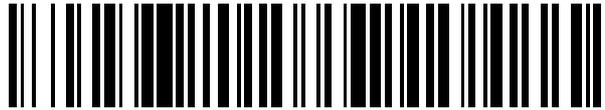


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 552 343**

51 Int. Cl.:

C12Q 1/68 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **13.08.2010 E 10855744 (8)**

97 Fecha y número de publicación de la concesión europea: **22.07.2015 EP 2604700**

54 Título: **Un procedimiento para el análisis de cromosomas de células**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
27.11.2015

73 Titular/es:

**BGI GENOMICS CO., LTD. (100.0%)
21 Hongan 3rd Street, Floor 7-14, Building 7, BGI
Park, Yantian District
Shenzhen, CN**

72 Inventor/es:

**ZHANG, XIUQING;
XUAN, ZHAOLING;
CHEN, FANG;
JIANG, FUMAN;
LIN, JINGRONG y
LI, PEIPEI**

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 552 343 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

Un procedimiento para el análisis de cromosomas de células

Campo de la invención

5 La presente invención se refiere a un procedimiento para el análisis de cromosomas celulares y, en particular, se refiere a un procedimiento para el análisis de cromosomas de células amnióticas mediante secuenciación.

Antecedentes de la invención

Aneuploidía cromosómica fetal

10 La aneuploidía cromosómica fetal es una afección en la que el número de cromosomas no es diplóntico en un genoma fetal. Normalmente, en un genoma humano en el que el cariotipo masculino es (46, XY) y el cariotipo femenino es (46, XX) hay 44 autosomas y 2 cromosomas sexuales. La aneuploidía cromosómica fetal puede referirse al estado de tener un cromosoma más que un feto diploide normal, es decir 47 cromosomas en el genoma fetal. Tómese la trisomía fetal 21 por ejemplo, en comparación con un feto diploide normal, el feto con trisomía tiene un cromosoma 21 extra con el cariotipo de 47, XX (o XY), + 21. Además, la aneuploidía cromosómica fetal se refiere al estado de falta de un cromosoma en comparación con el feto diploide normal, es decir 45 cromosomas en el
15 genoma fetal. Por ejemplo, el feto con síndrome de Turner con el cariotipo de 45, XO carece de un cromosoma, en relación con el feto diploide normal. La aneuploidía cromosómica fetal también se refiere al estado en el que una parte del cromosoma se ha perdido, por ejemplo, la trisomía translocada del 21 con el cariotipo de 45, XX, der (14; 21) (q10; q10), y el síndrome de Cri du chat con el cariotipo de 46, XX (XY), del (5) (p13).

20 De acuerdo con las estadísticas incompletas, la tasa de natalidad de los fetos con aneuploidía cromosómica es 1/160 en el mundo, en el que la tasa de nacimiento de fetos con trisomía del 21 (T21, síndrome de Down) es 1/800, la tasa de natalidad de la trisomía fetal del 18 (T18, síndrome de Edwards) es 1/6000, y la tasa de natalidad de la trisomía del 13 (T13, el síndrome de Patau) es 1/10.000. El desarrollo de fetos con otros tipos de aneuploidía cromosómica se estanca debido a que algunas etapas del desarrollo no pudieron llevarse a cabo, lo que da lugar a un aborto involuntario sin razones clínicas en las primeras etapas de la gestación (Deborah A. Driscoll, M.D., and Susan Gross, M.D., Prenatal Screening for Aneuploidy[J]. N Engl J Med 2009; 360:2556 – 62).

Situación actual del cultivo de las células amnióticas

30 Las células amnióticas son células epiteliales que flotan en el líquido amniótico, que derivan de la piel, los tractos digestivos y las vías respiratorias del feto. El procedimiento de cultivar células amnióticas, enriqueciendo el número de células nucleadas fetales, preparando la muestra cromosómica, y analizando el cariotipo cromosómico fetal es una referencia para el diagnóstico clínico tradicional de la anomalía cromosómica de los fetos. Esta técnica es fiable, precisa, y permite la observación de anomalías de número y estructura de los cromosomas. Sus desventajas, sin embargo, son que requiere mucho tiempo, tarda de 10 días a 3 semanas para producir el resultado, y la tasa de fallos del cultivo es de aproximadamente 1,00 % (THEIN AT, ABDEL-FATTAH SA, KYLE PM, y col., An Assessment of the Use of Interphase FISH with Chromosome Specific Probes as an Alternative to Cytogenetics in Prenatal
35 Diagnosis [J]. PrenatDiagn, 2000, 20(4): 275 – 280).

40 La razón de la alta tasa de fracaso del cultivo de las células amnióticas es que las células amnióticas están envejeciendo y las células picnóticas, lo que da lugar a un cultivo más difícil que el de otros tejidos (Changjun Ma, Yuania Chen, Peidan Huo; The Culture of Amniotic Cells and the Method of Preparing the Chromosomal Specimen Thereof [J]. Reproduction and Contraception, 1985, 5 (1): 53 – 4). Por lo tanto, el cultivo con éxito de células amnióticas desempeña un papel crucial en el proceso de detección de aneuploidía cromosómica. Debido a la alta exigencia del cultivo de células amnióticas, las relativamente pocas células viables en el líquido amniótico de algunas de las mujeres embarazadas, las relativamente pocas células recolectadas con fases de división, y la mala forma cromosómica, es difícil contar y analizar un número suficiente de células, y detectar cromosomas. Además, la amniocentesis es un riesgo de que aproximadamente un 2-3 % de las mujeres embarazadas sufren complicaciones, tales como contracciones uterinas, hinchazón abdominal, dolor, sangrado vaginal, infección, fugas de agua, o daño fetal. Sería inaceptable que se tuviera que pedir a una mujer embarazada que se sometiera a una segunda amniocentesis si el cultivo de células amnióticas falla o las células recogidas no son suficientes para contar y analizar. Además, la amniocentesis se realiza generalmente a las 16-20 semanas de gestación. Una vez que el cultivo ha fallado, es probable que el período de gestación de la mujer embarazada haya avanzado demasiado, por lo que podría tener que someterse a cordocentesis con un riesgo incluso mayor. Después de cultivar las células amnióticas, el análisis de cariotipo requiere mucha mano de obra y unos costes tales que muchos hospitales no pueden pagar el procedimiento, lo que causa grandes dificultades en la difusión clínica y la aplicación de la amniocentesis.

55 Con el continuo desarrollo de técnicas de secuenciación, se está aplicando cada vez más en la detección y el análisis del número de cromosomas. Dennis Lo y col., utilizaron la sangre periférica de una mujer embarazada como material experimental para examinar la anomalía del número de cromosomas por medio de secuenciación masiva basada en procedimientos estadísticos matemáticos (Y.M. Dennis Lo, y col., Quantitative Analysis of Fetal DNA in

Maternal Plasma and Serum: Implications for Noninvasive Prenatal Diagnosis. Am. J. Hum. Genet. 62:768 – 775, 1998). Pero este procedimiento no puede reemplazar por completo a la amniocentesis en la aplicación clínica, debido a algunos defectos que se producen en la técnica: los ADN sin células en el plasma son ADN fragmentados, no pueden formar un genoma completo por su cuenta, la aneuploidía, translocación o mosaicismo de los cromosomas aparte de los cromosomas 21, 18 y 13 no se puede detectar o tiene una precisión de detección baja. Rossa W. K. Chiu y col., describe, en el artículo "Non-invasive prenatal diagnosis of fetal chromosomal aneuploidy by massively parallel genomic sequencing of DNA in maternal plasma" (PNAS vol. 105, no. 51, 23 December 2008), un procedimiento independiente de locus de la utilización de la secuenciación genómica masivamente paralela para cuantificar las secuencias de ADN del plasma materno para la detección prenatal no invasiva de la trisomía fetal 21.

10 **Resumen de la invención**

A fin de superar los resultados perdidos de detección causados por la secuenciación de sangre periférica, y resolver la alta tasa de fracaso del cultivo de células amnióticas, la presente invención, en combinación con las ventajas del análisis de cariotipo por amniocentesis y el procedimiento de secuenciación de los ADN sin células en plasma, utiliza un procedimiento de detección de aneuploidía cromosómica basado en la secuenciación masiva de células amnióticas, incluyendo las etapas de extracción de células amnióticas, aislamiento del ADN, realización de secuenciación de alto rendimiento, análisis de los datos obtenidos y la adquisición de los resultados de detección.

