

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 989 171**

51 Int. Cl.:

G16B 30/00 (2009.01)

G16B 99/00 (2009.01)

A61B 5/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **16.05.2019 PCT/IB2019/054068**

87 Fecha y número de publicación internacional: **28.11.2019 WO19224668**

96 Fecha de presentación y número de la solicitud europea: **16.05.2019 E 19734884 (0)**

97 Fecha y número de publicación de la concesión europea: **03.07.2024 EP 3797418**

54 Título: **Método para determinar la probabilidad de riesgo de trastornos cromosómicos y genéticos a partir de ADN libre de origen fetal**

30 Prioridad:

23.05.2018 IT 201800005623

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

25.11.2024

73 Titular/es:

**ARTEMISIA S.P.A. (100.0%)
Viale Liegi 41
00198 Roma (RM), IT**

72 Inventor/es:

**DELLO RUSSO, CLAUDIO;
GIORLANDINO, CLAUDIO y
MESORACA, ALVARO**

74 Agente/Representante:

CARVAJAL Y URQUIJO, Isabel

ES 2 989 171 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Método para determinar la probabilidad de riesgo de trastornos cromosómicos y genéticos a partir de ADN libre de origen fetal

5

Campo de la invención

La presente invención se refiere a un método para determinar la probabilidad de riesgo de trastornos cromosómicos y genéticos a partir de ADN libre de origen fetal.

10

En particular, la presente invención se refiere a un método para determinar la probabilidad de riesgo de anomalía en el cromosoma de interés para determinar la probabilidad de microeliminación o riesgo de microduplicación en las regiones genómicas que componen el genoma.

15

Estado de la técnica

Las pruebas prenatales, incluidas la examinación prenatal y diagnóstico prenatal, permiten la detección de problemas con relación al desarrollo anatómico y fisiológico del cigoto, embrión o feto en la fase inicial de gestación.

20

Entre los métodos de examinación prenatal, de tipo no invasivo, se conoce un método de análisis de sangre, con base en la detección de ADN libre de origen fetal a partir de células presentes en la sangre materna, o ADNflic "ADN fetal libre de células", definido como "*Prueba prenatal no invasiva*" (NIPT). La lectura del ADN fetal, además de revelar el sexo del feto, nos permite reconocer la presencia de anomalías cromosómicas (errores en el número de cromosomas), en particular para el cromosoma 21 (síndrome de Down), y para otras trisomías más comunes (cromosomas 13, Patau, y 18, Edwards). La NIPT es una prueba predictiva y no diagnóstica capaz de identificar a las mujeres con alto riesgo de tener un hijo con enfermedades cromosómicas cuyo posible resultado positivo requiere pruebas de diagnóstico prenatal de tipo invasivo, tal como la amniocentesis o la villocentesis. La NIPT es una prueba que se puede realizar a partir de la décima semana de gestación y tiene una alta fiabilidad pero también una alta varianza en los resultados con relación a la pluralidad de pruebas en el mercado.

30

Sin embargo, la NIPT exhibe una pluralidad de límites, principalmente con relación a la sensibilidad no alta (identificación de aproximadamente 50% de las anomalías identificadas rutinariamente con pruebas de diagnóstico prenatal de tipo invasivo), imposibilidad de distinguir todas las aneuploidías, presencia de falsos positivos y negativos (entre 0,1% y 0,5%), necesidad de una cantidad adecuada de ADN fetal en la sangre materna (más del 5%), así como la imposibilidad de distinguir la condición del feto único en el caso de embarazo gemelar.

35

A partir de las primeras semanas de embarazo es posible detectar en la circulación sanguínea materna la presencia de células fetales intactas y ADN libre de origen fetal, lo que constituye una fracción variable en general comprendida entre el 3 y el 20% del ADN extracelular total detectable en la circulación materna, con una concentración que tiende a incrementar progresivamente durante el embarazo y puede variar en la presencia de aneuploidías fetales, en particular disminuyendo en la presencia de trisomía 13 y 18 e incrementando en la presencia de trisomía 21. El ADN libre de origen fetal está presente en la forma de fragmentos de tamaño reducido en comparación a aquellos que constituyen la fracción materna. La placenta y, en particular, las células de sincitio de trofoblasto apoptóticas, son la principal fuente de ADNflic, que entonces se remueve por completo de la circulación materna, probablemente a través de la excreción renal, a las pocas horas del parto.

40

Un punto crítico adicional del análisis de ADNflic es que representa, en promedio, el 10% del ADN total extraído del plasma, en tanto que la fracción predominante está representada por el ADN materno. Sin embargo, es posible realizar la prueba de detección prenatal de las aneuploidías fetales más comunes por el uso de tecnologías de alta sensibilidad tal como PCR digital, secuenciación masiva paralela (MPS), NGS basada en SNP, análisis digital de regiones seleccionadas (DANSR).

50

El resultado de la prueba de tipo no invasivo también está condicionado por el porcentaje de ADN fetal presente en el plasma que debe ser superior al 3-4%, puesto que cantidades más bajas pueden dar por resultado resultados falsos negativos. La cantidad relativa de ADN fetal se reduce en condiciones particulares como la edad gestacional inferior a la 10ª semana y un alto índice de masa corporal.

55

US2016/224724 divulga métodos y dispositivos para una evaluación no invasiva de variaciones genéticas en un feto y se considera que es la técnica anterior más cercana.

60

Breve descripción de la invención

El objeto de la presente invención es proporcionar un método de detección prenatal no invasivo para calcular las probabilidades de riesgo de anomalías genéticas y cromosómicas encontradas en el feto. Este método se implementa debido a un software peculiar que ofrece una mayor precisión de los resultados obtenibles en comparación a las metodologías de NIPT actualmente disponibles para los pacientes.

65

Los objetos anteriores se logran por un método para determinar la probabilidad de riesgo de trastornos cromosómicos y genéticos a partir del ADN libre de origen fetal de acuerdo con las reivindicaciones anexas.

