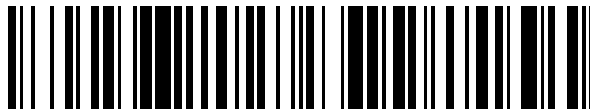


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 641 690**

21 Número de solicitud: 201630599

51 Int. Cl.:

C12Q 1/68 (2006.01)

C12N 15/10 (2006.01)

G01N 33/50 (2006.01)

12

SOLICITUD DE PATENTE

A1

22 Fecha de presentación:

09.05.2016

43 Fecha de publicación de la solicitud:

13.11.2017

71 Solicitantes:

**HEALTH IN CODE, S.L. (100.0%)
Edf. El Fortin, Hospital Marítimo de Oza, s/n
15006 AS XUBIAS (A Coruña) ES**

72 Inventor/es:

**GAYOSO BABÍO, Carmen María;
MARTÍNEZ DE ILÁRDUYA RUIZ DE
LARRAMENDI, Óskar y
LESENDE RODRÍGUEZ, Iván Aarón**

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

54 Título: **Método de identificación de mutaciones**

57 Resumen:

Método de identificación de mutaciones.

La presente invención se refiere a un método para la identificación de la posición de una mutación genética y al uso de dicho método para simplificar el cribado de dicha mutación genética.

ES 2 641 690 A1

DESCRIPCIÓN

Método de identificación de mutaciones

5 **Campo de la invención**

La presente invención pertenece al campo de las técnicas de biología molecular. En concreto, se refiere a un método para la identificación de la posición de una mutación genética y al uso de dicho método para simplificar el cribado de dicha mutación genética. La presente invención encuentra aplicación en aquellos campos que requieran el conocimiento preciso de una secuencia genética o del punto exacto de una mutación genética (clínica, 10 agronomía, biotecnología, etc.).

Antecedentes de la invención

15 La determinación de secuencias adyacentes a una región conocida del cromosoma es una tarea técnicamente complicada, y se han desarrollado diferentes metodologías para conseguirlo. Entre las técnicas descritas figuran "*ligation mediated PCR*" (LM-PCR o "*genome walking*"), PCR inversa (i-PCR), "*thermal asymmetric interlaced PCR*" (TAIL-PCR), 20 "*anchored PCR*" (a-PCR) o "*randomly primed PCR*" (rm-PCR). Todos estos métodos sufren de baja sensibilidad de detección o baja especificidad y además sólo son eficaces cuando el punto en el que ocurre la mutación en el genoma queda como máximo a unos pocos cientos de pares de bases de la secuencia conocida. Más recientemente se han desarrollado otras metodologías como "*Linear Amplification Mediated Polymerase Chain Reaction*" (LAM PCR), 25 que precisa la generación de un fragmento de doble hebra de DNA y la digestión de este fragmento. Posteriores modificaciones de esta técnica obvian la necesidad de esta digestión, que sustituyen por la digestión inicial del ADN genómico y la ligación de un adaptador de doble hebra. Todos estos métodos se basan en la amplificación exponencial del ADN.

30 Existen diferentes técnicas bien establecidas para la identificación de regiones del genoma que presentan pérdida o ganancia de material genético. Entre ellas se incluyen técnicas basadas en PCR (*multiplex ligation-dependent probe amplification*, MLPA) o en hibridación (*comparative genomic hybridization array* (CGH array), polimorfismo de nucleótido único, del inglés *single nucleotide polymorphism, array* (SNP arrays"), etc.). En función del diseño de 35 oligonucleótidos o sondas, estas técnicas permiten identificar regiones del genoma en las

que se produce pérdida o ganancia de material genético, aunque en ningún caso identifican los puntos exactos de comienzo o final de estas regiones ni tampoco, en el caso de inserciones, ofrecen ninguna información sobre la región del genoma en que se produce la inserción. En el caso de MLPA, esta técnica ofrece muy poca procesividad, dado que sólo es posible el análisis de un número limitado de regiones genéticas (normalmente exones). Además, es una técnica que ofrece problemas de aparición de falsos negativos.

En el caso de *SNP arrays* y *CGH array*, es posible el análisis del genoma completo de muestras sin conocimiento previo de la secuencia. Sin embargo, adolece de una capacidad de resolución relativamente baja ya que rara vez es capaz de detectar deleciones o inserciones de un tamaño menor a 50 Kb.

El uso de técnicas como MLPA (Stuppia et al., *Use of the MLPA assay in the molecular diagnosis of gene copy number alterations in human genetic diseases*. Int J Mol Sci. 2012; 13: 3245–76), *CGH array* (Lai et al., *Comparative analysis of algorithms for identifying amplifications and deletions in array CGH data*. Bioinformatics 2005; 21: 3763-70), *SNP arrays* (Zhang et al., *Evaluation of copy number variation detection for a SNP array platform*. BMC Bioinformatics 2014; 15: 50) o “*NGS targeted sequencing*” (Zhao et al., *Computational tools for copy number variation (CNV) detection using next-generation sequencing data: features and perspectives*. BMC Bioinformatics 2013; 14 (Suppl 11): S1) permite detectar la presencia de mutaciones por pérdida (deleciones) o ganancia (inserciones) de material genético. El resultado obtenido por cualquiera de estas técnicas delimita la región cromosómica mínima que con una certeza del 100% comprende la variante estructural, pero ninguna de ellas identifica los límites exactos de la variante estructural en cuestión. Aquella región de ácidos nucleicos que con una certeza del 100% comprende una mutación en cuestión, es referida de aquí en adelante como “región de ácidos nucleicos certera” o “región ANC”.

Existen además otras modificaciones genéticas que no implican cambio en la cantidad de material genético, como las translocaciones, y que por tanto no se detectan con los métodos mencionados. Para la detección de la región ANC de estas variantes estructurales se utilizan técnicas específicas, como *Southern blot*, análisis de cariotipo o hibridación fluorescente *in situ* (FISH) en translocaciones de grandes segmentos de cromosoma, o PCR a tiempo real (RT-PCR) o Δ -PCR para identificar otras translocaciones que pueden dar lugar a fusiones de genes.

En cuanto a las fusiones génicas, actualmente se considera que fusiones génicas, producidas por translocaciones cromosómicas, inversiones, deleciones, etc, tienen gran relevancia en cánceres epiteliales frecuentes, como los carcinomas de próstata o pulmón. Por ejemplo, la mayoría de los cánceres de próstata presentan una fusión regulada por andrógenos de alguno de los factores de transcripción de la familia de genes ETS. Clínicamente, algunos neoplasmas se clasifican o manejan de acuerdo con la presencia de un fusión génica en concreto: por ejemplo, leucemias promielocíticas que portan una fusión PML–RAR α del receptor α del ácido retinoico se tratan con ácido retinoico, mientras que la leucemias mieloides crónicas con presencia de la fusión BCR–ABL se tratan con el fármaco imatinib. Sin embargo, los ensayos que se llevan a cabo por RT-PCR exigen que ambos elementos fusionados sean conocidos y que a su vez den lugar a una variante previamente caracterizada, cuando en ocasiones únicamente se conoce uno de los genes que puede estar involucrado en la fusión génica. Dicho único gen conocido se correspondería por tanto con la región ANC.

La reciente aparición de tecnologías de ultrasecuenciación masiva (*“next generation sequencing”*, NGS) ha permitido enormes avances en el conocimiento de secuencias de ácidos nucleicos de multitud de organismos. A su vez, ha servido para poner de manifiesto fuentes de variabilidad genética que anteriormente pasaban con frecuencia desapercibidas; entre estas se incluyen variantes estructurales genómicas y variación en el número de copias de un gen (CNV).

Existen diferentes métodos de preparación de muestras para su secuenciación NGS. En una de sus formas, la metodología incluye un paso de “enriquecimiento de dianas” que consiste en la selección, por diferentes métodos (por ejemplo mediante captura con una solución de sondas específicas), de aquella región del genoma que interesa estudiar (por ejemplo para secuenciar un panel de genes concretos, para secuenciar sólo exomas, etc.), prescindiendo del resto del genoma. Estos métodos consiguen lecturas muy fiables de una región de interés pero virtualmente ninguna información del resto del genoma.

Otra opción para caracterizar este tipo de mutaciones consiste en la secuenciación del genoma completo (*“Whole Genome Sequence”*, WGS) de aquellas muestras en las que existe sospecha de modificaciones de este tipo. Si bien esta metodología ofrece resultados que cubren todo el genoma, existen zonas recalcitrantes en las que se localizan pocas o ninguna lectura. Además, la metodología WGS resulta cara y presenta problemas cuando la región de interés se localiza en regiones repetitivas del genoma, con lo que no garantiza la

correcta caracterización de variantes genéticas de este tipo.

