

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 699 743**

51 Int. Cl.:

C12Q 1/68 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **17.12.2011 PCT/CN2011/084165**

87 Fecha y número de publicación internacional: **20.06.2013 WO13086744**

96 Fecha de presentación y número de la solicitud europea: **17.12.2011 E 11877603 (8)**

97 Fecha y número de publicación de la concesión europea: **05.09.2018 EP 2792751**

54 Título: **Procedimiento y sistema para determinar si un genoma es anormal**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
12.02.2019

73 Titular/es:
**BGI GENOMICS CO., LTD. (100.0%)
Floors 7-14, Building No. 7, BGI Park, No. 21
Hongan 3rd Street, Yantian District
Shenzhen, Guangdong Province 518083, CN**

72 Inventor/es:
**QIU, YONG;
LIU, LIFU;
JIANG, HUI;
CHEN, FANG;
ZHANG, CHUNLEI;
WANG, JIAN;
WANG, JUN;
YANG, HUANMING y
ZHANG, XIUQING**

74 Agente/Representante:
CARPINTERO LÓPEZ, Mario

ES 2 699 743 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

Procedimiento y sistema para determinar si un genoma es anormal

Campo técnico

5 La presente invención se refiere al campo biomédico. Específicamente, se refiere a un procedimiento y sistema para determinar si existe una anomalía genómica, y más específicamente, la presente invención se refiere a un procedimiento de determinación de la secuencia genómica de glóbulos rojos nucleados fetales, un procedimiento para determinar si existe una anomalía genómica, y un sistema para determinar si existe una anomalía genómica.

Antecedentes de la técnica

10 El diagnóstico prenatal, también conocido como diagnóstico previo al nacimiento, se refiere a efectuar un diagnóstico de alta precisión sobre si un feto antes del nacimiento sufrirá ciertas enfermedades genéticas o malformaciones congénitas combinando los resultados de la detección genética y el examen por imágenes. Los procedimientos utilizados actualmente para el diagnóstico prenatal se clasifican principalmente en diagnóstico invasivo y diagnóstico no invasivo según la diferencia en los procedimientos de muestreo. Los diagnósticos invasivos incluyen principalmente amniocentesis (prueba del líquido amniótico), centesis coriónica, muestreo de sangre del cordón umbilical, fetoscopia, biopsia de embrión, etc. Actualmente, la amniocentesis y la centesis coriónica se aplican con relativa frecuencia. Para los diagnósticos invasivos, dado que las células o tejidos fetales se pueden muestrear directamente, se puede obtener un resultado preciso y confiable después de la detección genética. Sin embargo, debido al procedimiento de muestreo invasivo, se puede causar un daño potencial a la mujer embarazada y al feto, y en un caso grave, incluso puede producirse un aborto fetal o una infección intrauterina. Según las estadísticas, la amniocentesis y la centesis coriónica pueden llevar a una tasa de aborto de aproximadamente 1 % y de 1 %-2 %, respectivamente.

15 Eric Z. Chen y col., desvelan una secuenciación masivamente paralela de moléculas de ADN en el plasma de mujeres embarazadas que ha demostrado permitir una detección prenatal no invasiva y precisa de trisomía fetal 21. Sin embargo, si el enfoque de secuenciación es tan preciso para el diagnóstico prenatal no invasivo de trisomía 13 y 18 no está claro debido a la falta de datos de un gran conjunto de muestras (Eric Z. Chen y col.: "*Noninvasive Prenatal Diagnosis of Fetal Trisomy 18 and Trisomy 13 by Maternal Plasma DNA Sequencing*", PloS ONE, vol. 6, n.º 7, 6 julio de 2011, página e21791").

Actualmente, existe la necesidad de mejorar los procedimientos de diagnóstico prenatal.

Sumario de la invención

30 La presente invención pretende resolver al menos uno de los problemas técnicos existentes en la técnica anterior. Por lo tanto, la presente invención proporciona un procedimiento y un sistema capaces de determinar eficazmente si existe una anomalía genómica.

35 Según un aspecto de la presente invención, la presente invención proporciona un procedimiento para determinar si existe una anomalía genómica. Según las realizaciones de la presente invención, el procedimiento para determinar si existe una anomalía genómica comprende las etapas definidas en la reivindicación 1. Los inventores descubrieron que se puede determinar eficazmente si existe una anomalía genómica en los glóbulos rojos nucleados fetales separados de una muestra de una mujer embarazada utilizando un procedimiento según las realizaciones de la presente invención. El procedimiento puede ser para un fin no médico.

40 Según un segundo aspecto de la presente invención, la presente invención proporciona un sistema para determinar si existe una anomalía genómica. Según realizaciones de la presente invención, el sistema se define en la reivindicación 1. El procedimiento para determinar si existe una anomalía genómica según las realizaciones de la presente invención puede implementarse eficazmente utilizando el sistema para determinar la aneuploidía cromosómica en los glóbulos rojos nucleados, y por lo tanto, puede determinarse eficazmente si existe una anomalía genómica en los glóbulos rojos nucleados.

45 Según otro aspecto de la presente invención, la presente invención proporciona un procedimiento de determinación de la secuencia genómica de glóbulos rojos nucleados fetales. Según las realizaciones de la presente invención, el procedimiento comprende las etapas que consisten en: separar glóbulos rojos nucleados fetales de una muestra de una mujer embarazada; y secuenciar al menos una parte del genoma de dichos glóbulos rojos nucleados, para obtener un resultado de secuenciación. La información de la secuencia genómica de los glóbulos rojos nucleados se puede determinar de manera efectiva mediante este procedimiento, y la información de la secuencia del genoma fetal se puede determinar de ese modo.

50 Los aspectos y ventajas adicionales de la presente invención se darán en parte en la siguiente descripción, y en parte se harán evidentes a partir de la siguiente descripción, o se entenderán por la práctica de la presente invención.

Descripción de los dibujos

Los aspectos y ventajas mencionados anteriormente y/o adicionales de la presente invención se harán evidentes y fáciles de entender a partir de la descripción de las realizaciones junto con los siguientes dibujos, en los que:

La Figura 1 muestra un diagrama de flujo esquemático de un procedimiento para determinar si existe una anomalía genómica en células nucleadas según una realización de la presente invención.

La Figura 2 muestra un diagrama de flujo esquemático de un procedimiento para determinar si existe una anomalía genómica en células nucleadas según otra realización de la presente invención.

La Figura 3 muestra un diagrama esquemático de un sistema utilizado para determinar si existe una anomalía genómica en células nucleadas según una realización de la presente invención.

La Figura 4 muestra un diagrama esquemático de un dispositivo de separación de glóbulos rojos nucleados según una realización de la presente invención.

La Figura 5 muestra un diagrama esquemático de un sistema utilizado para determinar si existe una anomalía genómica en células nucleadas según otra realización de la presente invención.

La Figura 6 muestra un diagrama esquemático para un dispositivo de preparación de bibliotecas de secuenciación de genoma completo según una realización de la presente invención.

La Figura 7 muestra un resultado de detección de una biblioteca de ADN construida analizada por Agilent®Bioanalyzer 2100 según una realización de la presente invención. Brevemente, un producto de amplificación del genoma completo de células positivas aisladas (glóbulos rojos nucleados) se cortó mediante una onda ultrasónica, los fragmentos de ADN en la banda principal cortada fueron de aproximadamente 350 pb, las longitudes de los fragmentos después del ligamiento a un adaptador se incrementaron en aproximadamente 120 pb, y los fragmentos de 430-450 pb se recuperaron por escisión del gel. Se puede ver en la Figura 7 que el intervalo de fragmentos de las cuatro bibliotecas cumple con los requisitos, y la calidad de la biblioteca cumple con los requisitos de secuenciación, en la que GP9 es una muestra de ensayo e YH6 es una muestra de control (una muestra humana normal).

La Figura 8 muestra un resultado de análisis de los datos de secuenciación según una realización de la presente invención, en el que, (A) muestra la distribución del valor de GC de cada ventana y el número de datos de secuenciación alineados de forma única para una muestra a analizar; (B) muestra la curva ajustada por spline lisa de la relación entre el contenido de GC y el número de datos de secuenciación alineados de forma única; (C) muestra la distribución del coeficiente de ponderación correspondiente a la corrección de los datos de cada ventana en la muestra a analizar, en la que la ventana de cada contenido de GC corresponde a un valor de UR como un peso de corrección; y (D) muestra un diagrama de cajas que muestra los datos de secuenciación de cada cromosoma.

Realizaciones particulares

Las realizaciones de la presente invención se detallan a continuación, y los ejemplos de dichas realizaciones se muestran en los dibujos, en los que las etiquetas idénticas o similares representan los elementos idénticos o similares o elementos con las mismas funciones o funciones similares de la primera a la última. Las siguientes realizaciones descritas con referencia a los dibujos son ejemplares, solo se utilizan para explicar la presente invención y no pueden entenderse como una limitación de la presente invención.

Debe tenerse en cuenta que los términos "primero" y "segundo" se utilizan solo con el fin de describir, y no se pueden entender por indicar o implicar la importancia relativa o especificar implícitamente el número de características técnicas indicadas. Por lo tanto, las características definidas por "primero" y "segundo" pueden incluir explícita o implícitamente una o más de las características. Además, en la descripción de la presente invención, a menos que se indique lo contrario, el significado de "una pluralidad de" es dos o más.

1. Procedimiento para determinar si existe una anomalía genómica

Un aspecto de la presente invención se refiere a un procedimiento para determinar si existe una anomalía en un genoma. Con referencia a la Figura 1, según las realizaciones de la presente invención, el procedimiento comprende las etapas que consisten en: **E10: separar glóbulos rojos nucleados fetales de una muestra de una mujer embarazada.**

En la presente invención, la selección de la separación de glóbulos rojos nucleados fetales de una muestra de una mujer embarazada se realiza en base al siguiente descubrimiento de los inventores. Actualmente, los estudios sobre anomalías genéticas fetales se basan principalmente en la separación del ADN fetal libre de una mujer embarazada. Sin embargo, además del ADN libre de origen fetal, también existe una gran cantidad de ADN de origen materno en la sangre periférica de la mujer embarazada, y la mitad del ADN genómico fetal proviene de la madre, lo que hace que sea relativamente difícil determinar con precisión el origen del ADN en el presente. Además, el ADN fetal libre en la sangre periférica materna existe en un genoma incompleto, lo que puede aumentar en gran medida la probabilidad de falso negativo debido a la pérdida de una plantilla en el procedimiento de detección de un sitio génico específico. Por lo tanto, la detección de anomalías genéticas fetales utilizando ADN fetal libre tiene sus propios defectos. Los inventores descubrieron asimismo que, además de los ácidos nucleicos fetales libres, también existen células fetales intactas en la sangre periférica materna. Las células fetales libres en la sangre periférica de la mujer embarazada incluyen principalmente: células trofoblásticas, glóbulos blancos y glóbulos rojos nucleados fetales. Entre estas células, los inventores descubrieron que las células trofoblásticas tienden a conducir a un

diagnóstico erróneo debido a la existencia de dos formas de tales células, es decir, una forma multinucleada y una forma mononucleada. Los glóbulos blancos existirán en la sangre materna persistentemente después del nacimiento del feto, por lo que pueden interferir en la detección en el próximo embarazo. Los inventores descubrieron que los glóbulos rojos nucleados fetales tienen un ciclo de vida relativamente corto, desaparecerán a los 90 días posteriores al nacimiento del feto y no interferirán en la detección del próximo embarazo, y que los antígenos en la superficie de los glóbulos rojos nucleados son relativamente estables, y pueden ser fácilmente reconocidos y separados. Por lo tanto, una anomalía genómica fetal se puede determinar de manera efectiva utilizando los glóbulos rojos nucleados fetales. Los inventores descubrieron que el uso de glóbulos rojos nucleados fetales en la sangre periférica de una mujer embarazada para realizar un diagnóstico prenatal no invasivo junto con una secuenciación de alto rendimiento ha logrado resultados mucho más superiores que el diagnóstico prenatal no invasivo actualmente disponible utilizando el plasma de la mujer embarazada.

Según realizaciones de la presente invención, la muestra de una mujer embarazada como fuente de glóbulos rojos nucleados no está particularmente limitada. Según algunas realizaciones de la presente invención, dicha muestra de una mujer embarazada es preferentemente sangre periférica de la mujer embarazada. Por lo tanto, estas muestras se pueden obtener convenientemente de la mujer embarazada, y en la premisa de modo que el desarrollo fetal no se vea afectado. Los glóbulos rojos nucleados fetales se pueden obtener de la sangre periférica de la mujer embarazada para realizar la secuenciación del genoma completo, realizando un examen prenatal no invasivo. Además, el estadio del embarazo de una mujer embarazada como fuente de glóbulos rojos nucleados fetales no está particularmente limitado. Según realizaciones de la presente invención, una muestra de una mujer embarazada con una edad gestacional inferior a 20 semanas puede utilizarse para el aislamiento de glóbulos rojos nucleados fetales. Por ejemplo, una muestra de una mujer embarazada con una edad gestacional de 12-20 semanas puede ser adoptada como un objeto de investigación. Por lo tanto, los glóbulos rojos nucleados fetales se pueden aislar de manera más efectiva para un análisis adicional. Además, los inventores descubrieron sorprendentemente que con el uso del procedimiento de la presente invención, incluso si solo se obtiene 1 glóbulo rojo nucleado fetal, el análisis aún se puede realizar de manera efectiva.