Un aspecto de la invención se expone en la reivindicación 1. En un ejemplo, se proporciona un procedimiento de utilización de la técnica de secuenciación de alto rendimiento para analizar la información cromosómica de células de un sujeto, que comprende las etapas de:

- a. romper al azar el ADN genómico de las células, obtener los fragmentos de ADN con un cierto tamaño, y secuenciarlos;
- b. alinear estrictamente las secuencias de ADN secuenciadas en la etapa a la secuencia de referencia del genoma humano para obtener la información acerca de las secuencias de ADN localizadas en un cromosoma en particular;
- c. para el cromosoma particular N, determinar el número total de las secuencias asignadas a una región única del cromosoma, entre las secuencias de ADN anteriormente secuenciadas, haciendo, de este modo, el % de CrN para el cromosoma N, es decir, la relación del número total (S1) de las secuencias de asignada al único lugar del cromosoma N, entre las secuencias de ADN anteriormente secuenciadas, con el número total (S2) de las secuencias localizadas en todos los cromosomas, entre las secuencias de ADN secuenciadas anteriormente: % de CrN $\% = S1/S2$;
- d. comparar el % de CrN para el cromosoma N con el % de CrN para el cromosoma correspondiente procedente de células estándar para determinar si existe una diferencia entre el cromosoma de las células y la correspondiente de las células estándar.

Las células pueden ser, por ejemplo, células amnióticas, en las que las células amnióticas pueden ser células amnióticas no cultivadas o células amnióticas cultivadas. En una realización de la invención, para evitar el cultivo de células amnióticas, las células amnióticas son células amnióticas no cultivadas.

El ADN genómico de las células puede obtenerse mediante procedimientos tradicionales de aislamiento de ADN, tales como precipitación salina, cromatografía en columna, y SDS, preferiblemente mediante cromatografía en columna. La denominada cromatografía en columna implica el uso de tampón de lisis celular y la proteasa K para tratar células amnióticas o tejidos para exponer moléculas de ADN desnudas, haciéndolas pasar a través de una columna de membrana de sílice capaz de unirse a moléculas de ADN cargadas negativamente, a la que las moléculas de ADN genómico en las sistema son adsorbidas de forma reversible, eliminando las impurezas tales como proteínas o lípidos mediante tampones de lavado y diluyendo mediante tampones de purificación para obtener el ADN de las células amnióticas (para más detalles acerca de los principios y procedimientos específicos, consulte los manuales de producto para el producto N ° 56304 de Qiagen y el producto DP316 de Tiangen).

Las moléculas de ADN se rompen al azar mediante escisión por restricción, atomización, ultrasonidos o procedimiento HydroShear. El procedimiento HydroShear se utiliza, preferiblemente, (cuando la solución que contiene el ADN está fluyendo a través de un pasaje con una sección pequeña, el caudal se acelera, creando una fuerza suficiente para destruir de repente el ADN para producir fragmentos de ADN de diferentes tamaños dependiendo del caudal y el área de la sección. Para más detalles acerca de los principios y procedimientos específicos, consulte los manuales del producto de HydroShear de Life Sciences Wiki). De esta manera, las moléculas de ADN se rompen en fragmentos con un estrecho intervalo de tamaños, de los cuales las principales bandas generalmente van desde 200 pb a 300 pb de tamaño.

El procedimiento de secuenciación adoptado en la invención puede ser el procedimiento de secuenciación de segunda generación, tal como Illumina / Solexa o ABI/SOLiD. En una realización de la invención, el procedimiento de secuenciación es Illumina / Solexa y las secuencias resultantes son fragmentos de 35 pb de tamaño.

Cuando las moléculas de ADN que se van a examinar proceden de múltiples muestras, cada muestra se puede tener unido un índice de secuencia marcado diferente para su procesamiento durante el proceso de secuenciación (Micah Hamady, Jeffrey J Walker, J Kirk Harris y col., Error-correcting Barcoded Primers for Pyrosequencing Hundreds of Samples in Multiplex. Nature Methods, 2008, March, Vol.5 No.3).

5 La secuencia de referencia del genoma humano se produce después de la protección de las secuencias repetidas dentro de la secuencia del genoma humano, por ejemplo, la última versión de la secuencia de referencia del genoma humano en la base de datos del NCBI. En una realización de la invención, la secuencia del genoma humano es la secuencia de referencia del genoma humano, como se muestra en la versión 36 (NCBI Build 36) de la base de datos del NCBI.

10 La alineación estricta con la secuencia de referencia del genoma humano significa que el procedimiento adoptado de alineación es una alineación intolerante al fallo de única región situada en la secuencia de referencia del genoma humano. En una realización de la invención, se utilizó software de alineación Eland (un paquete de software proporcionado por Illumina), y el procedimiento adoptado fue una alineación absoluta de intolerancia a fallos.

15 Cuando dichas secuencias de ADN es una secuencia que es capaz estar localizada en una región única de la secuencia de referencia del genoma humano, se define como la única secuencia representada por *Unique reads*. En la invención, con el propósito de evitar la interferencia de las secuencias repetidas, es necesario eliminar las secuencias de ADN localizadas en las regiones de repeticiones en tándem y de repeticiones transposicionales dentro de la secuencia de referencia del genoma humano y se limita a tener en cuenta aquellas secuencias de ADN, es decir, las secuencias individuales, que pueden estar situadas en una región única. En general, de todas las secuencias de ADN secuenciadas, de aproximadamente un cuarto a un tercio de las secuencias de ADN son capaces de estar ubicadas en una región única del genoma, es decir, secuencias únicas. El número estadístico de estas secuencias únicas representa la distribución de las secuencias de ADN en los cromosomas genómicos.

20 Por lo tanto, las secuencias únicas pueden ayudar en la localización de cada secuencia de ADN que se produce por la rotura y la secuenciación de las moléculas de ADN aisladas a partir de células amnióticas en un cromosoma particular. % de CrN corresponde a los valores producidos por la normalización de las secuencias únicas encontradas en diferentes cromosomas, y los valores son meramente relevantes para el tamaño de un cromosoma particular en lugar de la cantidad de los datos que se están secuenciando. Así, los valores pueden usarse para analizar la información sobre 46 cromosomas individuales. Por lo tanto, el % CrN es el valor básico para realizar un análisis cromosómico.

25 Si existe una diferencia entre el número de un cromosoma particular en las muestras celulares y las células estándar se puede determinar mediante la elaboración de un diagrama de caja, en el que se determina que una muestra para la que % CrN corresponde a un valor atípico que va más allá de 1,5-3 veces o por encima de 3 veces el rango intercuartil se determina difiere de las células estándar en el número de cromosomas, es decir, aneuploidía.

30 La determinación de si existe una diferencia entre un cromosoma particular, respectivamente, en dichas muestras celulares y en las muestras celulares estándar puede llevarse a cabo mediante el uso de "puntuación z_{CrN} " para indicar la desviación del % CrN para dichas muestras celulares de % CrN para las muestras celulares estándar.

Específicamente, $z = \text{puntuación}_{\% \text{ de CrN}} (\% \text{ de CrN para un cromosoma particular a partir de muestras de detección} - \text{media del \% de CrN (media}_{\% \text{ de CrN}} \text{ para el cromosoma particular)}) / \text{desviación estándar del \% de CrN (SD}_{\% \text{ de CrN}})$.

35 Si la puntuación z_{CrN} es extremadamente grande o pequeña, significa que la desviación del número de cromosomas en la muestra de detección celular con respecto de la de la muestra normal es significativa. Cuando alcanza un determinado nivel de significación, se puede creer que hay una aparente diferencia entre el primer número y el último número.

40 El valor promedio de % de CrN para un cromosoma particular de las muestras celulares estándar puede determinarse de acuerdo con el % de CrN para el cromosoma de tal como al menos 10, 20, 30, 50, o 100 muestras celulares estándar.

Las muestras celulares estándar son las muestras de células humanas en las que el número de cromosomas es diploide. Una célula masculina normal tiene 44 autosomas y 2 cromosomas sexuales diferentes, (46, XY). Por otra parte, una célula normal femenina tiene 44 autosomas y 2 cromosomas sexuales idénticos, (46, XX).

45 La desviación estándar del % de CrN (S.D._% de CrN) para un cromosoma particular de las muestras celulares estándar puede determinarse de acuerdo con el % de CrN para los cromosomas, tal como al menos 10, 20, 30, 50, o 100 muestras celulares estándar.

50 En una realización de la invención, las muestras celulares estándar tienen 20 muestras de varones normales y 10 muestras de mujeres normales, numeradas, respectivamente, con 1, 2, . . . , 30, en las que los nº 1 – 20 son las muestras de detección de varones normales, los nº 21 – 30 son las muestras de detección de mujeres normales. El valor promedio del % de CrN (media_% de CrN) para las muestras celulares estándar se calcula de la siguiente

manera:

$$\text{Media_ \% CrN} = \frac{1}{30} \sum_{m=1}^{30} \text{CrX-M}$$

(en la que N representa los autosomas 1-22, M representa las muestras normales nº 1 – 30)

$$\text{Media_ \% CrX (varón)} = \frac{1}{20} \sum_{m=1}^{20} \text{CrX-M}$$

5 (M representa las muestras normales de varones nº 1 – 20)

$$\text{Media_ \% CrY (varón)} = \frac{1}{20} \sum_{m=1}^{20} \text{CrY_M}$$

(M representa las muestras normales de varones nº 1 – 20)

$$\text{Media_ \% CrX (mujer)} = \frac{1}{10} \sum_{m=21}^{30} \text{CrX_M}$$

(M representa las muestras normales de varones nº 21 – 30)

$$\text{Media_ \% CrY (mujer)} = \frac{1}{10} \sum_{m=21}^{30} \text{CrY_M}$$

10 (M representa las muestras normales de varones nº 21 – 30)

(Nota: debido a la fluctuación de la secuenciación y un gran número de huecos existentes en el cromosoma Y de la secuencia de referencia, tiene como resultado que incluso para las muestras de mujeres normales hay algunas secuencias de ADN alineadas con el cromosoma Y. En comparación con los varones, sin embargo, el % de CrN para las mujeres es mucho menor que el de los varones. En los ejemplos, el % de CrN para las mujeres es de aproximadamente 0,004, mientras que el % de CrN para los varones es de aproximadamente 0,114.)