Los autores de la presente invención han desarrollado un método que supera los problemas de las técnicas anteriores y permite una simplificación del método de identificación de mutaciones y cribado de las mismas.

Objeto de la invención

La presente invención se refiere en un primer aspecto a un método para identificar la posición de una mutación genética caracterizado por que comprende las siguientes etapas:

a) determinar la región ANC que comprende la mutación genética en estudio;
b) llevar a cabo una PCR, donde el molde de la PCR es ADN extraído de una muestra biológica tomada del sujeto en estudio y los cebadores son:

b1) un cebador degenerado, y

b2) un cebador específico que hibrida específicamente:

- dentro de la región ANC, a una distancia de 0,2-100 Kb de cualquiera de los extremos de la región ANC, si la mutación comprende una ganancia de material genético, una translocación o una fusión génica, y se extiende mediante la PCR hacia fuera de la región ANC, o

- fuera de la región ANC, a una distancia de 0,2-100 Kb de cualquiera de los extremos de la región ANC, si la mutación comprende una pérdida de material genético, y se extiende mediante la PCR hacia la región ANC,

y donde en la reacción de PCR la etapa de hibridación comprende dos hibridaciones consecutivas, una primera hibridación a una temperatura de hibridación igual a la T_m del cebador específico $\pm 10^\circ\text{C}$, y una segunda hibridación a una temperatura de hibridación igual a la T_m del cebador degenerado $\pm 5^\circ\text{C}$;

c) secuenciar el producto de amplificación obtenido en la etapa b); y

d) alinear la secuencia obtenida en la etapa c) con un genoma de referencia para identificar la posición de la mutación.

Un segundo aspecto de la presente invención se refiere a un método para cribar una mutación genética caracterizado porque comprende las siguientes etapas:

i) identificar la posición de la mutación genética siguiendo el método según el primer aspecto de la invención;

ii) llevar a cabo una PCR donde el ADN modelo es ADN extraído a partir de una muestra biológica tomada del sujeto en estudio y los cebadores son un cebador directo y un cebador

inverso que hibridan específicamente aguas arriba y aguas abajo, respectivamente, de la posición de la mutación identificada en la etapa i) de manera que se amplifica un producto de PCR que comprende dicha posición;

5 iii) analizar el tamaño del producto de PCR amplificado en ii), donde un producto de PCR del tamaño esperado según los cebadores directo e inverso utilizados en la etapa ii) es indicativo de la presencia de la mutación genética en estudio, y la ausencia de dicho producto de PCR es indicativa de ausencia de la mutación genética.

10 Un tercer aspecto de la presente invención se refiere a un método de diagnóstico de una patología asociada a una mutación genética que comprende llevar a cabo el método según el segundo aspecto de la invención, siendo el sujeto en estudio diagnosticado como afecto con la patología cuando el producto de PCR de la etapa iii) tiene el tamaño esperado según los cebadores directo e inverso utilizados en la etapa ii).

15 Un cuarto aspecto de la presente invención se refiere a un kit para llevar a cabo uno cualquiera de los métodos según el aspecto primero, segundo o tercero de la invención, que comprende un cebador degenerado b1), y un cebador específico b2), tal y como se han definido en el primer aspecto de la invención.

20 **Breve descripción de las figuras**

Figura 1: Esquema de la delección de los exones 59-66 del gen FBN1. DUT se refiere al gen deoxiuridina trifosfatasa, adyacente al gen FBN1 hacia el centrómero. Los números en cursiva representan el número de nucleótidos. Los números en negrita representan la posición en el cromosoma. El término “En^o” significa “exón n^o”, por ejemplo “E58” significa “exón 58”. La flecha FBN1-Ex58-FW denota el cebador específico, representando la cabeza de la flecha la dirección de extensión de dicho cebador.

Figura 2: Diagrama mostrando la predicción de la estructura secundaria a 30°C (panel B) del cebador específico de SEC ID N° 5. El panel A muestra la estructura a 60°C, donde la flecha marca la región de homología al intrón 45 del gen DMD.

Figura 3: Esquema de la duplicación de los exones 46 y 47 del gen DMD. Los números en cursiva representan el número de nucleótidos. Los números en negrita representan la posición en el cromosoma. El término “En^o” significa “exón n^o”, por ejemplo E46 significa Exón 46. La flecha DMDdup-Ex47-RV denota el cebador específico, representando la cabeza de la flecha la dirección de extensión de dicho cebador.

Descripción detallada de la invención

Como se usa en la presente solicitud, las formas en singular "un/uno", "una" y "el/la" incluyen sus correspondientes plurales a menos que el contexto indique claramente lo contrario. A
5 menos que se defina otra cosa, todos los términos técnicos y científicos usados en el presente documento tienen el mismo significado que un experto en la técnica a la que esta invención pertenece entiende habitualmente. Con el fin de facilitar la comprensión y de aclarar el significado de determinados términos en el contexto de la presente invención se aportan las siguientes **definiciones** y realizaciones particulares y preferentes de las
10 mismas, aplicables a todas las realizaciones de los distintos aspectos de la presente invención:

"**PCR**" significa reacción en cadena de la polimerasa. En general, consta de la repetición de incubaciones sucesivas a una temperatura de desnaturalización, una temperatura de
15 hibridación y una temperatura de polimerización, aunque existen incontables modificaciones sobre el método general. Se incluyen métodos basados en la PCR, tales como *touchdown* PCR (TD-PCR), "*nested PCR*", "*primer extension*", etc.

"**Producto de amplificación**" o "**Producto de PCR**" es la mezcla de ADNs en solución
20 obtenidos como resultado de una reacción de PCR.

"**Cebador**" o "*primer*" (de su denominación en inglés) es un oligonucleótido capaz de hibridar con la secuencia que se pretende amplificar, y que sirve como punto de partida para que la polimerasa comience la reacción de amplificación. Esta molécula puede incluir
25 modificaciones, como biotilación, fosforilación, adición de ácidos nucleicos bloqueados ("*locked nucleic acids*", LNA).

"**Cebador específico**" se refiere a un cebador que es capaz de unirse específicamente o hibridar específicamente con la secuencia que se pretende amplificar. En una realización
30 particular el cebador es un deoxioligonucleótido con un tamaño de 15-35 nucleótidos, preferentemente 18-25 nucleótidos.

"**Hibridar específicamente**" se refiere al reconocimiento entre moléculas de ácidos nucleicos que presentan complementariedad exacta (100% de complementariedad) entre
35 sus secuencias de nucleótidos.

"Cebador degenerado" se refiere a mezclas de oligonucleótidos que están compuestos por nucleótidos N (N es A, T, C o G). En particular la longitud de los oligonucleótidos es de 5-16 nucleótidos, preferentemente 5-11 y más preferiblemente 7 nucleótidos. En los oligonucleótidos, todas las N pueden ser el mismo nucleótido o pueden ser diferentes comprendiendo así el oligonucleótido variaciones en la secuencia en alguna de sus posiciones. En particular puede comprender variaciones en al menos 1, al menos 2, al menos 3, al menos 4, al menos 5, al menos 6, al menos 7, al menos 8, al menos 9, al menos 10, al menos 11, al menos 12, al menos 13, al menos 14, al menos 15 posiciones, según el tamaño del oligonucleótido. En una realización preferida el cebador degenerado comprende una mezcla de todos los posibles oligonucleótidos. En una realización preferente los oligonucleótidos están fosforilados (cebador degenerado fosforilado), preferiblemente están monofosforilados. Preferiblemente los oligonucleótidos están fosforilados en su extremo 5' (cebador degenerado fosforilado en su extremo 5').

"Cebador directo" es el cebador que se extiende en la PCR desde el codón de iniciación hacia el codón de terminación (codón Stop) del ADN molde.

"Cebador inverso" es el cebador que se extiende desde el codón de terminación hacia el codón de iniciación del ADN molde.

"Genoma de referencia" se refiere a una secuencia digital de ácidos nucleicos, previamente ensamblada como ejemplo representativo del conjunto de genes de una especie.

"Etapa de hibridación" se refiere a la fase de hibridación de la PCR. Esta fase se denomina también fase de "annealing", de emparejamiento o de alineamiento. La temperatura de fusión o annealing (T_m , "melting temperature") depende de varios factores y es relativamente específica para cada primer. La longitud de los primers y la secuencia son críticas en la designación de los parámetros de una amplificación, una fórmula simple para calcular la T_m es la siguiente: $T_m = 4(G+C) + 2(A+T)$, aunque hay muchas otras que también pueden utilizarse. En el caso de un cebador degenerado, la T_m se refiere a la T_m media de la mezcla de los oligonucleótidos compuestos por nucleótidos N.

