

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 903 103**

51 Int. Cl.:

C12Q 1/68 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **24.07.2015 PCT/CN2015/085109**

87 Fecha y número de publicación internacional: **28.01.2016 WO16011982**

96 Fecha de presentación y número de la solicitud europea: **24.07.2015 E 15825462 (3)**

97 Fecha y número de publicación de la concesión europea: **13.10.2021 EP 3178941**

54 Título: **Método para determinar la fracción de ácidos nucleicos fetales libres de células en una muestra de sangre periférica de una mujer embarazada y uso del mismo**

30 Prioridad:

25.07.2014 CN 201410359726

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

31.03.2022

73 Titular/es:

**BGI GENOMICS CO., LIMITED (100.0%)
Floors 7-14, Building No.7 BGI Park, No.21
Hongan 3rd Street Yantian District
Shenzhen, Guangdong 518083, CN**

72 Inventor/es:

**JIANG, FUMAN;
YUAN, YUYING;
WANG, WEI y
YIN, YE**

74 Agente/Representante:

SÁEZ MAESO, Ana

ES 2 903 103 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

Método para determinar la fracción de ácidos nucleicos fetales libres de células en una muestra de sangre periférica de una mujer embarazada y uso del mismo

Campo de la invención

5 La presente divulgación se refiere a determinar la fracción de ácidos nucleicos fetales libres de células en una muestra de sangre periférica de una mujer embarazada y usos de los mismos.

Antecedentes

10 Desde 1977, los investigadores han encontrado sucesivamente ADN derivado del cáncer en la sangre periférica de pacientes con tumores; también confirmaron la presencia de ADN fetal libre de células (ADN-flc) en plasma de mujeres embarazadas. La determinación de la fracción de ADN fetal libre de células en plasma de una mujer embarazada es de gran importancia.

15 Los documentos WO2013/132305 (The Chinese University de Hong Kong), WO2014/068075 (Genesupport SA) y Yu et al. (2014) Proc. Natl.Acad. Sci. 111, 8583-8588) ejemplifican divulgaciones que buscan hacer uso de la distribución del tamaño de fragmentos de ácidos nucleicos libres de células en muestras de sangre de mujeres embarazadas en el contexto de pruebas fetales no invasivas para detectar anomalías genéticas fetales. Sin embargo, los métodos actuales para determinar la fracción de ácidos nucleicos fetales libres de células en dichas muestras biológicas quedan por mejorar. Un objeto de la presente divulgación es proporcionar un método capaz de determinar de forma precisa y eficaz la fracción de ácidos nucleicos fetales libres de células en una muestra de sangre materna.

20 En la actualidad, existen dos direcciones principales para estimar la fracción de ADN fetal libre de células en sangre periférica: (1) aprovechar las diferentes respuestas de los fragmentos de ADN materno y los fragmentos de ADN fetal libre de células en las células mononucleares de la sangre periférica materna a la metilación de marcadores específicos; (2) seleccionar una pluralidad de sitios SNP representativos en base a las diferencias entre los sitios de polimorfismo de un solo nucleótido (SNP). Ambos métodos tienen una cierta limitación: El método (1) requiere una gran cantidad de plasma, y el método (2) requiere captura de sonda y alta profundidad de secuenciación, o requiere la obtención de información de los progenitores. Sin embargo, hasta ahora no se ha publicado ningún informe sobre la estimación de la fracción de ADN fetal libre de células con secuenciación del genoma completo a una profundidad de cobertura baja. Los estudios han demostrado que los fragmentos de ADN fetal libres de células en la circulación sanguínea materna, la mayoría de los cuales tienen menos de 313 pb, son generalmente más cortos que los fragmentos de ADN materno libres de células. Bajo esta inspiración, los inventores han inventado un procedimiento para estimar la fracción de ADN fetal libre de células en base a la secuenciación del plasma de la mujer embarazada, y cuyo método tiene amplias aplicaciones.

Resumen de la invención

La presente invención proporciona un método para determinar la fracción de ácidos nucleicos fetales libres de células en una muestra de sangre periférica de una mujer embarazada, que comprende:

35 (i) realizar la secuenciación sobre ácidos nucleicos libres de células contenidos en la muestra para obtener un resultado de la secuenciación que consiste en una pluralidad de datos de secuenciación;

(ii) determinar el número de ácidos nucleicos libres de células de una longitud que cae en un rango predeterminado en la muestra en base al resultado de la secuenciación; y

40 (iii) determinar la fracción de ácidos nucleicos fetales libres de células en la muestra en base al número de ácidos nucleicos libres de células de una longitud que cae en el rango predeterminado, en el que se determina dicho rango predeterminado antes de la etapa (i) mediante las siguientes etapas:

(a) determinar las longitudes de los ácidos nucleicos libres de células en una pluralidad de muestras de sangre periférica de control de mujeres embarazadas en cada una de las cuales se conoce la fracción de ácidos nucleicos fetales libres de células;

45 (b) establecer una pluralidad de rangos de longitud candidatos y determinar el porcentaje de ácidos nucleicos libres de células obtenidos de cada una de la pluralidad de muestras de control presentes en cada rango de longitud candidato;

50 (c) determinar un coeficiente de correlación para cada rango de longitud candidato en base al porcentaje de ácidos nucleicos libres de células obtenidos de cada una de la pluralidad de muestras de control presentes en cada rango de longitud candidato y la fracción conocida de ácidos nucleicos fetales libres de células en cada una de las muestras de control; y

(d) determinar al menos un rango de longitud candidato o una combinación de los rangos de longitud candidatos como el rango predeterminado en base al coeficiente de correlación máximo.

En las realizaciones de la presente divulgación, los ácidos nucleicos libres de células en la muestra biológica se secuencian mediante secuenciación de extremos emparejados, secuenciación de un solo extremo o secuenciación de una sola molécula. Por lo tanto, se pueden obtener fácilmente longitudes de los ácidos nucleicos libres de células, lo que conduce a las etapas posteriores.

5 En las realizaciones de la presente divulgación, los ácidos nucleicos libres de células son ADN.

10 En las realizaciones de la presente divulgación, determinar el número de ácidos nucleicos libres de células de una longitud que cae en el rango predeterminado en la muestra biológica en base al resultado de la secuenciación incluye adicionalmente: alinear el resultado de la secuenciación con un genoma de referencia, para construir un conjunto de datos que consiste en una pluralidad de lecturas mapeadas de forma única, donde cada lectura en el conjunto de datos se puede mapear a una única posición del genoma de referencia; determinar la longitud del ácido nucleico libre de células que corresponde a cada lectura mapeada de forma única en el conjunto de datos; y determinar el número de ácidos nucleicos libres de células de una longitud que cae en el rango predeterminado. Por lo tanto, el número de ácidos nucleicos libres de células de una longitud que cae en el rango predeterminado en la muestra biológica se puede determinar fácilmente, lo que da lugar a un resultado preciso y confiable y una buena reproducibilidad.

15 En las realizaciones de la presente divulgación, determinar la longitud del ácido nucleico libre de células que corresponde a cada lectura mapeada de forma única en el conjunto de datos incluye adicionalmente: determinar la longitud de cada lectura mapeada de forma única en el genoma de referencia como la longitud del ácido nucleico libre de células que corresponde a la lectura. Por lo tanto, la longitud del ácido nucleico libre de células que corresponde a cada lectura alineada de forma única en el conjunto de datos se puede determinar con precisión.

20 En las realizaciones de la presente divulgación, en el caso de que los ácidos nucleicos libres de células en la muestra biológica se secuencien mediante secuenciación de extremos emparejados, la determinación de la longitud del ácido nucleico libre de células que corresponde a cada lectura mapeada de forma única en el conjunto de datos incluye: determinar la posición que corresponde al genoma de referencia del extremo 5' del ácido nucleico libre de células en base a los datos de secuenciación para el extremo 5' de cada lectura mapeada de forma única obtenida en la secuenciación de extremos emparejados; determinar la posición que corresponde al genoma de referencia del extremo 3' del ácido nucleico libre de células en base a los datos de secuenciación en el otro extremo de la misma lectura mapeada de forma única obtenida en la secuenciación de extremos emparejados; y determinar la longitud del ácido nucleico libre de células en base a la posición del extremo 5' del ácido nucleico libre de células y la posición del extremo 3' del ácido nucleico libre de células. Por lo tanto, la longitud del ácido nucleico libre de células que corresponde a cada lectura mapeada de forma única en el conjunto de datos se puede determinar con precisión.

25 El rango predeterminado se determina en base a una pluralidad de muestras de control, en cada una de las cuales se conoce la fracción de ácidos nucleicos fetales libres de células. Por lo tanto, el rango predeterminado se puede determinar como un resultado preciso y confiable.

35 En las realizaciones de la presente divulgación, el rango predeterminado se determina en base a al menos 20 muestras de control.

En las realizaciones de la presente divulgación, el rango de longitud candidato es de un tramo de 5 pb a 20 pb.

En las realizaciones de la presente divulgación, el rango de longitud candidato es de un tramo de 1 pb a 20 pb.

En las realizaciones de la presente divulgación, la pluralidad de rangos de longitud candidatos tiene un tamaño de etapa de 1 pb a 2 pb.

40 En las realizaciones de la presente divulgación, determinar la fracción de ácidos nucleicos fetales libres de células en la muestra en base al número de ácidos nucleicos libres de células de una longitud que cae en el rango predeterminado incluye adicionalmente:

determinar el porcentaje de ácidos nucleicos libres de células presentes en el rango predeterminado en base al número de ácidos nucleicos libres de células de una longitud que cae en el rango predeterminado; y

45 determinar la fracción de ácidos nucleicos fetales libres de células en la muestra, en base al porcentaje de ácidos nucleicos libres de células presentes en el rango predeterminado, de acuerdo con una función predeterminada, en la que la función predeterminada se determina en base a la pluralidad de muestras de control. Por lo tanto, la fracción de ácidos nucleicos fetales libres de células en la muestra se puede determinar de manera eficiente, lo que da lugar a un resultado preciso y confiable y una buena reproducibilidad.

50 En las realizaciones de la presente divulgación, la función predeterminada se obtiene mediante las siguientes etapas:

(i) determinar el porcentaje de ácidos nucleicos libres de células obtenidos de cada muestra de control presente en el rango predeterminado; y

(ii) ajustar el porcentaje de ácidos nucleicos libres de células obtenido de cada muestra de control presente en el rango predeterminado con la fracción conocida de ácidos nucleicos fetales libres de células para determinar la función

predeterminada. Por lo tanto, la función predeterminada se puede determinar con precisión y eficacia, lo que conduce a las etapas posteriores.

5 En las realizaciones de la presente divulgación, el porcentaje de ácidos nucleicos libres de células obtenido de cada muestra de control presente en el rango predeterminado se ajusta con la fracción conocida de ácidos nucleicos fetales libres de células mediante un ajuste lineal.

A modo de ejemplo, el rango predeterminado es de 185 pb a 204 pb. A modo de ejemplo, como se ilustra en la ejemplificación, la función predeterminada puede ser $d = 0.0334 * p + 1.6657$, donde d representa la fracción de ácidos nucleicos fetales libres de células, y p representa el porcentaje de ácido nucleico libre de células presente en el rango predeterminado.

10 En las realizaciones de la presente divulgación, la muestra de control es una muestra de sangre periférica obtenida de una mujer embarazada con un feto masculino normal, en la que se sabe que la fracción de ácidos nucleicos fetales libres de células se determina por el cromosoma Y. Por lo tanto, se determina el rango predeterminado con precisión.

15 En las realizaciones de la presente divulgación, la fracción de ácidos nucleicos fetales libres de células en la muestra de control es una fracción de ADN fetal libre de células que se estima por el cromosoma Y. Por lo tanto, el rango predeterminado se puede determinar al utilizar eficientemente la fracción de ácidos nucleicos fetales libres de células de la muestra de control, y luego el número de ácidos nucleicos libres de células de una longitud que cae en el rango predeterminado y se puede determinar adicionalmente la fracción de ADN fetal libre de células en una muestra obtenida de una mujer embarazada bajo detección.

20 Cabe señalar que, el método para determinar la fracción de ácidos nucleicos fetales libres de células en una muestra de sangre materna de acuerdo con la presente divulgación tiene al menos las siguientes ventajas:

1) Universalidad: se pueden estimar las fracciones de ADN fetal libre de células (en particular feto femenino) en todas las muestras que cumplen con el control de calidad.

2) Se puede mejorar la precisión de la detección de NIPT.

25 3) Simplicidad operativa: las fracciones de ADN fetal libres de células se pueden estimar directamente simplemente utilizando datos fuera de línea.

30 En un aspecto adicional, la presente invención proporciona un método para determinar el sexo de gemelos. Dicho método incluye: realizar la secuenciación sobre ácidos nucleicos libres de células contenidos en una muestra de sangre periférica obtenida de una mujer embarazada con gemelos, para obtener un resultado de la secuenciación que consiste en una pluralidad de datos de secuenciación; determinar una primera fracción de ADN fetal libre de células en base a los datos de secuenciación mediante el método de la invención descrito anteriormente; determinar una segunda fracción de ADN fetal libre de células en base a los datos de secuenciación derivados del cromosoma Y en el resultado de la secuenciación; y determinar el sexo de los gemelos en base a la primera fracción de ADN fetal libre de células y la segunda fracción de ADN fetal libre de células. Los inventores han encontrado sorprendentemente que el sexo de los gemelos en una mujer embarazada se puede determinar de forma precisa y eficaz mediante este método.

35 En las realizaciones de la presente divulgación, la segunda fracción de ADN fetal libre de células se determina de acuerdo con la siguiente fórmula:

$$fra.cry = (cry.ER\% - Mujer.cry.ER\%) / (Hombre.cry.ER\% - Mujer.cry.ER\%) * 100\%$$

40 donde *fra.cry* representa la segunda fracción de ADN fetal libre de células, *cry.ER%* representa un porcentaje de los datos de secuenciación derivados del cromosoma Y en el resultado de la secuenciación de los datos de secuenciación totales; *Mujer.cry.ER%* representa un porcentaje promedio de los datos de secuenciación de ácidos nucleicos libres de células alineados con el cromosoma Y en una muestra de sangre periférica obtenida de una mujer embarazada predeterminada para estar con un feto femenino normal para los datos de secuenciación totales de los mismos; y *Hombre.cry.ER%* representa un porcentaje promedio de los datos de secuenciación de ácidos nucleicos libres de células derivados del cromosoma Y en una muestra de sangre periférica obtenida de un hombre sano para los datos de secuenciación totales de los mismos.

Por lo tanto, la segunda fracción de ADN fetal libre de células se puede determinar con precisión.

50 En las realizaciones de la presente divulgación, la determinación del sexo de los gemelos en base a la primera fracción de ADN fetal libre de células y la segunda fracción de ADN fetal libre de células incluye adicionalmente: (a) determinar una relación de la segunda fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células; y (b) determinar el sexo de los gemelos al comparar la relación determinada en (a) con un primer umbral predeterminado y un segundo umbral predeterminado, en el que el primer umbral se ha determinado en base a una pluralidad de muestras de control obtenidas de mujeres embarazadas conocidas por tener gemelos femeninos y el

segundo umbral se ha determinado en base a una pluralidad de muestras de control obtenidas de mujeres embarazadas que se sabe que tienen gemelos varones.

5 Ambos fetos de los gemelos se determinan como femeninos si la relación de la segunda fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células es menor que el primer umbral, ambos fetos de los gemelos se determinan como masculinos si la relación de la segunda fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células es mayor que el segundo umbral, y se determina que los gemelos incluyen un feto masculino y un feto femenino si la relación de la segunda fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células es igual al primer umbral o al segundo umbral, o entre el primer umbral y el segundo umbral.

10 En las realizaciones de la presente divulgación, el primer umbral es 0.35 y el segundo umbral es 0.7.

En otro aspecto, la presente invención proporciona un método para detectar una aneuploidía cromosómica de gemelos. Dicho método incluye:

15 realizar la secuenciación sobre ácidos nucleicos libres de células contenidos en una muestra de sangre periférica obtenida de una mujer embarazada con gemelos para obtener un resultado de la secuenciación que consiste en una pluralidad de datos de secuenciación;

determinar una primera fracción de ADN fetal libre de células en base a los datos de secuenciación mediante el método de la invención descrito anteriormente;

determinar una tercera fracción de ADN fetal libre de células en base a los datos de secuenciación derivados de un cromosoma predeterminado en el resultado de la secuenciación; y

20 determinar si los gemelos bajo detección tienen aneuploidía con respecto al cromosoma predeterminado en base a la primera fracción de ADN fetal libre de células y la tercera fracción de ADN fetal libre de células.

Por lo tanto, la aneuploidía cromosómica de los gemelos se puede detectar con precisión y eficacia.

En las realizaciones de la presente divulgación, la tercera fracción de ADN fetal libre de células se determina de acuerdo con la siguiente fórmula:

25 $fra.cri = 2*(cri.ER\%/ajuste.cri.ER\%-1)*100\%$, donde *fra.cri* representa la tercera fracción de ADN fetal libre de células, *i* representa un número de serie del cromosoma predeterminado e *i* es cualquier entero en el rango de 1 a 22; *cri.ER%* representa un porcentaje de los datos de secuenciación derivados del cromosoma predeterminado en el resultado de la secuenciación para los datos de secuenciación totales; *ajuste.cri.ER%* representa un porcentaje promedio de los datos de secuenciación de ácidos nucleicos libres de células derivados del cromosoma predeterminado en una muestra de sangre periférica obtenida de una mujer embarazada predeterminada para estar con gemelos normales para los datos de secuenciación totales de los mismos.

30

Por lo tanto, la tercera fracción de ADN fetal libre de células se puede determinar con precisión.

35 En las realizaciones de la presente divulgación, determinar si los gemelos bajo detección tienen aneuploidía con respecto al cromosoma predeterminado en base a la primera fracción de ADN fetal libre de células y la tercera fracción de ADN fetal libre de células incluye adicionalmente: (a) determinar una relación de la tercera fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células; y (b) determinar si los gemelos bajo detección tienen aneuploidía con respecto al cromosoma predeterminado al comparar la relación determinada en (a) con un tercer umbral predeterminado y un cuarto umbral predeterminado y en el que el tercer umbral se ha determinado en base a una pluralidad de muestras de control obtenidas de mujeres embarazadas con gemelos que se sabe que no tienen aneuploidía con respecto al cromosoma predeterminado, y el cuarto umbral se ha determinado con base en una pluralidad de muestras de control obtenidas de mujeres embarazadas con gemelos que se sabe que tienen aneuploidía con respecto al cromosoma predeterminado.

40 Se determina que ambos fetos de los gemelos no tienen aneuploidía con respecto al cromosoma predeterminado si la relación de la tercera fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células es menor que el tercer umbral, se determina que ambos fetos de los gemelos tienen aneuploidía con respecto al cromosoma predeterminado si la relación de la tercera fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células es mayor que el cuarto umbral, y se determina que un feto de los gemelos tiene aneuploidía con respecto al cromosoma predeterminado, mientras que se determina que el otro feto de los gemelos no tiene aneuploidía con respecto al cromosoma predeterminado si la relación de la tercera fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células es igual al tercer umbral o el cuarto umbral, o entre el tercer umbral y el cuarto umbral.

45 En las realizaciones de la presente divulgación, el tercer umbral es 0.35 y el cuarto umbral es 0.7.

50 En las realizaciones de la presente divulgación, el cromosoma predeterminado es al menos uno seleccionado de los cromosomas 18, 21 y 23.

En todavía un aspecto adicional, la presente invención proporciona un método para determinar una aneuploidía cromosómica de gemelos que incluye:

5 realizar la secuenciación sobre ácidos nucleicos libres de células contenidos en una muestra de sangre periférica obtenida de una mujer embarazada con gemelos, para obtener un resultado de la secuenciación que consiste en una pluralidad de datos de secuenciación y determinar a partir de ellos la fracción de ácidos nucleicos fetales libres de células como una primera fracción de ADN fetal libre de células mediante el método de la invención como se describió anteriormente;

10 determinar una fracción x_i del número de datos de secuenciación derivados del cromosoma i en el resultado de la secuenciación para los datos de secuenciación totales, donde i representa un número de serie del cromosoma, e i es cualquier entero en el rango de 1 a 22;

15 determinar una puntuación T del cromosoma i de acuerdo con $T_i = (X_i - \mu_i) / \sigma_i$, donde i representa el número de serie del cromosoma e i es cualquier entero en el rango de 1 a 22, μ_i representa un porcentaje promedio de los datos de secuenciación del cromosoma i seleccionado como un sistema de referencia en una base de datos de referencia para los datos de secuenciación totales del mismo, σ_i representa una desviación estándar de los porcentajes de los datos de secuenciación del cromosoma i seleccionado como un sistema de referencia en la base de datos de referencia para los datos de secuenciación totales del mismo,

20 determinar una puntuación L del cromosoma i de acuerdo con $L_i = \log(d(T_i, a)) / \log(d(T2_i, a))$, donde i representa el número de serie del cromosoma e i es cualquier entero en el rango de 1 a 22, $T2_i = (X_i - \mu_i * (1 - fra/2)) / \sigma_i$; $d(T_i, a)$ y $d(T2_i, a)$ representan la función de densidad de probabilidad de distribución t , donde a representa el grado de libertad y fra representa dicha primera fracción de ADN fetal libre de células;

25 graficar un diagrama de cuatro cuadrantes con T como coordenada vertical y L como coordenada horizontal mediante la zonificación con una primera línea recta donde $T =$ quinto umbral predeterminado y una segunda línea recta donde $L =$ sexto umbral predeterminado, en el que se determina que ambos fetos de los gemelos tienen un trisoma si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que cae en un primer cuadrante; se determina que un feto de los gemelos tiene un trisoma y se determina que el otro feto de los gemelos es normal si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que cae en un segundo cuadrante; se determina que ambos fetos de los gemelos son normales si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que cae en un tercer cuadrante y se determina que los gemelos han dado lugar a una fracción fetal baja si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que cae en un cuarto cuadrante, de tal manera que no se adopta el resultado. Los inventores han encontrado sorprendentemente que la determinación de si los gemelos tienen aneuploidía con respecto a un cromosoma predeterminado se puede lograr de forma precisa y eficaz mediante este método.

En otro aspecto, la presente invención proporciona un método para detectar quimera fetal que incluye:

35 realizar la secuenciación sobre ácidos nucleicos libres de células contenidos en una muestra de sangre periférica obtenida de una mujer embarazada con un feto, opcionalmente un feto masculino, para obtener un resultado de la secuenciación que consiste en una pluralidad de datos de secuenciación;

determinar una primera fracción de ADN fetal libre de células en base a los datos de secuenciación mediante el método de la invención descrito anteriormente,

40 determinar una tercera fracción de ADN fetal libre de células en base a los datos de secuenciación derivados de un cromosoma predeterminado en el resultado de la secuenciación; y

determinar si el feto bajo detección tiene quimera fetal con respecto al cromosoma predeterminado en base a la primera fracción de ADN fetal libre de células y la tercera fracción de ADN fetal libre de células.

Por lo tanto, se puede analizar con precisión si el feto bajo detección tiene quimera fetal con respecto a un cromosoma específico.

45 En las realizaciones de la presente divulgación, la tercera fracción de ADN fetal libre de células se determina mediante la siguiente fórmula:

50 $fra.cri = 2 * (cri.ER\% / ajuste.cri.ER\% - 1) * 100\%$, donde $fra.cri$ representa la tercera fracción de ADN fetal libre de células, i representa un número de serie del cromosoma predeterminado e i es cualquier entero en el rango de 1 a 22; $cri.ER\%$ representa un porcentaje de los datos de secuenciación derivados del cromosoma predeterminado en el resultado de la secuenciación de los datos de secuenciación totales; $ajuste.cri.ER\%$ representa un porcentaje promedio de los datos de secuenciación de ácidos nucleicos libres de células derivados del cromosoma predeterminado en una muestra de sangre periférica obtenida de una mujer embarazada predeterminada para estar con un feto para los datos de secuenciación totales del mismo. Por lo tanto, si el feto bajo detección tiene quimera fetal con respecto al cromosoma específico se puede analizar con una eficacia mejorada adicional.

5 En las realizaciones de la presente divulgación, determinar si el feto bajo detección tiene quimera fetal con respecto al cromosoma predeterminado en base a la primera fracción de ADN fetal libre de células y la tercera fracción de ADN fetal libre de células incluye adicionalmente: (a) determinar una relación de la tercera fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células; y (b) determinar si el feto bajo detección tiene quimera con respecto al cromosoma predeterminado al comparar la relación determinada en (a) con una pluralidad de umbrales predeterminados. Por lo tanto, si el feto bajo detección tiene quimera fetal con respecto al cromosoma específico se puede analizar con una eficacia mejorada adicional.

En las realizaciones de la presente divulgación, la pluralidad de umbrales predeterminados incluye al menos uno seleccionado de:

10 un séptimo umbral, determinado en base a una pluralidad de muestras de control con el cromosoma predeterminado que se sabe es un monosoma completo,

un octavo umbral, determinado en base a una pluralidad de muestras de control con el cromosoma predeterminado que se sabe es una quimera monosoma,

15 un noveno umbral, determinado en base a una pluralidad de muestras de control con el cromosoma predeterminado que se sabe que es normal,

un décimo umbral, determinado en base a una pluralidad de muestras de control con el cromosoma predeterminado que se sabe es un trisoma completo.

20 Se determina que el cromosoma predeterminado del feto bajo detección es un monosoma completo, si la relación de la tercera fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células es menor que el séptimo umbral;

se determina que el cromosoma predeterminado del feto bajo detección es una quimera monosoma, si la relación de la tercera fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células no es menor que el séptimo umbral ni mayor que el octavo umbral;

25 se determina que el cromosoma predeterminado del feto bajo detección es normal, si la relación de la tercera fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células es mayor que el octavo umbral y menor que el noveno umbral;

se determina que el cromosoma predeterminado del feto bajo detección es una quimera trisoma, si la relación de la tercera fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células no es menor que el noveno umbral ni superior al décimo umbral; y

30 se determina que el cromosoma predeterminado del feto bajo detección es un trisoma completo, si la relación de la tercera fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células es mayor que el décimo umbral.

En las realizaciones de la presente divulgación, el séptimo umbral es mayor que -1 y menor que 0, por ejemplo -0.85;

el octavo umbral es mayor que el séptimo umbral y menor que 0, por ejemplo, -0.3;

35 el noveno umbral es mayor que 0 y menor que 1, por ejemplo 0.3; y

el décimo umbral es mayor que el noveno umbral y menor que 1, por ejemplo 0.85.

Por lo tanto, si el feto bajo detección tiene quimera fetal con respecto a un cromosoma específico, se puede analizar con una eficacia mejorada adicional.

En un aspecto adicional, la presente invención proporciona un método para detectar quimera fetal que incluye:

40 realizar la secuenciación sobre los ácidos nucleicos libres de células contenidos en una muestra de sangre periférica obtenida de una mujer embarazada con un feto para obtener un resultado de la secuenciación que consiste en una pluralidad de datos de secuenciación y determinar a partir de ellos la fracción de ácidos nucleicos fetales libres de células como una primera fracción de ADN fetal libre de células mediante el método de la invención como se describió anteriormente;

45 determinar una fracción x_i del número de datos de secuenciación derivados del cromosoma i en el resultado de la secuenciación para los datos de secuenciación totales, donde i representa un número de serie del cromosoma, e i es cualquier entero en el rango de 1 a 22;

50 determinar una puntuación T del cromosoma i de acuerdo con $T_i(X_i - \mu_i) / \sigma_i$, donde i representa el número de serie del cromosoma e i es cualquier entero en el rango de 1 a 22, μ_i representa un valor promedio de porcentajes de datos de secuenciación del cromosoma i seleccionado como un sistema de referencia en una base de datos de referencia para

los datos de secuenciación totales del mismo, σ_i representa una desviación estándar de los porcentajes de los datos de secuenciación del cromosoma i seleccionado como un sistema de referencia en la base de datos de referencia para los datos de secuenciación totales del mismo;

5 determinar una puntuación L del cromosoma i de acuerdo con $L_i = \log(d(T_i, a)) / \log(d(T2_i, a))$, donde i representa el número de serie del cromosoma e i es cualquier entero en el rango de 1 a 22, $T2_i = (X_i - \mu_i * (1/\text{fra}/2)) / \sigma_i$; $d(T_i, a)$ y $d(T2_i, a)$ representan la función de densidad de probabilidad de distribución t , a representa el grado de libertad y fra representa dicha primera fracción de ADN fetal libre de células;

10 graficar un diagrama de cuatro cuadrantes con T como coordenada vertical y L como coordenada horizontal mediante zonificación con una tercera línea recta donde $T =$ undécimo umbral predeterminado y una cuarta línea recta donde $L =$ duodécimo umbral predeterminado, cuando la puntuación T no es mayor que 0,

en el que se determina que el feto tiene un monosoma completo o quimera monosoma con respecto al cromosoma predeterminado, si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que cae en un primer cuadrante;

15 se determina que el feto tiene una quimera monosoma con respecto al cromosoma predeterminado, si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que cae en un segundo cuadrante;

se determina que el feto es normal con respecto al cromosoma predeterminado, si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que cae en un tercer cuadrante y

se determina que el feto ha dado lugar a una fracción fetal baja si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que cae en un cuarto cuadrante, de tal manera que no se adopta el resultado,

20 graficar un diagrama de cuatro cuadrantes con T como coordenada vertical y L como coordenada horizontal mediante la zonificación con una quinta línea recta donde $T =$ decimotercero umbral predeterminado y una sexta línea recta donde $L =$ decimocuarto umbral predeterminado, cuando la puntuación T es mayor que 0,

25 en el que se determina que el feto tiene un trisoma o quimera trisoma completa con respecto al cromosoma predeterminado, si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que cae en un primer cuadrante;

se determina que el feto tiene una quimera trisoma con respecto al cromosoma predeterminado, si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que cae en un segundo cuadrante;

se determina que el feto es normal con respecto al cromosoma predeterminado, si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que cae en un tercer cuadrante;

30 se determina que el feto ha dado lugar a una fracción fetal baja si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que cae en un cuarto cuadrante, de tal manera que no se adopta el resultado,

opcionalmente, el undécimo umbral y el decimotercero umbral son cada uno independientemente 3, y el duodécimo umbral y el decimocuarto umbral son cada uno independientemente 1.