5 Descripción de la invención

La presente invención se refiere a un método de examen y análisis de fragmentos de ADN libre que circulan en la sangre materna, denominados ADN libre fetal o ADN_{fl}, derivados de trofoblasto (la estructura que forma la placenta), que, en la gran mayoría de los casos, siguen la composición del ADN fetal. En particular, la presente invención se refiere a un método para determinar la probabilidad de riesgo de trastornos cromosómicos y genéticos a partir del ADN libre de origen fetal.

Más específicamente, el método de acuerdo con la presente invención es un método de detección capaz de evaluar el riesgo de que el feto sea portador de anomalías cromosómicas. El método de acuerdo con la presente invención es capaz de obtener un número muy bajo de falsos negativos y falsos positivos, menor que 0,1%, en comparación a los métodos conocidos que, en su lugar, tienen valores que varían de 0,1% a 0,3% de los casos.

De acuerdo con la presente invención, el método para examinar y analizar el ADN fetal presente en una muestra de plasma materno, previamente tomada, utiliza una tecnología para la preparación de bibliotecas de ADN para verificar la presencia de anomalías cromosómicas. El ADN aislado de la muestra de sangre materna incluye material genético tanto materno como fetal. A continuación, se aísla el ADN_{flc} y se analiza la cantidad relativa de material genético que se alinea con los cromosomas estudiados.

El método para determinar la probabilidad de riesgo de trastornos cromosómicos y genéticos a partir del ADN libre de origen fetal, de acuerdo con la presente invención, que comprende una pluralidad de pasos, que se describen a continuación.

Primero está el paso de extraer el ADN libre de origen fetal, es decir, el ADN_{flc}, de una muestra de plasma materno previamente tomada, preferentemente contenida en tubos de ensayo adecuados.

Se puede utilizar un protocolo automatizado para la purificación de ADN total de sangre fetal, aislado de muestras de sangre materna. Por ejemplo, la tecnología de partículas magnéticas se puede utilizar para purificar ácidos nucleicos de alta calidad que están libres de proteínas, nucleasas y otras impurezas. Los ácidos nucleicos purificados se utilizan posterior y directamente en aplicaciones posteriores (tal como la preparación de bibliotecas de ADN).

A continuación, se lleva a cabo un paso de secuenciación del ADN libre de origen fetal.

El ADN extraído de la muestra de plasma materno se adiciona a una reacción de reparación final que incluye incubar la muestra de ADN con una mezcla enzimática. Los fragmentos se fosforilan en el extremo de 5' para permitir la ligadura posterior con los oligonucleótidos adaptadores. La reacción termina con un paso de incubación para desnaturalizar las enzimas de reparación finales. Al final de la reacción de reparación final, se obtiene la reacción de ligadura con los adaptadores. La reacción de ligadura con los adaptadores incluye una incubación de la muestra con oligonucleótidos. Estas enzimas unen los oligonucleótidos adaptadores a fragmentos de ADN y reparan la unión entre el fragmento de ADN y el adaptador unido. Al final de la reacción de ligadura con los adaptadores, se realiza un paso de limpieza utilizando los reactivos en esferas paramagnéticas. Se adicionan esferas paramagnéticas a la muestra unida a los adaptadores, para unir los fragmentos de ADN unidos a los adaptadores y aislarlos de los adaptadores no utilizados. Las esferas a las que se unen los fragmentos de ADN se separan de la solución por medio de un imán y se lavan con agua libre de nucleasas antes de la elución de ADN.

Después del paso de limpieza es posible llevar a cabo un paso de amplificación de ADN libre de origen fetal por medio de PCR, Reacción en Cadena de la Polimerasa, donde los fragmentos de ADN unidos a los adaptadores, ahora aislados, se pueden someter a la llamada amplificación por PCR. El paso de amplificación mencionado anteriormente se lleva a cabo en dirección de 3' del paso de extracción del ADN libre de origen fetal y en dirección de 5' del paso de secuenciación de este ADN libre de origen fetal extraído.

Los términos "reacción en cadena de la polimerasa" o "PCR" significan, en la presente invención, las técnicas de biología molecular que permiten la multiplicación (amplificación) de fragmentos de ácidos nucleicos de los cuales se conocen las secuencias iniciales y finales de nucleótidos. La amplificación por PCR permite obtener muy rápidamente in vitro la cantidad de material genético necesario para aplicaciones posteriores.

En particular, los fragmentos de ADN obtenidos en el paso de limpieza se someten a amplificación por PCR con reactivos de PCR. Para generar la biblioteca de ADN amplificado, se utiliza una mezcla de PCR que contiene enzima de ADN polimerasa de alta fidelidad y cebadores de PCR diseñados para unirse a secuencias en oligonucleótidos adaptadores.

Las bibliotecas de ADN se limpian con los reactivos en esferas paramagnéticas. Las esferas paramagnéticas se adicionan a las bibliotecas de ADN para unir los fragmentos de ADN unidos a los adaptadores y amplificados y para aislarlos de la mezcla de reacción y los cebadores de PCR no utilizados.

5 El paso de seleccionar el tamaño del grupo de bibliotecas de ADN sometidas a multiplexación tiene lugar utilizando los reactivos en esferas paramagnéticas. Las bibliotecas de ADN se seleccionan en función del tamaño, tal que los fragmentos caen dentro del intervalo de longitud de lectura adecuado para la plataforma de secuenciación. En primer lugar, las esferas paramagnéticas se adicionan al conjunto de bibliotecas de ADN multiplexadas para unirse a cualquier fragmento indeseable más grande presente en la muestra. Las esferas a las que se unen los fragmentos de ADN se separan de la solución por medio de un imán y se mantiene el sobrenadante que contiene los fragmentos de biblioteca de ADN del tamaño deseado. Las esferas paramagnéticas se adicionan entonces al sobrenadante para unirse a los fragmentos restantes de las bibliotecas de ADN seleccionadas con base en su tamaño.