Las "variantes estructurales" implican cambios en algunas partes de los cromosomas en lugar de cambios en el número de cromosomas o juegos de cromosomas en el genoma. Hay cuatro tipos comunes de mutaciones que resultan en variantes estructurales: deleciones e inserciones, por ejemplo duplicaciones (las cuales implican un cambio en la cantidad de

ADN en un cromosoma, pérdida y ganancia de material genético respectivamente), inversiones (que implican un cambio en la disposición de un segmento cromosómico) y translocaciones (que implican un cambio en la ubicación de un segmento cromosómico y pueden dar lugar a fusiones génicas). En la presente invención el término “**variante estructural**” comprende una variante estructural seleccionada del grupo formado por una variante estructural caracterizada por (que comprenden) una pérdida de material genético, una variante estructural caracterizada por una ganancia de material genético, una translocación, una fusión génica y combinaciones de las mismas.

“**Pérdida de material genético**” se refiere a la existencia de regiones del genoma en que al menos uno de los alelos muestra la desaparición de un segmento de ADN que sí aparece en dicho alelo en el genoma de referencia. Esta pérdida puede dar lugar a cambios en el fenotipo.

“**Ganancia de material genético**” se refiere a la existencia de regiones del genoma en que al menos uno de los alelos muestra la inserción de un segmento de ADN ausente en dicho alelo en el genoma de referencia. El segmento de ADN insertado puede ser de origen endógeno (duplicaciones, pseudogenes, retrotrasposones, etc.) o exógeno (transgenes, integración de virus, etc.). Esta ganancia de material genético puede dar lugar a cambios en el fenotipo en función del lugar de inserción o de la secuencia del elemento insertado. En el contexto de la presente invención, el término “**inserción**” se refiere a cualquiera de las variantes estructurales que conlleva una ganancia de material genético, en particular a las mencionadas en este párrafo: duplicación, pseudogen, retrotrasposón, transgen e integración de virus.

“**Translocación**” se refiere a la reorganización de regiones del genoma que da lugar a la presencia de un segmento de ADN en una región del genoma diferente a la de su localización en el genoma de referencia. Una translocación puede dar lugar a cambios en el fenotipo en función del lugar de inserción o escisión y de la secuencia del elemento insertado.

“**Fusión génica**” se refiere al resultado de la ruptura y fusión de un segmento de ADN que ocurre en secuencias contenidas dentro de genes. Ocurre generalmente como resultado de una translocación y puede dar lugar a un transcrito fusionado que produce una proteína quimérica. La fusión génica puede dar lugar a cambios en el fenotipo, en función de los genes afectados por el reordenamiento.

5 “**Región de ácidos nucleicos certera**” (Región ANC) se refiere a la región de ácidos nucleicos (secuencia de nucleótidos contiguos en una molécula de ácido nucleico) que con una certeza del 100% comprende una mutación genética en estudio. La región ANC se puede determinar a partir de material genético extraído de una muestra biológica tomada del sujeto en estudio. En particular, a partir de ADN o ARN extraído de una muestra biológica tomada del sujeto en estudio, preferentemente ADN. Como se ha indicado anteriormente, la región ANC es determinable a través de multitud de técnicas conocidas por el experto en la materia. Entre otras destacan MLPA, CGH *array*, SNP *array*, NGS, *Southern blot*, análisis de cariotipo, FISH, RT-PCR y Δ -PCR. El experto en la materia sabe igualmente cómo determinar la región que con un 100% de certeza comprende la mutación en estudio (región ANC).

15 “**Gen**” incluye no solo regiones que codifican productos génicos sino también regiones reguladoras, incluyendo, por ejemplo, promotores, regiones de terminación, secuencias reguladoras de la traducción (tales como sitios de unión al ribosoma y sitios internos de entrada al ribosoma), potenciadores, silenciadores, aislantes, elementos de frontera, orígenes de replicación, sitios de unión a la matriz y regiones de control del locus. El término “gen” incluye además todos los intrones y otras secuencias de ADN de corte y empalme del transcrito de ARNm, junto con variantes resultantes de los sitios de corte y empalme alternativos. El término “gen” incluye además cualquier porción de un gen, por ejemplo, cualquier porción de las regiones mencionadas anteriormente.

25 “**Sujeto**” se refiere a cualquier miembro de los reinos Monera, Protista, Fungi, Plantae o Animalia. En una realización preferente, el sujeto es un miembro de la subdivisión Magnoliophyta incluyendo, sin limitación, organismos modelo, tales como *Arabidopsis thaliana* o arroz, y especies de plantas monocotiledóneas y dicotiledóneas con interés agronómico, tales como, trigo y otros cereales, girasol, algodón, soja y otras leguminosas, plantas ornamentales, árboles frutales, coníferas y otras cosechas.

30 En otra realización preferente, el sujeto es cualquier miembro de la clase Mammalia, incluyendo, sin limitación, seres humanos y primates no humanos tales como chimpancés y otros simios y especies de monos; animales de granja, tales como ganado vacuno, ovejas, cerdos, cabras y caballos; mamíferos domésticos tales como perros y gatos; animales de laboratorio incluyendo roedores, tales como ratones, ratas y cobayas, y similares. El término no indica una edad o sexo concretos. Por lo tanto, los sujetos adultos y recién nacidos, así como los fetos, sean de sexo masculino o femenino, están destinados a su inclusión dentro

del alcance de este término. El sujeto es, preferentemente, un ser humano.

Antes de llevar a cabo cualquiera de los métodos de la presente invención, se puede conocer la presencia de la mutación en el sujeto en estudio, o puede haber sospecha de que el sujeto tenga la mutación, por ejemplo en base a datos clínicos, o se puede desconocer por completo si el sujeto tiene o no la mutación .

La “**muestra biológica**” tomada de un sujeto en estudio contiene cualquier material biológico que permita extracción de ácidos nucleicos y es un material que comprende material genético de un sujeto o de una mezcla de sujetos. En la presente invención, la muestra comprende material genético del sujeto en estudio. En una realización particular el material genético es ADN o ARN. ADN se refiere a cualquier tipo de ADN, como por ejemplo: ADN genómico (ADNg), ADN mitocondrial (ADNm), ADN complementario (ADNc), ADN circulante (ADNci). En una realización particular, el material genético es ADNg, ADNm, ADNc o ADNci. En una forma de realización preferida, el ADN es ADN genómico o ADNc, preferentemente ADN genómico. El aislamiento del ácido nucleico de la muestra puede realizarse por procedimientos estándar conocidos por la persona experta en la técnica, tales como los descritos en Sambrook *et al.*, (Molecular Cloning: A Laboratory Manual, New York: Cold Spring Harbor Press, 1989). En una realización particular, la muestra biológica es tejido fresco o fijado obtenido de biopsia o autopsia o un fluido biológico. En una forma de realización preferida de cualquiera de los métodos de la presente invención, la muestra biológica es un fluido biológico siendo así el método de la invención no invasivo, mínimamente invasivo o menos invasivo que los que requieren una muestra de tejido tomada mediante biopsia. El fluido biológico se selecciona del grupo formado por saliva, sangre, suero, plasma, orina, heces, eyaculado, médula ósea, exudado bucal o bucofaríngeo, líquido pleural, líquido peritoneal, líquido pericárdico, líquido cefalorraquídeo, líquido intraarticular, líquido amniótico, y mezclas de los mismos. Preferentemente, el fluido biológico se selecciona entre sangre, plasma, suero o saliva, más preferentemente la muestra biológica es sangre o saliva. En una realización adicional, la muestra biológica es una muestra materna que incluye ADN fetal, permitiendo así un análisis prenatal.

La presente invención se refiere en un primer aspecto a un método para identificar la posición de una mutación genética, caracterizado por que comprende las siguientes etapas:

a) determinar la región ANC que comprende la mutación en estudio;
b) llevar a cabo una PCR, donde el molde de la PCR es ADN extraído de una muestra biológica tomada del sujeto en estudio y los cebadores son:

b1) un cebador degenerado, y

b2) un cebador específico, que hibrida específicamente:

- dentro de la región ANC, a una distancia de 0,2-100 Kb de cualquiera de los extremos de la región ANC, si la mutación comprende una ganancia de material genético, una translocación o una fusión génica, y se extiende mediante la PCR (en dirección 5' a 3') hacia fuera de la región ANC, o

- fuera de la región ANC, a una distancia de 0,2-100 Kb de cualquiera de los extremos de dicha región, si la mutación comprende una pérdida de material genético, y se extiende mediante la PCR (en dirección 5' a 3') hacia la región ANC, y donde en la reacción de PCR la etapa de hibridación comprende dos hibridaciones consecutivas, una primera hibridación a una temperatura de hibridación igual a la T_m del cebador específico $\pm 10^\circ\text{C}$, y una segunda hibridación a una temperatura de hibridación igual a la T_m del cebador degenerado $\pm 5^\circ\text{C}$;

c) secuenciar el producto de amplificación obtenido en la etapa b); y

d) alinear la secuencia obtenida en la etapa c) con una secuencia de referencia para identificar la posición de la mutación.