15 Sobre la base de cada % de CrN medio (media_%de CrN) para las muestras celulares estándar obtenidas por el procedimiento descrito anteriormente, la desviación estándar del % de CrN (SD_% de CrN) se calcula con la siguiente fórmula:

$$\text{SD_ \%CrN} = \frac{1}{30} \sum_{m=1}^{30} (\text{CrN_M} - \text{media_CrN})^2$$

20 (en la que N representa los autosomas 1 – 22)

$$\text{S.D. \%CrX (varón)} = \frac{1}{20} \sum_{m=1}^{20} (\text{CrX_M} - \text{media_CrX})^2$$

(M representa las muestras normales de varones nº 1 – 20)

$$\text{S.D. \%CrY (varón)} = \frac{1}{20} \sum_{m=1}^{20} (\text{CrY_M} - \text{media_CrY})^2$$

25 (M representa las muestras normales de varones nº 1 – 20)

$$\text{S.D. \%CrX (mujer)} = \frac{1}{10} \sum_{m=21}^{30} (\text{CrX_M} - \text{media_CrX})^2$$

(M representa las muestras normales de varones nº 21 – 30)

$$S.D. \text{ \%Cr Y (mujer)} = \frac{1}{20} \sum_{m=1}^{20} (CrX_M - media_CrX)^2$$

(M representa las muestras normales de varones nº 21 – 30)

5 Puesto que falta un cromosoma X sustituido por el cromosoma Y entre los cromosomas masculinos en contraste con los cromosomas femeninos, y toda la longitud del cromosoma X es de aproximadamente 155M, mientras que el cromosoma Y es de aproximadamente 59M. En la detección de estos cromosomas sexuales, es necesario establecer un conjunto de curvas de distribución normales relativas al % de CrX o el % de CrY para diferentes agendas. El análisis más preciso para el cromosoma X se puede obtener de las diferentes curvas de distribución normal basadas en el programa.

10 En una realización de la invención, se seleccionaron 30 muestras celulares estándar para llevar a cabo el análisis cromosómico. A continuación se estableció una curva de distribución normal bajo el requisito de nivel de significación (tal como 0,1 %) para la distribución normal alcanzada en el caso de tener una diferencia entre el número de cromosomas simulados y que en las células estándar. De este modo, el caso del valor absoluto de la puntuación z_{CrN} inferior a 3 se definió por el número de cromosomas que es igual al de las células estándar. Sobre la base de los resultados anteriores, después se analizaron los cromosomas de las muestras de detección del siguiente modo:

Si el valor absoluto de la puntuación z equivale a 3, las muestras tienen una probabilidad del 99,9 % de no encontrarse entre la población normalmente distribuida, es decir, los valores atípicos. Esto significa que el número de cromosomas de las células detectadas difiere del de las células estándar, es decir, aneuploidía cromosómica.

20 Si el valor absoluto de la puntuación z es menor que 3, las muestras tienen una probabilidad del 99,9% de que sean muestras normales, lo que significa que el número de cromosomas de las células detectadas es la misma que el de las células estándar.

Si el valor absoluto de la puntuación z es superior a 3, las muestras tienen una probabilidad del 99,9% de que sean muestras anormales, lo que significa que el número de cromosomas de las células detectadas difiere del de las células estándar, es decir aneuploidía cromosómica.

25 Además, si el valor absoluto de la puntuación z es mayor que 3, para el caso específico de que se produce aneuploidía cromosómica en las células detectadas, el valor de referencia Z (valor de corte) se puede utilizar para determinarla. El valor de referencia Z se calcula con la siguiente fórmula:

$$Z = (media_ \%CrN \times 0,5 \times \%X) / SD_ \%CrN$$

30 Cuando N representa autosomas, la $media_ \% CrN$ y $SD_ \% CrN$ son las medias para todas las muestras de las células estándar. Cuando N representa cromosomas sexuales, la $media_ \% CrN$ y $SD_ \% CrN$ son las medias para las muestras de las células estándar del respectivo programa;

X puede ser cualquier número entero entre, inclusive, -100 y 100, tales como -100, -90, -80, -70, -60, -50, -40, -30, -20, -10, 0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100.

35 En una realización de la invención, cuando X equivale a 100, representa que las muestras de detección celulares tienen un cromosoma más que las células estándar. En una realización de la invención, cuando X equivale a -100, representa que las muestras de detección celulares tienen un cromosoma menos que las células estándar. Cuando X equivale a 50, lo que representa que las muestras de detección celulares tienen una mitad más de un cromosoma que las células estándar. En una realización de la invención, cuando X equivale a 50, representa que las muestras de detección celulares carecen de la mitad de un cromosoma en comparación con las células estándar.

40 En el cálculo del valor de referencia Z para los cromosomas sexuales, la $media_ \% CrN$ y la $SD_ \% CrN$ se calculan tanto para las muestras femeninas como para las muestras masculinas. Es decir:

Para las muestras masculinas, el valor de referencia Z alcanzado puede ser $(media_ \% CrN \text{ (varón)} \times 0,5 \times \% X) / SD_ \% CrN \text{ (varón)}$;

45 Para las muestras femeninas, el valor de referencia Z alcanzado puede ser $(media_ \% CrN \text{ (mujer)} \times 0,5 \times \% X) / SD_ \% CrN \text{ (mujer)}$;

Cuando el valor absoluto de la puntuación z_{CrN} es mayor que o igual a 3 y alcanza el valor absoluto del valor de referencia Z , hay una diferencia significativa entre el número de un cromosoma en particular en las células detectadas y el de las células estándar igual a $\% X$. Por ejemplo, cuando X equivale a 50, representa que las células

detectadas tienen una mitad más del cromosoma particular con respecto a las células estándar; cuando X equivale a 100, representa que las células detectadas tienen un cromosoma particular más que las células estándar.

En una realización de la invención, el procedimiento específico de análisis de los cromosomas de las células amnióticas incluye las etapas siguientes:

5 1. Aislamiento y secuenciación de ADN

Se construyó una biblioteca de acuerdo con el procedimiento estándar modificado de Illumina / Solexa de construcción de una biblioteca después de aislar el ADN de las células amnióticas de conformidad con el manual de Tiangen Micro Kit. Después se añadieron adaptadores para la secuenciación a ambos extremos de los fragmentos de DNA rotos al azar. Durante el proceso, una secuencia marcada diferente (índice) se adjunta a cada una de las muestras de tal modo que múltiples muestras puedan diferenciarse en los datos obtenidos de la secuenciación de una sola vez.

2. Alineación y estadística

Después de la secuenciación mediante el uso de la técnica de secuenciación de segunda generación, conocida como secuenciación de Illumina / Solexa (también se pueden usar otros procedimientos de secuenciación alternativos para lograr los efectos iguales o similares efectos), las secuencias de ADN fragmentadas de un tamaño especificado se produjeron para cada muestra, y se sometieron a alineación estrictamente con la secuencia de referencia del genoma humano. Por lo tanto, la información se obtuvo de la ubicación de las secuencias en las regiones correspondientes del genoma.

Una alineación restringida de tal modo se requirió, ya que no se pudo determinar a partir qué cromosoma se originaba una secuencia de ADN dada si se permitió la alineación tolerante a fallos o la alineación con múltiples regiones. Esto sería desfavorable para el análisis posterior de los datos.

El número total de las secuencias únicas localizadas en cada cromosoma se calculó tomando cada cromosoma como una unidad, por lo tanto haciendo que el % de CrN para cada cromosoma, es decir, la relación del número total (S1) de las secuencias entre las secuencias de ADN secuenciadas anteriormente, que se encuentran en el lugar único del cromosoma N, y el número total (S2) de las secuencias entre las secuencias de ADN secuenciadas anteriormente, que se encuentran en todos los cromosomas $\% \text{ de CrN} = S1/S2$.

Este es un procedimiento de normalización para diferentes muestras que tienen una cantidad diferente de secuenciación. Dado que las células amnióticas contienen toda la información de 46 cromosomas, en teoría el número total de las secuencias localizadas en un cromosoma dado es directamente proporcional a la longitud del cromosoma.