15 Las esferas se separan de la solución usando un imán y se lavan con agua libre de nucleasas antes de la elución de ADN. El grupo de bibliotecas de ADN sometidas a multiplexación y seleccionadas con base en el tamaño se cuantifica y se diluye para obtener la concentración adecuada para la reacción de secuenciación.

20 La muestra preparada entonces se coloca en una plataforma de secuenciación. La mezcla enzimática contiene las enzimas ADN ligasa y ADN polimerasa. Esta secuenciación se lleva a cabo preferentemente por técnicas de secuenciación de próxima generación NGS, pero también se puede lograr por técnicas de secuenciación individual.

25 Los términos "Secuenciación de Próxima Generación" o "NGS" o "Secuenciación Paralela" se refieren, en la presente invención, a las tecnologías que permiten secuenciar genomas grandes en un corto tiempo, preferentemente en el orden de semanas.

30 Por lo tanto, la secuenciación con tecnología de multiplexación de NGS se realiza de acuerdo con el protocolo para la plataforma de secuenciación pertinente con tecnología de NGS. Después de la secuenciación de las bibliotecas de ADN multiplexadas, los datos obtenidos se procesan por un software adecuado que calcula la probabilidad de que el feto pueda verse afectado o no por una de las trisomías examinadas, la aneuploidía fetal o para el estudio de microeliminaciones o síndromes de microduplicación.

35 El análisis prenatal no invasivo (también conocido como NIPT, pruebas prenatales no invasivas) utiliza ADNfcl (ADN libre de células) extraído de la sangre periférica materna para investigar trastornos genéticos en el feto. En contraste a los métodos invasivos, tal como la amniocentesis, el sistema de NIPT no presenta riesgos de aborto espontáneo.

40 Para realizar el análisis de NIPT, se extrae ADNfcl del plasma materno y se secuencia utilizando un sistema de NGS (Secuenciación de próxima generación).

45 Por lo tanto, sigue el paso de almacenar la secuenciación realizada en formato de FASTQ, que comprende las secuencias biológicas y las puntuaciones de calidad de las propias secuencias biológicas. En particular, después de la secuenciación, las máquinas retornan un archivo, llamado FASTQ, que contiene el almacenamiento de todas las secuencias obtenidas y las respectivas puntuaciones de calidad.

50 El término "formato de FASTQ" significa, en la presente invención, un formato basado en texto para almacenar tanto una secuencia biológica (usualmente la secuencia de nucleótidos) como las puntuaciones de calidad relativas. Tanto la letra de secuencia como la puntuación de calidad se codifican con un carácter ASCII individual para brevedad.

55 Por lo tanto, el archivo FASTQ contiene información sobre la calidad de la secuenciación, en particular, contiene información sobre la puntuación de calidad de Phred o la medición de la calidad de la identificación de nucleobases generadas automáticamente por la secuenciación de ADN. La puntuación de calidad de Phred se calcula como: $Q = -10\log(P)$.

En el siguiente paso, se realiza la conversión de secuenciación del formato FASTQ al formato SAM, en el que las secuencias biológicas se alinean con un genoma de referencia.

60 Los términos "Mapa de Alineación de Secuencia" o "SAM" significan, en la presente invención, un formato basado en texto para almacenar secuencias biológicas alineadas a una secuencia de referencia. Este formato se utiliza para almacenar datos, tal como secuencias de nucleótidos, generados por las tecnologías de NGS. El formato SAM se utiliza para contener datos mapeados dentro de las herramientas de análisis de genoma (GATK) y consta de un encabezado y una sección de alineación.

65 De acuerdo con el paso de convertir la secuenciación del formato FASTQ al formato SAM, en el que las secuencias biológicas se alinean con un genoma de referencia, se implementa utilizando el algoritmo Burrows-Wheeler.

La aplicación del algoritmo mencionado anteriormente se puede implementar, por ejemplo, por medio del paquete de software Bowtie2, preferentemente en una plataforma Linux/Unix. Sin embargo, los paquetes y la plataforma mencionados anteriormente no son vinculantes. Bowtie2 es una herramienta ultrarrápida y eficiente para alinear lecturas (o fragmentos de secuenciación producidos por sistemas de NGS) a un genoma de referencia, como se describe en más detalle más adelante. Bowtie2 indexa un genoma utilizando el algoritmo Burrows-Wheeler, que es el estándar de oro para las alineaciones de secuencias de NGS.

Para construir la referencia, debido a la cual se alinearán las lecturas, será necesario tener el genoma de referencia, es decir, la secuencia de ensamblaje para cada cromosoma en un archivo individual, para permitir que todos los archivos de ensamblaje se condensen en un archivo individual, o el genoma de referencia que se utilizará de acuerdo con la presente invención.

El término "LECTURA" significa, en la presente invención, una pequeña secuencia de ADN de síntesis obtenida de una reacción de secuenciación. En primer lugar, el ADN del paciente se rompe en muchos fragmentos pequeños, a cada uno de los cuales se adicionan secuencias nucleotídicas cortas (los llamados adaptadores) que sirven para anclar los fragmentos de ADN al soporte en el que tendrá lugar la reacción de secuenciación. El conjunto de fragmentos de ADN preparados por la adición de adaptadores constituye la biblioteca de secuenciación. Los fragmentos de la biblioteca de secuenciación se amplifican y secuencian para obtener numerosas copias complementarias de cada fragmento, definidas como lecturas.

Por lo tanto, se crea el índice de referencia real para la alineación, con la producción de seis archivos que constituirán el índice al que se referirá el software para la aplicación del método de acuerdo con la presente invención. Después de la construcción del índice, se alinearán las lecturas contenidas en los archivos FASTQ producidos por la secuenciación NGS de muestras de ADN extraídas del plasma materno.

La alineación es el proceso por el cual se descubre cómo y dónde las secuencias de lectura son similares a la secuencia de referencia, es decir, alineando algunos o todos los caracteres presentes en las lecturas con los de referencia.