La presente invención se refiere a métodos para caracterizar mutaciones genéticas, en concreto para identificar la posición exacta de dichas mutaciones genéticas, delimitando así los límites exactos de dichas mutaciones, y para cribar dichas mutaciones (detectar la presencia de dichas mutaciones). Los métodos aquí descritos pueden ser útiles en la caracterización de una variedad de mutaciones genéticas. Las mutaciones que pueden ser caracterizadas utilizando los métodos descritos en la presente invención incluyen, por ejemplo, una inserción, una delección, una duplicación, y un reordenamiento (por ejemplo, una translocación o una fusión génica), así como cualquier combinación de las mismas. La inserción puede ser de un transgen. La mutación genética puede ser una mutación en la línea germinal o una mutación somática.

Así, en una realización particular, la mutación genética es una variante estructural, más particularmente la variante estructural es seleccionada del grupo formado por una variante estructural caracterizada por una pérdida de material genético, una variante estructural caracterizada por una ganancia de material genético, una translocación, una fusión génica y combinaciones de las mismas. En una realización preferente la variante estructural se selecciona del grupo formado por inserción, delección, duplicación, transgen, translocación y fusión génica. Más preferentemente la variación estructural es una inserción, duplicación o delección.

La mutación en estudio puede ser una mutación recurrente que se ha asociado con uno o más tipos de enfermedades, como por ejemplo con cardiopatías familiares. En este caso, la etapa a) se puede llevar a cabo en una región particular del ADN, por ejemplo aquella que comprenda los genes asociados con patologías o la región codificante del genoma (el exoma), que únicamente es un 1% del genoma pero en el que está el 85% de las enfermedades hereditarias (Bick y Dimmock. *Whole exome and whole genome sequencing*, Curr Opin Pediatr 2011; 23: 594-600; Biesecker et al. *Exome sequencing: the expert view*, Genome Biol 2011;12:128), evitándose así el análisis del genoma completo.

En una realización particular la etapa a) del método de la invención se lleva a cabo a partir de material genético extraído de una muestra biológica tomada del sujeto en estudio, más particularmente dicho material genético es ADN o ARN. La etapa a) se puede llevar a cabo por cualquier método conocido por el experto en la materia que permita la determinación de la región ANC, como por ejemplo: MLPA, CGH, SNP-arrays, NGS, etc. Así, en una realización particular la etapa a) se lleva a cabo mediante una técnica seleccionada del grupo formado por MLPA, CGH *array*, SNP-*arrays*, NGS, *Southern blot*, análisis de cariotipo, FISH, RT-PCR y Δ -PCR. En una realización preferente según una cualquiera de las realizaciones anteriores, la etapa a) se lleva a cabo mediante MLPA, CGH *array*, SNP-*arrays*, NGS y más preferiblemente mediante NGS. En una realización preferente según una cualquiera de las realizaciones de este párrafo el material genético es ADN.

En una realización particular según una cualquiera de las realizaciones anteriores, en la etapa b) el molde de la PCR es ADN_g, ADN_c o ADN_{ci}, preferentemente ADN_g o ADN_c y más preferiblemente ADN_g. En el caso de fusiones génicas, el molde de PCR es preferentemente ADN_c.

En una realización preferente según una cualquiera de las realizaciones anteriores, el cebador degenerado tiene entre 5 y 11 nucleótidos, y más preferentemente 7 nucleótidos. Preferentemente, el cebador degenerado tiene la secuencia SEC ID N° 1 (NNNNNNN). El cebador degenerado puede comprender nucleótidos modificados, por ejemplo *locked nucleic acids* (LNAs). Igualmente, el cebador degenerado puede estar mono-, di- o tri-fosforilado. En una realización preferente según una cualquiera de las realizaciones anteriores del primer aspecto de la invención, el cebador degenerado está fosforilado, preferentemente monofosforilado, en su extremo 5'. En una realización preferente el cebador degenerado tiene la secuencia SEC ID N° 1 y está fosforilado en su extremo 5' (5'-pNNNNNNN-3').

En una realización particular según una cualquiera de las realizaciones anteriores, en la etapa b) la temperatura de hibridación de la segunda hibridación es de entre 20-40°C, preferiblemente 27-33°C y más preferiblemente 30°C.

5 En otra realización particular según una cualquiera de las realizaciones anteriores, en la etapa b) la segunda hibridación se lleva a cabo durante un tiempo de entre 1 y 10 segundos, preferentemente 1-3 segundos y más preferentemente 1 segundo. Ventajosamente, la hibridación breve limita la incorporación de oligonucleótidos degenerados en las cadenas de nueva síntesis y evita la aparición de productos de reacción inespecíficos.

10 En una realización preferida según una cualquiera de las realizaciones anteriores, en la etapa b) la segunda hibridación se lleva a cabo a 30°C durante 1 segundo, consiguiéndose así una amplificación con la menor cantidad de productos de reacción inespecíficos.

15 Así, la reacción de PCR incluye un paso de hibridación a baja temperatura durante un tiempo muy corto para permitir la incorporación de los cebadores degenerados a las cadenas de nueva síntesis. Por ello, y para reducir o evitar que el cebador específico se una a esa baja temperatura a otras regiones del genoma, el cebador específico puede incluir en su extremo 5' una secuencia complementaria a la de su extremo 3'. De este modo, se favorece la formación de estructuras en horquilla por hibridación entre ambos extremos del cebador a bajas temperaturas de *annealing*. Con esto se reduce la posibilidad de homología inespecífica del cebador específico con el ADN de estudio durante el paso de *annealing* a baja temperatura (segunda hibridación). La Figura 2 muestra la estructura secundaria predicha a 30°C (panel B) para un primer específico en el que se ha diseñado la extensión de 7 nucleótidos en su extremo 5' que favorecen la formación de estructuras intramoleculares por complementación de 9 nucleótidos. El tamaño de esta región con complementación puede variar en función del contenido GC del cebador, pero siempre de modo que forme estructuras secundarias a la temperatura de hibridación de la segunda hibridación que no sean estables a la temperatura de hibridación de la primera hibridación.

30 En una realización preferente, según una cualquiera de las realizaciones anteriores, el cebador específico comprende en el extremo 5' una secuencia complementaria a la de su extremo 3', preferentemente de 5-12 nucleótidos y más preferentemente de 7-9 nucleótidos. Así, se favorece la formación de estructuras secundarias en horquilla a bajas temperaturas de hibridación, como las requeridas para la hibridación del cebador degenerado, reduciendo la posibilidad de homología inespecífica con el ADN en estudio.

35

En una realización particular según una cualquiera de las realizaciones anteriores, el cebador específico tiene la secuencia SEC ID N° 2 o SEC ID N° 5.

5 Como se ha indicado anteriormente, el cebador específico se diseña de manera que hibride específicamente dentro o fuera de la región ANC según el tipo de mutación que se esté caracterizando. Dicho cebador se extiende mediante la PCR en dirección 5' a 3' mediante la PCR hacia la posición de la mutación en estudio. Así, en el caso en que la mutación comprende una ganancia de material genético, translocación o fusión génica, donde la
10 mutación se encuentra fuera de la ANC, el cebador específico hibrida dentro de la región ANC y se extiende hacia fuera de la región ANC, En el caso en que la mutación comprende una pérdida de material genético, el cebador específico hibrida fuera de la región ANC y se extiende hacia la región ANC. En concreto, el cebador específico hibrida específicamente en una región que se encuentra a 0,2-100 Kb de cualquiera de los extremos (3' o 5') de la
15 región ANC, dentro o fuera de la región ANC según la mutación (dentro de la región ANC en mutaciones que comprenden ganancia de materia genético, translocaciones o fusiones génicas y fuera de la región ANC en mutaciones que comprenden pérdida de materia genético). Más particularmente, el cebador específico hibrida específicamente a 0,5-80 Kb de los extremos de la región ANC. Preferentemente a 0,5-40 Kb y más preferiblemente a
20 0,5-20 Kb de los extremos de la región ANC.