Por lo tanto, las situaciones de quimera fetal se pueden analizar de manera eficiente.

35 Los aspectos y ventajas adicionales de las realizaciones de la presente divulgación se darán en parte en las siguientes descripciones, se harán evidentes en parte a partir de las siguientes descripciones, o se aprenderán de la práctica de las realizaciones de la presente divulgación.

Breve descripción de los dibujos

40 Estos y otros aspectos y ventajas de las realizaciones de la presente divulgación resultarán evidentes y se apreciarán más fácilmente a partir de las siguientes descripciones realizadas con referencia a los dibujos, en los que:

La Fig. 1 es un diagrama de flujo que muestra un método para determinar una fracción de ácidos nucleicos libres de células en una muestra biológica de acuerdo con una realización de la presente divulgación.

45 La Fig. 2 es un diagrama de flujo que muestra un método para determinar el número de ácidos nucleicos libres de células en una longitud que cae en un rango predeterminado de acuerdo con una realización de la presente divulgación.

La Fig. 3 es un diagrama de flujo que muestra un método para determinar la longitud del ácido nucleico libre de células de acuerdo con una realización de la presente divulgación.

La Fig. 4 es un diagrama de flujo que muestra un método para determinar un rango predeterminado de acuerdo con una realización de la presente divulgación.

- La Fig. 5 es un diagrama de flujo que muestra un método para determinar una fracción de ácidos nucleicos libres de células de una fuente predeterminada en una muestra biológica de acuerdo con una realización de la presente divulgación.
- 5 La Fig. 6 es un diagrama de flujo que muestra un método para determinar una función predeterminada de acuerdo con una realización de la presente divulgación.
- La Fig. 7 es un diagrama estructural de un dispositivo para determinar una fracción de ácidos nucleicos libres de células de una fuente predeterminada en una muestra biológica mediante un método de acuerdo con una realización de la presente divulgación.
- 10 La Fig. 8 es un diagrama estructural de un aparato de recuento para uso en llevar a cabo un método de acuerdo con una realización de la presente divulgación.
- La Fig. 9 es un diagrama estructural de una primera unidad de determinación de longitud para uso en llevar a cabo un método de acuerdo con una realización de la presente divulgación.
- La Fig. 10 es un diagrama estructural de un aparato de determinación de rango predeterminado para uso en llevar a cabo un método de acuerdo con una realización de la presente divulgación.
- 15 La Fig. 11 es un diagrama estructural de un aparato para determinar una fracción de ácidos nucleicos libres de células para uso en llevar a cabo un método de acuerdo con una realización de la presente divulgación.
- La Fig. 12 es un diagrama estructural de un aparato de determinación de función predeterminada para uso en llevar a cabo un método de acuerdo con una realización de la presente divulgación.
- 20 La Fig. 13 es un diagrama de ajuste lineal del coeficiente de correlación entre la fracción de ADN fetal libre de células estimada por el cromosoma Y y el porcentaje de moléculas de ADN presentes en el rango de longitud de 185 pb ~ 204 pb para cada muestra obtenida de 37 mujeres embarazadas que se sabe tienen un feto masculino normal, de acuerdo con una realización de la presente invención; y
- Las Fig. 14 al final son diagramas de cuatro cuadrantes para puntuaciones T y puntuaciones L de 11 muestras bajo detección de acuerdo con una realización de la presente invención.
- 25 Descripción detallada
- Las realizaciones de la presente divulgación se describirán en detalle a continuación, que son explicativas, ilustrativas y se utilizan para explicar en general la presente divulgación, por lo que no se interpretará que limitan la presente divulgación.
- 30 Método para determinar una fracción de ácidos nucleicos fetales libres de células en una muestra de sangre periférica de una mujer embarazada
- Como se indicó anteriormente en el presente documento, los inventores han encontrado sorprendentemente que la fracción de ácidos nucleicos fetales en una muestra de sangre periférica obtenida de una mujer embarazada se puede determinar de forma precisa y eficaz mediante el método de la presente divulgación.
- 35 La fracción de ácidos nucleicos fetales libres de células, significa la fracción del número de ácidos nucleicos fetales libres de células para el número total de ácidos nucleicos libres de células en la sangre periférica obtenida de la mujer embarazada, que a veces también se conoce como "fracción de ADN fetal libre de células en la sangre periférica obtenida de la mujer embarazada" o "fracción de ADN fetal libre de células".
- Con referencia a la Fig. 1, el método incluye las siguientes etapas.
- S100: secuenciación de ácido nucleico
- 40 Los ácidos nucleicos libres de células en la muestra biológica se secuencian para obtener un resultado de la secuenciación que consiste en una pluralidad de datos de secuenciación. La muestra biológica es una muestra de sangre periférica obtenida de una mujer embarazada. El ácido nucleico libre de células de la fuente predeterminada son ácidos nucleicos fetales libres de células en la muestra de sangre periférica. Cabe señalar que el término "datos de secuenciación" utilizado en el presente documento se refiere a "lecturas de secuencia", que corresponden a ácidos nucleicos sometidos a secuenciación. El resultado de la secuenciación incluye longitudes de los ácidos nucleicos libres de células.
- 45
- En las realizaciones de la presente divulgación, los ácidos nucleicos libres de células en la muestra biológica se secuencian mediante secuenciación de extremos emparejados, secuenciación de un solo extremo o secuenciación de una sola molécula. Por lo tanto, se pueden obtener fácilmente longitudes de los ácidos nucleicos libres de células, lo que conduce a las etapas posteriores.
- 50

S200: determinar el número de ácidos nucleicos libres de células en una longitud que cae en un rango predeterminado

El número de ácidos nucleicos libres de células en la longitud que cae en el rango predeterminado en la muestra biológica se determina en base al resultado de la secuenciación.

5 Cabe señalar que el término “longitud” utilizado en el presente documento se refiere a la “longitud del ácido nucleico (leído)” en pares de bases (pb).

En las realizaciones de la presente divulgación, con referencia a la Fig.2, S200 incluye adicionalmente las siguientes etapas:

10 S210: el resultado de la secuenciación se alinea con un genoma de referencia. Específicamente, el resultado de la secuenciación se alinea con el genoma de referencia, para construir un conjunto de datos que consiste en una pluralidad de lecturas mapeadas de forma única, donde cada lectura en el conjunto de datos se puede mapear a una única posición del genoma de referencia; preferiblemente, no hay lecturas mapeadas de forma incorrecta o como máximo una lectura mapeada de forma incorrecta o como máximo dos lecturas mal alineadas.

S220: se determina una longitud del ácido nucleico libre de células. Específicamente, se determina una longitud del ácido nucleico libre de células que corresponde a cada lectura mapeada de forma única en el conjunto de datos.

15 S230: se determina el número de ácidos nucleicos libres de células que caen en el rango predeterminado. Específicamente, se determina el número de ácidos nucleicos libres de células de una longitud que cae en el rango predeterminado.

20 Por lo tanto, el número de ácidos nucleicos libres de células de una longitud que cae en el rango predeterminado en la muestra biológica se puede determinar fácilmente, lo que da lugar a un resultado preciso y confiable y una buena reproducibilidad.

En las realizaciones de la presente divulgación, en S220, la longitud de cada lectura mapeada de forma única en el genoma de referencia se determina como la longitud del ácido nucleico libre de células que corresponde a la lectura. Por lo tanto, la longitud del ácido nucleico libre de células que corresponde a cada lectura mapeada de forma única en el conjunto de datos se puede determinar con precisión.

25 En las realizaciones de la presente divulgación, en el caso de que los ácidos nucleicos libres de células en la muestra biológica se secuencien mediante secuenciación de extremos emparejados, con referencia a la Fig. 3, S220 incluye las siguientes etapas.

30 S2210: se determina una posición, que corresponde al genoma de referencia, del extremo 5' del ácido nucleico libre de células. Específicamente, se determina la posición, que corresponde al genoma de referencia, del extremo 5' del ácido nucleico libre de células en base a los datos de secuenciación para el extremo 5' de cada lectura mapeada de forma única obtenida en la secuenciación de extremos emparejados.

35 S2220: se determina una posición, que corresponde al genoma de referencia, del extremo 3' del ácido nucleico libre de células. Específicamente, se determina la posición, que corresponde al genoma de referencia, del extremo 3' del ácido nucleico libre de células en base a los datos de secuenciación en el otro extremo de la misma lectura mapeada de forma única obtenida en la secuenciación de extremos emparejados.

S2230: se determina la longitud del ácido nucleico libre de células. Específicamente, la longitud del ácido nucleico libre de células se determina en base a la posición del extremo 5' del ácido nucleico libre de células y la posición del extremo 3' del ácido nucleico libre de células.

40 Por lo tanto, la longitud del ácido nucleico libre de células que corresponde a cada lectura alineada de forma única en el conjunto de datos se puede determinar con precisión.

S300: determinar la fracción de ácidos nucleicos fetales libres de células

Se determina la fracción de ácidos nucleicos fetales libres de células en la muestra en base al número de ácidos nucleicos libres de células de una longitud que cae en el rango predeterminado.

45 Adicionalmente, el método de acuerdo con la presente invención incluye adicionalmente determinar el rango predeterminado (S400, no mostrado en las Figuras). El rango predeterminado se determina en base a una pluralidad de muestras de control, en cada una de las cuales se conoce la fracción de ácidos nucleicos fetales libres de células. Por lo tanto, el rango predeterminado se puede determinar como un resultado preciso y confiable. En las realizaciones de la presente divulgación, el rango predeterminado se determina en base a al menos 20 muestras de control.

En las realizaciones de la presente divulgación, con referencia a la Fig. 4, S400 incluye las siguientes etapas.

50 S410: determinar las longitudes de los ácidos nucleicos libres de células en la pluralidad de muestras de control.

S420: determinar el porcentaje de ácidos nucleicos libres de células presentes en cada rango de longitud candidato. Específicamente, se establecen una pluralidad de rangos de longitud de candidatos, y se determina el porcentaje de ácidos nucleicos libres de células, obtenidos de cada una de la pluralidad de muestras de control, presentes en cada rango de longitud candidato.

5 S430: se determina un coeficiente de correlación. Específicamente, se determina un coeficiente de correlación para cada rango de longitud candidato, en base al porcentaje de ácidos nucleicos libres de células obtenidos de cada una de la pluralidad de muestras de control presentes en cada rango de longitud candidato, y la fracción conocida de ácidos nucleicos fetales libres de células en cada una de las muestras de control; y

10 S440: se determina el rango predeterminado. Específicamente, un rango de longitud candidato con el mayor coeficiente de correlación se determina como el rango predeterminado.

Por lo tanto, el rango predeterminado se puede determinar de manera precisa y eficiente.

En las realizaciones de la presente divulgación, el rango de longitud candidato es de un tramo de 1 pb a 20 pb.

En las realizaciones de la presente divulgación, la pluralidad de rangos de longitud candidatos tiene un tamaño de etapa de 1 pb a 2 pb.

15 En las realizaciones de la presente divulgación, con referencia a la Fig. 5, S300 incluye adicionalmente las siguientes etapas.

S310: se determina un porcentaje de los ácidos nucleicos libres de células presentes en el rango predeterminado. Específicamente, el porcentaje de ácidos nucleicos libres de células presentes en el rango predeterminado se determina en base al número de ácidos nucleicos libres de células en la longitud que cae en el rango predeterminado.

20 S320: se determina la fracción de ácidos nucleicos fetales libres de células en la muestra. Específicamente, la fracción de ácidos nucleicos fetales libres de células en la muestra se determina en base al porcentaje de ácidos nucleicos libres de células presentes en el rango predeterminado, de acuerdo con una función predeterminada, en la que la función predeterminada se determina en base a la pluralidad de muestras de control.

25 Por lo tanto, la fracción de ácidos nucleicos fetales libres de células en la muestra se puede determinar de manera eficiente, lo que da lugar a un resultado preciso y confiable y una buena reproducibilidad.

En las realizaciones de la presente divulgación, el método incluye adicionalmente determinar la función predeterminada (S500, no mostrada en las Figuras).

En algunas realizaciones específicas de la presente divulgación, con referencia a la Fig. 6, S500 incluye las siguientes etapas.

30 S510: se determina el porcentaje de ácidos nucleicos libres de células obtenido de cada muestra de control presente en el rango predeterminado.

S520: el porcentaje de ácidos nucleicos libres de células obtenido de cada muestra de control presente en el rango predeterminado se ajusta con la fracción conocida del ácido nucleico fetal libre de células para determinar la función predeterminada.

35 Por lo tanto, la función predeterminada se puede determinar con precisión y eficacia, lo que conduce a las etapas posteriores.

En las realizaciones de la presente divulgación, el porcentaje de ácidos nucleicos libres de células, obtenido de cada muestra de control, presente en el rango predeterminado se ajusta con la fracción conocida del ácido nucleico fetal libre de células mediante un ajuste lineal.

40 Como se indicó anteriormente, a modo de ejemplo, el rango predeterminado es de 185 pb a 204 pb. Como se ilustra mediante la ejemplificación, la función predeterminada puede ser $d = 0.0334 * p + 1.6657$, donde d representa la fracción de ácidos nucleicos fetales libres de células y p representa un porcentaje de ácido nucleico libre de células presente en el rango predeterminado. La fracción de ácidos nucleicos fetales libres de células en la muestra se puede determinar de manera eficiente en base a la función predeterminada, lo que da lugar a un resultado preciso y confiable y una
45 buena reproducibilidad.

Cabe señalar que la expresión “el porcentaje de ácidos nucleicos libres de células presentes en el rango predeterminado” se refiere al porcentaje del número de ácidos nucleicos libres de células distribuidos en un cierto rango de longitud predeterminado para el número total de ácidos nucleicos libres de células en la muestra biológica.

50 En las realizaciones de la presente divulgación, la muestra de control es una muestra de sangre periférica obtenida de una mujer embarazada con un feto masculino normal, en la que se sabe que la fracción de ácidos nucleicos fetales libres de células se determina por el cromosoma Y. Por lo tanto, el rango predeterminado se determina con precisión.

En las realizaciones de la presente divulgación, la muestra de control es una muestra de sangre periférica obtenida de una mujer embarazada con un feto normal. Por lo tanto, el rango predeterminado se determina con precisión.

5 En las realizaciones de la presente divulgación, la fracción de ácidos nucleicos fetales libres de células en la muestra de control es una fracción de ADN fetal libre de células que se estima por el cromosoma Y. Por lo tanto, el rango predeterminado se puede determinar al utilizar eficientemente la fracción de ácidos nucleicos fetales libres de células de la muestra de control, y luego se puede determinar adicionalmente el número de ácidos nucleicos libres de células en la longitud que cae en el rango predeterminado y la fracción de ADN fetal libre de células en una muestra obtenida de una mujer embarazada bajo detección.

En otras realizaciones de la presente divulgación, el método puede incluir adicionalmente las siguientes etapas.

10 1) Secuenciación del genoma completo (WGS): la muestra bajo detección se somete a secuenciación del genoma completo utilizando una plataforma de alto rendimiento. Los ADN fetales libres de células en plasma, que son relativamente cortos y en los que solo una pequeña cantidad supera los 300 pb de longitud, se secuencian mediante secuenciación de un solo extremo o secuenciación de extremos emparejados, ya que es necesario obtener longitudes de todos los ADN fetales libres de células y se requiere secuenciar la molécula completa de ADN libre de células si se realiza mediante secuenciación de un solo extremo.

15 2) Obtener lecturas mapeadas de forma única: las lecturas de una muestra de prueba se alinean con una secuencia del genoma de referencia.

3) Obtener la longitud del ADN correspondiente a cada lectura mapeada de forma única en base a la información de alineación de cada lectura mapeada de forma única.

20 4) Seleccionar uno o más rangos con alta correlación: uno o más rangos con alta correlación se seleccionan de acuerdo con la distribución de longitud de las moléculas de ADN.

5) Obtener una fórmula de función: se obtiene la fórmula de función entre un porcentaje de ADN presente en el uno o más rangos con alta correlación obtenido en 4) y la fracción de ADN fetal libre de células conocida.

25 6) Obtener un porcentaje de ADN presente en uno o más rangos seleccionados, es decir, el porcentaje de ADN presente en uno o más rangos de longitud.

7) Obtener una fracción de ADN fetal libre de células de la muestra bajo detección en base a la fórmula de función y el porcentaje de ADN en la muestra bajo detección presente en uno o más rangos de longitud.

Específicamente, la Etapa 4) incluye las siguientes etapas:

30 I. Seleccionar una muestra de control, es decir, una muestra en la que se conoce la fracción de ADN fetal libre de células.

II. Todas las muestras se someten a secuenciación WGS y la información de longitud de las moléculas de ADN representadas por cada lectura alineada de forma única se obtiene mediante la alineación única de la información obtenida mediante la alineación única de las lecturas para un cromosoma.

35 III. El número de moléculas de ADN con una longitud seleccionada de hasta Mbp (M representa un valor de longitud máxima, la molécula de ADN libre de células puede tener una longitud de hasta 400 pb) se obtiene para todas las muestras de control.

40 IV. Se obtiene una pluralidad de rangos de ventana (rangos de longitud) a través del movimiento de una ventana en una cierta longitud de ventana de acuerdo con un cierto tamaño de etapa, y un porcentaje de moléculas de ADN presentes en cada rango de ventana, es decir, se calcula un porcentaje de moléculas de ADN presentes en cada rango de longitud. Cabe señalar que el número de moléculas de ADN presentes en cada rango de ventana, es decir, distribuidas en cada rango de longitud dividido por el número total de moléculas de ADN, se define como el porcentaje de moléculas de ADN presentes en cada rango de ventana. Por ejemplo, se puede tomar 1 pb, 5 pb, 10 pb o 15 pb como ventana y cualquier tamaño seleccionado desde 1 pb hasta la longitud de la ventana se puede tomar como el tamaño de etapa. Específicamente, como ejemplo, si se tomó 5 pb como la ventana y 2 pb como el tamaño de etapa, entonces se pueden obtener las distribuciones de moléculas de ADN en [1 pb, 5 pb], [2 pb, 6 pb], [4 pb, 8 pb], [6 pb, 10 pb] y así sucesivamente. Como otro ejemplo, si se tomó 5 pb como ventana y 5 pb como tamaño de etapa, entonces se pueden obtener distribuciones de moléculas de ADN en [1 pb, 5 pb], [6 pb, 10 pb], [11 pb, 15 pb] y así sucesivamente. El "número total de moléculas de ADN" descrito anteriormente en el presente documento se refiere al número total de todas las moléculas de ADN con diferente longitud.

50 V. Se determina un rango de ventana o una combinación de rangos de ventana (es decir, un rango de longitud o una pluralidad de rangos de longitud) en el que el porcentaje de moléculas de ADN presentes está altamente correlacionado con la fracción de ADN fetal libre de células conocida, y se establece una fórmula de función

Dispositivo para determinar una fracción de ácidos nucleicos fetales libres de células en una muestra biológica

También se divulga ahora, pero no forma parte de la invención reivindicada, un dispositivo para determinar una fracción de ácidos nucleicos fetales libres de células en una muestra de sangre periférica de una mujer embarazada mediante un método de la invención. Con referencia a la Fig. 7, el dispositivo incluye: un aparato 100 de secuenciación, un aparato 200 de recuento y un aparato 300 para determinar una fracción de ácidos nucleicos fetales libres de células.

5 Específicamente, el aparato 100 de secuenciación se configura para secuenciar ácidos nucleicos libres de células contenidos en la muestra biológica, para obtener un resultado de la secuenciación que consiste en una pluralidad de datos de secuenciación. El aparato 200 de recuento se conecta al aparato 100 de secuenciación y se configura para determinar el número de ácidos nucleicos libres de células en una longitud que cae en un rango predeterminado en la muestra biológica en base al resultado de la secuenciación. El aparato 300 para determinar una fracción de ácidos nucleicos fetales libres de células se conecta al aparato 200 de recuento y se configura para determinar la fracción de ácidos nucleicos fetales libres de células en base al número de ácidos nucleicos libres de células de una longitud que cae en el rango predeterminado.

10 Como se indicó anteriormente, los ácidos nucleicos libres de células en la muestra biológica se secuencian mediante secuenciación de extremos emparejados, secuenciación de un solo extremo o secuenciación de una sola molécula. Por lo tanto, se pueden obtener fácilmente longitudes de los ácidos nucleicos libres de células, lo que conduce a las etapas posteriores.

15 En las realizaciones de la presente divulgación, con referencia a la Fig. 8, el aparato 200 de recuento incluye adicionalmente: una unidad 210 de alineación, una primera unidad 220 de determinación de longitud y una unidad 230 de determinación de número. Específicamente, la unidad 210 de alineación se configura para alinear el resultado de la secuenciación con un genoma de referencia para construir un conjunto de datos que consiste en una pluralidad de lecturas mapeadas de forma única, donde cada lectura en el conjunto de datos se puede mapear a una única posición del genoma de referencia. La primera unidad 220 de determinación de longitud se conecta a la unidad 210 de alineación y se configura para determinar la longitud del ácido nucleico libre de células que corresponde a cada lectura mapeada de forma única en el conjunto de datos. La unidad 230 de determinación de número se conecta a la primera unidad 220 de determinación de longitud y se configura para determinar el número de ácidos nucleicos libres de células de una longitud que cae en el rango predeterminado. Por lo tanto, el número de ácidos nucleicos libres de células de una longitud que cae en el rango predeterminado en la muestra biológica se puede determinar fácilmente, lo que da lugar a un resultado preciso y confiable y una buena reproducibilidad.

20 En las realizaciones de la presente divulgación, el primer aparato 220 de determinación de longitud se configura para determinar la longitud de cada lectura mapeada de forma única en el genoma de referencia como la longitud del ácido nucleico libre de células que corresponde a la lectura. Por lo tanto, la longitud del ácido nucleico libre de células que corresponde a cada lectura mapeada de forma única en el conjunto de datos se puede determinar con precisión.

25 En las realizaciones de la presente divulgación, con referencia a la Fig.9, en el caso de que los ácidos nucleicos libres de células en la muestra biológica se secuencien mediante secuenciación de extremos emparejados, la primera unidad 220 de determinación de longitud incluye adicionalmente: un módulo 2210 de determinación de posición del extremo 5', un módulo 2220 de determinación de posición de extremo 3' y un módulo 2230 de cálculo de longitud. Específicamente, el módulo 2210 de determinación de posición del extremo 5' se configura para determinar una posición que corresponde al genoma de referencia del extremo 5' del ácido nucleico libre de células, en base a los datos de secuenciación en un extremo de cada lectura mapeada de forma única obtenida en la secuenciación de extremos emparejados. El módulo 2220 de determinación de posición del extremo 3' se conecta al módulo 2210 de determinación de posición del extremo 5' y se configura para determinar una posición que corresponde al genoma de referencia del extremo 3' del ácido nucleico libre de células, en base a los datos de secuenciación. en el otro extremo de la misma lectura mapeada de forma única obtenida en la secuenciación del extremo emparejado. El módulo 2230 de cálculo de longitud se conecta al módulo 2220 de determinación de posición del extremo 3' y se configura para determinar la longitud del ácido nucleico libre de células en base a la posición del extremo 5' del ácido nucleico libre de células y la posición del extremo 3' del ácido nucleico libre de células. Por lo tanto, la longitud del ácido nucleico libre de células que corresponde a cada lectura mapeada de forma única en el conjunto de datos se puede determinar con precisión.

30 En las realizaciones de la presente divulgación, el dispositivo incluye adicionalmente un aparato 400 de determinación de rango predeterminado configurado para determinar el rango predeterminado en base a una pluralidad de muestras de control, en cada una de las cuales se conoce la fracción de ácidos nucleicos fetales libres de células, opcionalmente, el rango predeterminado se determina en base a al menos 20 muestras de control.

35 En las realizaciones de la presente divulgación, con referencia a la Fig. 10, el aparato 400 de determinación de rango predeterminado incluye adicionalmente: una segunda unidad 410 de determinación de longitud, una primera unidad 420 de determinación de porcentaje, una unidad 430 de determinación de coeficiente de correlación y una unidad 440 de determinación de rango predeterminado. Específicamente, la segunda unidad 410 de determinación de longitud se configura para determinar las longitudes de los ácidos nucleicos libres de células en la pluralidad de muestras de control. La primera unidad 420 de determinación de porcentaje se conecta a la segunda unidad 410 de determinación de longitud y se configura para establecer una pluralidad de rangos de longitud candidatos y determinar un porcentaje de los ácidos nucleicos libres de células, obtenidos de cada una de la pluralidad de muestras de control, presentes en

cada rango de longitud candidato. La unidad 430 de determinación de coeficiente de correlación se conecta a la primera unidad 420 de determinación de porcentaje y se configura para determinar un coeficiente de correlación entre cada rango de longitud candidato y la fracción del ácido nucleico fetal libre de células, en base al porcentaje de ácidos nucleicos libres de células, obtenidos de cada una de la pluralidad de muestras de control, presentes en cada rango de longitud candidato y la fracción de ácidos nucleicos fetales libres de células en las muestras de control. La unidad 440 de determinación de rango predeterminado se conecta a la unidad 430 de determinación de coeficiente de correlación y se configura para seleccionar un rango de longitud candidato con el mayor coeficiente de correlación como rango predeterminado. Por lo tanto, el rango predeterminado se puede determinar de manera precisa y eficiente.

En las realizaciones de la presente divulgación, el rango de longitud candidato es de un tramo de 1 pb a 20 pb.

En las realizaciones de la presente divulgación, la pluralidad de rangos de longitud candidatos tiene un tamaño de etapa de 1 pb a 2 pb.

En las realizaciones de la presente divulgación, con referencia a la Fig. 11, el aparato 300 para determinar la fracción de ácidos nucleicos fetales libres de células incluye adicionalmente: una segunda unidad 310 de determinación de porcentaje y una unidad 320 para calcular una fracción de ácidos nucleicos fetales libres de células. Específicamente, la segunda unidad 310 de determinación de porcentaje se configura para determinar un porcentaje de los ácidos nucleicos libres de células presentes en el rango predeterminado en base al número de ácidos nucleicos libres de células en la longitud que cae en el rango predeterminado. La unidad 320 para calcular una fracción de ácidos nucleicos fetales libres de células se conecta a la segunda unidad 310 de determinación de porcentaje y se configura para determinar la fracción de ácidos nucleicos fetales libres de células en la muestra biológica, en base al porcentaje de ácidos nucleicos libres de células presentes en el rango predeterminado, de acuerdo con una función predeterminada, en la que la función predeterminada se determina en base a la pluralidad de muestras de control. Por lo tanto, la fracción de ácidos nucleicos fetales libres de células en la muestra biológica se puede determinar de forma eficaz, lo que da lugar a un resultado preciso y confiable y una buena reproducibilidad.

En las realizaciones de la presente divulgación, el dispositivo incluye adicionalmente un aparato 500 de determinación de función predeterminada. Con referencia a la Fig. 12, el aparato 500 de determinación de función predeterminada incluye: una tercera unidad 510 de determinación de porcentaje y una unidad 520 de ajuste. Específicamente, la tercera unidad 510 de determinación de porcentaje se configura para determinar el porcentaje de ácidos nucleicos libres de células, obtenidos de cada muestra de control, presentes en el rango predeterminado. La unidad 520 de ajuste se conecta a la tercera unidad 510 de determinación de porcentaje y se configura para ajustar el porcentaje de ácidos nucleicos libres de células, obtenido de cada muestra de control, presente en el rango predeterminado, con la fracción conocida del ácido nucleico fetal libre de células para determinar la función predeterminada. Por lo tanto, la función predeterminada se puede determinar de forma precisa y fiable, lo que conduce a las etapas posteriores. En las realizaciones de la presente divulgación, el porcentaje de ácidos nucleicos libres de células, obtenido de cada muestra de control, presente en el rango predeterminado se ajusta con la fracción conocida del ácido nucleico fetal libre de células de la fuente predeterminada mediante un ajuste lineal..

Método para determinar el sexo de los gemelos.

Como se indicó anteriormente, en otro aspecto, la presente invención proporciona un método para determinar el sexo de gemelos. Dicho método incluye: realizar la secuenciación sobre ácidos nucleicos libres de células contenidos en una muestra de sangre periférica obtenida de una mujer embarazada con gemelos, para obtener un resultado de la secuenciación que consiste en una pluralidad de datos de secuenciación; determinar una primera fracción de ADN fetal libre de células en base a los datos de secuenciación, mediante el método de la invención anteriormente descrita; determinar una segunda fracción de ADN fetal libre de células en base a los datos de secuenciación derivados del cromosoma Y en el resultado de la secuenciación; y determinar el sexo de los gemelos en base a la primera fracción de ADN fetal libre de células y la segunda fracción de ADN fetal libre de células. Los inventores han encontrado sorprendentemente que el sexo de los gemelos en una mujer embarazada se puede determinar de forma precisa y eficaz mediante este método.