Por ejemplo, el cromosoma 1 es el cromosoma más grande (de aproximadamente 247 M) en el genoma humano, mientras que el cromosoma 21 es el cromosoma más pequeño (de aproximadamente 47 M) en el genoma humano, por lo tanto, dada una cierta cantidad total de secuenciación, los resultados de la secuenciación de células amnióticas diploides normales son casi un valor fijo. Aunque en algunas condiciones de secuenciación y experimentales, el % de CrN no es directamente proporcional al tamaño de los cromosomas, por lo general es un valor fijo.

3. Análisis de datos

Mediante un análisis de gráfico de caja se puede determinar directamente si es probable que las muestras de detección celulares difieran de las células estándar en el número de cromosomas. Las muestras con valores anormales se consideran directamente como muestras sospechosas, y las otras se consideran como muestras celulares estándar. El procedimiento detallado es el siguiente:

Con el fin de ayudar en el análisis de datos, se adopta un diagrama de caja (utilizado para diferenciar los valores anormales en estadística matemática) que implica el % de CrN producido anteriormente para determinar las muestras sospechosas. El procedimiento de dibujo específico es el siguiente:

- 45 1) calcular el cuartil superior (75 %), la mediana (50 %), y el cuartil inferior (25 %);
- 2) calcular la diferencia, el rango intercuartil (IQR), entre el cuartil superior y el cuartil inferior;
- 3) dibujar los rangos superiores e inferiores de un diagrama de caja, siendo el límite superior el cuartil superior y siendo el límite inferior el cuartil inferior y dibujar una línea horizontal en la que la mediana se encuentra dentro de la caja;
- 50 4) Los valores que son 1,5 veces mayores que el cuartil superior del rango intercuartil o 1,5 veces menor que el cuartil menor del rango intercuartil se clasifican como valores atípicos.
- 5) más allá de los valores atípicos, dibujar una línea horizontal a través de los dos puntos de valor más cercano al margen superior y al margen inferior, respectivamente, como un "bigote" del diagrama de caja.
- 55 6) valores atípicos extremos que van más allá de una distancia tres veces mayor que el rango intercuartil se representan con puntos de estrella; los valores atípicos leves que se encuentran dentro de 1,5-3 veces la

distancia del intercuartil se representan con puntos huecos.

5 La línea en negrita en el medio de la caja representa la mediana de los valores y los límites superior e inferior representan los cuartiles superior e inferior, respectivamente. Los valores atípicos se definen por los puntos que se desvían de 1,5 veces la distancia entre el cuartil superior y el cuartil inferior. Por ejemplo, cuando las muestras de detección son muestras celulares estándar, el % de CrN correspondiente a sus cromosomas es un valor fijo (por ejemplo, 1). Cuando el % de CrN correspondiente a sus cromosomas es 1,5, la diferencia se puede considerar en gran medida significativa, lo que hace que las muestras sean muestras sospechosas. Es decir, es probable que sean muestras diferentes de las células estándar en el número de cromosomas.

10 Si es necesario, la media del y la desviación estándar del % de CrN (SD_% CrN) se pueden determinar, respectivamente, mediante el % de CrN para un cromosoma particular correspondiente a las muestras celulares estándar. Por tanto, la puntuación z_CrN para el cromosoma de las muestras sospechosas se calcula con la siguiente fórmula:

Puntuación z_CrN= (% de CrN % para el cromosoma particular de las muestras – media del % de CrN) / S.D._del % CrN

15 Si el valor absoluto de la puntuación z_CrN es mayor o igual a 3, hay una diferencia entre el número de un cromosoma particular en las muestras celulares y el de las células estándar.

Además, para el caso específico de un número de cromosomas anormal en las células, la referencia al valor de referencia Z (valor de corte) se puede utilizar para determinarlo. El valor Z se calcula con la siguiente fórmula:

20 **Z = (media_% CrN x 0,5 x % X) / S.D._% CrN** Cuando N representa autosomas, la media_% CrN y la SD_% CrN son las media de todas las muestras de las células estándar. Cuando N representa cromosomas sexuales, la media_% CrN y la SD_% CrN son las medias para las muestras de las células estándar del respectivo programa;

Se asigna a X 50 o 100. Correspondientemente, cuando X es 100, que representa que las muestras de detección celulares tienen un cromosoma más que las células estándar. Cuando X es 50, lo que representa que las muestras de detección celulares tienen una mitad más de un cromosoma que las células estándar.

25 En el cálculo del valor de referencia Z para los cromosomas sexuales, la media_% CrN y la SD_% CrN son la media tanto para las muestras femeninas como para las muestras masculinas, es decir:

Para las muestras masculinas, el valor de referencia Z alcanzado es (media_% CrN (varón) × 0,5 × % X) / SD_% CrN (varón);

30 Para las muestras femeninas, el valor de referencia Z alcanzado es (media_% CrN (mujer) × 0,5 × % X) / SD_% CrN (mujer).

Cuando el valor absoluto de la puntuación z_CrN es mayor que o igual a 3 y alcanza el valor absoluto del valor de referencia Z, hay una diferencia de %X entre el número de un cromosoma en particular en las células y el de las células estándar.

Ventajas de la invención

35 La invención se puede utilizar para el análisis de células, tales como células amnióticas. De acuerdo con la invención, el ADN puede aislarse directamente a partir de células amnióticas para su detección sin un subcultivo, lo que disminuye en gran medida las dificultades de una unión tan difícil, un número insuficiente, o el fracaso del cultivo causado por el cultivo de células amnióticas

40 Mediante el uso de la característica de las células amnióticas que contienen toda la información genómica sobre un feto, la invención es capaz de hacer un análisis de la aneuploidía de todos los cromosomas de las células, en lugar de examinar sólo los cromosomas sexuales X e Y y los cromosomas 21, 18, 13.

45 Además, aunque el procedimiento de determinación involucrado en la invención, en comparación con las muestras de plasma, también depende de la distribución aproximadamente normal establecida sobre muestras celulares estándar, tal dependencia de las muestras celulares estándar se reduce considerablemente. Además, las muestras anormales pueden determinarse directamente a partir de anomalías de datos, suponiendo datos suficientes.

Al utilizar el procedimiento de la invención, un gran número de muestras celulares de detección pueden someterse a análisis por lotes. Cientos de miles de muestras de detección celulares se pueden detectar a la vez, y así ahorrar en gran medida trabajos y costes.

Breve descripción de las figuras

La figura 1 muestra un diagrama de caja representado de conformidad con 53 muestras celulares de detección, en la que la abscisa representa el número de cromosomas, y la ordenada representa el valor de % CrN.

5 La Figura 1A muestra los cromosomas 1-6, la figura 1B muestra los cromosomas 7 – 12, la figura 1C muestra los cromosomas 13-17, y la figura 1D muestra los cromosomas 18-22.

Descripción detallada de las realizaciones preferidas

10 La forma de realización de la invención se describirá con detalle en combinación con los siguientes ejemplos. Una persona experta en la técnica apreciará, sin embargo, que los siguientes ejemplos son meramente con la intención de hacer una descripción de la invención y no se considerarían como la limitación del alcance de la invención. Si las condiciones específicas no están especificadas en los ejemplos, estos ejemplos se realizan de acuerdo con las condiciones de uso habitual o con las recomendadas por los fabricantes. Si no se especifican las fuentes de los reactivos o equipos o instrumentos utilizados en la invención, tales como sus fabricantes, todos ellos están disponibles comercialmente. Los conectores utilizados para la secuenciación y el índice de secuencias marcadas proceden del kit Multiplexing Sample Preparation Oligonutide Kit proporcionado por Illumina.

15 En los siguientes paréntesis está el número de producto de los fabricantes para cada uno de los reactivos o kits.

Ejemplo 1: Análisis cromosómico de células amnióticas no cultivadas**1. Aislamiento y secuenciación de ADN**

20 El ADN de las células amnióticas se aisló de acuerdo con el procedimiento de manipulación de una pequeña cantidad de genoma de Tiangen Micro Kit (DP316) y se cuantificó con Qubit (Invitrogen, el Kit Quant-iT™ dsDNA HS Assay). La cantidad total de ADN aislado varió de 100 ng a 500 ng.

25 El ADN aislado era o todo el ADN genómico o ADN de tipo frotis degradado parcialmente. Se construyó una biblioteca de ADN conforme al procedimiento estándar de creación de bibliotecas proporcionada por el Illumina / Solexa modificado. Se añadieron adaptadores en ambos extremos de las moléculas de ADN rotas al azar y se unieron con diferentes índices de secuencia marcados. A continuación, estas moléculas se hibridaron con adaptadores complementarios en la superficie de una célula de flujo, y se dejaron agrupar en condiciones particulares. Se llevaron a cabo 36 ciclos de secuenciación en un analizador Illumina Genome Analyzer II, produciendo fragmentos de ADN con 35 pb.