Según realizaciones de la presente invención, un procedimiento de separación de glóbulos rojos nucleados de una muestra biológica, p. ej., sangre periférica, no está particularmente limitado. Según las realizaciones particulares de la presente invención, separar dichos glóbulos rojos nucleados de sangre periférica comprende además las etapas que consisten en:

En primer lugar, la centrifugación por gradiente se realiza en dicha sangre periférica utilizando un reactivo de gradiente de densidad, para obtener monocitos. Según realizaciones de la presente invención, el tipo de reactivo de gradiente de densidad no está particularmente limitado, y según los ejemplos particulares, polisucrosa, p. ej., Ficoll, se puede utilizar para formar un gradiente de densidad. Preferentemente, dicha centrifugación por gradiente se puede realizar a 800 X g durante 30 minutos.

Después de que se obtengan los monocitos, los glóbulos rojos nucleados se enriquecen a partir de los monocitos obtenidos utilizando perlas magnéticas que transportan un anticuerpo, en las que el anticuerpo transportado en las perlas magnéticas reconoce específicamente un antígeno en la superficie de los glóbulos rojos nucleados. Los glóbulos rojos nucleados se unirán a las perlas magnéticas por el anticuerpo, y posteriormente, los glóbulos rojos nucleados se pueden obtener por un tamizaje magnético.

Según una realización de la presente invención, después de que se obtengan dichos monocitos y antes de que dichos glóbulos rojos nucleados se enriquezcan utilizando las perlas magnéticas que transportan el anticuerpo anti-CD71, el procedimiento comprende además lavar dichos monocitos utilizando solución salina tamponada con fosfato (tampón PBS) que contiene albúmina de suero bovino al 1 % (BSA), para eliminar el reactivo de gradiente de densidad residual. Preferentemente, dicho tampón PBS contiene dihidrogenofosfato de potasio e hidrogenofosfato de sodio, pero está libre de iones de calcio e iones de magnesio, para mejorar así significativamente la eficacia del enriquecimiento de los glóbulos rojos nucleados. Según los ejemplos particulares de la presente invención, lavar dichos monocitos utilizando tampón PBS que contiene BSA al 1 % comprende además: mezclar dichos monocitos y dicho tampón PBS que contiene BSA al 1 %, para obtener una suspensión que contiene monocitos; y centrifugar dicha suspensión que contiene monocitos, preferentemente, centrifugar la suspensión a 200 X g durante 5 minutos, y descartar el sobrenadante, para obtener glóbulos rojos nucleados lavados. Preferentemente, dicho anticuerpo es un anticuerpo que reconoce específicamente CD71.

Los inventores descubrieron que al usar un procedimiento de separación de glóbulos rojos nucleados según realizaciones de la presente invención, los glóbulos rojos nucleados, particularmente los glóbulos rojos nucleados fetales; pueden ser separados eficazmente de la sangre periférica. Por lo tanto, la presente invención proporciona un procedimiento de separación de glóbulos rojos nucleados fetales de sangre periférica que es simple y fácil de llevar a cabo. Los expertos en la materia entenderán fácilmente que, en el procedimiento de separación de glóbulos rojos nucleados, también se pueden incluir otras etapas. Por ejemplo, según ejemplos particulares de la presente invención, el procedimiento de separación de glóbulos rojos nucleados comprende: tomar una cantidad apropiada de sangre periférica de una mujer embarazada, realizar una anti-coagulación con un agente anticoagulante, diluir la muestra de sangre proporcionalmente con PBS 0,1 M libre de iones calcio e iones magnesio, colocar la muestra diluida lentamente en un reactivo para la centrifugación por gradiente de densidad y realizar una centrifugación por gradiente de densidad a temperatura ambiente. Después de la centrifugación, se puede observar una capa de

monocitos, esta capa de células se pipetea cuidadosamente, se transfiere a un nuevo tubo de centrifuga, se vuelve a suspender con 3 volúmenes de tampón PBS que contiene BSA al 1 % y se centrifuga nuevamente a temperatura ambiente, el sobrenadante se desecha, el precipitado celular obtenido se lava adicionalmente con el mismo procedimiento dos veces para eliminar el líquido del gradiente de densidad residual y, finalmente, el precipitado celular se vuelve a suspender en PBS que contiene BSA al 0,1 % y se pipeta de manera uniforme. Luego, las perlas magnéticas que transportan un anticuerpo se añaden a una proporción de 20 microlitros/10⁶ células, y se centrifugan después de dejar reposar a 4 °C, el sobrenadante se desecha y el precipitado se vuelve a suspender en PBS que contiene BSA al 0,1 %; y se monta un sistema de clasificación de perlas magnéticas. Una columna de clasificación se humedece con 500 microlitros de tampón PBS que contiene BSA al 0,1 %, y después de vaciar el líquido, las células que se van a clasificar se cargan en la columna, y el líquido del efluente se recoge y se marca como células negativas. El tubo se humedece con PBS que contiene BSA al 0,1 % y, después de vaciar el líquido, se repite lo mismo dos veces. Se añade PBS/EDTA/BSA a la columna de clasificación, y después de que se vacíe el líquido, se repite lo mismo una vez más. Finalmente, se añade PBS que contiene BSA al 0,1 % a la columna de clasificación, se retira el campo magnético, el líquido se lava en un nuevo tubo de centrifuga y se obtienen glóbulos rojos nucleados.

E20: después de separar los glóbulos rojos nucleados fetales de una muestra de una mujer embarazada, se puede secuenciar al menos una parte del genoma de los glóbulos rojos nucleados y, por lo tanto, se puede obtener un resultado de secuenciación correspondiente a los objetos de secuenciación.

Según las realizaciones de la presente invención, después de separar los glóbulos rojos nucleados fetales, se puede secuenciar al menos una parte del genoma de los glóbulos rojos nucleados. Los expertos en la materia pueden seleccionar los objetos de secuenciación del genoma de los glóbulos rojos nucleados según los genes de interés, y de este modo obtener un resultado de secuenciación correspondiente a estos objetos de secuenciación. Según realizaciones de la presente invención, los expertos en la técnica pueden adoptar cualquier procedimiento conocido para seleccionar los objetos de secuenciación, p. ej., pueden seleccionar solo varios cromosomas en los mismos. Los expertos en la materia entenderán fácilmente que el genoma completo de los glóbulos rojos nucleados también se puede secuenciar directamente y que, una vez obtenido el resultado de la secuenciación, se seleccionan los datos de secuenciación de un sitio específico del resultado de secuenciación para su posterior análisis (véanse los detalles más adelante). Por conveniencia, el siguiente ejemplo ilustra la secuenciación del genoma completo de los glóbulos rojos nucleados.

Según las realizaciones de la presente invención, una vez obtenidos los glóbulos rojos nucleados, el procedimiento de secuenciación del genoma completo de los glóbulos rojos nucleados no está particularmente limitado. Según una realización de la presente invención, la secuenciación del genoma completo de los glóbulos rojos nucleados comprende además: en primer lugar, amplificar el genoma completo de los glóbulos rojos nucleados para obtener un genoma completo amplificado; posteriormente, construir una biblioteca de secuenciación del genoma completo utilizando el genoma completo amplificado; y finalmente, secuenciar la biblioteca de secuenciación del genoma completo, para obtener un resultado de secuenciación que contiene una pluralidad de datos de secuenciación. Por lo tanto, la información del genoma completo de los glóbulos rojos nucleados se puede adquirir de manera efectiva, lo que mejora aún más la eficacia de determinar si existe una anomalía genómica en los glóbulos rojos nucleados.

Los expertos en la materia pueden seleccionar diferentes procedimientos de construcción de una biblioteca de secuenciación de genoma completo según el protocolo específico para la técnica de secuenciación genómica aplicada. Para obtener detalles de la construcción de la biblioteca de secuenciación del genoma completo, véanse las instrucciones de la directiva proporcionadas por el fabricante del secuenciador, p. ej., the Illumina Corporation. Por ejemplo, véase Multiplexing Sample Preparation Guide (parte n.º 1005361; febrero de 2010) o Paired- End SamplePrep Guide (parte n.º 1005063; febrero de 2010) de Illumina Corporation.

Opcionalmente, un procedimiento según una realización de la presente invención puede comprender además una etapa que consiste en lisar dichos glóbulos rojos nucleados, para liberar el genoma completo de dichos glóbulos rojos nucleados. Según algunos ejemplos de la presente invención, el procedimiento que puede utilizarse para lisar glóbulos rojos nucleados y liberar el genoma completo no está particularmente limitado, siempre que el procedimiento pueda lisar, preferentemente lisar por completo, los glóbulos rojos nucleados. Según los ejemplos particulares de la presente invención, el tampón de lisis alcalina se puede utilizar para lisar dichos glóbulos rojos nucleados y liberar el genoma completo de dichos glóbulos rojos nucleados. Los inventores descubrieron que en este procedimiento, los glóbulos rojos nucleados pueden lisarse de manera efectiva y que se puede liberar el genoma completo, y que cuando se secuencian el genoma completo liberado, la tasa de precisión se puede mejorar, lo que mejora aún más la eficacia de determinación de aneuploidía cromosómica en los glóbulos rojos nucleados.

Según las realizaciones de la presente invención, el procedimiento de amplificación del genoma completo de glóbulos rojos nucleados no está particularmente limitado. Un procedimiento basado en PCR, p. ej., PEP-PCR, DOP-PCR y OmniPlex WGA, se puede utilizar, y un procedimiento no basado en PCR, p. ej., MDA (amplificación de desplazamiento múltiple), también se puede utilizar. Según los ejemplos particulares de la presente invención, preferentemente, se utiliza un procedimiento basado en PCR, p. ej., el procedimiento OmniPlex WGA. Los kits comercializados seleccionables incluyen, entre otros, GenomePlex de Sigma Aldrich, PicoPlex de Rubicon Genomics, REPLI-g de Qiagen, ilustra GenomiPhi de GE Healthcare, etc. Por lo tanto, según los ejemplos

particulares de la presente invención, antes de que se construya la biblioteca de secuenciación, OmniPlex WGA puede utilizarse para amplificar el genoma completo de los glóbulos rojos nucleados. El genoma completo se puede amplificar eficazmente, lo que mejora aún más la eficacia para determinar la aneuploidía cromosómica en los glóbulos rojos nucleados.

5 Según las realizaciones de la presente invención, el procedimiento de construcción de una biblioteca de secuenciación de genoma completo (también referida a veces como "biblioteca de ácidos nucleicos" o "biblioteca de secuenciación" en la presente memoria) utilizando dicho genoma completo amplificado comprende además:
 En primer lugar, fragmentar el genoma completo amplificado, para obtener fragmentos de ADN. Según realizaciones de la presente invención, el procedimiento de fragmentación del ADN obtenido no está particularmente limitado.
 10 Según algunos ejemplos particulares, la fragmentación se puede realizar por al menos uno seleccionado entre el grupo que consiste en atomización, procedimiento de cizallamiento ultrasónico, HydroShear y tratamiento de digestión enzimática. Preferentemente, el genoma completo amplificado se fragmenta utilizando un dispositivo de cizallamiento ultrasónico covaris. Según realizaciones de la presente invención, los fragmentos de ADN obtenidos después del tratamiento de fragmentación tienen una longitud de 200-400 pb, preferentemente 350 pb. Los
 15 inventores descubrieron que los fragmentos de ADN obtenidos de esta longitud se pueden utilizar de manera efectiva para la construcción de la biblioteca de ácidos nucleicos y su posterior manipulación.

Una vez obtenidos los fragmentos de ADN, se puede realizar una reparación de extremo en los fragmentos de ADN obtenidos, a fin de obtener fragmentos de ADN con extremos reparados. Según las realizaciones de la presente invención, la reparación de extremo se puede realizar en los fragmentos de ADN utilizando el fragmento Klenow, la
 20 ADN polimerasa T4 y la polinucleótido quinasa T4, en el que el fragmento Klenow tiene actividad polimerasa 5' -- +3' y actividad exonucleasa 3' -- + 5', pero carece de la actividad exonucleasa 5' -- +3', que puede así reparar eficazmente el extremo de los fragmentos de ADN.

Después de realizar la reparación de extremo en los fragmentos de ADN, se pueden añadir las bases A a los extremos 3' de los fragmentos de ADN con extremos reparados, para obtener fragmentos de ADN con el extremo A pegajoso. Según algunos ejemplos particulares de la presente invención, las bases A se pueden añadir a los
 25 extremos 3' de los fragmentos de ADN con extremos reparados utilizando el fragmento Klenow (3'-5' exo-), es decir, el fragmento Klenow carece de la actividad exonucleasa 3'→5', para obtener así eficazmente los fragmentos de ADN con el extremo pegajoso A.

Después de añadir las bases A al extremo, los fragmentos de ADN con el extremo pegajoso A se pueden ligar a un adaptador, para obtener un producto de ligamiento. Según realizaciones de la presente invención, los fragmentos de
 30 ADN con el extremo pegajoso A se pueden ligar al adaptador utilizando ADN ligasa T4, para obtener así eficazmente el producto de ligamiento. Según las realizaciones de la presente invención, se puede incluir una etiqueta adicional en el adaptador, y así las bibliotecas de secuenciación del genoma completo de una pluralidad de muestras de glóbulos rojos nucleados se puede construir simultáneamente de una manera conveniente, y las bibliotecas de
 35 secuenciación de la pluralidad de muestras se combinan y secuencian simultáneamente. Por lo tanto, una plataforma de secuenciación de alto rendimiento puede utilizarse completamente para ahorrar tiempo y reducir el costo de la secuenciación.

Una vez obtenido el producto de ligamiento, se realiza la amplificación por PCR en dicho producto de ligamiento, para obtener un segundo producto de amplificación; y dicho segundo producto de amplificación se purifica y recupera, para obtener un producto recuperado, y dicho producto recuperado forma dicha biblioteca de
 40 secuenciación del genoma completo.