En las realizaciones de la presente divulgación, la segunda fracción de ADN fetal libre de células se determina de acuerdo con la siguiente fórmula:

$$fra.cry=(cry.ER\%-Mujer.cry.ER\%)/(Hombre.cry.ER\%-Mujer.cry.ER\%)*100\%$$

donde *fra.cry* representa la segunda fracción de ADN fetal libre de células, *cry.ER%* representa un porcentaje de los datos de secuenciación derivados del cromosoma Y en el resultado de la secuenciación para los datos de secuenciación totales; *Mujer.cry.ER%* representa un porcentaje promedio de los datos de secuenciación de ácidos nucleicos libres de células alineados con el cromosoma Y en una muestra de sangre periférica obtenida de una mujer embarazada predeterminada para estar con un feto femenino normal para los datos de secuenciación totales de los mismos; y *Hombre.cry.ER%* representa un porcentaje promedio de los datos de secuenciación de ácidos nucleicos libres de células derivados del cromosoma Y en una muestra de sangre periférica obtenida de un hombre sano para los datos de secuenciación totales de los mismos.

Por lo tanto, la segunda fracción de ADN fetal libre de células se puede determinar con precisión.

En las realizaciones de la presente divulgación, determinar el sexo de los gemelos en base a la primera fracción de ADN fetal libre de células y la segunda fracción de ADN fetal libre de células incluye adicionalmente: (a) determinar una relación de la segunda fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células; y (b) determinar el sexo de los gemelos al comparar la relación determinada en (a) con un primer umbral predeterminado y un segundo umbral predeterminado.

El primer umbral se determina en base a una pluralidad de muestras de control obtenidas de mujeres embarazadas que se sabe tienen gemelos femeninos, y el segundo umbral se determina en base a una pluralidad de muestras de control obtenidas de mujeres embarazadas que se sabe tienen gemelos varones.

Ambos fetos de los gemelos se determinan como femeninos si la relación de la segunda fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células es menor que el primer umbral, ambos fetos de los gemelos se determinan como masculinos si la relación de la segunda fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células es mayor que el segundo umbral, y se determina que los gemelos incluyen un feto masculino y un feto femenino si la relación de la segunda fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células es igual al primer umbral o al segundo umbral, o entre el primer umbral y el segundo umbral.

En las realizaciones de la presente divulgación, el primer umbral es 0.35 y el segundo umbral es 0.7.

También se describe ahora, pero no forma parte de la invención reivindicada, un sistema para determinar el sexo de gemelos de acuerdo con un método de la invención. En las realizaciones de la presente divulgación, el sistema incluye:

un primer dispositivo de determinación de fracción de ADN fetal libre de células, que es el dispositivo anterior para determinar la fracción de ácidos nucleicos fetales libres de células en la muestra biológica, y configurado para secuenciar los ácidos nucleicos libres de células contenidos en una muestra de sangre periférica obtenida de una mujer embarazada con gemelos, para obtener un resultado de la secuenciación que consiste en una pluralidad de datos de secuenciación, y configurado para determinar una primera fracción de ADN fetal libre de células en base a los datos de secuenciación;

un segundo dispositivo de determinación de fracción de ADN fetal libre de células, configurado para determinar una segunda fracción de ADN fetal libre de células en base a los datos de secuenciación derivados del cromosoma Y en el resultado de la secuenciación; y

un dispositivo de determinación del sexo, configurado para determinar el sexo de los gemelos en función de la primera fracción de ADN fetal libre de células y la segunda fracción de ADN fetal libre de células.

En las realizaciones de la presente divulgación, el dispositivo de determinación del sexo incluye adicionalmente: una unidad de determinación de relación, configurada para determinar una relación de la segunda fracción de ADN fetal libre de células a la primera fracción de ADN fetal libre de células; y una unidad de comparación, configurada para comparar la relación determinada por la unidad de determinación de relación con un primer umbral predeterminado y un segundo umbral predeterminado, para determinar el sexo de los gemelos. Por lo tanto, el sexo de los gemelos se puede determinar de manera eficiente.

Método para detectar una aneuploidía cromosómica de gemelos.

Como se indicó anteriormente, en un aspecto adicional, la presente invención proporciona un método para detectar una aneuploidía cromosómica de gemelos. Dicho método incluye:

realizar la secuenciación sobre ácidos nucleicos libres de células contenidos en una muestra de sangre periférica obtenida de una mujer embarazada con gemelos, para obtener un resultado de la secuenciación que consiste en una pluralidad de datos de secuenciación;

determinar una primera fracción de ADN fetal libre de células, en base a los datos de secuenciación, mediante el método de la invención descrito anteriormente;

determinar una tercera fracción de ADN fetal libre de células, en base a los datos de secuenciación derivados de un cromosoma predeterminado en el resultado de la secuenciación; y

determinar si los gemelos bajo detección tienen aneuploidía con respecto al cromosoma predeterminado en base a la primera fracción de ADN fetal libre de células y la tercera fracción de ADN fetal libre de células.

Por lo tanto, la aneuploidía cromosómica de los gemelos se puede detectar con precisión y eficacia.

En las realizaciones de la presente divulgación, la tercera fracción de ADN fetal libre de células se determina de acuerdo con la siguiente fórmula:

$$fra.cri = 2*(cri.ER\%/ajuste.cri.ER\%-1)*100\%$$

donde *fra.chi* representa la tercera fracción de ADN fetal libre de células, *i* representa un número de serie del cromosoma predeterminado e *i* es cualquier entero en el rango de 1 a 22; *cri.ER%* representa un porcentaje de los datos de secuenciación derivados del cromosoma predeterminado en el resultado de la secuenciación para los datos de secuenciación totales; *ajuste.cri.ER%* representa un porcentaje promedio de los datos de secuenciación de ácidos nucleicos libres de células derivados del cromosoma predeterminado en una muestra de sangre periférica obtenida de una mujer embarazada predeterminada para estar con gemelos normales para los datos de secuenciación totales de los mismos. Por lo tanto, la tercera fracción de ADN fetal libre de células se puede determinar con precisión.

En las realizaciones de la presente divulgación, determinar si los gemelos bajo detección tienen aneuploidía con respecto al cromosoma predeterminado en base a la primera fracción de ADN fetal libre de células y la tercera fracción de ADN fetal libre de células incluye adicionalmente: (a) determinar una relación de la tercera fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células; y (b) determinar si los gemelos bajo detección tienen aneuploidía con respecto al cromosoma predeterminado al comparar la relación determinada en (a) con un tercer umbral predeterminado y un cuarto umbral predeterminado. Por lo tanto, la aneuploidía cromosómica de los gemelos se puede detectar de manera eficiente.

El tercer umbral se determina en base a una pluralidad de muestras de control obtenidas de mujeres embarazadas con gemelos que se sabe que no tienen aneuploidía con respecto al cromosoma predeterminado, y el cuarto umbral se determina en base a una pluralidad de muestras de control obtenidas de mujeres embarazadas con gemelos que se sabe tienen aneuploidía con respecto al cromosoma predeterminado.

Se determina que ambos fetos de los gemelos no tienen aneuploidía con respecto al cromosoma predeterminado si la relación de la tercera fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células es menor que el tercer umbral, se determina que ambos fetos de los gemelos tienen aneuploidía con respecto al cromosoma predeterminado si la relación de la tercera fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células es mayor que el cuarto umbral, y se determina que un feto de los gemelos tiene aneuploidía con respecto al cromosoma predeterminado, mientras que se determina que el otro feto de los gemelos no tiene aneuploidía con respecto al cromosoma predeterminado si la relación de la tercera fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células es igual al tercer umbral o el cuarto umbral, o entre el tercer umbral y el cuarto umbral.

En las realizaciones de la presente divulgación, el tercer umbral es 0.35 y el cuarto umbral es 0.7.

En las realizaciones de la presente divulgación, el cromosoma predeterminado es al menos uno seleccionado de los cromosomas 18, 21 y 23.

También se divulga ahora, pero no forma parte de la invención reivindicada, un sistema para determinar una aneuploidía cromosómica de gemelos de acuerdo con un método como el anterior. El sistema incluye:

un primer dispositivo de determinación de fracción de ADN fetal libre de células, es el dispositivo anterior para determinar la fracción de ácidos nucleicos libres de células en la muestra biológica, y configurado para secuenciar los ácidos nucleicos libres de células contenidos en una muestra de sangre periférica obtenida de una mujer embarazada con gemelos, para obtener un resultado de la secuenciación que consiste en una pluralidad de datos de secuenciación, y configurado para determinar una primera fracción de ADN fetal libre de células en base a los datos de secuenciación;

un tercer dispositivo de determinación de fracción de ADN fetal libre de células, configurado para determinar una tercera fracción de ADN fetal libre de células en base a los datos de secuenciación derivados de un cromosoma predeterminado en el resultado de la secuenciación; y un primer dispositivo de determinación de aneuploidía, configurado para determinar si los gemelos bajo detección tienen aneuploidía con respecto al cromosoma predeterminado en base a la primera fracción de ADN fetal libre de células y la tercera fracción de ADN fetal libre de células. Los inventores han encontrado sorprendentemente que la aneuploidía cromosómica de gemelos se puede detectar de forma precisa y eficaz mediante un sistema de este tipo de acuerdo con la presente divulgación.

En las realizaciones de la presente divulgación, el primer dispositivo de determinación de aneuploidía incluye adicionalmente:

una unidad de determinación de relación, configurada para determinar una relación de la tercera fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células; y

una unidad de comparación, configurada para comparar la relación determinada por la unidad de determinación de relación con un tercer umbral predeterminado y un cuarto umbral predeterminado, para determinar si los gemelos detectados tienen aneuploidía con respecto al cromosoma predeterminado. Por lo tanto, la aneuploidía cromosómica de los gemelos se puede detectar de manera eficiente.

En todavía otro aspecto adicional, la presente divulgación proporciona un método para determinar una aneuploidía cromosómica de gemelos, método que incluye:

- realizar la secuenciación sobre ácidos nucleicos libres de células contenidos en una muestra de sangre periférica obtenida de una mujer embarazada con gemelos, para obtener un resultado de la secuenciación que consiste en una pluralidad de datos de secuenciación y determinar a partir de ellos la fracción de ácidos nucleicos fetales libres de células como una primera fracción de ADN libre de células mediante el método de la invención como se describió anteriormente;
- 5 determinar una fracción x_i del número de datos de secuenciación derivados del cromosoma i en el resultado de la secuenciación para los datos de secuenciación totales, donde i representa un número de serie del cromosoma, e es cualquier entero en el rango de 1 a 22;
- 10 determinar una puntuación T del cromosoma i de acuerdo con $T_i = (X_i - \mu_i) / \sigma_i$, donde i representa el número de serie del cromosoma y es cualquier entero en el rango de 1 a 22, μ_i representa un porcentaje promedio de los datos de secuenciación del cromosoma i seleccionado como un sistema de referencia en una base de datos de referencia para los datos de secuenciación totales del mismo, σ_i representa una desviación estándar de los porcentajes de los datos de secuenciación del cromosoma i seleccionado como un sistema de referencia en la base de datos de referencia para los datos de secuenciación totales del mismo,
- 15 determinar una puntuación L del cromosoma i de acuerdo con $L_i = \log(d(T_i, a)) / \log(d(T2_i, a))$, donde i representa el número de serie del cromosoma e i es cualquier entero en el rango de 1 a 22, $T2_i = (X_i - \mu_i * (1 - fra/2)) / \sigma_i$; $d(T_i, a)$ y $d(T2_i, a)$ representan la función de densidad de probabilidad de distribución t , donde a representa el grado de libertad y fra representa dicha primera fracción de ADN fetal libre de células;
- 20 graficar un diagrama de cuatro cuadrantes con T como coordenada vertical y L como coordenada horizontal mediante la zonificación con una primera línea recta donde $T =$ quinto umbral predeterminado y una segunda línea recta donde $L =$ sexto umbral predeterminado, en el que se determina que ambos fetos de los gemelos tienen un trisoma si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que se encuentra en un primer cuadrante; se determina que un feto de los gemelos tiene un trisoma y se determina que el otro feto de los gemelos es normal si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que cae en un segundo cuadrante; se determina que ambos fetos de los gemelos son normales si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que se encuentra en un tercer cuadrante; se determina que los gemelos han dado lugar a una fracción fetal baja si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que cae en un cuarto cuadrante, de tal manera que no se adopta el resultado. Los inventores han encontrado sorprendentemente que la determinación de si los gemelos bajo detección tienen
- 25 aneuploidía con respecto a un cromosoma predeterminado se puede lograr de forma precisa y eficaz mediante este método.
- 30 También se divulga ahora, pero no forma parte de la invención reivindicada, un sistema para determinar una aneuploidía cromosómica de gemelos de acuerdo con la invención que incluye:
- 35 un dispositivo de determinación de valor x_i , configurado para secuenciar ácidos nucleicos libres de células contenidos en una muestra de sangre periférica obtenida de una mujer embarazada con gemelos, para obtener un resultado de la secuenciación que consiste en una pluralidad de datos de secuenciación, y configurado para determinar una fracción x_i del número de datos de secuenciación derivados del cromosoma i en el resultado de la secuenciación para los datos de secuenciación total, donde i representa un número de serie del cromosoma e i es cualquier entero en el rango de 1 a 22;
- 40 un dispositivo de determinación de puntuación T , configurado para determinar una puntuación T del cromosoma i de acuerdo con $T_i = (X_i - \mu_i) / \sigma_i$, donde i representa el número de serie del cromosoma e i es cualquier entero en el rango de 1 a 22, μ_i representa un porcentaje promedio de los datos de secuenciación del cromosoma i seleccionado como un sistema de referencia en una base de datos de referencia para los datos de secuenciación totales del mismo, σ_i representa una desviación estándar de los porcentajes de los datos de secuenciación del cromosoma i seleccionado como un sistema de referencia en la base de datos de referencia para los datos de secuenciación totales del mismo,
- 45 un dispositivo de determinación de puntuación L , configurado para determinar una puntuación L del cromosoma i de acuerdo con $L_i = \log(d(T_i, a)) / \log(d(T2_i, a))$, donde i representa el número de serie del cromosoma e i es cualquier entero en el rango de 1 a 22, $T2_i = (X_i - \mu_i * (1 - fra/2)) / \sigma_i$; $d(T_i, a)$ y $d(T2_i, a)$ representan la función de densidad de probabilidad de distribución t , donde a representa el grado de libertad, y fra representa una primera fracción de ADN fetal libre de células determinada por el método de la invención descrito anteriormente;
- 50 un segundo dispositivo de determinación de aneuploidía, configurado para graficar un diagrama de cuatro cuadrantes con T como coordenada vertical y L como coordenada horizontal mediante zonificación con una primera línea recta donde $T =$ quinto umbral predeterminado y una segunda línea recta donde $L =$ sexto umbral predeterminado, en el que se determina que ambos fetos de los gemelos tienen un trisoma si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que cae en un primer cuadrante; se determina que un feto de los gemelos tiene un trisoma y se determina que el otro feto de los gemelos es normal si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que cae en un segundo cuadrante; se determina que ambos fetos de los gemelos son normales si se determina que una muestra bajo detección tiene una puntuación T y una
- 55

puntuación L que cae en un tercer cuadrante; se determina que los gemelos tienen una fracción fetal baja si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que se encuentra en un cuarto cuadrante, de tal manera que no se adopta el resultado.

5 Cabe señalar que, la “base de datos de referencia” descrita en “ μ_i representa un valor promedio de porcentajes de datos de secuenciación del cromosoma i seleccionado como un sistema de referencia en una base de datos de referencia para los datos de secuenciación totales del mismo” se refiere a los ácidos nucleicos libres de células en una muestra de sangre periférica obtenida de una mujer embarazada con un feto normal (feto femenino, feto masculino, feto único o gemelos) o datos de secuenciación (lecturas).

10 Los “datos de secuenciación” expresados en “datos de secuenciación derivados del cromosoma Y” significa lecturas obtenidas en la secuenciación.

En algunas realizaciones específicas de la presente divulgación, los términos “ x_i ,” “ER_i” y “Cri.ER%” se pueden cambiar en el presente documento, es decir, x_i puede ser un resultado obtenido después de someterse a la corrección de GC. Específicamente, el contenido de ER y GC de cada cromosoma se puede ajustar utilizando datos conocidos obtenidos de muestras normales para obtener una fórmula de relación, es decir $ER_i = F_i(GC_i) + \varepsilon_i$. Se calcula un valor

15 medio \overline{ER}_i de ER. Para analizar una muestra, se calcula un valor ER después de la corrección de acuerdo con la siguiente fórmula y se basa en la fórmula de relación anterior y ER y GC de la muestra.

$$\overline{ER}_{ij} = \overline{ER}_i + \varepsilon_i = \overline{ER}_i + ER_y - f_i(GC_{ij})$$

Método para detectar quimera fetal

En otro aspecto, la presente invención proporciona un método para detectar quimera fetal que incluye:

20 realizar la secuenciación sobre ácidos nucleicos libres de células contenidos en una muestra de sangre periférica obtenida de una mujer embarazada con un feto, opcionalmente un feto masculino, para obtener un resultado de la secuenciación que consiste en una pluralidad de datos de secuenciación;

determinar una primera fracción de ADN fetal libre de células, en base a los datos de secuenciación, mediante el método de la invención descrito anteriormente,

25 determinar una tercera fracción de ADN fetal libre de células en base a los datos de secuenciación derivados de un cromosoma predeterminado en el resultado de la secuenciación; y

determinar si el feto bajo detección tiene quimera fetal con respecto al cromosoma predeterminado en base a la primera fracción de ADN fetal libre de células y la tercera fracción de ADN fetal libre de células.

30 Por lo tanto, se puede analizar con precisión si el feto bajo detección tiene quimera fetal con respecto a un cromosoma específico.

En las realizaciones de la presente divulgación, la tercera fracción de ADN fetal libre de células se determina mediante la siguiente fórmula:

35 $fra.cri = 2 * (cri.ER\% / ajuste.cri.ER\% - 1) * 100\%$, donde *fra.cri* representa la tercera fracción de ADN fetal libre de células, *i* representa un número de serie del cromosoma predeterminado e *i* es cualquier entero en el rango de 1 a 22; *cri.ER%* representa un porcentaje de los datos de secuenciación derivados del cromosoma predeterminado en el resultado de la secuenciación para los datos de secuenciación totales; *ajuste.cri.ER%* representa un porcentaje promedio de los datos de secuenciación de ácidos nucleicos libres de células derivados del cromosoma predeterminado en una muestra de sangre periférica obtenida de una mujer embarazada predeterminada para estar con un feto para los datos de secuenciación totales del mismo. Por lo tanto, se puede analizar si el feto bajo detección tiene quimera fetal con

40 respecto al cromosoma específico con una eficiencia mejorada adicionalmente.

En las realizaciones de la presente divulgación, determinar si el feto bajo detección tiene quimera fetal con respecto al cromosoma predeterminado en base a la primera fracción de ADN fetal libre de células y la tercera fracción de ADN fetal libre de células incluye adicionalmente: (a) determinar una relación de la tercera fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células; y (b) determinar si el feto bajo detección tiene quimera con respecto al cromosoma predeterminado al comparar la relación determinada en (a) con una pluralidad de umbrales predeterminados. Por lo tanto, se puede analizar si el feto bajo detección tiene quimera fetal con respecto a un cromosoma específico con una eficiencia mejorada adicional.

45

En las realizaciones de la presente divulgación, la pluralidad de umbrales predeterminados incluye al menos uno seleccionado de:

50 un séptimo umbral, determinado en base a una pluralidad de muestras de control con el cromosoma predeterminado que se sabe es un monosoma completo,

un octavo umbral, determinado en base a una pluralidad de muestras de control con el cromosoma predeterminado que se sabe es una quimera monosoma,

un noveno umbral, determinado en base a una pluralidad de muestras de control con el cromosoma predeterminado que se sabe que es normal,

- 5 un décimo umbral, determinado en base a una pluralidad de muestras de control con el cromosoma predeterminado que se sabe es un trisoma completo.

Se determina que el cromosoma predeterminado del feto bajo detección es un monosoma completo, si la relación de la tercera fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células es menor que el séptimo umbral;

- 10 se determina que el cromosoma predeterminado del feto bajo detección es una quimera monosoma, si la relación de la tercera fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células no es menor que el séptimo umbral ni mayor que el octavo umbral;

- 15 se determina que el cromosoma predeterminado del feto bajo detección es normal, si la relación de la tercera fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células es mayor que el octavo umbral y menor que el noveno umbral;

se determina que el cromosoma predeterminado del feto bajo detección es una quimera trisoma, si la relación de la tercera fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células no es menor que el noveno umbral ni superior al décimo umbral; y

- 20 se determina que el cromosoma predeterminado del feto bajo detección es un trisoma completo, si la relación de la tercera fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células es mayor que el décimo umbral.

En las realizaciones de la presente divulgación, el séptimo umbral es mayor que -1 y menor que 0, por ejemplo -0.85;

el octavo umbral es mayor que el séptimo umbral y menor que 0, por ejemplo, -0.3;

el noveno umbral es mayor que 0 y menor que 1, por ejemplo 0.3;

- 25 el décimo umbral es mayor que el noveno umbral y menor que 1, por ejemplo 0.85.

Por lo tanto, se puede analizar si el feto bajo detección tiene quimera fetal con respecto a un cromosoma específico, con una eficacia mejorada adicional.

Adicionalmente, ahora se divulga, pero no forma parte de la invención reivindicada, un sistema para detectar quimeras fetales como anteriormente. El sistema incluye:

- 30 un primer dispositivo de determinación de fracción de ADN fetal libre de células, es el dispositivo anterior para determinar la fracción de ácidos nucleicos libres de células en la muestra biológica, y configurado para secuenciar los ácidos nucleicos libres de células contenidos en una muestra de sangre periférica obtenida de una mujer embarazada con el feto, para obtener un resultado de la secuenciación que consiste en una pluralidad de datos de secuenciación, y configurado para determinar una primera fracción de ADN fetal libre de células en base a los datos de secuenciación;

- 35 un tercer dispositivo de determinación de fracción de ADN fetal libre de células, configurado para determinar una tercera fracción de ADN fetal libre de células en base a los datos de secuenciación derivados de un cromosoma predeterminado en el resultado de la secuenciación; y

- 40 un dispositivo de determinación de quimera, configurado para determinar si el feto bajo detección tiene quimera fetal con respecto al cromosoma predeterminado en base a la primera fracción de ADN fetal libre de células y la tercera fracción de ADN fetal libre de células.

Dicho sistema para detectar quimeras fetales puede incluir adicionalmente:

una unidad de determinación de relación, configurada para determinar una relación de la tercera fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células; y

- 45 una unidad de comparación, configurada para comparar la relación determinada por la unidad de determinación de relación con una pluralidad de umbrales predeterminados, para determinar si el feto bajo detección tiene quimera con respecto al cromosoma predeterminado.

En un aspecto adicional, la presente invención proporciona un método para detectar quimera fetal que incluye:

realizar la secuenciación sobre los ácidos nucleicos libres de células contenidos en una muestra de sangre periférica obtenida de una mujer embarazada con un feto, para obtener un resultado de la secuenciación que consiste en una

pluralidad de datos de secuenciación y determinar a partir de ellos la fracción de ácidos nucleicos fetales libres de células como una primera fracción de ADN fetal libre de células mediante el método de la invención descrito anteriormente;

5 determinar una fracción x_i del número de datos de secuenciación derivados del cromosoma i en el resultado de la secuenciación para los datos de secuenciación totales, donde i representa un número de serie del cromosoma, e i es cualquier entero en el rango de 1 a 22;

10 determinar una puntuación T del cromosoma i de acuerdo con $T_i = (X_i - \mu_i) / \sigma_i$, donde i representa el número de serie del cromosoma e i es cualquier entero en el rango de 1 a 22, μ_i representa un valor promedio de porcentajes de datos de secuenciación del cromosoma i seleccionado como un sistema de referencia en una base de datos de referencia para los datos de secuenciación totales del mismo, σ_i representa una desviación estándar de los porcentajes de los datos de secuenciación del cromosoma i seleccionado como un sistema de referencia en la base de datos de referencia para los datos de secuenciación totales del mismo;

15 determinar una puntuación L del cromosoma i de acuerdo con $L_i = \log(d(T_i, a)) / \log(d(T2_i, a))$, donde i representa el número de serie del cromosoma e i es cualquier entero en el rango de 1 a 22, $T2_i = (X_i - \mu_i * (1/2)) / \sigma_i$; $d(T_i, a)$ y $d(T2_i, a)$ representan la función de densidad de probabilidad de distribución t , a representa el grado de libertad y fra representa dicha primera fracción de ADN fetal libre de células;

graficar un diagrama de cuatro cuadrantes con T como coordenada vertical y L como coordenada horizontal mediante zonificación con una tercera línea recta donde $T =$ undécimo umbral predeterminado y una cuarta línea recta donde $L =$ duodécimo umbral predeterminado, cuando la puntuación T no es mayor que 0,

20 en el que se determina que el feto tiene un monosoma completo o quimera monosoma con respecto al cromosoma predeterminado, si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que cae en un primer cuadrante;

se determina que el feto tiene una quimera monosoma con respecto al cromosoma predeterminado, si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que cae en un segundo cuadrante;

25 se determina que el feto es normal con respecto al cromosoma predeterminado, si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que cae en un tercer cuadrante; y

se determina que el feto ha dado lugar a una fracción fetal baja si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que cae en un cuarto cuadrante, de tal manera que no se adopta el resultado,

30 graficar un diagrama de cuatro cuadrantes con T como coordenada vertical y L como coordenada horizontal mediante la zonificación con una quinta línea recta donde $T =$ decimotercero umbral predeterminado y una sexta línea recta donde $L =$ decimocuarto umbral predeterminado, cuando la puntuación T es mayor que 0,

en el que se determina que el feto tiene un trisoma o quimera trisoma completa con respecto al cromosoma predeterminado, si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que cae en un primer cuadrante;

35 se determina que el feto tiene una quimera trisoma con respecto al cromosoma predeterminado, si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que cae en un segundo cuadrante;

se determina que el feto es normal con respecto al cromosoma predeterminado, si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que cae en un tercer cuadrante; y

40 se determina que el feto ha dado lugar a una fracción fetal baja si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que cae en un cuarto cuadrante, de tal manera que no se adopta el resultado.

Opcionalmente, el undécimo umbral y el decimotercero umbral son cada uno independientemente 3, y el duodécimo umbral y el decimocuarto umbral son cada uno independientemente 1.

Por lo tanto, las situaciones de quimera fetal se pueden analizar de manera eficiente.

45 Cabe señalar que la expresión "feto masculino/feto femenino/feto normal" significa que el cromosoma del feto es normal, por ejemplo, "feto masculino normal" se refiere a un feto masculino con cromosomas normales. Más aún, "feto masculino /feto femenino/feto normal" se puede referir a un solo feto o gemelos, por ejemplo, "feto masculino normal" puede ser feto único normal o gemelos normales; el "feto normal" no limita el sexo del feto ni limita que el feto sea un solo feto o gemelos.

50 Las realizaciones de la presente divulgación se describirán en detalle a continuación con referencia a los ejemplos, pero aquellos expertos en la técnica deben apreciar que los siguientes ejemplos se utilizan simplemente para ilustrar la presente divulgación, por lo que no se interpretará como una limitación del alcance de la presente divulgación. Un ejemplo en el que no se especifique la condición específica se llevará a cabo bajo condiciones normales o

recomendadas por el fabricante. Los reactivos o instrumentos sin fabricante especificado son productos convencionales disponibles en el mercado.

Ejemplo 1

5 Las fracciones de ADN fetal libre de células en muestras de plasma obtenidas de 11 mujeres embarazadas bajo detección se estimaron de acuerdo con la invención como sigue.

1) Recolección y tratamiento de muestras

Se sometieron a separación de plasma muestras de sangre periférica de 2 ml, extraídas durante el embarazo de cada una de las 11 mujeres embarazadas bajo detección y de 37 mujeres embarazadas con fetos masculinos, para obtener una muestra de plasma de cada mujer embarazada.

10 2) Construcción de biblioteca

La biblioteca se construyó de acuerdo con los requisitos de construcción de bibliotecas de plasma de Complete Genomics Inc.

3) Secuenciación

15 El proceso de secuenciación se llevó a cabo siguiendo estrictamente el procedimiento operativo estándar de Complete Genomics Inc.

4) Análisis de datos

La distribución de la longitud de los fragmentos de ADN se analizó con las lecturas obtenidas en la secuenciación de extremos emparejados; el proceso fue como se muestra en la Fig.1. Las etapas específicas fueron las siguientes:

20 a) Se calcularon las longitudes de los fragmentos de ADN en la muestra de sangre periférica obtenida de cada una de las 11 mujeres embarazadas bajo detección y de las 37 mujeres embarazadas que se sabía que tenían fetos masculinos. Específicamente, se eligieron 19 pb en un extremo de cada lectura mapeada de forma única y 12 pb en el otro extremo de cada lectura mapeada de forma única para determinar las posiciones inicial y de extremo correspondientes al genoma de referencia de la lectura mapeada de forma única y, por lo tanto, la longitud del fragmento de ADN se obtuvo en base a las posiciones inicial y de extremo correspondientes al genoma de referencia de la lectura alineada de forma única.

30 b) Para la muestra de sangre periférica obtenida de cada una de las 37 mujeres embarazadas que se sabe tienen un feto masculino, se obtuvieron una pluralidad de rangos de ventana (rango de longitud) a través del movimiento de una ventana en una cierta longitud de ventana de acuerdo con un cierto tamaño de etapa; se calculó el porcentaje de moléculas de ADN presentes en cada rango de ventana, es decir, el porcentaje de moléculas de ADN presentes en cada rango de longitud. Cabe señalar que el número de moléculas de ADN presentes en cada rango de ventana, es decir, distribuidas en cada rango de longitud, dividido por el número total de moléculas de ADN, se definió como el porcentaje de moléculas de ADN presentes en cada rango de ventana. Por ejemplo, se puede tomar 1 pb, 5 pb, 10 pb o 15 pb como ventana y cualquier tamaño seleccionado desde 1 pb hasta la longitud de la ventana se puede tomar como tamaño de etapa. Específicamente, como ejemplo, si se tomó 5 pb como la ventana y 2 pb como el tamaño de etapa, entonces se pueden obtener las distribuciones de moléculas de ADN en [1 pb, 5 pb], [2 pb, 6 pb], [4 pb, 8 pb], [6 pb, 10 pb] y así sucesivamente. Como otro ejemplo, si se tomó 5 pb como ventana y 5 pb como tamaño de etapa, entonces se pueden obtener distribuciones de moléculas de ADN en [1 pb, 5 pb], [6 pb, 10 pb], [11 pb, 15 pb] y así sucesivamente.