En un paso de conversión adicional, la secuenciación se convierte del formato SAM al formato BAM, en el que los datos contenidos en el formato SAM se comprimen por conversión en código binario.

Los términos "Mapa de Alineación Binario" o "BAM" significan, en la presente invención, la representación binaria comprimida y sin pérdidas del Mapa de Alineación de Secuencia.

La conversión del formato SAM al formato BAM es esencial para mejorar el rendimiento de todos los demás paquetes de software que se utilizarán para el método de acuerdo con la presente invención.

El método de acuerdo con la presente invención puede comprender además el paso de definir un formato de índice BAI a partir de este formato BAM comprimido, en donde el formato de índice BAI es adecuado para la ejecución de cada paquete de software que funciona en un formato BAM comprimido.

En este punto, es posible llevar a cabo los pasos del método para la ejecución real del análisis de NIPT.

Con el método de acuerdo con la presente invención es posible llevar a cabo el análisis anterior en los cromosomas 13, 18 y 21 para la determinación de trisomías de los cromosomas mencionados anteriormente y/o para la determinación del sexo fetal y la aneuploidía de los cromosomas sexuales X, Y y/o para la determinación de las trisomías de estos autosomas con relación a uno o más de los cromosomas 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 17, 19, 20, 22. Estas pruebas se pueden llevar a cabo, por ejemplo, en el paso de cálculo de puntuación z, como se describe en mayor detalle a continuación, en un cromosoma de interés con respecto a la muestra de referencia predeterminada.

Además, con el método de acuerdo con la presente invención es posible llevar a cabo el análisis mencionado anteriormente para la determinación de uno o más del síndrome de Wolf-Hirschhorn, síndrome de Jacobsen, síndrome de eliminación de 1p36, síndrome de Angelman, síndrome de DiGeorge, síndrome de Cri-du-chat, síndrome de Langer-Giedion, síndrome de Smith-Magenis, síndrome de Prader-Willi, síndrome de Williams, síndrome de Koolen-de-Vries, síndrome de HNPP, síndrome de eliminación de 18q, síndrome de WAGR, síndrome de Potocki-Shaffer, síndrome de eliminación de 1q21.1, síndrome de Kleefstra, síndrome de Phelan-Mcdermid, síndrome de Rubinstein-Taybi, síndrome de Alagille y síndrome de Miller-Dieker. Estas pruebas pueden, por ejemplo, llevarse a cabo en el paso de determinar la probabilidad de riesgo de microeliminación o microduplicación, como se describe en mayor detalle a continuación, en las regiones genómicas MOSAICOS que componen el genoma.

La invención proporciona un método adicional para determinar la probabilidad de riesgo de trastornos cromosómicos y genéticos a partir del ADN libre de origen fetal, donde la determinación de la probabilidad de riesgo de:

1. Fibrosis quística (GEN CFTR)
2. Sordera congénita (GEN GJB2)
3. Talasemia Beta (GEN HBB)
4. Hiperplasia suprarrenal congénita clásica debido a deficiencia de 21-hidroxilasa (GEN CYP21A2)
- 5 5. Hemocromatosis (GEN HFE)
6. Acondroplasia (GEN FGFR3)
7. Hipocondroplasia (GEN FGFR3)
8. Displasia disatófora (GEN FGFR3)
9. Síndrome de Apert (GEN FGFR2)
- 10 10. Síndrome de Crouzon (GEN FGFR2)
11. Síndrome de Leopard (GEN PTPN11)
12. Síndrome de Noonan (GENES PTPN11/SOS1/RAF1)
13. Fenilcetonuria (GEN PAH)
14. Síndrome de Rett (GEN MECP2)
- 15 15. Se obtiene riñón poliquistico recesivo (GEN PKHD1).

La invención también proporciona el estudio del riesgo de parto prematuro, que ahora es la causa más frecuente de mortalidad y morbilidad fetal neonatal. Es un estudio que se realiza por el examen de una serie de polimorfismos de susceptibilidad genética (SNP) (ver solicitud de patente núm. 102017000134611 del 23/11/2017).

La invención también proporciona el estudio, en la sangre materna, en los primeros tres meses, de la infección citomegálica y toxoplásmica. La infección citomegálica materna tiende a ser asintomática y rara vez se diagnostica a los pacientes solo con síntomas clínicos. Para la mayoría de las infecciones, la evidencia de seroconversión materna (definida como una conversión de IgM negativa a positiva o un incremento de 4 veces en el título de anticuerpos IgG durante 4-6 semanas) es suficiente para confirmar el diagnóstico de una infección primaria. Sin embargo, la precisión de la IgM materna anti-CMV para predecir la infección materna primaria se complica por el hecho de que los anticuerpos IgM pueden persistir durante meses o incluso años después de la infección primaria, y también se pueden encontrar en el entorno de reactivación o reinfección con una cepa diferente de CMV.12.

Otro método para determinar el momento de la infección materna por CMV es medir la avidéz del anticuerpo, que se refiere a la fuerza de la unión del anticuerpo a un antígeno diana. En tanto que la respuesta inmune a un antígeno particular madura con el tiempo, la avidéz incrementa. Por lo tanto, la detección de IgG anti-CMV de baja avidéz al principio del embarazo sugiere una infección aguda reciente y se puede utilizar para identificar a las mujeres embarazadas con riesgo incrementado de tener un feto infectado. En contraste, la presencia de anticuerpos de alta avidéz a las 12-16 semanas de gestación indica una infección pasada, probablemente antes de la concepción.

Por lo tanto, el mejor método es la búsqueda directa del virus en la sangre materna (posiblemente completada con investigación en la orina). Es por esto que el tamizaje total Fetal-materno incluye este examen fundamental para la gestión del embarazo.

Para la infección toxoplásmica materna, conocer la presencia de una infección en el compartimento materno de forma temprana permite establecer una terapia adecuada con una justificación terapéutica verdadera y absoluta.