En una realización preferida, cuando la mutación es una variante estructural caracterizada por una ganancia de material genético, el cebador específico hibrida a 0,2-10 Kb del extremo 5' o 3', dentro de la región ANC de ganancia de material genético.

25 En la presente invención, en la etapa b) la PCR se lleva a cabo con únicamente los cebadores b1) y b2). Como resultado del uso de un único cebador específico, se obtiene una amplificación lineal. Además, en la realización preferente en la que el cebador degenerado está fosforilado, la incorporación de dicho cebador degenerado fosforilado a la
30 reacción permite primar las cadenas lineales y generar regiones locales de doble hebra sobre las que la polimerasa regenera un ácido nucleico de doble cadena. Esto es una importante ventaja ya que los métodos establecidos para preparación de muestras para secuenciación NGS utilizan ADN de doble cadena como material de partida. Así, en una
35 realización preferente, según una cualquiera de las realizaciones anteriores, la secuenciación de la etapa c) se lleva a cabo mediante NGS. Y en otra realización preferente según una cualquiera de las realizaciones anteriores, en la etapa b) el cebador degenerado

está fosforilado en su extremo 5', preferentemente monofosforilado, y la secuenciación de la etapa c) se lleva a cabo mediante NGS.

5 En una realización particular, el método del primer aspecto de la invención comprende, tras la etapa d) en la que se identifica la posición de la mutación, una etapa e) en la que se lleva a cabo el diseño de cebadores que flanqueen la posición de la mutación genética, una reacción de PCR con dichos cebadores y ADN extraído de una muestra biológica del sujeto en estudio y una secuenciación de Sanger del producto amplificado. Así la etapa e) comprende la reconfirmación de la posición identificada en la etapa d) mediante
10 secuenciación Sanger. La secuenciación por el método descrito por Sanger et al en 1977 (Sanger *et al.*, *DNA sequencing with chain-terminating inhibitors*. Proc Natl Acad Sci USA 1977; 74: 5463-7) es de estándar clínico, por lo que esta etapa es recomendable cuando el método de la invención se usa para detectar o diagnosticar enfermedades caracterizadas por la mutación genética en estudio.

15 Las anomalías genéticas, como la duplicación, delección, translocación cromosómica, fusión génica y mutación puntual a menudo conducen a condiciones patológicas. Algunas enfermedades, como el cáncer, se deben a anomalías genéticas adquiridas en unas pocas células durante la vida, mientras que en otras enfermedades la anormalidad genética está presente en todas las células del cuerpo y presente desde la concepción, como ocurre en
20 algunas cardiopatías familiares. Así, en una realización particular la mutación genética en estudio es una mutación que está asociada al desarrollo de una patología, como por ejemplo, cardiopatías, cáncer, trastornos del desarrollo neurológico, enfermedad de Crohn, artritis reumatoide, diabetes de tipo 1 y tipo 2, etc. En una realización preferente, la patología es una cardiopatía familiar, más preferentemente es una cardiomiopatía o canalopatía.

En una realización particular según una cualquiera de las realizaciones anteriores, la mutación es una delección. Más particularmente, la delección es de un gen cuya delección está asociada a una patología, preferentemente una cardiopatía familiar.

30 En una realización particular según una cualquiera de las realizaciones anteriores, la mutación es una inserción. Más particularmente, la inserción es una duplicación de al menos un gen o al menos una parte de un gen y dicha duplicación está asociada a una patología, preferentemente una cardiopatía familiar.

35

En una realización particular, la mutación es una delección en el gen FBN1 y el cebador específico tiene la secuencia SEC ID N° 2. En otra realización particular, la mutación es una duplicación del gen DMD y el cebador específico tiene la secuencia SEC ID N° 5. La delección del gen FBN1 está asociada al síndrome de Marfan y la duplicación del gen DMD está asociada a la distrofia muscular y cardiomiopatía dilatada.

La metodología de la presente invención de acuerdo al primer aspecto de la invención, permite conocer la secuencia de regiones que alcanzan varias decenas de kilobases desde una secuencia conocida, con mayor extensión y especificidad que los métodos anteriormente descritos en el estado de la técnica. Para llevar a cabo el método de la invención basta con disponer de una secuencia conocida sobre la que diseñar un único oligonucleótido específico y no se precisa ningún conocimiento previo sobre la secuencia que se pretende caracterizar.

Como se ha indicado anteriormente, los métodos del estado de la técnica se basan en la amplificación exponencial del ADN. Sorprendentemente, el método de la presente invención elimina la necesidad de este paso e integra en la reacción de PCR un reactivo adicional que regenera la doble hebra sin necesidad de añadir adaptadores. Se obvia además la necesidad de sintetizar oligonucleótidos con modificaciones que favorecen la captura de los productos de PCR (por ejemplo, biotilación), ya que la secuenciación de la etapa c), en particular mediante NGS, aporta la sensibilidad necesaria para identificar fragmentos de ADN que cubran los límites exactos de la modificación genética. Así, en una realización preferente, según una cualquiera de las realizaciones anteriores según el primer aspecto de la invención, el cebador específico no comprende modificaciones.

Así, el método de la presente invención encuentra especial interés para su aplicación sobre muestras de sujetos en las que, tras la aplicación de alguna o varias de las técnicas de las citadas en la sección de antecedente u otras, se identifique pérdida o ganancia de material genético. La aplicación de esta invención permite la caracterización exacta de los límites exactos de la delección, inserción, translocación o fusión génica así como, en caso de inserción, translocación y fusión génica, la determinación del lugar exacto del genoma en que ésta se produce.

La invención encuentra también aplicación en la caracterización de lugares de inserción de transgenes en organismos en los que, por medio de manipulación genética, se introduce en su genoma una secuencia exógena conocida.

La invención es también de aplicación en la caracterización de fusiones génicas. Sorprendentemente, y a diferencia de las técnicas en uso, la presente invención sirve para la identificación y caracterización de fusiones a partir de la secuencia de sólo uno de los genes implicados en la fusión. Por ejemplo, los ensayos que se llevan a cabo por RT-PCR para detectar fusiones génicas, exigen que ambos elementos fusionados sean conocidos y que a su vez den lugar a una variante previamente caracterizada. En la presente invención, se puede caracterizar la fusión génica a partir del único gen conocido e implicado en dicha fusión génica. En este caso, la región ANC se corresponde con dicho único gen conocido. La presente invención supone así una importante ventaja, ya que como se ha indicado anteriormente actualmente se considera que fusiones génicas, producidas por translocaciones cromosómicas, inversiones, deleciones, etc, tienen gran relevancia en cánceres epiteliales frecuentes, como los carcinomas de próstata o pulmón.

Una vez identificada la posición de la mutación siguiendo el método del primer aspecto de la presente invención, resulta posible diseñar cebadores que permitan amplificar específicamente el punto exacto de la mutación. Y una vez se dispone de estos cebadores, el estudio de presencia o ausencia de la mutación en individuos emparentados se limita a PCR y secuenciación Sanger, método establecido desde hace años y mucho más económico y menos exigente técnicamente que NGS. Así, en un segundo aspecto la presente invención se refiere a un método para cribar una mutación genética caracterizado por que comprende las siguientes etapas:

i) identificar la posición de la mutación genética siguiendo el método de la invención según una cualquiera de las realizaciones del primer aspecto de la invención;

ii) llevar a cabo una PCR donde el ADN modelo es ADN extraído a partir de una muestra tomada de un sujeto en estudio y los cebadores son un cebador directo y un cebador inverso que hibridan específicamente aguas arriba y aguas abajo, respectivamente, de la posición exacta de la mutación de manera que se amplifica un producto de PCR que comprende dicha posición;

iii) analizar el tamaño del producto de PCR amplificado en ii), donde un producto de PCR del tamaño esperado según los cebadores directo e inverso de la etapa ii) (es decir, tamaño esperado según la posición en la que hibriden dichos cebadores) es indicativo de la presencia de la mutación genética, y la ausencia de dicho producto de PCR con el tamaño esperado es indicativa de ausencia de la mutación genética.

En una realización particular del segundo aspecto de la invención según una cualquiera de las realizaciones anteriores, cuando el análisis de la etapa iii) indica la presencia de la mutación genética, el método comprende además una etapa iv), en la que se secuencian el producto de PCR obtenido en la etapa ii).

5

En otra realización particular del segundo aspecto de la invención según una cualquiera de las realizaciones anteriores, alternativamente a la etapa iii) se lleva a cabo una etapa iv), en la que se secuencian el producto de PCR obtenido en la etapa ii). En este caso, el resultado de la secuenciación mostrará la presencia o no de la mutación genética en estudio.