40 c) Se selecciona/ron uno o más rangos, en los que los porcentajes de moléculas de ADN presentes para muestras de sangre periférica obtenidas de 37 mujeres embarazadas que se sabía tenían fetos masculinos estaban fuertemente correlacionados con la fracción fetal conocida. Para un cierto rango de longitud, las fracciones de ADN fetal libre de células en muestras de sangre periférica obtenidas de 37 mujeres embarazadas que se sabe que tienen fetos masculinos se estimaron mediante el cromosoma Y (para el método de estimación específico, se puede hacer referencia a Jiang et al. Noninvasive Fetal Trisomy (NIFTY) test: an advanced noninvasive prenatal diagnosis methodology for fetal autosomal and sex chromosomal aneuploidies. BMC Med Genomics. 2012 Dec 1;5:57 y se calculó un coeficiente de correlación entre cada fracción de ADN fetal libre de células y el porcentaje de moléculas de ADN en una longitud que cae en M. Además, se seleccionó un rango de longitud M con un valor absoluto máximo de coeficiente de correlación, por ejemplo, $M = 185\sim 204$ pb, el coeficiente de correlación $R = -0.87$, como se muestra en la Fig. 13; o $M = 121\sim 150$ pb, el coeficiente de correlación $R = -0.6199$.

50 d) Se determinó la relación funcional entre el porcentaje de moléculas de ADN en una muestra de sangre periférica obtenida de una mujer embarazada, presente en el rango de longitud M y la fracción de ADN fetal libre de células (registrada como d). Específicamente, con respecto a las 37 muestras anteriores con fracciones conocidas de ADN fetal libre de células d (muestras obtenidas de 37 mujeres embarazadas que se sabe que tienen fetos masculinos), se trazó un gráfico de ajuste lineal utilizando los porcentajes $p(i = 1, 2, \dots, 37)$ de moléculas de ADN presentes en el rango

de longitud 185~204 pb y fracciones de ADN fetal libre de células d_i ($i = 1, 2, \dots, 37$), de manera que se obtuvo una fórmula de relación entre ellos: $d = a * p + b$, donde $d = 0.0334 * p + 1.6657$.

5 e) Se obtuvo la distribución de longitud de los fragmentos de ADN y el porcentaje de moléculas de ADN presentes en M para cada muestra obtenida de las 11 mujeres embarazadas bajo detección. Se obtuvo el porcentaje p de fragmentos de ADN presentes en el rango de longitud de 185 pb~204 pb para cada muestra obtenida de las mujeres embarazadas bajo detección, y los resultados se muestran en la Tabla 1 a continuación.

10 f) Se estimaron las fracciones de ADN fetal libre de células de las muestras bajo detección. Se calculó la fracción de ADN fetal libre de células d_j de cada muestra bajo detección directamente de acuerdo con el porcentaje p_j (j era una marca de la muestra bajo detección) de fragmentos de ADN presentes en el rango de longitud de 185 pb~204 pb obtenidos como anteriormente para cada muestra bajo detección y la fórmula de relación $d = a * p + b$.

g) Los resultados estimados de las fracciones de ADN fetal libre de células de las muestras bajo detección se muestran en la Tabla 1; para las muestras obtenidas de 37 mujeres embarazadas que se sabe que tienen fetos masculinos, las fracciones de ADN fetal libre de células estimadas por crY estaban básicamente de acuerdo con aquellas obtenidas por el método de la invención.

15 Tabla 1

Porcentaje de moléculas de ADN presentes en 185 pb ~204 pb (%)	Fracción de ADN fetal libre de células (estimada por crY)	Fracción de ADN fetal libre de células (estimada por el porcentaje de moléculas de ADN presentes en 185 pb~204 pb, es decir, el método de la presente invención)	Muestra
45.33932509	0.14872	0.145567224	muestra de prueba
45.85055949	0.12618	0.129861913	muestra de prueba
47.80808312	0.06752	0.06972606	muestra de prueba
46.78175523	0.10438	0.101255234	muestra de prueba
47.76095629	0.07148	0.071173814	muestra de prueba
45.18367884	0.15494	0.150348735	muestra de prueba
45.61282777	0.13413	0.13716512	muestra de prueba
48.50026322	0.04923	0.04846203	muestra de prueba
47.30076126	0.08555	0.085311176	muestra de prueba
47.12684798	0.09149	0.090653857	muestra de prueba
47.59287602	0.07892	0.076337302	muestra de prueba

Ejemplo 2:

20 La determinación del sexo de los gemelos y la detección de la aneuploidía cromosómica de los gemelos se llevaron a cabo utilizando muestras de sangre periférica obtenidas de 11 mujeres embarazadas con gemelos como se describe en el Ejemplo 1, de acuerdo con los métodos divulgados anteriormente para determinar el sexo de los gemelos y determinar una aneuploidía cromosómica de gemelos y en base a los resultados de las fracciones de ADN fetal libre de células obtenidas en el Ejemplo 1.

1. Determinación del sexo del feto de la mujer embarazada bajo detección

25 Se determinó el sexo de cada feto de 11 mujeres embarazadas bajo detección en base a los resultados de las fracciones de ADN fetal libre de células determinadas en el Ejemplo 1 y de acuerdo con las siguientes etapas:

a) se calculó $fra.cry$ mediante el cromosoma Y;

b) se estimó $fra.tamaño$ mediante las diferencias dentro de una región predeterminada entre la madre y el feto;

c) Estándar de determinación:

I. Si un valor de fra.cry/fra.tamaño es menor que 0.35, ambos fetos de los gemelos son mujeres;

II. Si error! No se encuentra ninguna fuente de referencia. Si el valor de fra.cry/fra.tamaño no es menor que 0.35 ni mayor que 0.7, los gemelos incluyen un feto masculino y un feto femenino;

5 III. Si error! No se encuentra ninguna fuente de referencia. Si el valor de fra.cry/fra.tamaño es mayor que 0.7, ambos fetos de los gemelos son masculinos.

Los resultados se muestran en la siguiente tabla:

Resultados de detección del sexo del feto de 11 mujeres embarazadas bajo detección

Muestra No.	fra.cry	fra.tamaño	fra.cry/fra.tamaño	Resultado de la detección	Resultados de cariotipo
S1	0.131431	0.118	1.113822	[Maculino,Maculino]	[Maculino,Maculino]
S2	-0.00014	0.126652	-0.00114	[Femenino,Femenino]	[Femenino,Femenino]
S3	-0.00019	0.125	-0.00151	[Femenino,Femenino]	[Femenino,Femenino]
S4	0.141033	0.150125	0.939435	[Maculino,Maculino]	[Maculino,Maculino]
S5	0.071202	0.088074	0.808433	[Maculino,Maculino]	[Maculino,Maculino]
S6	0.06227	0.133	0.468195	[Femenino,Maculino]	[Femenino,Maculino]
S7	-0.00044	0.132	-0.0033	[Femenino,Femenino]	[Femenino,Femenino]
S8	-0.00396	0.074	-0.05353	[Femenino,Femenino]	[Femenino,Femenino]
S9	0.185233	0.172	1.076936	[Maculino,Maculino]	[Maculino,Maculino]
S10	0.072432	0.086	0.842233	[Maculino,Maculino]	[Maculino,Maculino]
S11	0.000339	0.092792	0.003653	[Femenino,Femenino]	[Femenino,Femenino]

10 1. Determinación de aneuploidía cromosómica de gemelos por fracciones fetales.

La aneuploidía cromosómica de gemelos de 11 mujeres embarazadas bajo detección se determinó mediante fracciones fetales, en base a los resultados de las fracciones de ADN fetal libre de células determinadas en el Ejemplo 1 y de acuerdo con las siguientes etapas:

a) se calculó fra.cry mediante el cromosoma i (i = 13,18,21);

15 b) se estimó fra.tamaño mediante las diferencias dentro de una región predeterminada entre la madre y el feto;

c) Estándar de determinación:

I. Si error! No se encuentra ninguna fuente de referencia. Si el valor de fra.cry/fra.tamaño es menor que 0.35, el cromosoma i en ambos fetos de los gemelos es normal;

20 II. Si error! No se encuentra ninguna fuente de referencia. Si el valor de fra.cry/fra.tamaño no es menor que 0.35 ni mayor que 0.7, el cromosoma i en un feto de los gemelos es trisoma y en el otro feto de los gemelos es normal;

III. Si error! No se encuentra ninguna fuente de referencia. Si el valor de fra.cry/fra.tamaño es mayor que 0.7, el cromosoma i en ambos fetos de los gemelos es trisoma.

Los resultados se muestran en la siguiente tabla:

25 Resultados de la determinación de aneuploidía cromosómica de gemelos de 11 mujeres embarazadas bajo detección por fracciones fetales

Muestra No.	fra.cr13	fra.cr18	fra.cr21	fra.tamaño	fra.cr13/fra.tamaño	fra.cr18/fra.tamaño	fra.cr21/fra.tamaño	Resultado de la detección	Resultados de cariotipo
S1	0.113979	0.017835	-0.01762	0.118	0.965921	0.151144	-0.14933	[T13,T13]	[T13,T13]
S2	-0.00586	0.004833	0.009971	0.126652	-0.0463	0.03816	0.078728	[normal, normal]	[normal, normal]
S3	0.001131	0.003996	0.121607	0.125	0.009048	0.031968	0.972856	[T21,T21]	[T21,T21]
S4	-0.01175	-0.01269	0.004841	0.150125	-0.07829	-0.08455	0.032246	[normal, normal]	[normal, normal]
S5	2.20E-05	-0.01032	0.011321	0.088074	0.00025	-0.11721	0.128539	[normal, normal]	[normal, normal]
S6	0.040963	-0.00491	-0.00723	0.133	0.307992	-0.03693	-0.05438	[T13,normal]	[T13,normal]
S7	0.005557	-0.01147	0.0754	0.132	0.042098	-0.08688	0.571212	[T21,normal]	[T21,normal]
S8	0.014679	0.031445	-0.01884	0.074	0.198365	0.424932	-0.25462	[T18,normal]	[T18,normal]
S9	0.010402	0.082992	0.007482	0.172	0.060477	0.482512	0.0435	[T18,normal]	[T18,normal]
S10	0.001905	0.074566	-0.00051	0.086	0.022151	0.867052	-0.00592	[T18,T18]	[T18,T18]
S11	0.014591	-0.00375	-0.01728	0.092792	0.157244	-0.04041	-0.1862	[normal, normal]	[normal, normal]

2. Determinación de aneuploidía cromosómica de gemelos de mujer embarazada bajo detección por puntuación T y puntuación L

La aneuploidía cromosómica de gemelos de 11 mujeres embarazadas bajo detección se determinó mediante la puntuación T y la puntuación L, en base a los resultados de las fracciones de ADN fetal libre de células determinadas en el Ejemplo 1 y de acuerdo con las siguientes etapas:

a) Secuenciación del genoma completo (WGS): la muestra bajo detección se sometió a secuenciación del genoma completo utilizando la plataforma de alto rendimiento de Illumina.

b) Se obtuvo información sobre la posición de las lecturas efectivas: las lecturas de una muestra de prueba se mapearon de forma única en la secuencia del genoma de referencia hg19 para obtener información de posición, que corresponde al genoma de referencia, de las lecturas mapeadas de forma única.

c) Se obtuvo el porcentaje de lecturas efectivas: Se obtuvo el porcentaje de lecturas efectivas en cada cromosoma con respecto a las lecturas efectivas totales obtenidas en b).

d) Corrección de GC:

Los contenidos de ER y GC de cada cromosoma se ajustaron al utilizar datos conocidos obtenidos de muestras normales para obtener una fórmula de relación: $ER_i = F(GC_i) + \varepsilon_i$, y un valor medio \overline{ER}_i de ER se calculó. Para analizar una muestra, se calculó un valor ER después de la corrección de acuerdo con la fórmula de relación anterior y ER y GC de la muestra.

$$\overline{ER}_{ij} = \overline{ER}_i + \varepsilon_i = \overline{ER}_i + ER_{ij} - f_i(GC_{ij})$$

e) Se calculó la puntuación T: $T_i = (X_i - \mu_i) / \sigma_i$

donde i : un número de serie del cromosoma ($i = 13, 18, 21$);

X_i : porcentaje de lecturas efectivas del cromosoma i en una muestra analítica;

μ_i : porcentaje promedio de lecturas efectivas del cromosoma i seleccionado como un sistema de referencia en una base de datos de referencia;

σ_i : una desviación estándar de los porcentajes de lecturas efectivas del cromosoma i seleccionado como un sistema de referencia en la base de datos de referencia;

f) Se calculó la puntuación L:

La puntuación L del cromosoma i se determinó de acuerdo con la fórmula:

$L_i = \log(d(T_i, a)) / \log(d(T2_i, a))$, donde i representa el número de serie del cromosoma, $T2_i = (X_i - \mu_i * (1 - fra/2)) / \sigma_i$; $d(T_i, a)$ y $d(T2_i, a)$ representan la función de densidad de probabilidad de distribución t, a representa el grado de libertad, fra representa la primera fracción de ADN fetal libre de células determinada por el método en el Ejemplo 1.

g) Se graficó un diagrama de cuatro cuadrantes con T como coordenada vertical y L como coordenada horizontal mediante la zonificación con una primera línea recta donde $T = 3$ y una segunda línea recta donde $L = 0.8$ (se determinó una muestra con fracción fetal $< 5\%$ para no cumplir con el control de calidad), los detalles se describieron de la siguiente manera:

I. Se determinó que ambos fetos de los gemelos tenían un trisoma si se determinaba que una muestra bajo detección tenía una puntuación T y una puntuación L ($T > 3$, $L > 0.8$) que cae en un primer cuadrante;

II. Se determinó que un feto de los gemelos tenía un trisoma y que el otro feto de los gemelos era normal si se determinaba que una muestra bajo detección tenía una puntuación T y una puntuación L ($T > 3$, $L \leq 0.8$) que cae en un segundo cuadrante;

III. Se determinó que ambos fetos de los gemelos eran normales si se determinaba que una muestra bajo detección tenía una puntuación T y una puntuación L ($T \leq 3$, $L \leq 0.8$) que cae en un tercer cuadrante;

IV. Se determinó que los gemelos habían dado lugar a una fracción fetal baja si se determinaba que una muestra bajo detección tenía una puntuación T y una puntuación L ($T \leq 3$, $L > 0.8$) que cae en un cuarto cuadrante; dicha muestra no cumplió con el control de calidad.

Los diagramas de cuatro cuadrantes graficados con puntuaciones T y puntuaciones L para las 11 muestras probadas se muestran en las Fig. 14 al finalizar. Se puede ver a partir de estas cifras que la aneuploidía cromosómica de los

gemelos de las 11 mujeres embarazadas bajo detección, determinada por la puntuación T y la puntuación L, está en conformidad con la determinada en la etapa 2 por la fracción fetal.

Ejemplo 3: Detección de quimeras

5 En los siguientes ejemplos, se utilizó una mezcla de fragmentos de ADN obtenidos de un tejido de aborto y plasma obtenido de una mujer no embarazada en cierta proporción para simular una muestra obtenida de una mujer embarazada. Se detectó la anomalía del número de cromosomas (trisoma, monosoma completo, quimera trisoma, quimera monosoma) del feto (masculino) de acuerdo con el siguiente método, que incluye las siguientes etapas.

1) Secuenciación del genoma completo (WGS): la muestra bajo detección se sometió a secuenciación del genoma completo utilizando la plataforma de alto rendimiento.

10 2) Se obtuvo información de posición de lecturas efectivas: las lecturas de una muestra de prueba se alinearon de forma única con la secuencia del genoma de referencia para obtener información de posición, que corresponde al genoma de referencia, de las lecturas mapeadas de forma única.

15 3) Se obtuvo la fracción de lecturas mapeadas de forma única en cada cromosoma y el porcentaje del número de guaninas (bases G) y citosinas (bases C) de lecturas mapeadas de forma única con respecto al número total de bases en cada cromosoma: se obtuvieron el porcentaje del número de lecturas efectivas en cada cromosoma en una muestra bajo detección al número total de lecturas efectivas del mismo y el porcentaje del número de bases G y C en lecturas efectivas en cada cromosoma al número total de bases del mismo, utilizando información de posición e información de base de lecturas efectivas.

4) Se calculó la fracción de ADN mediante el cromosoma i , que se calculó como $fra.cri$,

20 $fra.cri = 2^{*}(crt.ER\%/ajuste.cri.ER\%-1)*100\%$, donde $fra.cri$ representa una fracción de ADN fetal libre de células, i representa un número de serie de un cromosoma predeterminado e i es cualquier entero en el rango de 1 a 22;

$cri.ER\%$: $ER\%$ (abreviatura de tasa de lecturas efectivas, porcentaje de lecturas mapeadas de forma única) del cromosoma i en una muestra;

$ajuste.cri.ER\%$: un valor teórico de $ER\%$ del cromosoma i en una muestra normal;

25 5) Se calculó la fracción fetal mediante el cromosoma Y, que se calculó como $fra.cry$;

$fra.cry = (cry.ER\%-Mujer.cry.ER\%)/(Hombre.cry.ER\%-Mujer.cry.ER\%)*100\%$, donde $cry.ER\%$: $ER\%$ (abreviatura de tasa de lecturas efectivas, porcentaje de lecturas mapeadas de forma única) del cromosoma Y en una muestra bajo detección;

30 $Mujer.cry.ER\%$: un promedio de $ER\%$ del cromosoma Y en una muestra bajo detección obtenida de una mujer embarazada con un feto femenino;

$Hombre.cry.ER\%$: un promedio de $ER\%$ del cromosoma Y en una muestra bajo detección obtenida de un hombre.

6) Estándar de determinación:

I. Si $fra.cri/fra.cry < A1$ ($A1$ es una cierta constante y $A1 > -1$, tal como -0.85), el cromosoma i de un feto es un monosoma completo;

35 II. Si $fra.cri/fra.cry \in [A1, A2]$ ($A1$ y $A2$ son ciertas constantes y $-1 < A1 < A2 < 0$, tal como $[-0.85, -0.3]$), el cromosoma i del feto es una quimera monosoma;

III. Si $fra.cri/fra.cry \in [A2, A3]$ ($A2$ y $A3$ son ciertas constantes y $A2 < 0 < A3$, tal como $[-0.3, 0.3]$), el cromosoma i del feto es normal;

40 IV. Si $fra.cri/fra.cry \in [A3, A4]$ ($A3$ y $A4$ son ciertas constantes y $0 < A3 < A4 < 1$, tal como $[0.3, 0.85]$), el cromosoma i del feto es una quimera trisoma;

V. Si $fra.cri/fra.cry > A4$ ($A4$ es una cierta constante, tal como 0.85), el cromosoma i de un feto es un trisoma completo.

3.1 Se investigó la aneuploidía cromosómica en 19 muestras (M1-M19), cada una de las cuales era una mezcla de fragmentos de ADN obtenidos de un tejido de aborto y plasma obtenido de una mujer no embarazada (detalles mostrados en la tabla A) y 2 muestras de plasma (N1, N2), respectivamente, obtenidas de 2 mujeres embarazadas con un feto masculino (el feto masculino en una mujer embarazada tenía una quimera T18 y el feto masculino en la otra mujer embarazada tenía una quimera T21).

1) Recolección y tratamiento de muestras

Se extrajeron 2 ml de sangre periférica para la separación del plasma.

2) Construcción de biblioteca

La biblioteca se construyó de acuerdo con los requisitos de construcción de bibliotecas de plasma de Complete Genomics Inc.

3) Secuenciación

5 El proceso de secuenciación se llevó a cabo siguiendo estrictamente el procedimiento operativo estándar de Complete Genomics Inc.

4) Análisis de datos

10 a) Secuenciación del genoma completo (WGS): la muestra bajo detección se sometió a secuenciación del genoma completo utilizando la plataforma de alto rendimiento (las longitudes de todas las moléculas de ADN libres de células se obtuvieron mediante secuenciación de un solo extremo o secuenciación de extremos pareados, lo cual fue importante. Se requirió secuenciar toda la molécula de ADN libre de células si se realizaba mediante secuenciación de un solo extremo)

15 b) Se obtuvo información de posición de lecturas efectivas: las lecturas de una muestra de prueba se alinearon de forma única con la secuencia del genoma de referencia para obtener información de posición, que corresponde al genoma de referencia, de las lecturas mapeadas de forma única.

c) Se obtuvo el porcentaje de lecturas efectivas: Se obtuvo el porcentaje de lecturas efectivas en cada cromosoma con respecto a las lecturas efectivas totales obtenidas en b).

d) Corrección de GC:

20 Los contenidos de ER y GC de cada cromosoma se ajustaron al utilizar datos conocidos obtenidos de muestras normales para obtener una fórmula de relación: $ER_i = F_i(GC_i) + \varepsilon_i$, y se calculó un valor medio \overline{ER}_i de ER. Para analizar una muestra, se calculó un valor ER después de la corrección de acuerdo con la fórmula de relación anterior y ER y GC de la muestra.

$$\overline{ER}_{ij} = \overline{ER}_i + \varepsilon_i = \overline{ER}_i + ER_{ij} - f_i(GC_{ij})$$

e) se calculó fra.cri mediante el cromosoma i (i = 13, 18, 21);

25 f) se calculó fra.cry mediante el cromosoma Y;

g) se calculó fra.tamaño mediante el método para determinar la fracción de ADN fetal libre de células de la presente divulgación;

h) Se calculó la puntuación T: $T_i = (X_i - \mu_i) / \sigma_i$,

donde i: un número de serie del cromosoma (i = 1, 2... 22);

30 X_i : un porcentaje de lecturas efectivas del cromosoma i en una muestra analítica;

μ_i : porcentaje promedio de lecturas efectivas del cromosoma i seleccionado como un sistema de referencia en una base de datos de referencia;

σ_i : una desviación estándar de los porcentajes de lecturas efectivas del cromosoma i seleccionado como un sistema de referencia en la base de datos de referencia;

35 i) Se calculó la puntuación L:

En primer lugar se calculó el valor de T2: $T2_i = (X_i - \mu_i * (1 + fra/2)) / \sigma_i$,

Luego se calculó la puntuación L: $L_i = \log(d(T_i, a)) / \log(d(T2_i, a))$,

donde $d(T_i, a)$ y $d(T2_i, a)$ representan la función de densidad de probabilidad de distribución t, a representa el grado de libertad, fra representa fra.cry o fra.tamaño.

40 ♦ La aneuploidía cromosómica de una muestra bajo detección se determinó mediante fra.cry que se estimó por el cromosoma Y (como se muestra en la tabla A).

a) Estándar de determinación:

I. Si fra.cri/fra.cry < -0.85, el cromosoma i de un feto es un monosoma completo;

- II. Si $fra.cri/fra.cry \in [-0.85, -0.3]$, el cromosoma i de un feto es una quimera monosoma;
- III. Si $fra.cri/fra.cry \in [-0.3, 0.3]$, el cromosoma i de un feto es normal;
- IV. Si $fra.cri/fra.cry \in [0.3, 0.85]$, el cromosoma i de un feto es una quimera trisoma;
- V. Si $fra.cri/fra.cry > 0.85$, el cromosoma i de un feto es un trisoma completo.
- 5 ♦ La aneuploidía cromosómica de una muestra bajo detección se determinó mediante la puntuación T y la puntuación L (se estimó $fra.cry$ mediante el cromosoma Y) (como se muestra en la tabla B)
- a) Se graficó un diagrama de cuatro cuadrantes en base a las puntuaciones T y puntuaciones L;
- b) Cuando $T \leq 0$, se graficó un diagrama de cuatro cuadrantes con T como coordenada vertical y el valor absoluto de L como coordenada horizontal mediante zonificación con una línea recta donde $T = 3$ y una línea recta donde $L = 1$ (se determinó que una muestra con fracción fetal $< 5\%$ no cumplía con el control de calidad),
- 10 I. Se determinó que el feto tenía un monosoma o quimera monosoma si se determinaba que una muestra bajo detección tenía una puntuación T y una puntuación L ($T > 3, L > 1$) que cae en un primer cuadrante;
- II. Se determinó que el feto tenía una quimera monosoma si se determinaba que una muestra bajo detección tenía una puntuación T y una puntuación L ($T > 3, L \leq 1$) que cae en un segundo cuadrante;
- 15 III. Se determinó que el feto era normal si se determinaba que una muestra bajo detección tenía una puntuación T y una puntuación L ($T \leq 3, L \leq 1$) que cae en un tercer cuadrante;
- IV. Se determinó que el feto había dado lugar a una fracción fetal baja si se determinaba que una muestra bajo detección tenía una puntuación T y una puntuación L ($T \leq 3, L > 1$) que cae en un cuarto cuadrante; dicha muestra no cumplió con el control de calidad.
- 20 cuando $T > 0$, se graficó un diagrama de cuatro cuadrantes con T como coordenada vertical y L como coordenada horizontal mediante la zonificación con una línea recta donde $T = 3$ y una línea recta donde $L = 0.8$ (se determinó que una muestra con fracción fetal $< 5\%$ no cumplía con el control de calidad),
- I. Se determinó que el feto tenía un trisoma o una quimera trisoma si se determinaba que una muestra bajo detección tenía una puntuación T y una puntuación L ($T > 3, L > 1$) que cae en un primer cuadrante;
- 25 II. Se determinó que el feto tenía una quimera trisoma si se determinaba que una muestra bajo detección tenía una puntuación T y una puntuación L ($T > 3, L \leq 1$) que cae en un segundo cuadrante;
- III. Se determinó que el feto era normal si se determinaba que una muestra bajo detección tenía una puntuación T y una puntuación L ($T \leq 3, L \leq 1$) que cae en un tercer cuadrante;
- 30 IV. Se determinó que el feto había dado lugar a una fracción fetal baja si se determinaba que una muestra bajo detección tenía una puntuación T y una puntuación L ($T \leq 3, L > 1$) que cae en un cuarto cuadrante; dicha muestra no cumplió con el control de calidad.
- ♦ La aneuploidía cromosómica de una muestra bajo detección se determinó mediante $fra.tamaño$ que se estimó mediante el método para determinar la fracción de ADN fetal libre de células de la presente divulgación (como se muestra en la tabla C)
- 35 b) Estándar de determinación:
- I. Si $fra.cri/fra.cry < -0.85$, el cromosoma i de un feto es un monosoma completo;
- II. Si $fra.cri/fra.cry \in [-0.85, -0.3]$, el cromosoma i de un feto es una quimera monosoma;
- III. Si $fra.cri/fra.cry \in [-0.3, 0.3]$, el cromosoma i de un feto es normal;
- IV. Si $fra.cri/fra.cry \in [0.3, 0.85]$, el cromosoma i de un feto es una quimera trisoma;
- 40 V. Si $fra.cri/fra.cry > 0.85$, el cromosoma i de un feto es un trisoma completo
- ♦ La aneuploidía cromosómica de una muestra bajo detección se determinó mediante la puntuación T y la puntuación L (se estimó $fra.tamaño$ mediante el método para determinar la fracción de ADN fetal libre de células de la presente divulgación) (como se muestra en la tabla D, Fig.2)
- a) Se graficó un diagrama de cuatro cuadrantes en base a las puntuaciones T y puntuaciones L;

b) Cuando $T \leq 0$, se graficó un diagrama de cuatro cuadrantes con T como coordenada vertical y el valor absoluto de L como coordenada horizontal mediante zonificación con una línea recta donde $T = 3$ y una línea recta donde $L = 1$ (se determinó que una muestra con fracción fetal $< 5\%$ no cumplía con el control de calidad),

5 I. Se determinó que el feto tenía un monosoma o una quimera monosoma si se determinaba que una muestra bajo detección tenía una puntuación T y una puntuación L ($T > 3, L > 1$) que cae en un primer cuadrante;

II. Se determinó que el feto tenía una quimera monosoma si se determinaba que una muestra bajo detección tenía una puntuación T y una puntuación L ($T > 3, L \leq 1$) que cae en un segundo cuadrante;

III. Se determinó que el feto era normal si se determinaba que una muestra bajo detección tenía una puntuación T y una puntuación L ($T \leq 3, L \leq 1$) que cae en un tercer cuadrante;

10 IV. Se determinó que el feto tenía una fracción fetal baja si se determinaba que una muestra bajo detección tenía una puntuación T y una puntuación L ($T \leq 3, L > 1$) que cae en un cuarto cuadrante; dicha muestra no cumplió con el control de calidad;

15 Cuando $T > 0$, se graficó un diagrama de cuatro cuadrantes con T como coordenada vertical y L como coordenada horizontal mediante la zonificación con una línea recta donde $T = 3$ y una línea recta donde $L = 1$ (se determinó que una muestra con fracción fetal $< 5\%$ no cumplía con el control de calidad),

I. Se determinó que el feto tenía un trisoma o una quimera trisoma si se determinaba que una muestra bajo detección tenía una puntuación T y una puntuación L ($T > 3, L > 1$) que cae en un primer cuadrante;

II. Se determinó que el feto tenía una quimera trisoma si se determinaba que una muestra bajo detección tenía una puntuación T y una puntuación L ($T > 3, L \leq 1$) que cae en un segundo cuadrante;

20 III. Se determinó que el feto era normal si se determinaba que una muestra bajo detección tenía una puntuación T y una puntuación L ($T \leq 3, L \leq 1$) que cae en un tercer cuadrante;

IV. Se determinó que el feto había dado lugar a una fracción fetal baja si se determinaba que una muestra bajo detección tenía una puntuación T y una puntuación L ($T \leq 3, L > 1$) que cae en un cuarto cuadrante; dicha muestra no cumplió con el control de calidad.