Todas estas pruebas se pueden realizar al implementar el método de acuerdo con la presente invención con paquetes de software adecuados preferiblemente implementados en software estadístico. Algunos factores influyen en el rendimiento del análisis, por ejemplo, la fracción fetal, o la proporción entre las lecturas obtenidas de la secuenciación de ADNflc con respecto al ADN materno, la longitud de las lecturas y la cobertura, es decir, la cobertura de las lecturas con respecto al genoma de referencia.

Entonces, sigue el paso de contar y categorizar la LECTURA secuenciada del formato SAM, en el que cada LECTURA se categoriza como mapeada en cada cromosoma o mapeada en todos los autosomas.

Se pueden utilizar paquetes de software apropiados para analizar archivos BAM, contar lecturas y categorizarlos. Las lecturas no procesadas de las tecnologías de secuenciación de alto rendimiento en general no siguen las estadísticas de conteo teóricas. El exceso de varianza, producido en parte por los pasos de PCR, es con relación al contenido de GC de la secuencia de ADN. Este exceso de varianza se normaliza por el método de valor cromosómico normalizado de NCV y por la ponderación de conteo por categoría del contenido de GC. La NCV calcula la relación entre los recuentos mapeados en el cromosoma de interés y otro cromosoma, en tanto que para la ponderación de los conteos por categoría en GC, el conteo de las lecturas primero se determina para cada categoría, entonces el conteo promedio se calcula en función del contenido en GC. También es posible, a través de este análisis, enmascarar las áreas repetidas y de baja complejidad del genoma, a fin de tener en cuenta sólo las lecturas mapeadas de manera única.

Posteriormente, se lleva a cabo el paso de cálculo de la relación entre el número de LECTURAS mapeadas en cada cromosoma y el número de LECTURAS mapeadas en todos los autosomas. Para cada muestra, se calcula

la relación de las lecturas mapeadas en cada cromosoma con respecto a las lecturas mapeadas en todos los autosomas.

5 Después de la remoción de los artefactos de las lecturas, en el paso de definir una muestra de referencia predeterminada, se crea una referencia utilizando una serie de muestras de referencia con resultados conocidos y sexo fetal.

10 Entonces, en el paso de cálculo de la puntuación z en un cromosoma de interés con respecto a la muestra de referencia predeterminada. Por ejemplo, se puede calcular la puntuación z en los cromosomas 13, 18 y 21.

Finalmente, en un paso adicional, la probabilidad de riesgo de anomalía en el cromosoma de interés se determina si la puntuación z calculada es mayor que un valor predeterminado. Normalmente, en el caso de puntuaciones z con un valor mayor de tres, se declara una anomalía cromosómica en los cromosomas analizados.

15 Para la determinación del sexo fetal y la monosomía X, en su lugar, se calculará el puntuación z para el cromosoma X y para el cromosoma Y y, como se describe anteriormente, si el valor de puntuación z calculado es mayor que 3 para el cromosoma Y, la muestra se declarará sexo fetal masculino, en tanto que si es mayor que -2 para el cromosoma X se declarará sexo fetal femenino.

20 Finalmente, el método de acuerdo con la presente invención comprende además el paso de verificar el ADN libre de origen fetal por PCR digital si en el paso de determinar la probabilidad de riesgo de anomalía en el cromosoma de interés el valor predeterminado es excedido por esta puntuación z.

25 La reacción en cadena de polimerasa digital (PCR digital, dPCR) es una mejora biotecnológica de los métodos convencionales de reacción en cadena de polimerasa que se pueden utilizar para cuantificar y amplificar directamente los filamentos de ácido nucleico. La dPCR mide cuantitativamente una cierta variable y también realiza una reacción individual dentro de una muestra que se separa en una gran cantidad de particiones y la reacción se realiza individualmente en cada partición. Esta separación permite una recolección más fiable y una medición sensible de las cantidades de ácido nucleico. Esta precisión hace que la verificación de una prueba positiva por NGS sea aún más fiable, incrementando la sensibilidad del método en cuestión y disminuyendo la proporción de falsos positivos.

35 Por lo tanto, PCR digital es una aplicación en la detección y análisis cuantitativo de pequeñas unidades de ADN, proporcionando un análisis cuantitativo altamente sensible y la detección de moléculas de ácido nucleico diana por el procesamiento simultáneo de un número diferente de muestras. Junto con la secuenciación NGS, se utilizó PCR digital para mejorar el resultado de una prueba NIPT y reducir significativamente el porcentaje de falsos positivos.

40 El procedimiento para la búsqueda de microeliminaciones y microduplicaciones es extremadamente similar a ese ilustrado para la búsqueda de anomalías cromosómicas, de acuerdo con lo que se describe en mayor detalle a continuación.

45 Se lleva a cabo un paso de definición de un conjunto de regiones genómicas MOSAICOS que componen el genoma en formato BAM, en el que los datos se comprimen por conversión a código binario. El uso de azulejos permite definir categorías que tienen anchos variables y, por lo tanto, permiten una mayor flexibilidad de la función.

50 En un paso posterior de definición de una región genómica de referencia en formato BAM, los datos se comprimen por conversión en código binario, en el que los parámetros para la realización del formato BAM de la región genómica de referencia son los mismos que el formato BAM de las regiones genómicas MOSAICOS, tal que se elimine cualquier diferencia entre los dos archivos que implicaría la presencia de numerosos casos de falsos positivos.

Por lo tanto, se lleva a cabo un paso de comparación del formato BAM de las regiones genómicas MOSAICOS con el formato BAM de la región genómica de referencia.

55 Finalmente, se realiza un paso de determinación de la probabilidad de microeliminación o riesgo de microduplicación en las regiones genómicas MOSAICOS que componen el genoma.