30 Específicamente, se utilizó Diagenode Bioruptor para romper al azar aproximadamente 100-500ng de ADN aislado de células amnióticas en fragmentos de 300 pb. 100-500ng del ADN roto inicialmente se utilizaron para construir una biblioteca bajo Illumina / Solexa. Véase en la técnica anterior el un procedimiento detallado (manual de Illumina / Solexa para la creación de la biblioteca estándar proporcionado por <http://www.illumina.com/>). El tamaño de la biblioteca de ADN se determinó por medio del 2100 Bioanalyzer (Agilent) y los fragmentos insertados fueron de 300 pb. Después de la cuantificación exacta mediante QPCR, se realizó la secuenciación.

35 En el ejemplo, la secuenciación por lotes se realizó en 53 muestras de ADN aisladas de las células amnióticas de acuerdo a Cluster Station and GA II x (SE sequencing) publicados oficialmente por Illumina / Solexa.

2. Alineación y estadística

40 Consúltase en la técnica anterior (véase el procedimiento Pipeline relativo al manual proporcionado en <http://www.illumina.com/>), la información sobre secuencias obtenida en la etapa 1 se sometió a un solo proceso Pipeline y se eliminaron las secuencias con baja calidad, lo que en última instancia tuvo como resultado un resultado de la alineación de ELAND contra la secuencia de referencia del genoma humano de NCBI versión 36. A continuación, se analizó estadísticamente el número de las secuencias únicas localizadas en los cromosomas.

Se calculó el % de CrN para 22 cromosomas y los cromosomas X / Y, respectivamente, de 53 muestras y un diagrama de caja (ver figura 2) se ha elaborado en base a los datos El % de CrN para un cromosoma N particular en una muestra concreta M se calcula con la fórmula siguiente:

45 **Porcentaje e un cromosoma particular en la muestra de detección M, el % de CrN= número total de las secuencias únicas contenidas en la muestra M y localizadas en el correspondiente cromosoma de la secuencia de referencia a través del alineamiento (81)/el número total de las secuencias únicas contenidas en la muestra M y localizadas en todos los cromosomas de la secuencia de referencia mediante alineación (82).**

3. Análisis de datos

De acuerdo con el diagrama de caja dibujado en la etapa 2, primero se determinó si existía un valor atípico. Es decir, en comparación con los límites superior e inferior, si se una muestra sospechosa se ha desviado lejos del punto que

era 1,5 veces la diferencia entre el cuartil superior y el cuartil inferior, era probable que difiriera de las muestras estándar en el número de cromosomas.

5 Específicamente, se observó la distribución del diagrama de caja, y se detectaron 8 muestras sospechosas (nº de muestras P1–P8). Una distribución normal se estableció mediante el uso de muestras estándar como los datos relativos a 20 varones normales y 10 mujeres normales, elegida al azar a partir de las 45 muestras celulares estándar que quedan después de retirar las muestras sospechosas. El % de CrN medio (medio_ % de CrN) para cada cromosoma se designa por la media_% CrN y la desviación estándar se da en la tabla 1 (S.D. _% CrN).

Tabla 1: % de CrN, media y desviación estándar (S.D.) para cada cromosoma en células normales

Número	% Cr1	% Cr2	% Cr3	% Cr4	% Cr5	% Cr6	% Cr7	% Cr8	% Cr9	% Cr10	% Cr11	% Cr12
1	7,850	9,295	7,542	7,381	6,876	6,660	5,499	5,468	3,942	4,747	4,603	4,744
2	7,965	9,088	7,483	6,907	6,686	6,482	5,494	5,331	4,040	4,834	4,645	4,699
3	7,935	9,121	7,414	6,989	6,649	6,491	5,500	5,324	4,035	4,834	4,704	4,695
4	7,866	9,237	7,618	7,353	6,875	6,585	5,530	5,424	3,976	4,722	4,546	4,743
5	7,847	9,179	7,371	7,100	6,784	6,509	5,535	5,374	3,988	4,843	4,587	4,721
6	7,752	9,247	7,617	7,337	6,871	6,600	5,573	5,401	3,960	4,776	4,564	4,742
7	7,920	9,149	7,501	7,178	6,826	6,607	5,515	5,360	4,003	4,792	4,598	4,748
8	8,089	8,953	7,289	6,614	6,532	6,317	5,462	5,237	4,059	4,922	4,711	4,698
9	8,155	9,005	7,190	6,459	6,565	6,267	5,529	5,238	4,154	4,950	4,819	4,573
10	7,961	9,133	7,362	6,768	6,627	6,418	5,520	5,280	3,987	4,822	4,674	4,734
11	8,079	8,980	7,217	6,602	6,493	6,294	5,504	5,219	4,078	4,921	4,770	4,712
12	7,953	9,205	7,499	7,173	6,786	6,533	5,535	5,343	3,977	4,789	4,654	4,664
13	7,986	9,051	7,360	6,848	6,701	6,446	5,541	5,376	4,065	4,822	4,701	4,689
14	8,111	9,040	7,210	6,905	6,548	6,364	5,472	5,341	4,012	4,884	4,725	4,677
15	8,032	9,002	7,325	6,818	6,571	6,369	5,537	5,345	4,064	4,913	4,672	4,682
16	8,075	8,977	7,199	6,664	6,515	6,349	5,480	5,311	4,068	4,856	4,741	4,703
17	7,878	9,184	7,502	7,221	6,793	6,592	5,523	5,405	3,990	4,763	4,588	4,727
18	7,873	9,165	7,502	7,194	6,775	6,553	5,496	5,384	4,025	4,786	4,629	4,755
19	7,911	9,119	7,574	7,286	6,830	6,507	5,539	5,365	3,949	4,781	4,577	4,695
20	8,013	9,186	7,394	6,822	6,634	6,384	5,475	5,297	4,033	4,849	4,657	4,678
21	7,739	8,991	7,232	6,822	6,503	6,260	5,374	5,214	3,941	4,797	4,629	4,591
22	7,887	8,962	7,178	6,730	6,471	6,259	5,353	5,215	3,978	4,754	4,608	4,566
23	7,921	8,903	7,258	6,803	6,438	6,285	5,372	5,234	3,973	4,739	4,624	4,626
24	7,807	8,900	7,271	6,954	6,526	6,320	5,357	5,283	3,968	4,772	4,585	4,609
25	7,681	9,020	7,359	7,172	6,685	6,440	5,421	5,255	3,922	4,667	4,502	4,665
26	7,892	8,827	7,205	6,728	6,529	6,305	5,363	5,123	3,914	4,847	4,616	4,557
27	8,071	8,730	6,993	6,186	6,159	6,034	5,267	5,150	4,044	4,919	4,790	4,556
28	7,878	8,771	7,059	6,452	6,418	6,210	5,318	5,208	3,962	4,912	4,593	4,613
29	7,803	8,985	7,335	6,934	6,613	6,382	5,428	5,268	3,892	4,744	4,510	4,658

(continuación)