10

En una realización preferente según una cualquiera de las realizaciones particulares de los dos párrafos anteriores, el producto obtenido en la etapa ii) es purificado y concentrado antes de ser secuenciado en la etapa iv). En una realización preferente, la secuenciación de la etapa iv) es una secuenciación Sanger, que como se ha indicado anteriormente es una secuenciación de estándar clínico.

15

Como se ha indicado anteriormente, muchas de las mutaciones genéticas caracterizables por los métodos de la presente invención están asociadas a patologías. Así, en una realización particular según una cualquiera de las realizaciones del segundo aspecto de la invención, la mutación genética se asocia con una patología, como por ejemplo, cardiopatías, cáncer, trastornos del desarrollo neurológico, enfermedad de Crohn, artritis reumatoide, diabetes de tipo 1 y tipo 2, etc. En una realización preferente, la patología es una cardiopatía familiar, más preferentemente es una cardiomiopatía o canalopatía.

20

Puesto que el método de acuerdo al segundo aspecto de la presente invención permite determinar la presencia o no de una mutación y dicha mutación puede estar asociada a una patología, dicho método puede usarse en el diagnóstico de dicha patología. Así, en un tercer aspecto, la presente invención se refiere a un método de diagnóstico de una patología asociada a una mutación genética dada que comprende las etapas del método según una cualquiera de las realizaciones del segundo aspecto de la invención en las que se lleva a cabo la etapa iii), siendo el sujeto en estudio diagnosticado como afecto con la patología cuando el producto de PCR de la etapa iii) tiene el tamaño esperado según los cebadores directo e inverso utilizados en la etapa ii).

30

El tercer aspecto de la invención también se refiere a un método de diagnóstico de una patología asociada a una mutación genética dada que comprende las etapas del método

35

según una cualquiera de las realizaciones del segundo aspecto de la invención en las que se lleva a cabo la etapa iv) en lugar de la etapa iii), siendo el sujeto en estudio diagnosticado como afecto con la patología cuando la secuenciación de la etapa iv) revela la presencia de la mutación genética.

5

El tercer aspecto de la invención se refiere también al uso de un método según una cualquiera de las realizaciones del aspecto primero o segundo de la presente invención, para el diagnóstico de una patología asociada a una mutación genética.

10

En una realización preferente según cualquiera de las realizaciones del tercer aspecto de la invención, la patología se selecciona del grupo formado por cardiopatías, cáncer, trastornos del desarrollo neurológico, enfermedad de Crohn, artritis reumatoide, diabetes de tipo 1 y diabetes de tipo 2. En una realización preferente, la patología es una cardiopatía familiar, más preferentemente es una cardiomiopatía o canalopatía. En otra realización preferente la patología se selecciona del grupo formado por síndrome de Marfan, distrofia muscular y cardiomiopatía dilatada.

15

Otras aplicaciones o usos de los métodos de la presente invención incluyen la identificación de la secuencia de un haplotipo a lo largo de un segmento del cromosoma (como, por ejemplo, para determinar el origen de diversas mutaciones en el caso de heterocigotos compuestos), o la determinación de puntos de inserción de ADN en materiales biológicos utilizados en terapia génica, producción de organismos transgénicos, etc.

20

Por último, en un cuarto aspecto, la presente invención se refiere a un kit para llevar a cabo el método según los aspectos primero, segundo o tercero, que comprende:

25

- un cebador degenerado b1), tal y como se ha definido en cualquiera de las realizaciones particulares del primer aspecto de la invención, preferentemente de secuencia SEC ID N° 1,
- un cebador específico b2) tal y como se ha definido en el primer aspecto de la invención, preferentemente de secuencia SEC ID N° 2 o SEC ID N° 5.

30

En una realización preferente según una cualquier de las realizaciones del párrafo anterior el cebador degenerado es un cebador degenerado fosforilado en su extremo 5'.

Ejemplos

35

A continuación se detallan unos ejemplos concretos de realización de la invención que sirven para ilustrar la invención sin limitar el alcance de la misma.

EJEMPLO 1: Caracterización de la delección de los exones 59-66 del gen FBN1.

1.1.- Identificación de la región ANC - pérdida de material genético (delección)

Se extrajo el ADN a partir de 1 mL de sangre en un extractor automático de ácidos nucleicos
 5 QIAAsymphony SP (QIAGEN) utilizando el kit QIAAsymphony DSP DNA Midi Kit.

Se obtuvieron muestras de distintos sujetos, en concreto, el sujeto en estudio y otros sujetos
 sin la cardiopatía en estudio. Mutaciones en el gen FBN1 se asocian al desarrollo de
 síndrome de Marfan, patología coherente con los datos clínicos del sujeto en estudio.

Las muestras se prepararon utilizando el kit “SureSelectXT Target Enrichment System for
 Illumina Paired-End Multiplexed Sequencing Library” (Agilent, SureSelectXT Target
 Enrichment System for Illumina Paired-End Multiplexed Sequencing Library Protocol Version
 B4, August 2015 SureSelect platform manufactured with Agilent SurePrint Technology), por
 15 hibridación con una solución de sondas homólogas a la región exónica de 214 genes, y se
 secuenciaron en un equipo HiSeq 1500 (Illumina). La comparación de la profundidad de
 cobertura entre muestras secuenciadas en paralelo (Yoon et al., *Sensitive and accurate
 detection of copy number variants using read depth of coverage*. Genome Res. 2009; 19:
 1586–92) permitió identificar, en una de las muestras de ADN, la delección de los últimos 8
 20 exones (exones 59-66) del gen FBN1. La delección comenzaba en un punto indeterminado
 del intrón 58 y que se extendía hacia el espacio intergénico aguas abajo del gen FBN1
 (región ANC = chr15:48703095-48718147; genoma de referencia hg19). Así, la región ANC
 comprende los exones 59-66 del gen FBN1.

Mutaciones en este gen se asocian al desarrollo de síndrome de Marfan, patología
 coherente con los datos clínicos del sujeto en estudio de ADN. La delección se extiende hacia
 el espacio intergénico, en zona no cubierta por el método de enriquecimiento utilizado en la
 preparación de la muestra. Por tanto, la delección podría abarcar desde 14 Kb hasta 6 Mb
 (distancia a la que se localiza el gen más próximo presente en la región enriquecida). Se
 30 utiliza el método de la invención para caracterizar el alcance exacto de la mutación.

1.2.- Amplificación

1.2.1.- Composición de la reacción de PCR:

10X Long PCR buffer con 15 mM MgCl ₂	5 µL
dNTP Mix, 2 mM cada uno	5 µl (0,2 mM cada uno).
Cebador específico FBN1-Ex58-FW	1 µM

ADN molde	50 ng
N7*	0,2 µM
Enzima Long PCR (Thermo Scientific, K0181)	2,5 U
H ₂ O libre de nucleasas	hasta 50 µL

5

(*): N7 = mezcla de cebadores degenerados fosforilados en 5' (SEC ID N° 1 fosforilada en su extremo 5': 5'-pNNNNNNN-3').

Cebador FBN1-Ex58-FW: GCTTTCCCCTCTTGCTTCTTCT (SEC ID N° 2), hibrida fuera de la región ANC y se extiende hacia la región ANC.

10

1.2.2.- Condiciones de la reacción de PCR:

3 min, 94°C;

10 ciclos [94°C, 20s; 60°C 30 s; 30°C, 1 s; 68°C, 20 min];

25 ciclos [94°C, 20s; 60°C; 30 s 30°C, 1 s; 68°C, 20 min, con una extensión de 15 s por ciclo]

15

Extensión final de 10 min a 68°C.

Mezcla de reacción y condiciones de PCR optimizadas para amplificar 30 Kb.

20

1.3.- Preparación de los productos de amplificación para secuenciación:

1.3.1.- Eliminación de primers y fragmentos de bajo peso molecular utilizando Agencourt AMPure XP (Beckman Coulter, A63881) siguiendo las instrucciones del fabricante.

1.3.2.- Preparación de librerías con los kits NEXTflex™ Rapid DNA Sequencing Kit (Bioo Scientific, 5144-02) y NEXTflex™ DNA Barcodes-96 (Bioo Scientific, 514105) siguiendo las instrucciones del fabricante.

25

1.3.3.- Cuantificación de muestra y preparación de un *pool* para la secuenciación de las librerías.

1.4.- Secuenciación NGS de las librerías:

30

La secuenciación NGS se realizó en un ultrasecuenciador Illumina Hi-Seq 1500. La secuenciación se llevó a cabo en modo Rapid y formato "Paired-end" a 2x100 (HISEQ Rapid SBS Kit v2).