60 Una aplicación adicional del método de acuerdo con la presente invención se refiere al estudio del gen CFTR en el ADN genómico gestacional para la examinación de fibrosis quística. La Fibrosis Quística (FQ) es una enfermedad de transmisión recesiva autosómica multisistémica genética en la que el transporte defectuoso de cloro a través de las membranas celulares provoca secreciones deshidratadas y viscosas con el consiguiente espesamiento de moco en los bronquios, espesamiento del jugo pancreático y altos niveles de cloro en el sudor. Los órganos frecuentemente afectados son el hígado, el intestino, el sistema reproductor y los pulmones, donde el moco particularmente denso conduce a problemas respiratorios graves y las consiguientes infecciones. Los niños con Fibrosis Quística tienen dos padres que son al menos portadores sanos (o heterocigotos) de una mutación genética y, en particular, entre dos portadores sanos existe una probabilidad del 25% de que el feto pueda verse afectado

por la enfermedad. Hasta la fecha, se reconocen más de 1500 mutaciones del gen de la fibrosis quística (CFTR), muchas de ellas son extremadamente raras y otras aún se desconocen. Con el método de acuerdo con la presente invención es posible realizar el análisis del gen materno a través de una examinación llamada 1er nivel que permite analizar las mutaciones más comunes y frecuentes logrando identificar aproximadamente el 83% de los portadores. La frecuencia estimada, en la población italiana, de portadores sanos (a menudo sin estar consciente de serlo) es de 1 en 25 - 30, la de los nacidos afectados es de 1 en 2500 - 3000.

Por lo tanto, el método comprende los pasos de extraer el ADN materno, la amplificación específica del gen CFTR por medio de UN KIT de diagnóstico que permite la detección simultánea de mutaciones del gen que codifica para la proteína CFTR utilizando la metodología de transferencia puntual y la secuenciación automática de fragmentos como se describe anteriormente. La prueba permite la identificación de las principales variantes de la proteína de CFTR (Regulador de la Conductancia Transmembrana de la Fibrosis Quística) utilizando la tecnología de transferencia en mancha a partir de ADN extraído de sangre completa y mancha y el posterior análisis de fragmentos en un secuenciador automático. En cada tubo se realiza una amplificación específica y simultánea de todas las mutaciones buscadas. Esta compleja metodología requiere una aplicación rigurosa del protocolo.

Mutaciones analizadas: 711+1G-T, 621+1G-T, 1717-1G-A, 3849+10kbC-T, 2789+5G-A, G542X, G85E, G551D, R553X, N1303K, R117H, R1162X, L1077P, L1065P, W1282X, R347P, I507del, T338I, F508del, 1677delTA, 2183AA-G, S549R, Q552X, 852del22, R1066H, G1244E, 1259insA, D1152H, 711+5G-A, R1158X, 4382delA, 4016insT, A455E, 1706del117, 1502T, 3199del6, S912X.

En este caso, en el paso de determinación del riesgo es posible atribuir una probabilidad de alto riesgo de que el sujeto sea heterocigoto o portador sano, cuando se encuentra una mutación entre las analizadas en el gen materno de CFTR, o negativa, cuando ninguna de las mutaciones entre las analizadas se encuentra en el gen materno de CFTR.

Una aplicación adicional del método de acuerdo con la presente invención se refiere al estudio del gen SMN1 en el ADN genómico gestacional para la examinación de la Atrofia Muscular Espinal. La atrofia muscular espinal (SMA) es un grupo de enfermedades genéticas que causan debilidad y deterioro en los músculos voluntarios de bebés y niños y, más raramente, en adultos. Es una de las afecciones genéticas más comunes que afectan a los niños. Se estima que uno de cada 6.000 niños en todo el mundo nació con SMA.

En más del 95 por ciento de los casos, la SMA se provoca por la producción inadecuada de una proteína llamada proteína de supervivencia de la neurona motora (SMN) esencial para las neuronas motoras. SMN se produce por SMN1 y en menor medida por SMN2. Estos genes están en el cromosoma 5. Habitualmente, las personas tienen dos copias del gen SMN1 y hasta dos copias del gen SMN2 en cada una de sus células de neuronas motoras. En las personas con atrofia muscular espinal, ambas copias del gen SMN1 están alteradas o faltan. Tener copias adicionales (tres o más) del gen SMN2 se asocia con una enfermedad más leve que compensa parcialmente la falta de SMN1. En raras ocasiones, la SMA se provoca por mutaciones en genes distintos de SMN (no en el cromosoma 5).

Las neuronas motoras se ubican en el cuerno anterior de la médula espinal y controlan directamente los músculos esqueléticos del cuerpo. Sin una proteína de SMN adecuada, las neuronas motoras de la médula espinal comienzan a contraerse y morir. Cuando esto sucede, el cerebro del niño no puede controlar los músculos voluntarios del cuerpo, especialmente los de los brazos y las piernas y los de la cabeza y el cuello. Los músculos comienzan a debilitarse y dispersarse. Esto afecta a movimientos como caminar, gatear, revisar la cabeza y el cuello, tragar y respirar.

Hay cuatro tipos diferentes de atrofia muscular espinal. La clasificación está determinada por las etapas de desarrollo que el niño ha alcanzado en el momento de la aparición de la enfermedad. Los tipos I y II son los más comunes.