Después de construir la biblioteca de secuenciación del genoma completo, según las realizaciones de la presente invención, dicha biblioteca de secuenciación del genoma completo puede secuenciarse. Los expertos en la materia entenderán fácilmente que la etapa de secuenciación en la presente invención se puede realizar por cualquier
 45 procedimiento de secuenciación, que incluye, entre otros, el procedimiento de terminación de cadena dideoxi; preferentemente procedimientos de secuenciación de alto rendimiento. Por lo tanto, las características de secuenciación de alto rendimiento y profunda de estos dispositivos de secuencia pueden utilizarse para mejorar aún más la eficacia de la determinación de la aneuploidía cromosómica en los glóbulos rojos nucleados. Dichos procedimientos de secuenciación de alto rendimiento incluyen, entre otros, la técnica de secuenciación de segunda
 50 generación o la técnica de secuenciación de una sola molécula.

Las plataformas de secuenciación de segunda generación (Metzker ML. *Sequencing technologies-the next generation. Nat Rev Genet.* Enero de 2010; 11(1):31-46) incluyen, entre otros, plataformas de secuenciación Illumina-Solexa (GA™, HiSeq2000™, etc.), ABI-Solid y Roche-454 (pirosecuenciación). Las plataformas (técnicas) para la secuenciación de una sola molécula incluyen, entre otros, la verdadera técnica de secuenciación de una sola
 55 molécula (True Single Molecule DNA sequencing) de Helicos Corporation, la técnica de secuenciación de una sola molécula en tiempo real (SMRT™) de Pacific Biosciences Corporation, y la técnica de secuenciación de nanoporos de Oxford Nanopore Technologies Corporation, etc. (Rusk, Nicole (01-04-2009). *Cheap Third-Generation Sequencing. Nature Methods* 6 (4): 244-245).

Con el desarrollo continuo de la tecnología de secuenciación, los expertos en la técnica entenderían que la

secuenciación del genoma completo también se puede realizar utilizando otros procedimientos y dispositivos de secuenciación. Según las realizaciones de la presente invención, las longitudes de los datos de secuenciación obtenidos por la secuenciación del genoma completo no están particularmente limitadas. Según un ejemplo particular de la presente invención, la longitud promedio de dicha pluralidad de datos de secuenciación es de aproximadamente 50 pb. Los inventores descubrieron que cuando la longitud promedio de los datos de secuenciación es de aproximadamente 50 pb, el análisis de los datos de secuenciación se puede facilitar enormemente, la eficacia del análisis se mejora y, al mismo tiempo, el costo del análisis se puede reducir significativamente. Por lo tanto, la eficacia de la determinación de la aneuploidía cromosómica en los glóbulos rojos nucleados se mejora aún más, y el costo de la determinación de la aneuploidía cromosómica en los glóbulos rojos nucleados se reduce. La expresión "longitud promedio", como se utiliza en la presente memoria, se refiere al valor promedio de los valores numéricos de las longitudes de todos los datos de secuenciación.

E30: una vez obtenido el resultado de la secuenciación, se determina si existe una anomalía genómica en los glóbulos rojos nucleados en función del resultado de secuenciación obtenido.

La expresión "anomalía genómica", como se utiliza en la presente memoria, debe entenderse en un sentido amplio, que puede referirse a cualquier cambio en la secuencia genómica, p. ej., aneuploidía cromosómica, variación estructural, mutación de nucleótido único y otras variaciones genéticas (www.en.wikipedia.org/wiki/Genetic_variation), que también puede ser un cambio en un sitio de modificación genómica, p. ej., el nivel de metilación, etc. Según las realizaciones de la presente invención, la anomalía genómica estudiada es al menos una seleccionada entre el grupo que consiste en aneuploidía cromosómica y una mutación en una región predeterminada. En las realizaciones de la presente invención, la mutación en una región predeterminada se refiere a la variación estructural (www.en.wikipedia.org/wiki/Structural_variation) o mutación de un solo nucleótido (SNP, www.en.wikipedia.org/wiki/Single-nucleotide_polymorphism).

Según las realizaciones de la presente invención, determinar una mutación en una región predeterminada en el genoma puede comprender además:

En primer lugar, determinar la secuencia de ácido nucleico de la región predeterminada en dichos glóbulos rojos nucleados basándose en el resultado de la secuenciación. Los expertos en la materia pueden utilizar cualquier procedimiento conocido para determinar la secuencia de ácido nucleico de la región predeterminada. Por ejemplo, puede utilizarse un procedimiento conocido para ensamblar datos de secuenciación de una región específica en el resultado de la secuenciación, obteniendo así la secuencia de ácido nucleico de una secuencia predeterminada.

Después de obtener la secuencia de ácido nucleico de la secuencia predeterminada, la secuencia de ácido nucleico de la región predeterminada en dichos glóbulos rojos nucleados se alinea con una secuencia de ácido nucleico de control; preferentemente, dicha secuencia de ácido nucleico de control es una secuencia genómica humana normal. Posteriormente, sobre la base de un resultado de alineamiento, se puede determinar si existe una anomalía en la región predeterminada en los glóbulos rojos nucleados. Según realizaciones de la presente invención, la mutación en una región predeterminada que puede ser detectada por el procedimiento incluye al menos una seleccionada entre el grupo que consiste en mutación por inserción, mutación por delección, mutación por sustitución, mutación por inversión, variación del número de copias, mutación por translocación y polimorfismo de un único nucleótido simple.

Con referencia a la Figura 2, según una realización de la presente invención, el procedimiento de determinación de la aneuploidía cromosómica comprende las etapas que consisten en:

E100: en primer lugar, secuenciar el genoma completo de los glóbulos rojos nucleados para obtener un primer resultado de secuenciación. La descripción sobre la secuenciación del genoma completo se ha detallado anteriormente, por lo tanto, no se repite.

E200: dividir la secuencia conocida de un primer cromosoma en ventanas

Para analizar los datos de secuenciación de los glóbulos rojos nucleados, la secuencia conocida del primer cromosoma se divide primero en ventanas, y cada una de estas ventanas tiene, independientemente, una longitud predeterminada de las secuencias, respectivamente. Según una realización de la presente invención, las longitudes de secuencias dentro de estas ventanas pueden ser iguales, también pueden ser diferentes y no están particularmente limitadas. Según una realización de la presente invención, las longitudes predeterminadas de secuencias dentro de dicha pluralidad de ventanas son las mismas. Preferentemente, las longitudes predeterminadas de secuencias dentro de dicha pluralidad de ventanas son todas de 60 kB. Por lo tanto, se puede mejorar la eficacia de la determinación de la aneuploidía cromosómica en los glóbulos rojos nucleados. Los expertos en la materia pueden seleccionar el intervalo de secuencia del primer cromosoma cubierto por las ventanas según sea necesario.

E300: determinar el número de datos de secuencia que caen en cada ventana.

Una vez obtenidos los datos de secuenciación, los datos de secuenciación obtenidos se alinean con la secuencia conocida del primer cromosoma para dividir así los datos de secuenciación obtenidos en las ventanas con longitudes predeterminadas, respectivamente. Los expertos en la materia entenderían que cualquier procedimiento y medio conocidos pueden utilizarse para realizar el alineamiento de secuencias y para calcular el número total de estos datos de secuenciación. Por ejemplo, el software proporcionado por el fabricante del secuenciador puede ser adoptado para realizar el análisis, p. ej., SOAP v2.20. Según una realización de la presente invención, dichos datos de secuenciación que caen en cada ventana son datos de secuenciación

alineados de forma única. Por lo tanto, a través de la selección de los datos de secuenciación, la eficacia de la determinación de la aneuploidía cromosómica en los glóbulos rojos nucleados puede mejorarse aún más. La expresión "datos de secuenciación alineados de manera única", también referida como "lectura única", como se utiliza en la presente memoria, se refiere a los datos de secuenciación que pueden combinarse perfectamente y alinearse exitosamente solo una vez con un genoma de referencia cuando los datos de secuenciación están

alineados con una secuencia cromosómica conocida, p. ej., el genoma humano Hg19. El término "primer cromosoma", como se utiliza en la presente memoria, debe entenderse en un sentido amplio, que puede referirse a cualquier cromosoma objetivo que se espera estudiar. Su número no se limita a un solo cromosoma, se pueden analizar más e incluso todos los cromosomas simultáneamente. Según las realizaciones de la presente invención, el primer cromosoma puede ser cualquier cromosoma seleccionado entre los cromosomas humanos 1-23. Según realizaciones de la presente invención, preferentemente, el "primer cromosoma" es al menos uno seleccionado entre el grupo que consiste en el cromosoma humano 21, el cromosoma 18, el cromosoma 13, el cromosoma X y el cromosoma Y. Por lo tanto, las enfermedades cromosómicas humanas comunes se pueden determinar eficazmente, por ejemplo, para el pronóstico de enfermedades genéticas fetales. Por lo tanto, el procedimiento de determinación de la aneuploidía cromosómica en glóbulos rojos nucleados según las realizaciones de la presente invención se puede aplicar muy eficazmente en el cribado preimplantacional (PGS) y el diagnóstico preimplantacional (PGD) en el campo de la reproducción in vitro, prueba prenatal en células nucleadas fetales, etc. Por lo tanto, se puede predecir rápidamente si existe una anomalía cromosómica en un feto mediante el simple aislamiento de glóbulos rojos nucleados, lo que puede evitar el caso de que el feto sufra una enfermedad genética grave. La expresión "se puede alinear con el primer cromosoma", como se utiliza en la presente memoria, se refiere a que a través del alineamiento de los datos de secuenciación con la secuencia conocida del primer cromosoma del genoma de referencia, los datos de la secuencia se pueden alinear con la secuencia conocida del primer cromosoma, de este modo, se determina que estos datos de secuenciación se derivan del primer cromosoma.

E400: determinar un primer parámetro basado en el número de datos de secuencia que cae en cada ventana.

En el resultado de la secuenciación de la secuenciación del genoma completo de los glóbulos rojos nucleados 14, el número de datos de secuenciación para un cromosoma específico tiene una correlación positiva con el contenido de este cromosoma en el genoma completo. Por lo tanto, a través del análisis del número de datos de secuenciación derivados de un cromosoma específico en el resultado de la secuenciación y el número total derivado de la secuenciación del genoma completo, el cromosoma específico se puede analizar de manera efectiva. Para este fin, se puede determinar un primer parámetro mediante el análisis del número de datos de secuenciación que cae en cada ventana del primer cromosoma. Para hacer que el primer parámetro sea capaz de reflejar verdaderamente si existe aneuploidía en el primer cromosoma, según una realización de la presente invención, el primer parámetro basado en el número de datos de secuenciación que cae en cada ventana se determina mediante un procedimiento que comprende además: establecer un coeficiente de ponderación predeterminado (también referido en ocasiones en la presente memoria como "coeficiente ponderado") para el número de datos de secuenciación que cae en cada ventana, respectivamente; y según el coeficiente de ponderación, realizar un promedio ponderado en el número de datos de secuenciación que cae en cada ventana para obtener la mediana de dicho primer cromosoma, la mediana obtenida formando el primer parámetro del primer cromosoma.

Los inventores descubrieron que el primer parámetro obtenido de esta manera puede encarnar eficazmente el número de datos de secuenciación del primer cromosoma en los datos de secuenciación y, por lo tanto, la eficacia de la determinación de la aneuploidía cromosómica se puede mejorar aún más.

Según las realizaciones de la presente invención, el coeficiente de ponderación se establece para poder eliminar la distorsión de datos causada por algunos errores que pueden existir en el procedimiento de secuenciación. Según un ejemplo de la presente invención, el coeficiente de ponderación predeterminado se obtiene asociando el número de datos de secuencia que cae en cada ventana con el contenido de GC de la ventana respectiva. Por lo tanto, los errores resultantes debido al sesgo de la técnica de secuenciación hacia una región con un alto contenido de GC se pueden eliminar de manera efectiva. Por ejemplo, según una realización de la presente invención, el coeficiente de ponderación se puede obtener a través del siguiente procedimiento:

En primer lugar, cada cromosoma del genoma humano completo se divide en ventanas con una longitud fija de 60 kB respectivamente, y se registran la posición inicial y la posición final de cada ventana y el valor promedio de GC (registrado como GC_{ref}) se obtiene mediante estadísticas. Después de obtener los datos de secuenciación de una muestra a analizar, las secuencias en el resultado de la secuenciación se alinean con el genoma. Los datos de secuenciación que se combinan perfectamente y se alinean con éxito solo una vez, es decir; se extraen datos de secuenciación alineados de forma única, y se obtiene la información acerca de los sitios de todos los datos de secuenciación alineados de forma única. Se cuenta el número de datos de secuenciación alineados de forma única (registrados como $UR_{muestra}$) en cada ventana para cada cromosoma en el genoma correspondiente al resultado de alineamiento de la muestra a analizar. Los $UR_{muestra}$ y GC_{ref} resultantes se representan para obtener la Figura 8A. Como se muestra en la Figura 8A, el sesgo de GC introducido por el secuenciador causó una mayor distribución de los datos de secuenciación con un contenido de GC de aproximadamente en la región [0,35, 0,55]. Los puntos distintos en la Figura 8A están ajustados con el procedimiento de spline liso en una curva lisa para obtener la Figura 8B, que es un diagrama que refleja la relación entre el contenido de GC y el número de datos de secuenciación. La Figura 8B muestra todas las ventanas divididas por una longitud de etapa de 1 % del valor promedio de GC. El número de datos de secuenciación alineados de forma única, es decir, $M_{ajustado}$, en una ventana correspondiente a

cada valor promedio de GC se puede obtener a partir de los datos ajustados. Según la fórmula $W_{GC} = M/M_{ajustado}$, se obtiene el coeficiente de ponderación predeterminado W_{GC} , y véase en la Figura 8C para la distribución del coeficiente de ponderación, la que M es el número de datos de secuenciación alineados de forma única de la muestra que se analizará de la caída en ventanas de un valor promedio de GC igual.

