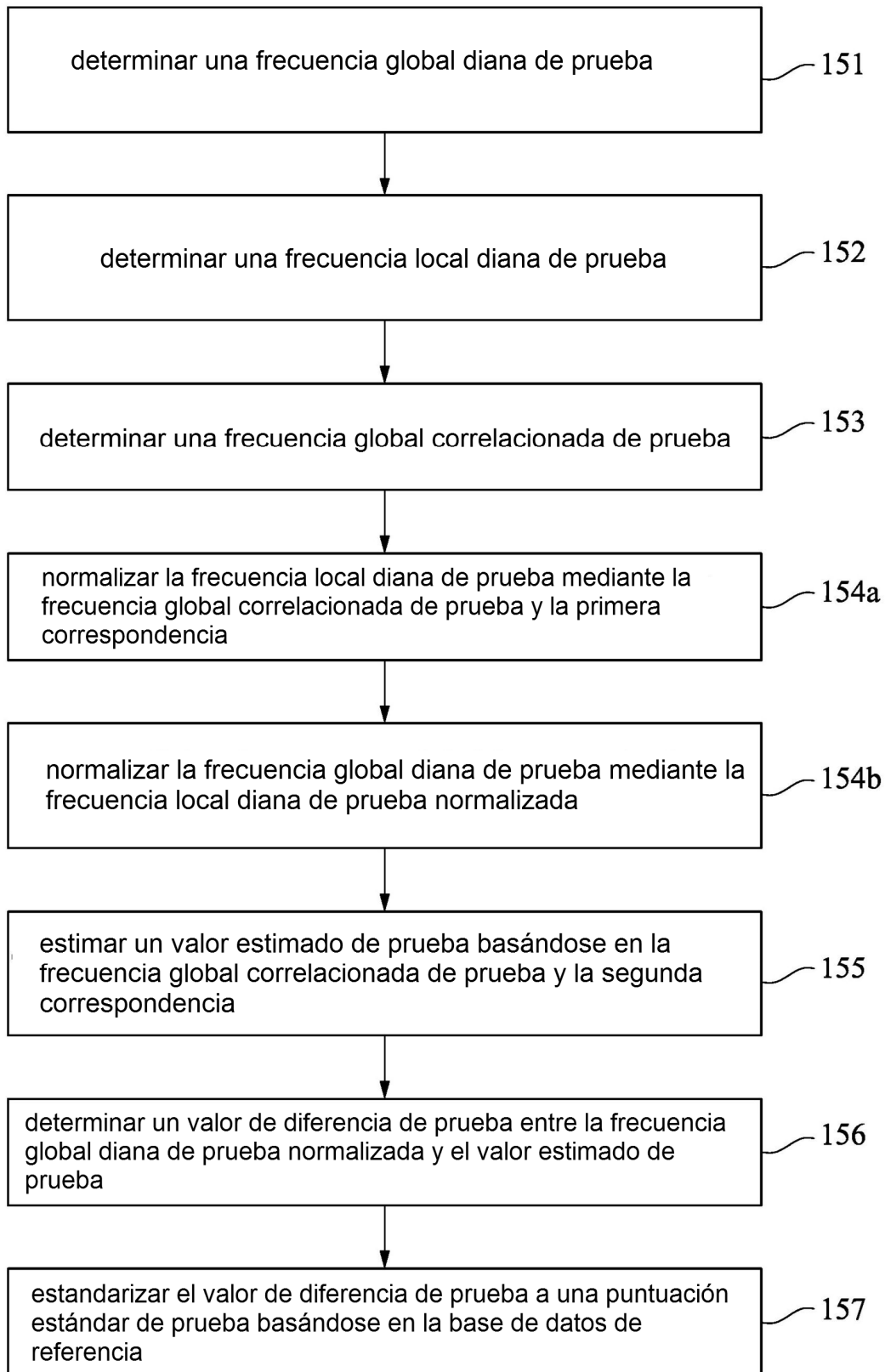


Fig. 11

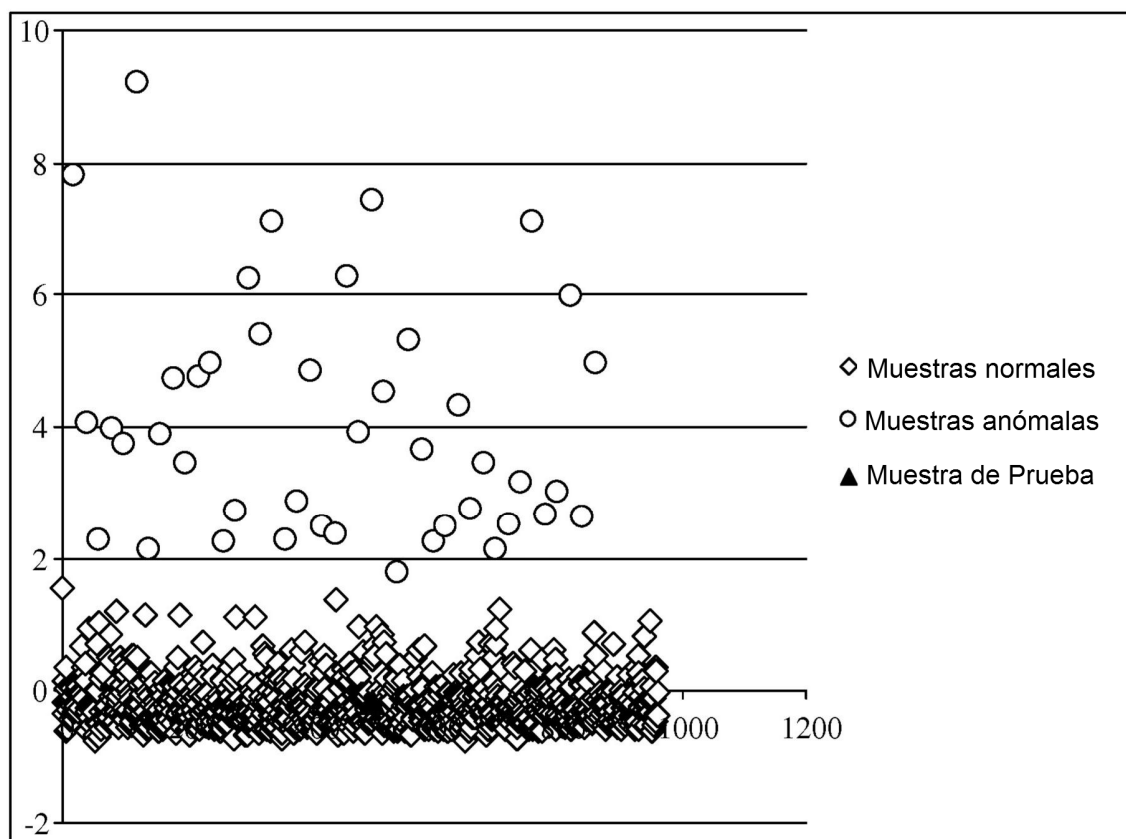


Fig. 12