Los tipos de atrofia muscular espinal (SMA) son:

SMA tipo 1 (grave): este tipo también se denomina enfermedad de Werdnig-Hoffmann. Es el tipo más grave y común de SMA. Suele ser evidente al nacer o en los primeros meses (0-6 meses). Los síntomas incluyen extremidades flexibles y movimiento débil del tronco. Los niños con este tipo usualmente tienen una capacidad muy limitada para moverse. También tendrán dificultades para comer y tragar, mantener la cabeza alta y respirar. La SMA tipo 1 progresa rápidamente, con un debilitamiento de los músculos que conduce a infecciones respiratorias frecuentes y, usualmente, a la muerte a los 2 años. Los bebés con SMA tipo 1 nunca pueden sentarse. SMA tipo 2 (intermedia): los síntomas suelen aparecer entre los 7 y los 18 meses. La tasa de progresión puede variar en gran medida. La enfermedad afecta a las piernas del niño más que a sus brazos. Los niños con SMA tipo 2 nunca pueden pararse. Las infecciones respiratorias también son comunes con este tipo de SMA. La esperanza de vida puede variar desde la primera infancia hasta la edad adulta, dependiendo de la gravedad de la condición del paciente. SMA tipo 3 (leve): este tipo de SMA también se llama Kugelberg-Welander o atrofia muscular espinal juvenil. Los

- 5 síntomas pueden aparecer por primera vez en un amplio intervalo de años, desde los 18 meses hasta el comienzo de la edad adulta. Los pacientes con SMA tipo 3 pueden ponerse de pie y caminar, pero pueden tener dificultades para levantarse de una posición sentada. También pueden tener debilidad muscular leve y tienen un mayor riesgo de infecciones respiratorias. La mayoría de los pacientes con SMA tipo 3 tienen una esperanza de vida cercana a lo normal.
- SMA tipo 4 (adulto): los síntomas de este tipo raro de SMA no suelen surgir hasta la segunda o tercera década de vida. Los pacientes con SMA tipo 4 pueden caminar durante la edad adulta, pero usualmente experimentan debilidad muscular progresivamente lenta y otros síntomas habituales de la SMA.
- 10 Una aplicación adicional del método de acuerdo con la presente invención se refiere al estudio de los principales genes asociados con la trombofilia hereditaria (Factor V de Leiden, Factor II, MTHFR) y del riesgo de preeclampsia, en el ADN genómico de la mujer embarazada.
- 15 Sin embargo, otra aplicación del método de acuerdo con la presente invención se refiere al estudio de Rh fetal. La determinación prenatal del genotipo de Rh fetal podría conducir a una reducción sustancial en el uso de inmunoglobulina anti-D y la prevención de la exposición innecesaria de las mujeres embarazadas portadoras de un feto RhD negativo. El grupo sanguíneo Rh es uno de los grupos sanguíneos más complejos y polimórficos conocidos en humanos. Los genes homólogos RHD y RHCE codifican para dos proteínas Rh distintas que se encuentran en el cromosoma 1. Estos genes tienen cada uno 10 exones y 9 intrones y comparten la homología al 20 96,8% a nivel de nucleótido.
- Aproximadamente el 50% de los casos de aloinmunización materna provienen de anticuerpos maternos anti-D (IgG) que atraviesan la placenta y se unen a glóbulos rojos fetales RHD positivos que conducen a su destrucción después de la anemia y en algunos casos contribuyen a la hidropesía fetal y la muerte fetal intrauterina o la muerte neonatal.
- 25 Al determinar el genotipo Rh fetal, se puede administrar inmunoglobulina anti-D intencionalmente a mujeres embarazadas RhD Rh negativo que portan Rh positivo del feto. Otra ventaja de la genotipificación de RhD fetal se refiere a las mujeres que están inmunizadas contra el antígeno D por diversas razones y que tienen anti-D en suero.
- 30 Por lo tanto, el método de acuerdo con la presente invención permite la genotipificación de RHD fetal a partir de ADN fetal libre de células (ADNflc) en el plasma de mujeres embarazadas. La genotipificación de RhD fetal se realiza utilizando la técnica de PCR en tiempo real con alta precisión, especificidad y sensibilidad.
- 35 El método de examen prenatal objeto de la presente descripción, al ser una prueba con base en la aplicación de análisis moleculares combinados, Secuenciación de Próxima Generación, PCR Digital, PCR en Tiempo Real, ARMS, Transferencia en mancha, reduce drásticamente el porcentaje de falsos positivos (<0,1%), e incrementa el poder diagnóstico de la prueba NIPT.
- 40 El método de acuerdo con la presente invención permite el estudio de las anomalías cromosómicas mencionadas anteriormente por el análisis de muestras de sangre materna y la detección de aneuploidías fetales por medio de secuenciación de próxima generación (NGS, Secuenciación de Próxima Generación), aplicando los pasos descritos para el método y confirmando los resultados positivos con PCR digital. La prueba, que se ofrece a las mujeres embarazadas que han pasado la décima semana de gestación, les permite limitar el uso de exámenes prenatales invasivos (amniocentesis y villocentesis) recurriendo a estas maniobras cuando sea necesario. A 45 petición del paciente, es posible buscar HR fetal por PCR en Tiempo Real, en tanto que con el ARMS y la técnica de transferencia en mancha se analizan las mutaciones principales del gen CFTR asociadas a Fibrosis Quística.

REIVINDICACIONES

1. Método para determinar la probabilidad de riesgo de trastornos cromosómicos y genéticos a partir del ADN libre de origen fetal que comprende los siguientes pasos:

5 - extraer este ADN libre de origen fetal de una muestra de plasma materno previamente recolectada;
 - secuenciar este ADN libre de origen fetal extraído;
 - obtener los datos de secuenciación en formato FASTQ que comprenden las secuencias biológicas y las puntuaciones de calidad de estas secuencias biológicas;
 10 este método se caracteriza porque comprende los siguientes pasos:

- convertir esta secuenciación de este formato FASTQ al formato SAM, en el que estas secuencias biológicas se alinean con un genoma de referencia;
 - convertir esta secuenciación del formato SAM al formato BAM, en el que los datos contenidos en este formato SAM se comprimen por conversión en código binario;
 15 - contar y categorizar la LECTURA secuenciada de este formato SAM, en el que cada LECTURA se categoriza como mapeada en cada cromosoma o mapeada en todos los autosomas;
 - definir una muestra de referencia predeterminada;
 - calcular la puntuación z en un cromosoma de interés con respecto a esta muestra de referencia predeterminada;
 20 determinar la probabilidad de riesgo de anomalía en este cromosoma de interés si esta puntuación z es más alta que un valor predeterminado.