5 Posteriormente, después de obtener el primer parámetro, se puede determinar si los glóbulos rojos nucleados tienen aneuploidía para dicho primer cromosoma comparando el primer parámetro con un parámetro de control predeterminado. El término "predeterminado", como se utiliza en la presente memoria, debe entenderse en un sentido amplio, que puede determinarse por adelantado mediante un experimento, y también puede obtenerse realizando un experimento paralelo cuando se analiza una muestra biológica. La expresión "experimento paralelo",
 10 como se utiliza en la presente memoria, debe entenderse en un sentido amplio, que no solo puede referirse a la secuenciación y al análisis simultáneos de una muestra desconocida y una muestra conocida, sino que también puede referirse a la secuenciación y al análisis realizados sucesivamente en la misma condición. Por ejemplo, un primer parámetro para la muestra obtenida al analizar una muestra de glóbulos rojos nucleados que se sabe que tiene aneuploidía o una muestra de glóbulos rojos nucleados que se sabe que no tiene aneuploidía se puede utilizar como parámetro de control.
 15

Además, los inventores descubrieron que si el primer cromosoma tiene aneuploidía se puede determinar comparando y analizando estadísticamente el número de datos de secuenciación de diferentes cromosomas en el mismo análisis de secuenciación. Por lo tanto, según una realización de la presente invención, si los glóbulos rojos nucleados tienen aneuploidía para el primer cromosoma basado en el primer parámetro se puede determinar por un procedimiento que comprende además: realizar el mismo tratamiento en un segundo cromosoma al igual que en el primer cromosoma para obtener la media de dicho segundo cromosoma; realizar la prueba del valor de t en la media del primer cromosoma y la media del segundo cromosoma, para obtener la diferencia entre dicho primer cromosoma y dicho segundo cromosoma; y comparar la diferencia obtenida con un primer umbral predeterminado y un segundo umbral, y si la diferencia obtenida es inferior a los umbrales predeterminados, entonces determinar que los glóbulos rojos nucleados tienen aneuploidía para el primer cromosoma, y si la diferencia obtenida es superior a dichos umbrales predeterminados, entonces determinar que las células nucleadas no tienen aneuploidía para el primer cromosoma. La diferencia entre el primer cromosoma y otros cromosomas también se puede determinar para mejorar la eficacia de la determinación de la aneuploidía cromosómica.
 20
 25

La expresión "segundo cromosoma", como se utiliza en la presente memoria, debe entenderse en un sentido amplio, que puede referirse a cualquier cromosoma objetivo que se espera que se estudie, su número no se limita a un solo cromosoma, más e incluso a todos los cromosomas diferentes al primer cromosoma pueden ser analizados simultáneamente. Según las realizaciones de la presente invención, el segundo cromosoma puede ser cualquier cromosoma entre los cromosomas humanos, y para los glóbulos rojos nucleados fetales, el segundo cromosoma es preferentemente cualquier cromosoma seleccionado entre los cromosomas humanos 1-23. Según realizaciones de la presente invención, preferentemente, el segundo cromosoma es uno cualquiera seleccionado entre el grupo que consiste en los cromosomas 1-12 del genoma humano. Debido a que la aneuploidía cromosómica no existe en estos cromosomas, estos cromosomas se pueden utilizar de manera efectiva como referencia para someter a ensayo el primer cromosoma para mejorar la eficacia del ensayo.
 30
 35

Según una realización de la presente invención, la siguiente fórmula se utiliza para realizar la prueba del valor t en la mediana de un primer cromosoma y la mediana de un segundo cromosoma,
 40

$$T_{i,j} = \frac{\mu_i - \mu_j}{\sqrt{\frac{\sigma_i^2}{n_i} + \frac{\sigma_j^2}{n_j}}} \quad (\text{fórmula I})$$

en la que $T_{i,j}$ representa la diferencia entre el primer cromosoma y el segundo cromosoma, μ_i representa la mediana del primer cromosoma, μ_j representa la mediana del segundo cromosoma, σ_i representa la desviación típica de la distribución del número de datos de secuenciación en cada ventana en el cromosoma, σ_j representa la desviación típica de la distribución del número de datos de secuenciación en cada ventana en el segundo cromosoma, n_i representa el número de ventanas en el primer cromosoma, y n_j representa el número de ventanas en el segundo cromosoma.
 45

Según las realizaciones de la presente invención, se utiliza como un umbral el valor de los umbrales predeterminados se puede obtener por experiencia, o el correspondiente valor de la prueba t obtenido por medio de una prueba de antemano de una muestra de glóbulos rojos nucleados que se sabe que tienen aneuploidía o una muestra de glóbulos rojos que no se saben que tienen aneuploidía. Preferentemente, el primer umbral predeterminado es -4 o menos, y el segundo umbral es -3,5 o superior.
 50

A diferencia de la euploidía cromosómica, el término "aneuploidía", como se utiliza en la presente memoria, se refiere a la falta o adición de uno o varios cromosomas al genoma. En general, hay dos cromatinas para cada

5 cromosoma en una célula normal. Sin embargo, los gametos con un número anormal de cromosomas se forman debido a una no disyunción o a una disyunción demasiado temprana de un par de cromosomas homólogos durante la meiosis, y la unión de tales gametos entre sí o con gametos normales generará varias células aneuploides. Además, las células aneuploides, tales como las células tumorales con una tasa de mutación muy alta, etc., también pueden generarse durante la división celular somática.

2. Sistema para determinar si existe una anomalía genómica en los glóbulos rojos nucleados

10 Según otro aspecto de la presente invención, la presente invención proporciona un sistema 1000 para determinar si existe una anomalía genómica. Con referencia a la Figura 3, según una realización de la presente invención, el sistema 1000 comprende: un dispositivo 100 de separación de glóbulos rojos nucleados, un dispositivo de secuenciación 200 y un dispositivo 300 de análisis de los resultados de secuenciación. Según una realización de la presente invención, el dispositivo 100 de separación de glóbulos rojos nucleados se utiliza para separar glóbulos rojos nucleados fetales de una muestra de una mujer embarazada. El dispositivo de secuenciación 200 se utiliza para secuenciar al menos una parte del genoma de los glóbulos rojos nucleados, para obtener un resultado de secuenciación. El dispositivo 300 de análisis de resultados de secuenciación está conectado al dispositivo de secuenciación 200, para recibir el resultado de la secuenciación del dispositivo de secuenciación 200, y determinar si existe una anomalía genómica en los glóbulos rojos nucleados separados basándose en el resultado de secuenciación obtenido.

20 Según una realización de la presente invención, el dispositivo 100 de separación de glóbulos rojos nucleados puede comprender además una unidad de separación de monocitos 101 y una unidad de enriquecimiento magnético 102. Según una realización de la presente invención, la unidad de separación de monocitos 101 es adecuada para realizar una centrifugación por gradiente en la muestra de una mujer embarazada utilizando un reactivo de gradiente de densidad, para obtener monocitos, en el que la muestra de una mujer embarazada es sangre periférica de la mujer embarazada. La unidad de enriquecimiento magnético 102 está conectada a la unidad de separación de monocitos 101, y es adecuada para separar glóbulos rojos nucleados de monocitos utilizando perlas magnéticas que transportan un anticuerpo, en el que dicho anticuerpo reconoce específicamente un antígeno en la superficie de los glóbulos rojos nucleados. Así, asistidos por el dispositivo 100 de separación de glóbulos rojos nucleados, los glóbulos rojos nucleados se pueden separar de manera eficaz a través del siguiente procedimiento:

25 En primer lugar, la centrifugación por gradiente se realiza en la sangre periférica materna utilizando un reactivo de gradiente de densidad, para obtener monocitos. Después de obtener los monocitos, los glóbulos rojos nucleados se enriquecen a partir de los monocitos obtenidos utilizando perlas magnéticas que transportan un anticuerpo, en las el anticuerpo transportado en las perlas magnéticas reconoce específicamente un antígeno en la superficie de los glóbulos rojos nucleados para así unir los glóbulos rojos nucleados a las perlas magnéticas a través del anticuerpo. Posteriormente, los glóbulos rojos nucleados se pueden obtener a través de un tamizado magnético.

30 Según los ejemplos particulares de la presente invención, el dispositivo 100 de separación de glóbulos rojos nucleados mencionado anteriormente es adecuado para realizar las siguientes operaciones: tomar una cantidad apropiada de sangre periférica de una mujer embarazada; realizar una anti-coagulación con un agente anticoagulante; diluir la muestra de sangre proporcionalmente con PBS 0,1 M libre de iones calcio e iones magnesio; colocar la muestra diluida lentamente en un reactivo para la centrifugación por gradiente de densidad; y realizar una centrifugación por gradiente de densidad a temperatura ambiente. Después de la centrifugación, se observa una capa de monocitos; esta capa de células se pipetea cuidadosamente, se transfiere a un nuevo tubo de centrífuga, se vuelve a suspender con 3 volúmenes de tampón PBS que contiene BSA al 1 % y se centrifuga nuevamente a temperatura ambiente; el sobrenadante se desecha; el precipitado celular obtenido se lava adicionalmente con el mismo procedimiento dos veces para eliminar el líquido del gradiente de densidad residual; y finalmente, el precipitado celular se vuelve a suspender en PBS que contiene BSA al 0,1 % y se pipeta de manera uniforme. Luego, las perlas magnéticas que transportan un anticuerpo se añaden a las células suspendidas en una proporción de 20 microlitros/10⁶ células; la mezcla de las células y las perlas se centrifuga después de reposar a 4 °C; el sobrenadante se desecha; y el precipitado se vuelve a suspender en PBS que contiene BSA al 0,1 %. Se ensambla un sistema de clasificación de perlas magnéticas: una columna de clasificación se humedece con 500 microlitros de tampón PBS que contiene BSA al 0,1 %; después de vaciar el líquido, las células que se van a clasificar se cargan en la columna; y el líquido del efluente se recoge y se etiqueta como células negativas. El tubo se humedece con PBS que contiene BSA al 0,1 % y, una vez que se vacía el líquido, se repite lo mismo dos veces; PBS/EDTA/BSA se añade a la columna de clasificación, y después de que el líquido se vacíe, se repite lo mismo una vez. Finalmente, se añade PBS que contiene BSA al 0,1 % a la columna de clasificación; se retira el campo magnético; el líquido se lava en un nuevo tubo de centrífuga; y se obtienen glóbulos rojos nucleados. El procedimiento de separación de glóbulos rojos nucleados se ha detallado anteriormente y no se repite.

35 Además, según las realizaciones de la presente invención, con referencia a la Figura 5, el sistema puede comprender además un dispositivo 400 de preparación de biblioteca de secuenciación de genoma completo. Según las realizaciones de la presente invención, el dispositivo 400 de preparación de la biblioteca de secuenciación del genoma completo está conectado al dispositivo de secuenciación 200, y proporciona una biblioteca de secuenciación del genoma completo para la secuenciación mediante el dispositivo de secuenciación 200. Con referencia a la Figura 6, según una realización de la presente invención, el dispositivo 400 de preparación de biblioteca de secuenciación de genoma completo puede comprender además una unidad 400 de lisis de glóbulos

rojos nucleada, una unidad 402 de amplificación de genoma completo y una unidad 403 de construcción de biblioteca de secuenciación. Según realizaciones de la presente invención, la unidad 401 de lisis de glóbulos rojos nucleados está conectada al dispositivo 100 de separación de glóbulos rojos nucleados, y recibe y lisa los glóbulos rojos nucleados separados para liberar el genoma completo de los glóbulos rojos nucleados. La unidad 402 de amplificación del genoma completo está conectada a la unidad 401 de lisis de glóbulos rojos nucleados, y se utiliza para amplificar el genoma completo de los glóbulos rojos nucleados, para obtener un genoma completo amplificado. La unidad 402 de construcción de la biblioteca de secuenciación se utiliza para recibir el genoma completo amplificado y para construir la biblioteca de secuenciación del genoma completo utilizando el genoma completo amplificado. El dispositivo de preparación de la biblioteca de secuenciación del genoma completo puede construir de manera efectiva una biblioteca de secuenciación de glóbulos rojos nucleados.

El término "conectar", como se utiliza en la presente memoria, debe entenderse en un sentido amplio, que no solo puede referirse a una conexión directa, sino que también puede referirse a una conexión indirecta, e incluso se puede utilizar el mismo contenedor o equipo, siempre que se pueda realizar un acoplamiento funcional. Por ejemplo, la unidad 302 de lisis de glóbulos rojos nucleados y la unidad 303 de amplificación del genoma completo pueden estar en el mismo equipo, es decir, una vez realizada la lisis de los glóbulos rojos nucleados, el tratamiento de amplificación del genoma completo se puede realizar en el mismo equipo o contenedor, y el genoma completo liberado no necesita ser enviado a otro equipo o contenedores, siempre que la condición (incluida la condición de reacción y la composición del sistema de reacción) en el equipo se convierta en la adecuada para realizar la reacción de amplificación del genoma completo. Por consiguiente, se realiza un acoplamiento funcional de la unidad 302 de lisis de glóbulos rojos nucleados y la unidad 303 de amplificación del genoma completo, que se considera que está abarcada en el término "conectar".