25 En esta realización, una muestra negativa era una muestra de plasma obtenida de una mujer normal que no estaba embarazada; se preparó una muestra positiva al mezclar fragmentos de ADN que se obtuvieron al romper aleatoriamente ADN de un tejido de aborto de acuerdo con un tamaño que variaba de 150 pb a 200 pb y plasma obtenido de una mujer normal que no estaba embarazada; (T21, T18 representan cada uno un feto masculino; T13 representa un feto femenino); se preparó una muestra quimérica positiva al mezclar fragmentos de ADN de tejido placentario (que se obtuvieron al romper aleatoriamente ADN de un tejido placentario de acuerdo con un tamaño que varía entre 150 pb y 200 pb), fragmentos de ADN de estirpe celular china (que se obtuvieron al romper aleatoriamente ADN de la estirpe celular china de acuerdo con un tamaño que varía de 150 pb a 200 pb) y plasma obtenido de una mujer normal; (T21, T18 representan cada uno un feto masculino; T13 representa un feto femenino).

Muestra No.	Cariotipo del tejido de aborto	Fracción fetal	Relación quimérica	Forma de expresión
M1	trisoma 13	3.5%	0	T13-3.5%
M2	trisoma 13	5%	0	T13-5%
M3	trisoma 13	8%	0	T13-8%
M4	trisoma 13	8%	0	T13-8%
M5	trisoma 18	10%	30%	T18-10%-30%
M6	trisoma 18	10%	70%	T18-10%-70%
M7	trisoma 18	10%	70%	T18-10%-70%
M8	trisoma 18	10%	70%	T18-10%-70%
M9	trisoma 18	3.5%	0	T18-3.5%
M10	trisoma 18	5%	0	T18-5%
M11	trisoma 18	5%	0	T18-5%
M12	trisoma 18	8%	0	T18-8%
M13	trisoma 18	10%	30%	T21-10%-30%
M14	trisoma 18	10%-	70%	T21-10%-70%
M15	trisoma 18	10%	70%	T21-10%-70%
M16	trisoma 18	10%	70%	T21-10%-70%
M17	trisoma 18	3.5%	0	T21-3.5%
M18	trisoma 18	5%	0	T21-5%
M19	trisoma 18	8%	0	T21-8%

ES 2 903 103 T3

Muestra No.	Resultados de cariotipo
N1	47,XN+18[3]/46,XN[20]
N2	47,XN+21[301]/46,XN[18]

Tabla A Detección de aneuploidia cromosómica por fracción de ADN mixto estimada por el cromosoma Y

Muestra No.	fra.cr13	fra.cr18	fra.cr21	fra.cry	fra.cr13/fra.cry	fra.cr18/fra.cry	fra.cr21/fra.cry	Resultado de la detección
M1	0.03531	-4.8E-06	-0.00742	0.00069				T13-3.5%
M2	0.04731	-0.01207	0.007764	-0.00002				T13-5%
M3	0.08136	-0.00477	-0.00129	-0.0002				T13-8%
M4	0.0796	-0.00738	-0.0013	-0.00157				T13-8%
M5	-0.01079	0.03401	-0.00442	0.105215	-0.10259	0.323243	-0.04197	T18-10%-30%
M6	-0.00192	0.069573	-0.00043	0.090435	-0.02127	0.769319	-0.00471	T18-10%-70%
M7	0.001756	0.058363	-0.00968	0.087785	0.020009	0.664844	-0.11022	T18-10%-70%
M8	-0.01081	0.082063	-0.00917	0.106115	-0.1019	0.773343	-0.08638	T18-10%-70%
M9	0.007816	0.0341	-0.00038	0.03723	0.209951	0.915928	-0.01009	T18-3.5%
M10	0.013116	0.0512	-0.00607	0.04843	0.270834	1.057196	-0.12525	T18-5%
M11	0.000956	0.0513	-0.01176	0.0566	0.016899	0.90636	-0.2077	T18-5%
M12	-0.00237	0.07693	-0.004	0.08877	-0.02674	0.866622	-0.04501	T18-8%
M13	0.001246	-0.01028	0.04323	0.110455	0.011285	-0.09311	0.391381	T21-10%-30%
M14	0.000266	-0.01586	0.06554	0.092805	0.002871	-0.17095	0.706212	T21-10%-70%
M15	-0.00041	0.004795	0.07461	0.110035	-0.00376	0.043578	0.678057	T21-10%-70%
M16	-0.00235	-0.00759	0.06985	0.097155	-0.02422	-0.07817	0.718954	T21-10%-70%
M17	0.005776	0.010345	0.0481	0.03478	0.166086	0.297445	1.382979	T21-3.5%
M18	0.007786	0.006345	0.0557	0.04853	0.160447	0.130747	1.147744	T21-5%
M19	0.009846	-0.00828	0.07449	0.06855	0.143639	-0.12086	1.086652	T21-8%
N1	0.007053	0.043212	0.000938	0.110337	0.063922	0.391637	0.008501	47.XN+18[3]/46.XN[20]
N2	-0.01105	0.002921	0.081665	0.167834	-0.06581	0.017404	0.486582	47.XN+21[30]/46.XN[18]

Nota:

fra.cr13: representa una fracción de ADN mezclada estimada mediante el cromosoma 13;
fra.cr18: representa una fracción de ADN mezclada estimada mediante el cromosoma 18;
fra.cr21: representa una fracción de ADN mezclada estimada mediante el cromosoma 21;
fra.cry: representa una fracción de ADN mezclada calculada por el cromosoma Y;
T21-10%-30%: fragmentos de ADN obtenidos de una estirpe celular con trisoma 21 mezclado con plasma femenino de acuerdo con una fracción de 10% y una relación quimérica de 30%.

Tabla B Detección de aneuploidía cromosómica por puntuación T y puntaje L (la fracción de ADN mixto fue estimada por el cromosoma Y)

Muestra No.	T.cr13	L.cr13	T.cr18	L.cr18	T.cr21	L.cr21	fra.cry	Resultado de la detección
M1	4.345751		-0.00053		-0.60057		0.00069	T13-3.5%
M2	5.436515		-1.216		0.52145		-0.00002	T13-5%
M3	10.46607		-0.47762		-0.09072		-0.0002	T13-8%
M4	9.126057		-0.70384		-0.08565		-0.00157	T13-8%
M5	-1.43378	0.018171	4.16099	0.292408	-0.39025	0.023403	0.105215	T18-10%-30%
M6	-0.20845	0.015392	7.977908	12.56388	-0.03313	0.031699	0.090435	T18-10%-70%
M7	0.210445	0.015766	7.151443	4.550099	-0.82976	0.035673	0.087785	T18-10%-70%
M8	-1.19968	0.019721	8.763087	10.86877	-0.7035	0.030472	0.106115	T18-10%-70%
M9	0.963351	0.189495	3.96625	9.33211	-0.03142	0.162273	0.03723	T18-3.5%
M10	1.8534	0.251063	6.429505	19.56264	-0.51602	0.082144	0.04843	T18-5%
M11	0.112697	0.031697	6.659763	23.63495	-1.01942	0.072816	0.0566	T18-5%
M12	-0.24379	0.016926	8.332999	25.4093	-0.2922	0.035987	0.08877	T18-8%
M13	0.164157	0.010591	-1.22048	0.017938	3.76578	0.491369	0.110455	T21-10%-30%
M14	0.037473	0.013915	-1.90428	0.036927	5.667693	4.554072	0.092805	T21-10%-70%
M15	-0.05039	0.011486	0.536781	0.017359	5.897363	3.910694	0.110035	T21-10%-70%
M16	-0.28883	0.01242	-0.86821	0.018234	6.167815	5.406755	0.097155	T21-10%-70%
M17	0.581767	0.19567	1.077291	0.427888	3.582341	4.827084	0.03478	T21-3.5%
M18	0.93815	0.145201	0.654071	0.129238	3.958377	8.227322	0.04853	T21-5%
M19	1.201005	0.051788	-1.03581	0.032569	6.795784	21.00502	0.06855	T21-8%
N1	0.703062	0.022374	5.418144	0.435182	0.081371	0.02066	0.110337	47.XN+18 3 46.XN 20
N2	-0.97092	0.012246	0.355485	0.005614	7.170205	0.903486	0.167834	47.XN+21 30 46.XN 18

Nota:

T.cr13: puntuación T del cromosoma 13;

L.cr13: puntuación L del cromosoma 13;

T.cr18: puntuación T del cromosoma 18;

L.cr18: puntuación L del cromosoma 18;

T.cr21: puntuación T del cromosoma 21;

L.cr21: puntuación L del cromosoma 21;

fra.cry: fracción de ADN fetal mezclado estimada mediante el cromosoma Y;

T21-10%-30%: fragmentos de ADN obtenidos de una estirpe celular con trisoma 21 mezclado con plasma femenino de acuerdo con una fracción de 10% y una relación química de 30%

Tabla C Detección de aneuploidia cromosómica mediante fra. tamaño

Muestra No.	fra.cr13	fra.cr18	fra.cr21	fra.tamaño	fra.cr13/fra.tamaño	fra.cr18/fra.tamaño	fra.cr21/fra.tamaño	Resultado de la detección
M1	0.03531	-4.8E-06	-0.00742	0.0369	0.956911	-0.00013	-0.20097	T13-3.5%
M2	0.04731	-0.01207	0.007764	0.05519	0.857221	-0.21879	0.140682	T13-5%
M3	0.08136	-0.00477	-0.00129	0.08462	0.961475	-0.05643	-0.01519	T13-8%
M4	0.0796	-0.00738	-0.0013	0.07332	1.085652	-0.10072	-0.01767	T13-8%
M5	-0.01079	0.03401	-0.00442	0.1071	-0.10078	0.317554	-0.04123	T18-10%-30%
M6	-0.00192	0.069573	-0.00043	0.0928	-0.02073	0.749713	-0.00459	T18-10%-70%
M7	0.001756	0.058363	-0.00968	0.1058	0.016602	0.551638	-0.09145	T18-10%-70%
M8	-0.01081	0.082063	-0.00917	0.1233	-0.0877	0.665558	-0.07434	T18-10%-70%
M9	0.007816	0.0341	-0.00038	0.0432	0.180937	0.789352	-0.0087	T18-3.5%
M10	0.013116	0.0512	-0.00607	0.05688	0.230599	0.900141	-0.10664	T18-5%
M11	0.000956	0.0513	-0.01176	0.0438	0.021838	1.171233	-0.2684	T18-5%
M12	-0.00237	0.07693	-0.004	0.06726	-0.03529	1.14377	-0.05941	T18-8%
M13	0.001246	-0.01028	0.04323	0.0705	0.017681	-0.14588	0.613191	T21-10%-30%
M14	0.000266	-0.01586	0.06554	0.1056	0.002524	-0.15024	0.620644	T21-10%-70%
M15	-0.00041	0.004795	0.07461	0.1205	-0.00343	0.039794	0.61917	T21-10%-70%
M16	-0.00235	-0.00759	0.06985	0.1165	-0.0202	-0.06519	0.599571	T21-10%-70%
M17	0.005776	0.010345	0.0481	0.0409	0.141234	0.252938	1.176039	T21-3.5%
M18	0.007786	0.006345	0.0557	0.04543	0.171395	0.139669	1.226062	T21-5%
M19	0.009846	-0.00828	0.07449	0.07848	0.125465	-0.10557	0.949159	T21-8%
N1	0.007053	0.043212	0.000938	0.102547	0.068778	0.421387	0.009147	47,XN+18[3]/46,XN[20]
N2	-0.01105	0.002921	0.081665	0.198228	-0.05572	0.014736	0.411975	47,XN+21[30]/46,XN[18]

Nota:

fra.cr13: representa una fracción de ADN mezclada estimada mediante el cromosoma 13;

fra.cr18: representa una fracción de ADN mezclada estimada mediante el cromosoma 18;

fra.cr21: representa una fracción de ADN mezclada estimada mediante el cromosoma 21;

fra.tamaño: representa una fracción de ADN mezclada estimada mediante el método de la presente divulgación;

T21-10%-30%: fragmentos de ADN obtenidos de una estirpe celular con trisoma 21 mezclado con plasma femenino de acuerdo con una fracción de 10% y una relación quimérica de 30%.

Tabla D Detección de aneuploidía cromosómica por puntuación T y puntuación L (fra. tamaño)

Muestra No.	T.cr13	L.cr13	T.cr18	L.cr18	T.cr21	L.21	fra.tamaño	Resultado de la detección
M1	4.345751	10.86269	-0.00053	0.06264	-0.60057	0.110064	0.0369	T13-3.5%
M2	5.436515	13.40864	-1.216	0.040421	0.52145	0.080194	0.05519	T13-5%
M3	10.46607	36.3034	-0.47762	0.013724	-0.09072	0.026097	0.08462	T13-8%
M4	9.126057	29.68988	-0.70384	0.019292	-0.08565	0.033903	0.07332	T13-8%
M5	-1.43378	0.015183	4.16099	0.216731	-0.39025	0.019427	0.1071	T18-10%-30%
M6	-0.20845	0.01467	7.977908	10.52499	-0.03313	0.030232	0.0928	T18-10%-70%
M7	0.210445	0.011014	7.151443	1.813942	-0.82976	0.026168	0.1058	T18-10%-70%
M8	-1.19968	0.015259	8.763087	4.176647	-0.7035	0.023665	0.1233	T18-10%-70%
M9	0.963351	0.136353	3.96625	6.810597	-0.03142	0.126392	0.0432	T18-3.5%
M10	1.8534	0.166895	6.429505	20.92116	-0.51602	0.062989	0.05688	T18-5%
M11	0.112697	0.051911	6.659763	13.06329	-1.01942	0.107019	0.0438	T18-5%
M12	-0.24379	0.028322	8.332999	19.64862	-0.2922	0.058661	0.06726	T18-8%
M13	0.164157	0.025146	-1.22048	0.037995	3.76578	2.372928	0.0705	T21-10%-30%
M14	0.037473	0.010899	-1.90428	0.030091	5.667693	2.443191	0.1056	T21-10%-70%
M15	-0.05039	0.009693	0.536781	0.014533	5.897363	2.522845	0.1205	T21-10%-70%
M16	-0.28883	0.008906	-0.86821	0.013317	6.167815	2.195736	0.1165	T21-10%-70%
N1	0.703062	0.030365	5.418144	0.576578	0.081371	0.02724	0.102547	47,XN+18[3]/46,XN[20]
N2	-0.97092	0.013866	0.355485	0.007248	7.170205	0.553012	0.198228	47,XN+21[30]/46,XN[18]

Nota:

T.cr13: puntuación T del cromosoma 13;

L.cr13: puntuación L del cromosoma 13;

T.cr18: puntuación T del cromosoma 18;

L.cr18: puntuación L del cromosoma 18;

T.cr21: puntuación T del cromosoma 21;

L.cr21: puntuación L del cromosoma 21;

fra.tamaño: fracción de ADN mezclado estimada mediante el método de la presente divulgación;

T21-10%-30%: fragmentos de ADN obtenidos de una estirpe celular con trisoma 21 mezclado con plasma femenino de acuerdo con una fracción de 10% y una relación quimérica de 30%.

5 La referencia en esta especificación a “una realización”, “algunas realizaciones”, “una realización”, “otro ejemplo”, “un ejemplo”, “un ejemplo específico” o “algunos ejemplos” significa que un rasgo, estructura, material o característica descritos en relación con la realización o ejemplo se incluye en al menos una realización o ejemplo de la presente divulgación. Por lo tanto, las apariciones de expresiones tales como “en algunas realizaciones”, “en una realización”, “en una realización”, “en otro ejemplo”, “en un ejemplo”, “en un ejemplo específico” o “en algunos ejemplos”, en varios lugares a lo largo de esta especificación no se refieren necesariamente a la misma realización o ejemplo de la presente divulgación. Adicionalmente, los rasgos, estructuras, materiales o características particulares se pueden combinarse de cualquier manera adecuada en una o más realizaciones o ejemplos.

REIVINDICACIONES

1. Un método para determinar la fracción de ácidos nucleicos fetales libres de células en una muestra de sangre periférica de una mujer embarazada, que comprende:

5 (i) realizar la secuenciación sobre ácidos nucleicos libres de células contenidos en la muestra para obtener un resultado de la secuenciación que consiste en una pluralidad de datos de secuenciación;

(ii) determinar el número de ácidos nucleicos libres de células de una longitud que cae en un rango predeterminado en la muestra en base al resultado de la secuenciación; y

10 (iii) determinar la fracción de ácidos nucleicos fetales libres de células en la muestra en base al número de ácidos nucleicos libres de células de una longitud que cae en el rango predeterminado, en el que se determina dicho rango predeterminado antes de la etapa (i) mediante las siguientes etapas:

(a) determinar las longitudes de los ácidos nucleicos libres de células en una pluralidad de muestras de sangre periférica de control de mujeres embarazadas en cada una de las cuales se conoce la fracción de ácidos nucleicos fetales libres de células;

15 (b) establecer una pluralidad de rangos de longitud candidatos y determinar el porcentaje de ácidos nucleicos libres de células obtenidos de cada una de la pluralidad de muestras de control presentes en cada rango de longitud candidato;

20 (c) determinar un coeficiente de correlación para cada rango de longitud candidato en base al porcentaje de ácidos nucleicos libres de células obtenidos de cada una de la pluralidad de muestras de control presentes en cada rango de longitud candidato y la fracción conocida de ácidos nucleicos fetales libres de células en cada una de las muestras de control; y

(d) determinar al menos un rango de longitud candidato o una combinación de los rangos de longitud candidatos como el rango predeterminado en base al coeficiente de correlación máximo.

25 2. Un método de acuerdo con la reivindicación 1, en el que determinar el número de ácidos nucleicos libres de células de una longitud que cae en el rango predeterminado en la muestra en base al resultado de la secuenciación comprende adicionalmente:

(a) alinear el resultado de la secuenciación con un genoma de referencia, para construir un conjunto de datos que consiste en una pluralidad de lecturas mapeadas de forma única, donde cada lectura en el conjunto de datos se puede mapear a una única posición del genoma de referencia;

30 (b) determinar la longitud del ácido nucleico libre de células que corresponde a cada lectura mapeada de forma única en el conjunto de datos; y

(c) determinar el número de ácidos nucleicos libres de células de una longitud que cae en el rango predeterminado.

3. Un método de acuerdo con la reivindicación 2, en el que la determinación de la longitud del ácido nucleico libre de células que corresponde a cada lectura mapeada de forma única en el conjunto de datos comprende adicionalmente:

35 determinar la longitud de cada lectura mapeada de forma única en el genoma de referencia como la longitud del ácido nucleico libre de células que corresponde a la lectura.

4. Un método de acuerdo con de la reivindicación 2, en el que los ácidos nucleicos libres de células de la muestra se secuencian mediante secuenciación de extremos emparejados y la determinación de la longitud del ácido nucleico libre de células que corresponde a cada lectura mapeada de forma única en el conjunto de datos comprende adicionalmente:

40 determinar la posición que corresponde al genoma de referencia del extremo 5' del ácido nucleico libre de células en base a los datos de secuenciación para el extremo 5' de cada lectura mapeada de forma única obtenida en la secuenciación de extremos emparejados;

45 determinar la posición que corresponde al genoma de referencia del extremo 3' del ácido nucleico libre de células en base a los datos de secuenciación en el otro extremo de la misma lectura mapeada de forma única obtenida en la secuenciación de extremos emparejados; y

determinar la longitud del ácido nucleico libre de células en base a las posiciones del extremo 5' y el extremo 3' del ácido nucleico libre de células.

50 5. Un método de acuerdo con la reivindicación 1, en el que la determinación de la fracción de ácidos nucleicos fetales libres de células en la muestra en base al número de ácidos nucleicos libres de células de una longitud que cae en el rango predeterminado comprende adicionalmente:

determinar el porcentaje de ácidos nucleicos libres de células presentes en el rango predeterminado en base al número de ácidos nucleicos libres de células de una longitud que cae en el rango predeterminado; y

determinar la fracción de ácidos nucleicos fetales libres de células en la muestra, en base al porcentaje de ácidos nucleicos libres de células presentes en el rango predeterminado de acuerdo con una función predeterminada,

5 en el que la función predeterminada se determina en base a la pluralidad de muestras de control.

6. Un método de acuerdo con la reivindicación 5, en el que la función predeterminada se obtiene mediante las siguientes etapas:

(i) determinar el porcentaje de ácidos nucleicos libres de células obtenidos de cada muestra de control presente en el rango predeterminado; y

10 (ii) ajustar el porcentaje de ácidos nucleicos libres de células obtenido de cada muestra de control presente en el rango predeterminado con la fracción conocida de ácidos nucleicos fetales libres de células para determinar dicha función; opcionalmente, en el que el porcentaje de ácidos nucleicos libres de células obtenido de cada muestra de control presente en el rango predeterminado se ajusta con la fracción conocida de ácidos nucleicos fetales libres de células mediante un ajuste lineal.

15 7. Un método de acuerdo con de la reivindicación 1, en el que el rango predeterminado es de 185 pb a 204 pb.

8. Un método de acuerdo con de la reivindicación 5, en el que la función predeterminada es $d = 0.0334 * p + 1.6657$, donde d representa una fracción de ácidos nucleicos fetales libres de células, y p representa el porcentaje de ácido nucleico libre de células presente en el rango predeterminado.

20 9. Un método de acuerdo con la reivindicación 8, en el que las muestras de control son muestras de sangre periférica obtenidas de mujeres embarazadas con un feto masculino normal, en el que se sabe que la fracción de ácidos nucleicos libres de células que son ácidos nucleicos fetales libres de células se determina por el cromosoma Y.

10. Un método de acuerdo con de la reivindicación 1, en el que dicha fracción de ácidos nucleicos fetales libres de células se determina en una muestra de sangre periférica obtenida de una mujer embarazada con gemelos como una primera fracción de ADN fetal libre de células y que adicionalmente comprende:

25 (a) determinar una segunda fracción de ADN fetal libre de células en base a los datos de secuenciación derivados del cromosoma Y en el resultado de la secuenciación; y

(b) determinar el sexo de los gemelos en base a la primera fracción de ADN fetal libre de células y dicha segunda fracción de ADN fetal libre de células.

30 11. Un método de acuerdo con la reivindicación 10, en el que dicha segunda fracción de ADN fetal libre de células se determina de acuerdo con la siguiente fórmula:

$$Fra.cry = (cry.ER\% - Mujer.cry.ER\%) / (Hombre.cry.ER\% - Mujer.cry.ER\%) * 100\% ,$$

donde $fra.cry$ representa dicha segunda fracción de ADN fetal libre de células, $cry.ER\%$ representa un porcentaje de los datos de secuenciación derivados del cromosoma Y en el resultado de la secuenciación para los datos de secuenciación totales;

35 $Mujer.cry.ER\%$ representa un porcentaje promedio de los datos de secuenciación de ácidos nucleicos libres de células que se alinean con el cromosoma Y en una muestra de sangre periférica obtenida de una mujer embarazada predeterminada para estar con un feto femenino normal para los datos de secuenciación totales de los mismos; y $Hombre.cry.ER\%$ representa un porcentaje promedio de los datos de secuenciación de ácidos nucleicos libres de células derivados del cromosoma Y en una muestra de sangre periférica obtenida de un hombre sano para los datos de secuenciación totales de los mismos.

40 12. Un método de acuerdo con de la reivindicación 10 o la reivindicación 11, en el que determinar el sexo de los gemelos en base a la primera fracción de ADN fetal libre de células y dicha segunda fracción de ADN fetal libre de células comprende adicionalmente:

45 (a) determinar una relación de dicha segunda fracción de ADN fetal libre de células a la primera fracción de ADN fetal libre de células; y

(b) determinar el sexo de los gemelos al comparar la relación determinada en (a) con un primer umbral predeterminado y un segundo umbral predeterminado,

en el que el primer umbral se ha determinado en base a una pluralidad de muestras de control obtenidas de mujeres embarazadas que se sabe que tienen gemelos femeninos y el segundo umbral se ha determinado en base a una pluralidad de muestras de control obtenidas de mujeres embarazadas que se sabe que tienen gemelos varones,

50

y en el que

ambos fetos de los gemelos se determinan como femeninos si la relación de dicha segunda fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células es menor que el primer umbral,

5 ambos fetos de los gemelos se determinan como masculinos si la relación de dicha segunda fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células es mayor que el segundo umbral, y

se determina que los gemelos incluyen un feto masculino y un feto femenino si la relación de dicha segunda fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células es igual al primer umbral o al segundo umbral, o entre el primer umbral y el segundo umbral.

13. Un método de acuerdo con de la reivindicación 12, en el que el primer umbral es 0.35 y el segundo umbral es 0.7.

10 14. Un método de acuerdo con de la reivindicación 1, en el que dicha fracción de ácidos nucleicos fetales libres de células se determina en una muestra de sangre periférica obtenida de una mujer embarazada con gemelos como una primera fracción de ADN fetal libre de células y que adicionalmente comprende:

(a) determinar una tercera fracción de ADN fetal libre de células, en base a los datos de secuenciación derivados de un cromosoma predeterminado en el resultado de la secuenciación; y

15 (b) determinar si los gemelos bajo detección tienen aneuploidía con respecto al cromosoma predeterminado en base a la primera fracción de ADN fetal libre de células y dicha tercera fracción de ADN fetal libre de células.

15. Un método de acuerdo con la reivindicación 14, en el que dicha tercera fracción de ADN fetal libre de células se determina de acuerdo con la siguiente fórmula:

20
$$fra.cri = 2 * (cri.ER\% / ajuste.cri.ER\% - 1) * 100\%$$
, donde *fra.cri* representa dicha tercera fracción de ADN fetal libre de células, *i* representa un número de serie del cromosoma predeterminado e *i* es cualquier entero en el rango de 1 a 22; *cri.ER%* representa un porcentaje de los datos de secuenciación derivados del cromosoma predeterminado en el resultado de la secuenciación a los datos de secuenciación totales; *ajuste.cri.ER%* representa un porcentaje promedio de los datos de secuenciación de ácidos nucleicos libres de células derivados del cromosoma predeterminado en una muestra de sangre periférica obtenida de una mujer embarazada predeterminada para estar con gemelos normales para los datos de secuenciación totales de los mismos.

16. Un método de acuerdo con de la reivindicación 14 o la reivindicación 15, en el que determinar si los gemelos bajo detección tienen aneuploidía con respecto al cromosoma predeterminado en base a la primera fracción de ADN fetal libre de células y dicha tercera fracción de ADN fetal libre de células comprende adicionalmente:

30 (a) determinar una relación de dicha tercera fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células; y

(b) determinar si los gemelos bajo detección tienen aneuploidía con respecto al cromosoma predeterminado al comparar la relación determinada en (a) con un tercer umbral predeterminado y un cuarto umbral predeterminado,

y en el que el tercer umbral se ha determinado en base a una pluralidad de muestras de control obtenidas de mujeres embarazadas con gemelos que se sabe que no tienen aneuploidía con respecto al cromosoma predeterminado y el cuarto umbral se ha determinado en base a una pluralidad de muestras de control obtenidas de mujeres embarazadas con gemelos que se sabe que tienen aneuploidía con respecto al cromosoma predeterminado,

y en el que

40 se determina que ambos fetos de los gemelos no tienen aneuploidía con respecto al cromosoma predeterminado si la relación de dicha tercera fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células es menor que el tercer umbral,

se determina que ambos fetos de los gemelos tienen aneuploidía con respecto al cromosoma predeterminado si la relación de dicha tercera fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células es mayor que el cuarto umbral, y

45 se determina que un feto de los gemelos tiene aneuploidía con respecto al cromosoma predeterminado, mientras que se determina que el otro feto de los gemelos no tiene aneuploidía con respecto al cromosoma predeterminado si la relación de dicha tercera fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células es igual al tercer umbral o al cuarto umbral, o entre el tercer umbral y el cuarto umbral.

17. Un método de acuerdo con de la reivindicación 1, en el que el tercer umbral es 0.35 y el cuarto umbral es 0.7.

50 18. Un método de acuerdo con la reivindicación 14, en el que el cromosoma predeterminado es al menos uno seleccionado de los cromosomas 18, 21 y 23.