1.5.- Análisis bioinformático

35

El demultiplexado de las librerías se hizo con el software CASAVA 1.8.2 (Illumina). Los ficheros correspondientes a cada muestra (en formato Fastq) se alinearon frente a la

secuencia de referencia humana GRCh37/hg19 utilizando el programa informático BWA-MEM.

Este análisis permitió detectar varias secuencias quiméricas no duplicadas y localizadas aguas arriba de la deleción (Figura 1). La comparación de estas secuencias quiméricas con el genoma de referencia utilizando la herramienta BLAT (UCSC Genome Browser) mostró que dichas secuencias incluían, cubriendo ambos extremos, una aparente deleción de ~73,8 Kb (Figura 1).

1.6.- Confirmación y cribado

Para confirmar que dicha deleción era real, se diseñaron y sintetizaron oligonucleótidos a ambos extremos 3' y 5' de la deleción. Una reacción de PCR sobre el ADN de la muestra que presentaba la deleción y con los cebadores directo (FBN1_del8-FW) e inverso (FBN1_del8-RV) de secuencia SEC ID N° 3 (TGGAAGCACAAGCTCCTT) y SEC ID N° 4 (CCAGGCAAGTGTCAGCATTA), respectivamente, produjo un amplicón de ~500 pb que no aparecía en muestras control. Este amplicón se secuenció por medio de secuenciación Sanger y confirmó la presencia de una deleción con coordenadas NC_000015.9:g.48645275_48719058del (marcada con * en la Figura 1).

De este modo, el estudio de la posible presencia de la mutación en muestras de parientes del sujeto en estudio queda simplificada al análisis del amplicón obtenido a partir de los cebadores que flanquean la deleción, como por ejemplo los oligonucleótidos de SEC ID N° 3 y SEC ID N° 4.

EJEMPLO 2: Caracterización de la duplicación de los exones 46 y 47 del gen DMD.

2.1.- Identificación de la región ANC - ganancia de material genético (inserción)

Se llevó a cabo como en el apartado 1.1, pero se analizó la duplicación de los exones 46 y 47 y de parte de los intrones adyacentes del gen DMD. Mutaciones en este gen se asocian al desarrollo de distrofia muscular y cardiomiopatía dilatada, patología coherente con los datos clínicos del donante de ADN. La duplicación incluye los exones 46 y 47 completos y parte de los intrones 45 y 47, pero no los exones 45 ó 48 (región ANC = chrX:31946705-31951957; genoma de referencia, hg19), por lo que podría abarcar desde 5,2 Kb hasta 100 Kb. No se obtiene información sobre el punto de inserción en el genoma. Se utiliza el método de la invención para caracterizar el alcance exacto y la localización de la mutación.

2.2.- Amplificación

En este caso, se diseña un oligonucleótido que incluye en su región 5' una secuencia

complementaria a la del extremo 3' para minimizar la posibilidad de hibridación inespecífica a bajas temperaturas de “*annealing*” (segunda hibridación).

Se utiliza el cebador específico:

5 Cebador DMDdup-Ex47-RV 5'- **CACATAGTTGTTTTGTTGTCTTTTGGGA**ACTATGTG****
(SEC ID N° 5).

La secuencia subrayada (posiciones 8-36 del cebador) es complementaria a las coordenadas NC_000023.10:g. 31951631-31951659, contenidas dentro de la zona duplicada. Las posiciones 1-9 del oligonucleótido son complementarias a las posiciones 28-36 (ambas en negrita). Se predice según diferentes modelos (RNAstructure: Reuter & Mathews, *RNAstructure: Software for RNA secondary structure prediction and analysis*. BMC Bioinformatics 2010; 11: 129; RNAfold: Hofacker et al., *Fast folding and comparison of RNA secondary structures*. Monatsh Chemie 1994; 125: 167-8) que este cebador presenta estructura secundaria a 30°C (temperatura de la segunda hibridación) (ver Figura 2). Los mismos programas predicen ausencia de estructura secundaria a 60°C (temperatura de la primera hibridación).

La reacción de PCR fue la misma (a excepción del cebador específico) que en el Ejemplo 1 y se llevó a cabo en las mismas condiciones que en el Ejemplo 1.

2.3-2.5

La preparación de los productos de amplificación, secuenciación NGS y análisis bioinformático se llevaron a cabo como se ha explicado en el Ejemplo 1.

25 El análisis permitió detectar secuencias apareadas (identificadas por secuenciación “*paired-end*”) no duplicadas y que resultan compatibles con la presencia de una duplicación en tándem (Figura 3). La comparación de estas secuencias con el genoma de referencia utilizando la herramienta BLAT (Kent, *BLAT--the BLAST-like alignment tool*. Genome Res 2002; 12: 656-64. UCSC Genome Browser) sugirió que dichas secuencias abarcaban los extremos de una posible duplicación en tándem de ~32,4 Kb que incluye los exones 46 y 47 de DMD (Figura 3).

2.6.- Confirmación y cribado

35 Para confirmar que dicha duplicación era real, se diseñaron y sintetizaron cebadores a ambos extremos de la posible región duplicada. Una reacción de PCR sobre el ADN de la muestra que presentaba la duplicación con los oligonucleótidos DMD_46-47dup-FW de SEC

ID N° 6 (CAGTTGGCAGAGAAAACACG) y DMD_46-47dup-RV de SEC ID N° 7 (TATCGCTTTGCCTACGCTCT) produjo un amplicón de ~500 pb que no aparecía en muestras control. Este amplicón se secuenció por medio de secuenciación Sanger y se confirmó la presencia de una duplicación en tándem de la región cromosómica que
5 corresponde con las coordenadas NC_000023.10:g.31920489_31953173dup.

De este modo, el estudio de la posible presencia de la mutación en muestras de parientes del sujeto en estudio queda simplificado al análisis del amplicón obtenido a partir de los cebadores SEC ID N° 6 y 7 que flanquean el punto de inserción de la duplicación.

REIVINDICACIONES

1.- Método para identificar la posición de una mutación genética caracterizado por que comprende las siguientes etapas:

- 5 a) determinar la región ANC que comprende la mutación genética en estudio;
- b) llevar a cabo una PCR, donde el molde de la PCR es ADN extraído de una muestra biológica tomada del sujeto en estudio y los cebadores son:
- b1) un cebador degenerado, y
- b2) un cebador específico que hibrida específicamente:
- 10 - dentro de la región ANC, a una distancia de 0,2-100 Kb de cualquiera de los extremos de la región ANC, si la mutación comprende una ganancia de material genético, una translocación o una fusión génica, y se extiende mediante la PCR hacia fuera de la región ANC, o
- fuera de la región ANC, a una distancia de 0,2-100 Kb de cualquiera de los extremos de la región ANC, si la mutación comprende una pérdida de material genético, y se extiende mediante la PCR hacia la región ANC,
- 15 y donde en la reacción de PCR la etapa de hibridación comprende dos hibridaciones consecutivas, una primera hibridación a una temperatura de hibridación igual a la T_m del cebador específico $\pm 10^\circ\text{C}$, y una segunda hibridación a una temperatura de hibridación igual a la T_m del cebador degenerado $\pm 5^\circ\text{C}$;
- 20 c) secuenciar el producto de amplificación obtenido en la etapa b); y
- d) alinear la secuencia obtenida en la etapa c) con un genoma de referencia para identificar la posición de la mutación.

25 2.- Método según la reivindicación anterior, donde la etapa a) se lleva a cabo a partir de ADN o RNA extraído de una muestra biológica tomada del sujeto en estudio.

3.- Método según la reivindicación anterior, donde la etapa a) se lleva a cabo a partir de ADN y mediante una técnica seleccionada del grupo formado por MLPA, CGH *array*, SNP *array*, NGS, *Southern blot*, análisis de cariotipo, FISH, RT-PCR y Δ -PCR, preferentemente MLPA, CGH *array*, SNP *array* y NGS y más preferiblemente NGS.

30

4.- Método según una cualquiera de las reivindicaciones anteriores, donde en la etapa b) el molde de la PCR es ADN_g, ADN_c o ADN_{ci}, preferentemente ADN_g o ADN_c y más preferiblemente ADN_g.

35

5.- Método según una cualquiera de las reivindicaciones anteriores, donde en la etapa b) el cebador degenerado es un cebador degenerado fosforilado en su extremo 5', preferentemente mono-fosforilado.

5 6.- Método según una cualquiera de las reivindicaciones anteriores, donde el cebador específico comprende en el extremo 5' una secuencia complementaria a la de su extremo 3', preferentemente de 5-12 nucleótidos y más preferentemente de 7-9 nucleótidos.