Los expertos en la materia pueden seleccionar diferentes procedimientos y equipos para construir una biblioteca de secuenciación de genoma completo según un protocolo específico para la técnica de secuenciación genómica aplicada, y para detalles de la construcción de la biblioteca de secuenciación de genoma completo, se puede hacer referencia a las reglas directivas proporcionadas por el fabricante del secuenciador, p. ej., Illumina Corporation. Según una realización de la presente invención, la unidad 303 de amplificación del genoma completo comprende un dispositivo adecuado para amplificar dicho genoma completo utilizando el procedimiento OmniPlex WGA. Por lo tanto, el genoma completo se puede amplificar de manera efectiva y, por lo tanto, se mejora aún más la eficacia de la determinación de una anomalía genómica en los glóbulos rojos nucleados.

Según una realización de la presente invención, el dispositivo 100 de secuenciación de genoma completo incluye al menos uno seleccionado entre el grupo que consiste en illumina-Solexa, ABI-Solid, Roche-454 y un dispositivo de secuenciación de molécula única. Por lo tanto, se pueden utilizar las características de estos dispositivos de secuenciación de alto rendimiento y secuenciación profunda, y por lo tanto se mejora la eficacia de la determinación de la aneuploidía cromosómica en los glóbulos rojos nucleados. Los expertos en la materia entenderán fácilmente que la secuenciación del genoma completo también se puede realizar utilizando otros procedimientos y dispositivos de secuenciación, p. ej., la técnica de secuenciación de tercera generación y técnicas de secuenciación más avanzadas que pueden desarrollarse posteriormente. Según las realizaciones de la presente invención, las longitudes de los datos de secuenciación obtenidos por la secuenciación del genoma completo no están particularmente limitadas.

Como se ha mencionado anteriormente, según las realizaciones de la presente invención, una vez obtenido el resultado de la secuenciación genómica de los glóbulos rojos nucleados, se puede analizar la aneuploidía cromosómica. Por lo tanto, según las realizaciones de la presente invención, el dispositivo 300 de análisis de resultados de secuenciación puede ser adecuado para ejecutar las siguientes operaciones: en primer lugar, dividir la secuencia conocida de un primer cromosoma en una pluralidad de ventanas, la pluralidad de ventanas tiene independientemente una longitud predeterminada, respectivamente; posteriormente, alinear los datos de secuenciación en dicho resultado de secuenciación a la secuencia conocida del primer cromosoma, para obtener el número de datos de secuenciación que caen en cada ventana; finalmente, sobre la base de obtener el número de datos de secuenciación que caen en cada ventana, determinar un primer parámetro; y sobre la base de dicho primer parámetro, determinar si dichos glóbulos rojos nucleados tienen aneuploidía para dicho primer cromosoma. Si los glóbulos rojos nucleados tienen aneuploidía cromosómica se puede determinar de manera efectiva utilizando el sistema 1000.

Según una realización de la presente invención, el dispositivo 300 de análisis de resultados de secuenciación comprende además una unidad de alineamiento de secuencias (no mostrada en las figuras). La unidad de alineamiento de secuencias se utiliza para alinear el resultado de la secuenciación con la información sobre la secuencia genómica conocida, a fin de obtener todos los datos de secuenciación que pueden alinearse con el genoma de referencia y obtener datos de secuenciación del primer cromosoma. De este modo, los datos de secuenciación de un cromosoma específico se pueden determinar de manera efectiva y, por lo tanto, la eficacia de la determinación de la aneuploidía cromosómica en los glóbulos rojos nucleados se mejora aún más. La expresión "primer cromosoma", utilizado en la presente memoria, debe entenderse en un sentido amplio, lo mismo puede referirse a cualquier cromosoma objetivo que se espera que se estudie, su número no se limita a un solo cromosoma, se pueden analizar más e incluso todos los cromosomas simultáneamente. Según las realizaciones de la presente invención, el primer cromosoma puede ser cualquier cromosoma entre los cromosomas humanos, p. ej.,

puede ser al menos uno seleccionado entre el grupo que consiste en el cromosoma humano 21, el cromosoma 18, el cromosoma 13, el cromosoma X y el cromosoma Y. Por consiguiente, las enfermedades cromosómicas humanas comunes pueden determinarse eficazmente, por ejemplo, se pueden predecir enfermedades genéticas fetales.

5 Por lo tanto, el procedimiento de determinación de la aneuploidía cromosómica en glóbulos rojos nucleados según las realizaciones de la presente invención se puede aplicar muy eficazmente en el cribado preimplantacional (PGS) y el diagnóstico preimplantacional (PGD) en el campo de la reproducción *in vitro*, y pruebas prenatales en células nucleadas fetales, etc. Se puede predecir rápidamente si existe una anomalía cromosómica en un feto mediante la simple extracción de glóbulos rojos nucleados, lo que evita el caso de que el feto sufra una enfermedad genética grave.

10 El primer parámetro y otras características se han detallado anteriormente, y no se repiten aquí. Se debe tener en cuenta lo siguiente:

Según una realización de la presente invención, el dispositivo 300 de análisis de resultados de secuenciación puede comprender además una unidad que sea adecuada para determinar un primer parámetro basado en el número de datos de secuenciación que caen en cada ventana, a través de las siguientes etapas: establecer un coeficiente de ponderación predeterminado para el número de datos de secuencia que caen en cada ventana, respectivamente; y según dicho coeficiente de ponderación, realizar promedios ponderados en dicho número de datos de secuenciación que caen en cada ventana para obtener la mediana de dicho primer cromosoma, que forma el primer parámetro de dicho primer cromosoma.

20 Según una realización de la presente invención, el dispositivo 300 de análisis de resultados de secuenciación puede comprender además una unidad que sea adecuada para determinar si dichos glóbulos rojos nucleados tienen aneuploidía para dicho primer cromosoma basándose en el primer parámetro, a través de las siguientes etapas: realizar el mismo tratamiento en un segundo cromosoma al igual que en dicho primer cromosoma, para obtener la mediana de dicho segundo cromosoma; realizar la prueba del valor t en la mediana de dicho primer cromosoma y la mediana de dicho segundo cromosoma, para obtener la diferencia entre dicho primer cromosoma y dicho segundo cromosoma; y comparar dicha diferencia con un umbral predeterminado, y si dicha diferencia es menor que dicho umbral predeterminado, entonces determinar que dichos glóbulos rojos nucleados tienen aneuploidía para el primer cromosoma.

25 Según una realización de la presente invención, el dispositivo 300 de análisis de resultados de secuenciación puede comprender además una unidad que sea adecuada para utilizar la siguiente fórmula para realizar la prueba del valor t en la mediana de dicho primer cromosoma y la mediana de dicho segundo cromosoma,

$$T_{i,j} = \frac{\mu_i - \mu_j}{\sqrt{\frac{\sigma_i^2}{n_i} + \frac{\sigma_j^2}{n_j}}}$$

35 en la que $T_{i,j}$ representa la diferencia entre dicho primer cromosoma y dicho segundo cromosoma, μ_i representa la mediana del primer cromosoma, μ_j representa la mediana del segundo cromosoma, σ_i representa la desviación típica de la distribución del número de datos de secuenciación en cada ventana en el cromosoma, σ_j representa la desviación típica de la distribución del número de datos de secuenciación en cada ventana en el segundo cromosoma, n_i representa el número de ventanas en el primer cromosoma, y n_j representa el número de ventanas en el segundo cromosoma.

40 Además, según las realizaciones de la presente invención, el dispositivo 300 de análisis de resultados de secuenciación es adecuado para determinar si existe una mutación en una región predeterminada en el genoma. Por lo tanto, según una realización de la presente invención, el dispositivo 300 de análisis de resultados de secuenciación puede comprender además: una unidad de determinación para una secuencia de ácido nucleico, una unidad de alineamiento y una unidad de determinación de anomalías. La unidad de determinación para una región de ácido nucleico predeterminada es adecuada para determinar la secuencia de ácido nucleico de la región predeterminada en dichos glóbulos rojos nucleados basándose en el resultado de la secuenciación. La unidad de alineamiento está conectada a dicha unidad de determinación de secuencia de ácido nucleico, y es adecuada para alinear la secuencia de ácido nucleico de la región predeterminada en dichos glóbulos rojos nucleados con una secuencia de ácido nucleico de control. Preferentemente, la secuencia de ácido nucleico de control se almacena en la unidad de alineamiento. Más preferentemente, la secuencia de ácido nucleico de control es la secuencia del genoma humano normal. La unidad de determinación de anomalías está conectada a dicha unidad de alineamiento y es adecuada para determinar si existe una anomalía en la región predeterminada en dichos glóbulos rojos nucleados en función del resultado de dicho alineamiento. El procedimiento y el detalle para determinar una mutación en una región predeterminada se han detallado anteriormente y no se repiten aquí.

Otras características y ventajas descritas en los procedimientos relevantes anteriores también son adecuadas para el sistema para determinar si existe una anomalía en los glóbulos rojos nucleados, y no se repiten aquí.

3. Procedimiento de determinación de la secuencia genómica de glóbulos rojos nucleados fetales

Otro aspecto de la presente invención se refiere a un procedimiento de determinación de la secuencia genómica de glóbulos rojos nucleados fetales, que comprende las siguientes etapas:

5 En primer lugar, los glóbulos rojos nucleados fetales se separan de una muestra de una mujer embarazada. Una vez que se obtienen los glóbulos rojos nucleados fetales, se secuencian al menos una parte del genoma de los glóbulos rojos nucleados separados, para obtener un resultado de secuenciación. Finalmente, sobre la base del resultado de secuenciación obtenido, se determina la secuencia genómica de los glóbulos rojos nucleados fetales.

10 La expresión "secuencia genómica", como se utiliza en la presente memoria, debe entenderse ampliamente, es decir, puede ser la secuencia del genoma completo y también puede ser la secuencia de una parte del genoma. La separación de los glóbulos rojos nucleados fetales de una muestra de una mujer embarazada y la secuenciación de al menos una parte del genoma de los glóbulos rojos nucleados se han detallado anteriormente, y no se repiten aquí. Lo que se debe tener en cuenta es:

15 Según una realización de la presente invención, la muestra de una mujer embarazada es sangre periférica de la mujer embarazada. Según otra realización de la presente invención, la edad gestacional de dicha mujer embarazada es de 12-20 semanas. Según otra realización de la presente invención, separar los glóbulos rojos nucleados fetales de la sangre periférica de dicha mujer embarazada comprende además: realizar una centrifugación por gradiente en dicha sangre periférica utilizando un reactivo de gradiente de densidad, para obtener monocitos; y separar los glóbulos rojos nucleados de dichos monocitos utilizando perlas magnéticas que transportan un anticuerpo, en las que dicho anticuerpo reconoce específicamente un antígeno en la superficie de los glóbulos rojos nucleados. Según
 20 una realización de la presente invención, dicho reactivo de gradiente de densidad es polisacárida, opcionalmente, dicha centrifugación por gradiente se realiza a 800 X g durante 30 minutos. Según una realización de la presente invención, después de que se obtienen dichos monocitos y antes de que se separen dichos glóbulos rojos nucleados utilizando las perlas magnéticas que transportan el anticuerpo anti-CD71, el procedimiento comprende además el lavado de dichos monocitos utilizando un tampón PBS que contiene BSA al 1 %, de modo que para eliminar el
 25 reactivo de gradiente de densidad residual, preferentemente, dicho tampón PBS está libre de iones calcio e iones magnesio. Según una realización de la presente invención, lavar dichos monocitos utilizando tampón PBS que contiene BSA al 1 % comprende además: mezclar dichos monocitos con dicho tampón PBS que contiene BSA al 1 %, para obtener una suspensión que contiene monocitos; y centrifugar dicha suspensión que contiene monocitos, preferentemente, a 200 X g durante 5 minutos, y descartar el sobrenadante, para obtener glóbulos rojos nucleados
 30 lavados. Según una realización de la presente invención, dicho anticuerpo es un anticuerpo que reconoce específicamente CD71. Según una realización de la presente invención, se secuencian el genoma completo de dichos glóbulos rojos nucleados. Según una realización de la presente invención, la secuenciación del genoma completo de dichos glóbulos rojos nucleados comprende además: amplificar el genoma completo de dichos glóbulos rojos nucleados para obtener un genoma completo amplificado; construir una biblioteca de secuenciación del genoma completo utilizando dicho genoma completo amplificado; y secuenciar dicha biblioteca de secuenciación de genoma completo, para obtener un resultado de secuenciación que consiste en una pluralidad de datos de secuenciación. Según una
 35 realización de la presente invención, el genoma completo de dichos glóbulos rojos nucleados se amplifica a través del procedimiento OmniPlex WGA. Según una realización de la presente invención, construir una biblioteca de secuenciación de genoma completo utilizando dicho genoma completo amplificado comprende además: fragmentar dicho genoma completo amplificado, para obtener fragmentos de ADN; realizar una reparación de extremo en dichos fragmentos de ADN, a fin de obtener fragmentos de ADN con extremos reparados; añadir bases A a los extremos 3' de dichos fragmentos de ADN con extremos reparados, para obtener fragmentos de ADN con el extremo pegajoso A; ligar dichos fragmentos de ADN con el extremo pegajoso A a un adaptador, para obtener un producto de
 45 ligamiento; realizar la amplificación por PCR en dicho producto de ligamiento, para obtener un segundo producto de amplificación; y purificar y recuperar dicho segundo producto de amplificación, para obtener un producto recuperado, y dicho producto recuperado forma dicha biblioteca de secuenciación del genoma completa. Según una realización de la presente invención, la fragmentación de dicho genoma completo amplificado se realiza a través de un dispositivo de corte Covaris. Según una realización de la presente invención, las longitudes de dichos fragmentos de
 50 ADN son de aproximadamente 350 pb. Según una realización de la presente invención, la realización de la reparación de extremo en dichos fragmentos de ADN se realiza utilizando el fragmento Klenow, la ADN polimerasa T4 y la polinucleótido quinasa T4, y dicho fragmento Klenow tiene actividad polimerasa 5'→ 3' y actividad exonucleasa 3'→ 5', pero carece de actividad exonucleasa 5'→ 3'. Según una realización de la presente invención, la adición de las bases A a los extremos 3' de dichos fragmentos de ADN con extremos reparados se realiza utilizando el fragmento Klenow (3'-5' exo-). Según una realización de la presente invención, el ligamiento de dichos fragmentos de ADN con el extremo pegajoso A a un adaptador se realiza utilizando ADN ligasa T4. Según una
 55 realización de la presente invención, dicha secuenciación se realiza utilizando al menos una seleccionada entre el grupo que consiste en HiSeq2000, SOLiD, 454 y un dispositivo de secuenciación de una sola molécula. Las ventajas de estas características se han detallado anteriormente, y no se repiten.