19. Un método de acuerdo con de la reivindicación 1, en el que dicha fracción de ácidos nucleicos fetales libres de células se determina en una muestra de sangre periférica obtenida de una mujer embarazada con gemelos como una primera fracción de ADN fetal libre de células y que adicionalmente comprende:

5 (a) determinar una fracción x_i del número de datos de secuenciación derivados del cromosoma i en el resultado de la secuenciación para los datos de secuenciación totales, donde i representa un número de serie del cromosoma, e es cualquier entero en el rango de 1 a 22;

10 (b) determinar una puntuación T del cromosoma i de acuerdo con $T_i = (X_i - \mu_i) / \sigma_i$, donde i representa el número de serie del cromosoma e y es cualquier entero en el rango de 1 a 22, μ_i representa un porcentaje promedio de los datos de secuenciación del cromosoma i seleccionado como un sistema de referencia en una base de datos de referencia para los datos de secuenciación totales del mismo, σ_i representa una desviación estándar de los porcentajes de los datos de secuenciación del cromosoma i seleccionado como un sistema de referencia en la base de datos de referencia para los datos de secuenciación totales del mismo,

15 (c) determinar una puntuación L del cromosoma i de acuerdo con $L_i = \log(d(T_i, a)) / \log(d(T2_i, a))$, donde i representa el número de serie del cromosoma e y es cualquier entero en el rango de 1 a 22, $T2_i = (X_i - \mu_i * (1 + fra/2)) / \sigma_i$; $d(T_i, a)$ y $d(T2_i, a)$ representan la función de densidad de probabilidad de distribución t , donde a representa el grado de libertad, fra representa dicha primera fracción de ADN fetal libre de células,

(d) graficar un diagrama de cuatro cuadrantes con T como coordenada vertical y L como coordenada horizontal mediante zonificación con una primera línea recta donde $T =$ quinto umbral predeterminado y una segunda línea recta donde $L =$ sexto umbral predeterminado,

20 en el que

se determina que ambos fetos de los gemelos tienen un trisoma si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que cae en un primer cuadrante;

25 se determina que un feto de los gemelos tiene un trisoma y se determina que el otro feto de los gemelos es normal si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que cae en un segundo cuadrante;

se determina que ambos fetos de los gemelos son normales si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que cae en un tercer cuadrante, y

se determina que los gemelos han dado lugar a una fracción fetal baja si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que cae en un cuarto cuadrante de tal manera que no se adopta el resultado.

30 20. Un método de acuerdo con de la reivindicación 1, en el que dicha fracción de ácidos nucleicos fetales libres de células se determina en una muestra de sangre periférica obtenida de una mujer embarazada con un feto como una primera fracción de ADN fetal libre de células y que adicionalmente comprende:

(a) determinar una tercera fracción de ADN fetal libre de células en base a los datos de secuenciación derivados de un cromosoma predeterminado en el resultado de la secuenciación; y

35 (b) determinar si el feto bajo detección tiene quimera fetal con respecto al cromosoma predeterminado en base a la primera fracción de ADN fetal libre de células y dicha tercera fracción de ADN fetal libre de células.

21. Un método de acuerdo con la reivindicación 20, en el que dicha tercera fracción de ADN fetal libre de células se determina mediante la siguiente fórmula:

40 $fra.cri = 2 * (cri.ER\% / ajuste.cri.ER\% - 1) * 100\%$, donde $fra.cri$ representa dicha tercera fracción de ADN fetal libre de células, i representa un número de serie del cromosoma predeterminado y e es cualquier entero en el rango de 1 a 22; $cri.ER\%$ representa un porcentaje de los datos de secuenciación derivados del cromosoma predeterminado en el resultado de la secuenciación a los datos de secuenciación totales; $ajuste.cri.ER\%$ representa un porcentaje promedio de los datos de secuenciación de ácidos nucleicos libres de células derivados del cromosoma predeterminado en una muestra de sangre periférica obtenida de una mujer embarazada predeterminada para estar con un feto normal para los datos de secuenciación totales de los mismos.

45 22. Un método de acuerdo con de la reivindicación 20 o la reivindicación 21, en el que determinar si el feto bajo detección tiene quimera fetal con respecto al cromosoma predeterminado en base a la primera fracción de ADN fetal libre de células y dicha tercera fracción de ADN fetal libre de células comprende adicionalmente:

50 (a) determinar una relación de dicha tercera fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células; y

(b) determinar si el feto bajo detección tiene quimera con respecto al cromosoma predeterminado al comparar la relación determinada en (a) con una pluralidad de umbrales predeterminados.

23. Un método de acuerdo con de la reivindicación 22, en el que la pluralidad de umbrales predeterminados comprende al menos uno seleccionado de:

un séptimo umbral determinado en base a una pluralidad de muestras de control con el cromosoma predeterminado que se sabe es un monosoma completo,

5 un octavo umbral determinado en base a una pluralidad de muestras de control con el cromosoma predeterminado que se sabe es una quimera monosoma,

un noveno umbral determinado en base a una pluralidad de muestras de control con el cromosoma predeterminado que se sabe que es normal,

10 un décimo umbral determinado en base a una pluralidad de muestras de control con el cromosoma predeterminado que se sabe es un trisoma completo,

y en el que

se determina que el cromosoma predeterminado del feto bajo detección es un monosoma completo si la relación de dicha tercera fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células es menor que el séptimo umbral;

15 se determina que el cromosoma predeterminado del feto bajo detección es una quimera monosoma si la relación de dicha tercera fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células no es menor que el séptimo umbral ni mayor que el octavo umbral;

20 se determina que el cromosoma predeterminado del feto bajo detección es normal si la relación de dicha tercera fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células es mayor que el octavo umbral y menor que el noveno umbral;

se determina que el cromosoma predeterminado del feto bajo detección es una quimera trisoma, si la relación de dicha tercera fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células no es menor que el noveno umbral ni mayor que el décimo umbral; y

25 se determina que el cromosoma predeterminado del feto bajo detección es un trisoma completo, si la relación de dicha tercera fracción de ADN fetal libre de células con la primera fracción de ADN fetal libre de células es mayor que el décimo umbral.

24. Un método de acuerdo con la reivindicación 23, en el que

el séptimo umbral es mayor que -1 y menor que 0, por ejemplo -0.85;

el octavo umbral es mayor que el séptimo umbral y menor que 0, por ejemplo, -0.3;

30 el noveno umbral es mayor que 0 y menor que 1, por ejemplo 0.3; y

el décimo umbral es mayor que el noveno umbral y menor que 1, por ejemplo 0.85.

25. Un método de acuerdo con de la reivindicación 1, en el que dicha fracción de ácidos nucleicos fetales libres de células se determina en una muestra de sangre periférica obtenida de una mujer embarazada con un feto como una primera fracción de ADN fetal libre de células y que adicionalmente comprende:

35 (a) determinar una fracción x_i del número de datos de secuenciación derivados del cromosoma i en el resultado de la secuenciación para los datos de secuenciación totales, donde i representa un número de serie del cromosoma, e i es cualquier entero en el rango de 1 a 22;

(b) determinar una puntuación T del cromosoma i de acuerdo con $T_i = (X_i - \mu_i) / \sigma_i$, donde i representa el número de serie del cromosoma e i es cualquier entero en el rango de 1 a 22, μ_i representa un valor promedio de porcentajes de datos de secuenciación del cromosoma i seleccionado como un sistema de referencia en una base de datos de referencia para los datos de secuenciación totales del mismo, σ_i representa una desviación estándar de los porcentajes de los datos de secuenciación del cromosoma i seleccionado como el sistema de referencia en la base de datos de referencia para los datos de secuenciación totales del mismo,

(c) determinar una puntuación L del cromosoma i de acuerdo con

45 $L_i = \log(d(T_i, a)) / \log(d(T2_i, a))$, donde i representa el número de serie del cromosoma e i es cualquier entero en el rango de 1 a 22,

$T2_i = (X_i - \mu_i * (1 + fra/2)) / \sigma_i$; $d(T_i, a)$ y $d(T2_i, a)$ representa la función de densidad de probabilidad de distribución t , donde a representa el grado de libertad, fra representa dicha primera fracción de ADN fetal libre de células;

(d) graficar un diagrama de cuatro cuadrantes con T como coordenada vertical y L como coordenada horizontal mediante zonificación con una tercera línea recta donde T = undécimo umbral predeterminado y una cuarta línea recta donde L = duodécimo umbral predeterminado, cuando la puntuación T no es mayor que 0,

en el que

- 5 se determina que el feto tiene un monosoma completo o quimera monosoma con respecto al cromosoma predeterminado si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que cae en un primer cuadrante;
- se determina que el feto tiene una quimera monosoma con respecto al cromosoma predeterminado si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que cae en un segundo cuadrante;
- 10 se determina que el feto es normal con respecto al cromosoma predeterminado, si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que cae en un tercer cuadrante; y
- se determina que el feto ha dado lugar a una fracción fetal baja si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que cae en un cuarto cuadrante de tal manera que no se adopta el resultado,
- 15 (e) graficar un diagrama de cuatro cuadrantes con T como coordenada vertical y L como coordenada horizontal mediante zonificación con una primera línea recta donde T = decimotercero umbral predeterminado y una segunda línea recta donde L = decimocuarto umbral predeterminado, cuando la puntuación T es mayor que 0, en el que
- se determina que el feto tiene un trisoma o quimera trisoma completo con respecto al cromosoma predeterminado si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que cae en un primer cuadrante;
- 20 se determina que el feto tiene una quimera trisoma con respecto al cromosoma predeterminado si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que cae en un segundo cuadrante;
- se determina que el feto es normal con respecto al cromosoma predeterminado si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que cae en un tercer cuadrante; y
- 25 se determina que el feto ha dado lugar a una fracción fetal baja si se determina que una muestra bajo detección tiene una puntuación T y una puntuación L que cae en un cuarto cuadrante de tal manera que no se adopta el resultado.
26. Un método de acuerdo con la reivindicación 25, en el que el undécimo umbral y el decimotercero umbral son cada uno independientemente 3 y el duodécimo umbral y el decimocuarto umbral son cada uno independientemente 1.

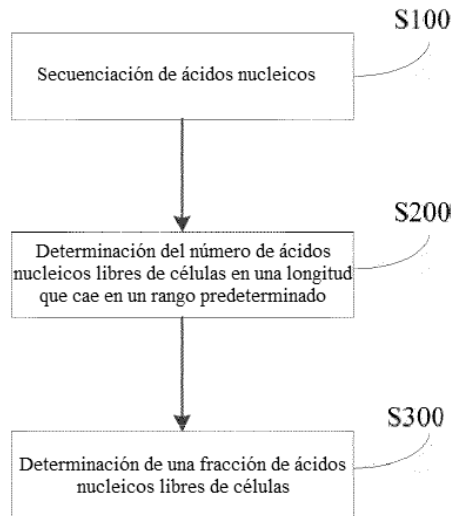


Fig. 1

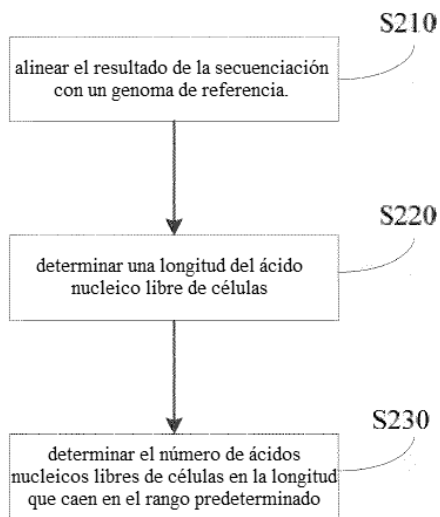


Fig. 2

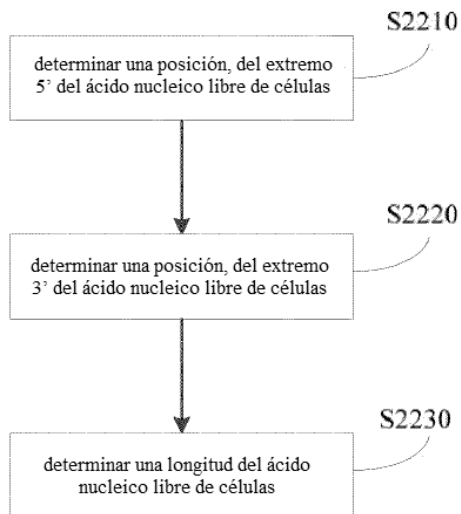


Fig. 3

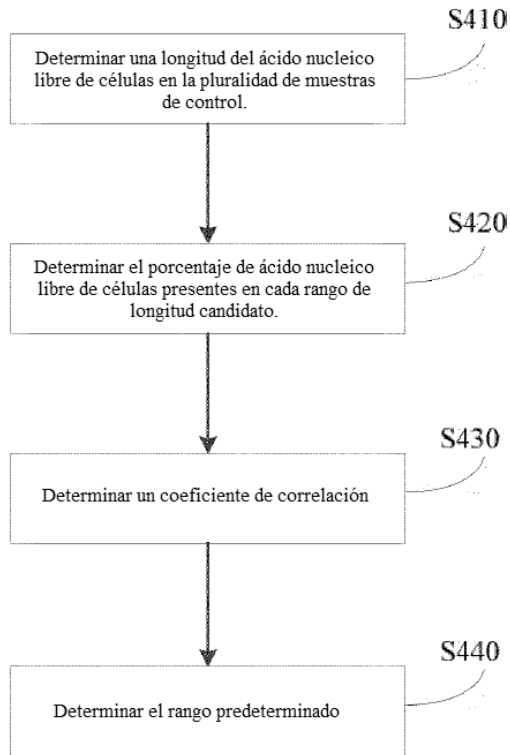


Fig. 4

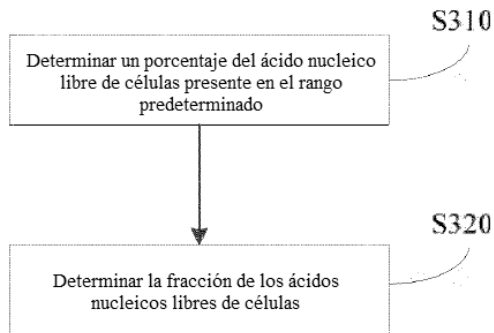


Fig. 5

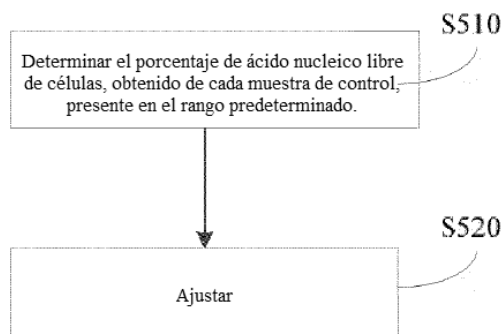


Fig. 6

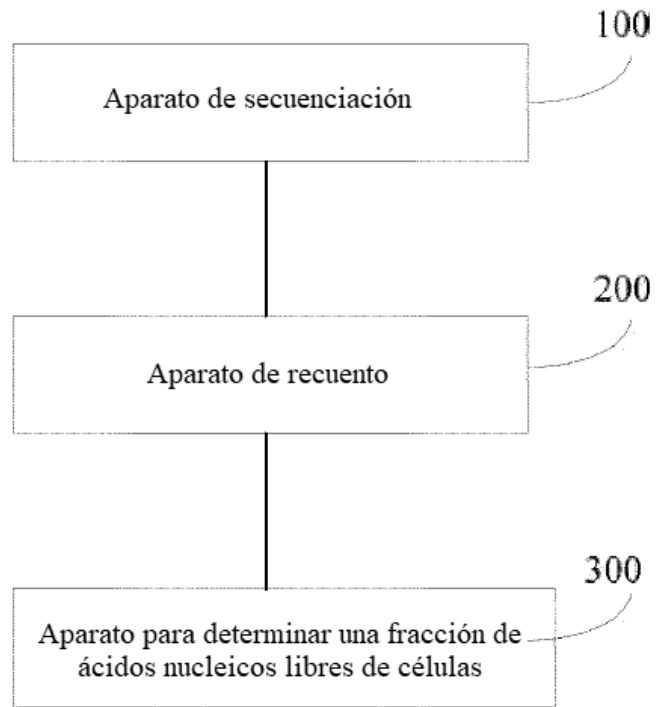


Fig. 7

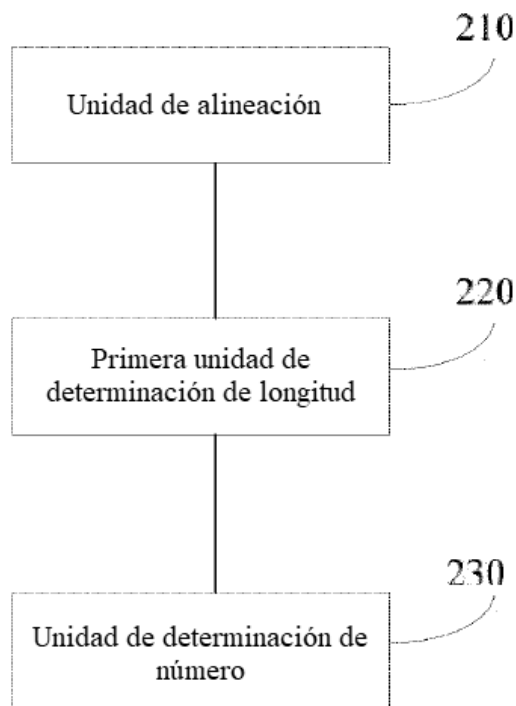


Fig. 8

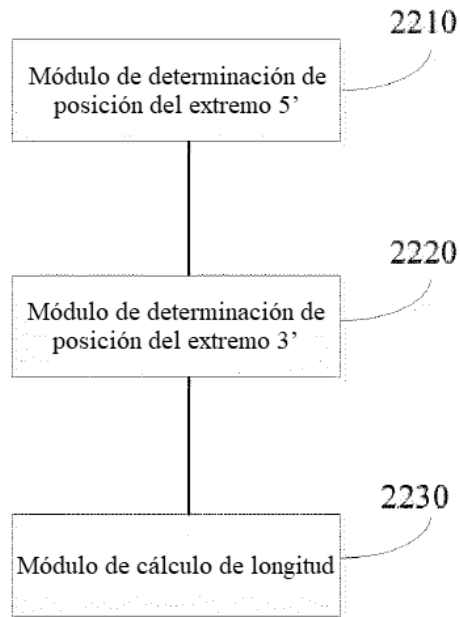


Fig. 9

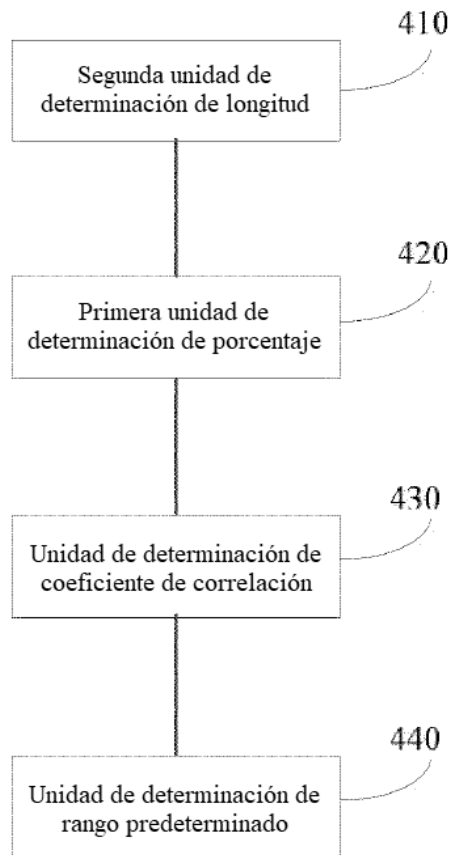


Fig. 10

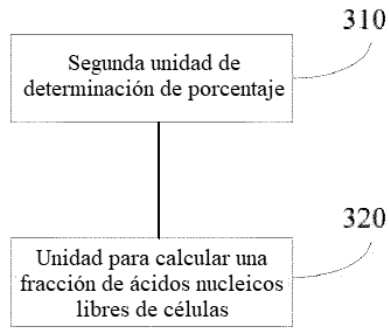


Fig. 11

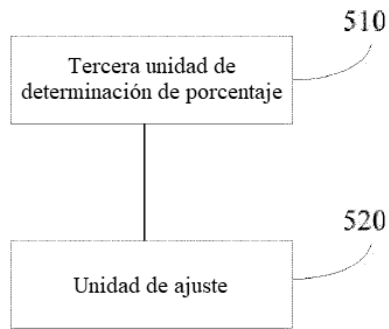


Fig. 12

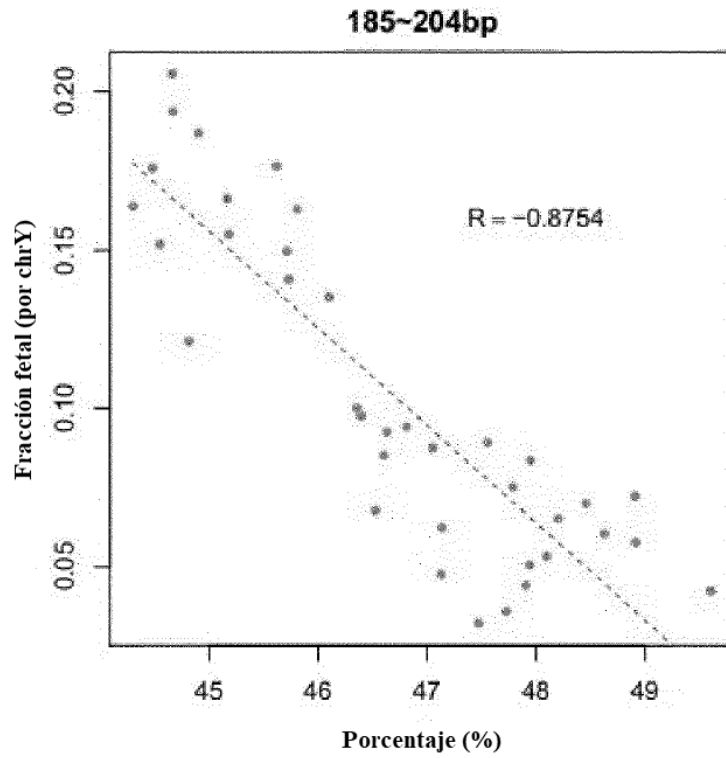
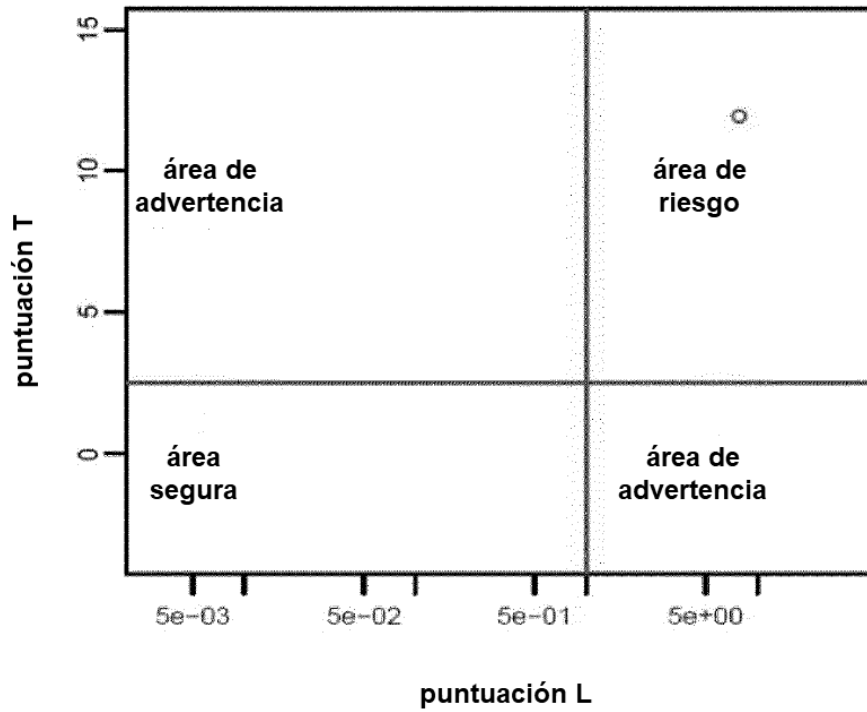


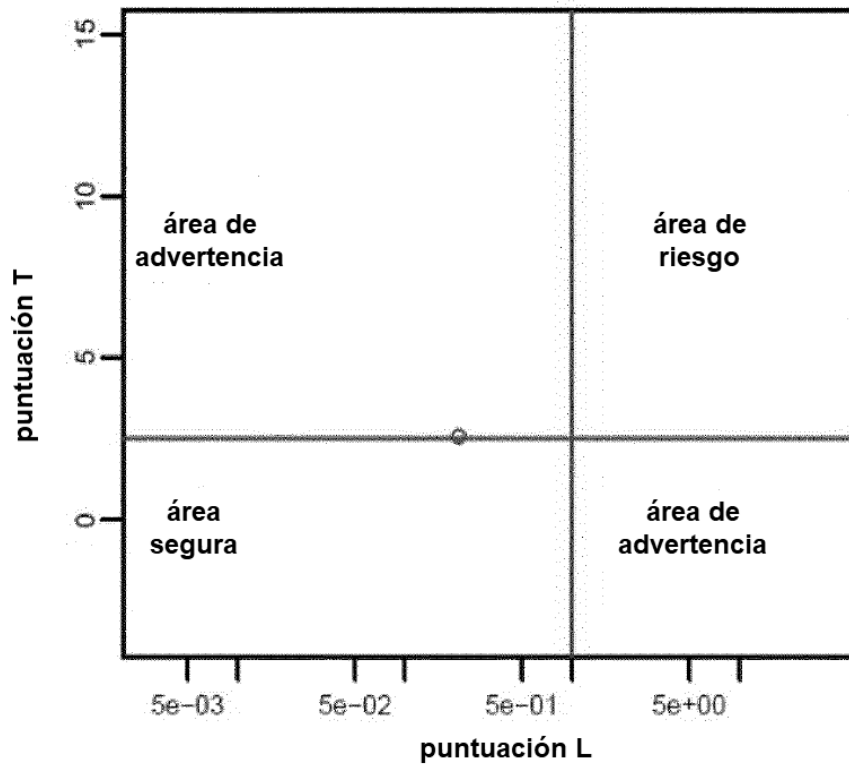
Fig. 13

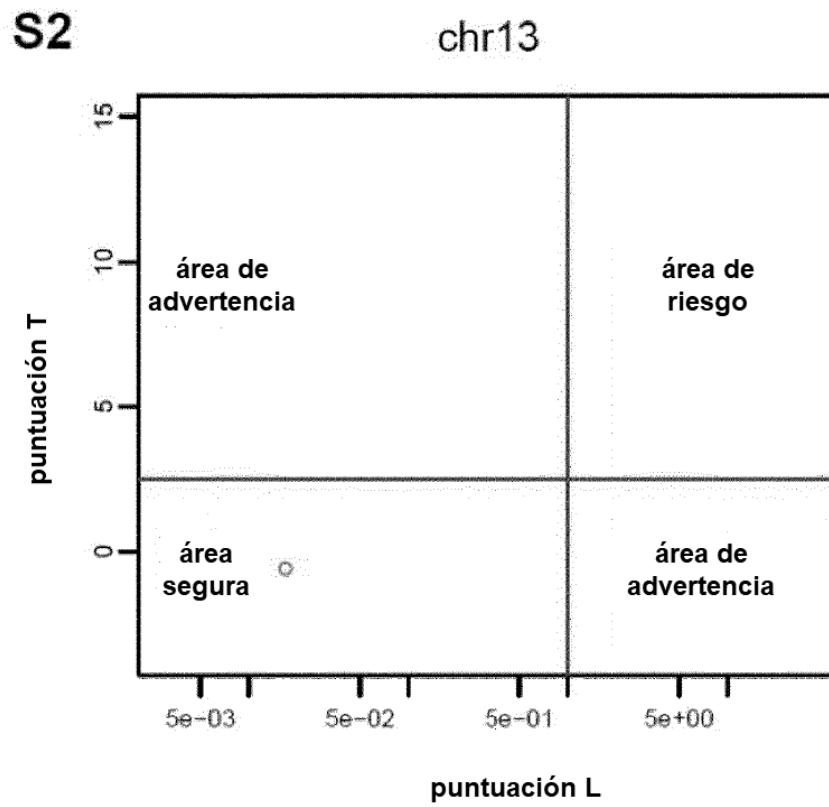
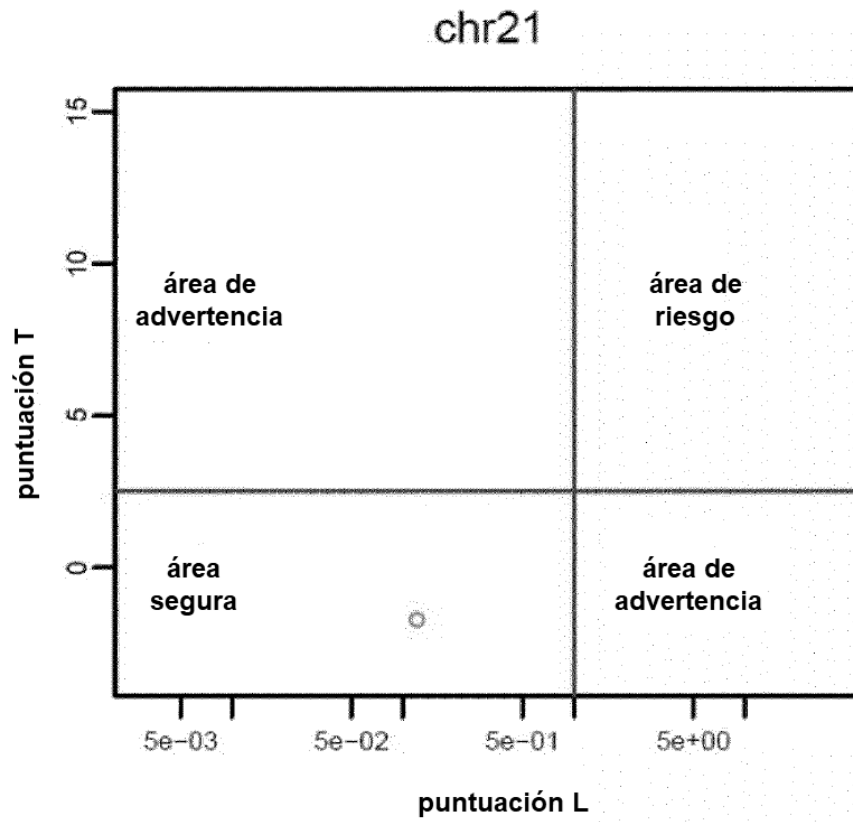
S1