Número	% Cr1	% Cr2	% Cr3	% Cr4	% Cr5	% Cr6	% Cr7	% Cr8	% Cr9	% Cr10	% Cr11	% Cr12
30	7,992	8,818	7,126	6,540	6,363	6,167	5,329	5,089	4,028	4,820	4,689	4,548
media	7,931	9,041	7,339	6,898	6,621	6,400	5,461	5,295	4,001	4,819	4,644	4,669
S.D.	0,116	0,146	0,163	0,300	0,172	0,148	0,082	0,091	0,057	0,069	0,078	0,065
numero	% Cr13	% Cr14	% Cr15	% Cr16	% Cr17	% Cr18	% Cr19	% Cr20	% Cr21	% Cr22	% CrX	% CrY
1	3,917	3,294	2,762	2,401	2,380	3,051	1,120	1,966	1,295	0,920	2,163	0,126
2	3,743	3,279	2,885	2,595	2,650	2,980	1,431	2,148	1,296	1,088	2,128	0,123
3	3,770	3,268	2,914	2,565	2,602	3,027	1,434	2,112	1,270	1,073	2,154	0,119
4	3,944	3,254	2,801	2,416	2,402	3,080	1,140	2,000	1,279	0,914	2,170	0,126
5	3,880	3,234	2,866	2,545	2,576	3,049	1,354	2,064	1,275	1,050	2,154	0,112
6	3,909	3,322	2,792	2,393	2,389	3,089	1,139	1,966	1,309	0,910	2,217	0,127
7	3,868	3,305	2,833	2,451	2,481	3,002	1,266	2,033	1,327	0,942	2,165	0,129
8	3,585	3,258	2,943	2,800	2,863	2,996	1,685	2,262	1,296	1,223	2,091	0,116
9	3,590	3,205	2,926	2,860	2,960	2,879	1,669	2,271	1,339	1,240	2,046	0,111
10	3,714	3,305	2,873	2,677	2,767	3,001	1,539	2,168	1,307	1,154	2,085	0,122
11	3,638	3,303	2,927	2,787	2,872	2,947	1,687	2,244	1,294	1,275	2,041	0,115
12	3,823	3,285	2,829	2,493	2,504	3,012	1,323	2,068	1,293	1,025	2,116	0,120
13	3,734	3,282	2,819	2,676	2,670	3,002	1,479	2,100	1,299	1,097	2,135	0,121
14	3,647	3,274	2,894	2,746	2,758	2,939	1,501	2,256	1,308	1,147	2,120	0,121
15	3,720	3,262	2,891	2,663	2,730	2,976	1,551	2,176	1,287	1,164	2,122	0,130
16	3,584	3,269	2,970	2,767	2,868	2,965	1,667	2,186	1,270	1,277	2,125	0,114
17	3,864	3,246	2,827	2,475	2,476	3,036	1,228	2,024	1,297	0,963	2,277	0,120
18	3,853	3,315	2,844	2,506	2,489	3,040	1,229	2,071	1,299	0,963	2,129	0,124
19	3,950	3,312	2,845	2,457	2,484	3,065	1,234	2,031	1,265	0,962	2,136	0,125
20	3,731	3,285	2,886	2,645	2,720	2,906	1,485	2,155	1,299	1,151	2,200	0,114
21	3,690	3,224	2,829	2,522	2,618	2,919	1,384	2,119	1,269	1,107	4,224	0,003
22	3,642	3,223	2,877	2,578	2,673	2,910	1,422	2,137	1,259	1,087	4,226	0,004
23	3,680	3,234	2,824	2,536	2,585	2,968	1,380	2,098	1,265	1,062	4,187	0,004
24	3,707	3,223	2,809	2,501	2,580	2,933	1,372	2,039	1,262	1,042	4,177	0,004
25	3,829	3,214	2,740	2,388	2,390	2,978	1,211	1,975	1,259	0,929	4,293	0,004
26	3,679	3,232	2,859	2,579	2,707	2,918	1,485	2,081	1,262	1,129	4,159	0,003
27	3,404	3,207	2,912	2,899	2,968	2,893	1,802	2,344	1,263	1,330	4,077	0,003

(continuación)

número	% Cr13	% Cr14	% Cr15	% Cr16	% Cr17	% Cr18	% Cr19	% Cr20	% Cr21	% Cr22	%CrX	% CrY
28	3,525	3,190	2,869	2,771	2,852	2,903	1,647	2,215	1,278	1,227	4,127	0,003
29	3,786	3,177	2,778	2,456	2,498	2,980	1,258	2,037	1,271	0,986	4,216	0,003
30	3,569	3,224	2,913	2,699	2,816	2,898	1,603	2,168	1,284	1,179	4,133	0,003
media	3,733	3,257	2,858	2,595	2,644	2,978	1,424	2,117	1,286	1,087	—	—
S.D.	0,135	0,040	0,055	0,147	0,175	0,060	0,187	0,099	0,021	0,120	—	—
media-% Cr X/Y-V	—	—	—	—	—	—	—	—	—	—	2,139	0,121
S.D.-% Cr X/Y-V	—	—	—	—	—	—	—	—	—	—	0,055	0,006
media-% Cr X/Y-M	—	—	—	—	—	—	—	—	—	—	4,182	0,003
S.D.-% Cr X/Y-M	—	—	—	—	—	—	—	—	—	—	0,062	0,001

Por otra parte, con el fin de examinar si existía el caso de medio o un cromosoma más en las muestras sospechosas, se asignó a X un valor de 50 o 100 y se calculó el correspondiente valor de referencia cromosómica Z (valor de corte) (véase la tabla 2):

$$Z = (\text{media_ \% Cr N} \times 0,5 \times \% X) / \text{S.D._ \% Cr N}, \text{ en la que N representa los cromosomas } 1 \sim 22, X \text{ es } 50 \text{ o } 100.$$

5 Tabla 2: Determinación del valor de referencia Z (valor de corte) para trisomía en las células de detección

Número de cromosoma	media_ % CrN	S.D._ % CrN	Valor de referencia Z	
			por encima del 50 %	por encima del 100 %
Cr1	7,9307295	0,1159668	17,0969881	34,1939762
Cr2	9,0408152	0,1458970	15,4917815	30,9835629
Cr3	7,3394728	0,1633916	11,2298783	22,4597567
Cr4	6,8980316	0,3000703	5,7470125	11,4940249
Cr5	6,6213096	0,1717845	9,6360687	19,2721373
Cr6	6,3996516	0,1482759	10,7901077	21,5802154
Cr7	5,4613809	0,0819567	16,6593481	33,3186962
Cr8	5,2953272	0,0912159	14,5131676	29,0263351
Cr9	4,0009382	0,0572472	17,4721938	34,9443876
Cr10	4,8192433	0,0693516	17,3725054	34,7450108
Cr11	4,6436180	0,0780559	14,8727280	29,7454561
Cr12	4,6690652	0,0647597	18,0245772	36,0491544
Cr13	3,7325287	0,1346539	6,9298575	13,8597151
Cr14	3,2568057	0,0398485	20,4324240	40,8648480
Cr15	2,8578247	0,0553827	12,9003599	25,8007199
Cr16	2,5948409	0,1465308	4,4271266	8,8542531
Cr17	2,6442999	0,1753843	3,7692929	7,5385857
Cr18	2,9780305	0,0598664	12,4361514	24,8723027
Cr19	1,4242524	0,1868697	1,9054083	3,8108167
Cr20	2,1171964	0,0985564	5,3705200	10,7410400
Cr21	1,2858576	0,0205784	15,6214179	31,2428358
Cr22	1,0872769	0,1204085	2,2574758	4,5149516
CrX-M	4,1819271	0,0615940	-16,9737663	-33,9475326
CrY-M	0,0034667	0,0006359	/	/
CrX-V	2,1387665	0,0545815	9,7962061	19,5924123
CrY-V	0,1207910	0,0055482	5,4468013	10,8856025

La puntuación z_CrN para cada cromosoma en las muestras sospechosas se calculó con la siguiente fórmula:

$$\text{Puntuación } z_{\text{Cr N}} = (\% \text{Cr N para un cromosoma dado en las muestras de detección} - \text{media_ \% Cr N}) / \text{S.D._ \% Cr N}.$$

Tabla 3: Puntuación z_CrN para cada cromosoma en las muestras sospechosas

Número de cromosoma	Media	SD	P1	P2	P3	P4	P5	P6	P7	P8
Cr1	7,931	0,116	-1,566	0,349	-0,274	2,466	-0,537	-0,733	-0,028	-2,100
Ch2	9,041	0,146	-0,843	0,799	-1,326	-1,063	-0,771	-0,605	-0,341	-0,856
Cr3	7,339	0,163	-0,487	0,498	-1,781	-0,924	-0,346	-0,394	0,256	0,129
Cr4	6,898	0,300	0,238	0,988	-1,045	-1,548	-0,469	-0,177	0,151	0,774
Cr5	6,621	0,172	-0,726	-0,122	-1,514	-1,370	-0,522	-0,501	-0,137	-0,071
Cr6	6,400	0,148	-0,572	0,174	-0,929	-1,521	-0,509	-0,181	-0,281	-0,172
Cr7	5,461	0,082	-0,582	-0,312	-1,889	-1,166	-0,674	0,073	-0,292	-1,247
Cr8	5,295	0,091	-0,618	0,492	-2,122	-2,192	-0,863	-0,406	-1,501	-0,336
Cr9	4,001	0,057	-1,659	0,700	0,747	-2,107	-0,038	-0,916	-0,546	-1,032
Cr10	4,819	0,069	-0,644	-0,051	-0,385	0,131	-0,174	-0,849	-0,997	-0,712
Cr11	4,644	0,078	-2,056	-0,749	-0,513	1,071	-0,147	-0,764	-1,099	-2,300
Cr12	4,669	0,065	-1,121	-0,489	-1,854	-2,140	-0,616	-0,562	-0,137	-0,682
Cr13	3,733	0,135	-0,163	0,544	-1,460	-2,573	-0,515	-0,102	0,514	7,393
Cr14	3,257	0,040	-1,301	-1,548	-1,184	-1,124	-0,268	-1,574	-0,683	-0,763
Cr15	2,858	0,055	-1,966	-1,048	-0,421	0,865	0,465	0,044	1,002	-0,542
Cr16	2,595	0,147	-0,963	-0,739	0,305	0,779	-0,096	-0,336	-0,405	-1,787
Cr17	2,644	0,175	-0,913	-0,766	0,394	0,846	0,278	-0,236	-0,561	-1,736
Cr18	2,978	0,060	-0,111	0,506	-2,574	-0,127	23,853	26,141	24,438	-0,951
Cr19	1,424	0,187	-0,586	-0,664	0,660	0,340	0,191	-0,141	-0,608	-1,489
Cr20	2,117	0,099	-0,907	-0,697	0,211	0,210	-0,002	-0,018	-0,659	-0,023
Cr21	1,286	0,021	27,203	31,507	29,820	21,831	-0,305	-1,262	-0,729	-0,750
Cr22	1,087	0,120	-0,604	-0,743	0,616	0,747	-0,026	-0,220	-0,467	-1,310
CrX-V	2,139	0,055	/	0,650	/	/	-0,967	-0,678	-1,963	/
CrY-V	0,121	0,006	/	-1,117	/	/	-0,551	0,215	-0,197	/
CrXM	4,182	0,062	-1,347	/	-1,428	-2,272	/	/	/	0,700
CrY-M	0,003	0,001	0,881	/	-0,741	0,616	/	/	/	-0,303