60 Otras características y ventajas descritas en el procedimiento para determinar si existe una anomalía genómica en los cromosomas de los glóbulos rojos nucleados anteriores también son adecuadas para el procedimiento de determinación de la secuencia genómica de los glóbulos rojos nucleados fetales, y no se repiten aquí.

La presente invención se describe a continuación a través de realizaciones particulares. Es necesario tener en

cuenta que estas realizaciones son solo para fines de ilustración, por lo que no se debe considerar que limitan la presente invención de ninguna manera.

Realización 1:

Materiales experimentales

5 Las muestras de sangre periférica se obtuvieron de mujeres embarazadas con alto riesgo de bebés con síndrome de Down, todos los cuales ya tenían un resultado final clínico. Si no está especialmente indicado, todos los demás materiales de prueba fueron reactivos preparados por procedimientos convencionales en la técnica o reactivos disponibles comercialmente.

Procedimiento experimental

10 1. Separación de glóbulos rojos nucleados

Se extrajo sangre periférica de mujeres embarazadas (3 ml). EDTA fue seleccionado como un agente anticoagulante. La muestra de sangre se diluyó con solución de tampón fosfato 0,1 M (PBS) libre de Ca^{2+} y Mg^{2+} en una proporción de 1:1. La muestra diluida se colocó lentamente en 3 ml del reactivo Ficoll (un producto de Sigma Corporation, EE. UU.) con una densidad de 1,077, y se realizó una centrifugación por gradiente de densidad a temperatura ambiente, 800 g X 30 min. Después de la centrifugación, se observó una capa de monocitos. La capa de células se pipeteó cuidadosamente y se transfirió a un nuevo tubo de centrifuga de 1,5 ml. Los monocitos obtenidos se volvieron a suspender con 3 volúmenes de PBS que contenían BSA al 1 % y se centrifugaron a temperatura ambiente 200 g X 5 min. El sobrenadante se descartó y el precipitado celular se lavó adicionalmente con el mismo procedimiento dos veces para eliminar el líquido del gradiente de densidad residual. Finalmente, el precipitado celular se volvió a suspender en 300 microlitros de PBS que contenía BSA al 0,1 % y se pipeteó uniformemente. Las células se contaron y luego se añadieron perlas magnéticas que transportan el anticuerpo anti-CD71 (Miltenyi Biotec Corporation, Alemania) en una proporción de 20 microlitros/ 10^6 células, y se mantuvieron a 4 °C durante 15 minutos. Se realizó centrifugación, 300 g X 10 min. El sobrenadante se descartó y el precipitado se volvió a suspender en 500 microlitros de PBS que contenía BSA al 0,1 %. Se montó un sistema de clasificación de perlas magnéticas (Miltenyi Biotec Corporation, Alemania). Una columna de clasificación se humedeció con 500 microlitros de PBS que contenía BSA al 0,1 %. Después de vaciar el PBS, las células a clasificar se cargaron en la columna, y el líquido del efluente se recogió y se marcó como células negativas. El tubo se humedeció con 500 microlitros de PBS que contenía BSA al 0,1 % y, después de vaciar el líquido, se repitió el mismo procedimiento dos veces. Se añadió PBS/EDTA/BSA (500 microlitros) a la columna de clasificación, y después del líquido, se repitió el mismo procedimiento una vez. Finalmente, se añadió 1 ml de PBS que contenía BSA al 0,1 % a la columna de clasificación, la columna de clasificación se retiró del campo magnético y las células se lavaron en un tubo de 15 ml y se marcaron como células positivas. Las células positivas se centrifugaron y se concentraron, solo se retuvieron alrededor de 100 microlitros de la capa más baja y se prepararon para su uso, y los glóbulos rojos nucleados obtenidos se marcaron como GP9.

35 2. Construcción de bibliotecas de secuenciación de genoma completo

La amplificación del genoma completo se realizó en todos los glóbulos rojos nucleados separados. El kit de amplificación de genoma completo de célula única GenomePlex se seleccionó para realizar la amplificación de genoma completo en este experimento. El procedimiento operativo se realizó según las instrucciones proporcionadas por el fabricante, Sigma Corporation. El producto de amplificación se cortó mediante un dispositivo de corte Covaris siguiendo estrictamente las instrucciones anexas al dispositivo de corte para obtener un producto cortado, es decir, fragmentos de ADN, con la banda principal cortada concentrada a aproximadamente 350 pb.

La reparación de extremo se realizó en el producto cortado y las bases A se añadieron a un extremo utilizando el procedimiento específico de la siguiente manera:

La reacción para la reparación de extremo se realizó con el siguiente sistema:

10Xtampón polinucleótido quinasa T4	10 µl
dNTPs (10 mM)	4 µl
ADN polimerasa T4	5 µl
Fragmento Klenow	1 µl
Polinucleótido quinasa T4	5 µl
Fragmentos de ADN procedentes del producto cortado	30 µl
ddH ₂ O añadido hasta 100 µl	

45 Después de la reacción a 20 °C durante 30 minutos, el producto con extremos reparados se recuperó utilizando el kit de purificación de PCR (QIAGEN). El producto obtenido se disolvió finalmente en 34 µl de tampón EB.

ES 2 699 743 T3

La reacción para añadir las bases A a un extremo se llevó a cabo con el siguiente sistema:

10Xtampón Klenow	5 µl
dATP (1 mM)	10 µl
Klenow (3'-5'exo ⁻)	3 µl
ADN	32 µl

Después de la incubación a 37 °C durante 30 minutos, el ADN con la base A añadida a un extremo se purificó con el kit de purificación por PCR MinElute® (QIAGEN) y el producto se disolvió en 12 µl de EB.

5 La reacción de ligamiento para el adaptador fue la siguiente:

2x tampón de ligamiento rápido de ADN	25 µl
PEI Un adaptador oligomix (20 µM)	10 µl
ADN ligasa T4	5 µl
ADN con bases A añadidas a un extremo	10 µl

Después de la reacción a 20 °C durante 15 minutos, el producto de ligamiento se recuperó utilizando el kit de purificación de PCR (QIAGEN). El producto se disolvió finalmente en 32 µl de tampón EB.

El sistema de reacción de PCR se preparó en un tubo de 0,2 ml de PCR:

Muestra	10 µl
Mezcla de ADN polimerasa de fusión	25 µl
Cebador* de PCR (10 pmol/lil)	1 µl
Índice del cebador** N (10 pmol/µl)	1 µl
Agua UltraPure™	13 µl

10

*La secuencia del cebador de PCR fue

AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCG
ATCT

**La secuencia del índice del cebador N fue

CAAGCAGAAGACGGCATAACGAGATCGTGATGTGACTGGAGTTCAGACGTGTGC
TCTTCCGATCT

15 El programa de reacción fue el siguiente:

98°C	30 s	}	10 ciclos
98°C	10 s		
65°C	30 s		
72°C	30 s		
72°C	5 min		
4°C	mantenido		

El producto de PCR se recuperó utilizando el kit de purificación de PCR (QIAGEN). La muestra se disolvió finalmente en 22 µl de tampón EB.

20 Entre ellos, el índice N era una secuencia de etiqueta única de 8 pb en cada biblioteca, respectivamente. Las bibliotecas construidas fueron probadas por Agilent®Bioanalyzer 2100 y los resultados se muestran en la Figura 8.

Como se muestra en la Figura 8, el intervalo de distribución de fragmentos de las bibliotecas construidas cumplió con los requisitos. Las bibliotecas se cuantificaron adicionalmente mediante el procedimiento de Q-PCR respectivamente. Después de ser calificadas, las bibliotecas se mezclaron en el mismo carril en una celda de flujo para realizar la secuenciación *in silico*, y para ahorrar el costo, se seleccionó la secuenciación de extremo único. En este ejemplo, se seleccionó el instrumento de secuenciación illumina@HiSeq2000™ y un procedimiento para realizar la secuenciación, en el que el ajuste de los parámetros y el procedimiento operativo del instrumento se realizaron en estricta conformidad con el manual operativo proporcionado por el fabricante Illumin@ (disponible en www.illumina.com/support/documentation.ilmn). En este ejemplo, se utilizó el secuenciador HiSeq2000™, y el número de ciclos de secuenciación fue el índice PE91 (es decir, la secuenciación del índice de 91 pb en el extremo del par).

3. Análisis de datos

Los resultados de la secuenciación obtenidos mediante la secuenciación se muestran en la Tabla 1, en la que el número total de datos de secuenciación (lecturas) obtenidos mediante la secuenciación de la muestra GP9 fue 13.407,381, el número de aquellos que pudieron alinearse con el genoma de referencia (HG19) fue 9.217,701, la tasa de alineamiento fue de 68,70 %, el número de datos de secuenciación que podrían alinearse de manera única con la secuencia de referencia fue de 7.341,230, y la tasa de alineamiento única fue de 80 %.

Tabla 1 Datos de secuenciación

Nombre de la muestra	Número total de los datos de secuenciación	Número de datos de secuenciación que podría alinearse	Tasa de alineamiento	Número de datos de secuenciación que podría alinearse de manera única	Tasa de alineamiento único
GP9	13407381	9217701	68,70 %	7341230	80 %

Después del análisis preliminar de los datos, la longitud de lectura de secuenciación se alineó con la secuencia de referencia hg19 utilizando SOAP v2.20, y el procedimiento de alineamiento fue el siguiente:

Se obtuvo un coeficiente de ponderación por el siguiente procedimiento:

En primer lugar, cada cromosoma del genoma completo humano se dividió en ventanas con una longitud fija de 60 kB respectivamente, y la posición inicial y la posición final de cada ventana se registraron y el valor promedio de GC (registrado como GC_{ref}) se obtuvo mediante estadísticas. Después de que se obtuvieran los datos de secuenciación de una muestra a analizar, las secuencias en el resultado de la secuenciación se alinearon con el genoma. Los datos de secuenciación que se combinaron perfectamente y se alinearon correctamente solo una vez, es decir; se eliminaron los datos de secuenciación alineados de forma única, y se obtuvo la información sobre los sitios de todos los datos de secuenciación alineados de forma única. Se contó el número de datos de secuenciación alineados de forma única (registrados como $UR_{muestra}$) en cada ventana en cada cromosoma en el genoma correspondiente al resultado de alineamiento de la muestra a analizar, la $UR_{muestra}$ y GC_{ref} obtenidos se representaron para obtener la Figura 8A. Como se muestra en la Figura 8A, el sesgo de GC introducido a través del secuenciador causó una mayor distribución de los datos de secuenciación con GC aproximadamente en la región [0,35, 0,55]. Los puntos distintos en la Figura 8A se ajustaron con el procedimiento de spline liso en una curva lisa, es decir, la Figura 8B, que era un diagrama que refleja la relación entre el contenido de GC y el número de datos de secuenciación. En la Figura 8B, todas las ventanas se dividieron en la longitud de la etapa de 1 % del valor promedio de GC. El número de datos de secuenciación alineados de forma única, es decir, $M_{ajustado}$, en una ventana correspondiente a cada valor promedio de GC podría obtenerse a partir de los datos ajustados. Según la fórmula $W_{GC} = M/M_{ajustado}$, se obtuvo el coeficiente de ponderación predeterminado W_{GC} , y véase la Figura 8C para la distribución del coeficiente de ponderación, en la que M es el número de datos de secuenciación alineados de forma única de la muestra a analizar que se caen en una ventana de un valor promedio de GC igual.

La cantidad de datos de secuenciación alineados de forma única (registrados como UR_{bin}) que caen en las ventanas de cada cromosoma en el genoma para la muestra a analizar se corrigió utilizando el coeficiente de ponderación obtenido, $UR_{ajustado}$ (correspondiente a cada ventana) = $UR_{bin} * W_{GC}$, se calculó la mediana UR_i ($i = 1, 2, \dots, 22$) del número corregido de datos de secuenciación alineados de forma única para cada cromosoma, y se dibujó y obtuvo un diagrama de caja, Figura 8D. En la Figura 8D se puede ver que el número de datos de secuenciación alineados de manera única que corresponden al cromosoma 21 en la muestra GP9 fue obviamente mayor que los correspondientes a otros cromosomas, y se pudo determinar que existía aneuploidía en el cromosoma 21.