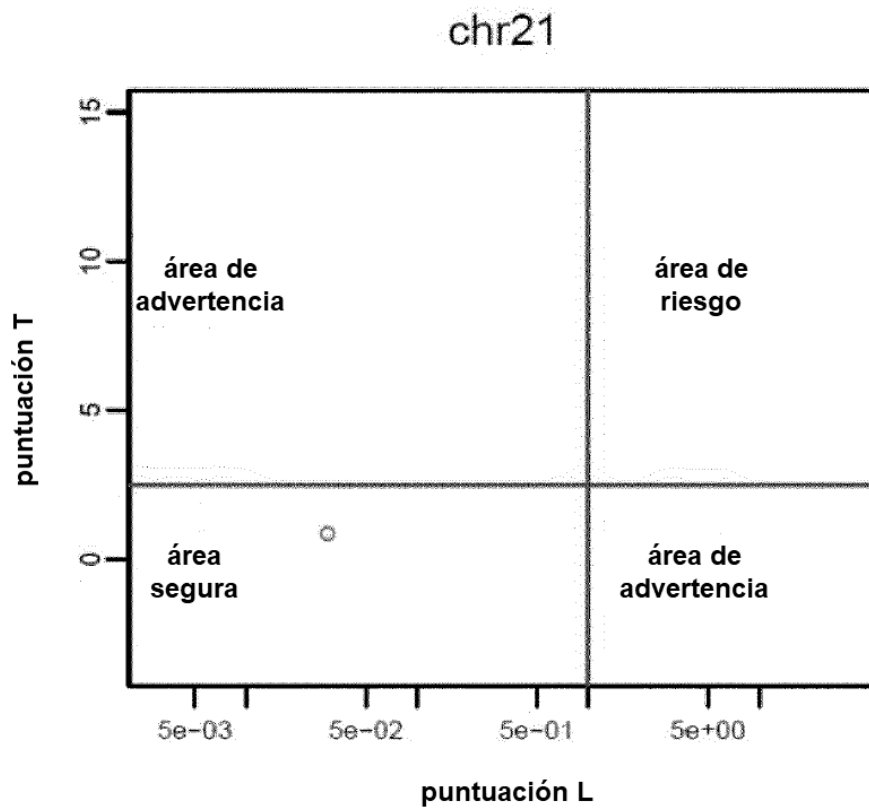
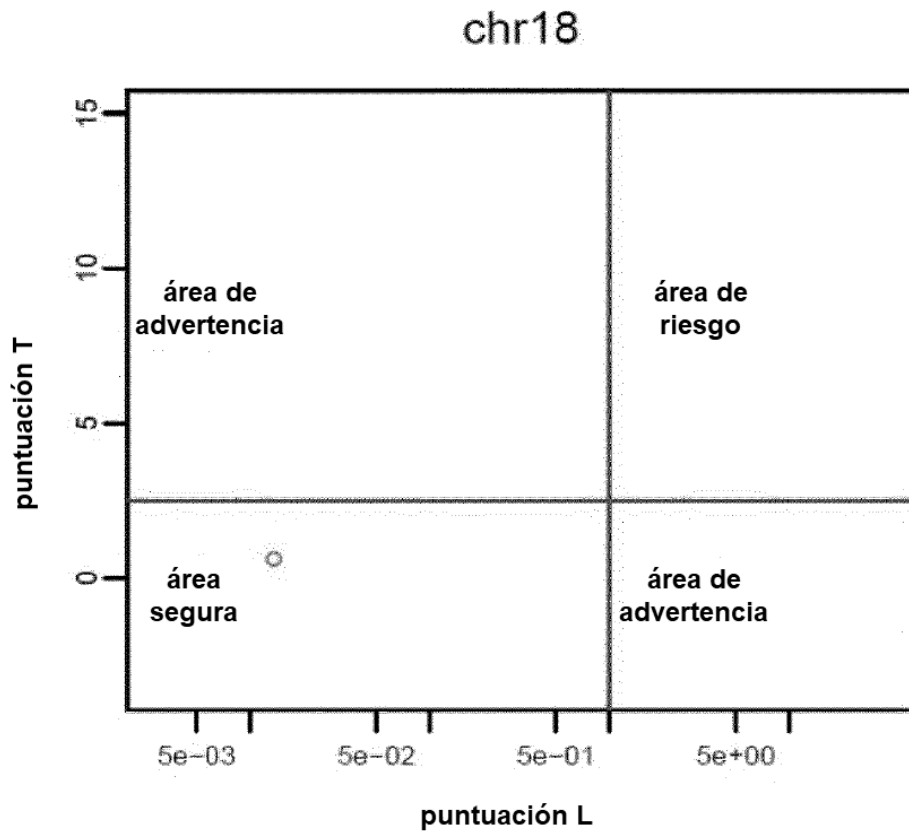
chr13



chr18

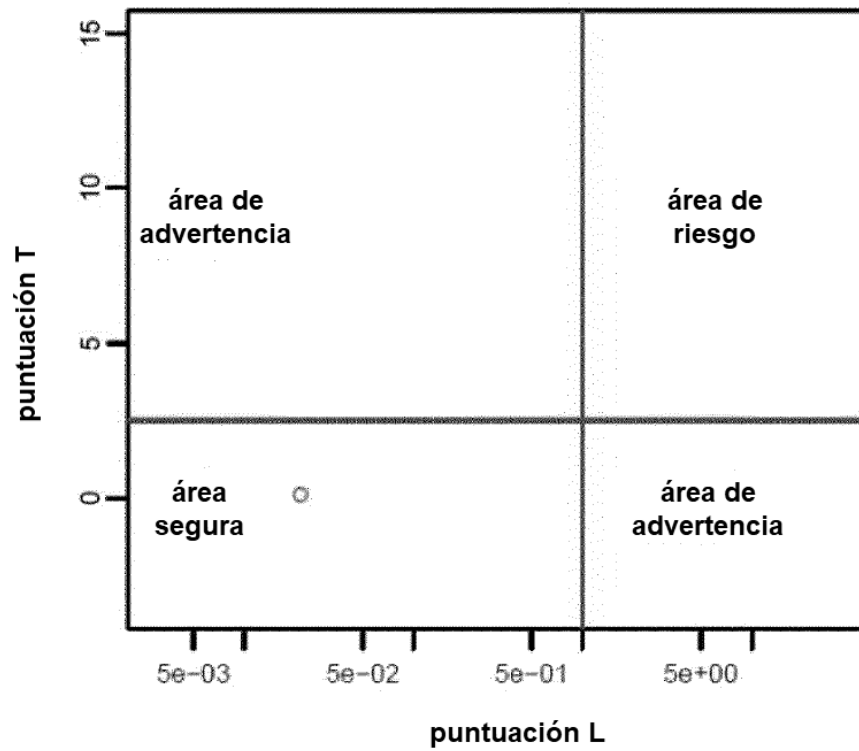




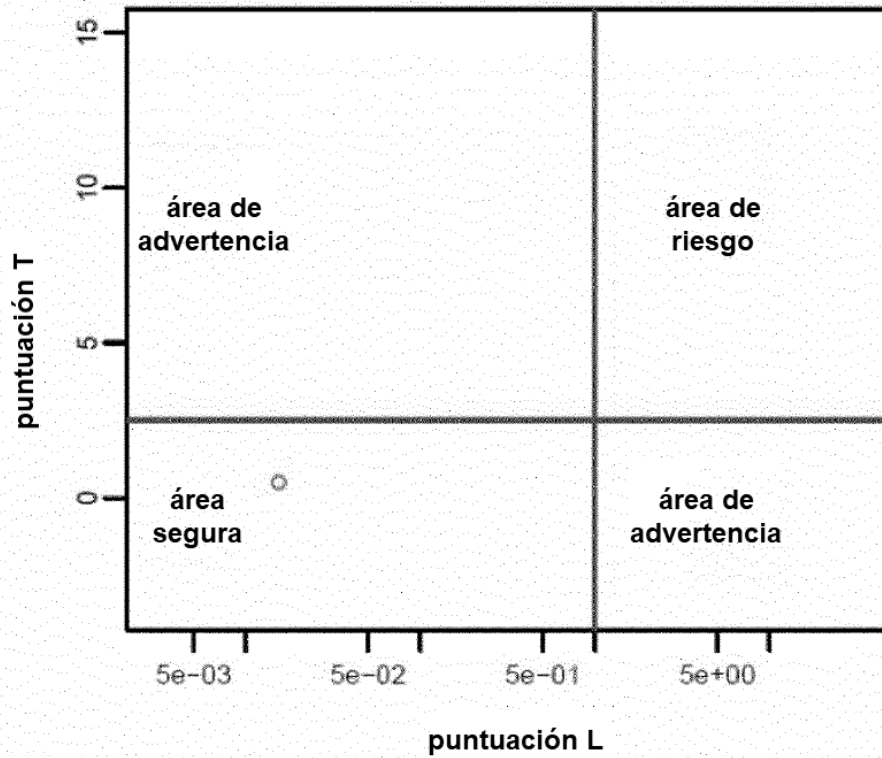


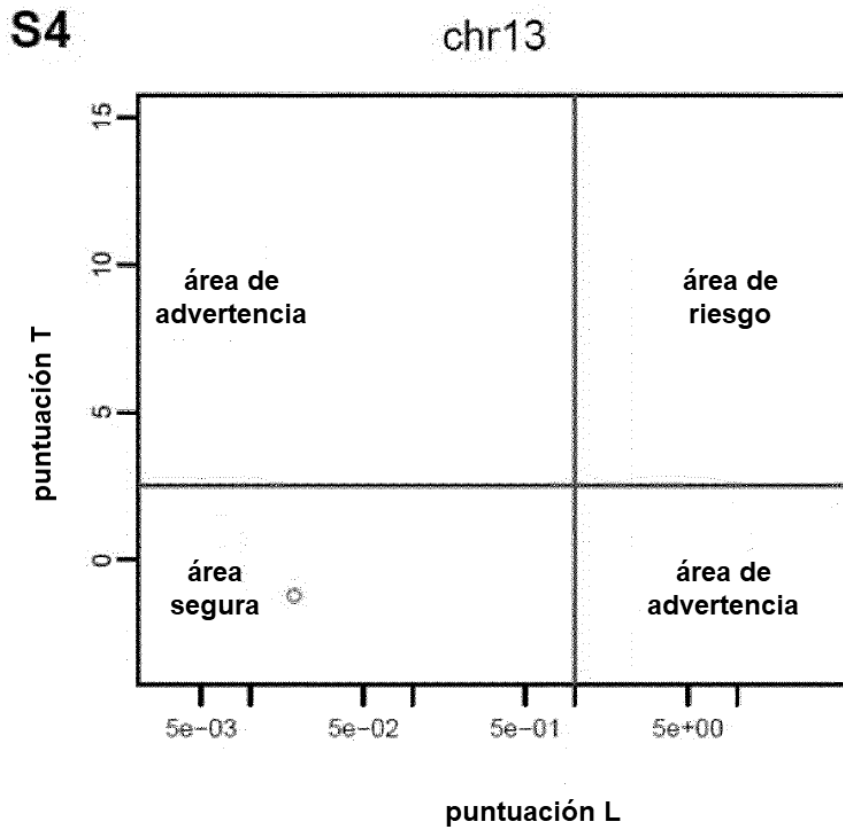
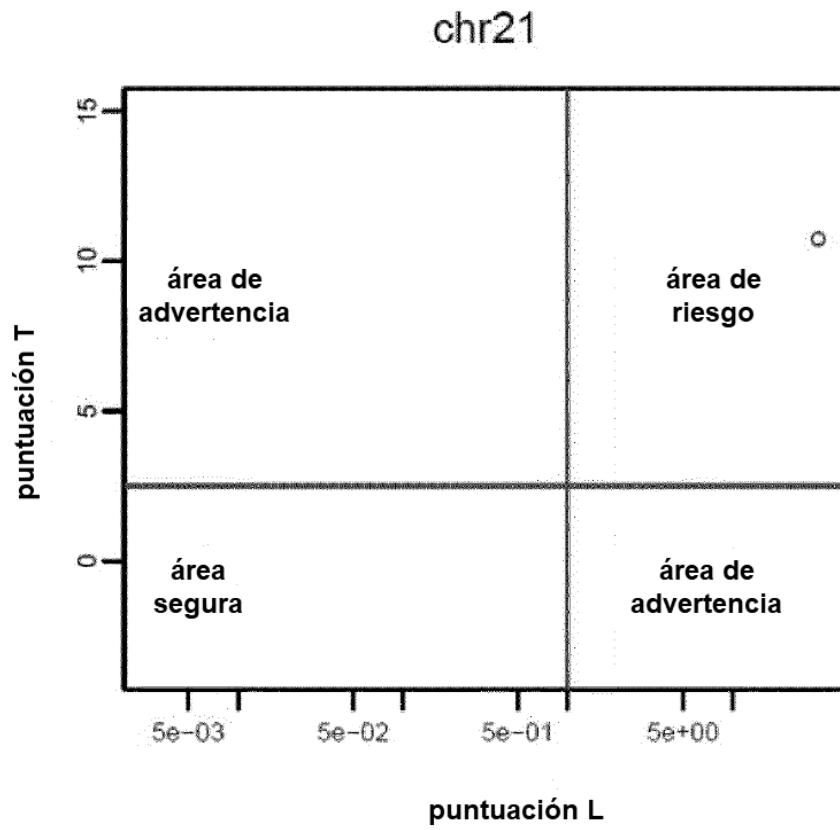
S3

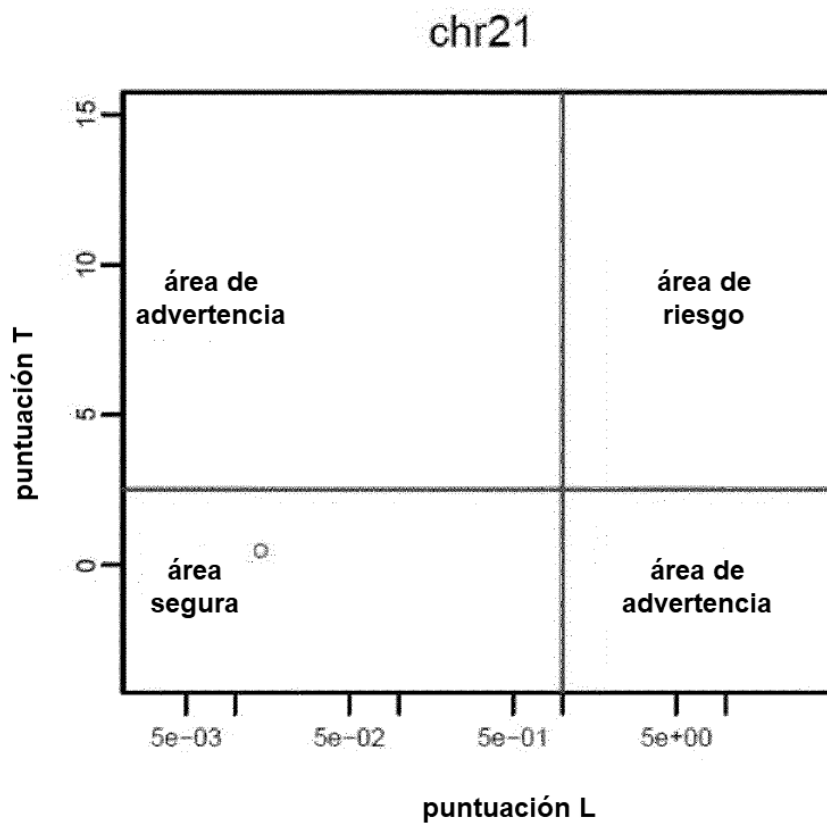
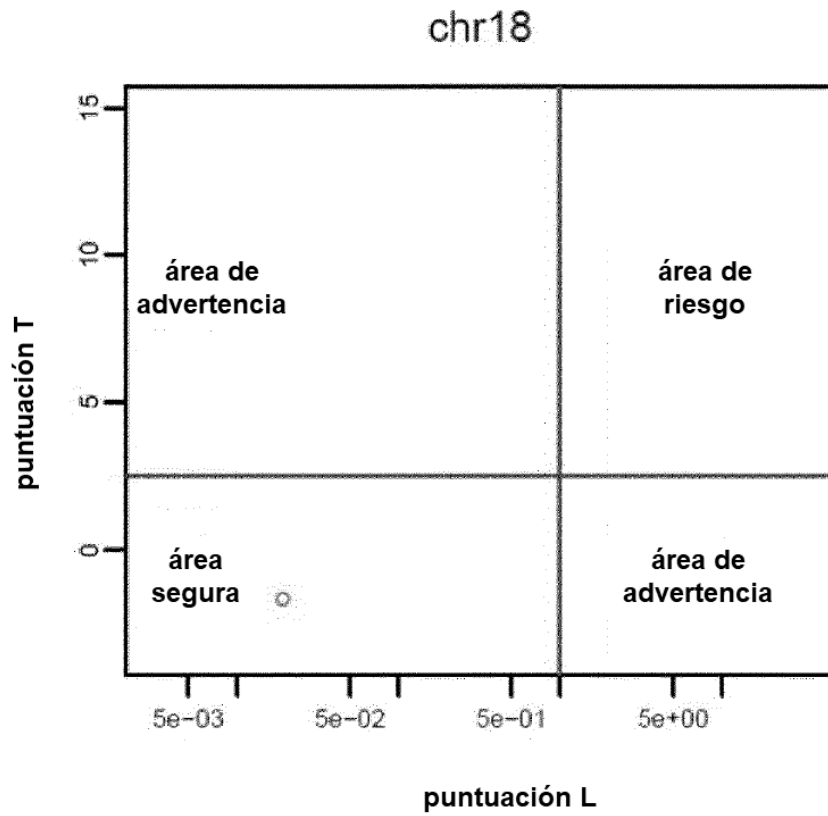
chr13



chr18

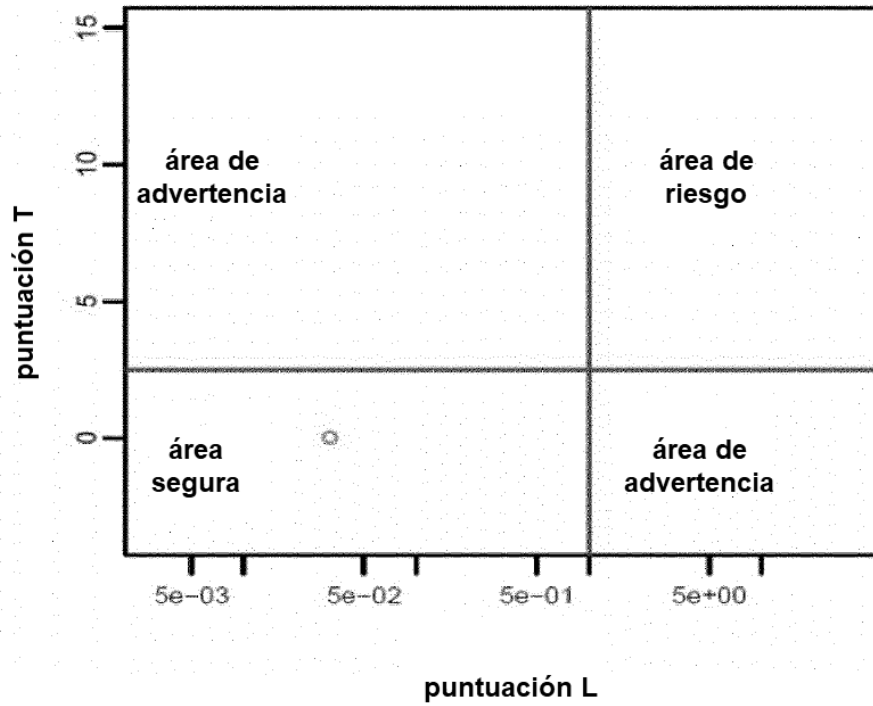




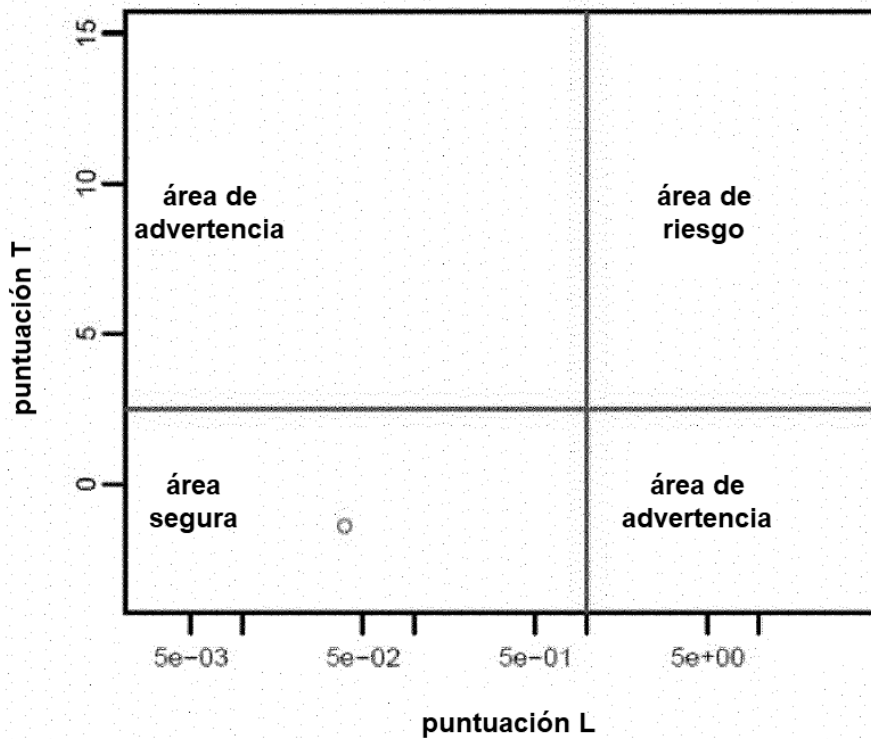


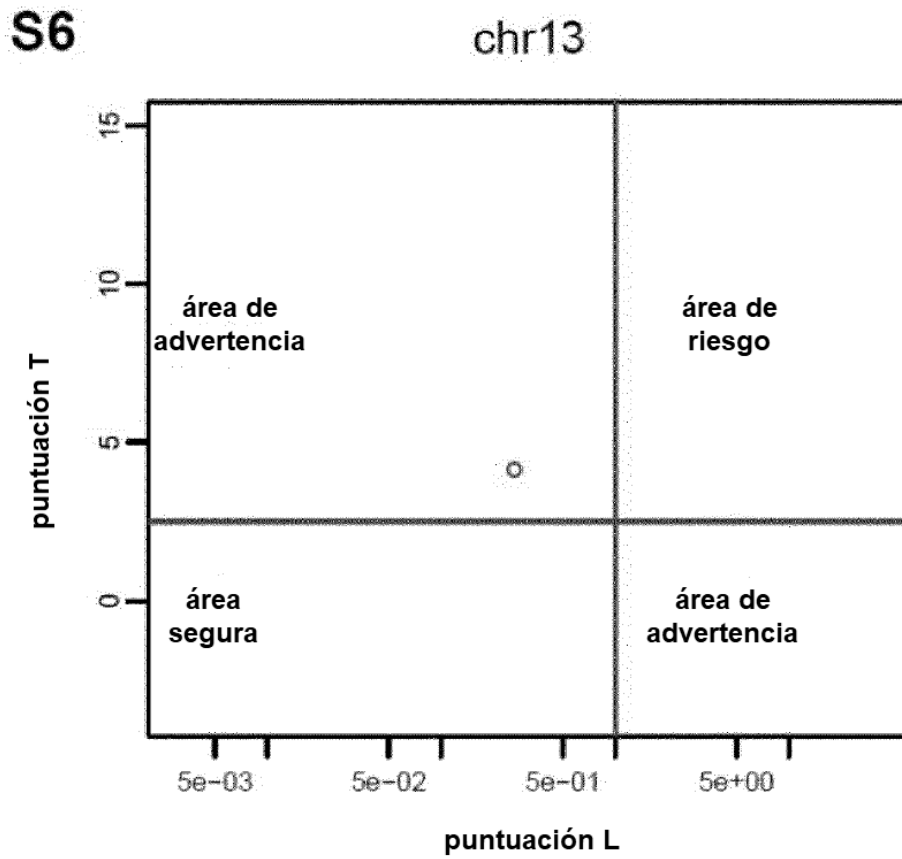
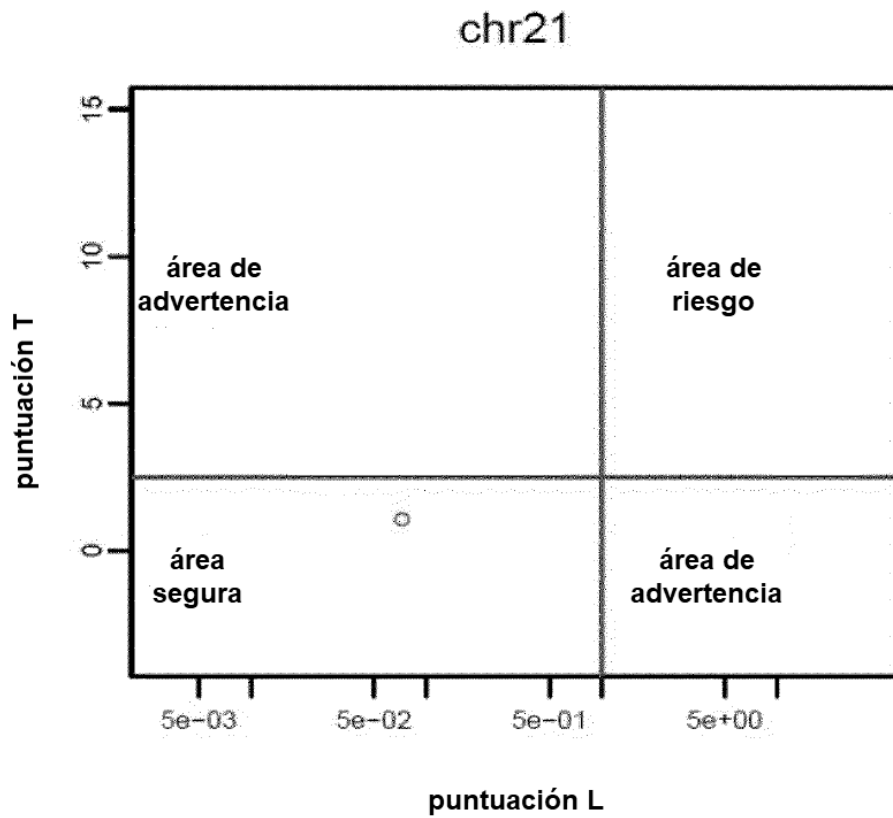
S5