5 Como se muestra en el análisis anterior, las muestras sospechosas fueron 8 en total entre las 53 muestras de detección de células amnióticas, en el que, para los cromosomas en cada una de las muestras sospechosas, se detectaron 8 anomalías del número de cromosoma con el valor absoluto de una puntuación z_CrN mayor de 3 fueron detectados (véase la tabla 3). En concreto, fueron los siguientes:

- 10 1) Cr21 para P1, Cr21 para P2, Ch21 para P3, y Cr21 para P4;
 2) Ch18 para P5, Cr18 para P6, y Cr18 para P7; y
 3) Cr13 para P8.

15 Mediante la comprobación del valor Z obtenida cuando X = 100 en la Tabla 2 se determinó que el número de cromosoma 21 en las muestras P1-P4 y el número del cromosoma 18 en las muestras P5-P7 era uno más que el número de cromosomas correspondientes en las células estándar, respectivamente, y el número del cromosoma 13 en P8 era medio más que el número de cromosomas correspondiente en las células estándar. Es decir, P1-P4 fueron T21 (síndrome de Down), y P5-P7 fueron T18 (síndrome de Edwards) y P8 fue T13 mosaico (síndrome del mosaico de Patau). Los resultados eran completamente coherentes con los resultados del análisis tradicional del cariotipo cromosómico.

Ejemplo 2

6 muestras adicionales (Q1-Q6) de células amnióticas se trataron y se secuenciaron de la misma forma que el anterior para producir los datos para el análisis. La puntuación z_{CrN} se calculó en la media_% CrN y S.D. _% CrN de 30 muestras celulares estándar en el ejemplo 1. Se identificaron 3 muestras positivas en las 6 muestras.

5

Tabla 4: El % de CrN se 6 muestras de detección (Q1–Q6)

	Q1	Q2	Q3	Q4	Q5	Q6
Cr1 %	7,900099	7,781541	7,965013	7,937310	7,835625	7,756449
Cr2 %	9,195581	8,969471	8,998068	9,137041	8,836921	9,014913
Cr3 %	7,389485	7,365766	7,389563	7,452117	7,134356	7,378118
Cr4 %	7,090694	7,005976	6,921334	7,112517	6,510565	7,058824
Cr5 %	6,759707	6,600836	6,604984	6,768853	6,357255	6,605637
Cr6 %	6,541994	6,468799	6,461957	6,545516	6,170975	6,376054
Cr7 %	5,562187	5,423140	5,522745	5,521768	5,342112	5,403700
Cr8 %	5,387074	5,344078	5,357094	5,318220	5,176933	5,275211
Cr9 %	3,946516	3,924984	4,061791	4,037918	4,007161	3,958868
Cr10 %	4,831699	4,680082	4,876470	4,845395	4,798439	4,700673
Cr11 %	4,634992	4,541423	4,682686	4,637077	4,603257	4,462601
Cr12 %	4,727456	4,552861	4,734034	4,700509	4,571935	4,594756
Cr13 %	3,871131	3,749202	3,677764	3,875716	3,475357	3,776552
Cr14 %	3,261377	3,247342	3,285671	3,281681	3,238633	3,207609
Cr15 %	2,875226	2,782605	2,926826	2,866104	2,830466	2,757703
Cr16 %	2,516559	2,443884	2,624248	2,524383	2,665946	2,413713
Cr17 %	2,519897	2,481007	2,684055	2,561488	2,725292	2,495129
Cr18 %	3,026389	2,939323	3,027751	2,994205	2,893438	2,985936
Cr19 %	1,291162	1,240504	1,479644	1,317938	1,482326	1,235699
Cr20 %	2,096249	2,063225	2,174512	2,076472	2,173705	2,049467
Cr21 %	1,290966	1,267192	1,297270	1,291611	1,896099	1,297050
Cr22 %	0,992960	0,989674	1,125718	1,017386	1,164497	0,995698
CrX %	2,173990	4,134076	2,117655	2,177186	4,104752	4,197133
CrY %	0,116611	0,003010	0,003148	0,001590	0,003956	0,002508

Tabla 5: La puntuación z_{CrN} para 6 muestras calculada a partir de la media_% Cr N y la S.D. _% Cr N de 30 muestras negativas en el ejemplo 1

Número de cromosoma	media	SD	Q1	Q2	Q3	Q4	Q5	Q6
Cr1	7,9310	0,1160	-0,2664	-1,2884	0,2932	0,0544	-0,8222	-1,5048
Cr2	9,0410	0,1460	1,0588	-0,4899	-0,2941	0,6578	-1,3978	-0,1787
Cr3	7,3390	0,1630	0,3097	0,1642	0,3102	0,6940	-1,2555	0,2400
Cr4	6,8980	0,3000	0,6423	0,3599	0,0778	0,7151	-1,2914	0,5361
Cr5	6,6210	0,1720	0,8064	-0,1172	-0,0931	0,8596	-1,5334	-0,0893
Cr6	6,4000	0,1480	0,9594	0,4649	0,4186	0,9832	-1,5475	-0,1618
Cr7	5,4610	0,0820	1,2340	-0,4617	0,7530	0,7411	-1,4499	-0,6988
Cr8	5,2950	0,0910	1,0118	0,5393	0,6823	0,2552	-1,2974	-0,2175

(continuación)

Número de cromosoma	media	SD	Q1	Q2	Q3	Q4	Q5	Q6
Cr9	4,0010	0,0570	-0,9559	-1,3336	1,0665	0,6477	0,1081	-0,7392
Cr10	4,8190	0,0690	0,1840	-2,0133	0,8329	0,3825	-0,2980	-1,7149
Cr11	4,6440	0,0780	-0,1155	-1,3151	0,4960	-0,0888	-0,5223	-2,3256
Cr12	4,6690	0,0650	0,8993	-1,7868	1,0005	0,4848	-1,4933	-1,1422
Cr13	3,7330	0,1350	1,0232	0,1200	-0,4092	1,0572	-1,9085	0,3226
Cr14	3,2570	0,0400	0,1094	-0,2414	0,7168	0,6170	-0,4592	-1,2348
Cr15	2,8580	0,0550	0,3132	-1,3708	1,2514	0,1473	-0,5006	-1,8236
Cr16	2,5950	0,1470	-0,5336	-1,0280	0,1990	-0,4804	0,4826	-1,2332
Cr17	2,6440	0,1750	-0,7092	-0,9314	0,2289	-0,4715	0,4645	-0,8507
Cr18	2,9780	0,0600	0,8065	-0,6446	0,8292	0,2701	-1,4094	0,1323
Cr19	1,4240	0,1870	-0,7104	-0,9813	0,2976	-0,5672	0,3119	-1,0070
Cr20	2,1170	0,0990	-0,2096	-0,5432	0,5809	-0,4094	0,5728	-0,6822
Cr21	1,2860	0,0210	0,2365	-0,8956	0,5367	0,2672	29,0523	0,5262
Cr22	1,0870	0,1200	-0,7837	-0,8111	0,3226	-0,5801	0,6458	-0,7609
CrX-V	2,1390	0,0550	0,6362	/	-0,3881	0,6943	/	/
CrY-M	0,1210	0,0060	-0,7315	/	-19,6420	-19,9016	/	/
CrX-V	4,1820	0,0620	/	-0,7730	-33,2958	-32,3357	-1,2459	0,2441
CrY-V	0,0030	0,0010	/	0,0100	-1,4100	-1,41/	0,9564	-0,4919

5 Como se ve por los resultados, Q5 tenía una copia extra del cromosoma 21 en comparación con las células estándar, que era T21; Q3, Q4 carecían de una copia del cromosoma X, que era 45XO (síndrome de Turner). Los resultados eran completamente coherentes con los resultados del análisis tradicional del cariotipo cromosómico.

Aunque los ejemplos de la invención se han descrito con gran detalle, un experto en la técnica comprenderá que, de acuerdo con todas las enseñanzas divulgadas pueden realizarse diversas modificaciones y sustituciones de esos detalles. Todo el alcance de la invención está definido por las reivindicaciones adjuntas y sus equivalentes.