Además, la prueba del valor t se realizó en los cromosomas 13, 18 y 21 en los que la aneuploidía es común. Para estas tres aneuploides relativamente comunes, se realiza la prueba del valor t en los volúmenes de datos de estos tres cromosomas según la siguiente fórmula,

$$T_{i,i} = \frac{\mu_i - \mu_j}{\sqrt{\frac{\sigma_i^2}{n_i} + \frac{\sigma_j^2}{n_j}}} \quad (\text{fórmula 1})$$

en la que μ es UR ($i = 13, 18, 21; j = 1, 2, \dots, 22$) para algunos cromosomas de la muestra a analizar; i y j representan los cromosomas i, j ; σ representa la desviación típica de la distribución de UR para algunos cromosomas; y n representa el número de ventanas en algunos cromosomas.

- 5 El cromosoma i se comparó como se ha mencionado anteriormente con los cromosomas 1-12 respectivamente y la diferencia promedio (valor t) se obtuvo según la siguiente fórmula,

$$T_i = \frac{1}{12} * \sum_{j=1}^{12} T_{i,j}$$

10 Se seleccionaron los cromosomas 1-12, ya que la fluctuación de los datos de los cromosomas 1-12 fue relativamente pequeña, y los cromosomas de nuestro interés (incluidos cr13, cr18 y cr21) se excluyeron en la medida de lo posible. Los umbrales establecidos en este procedimiento (para autosomas) fueron: valor $t \leq -4$, lo que significa que el cromosoma correspondiente tenía un alto riesgo de trisomía; $-4 < \text{valor } t \leq -3,5$ significa que el resultado de la detección era incierto, y las muestras debían recuperarse para cargarse en el instrumento o las bibliotecas debían reconstruirse para cargarse en el instrumento para determinar el resultado de detección y el valor $t > -3,5$ significa que el resultado de la detección tenía un bajo riesgo. Se puede ver en la Tabla 2 que entre los valores de t correspondientes a los cromosomas 13, 18 y 21 de la muestra GP9, el valor de t para el cromosoma 21 ya había excedido los criterios de alto riesgo de trisomía 21.

Tabla 2: Resultados de la prueba del valor de t

Nombre de la muestra	valor de t para el cromosoma 13	valor de t para el cromosoma 18	valor de t para el cromosoma 21	Resultado
GP9	0,96287409	-2,206632741	-34,59904702	trisomía 21

20 En la descripción de la presente memoria descriptiva, mediante la descripción de los términos de referencia "una realización", "algunas realizaciones", "ejemplo", "ejemplo particular" o "algunos ejemplos" y otros se tiene por objeto que los distintivos, estructuras, materiales o características descritos en conjunto con la realización o ejemplo están contenidos en al menos una realización o ejemplo de la presente invención. En la presente memoria descriptiva, la expresión esquemática de los términos mencionados anteriormente no se refiere necesariamente a la misma realización o ejemplo.

25

REIVINDICACIONES

1. Un procedimiento para determinar si existe una anomalía genómica en un feto, que comprende las etapas de:

- a) separar glóbulos rojos nucleados fetales de una muestra obtenida de una mujer embarazada;
 - b) amplificar un genoma completo de los glóbulos rojos nucleados para obtener un genoma completo amplificado, construir una biblioteca de secuenciación del genoma completo utilizando el genoma completo amplificado y secuenciar la biblioteca de secuenciación del genoma completo, para obtener un resultado de secuenciación que contiene una pluralidad de datos de secuenciación;
 - c) determinar si la anomalía genómica existe en el feto mediante la determinación de la anomalía genómica en los glóbulos rojos nucleados basándose en el resultado de la secuenciación,
- en el que la anomalía genómica es una aneuploidía cromosómica, y la anomalía genómica en los glóbulos rojos nucleados se determina de la siguiente manera:

- c1) dividir la secuencia conocida de un primer cromosoma en una pluralidad de ventanas, la pluralidad de ventanas tiene independientemente una longitud predeterminada, respectivamente;
 - c2) alinear los datos de secuenciación en el resultado de la secuenciación con la secuencia conocida del primer cromosoma, para obtener el número de datos de secuenciación que caen en cada ventana;
 - c3) determinar un primer parámetro en base al número de datos de secuencia que caen en cada una de las ventanas;
 - c4) determinar si los glóbulos rojos nucleados tienen aneuploidía para el primer cromosoma en base al primer parámetro,
- en el que el primer parámetro se determina por:

- c4-i) establecer un coeficiente de ponderación predeterminado para el número de datos de secuencia que caen en cada ventana, respectivamente; y
- c4-ii) según el coeficiente de ponderación, realizar un promedio ponderado en el número de datos de secuenciación que caen en cada una de las ventanas para obtener la mediana del primer cromosoma, la mediana del primer cromosoma que forma el primer parámetro del primer cromosoma,

en el que el coeficiente de ponderación predeterminado se obtiene asociando el número de datos de secuenciación que caen en cada ventana con el contenido de GC de la ventana respectiva,

c5) realizar el mismo tratamiento en un segundo cromosoma al igual que en el primer cromosoma, para obtener la mediana del segundo cromosoma, en el que

el primer cromosoma es al menos uno seleccionado entre el grupo que consiste en cromosoma 21, cromosoma 18, cromosoma 13, cromosoma X y cromosoma Y humanos; y
 el segundo cromosoma es uno cualquiera seleccionado entre el grupo que consiste en los cromosomas 1-12 en el genoma humano;

- c6) realizar la prueba del valor t en la mediana del primer cromosoma y la mediana del segundo cromosoma para obtener una diferencia entre el primer cromosoma y el segundo cromosoma; y
- c7) comparar la diferencia con un primer umbral y un segundo umbral predeterminados, y si la diferencia es inferior a los umbrales predeterminados, entonces concluir que los glóbulos rojos nucleados tienen aneuploidía para el primer cromosoma, y si la diferencia es superior a los umbrales predeterminados, entonces concluir que los glóbulos rojos nucleados no tienen aneuploidía para el primer cromosoma.

2. El procedimiento según la reivindicación 1, en el que la muestra de la mujer embarazada es sangre periférica de la mujer embarazada, y la edad gestacional de la mujer embarazada es de 12 a 20 semanas, preferentemente la etapa que consiste en separar los glóbulos rojos nucleados fetales de la sangre periférica de la mujer embarazada comprende además:

- realizar una centrifugación por gradiente en la sangre periférica utilizando un reactivo de gradiente de densidad, para obtener monocitos; y
- enriquecer los glóbulos rojos nucleados de los monocitos utilizando perlas magnéticas que transportan un anticuerpo, en el que el anticuerpo reconoce específicamente un antígeno en la superficie de los glóbulos rojos nucleados,

preferentemente el anticuerpo es un anticuerpo que reconoce específicamente CD71.

3. El procedimiento según la reivindicación 1, en el que la etapa de construir la biblioteca de secuenciación del genoma completo utilizando el genoma completo amplificado comprende además:

- fragmentar el genoma completo amplificado, para obtener fragmentos de ADN;
- realizar una reparación de extremo en los fragmentos de ADN, a fin de obtener fragmentos de ADN con extremos reparados;
- añadir bases A a los extremos 3' de los fragmentos de ADN con extremos reparados, para obtener fragmentos de ADN con un extremo pegajoso A;

ligar los fragmentos de ADN con el extremo pegajoso A a un adaptador, para obtener un producto de ligamento; realizar la amplificación por PCR en el producto de ligamento, para obtener un segundo producto de amplificación; y

5 purificar y recuperar el segundo producto de amplificación, para obtener un producto recuperado, y el producto recuperado que forma la biblioteca de secuenciación del genoma completo.

4. El procedimiento según la reivindicación 1, en el que las longitudes predeterminadas de la pluralidad de ventanas son las mismas, preferentemente, las longitudes predeterminadas de la pluralidad de ventanas son todas de 60 kB.

5. El procedimiento según la reivindicación 1, en el que los datos de secuenciación que caen en cada una de las ventanas son datos de secuenciación alineados de forma única.

10 6. El procedimiento según la reivindicación 1, en el que la siguiente fórmula se utiliza para realizar la prueba del valor t en la mediana del primer cromosoma y la mediana del segundo cromosoma,

$$T_{i,j} = \frac{\mu_i - \mu_j}{\sqrt{\frac{\sigma_i^2}{n_i} + \frac{\sigma_j^2}{n_j}}}$$

15 en la que $T_{i,j}$ representa la diferencia entre el primer cromosoma y el segundo cromosoma, μ_i representa la mediana del primer cromosoma, μ_j representa la mediana del segundo cromosoma, σ_i representa la desviación típica de la distribución del número de datos de secuenciación en cada ventana en el cromosoma, σ_j representa la desviación típica de la distribución del número de datos de secuenciación en cada ventana en el segundo cromosoma, n_i representa el número de ventanas en el primer cromosoma, y n_j representa el número de ventanas en el segundo cromosoma.

20 7. El procedimiento según la reivindicación 1, en el que el primer umbral predeterminado es -4 o menor, y el segundo umbral es -3,5 o superior.

8. El procedimiento según la reivindicación 1, en el que la anomalía genómica es una mutación en una región predeterminada, y la anomalía genómica que existe en los glóbulos rojos nucleados en base del resultado de secuenciación se determina mediante un procedimiento que comprende:

25 determinar la secuencia de ácido nucleico de la región predeterminada en dichos glóbulos rojos nucleados en base al resultado de secuenciación;

alinear la secuencia de ácido nucleico de la región predeterminada en los glóbulos rojos nucleados con una secuencia de ácido nucleico de control, preferentemente, la secuencia de ácido nucleico de control es una secuencia genómica humana normal; y

30 determinar si existe una anomalía en la región predeterminada en los glóbulos rojos nucleados en base al resultado del alineamiento.

9. Un sistema para determinar si existe una anomalía genómica, que comprende:

un dispositivo de separación de glóbulos rojos nucleados, el dispositivo de separación de glóbulos rojos nucleados se utiliza para separar glóbulos rojos nucleados fetales de una muestra obtenida de una mujer embarazada;

35 un dispositivo de secuenciación, el dispositivo de secuenciación se utiliza para amplificar un genoma completo de los glóbulos rojos nucleados para obtener un genoma completo amplificado, construir una biblioteca de secuenciación de genoma completo utilizando el genoma completo amplificado, y secuenciar la biblioteca de secuenciación del genoma completo, para obtener un resultado de secuenciación que contiene una pluralidad de datos de secuenciación; y

40 un dispositivo de análisis de resultados de secuenciación, el dispositivo de análisis de resultados de secuenciación está conectado al dispositivo de secuenciación, para recibir el resultado de secuenciación del dispositivo de secuenciación, y determinar si existe una anomalía genómica en dichos glóbulos rojos nucleados en base al resultado de secuenciación,

45 en el que la anomalía genómica es una aneuploidía cromosómica, y el dispositivo de análisis de resultados de secuenciación es adecuado para las siguientes operaciones:

c1) dividir la secuencia conocida de un primer cromosoma en una pluralidad de ventanas, la pluralidad de ventanas tiene independientemente una longitud predeterminada, respectivamente;

c2) alinear los datos de secuenciación en el resultado de la secuenciación con la secuencia conocida del primer cromosoma, para obtener el número de datos de secuenciación que caen en cada ventana;

50 c3) determinar un primer parámetro en base al número de datos de secuenciación que caen en cada una de las ventanas; y

c4) determinar si los glóbulos rojos nucleados tienen una aneuploidía para el primer cromosoma en base al

primer parámetro,
en el que el primer parámetro se determina por:

c4-i) establecer un coeficiente de ponderación predeterminado para el número de datos de secuenciación que caen en cada una de las ventanas, respectivamente; y

5 c4-ii) según el coeficiente de ponderación, realizar un promedio ponderado en el número de datos de secuenciación que caen en cada una de las ventanas para obtener la mediana del primer cromosoma, la mediana del primer cromosoma que forma el primer parámetro del primer cromosoma,

en el que el coeficiente de ponderación predeterminado se obtiene asociando el número de datos de secuenciación que caen en cada ventana con el contenido de GC de la ventana respectiva,

10 c5) realizar el mismo tratamiento en un segundo cromosoma al igual que en el primer cromosoma, para obtener la mediana del segundo cromosoma, en el que

el primer cromosoma es al menos uno seleccionado entre el grupo que consiste en cromosoma 21, cromosoma 18, cromosoma 13, cromosoma X y cromosoma Y humanos; y

15 el segundo cromosoma es uno cualquiera seleccionado entre el grupo que consiste en los cromosomas 1-12 en el genoma humano;

c6) realizar la prueba del valor t en la mediana del primer cromosoma y la mediana del segundo cromosoma para obtener una diferencia entre el primer cromosoma y el segundo cromosoma; y

20 c7) comparar la diferencia con un primer umbral y un segundo umbral predeterminados, y si la diferencia es inferior a los umbrales predeterminados, entonces concluir que los glóbulos rojos nucleados tienen una aneuploidía para el primer cromosoma, y si la diferencia es superior a los umbrales predeterminados, entonces concluir que los glóbulos rojos nucleados no tienen una aneuploidía para el primer cromosoma, en el que el dispositivo de separación de glóbulos rojos nucleados comprende además:

una unidad de separación de monocitos, la unidad de separación de monocitos es adecuada para realizar una centrifugación por gradiente en la muestra de una mujer embarazada utilizando un reactivo de gradiente de densidad, para obtener monocitos, en el que la muestra de la mujer embarazada es preferentemente sangre periférica de la mujer embarazada; y

25 una unidad de enriquecimiento magnético, la unidad de enriquecimiento magnético está conectada a la unidad de separación de monocitos, y es adecuada para enriquecer los glóbulos rojos nucleados de los monocitos utilizando perlas magnéticas que transportan un anticuerpo, en el que el anticuerpo reconoce específicamente un antígeno en la superficie de los glóbulos rojos nucleados.