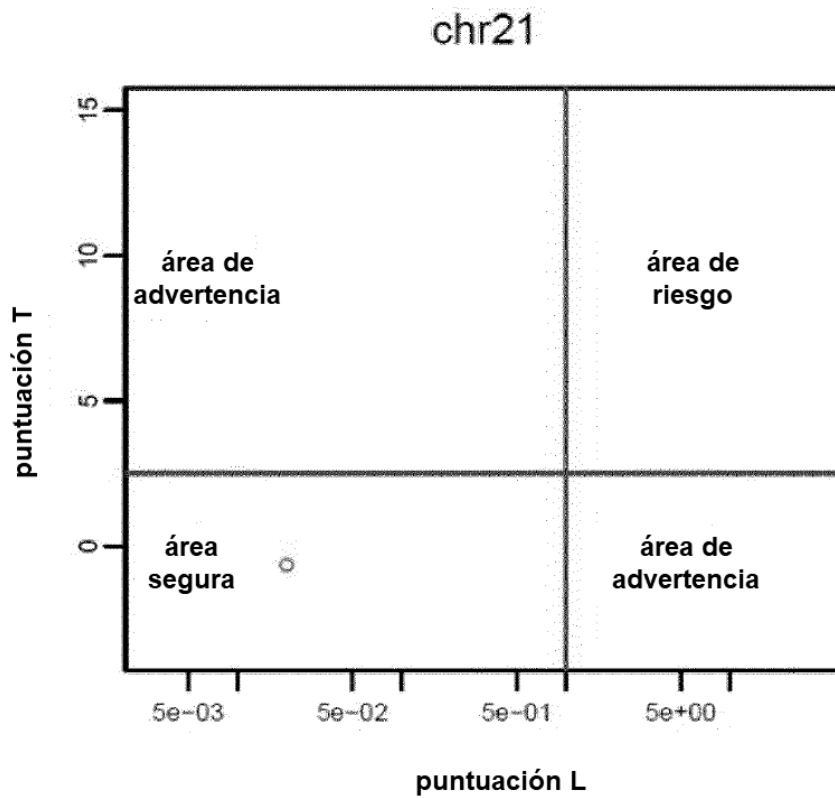
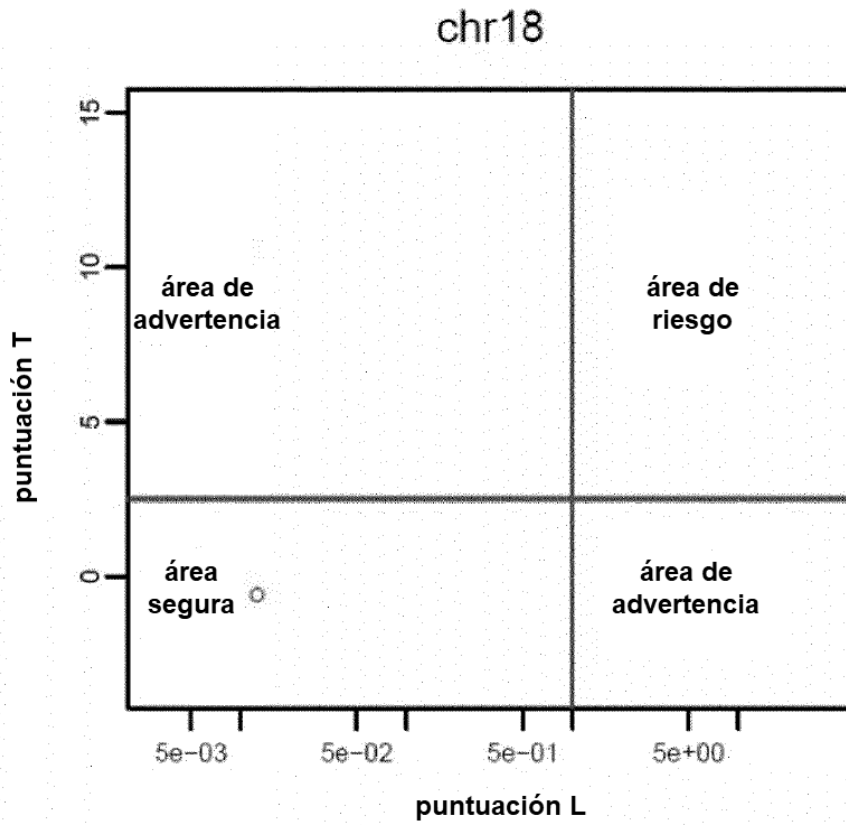
chr13



chr18

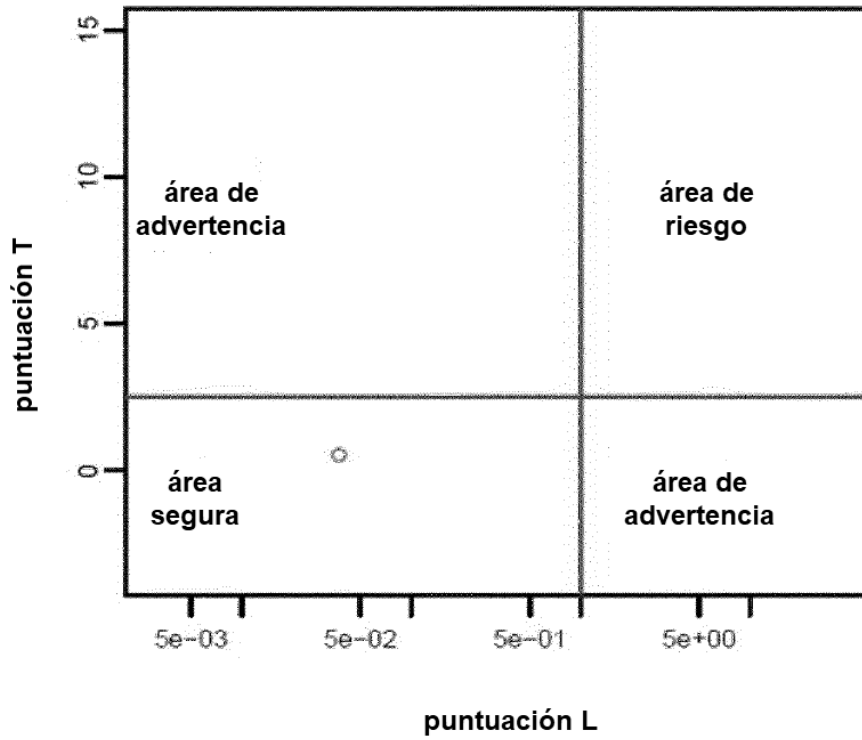




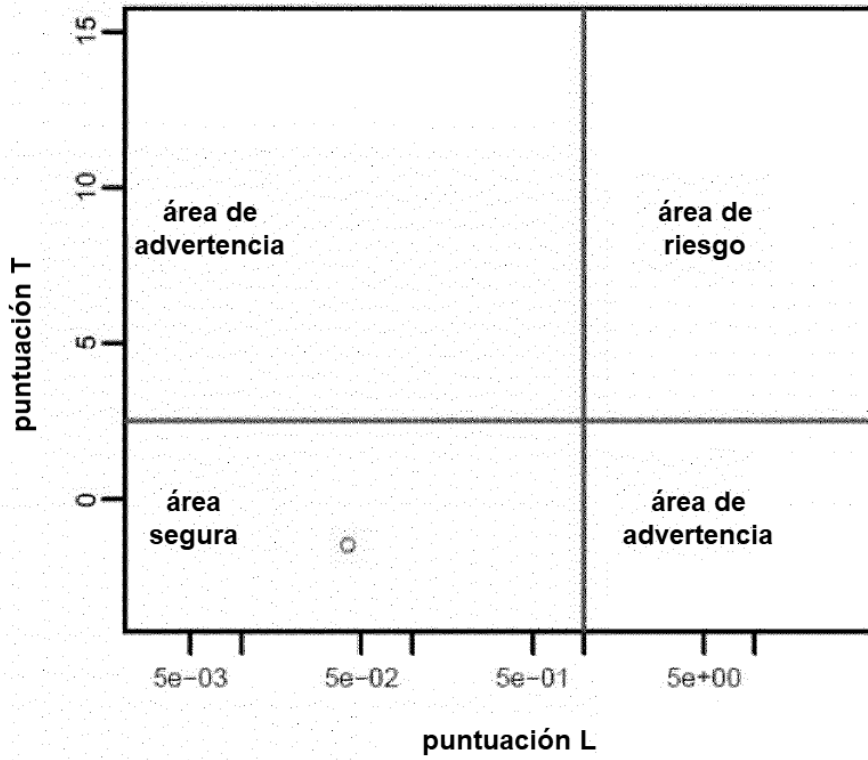


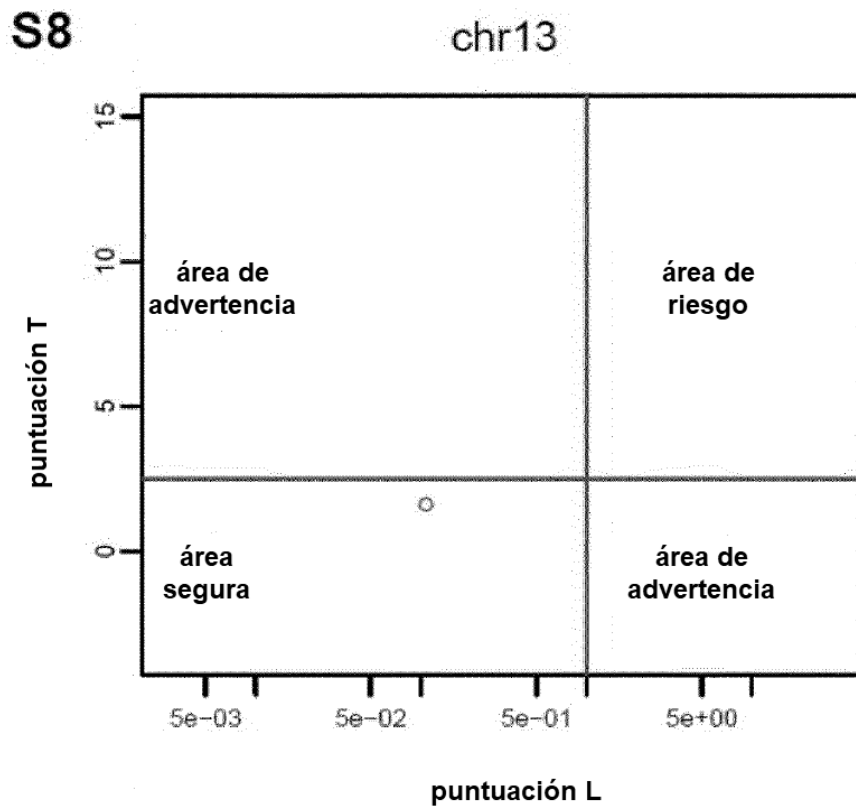
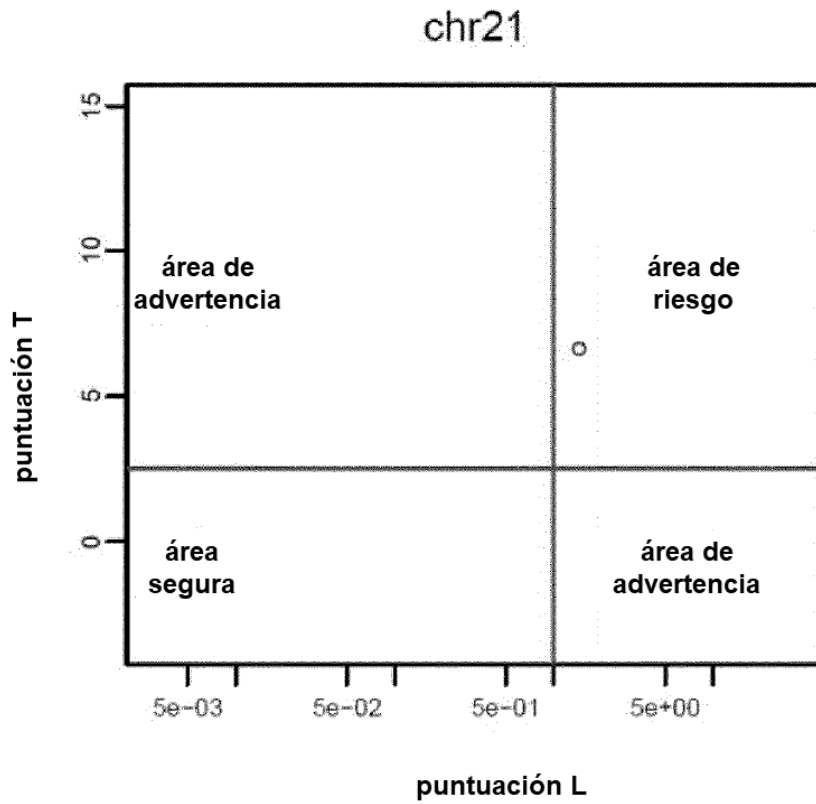
S7

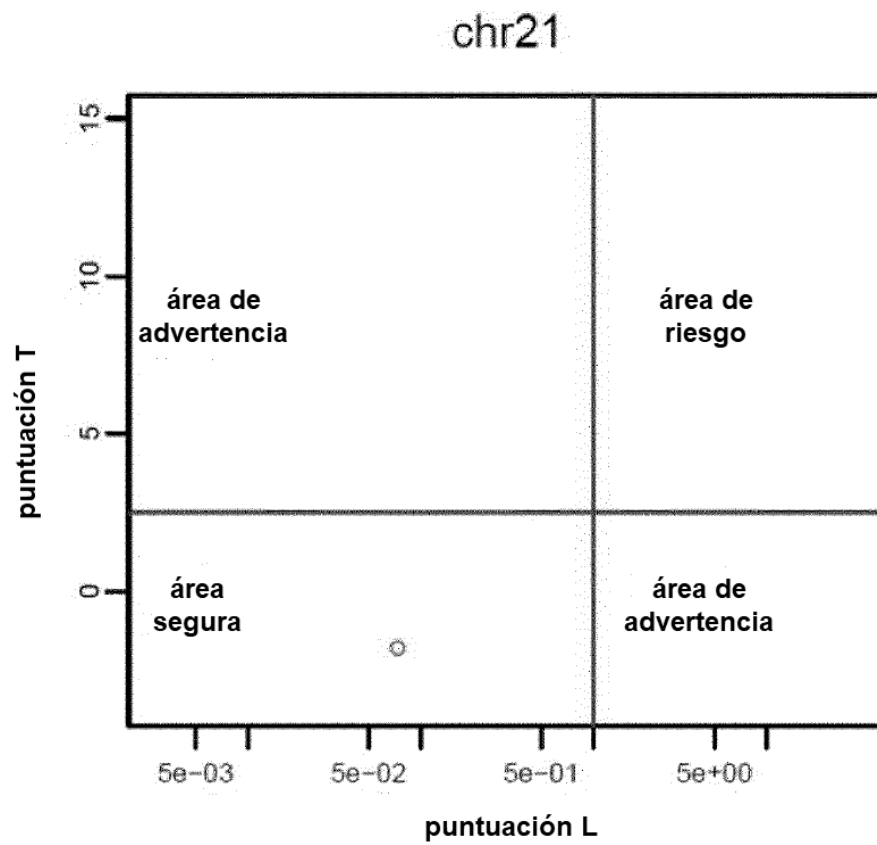
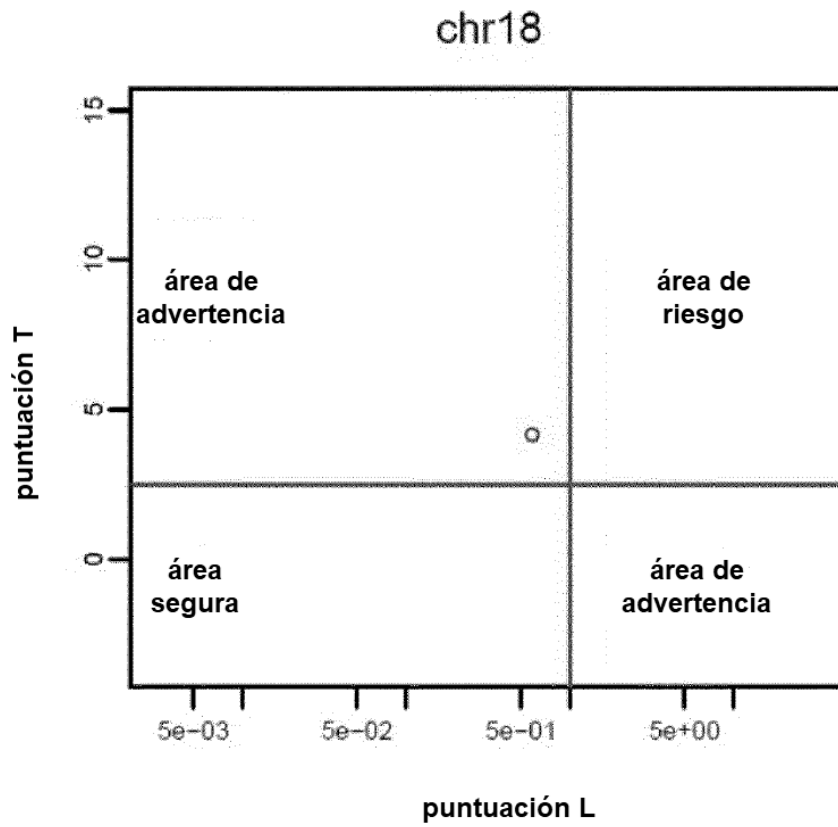
chr13



chr18

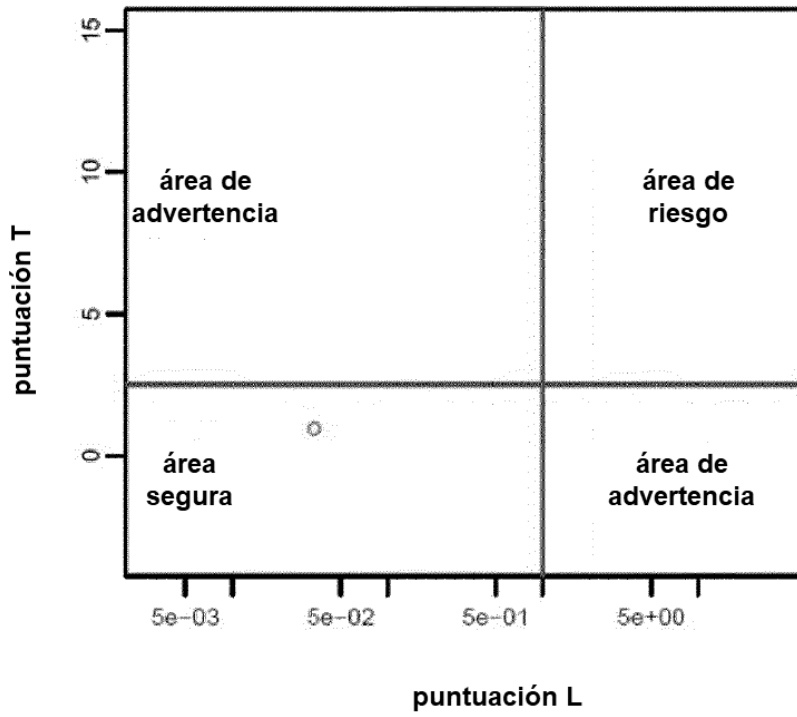




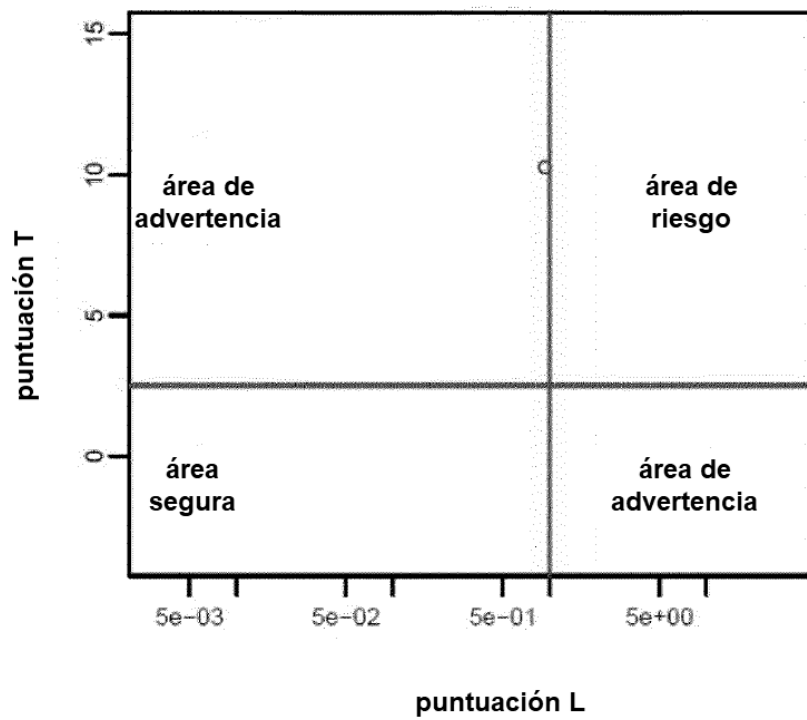


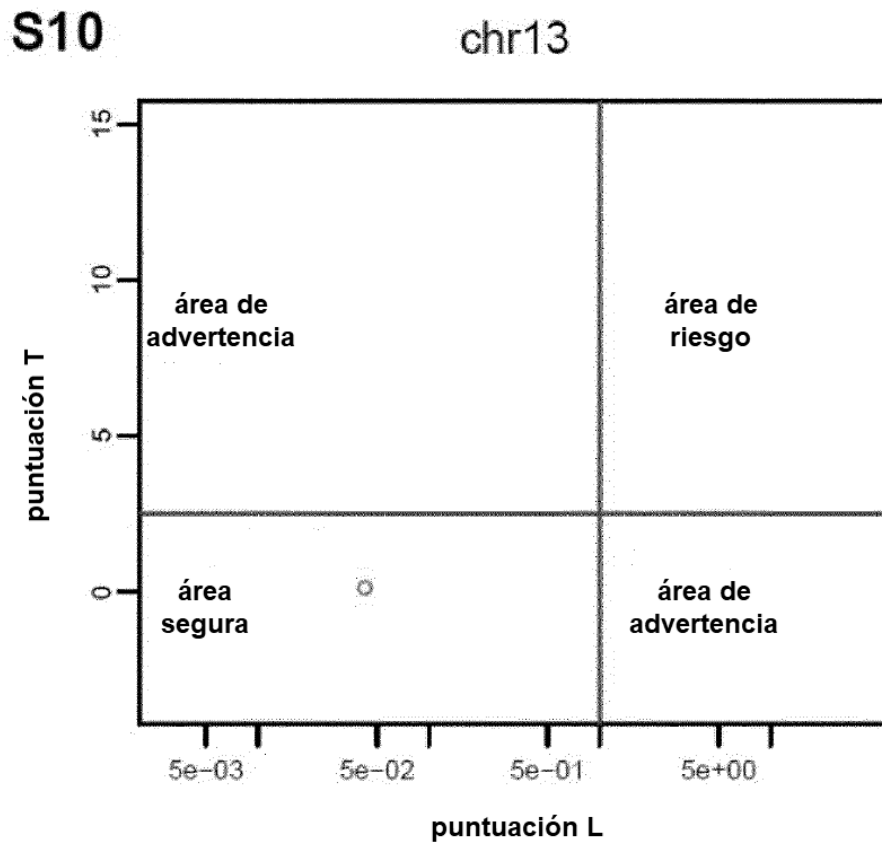
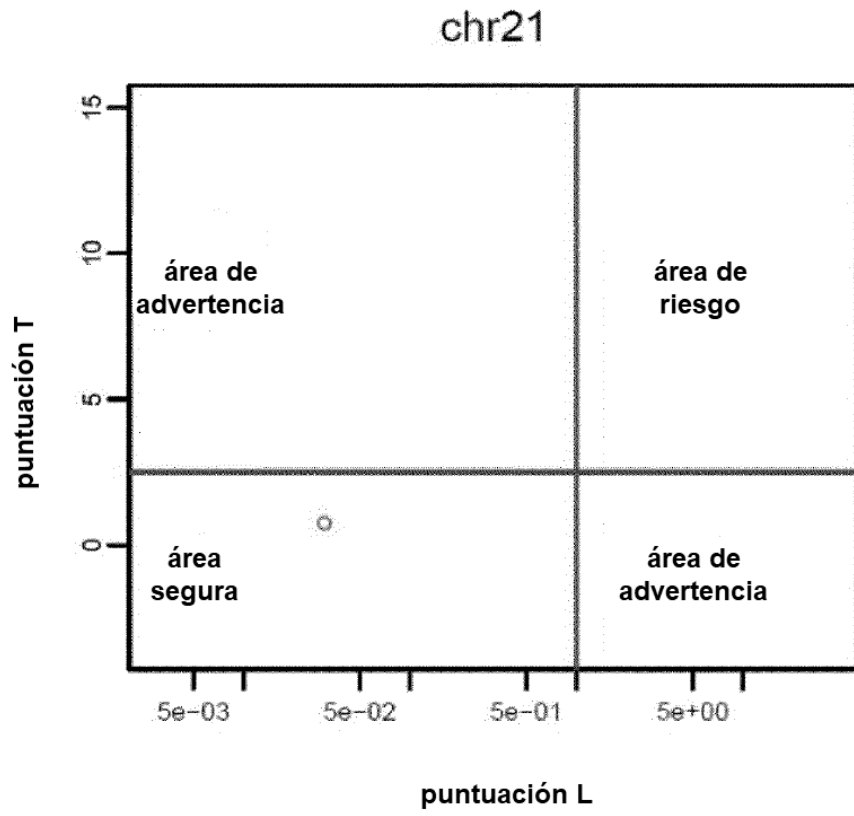
S9

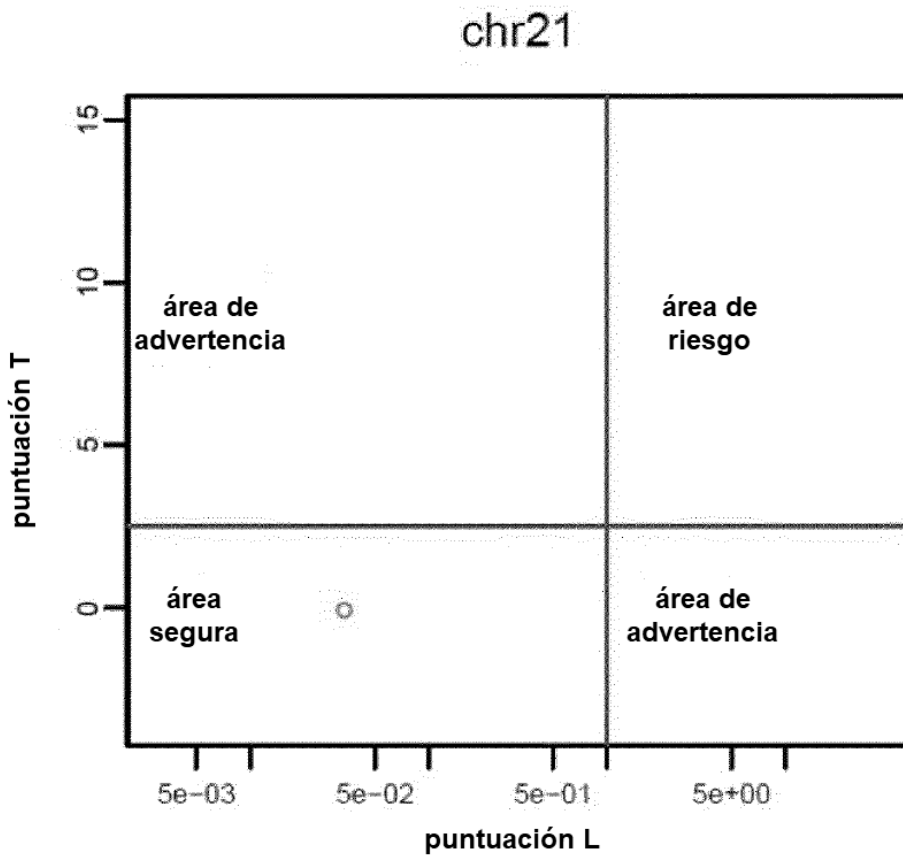
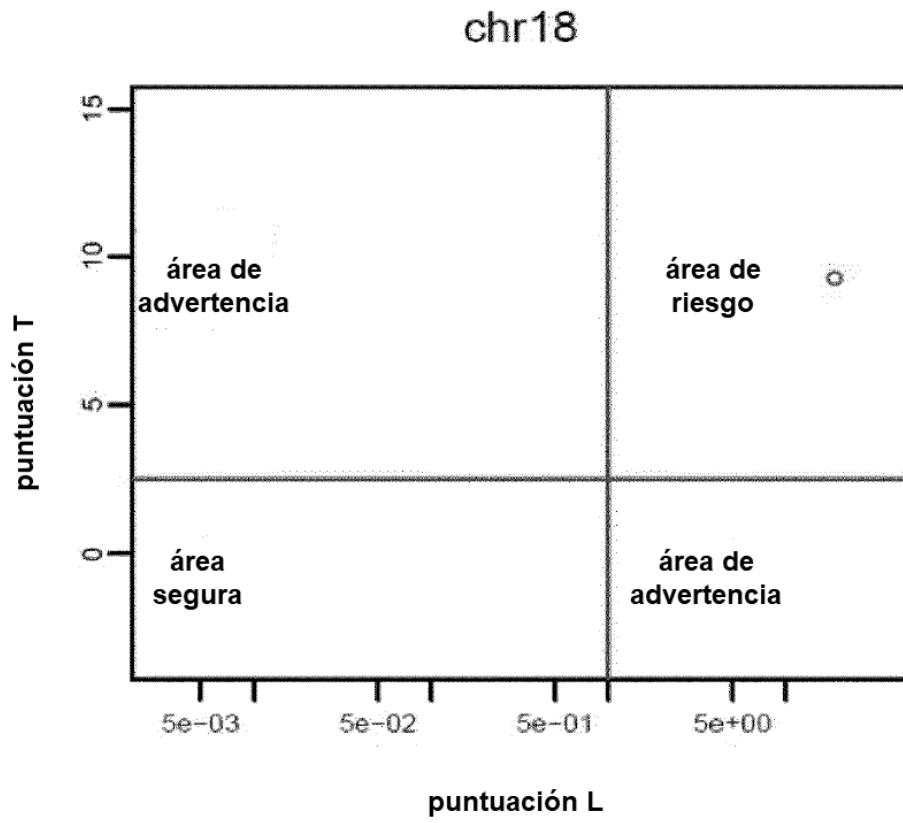
chr13



chr18

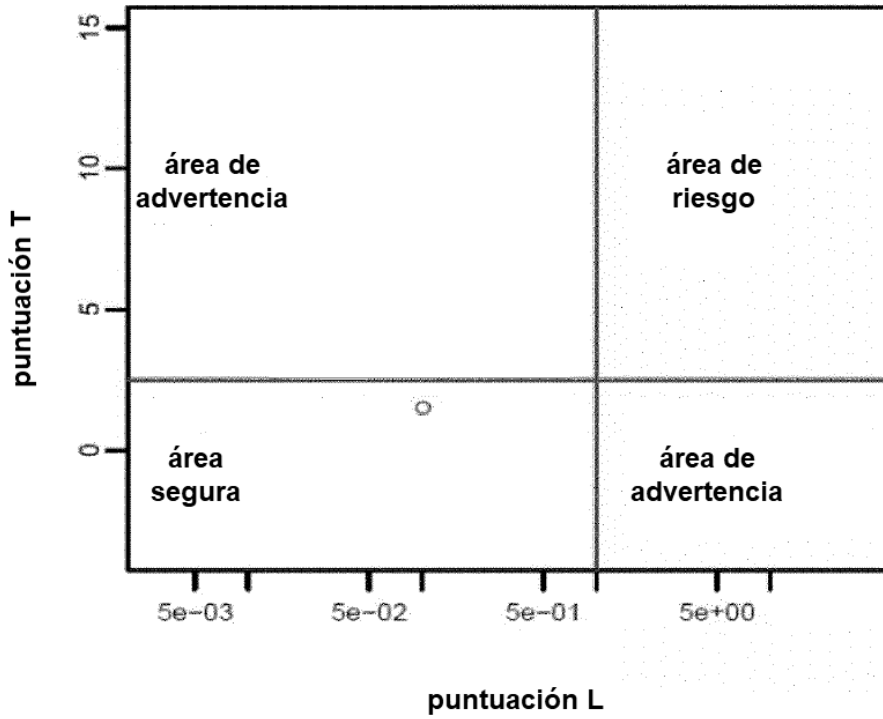






S11

chr13



chr18