REIVINDICACIONES

1. Un procedimiento de análisis de la información cromosómica de las células muestra, que incluye las etapas de:

romper al azar un genoma de ADN de las células muestra para obtener fragmentos de ADN de un tamaño determinado;

5 secuenciar los fragmentos de ADN para obtener secuencias de ADN;
 alinear estrictamente las secuencias de ADN resultantes contra una secuencia de referencia del genoma humano para obtener información sobre la ubicación de las secuencias de ADN en los cromosomas específicos;
 determinar un número total de secuencias de ADN resultantes que se encuentran en regiones únicas en un cromosoma N particular;

10 calcular un valor del % de CrN para el cromosoma N, en el que % de CrN = el número total de secuencias de ADN resultantes situadas en las regiones únicas en el cromosoma N / un número total de secuencias de ADN resultantes situadas en regiones únicas en todos los cromosomas; y

determinar si existe una diferencia entre el número de cromosomas en las células muestra y el número de cromosomas en las células estándar en base al valor calculado de % de CN y un valor de referencia de % de CrN para las células estándar,

15 en el que la determinación se realiza mediante la comparación de una "puntuación z_CrN" contra un valor de referencia Z,

en el que la puntuación z CrN = (el % de CrN calculado de las células muestra - un valor medio del % de CrN) / una desviación estándar del % de CrN, el valor medio del % de CrN determinado en base al % de CrN para al menos 10 muestras de células estándar y la desviación estándar del % de CrN determinada en base a un valor medio de los valores medios del % de CrN para al menos 10 muestras de celulares estándar,

20 en el que el valor de referencia Z = (media de % de CrN × 0,5 × % X) / desviación estándar de % de CrN, siendo X un número entero entre -100 y 100 ambos incluidos, y

25 En el que un valor absoluto de la puntuación z_CrN es mayor que o igual a 3 y alcanza un valor absoluto del valor de referencia Z, se determina que hay una diferencia del % de X entre el número de cromosomas en las células muestra y el de las células estándar.

2. El procedimiento de la reivindicación 1, en el que la alineación estricta se realiza en base a una alineación intolerante a fallos de las regiones únicas localizadas en la secuencia de referencia del genoma humano, y en el que la secuencia de referencia del genoma humano se obtiene mediante la protección de secuencias repetidas dentro de la secuencia del genoma humano.

3. El procedimiento de la reivindicación 1, en el que la determinación de si existe una diferencia entre el número de cromosomas en las células muestra y el de las células estándar se lleva a cabo mediante análisis con gráfico de cajas.

35 4. El procedimiento de la reivindicación 3, en el que se determina que existe una diferencia si el % de CrN calculado corresponde a un valor atípico que supera un rango intercuartil en más de 1,5 veces.

5. El procedimiento de la reivindicación 3, en el que se determina que existe una diferencia si el % de CrN calculad corresponde a un valor atípico que supera un rango intercuartil en más de 3 veces.

40 6. El procedimiento de la reivindicación 1, en el que si un valor absoluto de la puntuación z_CrN z es mayor o igual a 3, se determina que existe una diferencia entre el número de cromosomas en las células muestra y el de las células estándar.

7. El procedimiento de cualquier reivindicación precedente, en el que las células muestra son células amnióticas, tales como células amnióticas cultivadas o no cultivadas.

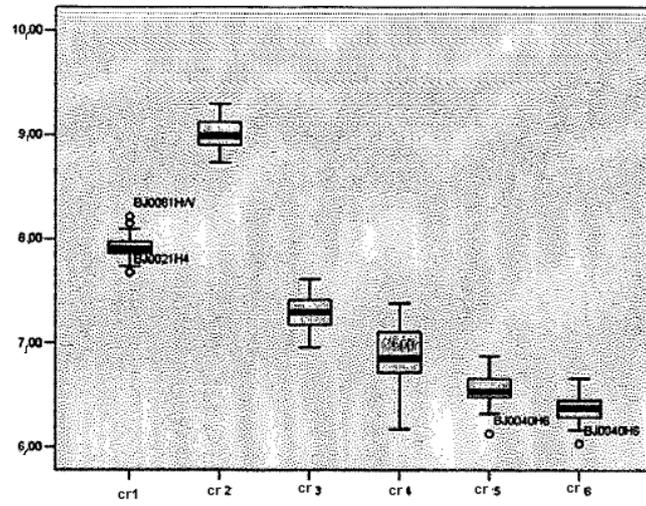


Fig. 1A

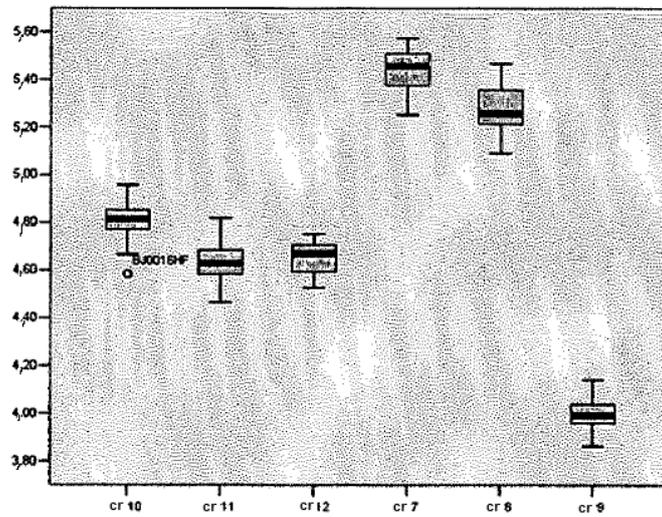


Fig. 1B

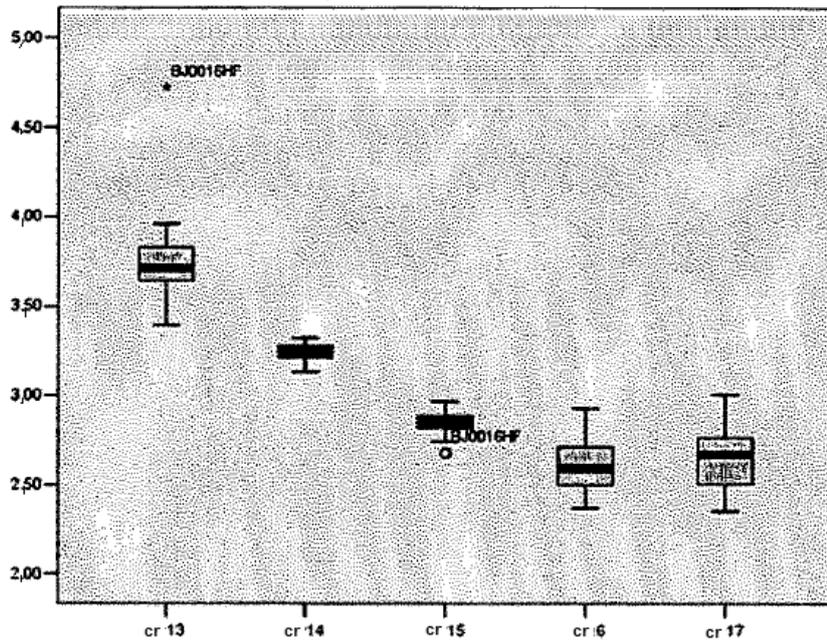


Fig. 1C

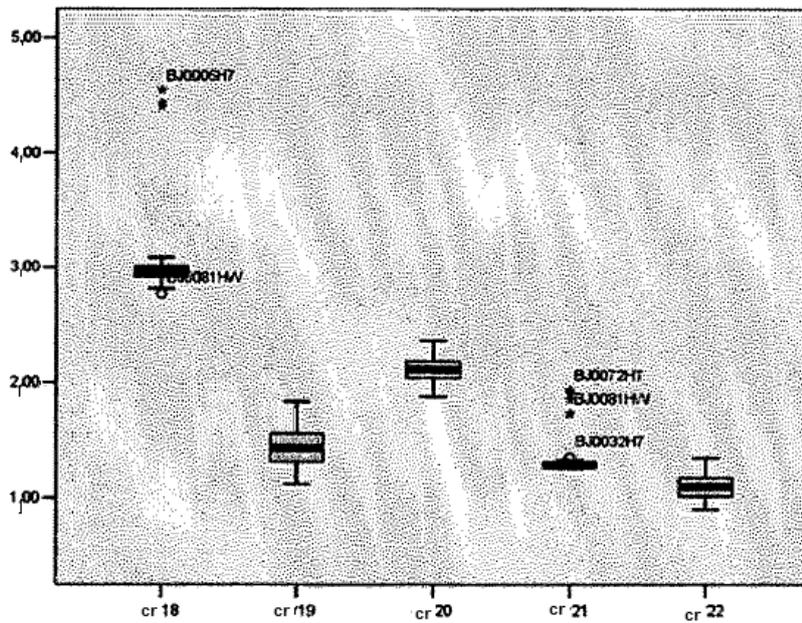


Fig. 1D