10. El sistema según la reivindicación 9, que comprende además un dispositivo de preparación de biblioteca de secuenciación del genoma completo, el dispositivo de biblioteca de secuenciación del genoma completo está conectado al dispositivo de secuenciación, y proporciona una biblioteca de secuenciación del genoma completo utilizada para la secuenciación del dispositivo de secuenciación,

35 en el que el dispositivo de preparación de la biblioteca de secuenciación del genoma completo comprende además:

una unidad de lisis de glóbulos rojos nucleados, la unidad de lisis de glóbulos rojos nucleados se conecta al dispositivo de separación de glóbulos rojos nucleados, y recibe y lisa los glóbulos rojos nucleados, para liberar el genoma completo de los glóbulos rojos nucleados;

40 una unidad de amplificación del genoma completo, la unidad de amplificación del genoma completo está conectada a la unidad de lisis de glóbulos rojos nucleados y se utiliza para amplificar el genoma completo de los glóbulos rojos nucleados, a fin de obtener un genoma completo amplificado; y

una unidad de construcción de biblioteca de secuenciación, la unidad de construcción de biblioteca de secuenciación se utiliza para recibir el genoma completo amplificado, y construir la biblioteca de secuenciación del genoma completo utilizando el genoma completo amplificado.

45 11. El sistema según la reivindicación 9, en el que el dispositivo de análisis de resultados de secuenciación comprende además:

una unidad de determinación de secuencia de ácido nucleico, la unidad de determinación de secuencia de ácido nucleico en una región predeterminada es adecuada en base al resultado de secuenciación, determinar la secuencia de ácido nucleico de la región predeterminada en los glóbulos rojos nucleados;

50 una unidad de alineamiento, la unidad de alineamiento está conectada a la unidad de determinación de secuencia de ácido nucleico, y es adecuada para el alineamiento de una secuencia de ácido nucleico de la región predeterminada en los glóbulos rojos nucleados con una secuencia de ácido nucleico de control; y

una unidad de determinación de anomalías, la unidad de determinación de anomalías está conectada a la unidad de alineamiento y es adecuada para determinar si existe una anomalía en la región predeterminada en los glóbulos rojos nucleados en base al resultado de alineamiento.

55

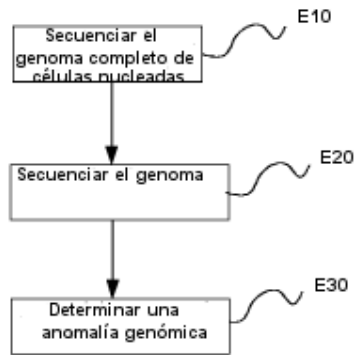


Fig. 1

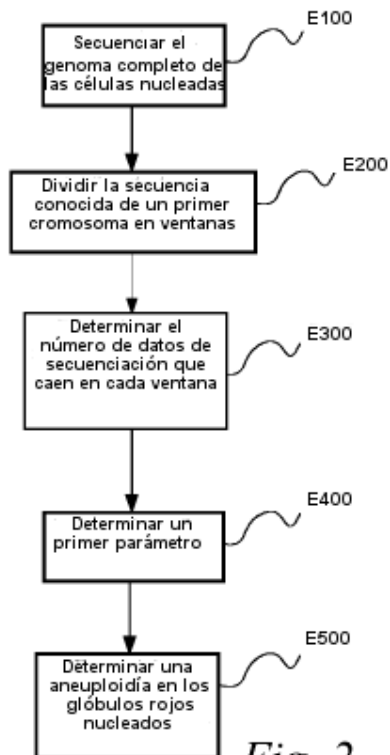


Fig. 2



Fig. 3

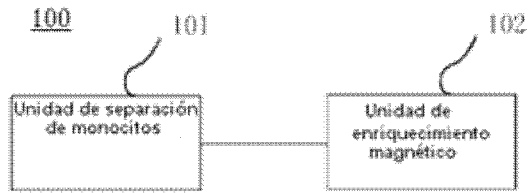


Fig. 4

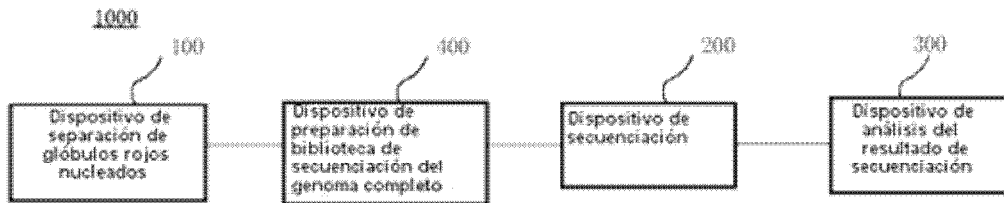


Fig. 5

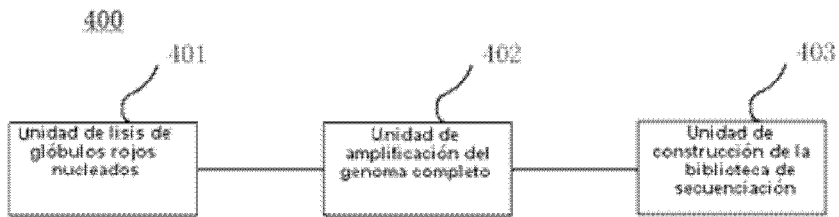
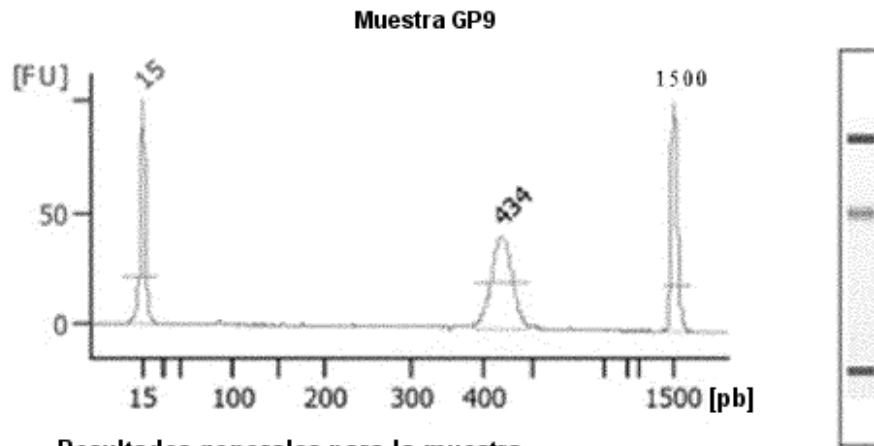


Fig. 6



Resultados generales para la muestra

Picos descubiertos		1
Tamaño del valor de pico	Concentración de masa	Concentración de masa
434 pb	3,6ng/ μ I	12,5nmol/I

Fig. 7

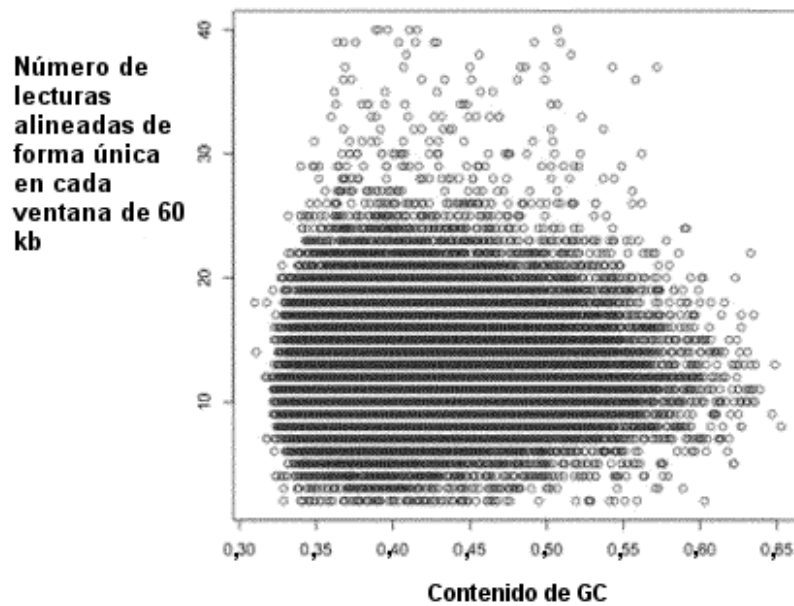


Fig. 8A

Número
corregido de
las lecturas
alineadas de
forma única en
cada ventana
de 60 kb

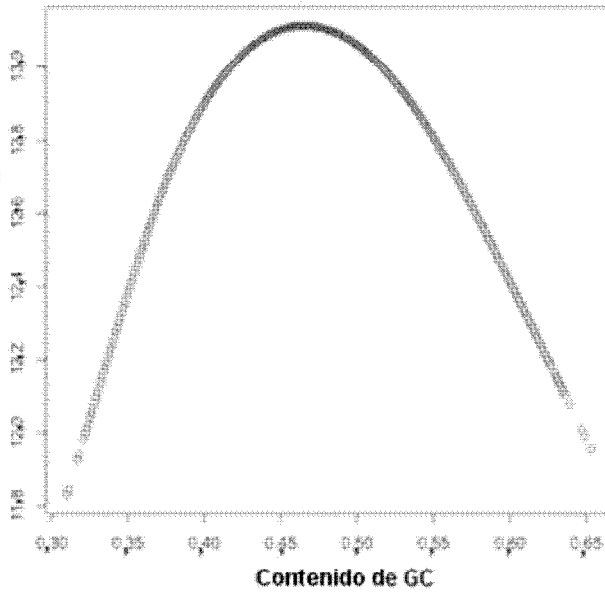


Fig. 8B

Peso

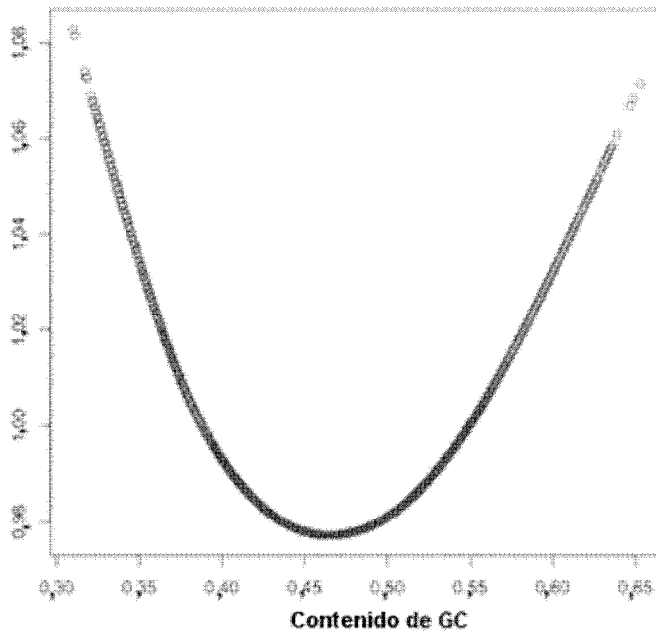


Fig. 8C

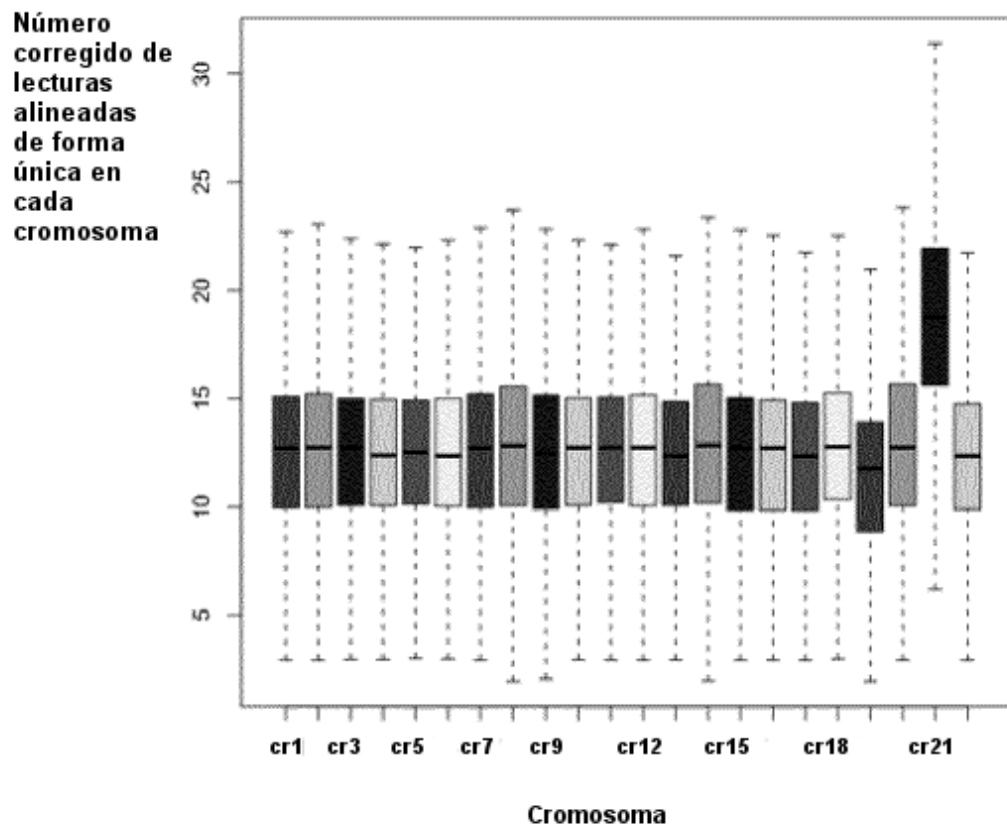


Fig. 8D