2. Método para determinar la probabilidad de riesgo de trastornos cromosómicos y genéticos a partir del ADN libre de origen fetal de acuerdo con la reivindicación 1, donde el método comprende además los pasos de:

- definir un conjunto de regiones genómicas MOSAICOS que componen el genoma en formato BAM, en el que los datos se comprimen por conversión en código binario;
 - definir una región genómica de referencia en formato BAM, en la que los datos se comprimen por conversión en código binario, en el que los parámetros para implementar este formato BAM de esta región genómica de referencia son los mismos que este formato BAM de estas regiones genómicas MOSAICOS;
 30 - comparar este formato BAM de estas regiones genómicas MOSAICOS con este formato BAM de esta región genómica de referencia;
 - determinar la probabilidad de microeliminación o riesgo de microduplicación en estas regiones genómicas MOSAICOS que componen el genoma.
 35

3. Método para determinar la probabilidad de riesgo de trastornos cromosómicos y genéticos a partir del ADN libre de origen fetal de acuerdo con la reivindicación 1 o 2, donde el método comprende además el paso de definir un formato de índice BAI a partir del formato BAM comprimido, en el que el formato de índice BAI se adapta a la ejecución de cada paquete de software que funciona en este formato BAM comprimido.
 40

4. Método para determinar la probabilidad de riesgo de trastornos cromosómicos y genéticos a partir del ADN libre de origen fetal de acuerdo con una o más de las reivindicaciones 1 a 3, donde este paso de convertir esta secuenciación de este formato FASTQ al formato SAM, en el que estas secuencias biológicas están alineadas con un genoma de referencia, se implementa por el algoritmo Burrows-Wheeler.
 45

5. Método para determinar la probabilidad de riesgo de trastornos cromosómicos y genéticos a partir del ADN libre de origen fetal de acuerdo con una o más de las reivindicaciones 1 a 4, donde este paso de secuenciación de este ADN libre de origen fetal extraído se lleva a cabo por técnicas de Secuenciación de Próxima Generación NGS.
 50

6. Método para determinar la probabilidad de riesgo de trastornos cromosómicos y genéticos a partir del ADN libre de origen fetal de acuerdo con una o más de las reivindicaciones 1 a 5, donde este método comprende además el paso de amplificar este ADN libre de origen fetal por PCR de reacción en cadena de la polimerasa, en el que este paso de amplificación se lleva a cabo corriente abajo de este paso de extracción de este ADN libre de origen fetal y corriente arriba de este paso de secuenciación de este ADN libre de origen fetal extraído.
 55

7. Método para determinar la probabilidad de riesgo de trastornos cromosómicos y genéticos a partir del ADN libre de origen fetal de acuerdo con una o más de las reivindicaciones 1 a 6, donde este método comprende además el paso de verificar este ADN libre de origen fetal por PCR digital si en este paso de determinar la probabilidad de riesgo de anomalía en este cromosoma de interés, este valor predeterminado es excedido por esta puntuación z.
 60

8. Método para determinar la probabilidad de riesgo de trastornos cromosómicos y genéticos a partir de ADN libre de origen fetal de acuerdo con una o más de las reivindicaciones 1 a 7, donde este paso de calcular esta puntuación z en un cromosoma de interés con respecto a esta muestra de referencia predeterminada se lleva a cabo en los cromosomas 13, 18 y 21 para la determinación de trisomías de los cromosomas mencionados anteriormente y/o para la determinación del sexo fetal y aneuploidía de los cromosomas sexuales X, Y y/o para la determinación de
 65

las trisomías de estos autosomas con relación a uno o más de los cromosomas 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 17, 19, 20, 22.

5 9. Método para determinar la probabilidad de riesgo de trastornos cromosómicos y genéticos a partir del ADN libre de origen fetal de acuerdo con una o más de las reivindicaciones 2 a 7, donde este paso de determinar la probabilidad de riesgo de microeliminación o microduplicación en estas regiones genómicas que componen el genoma es con relación a uno o más de Síndrome de Wolf-Hirschhorn, Síndrome de Jacobsen, Síndrome de eliminación de 1p36, Síndrome de Angelman, Síndrome de DiGeorge, Síndrome de Cri-du-chat, Síndrome de Langer-Giedion, Síndrome de Smith-Magenis, Síndrome de Prader-Willi, Síndrome de Williams, Síndrome de Koolen-de-Vries, Síndrome de HNPP, Síndrome de eliminación de 18q, Síndrome de WAGR, Síndrome de Potocki-Shaffer, Síndrome de eliminación de 1q21.1, Kleeftstra, Síndrome de Phelan-Mcdermid, Síndrome de Rubinstein-Taybi, Síndrome de Alagille y Síndrome de Miller-Dieker.

15 10. Método para determinar la probabilidad de riesgo de trastornos cromosómicos y genéticos a partir del ADN libre de origen fetal de acuerdo con una o más de las reivindicaciones 2 a 7, donde este paso se obtiene para determinar la probabilidad de riesgo de

1. Fibrosis quística (GEN CFTR)
2. Sordera congénita (GEN GJB2)
- 20 3. Talasemia Beta (GEN HBB)
4. Hiperplasia suprarrenal congénita clásica debido a deficiencia de 21-hidroxilasa (GEN CYP21A2)
5. Hemocromatosis (GEN HFE)
6. Acondroplasia (GEN FGFR3)
7. Hipocondroplasia (GEN FGFR3)
- 25 8. Displasia disatófora (GEN FGFR3)
9. Síndrome de Apert (GEN FGFR2)
10. Síndrome de Crouzon (GEN FGFR2)
11. Síndrome de Leopard (GEN PTPN11)
12. Síndrome de Noonan (GENES PTPN11/SOS1/RAF1)
- 30 13. Fenilcetonuria (GEN PAH)
14. Síndrome de Rett (GEN MECP2)
15. Riñón poliquístico recesivo (GEN PKHD1).