10 7.- Método según una cualquiera de las reivindicaciones anteriores, donde el cebador degenerado tiene entre 5 y 16 nucleótidos, preferiblemente entre 5 y 11, y más preferentemente 7 nucleótidos.

15 8.- Método según una cualquiera de las reivindicaciones anteriores, donde la temperatura de hibridación de la segunda hibridación es de entre 20-40°C, preferiblemente 27-33°C y más preferiblemente 30°C.

20 9.- Método según una cualquiera de las reivindicaciones anteriores, donde la segunda hibridación se lleva a cabo durante un tiempo de entre 1 y 10 segundos, preferentemente 1-3 segundos y más preferentemente 1 segundo.

10.- Método según una cualquiera de las reivindicaciones anteriores, donde el cebador específico hibrida a 0,5-80 Kb, preferentemente 0,5-40 Kb y más preferiblemente 0,5-20 Kb del extremo 3' o 5' de la región ANC.

25 11.- Método según una cualquiera de las reivindicaciones anteriores, donde el cebador degenerado tiene la secuencia SEC ID N° 1.

12.- Método según una cualquiera de las reivindicaciones anteriores, donde el cebador específico tiene la secuencia SEC ID N° 2 o SEC ID N° 5.

30 13.- Método según una cualquiera de las reivindicaciones anteriores, donde la secuenciación de la etapa c) se lleva a cabo mediante NGS.

35 14.- Método según una cualquiera de las reivindicaciones anteriores, donde tras la etapa d) en la que se identifica la posición de la mutación, se lleva a cabo una etapa e) en la que se reconfirma dicha posición mediante secuenciación Sanger.

15.- Método según una cualquiera de las reivindicaciones anteriores, donde la mutación genética es una variante estructural.

5 16.- Método según la reivindicación anterior, donde la variante estructural es seleccionada del grupo formado por una variante estructural caracterizada por una pérdida de material genético, una variante estructural caracterizada por una ganancia de material genético, una translocación, una fusión génica y combinaciones de las mismas.

10 17.- Método según una cualquiera de las reivindicaciones anteriores, donde la mutación genética es seleccionada del grupo formado por inserción, delección, duplicación, transgen, translocación y fusión génica, preferentemente es inserción, duplicación o delección.

15 18.- Método para cribar una mutación genética caracterizado por que comprende las siguientes etapas:

i) identificar la posición de una mutación genética siguiendo el método según una cualquiera de las reivindicaciones 1-17;

20 ii) llevar a cabo una PCR donde el ADN modelo es ADN extraído a partir de una muestra tomada de un sujeto en estudio y los cebadores son un cebador directo y un cebador inverso que hibridan específicamente aguas arriba y aguas abajo, respectivamente, de la posición exacta de la mutación en estudio;

25 iii) analizar del tamaño del producto de PCR amplificado en ii), donde un producto de PCR del tamaño esperado según los cebadores directo e inverso de la etapa ii) es indicativo de la presencia de la mutación genética, y la ausencia de dicho producto de PCR es indicativa de ausencia de la mutación genética.

19.- Método según la reivindicación 18, que comprende además una etapa iv), posterior a la etapa iii), en la que se secuencian el producto de PCR obtenido en la etapa ii).

30 20.- Método según la reivindicación 19, donde la secuenciación es una secuenciación Sanger.

35 21.- Método según la reivindicación 18, que comprende en lugar de la etapa iii) una etapa iv) en la que se secuencian el producto de PCR obtenido en la etapa ii), donde el resultado de la secuenciación mostrará la presencia o ausencia de la mutación genética en estudio.

22.- Método según la reivindicación 21, donde la secuenciación es una secuenciación Sanger.

5 23.- Método de diagnóstico de una patología asociada a una mutación genética caracterizado por que comprende las etapas del método según una cualquiera de las reivindicaciones 18-20, siendo el sujeto en estudio diagnosticado como afecto con la patología cuando el producto de PCR de la etapa iii) tiene el tamaño esperado según los cebadores directo e inverso utilizados en la etapa ii).

10 24.- Método de diagnóstico de una patología asociada a una mutación genética caracterizado por que comprende las etapas del método según una cualquiera de las reivindicaciones 21-22, siendo el sujeto en estudio diagnosticado como afecto con la patología cuando la secuenciación de la etapa iv) revela la presencia de la mutación genética.

15 25.- Método según una cualquiera de las reivindicaciones anteriores, donde la muestra biológica se selecciona del grupo formado por tejido fresco o fijado obtenido de biopsia o autopsia, sangre, plasma, suero, orina, heces, saliva, eyaculado, médula ósea, líquido cefalorraquídeo, líquido ascítico, líquido sinovial, líquido amniótico, exudado bucal o bucal-faríngeo, líquido pleural, líquido peritoneal, líquido pericárdico, líquido cefalorraquídeo, 20 líquido intraarticular, y mezclas de los mismos.

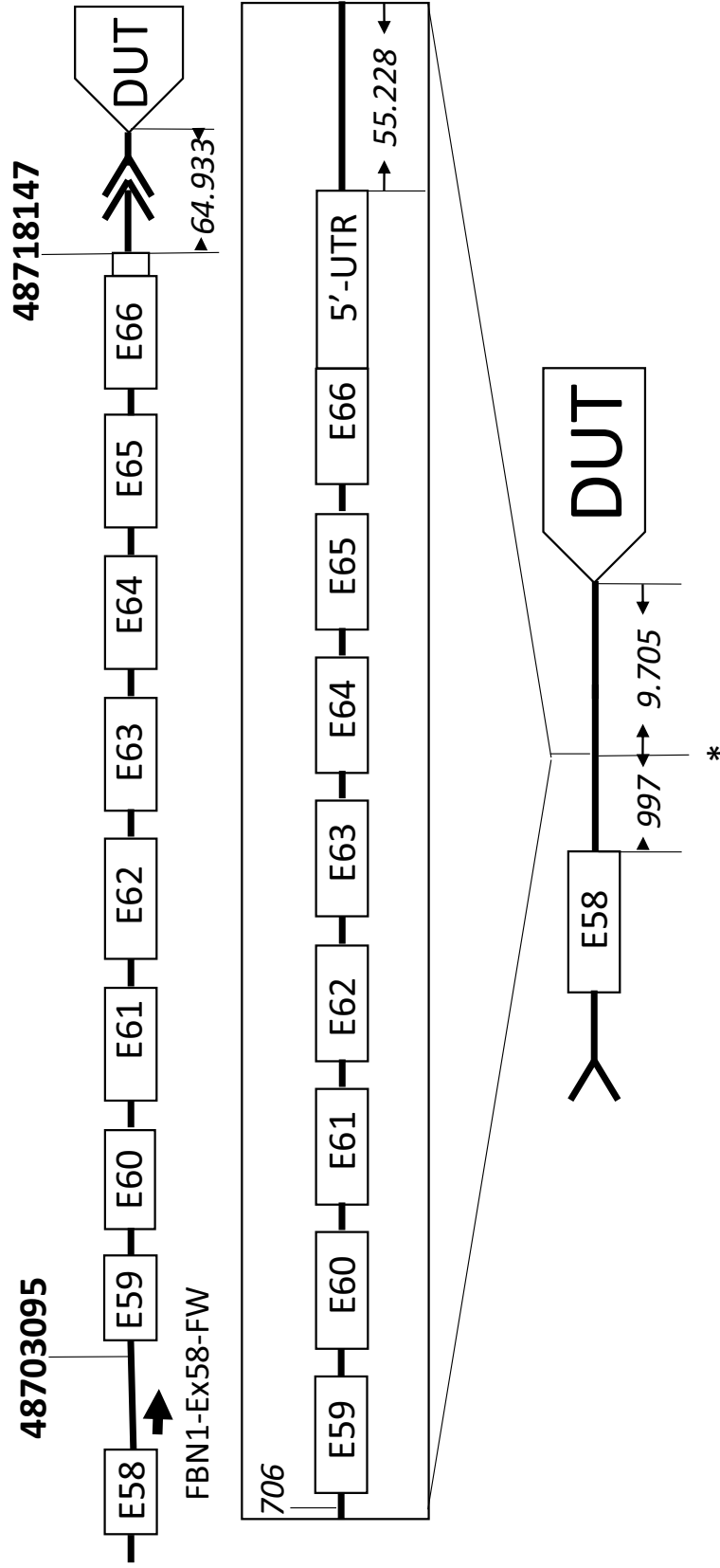


FIG 1 48645275-48719058 del

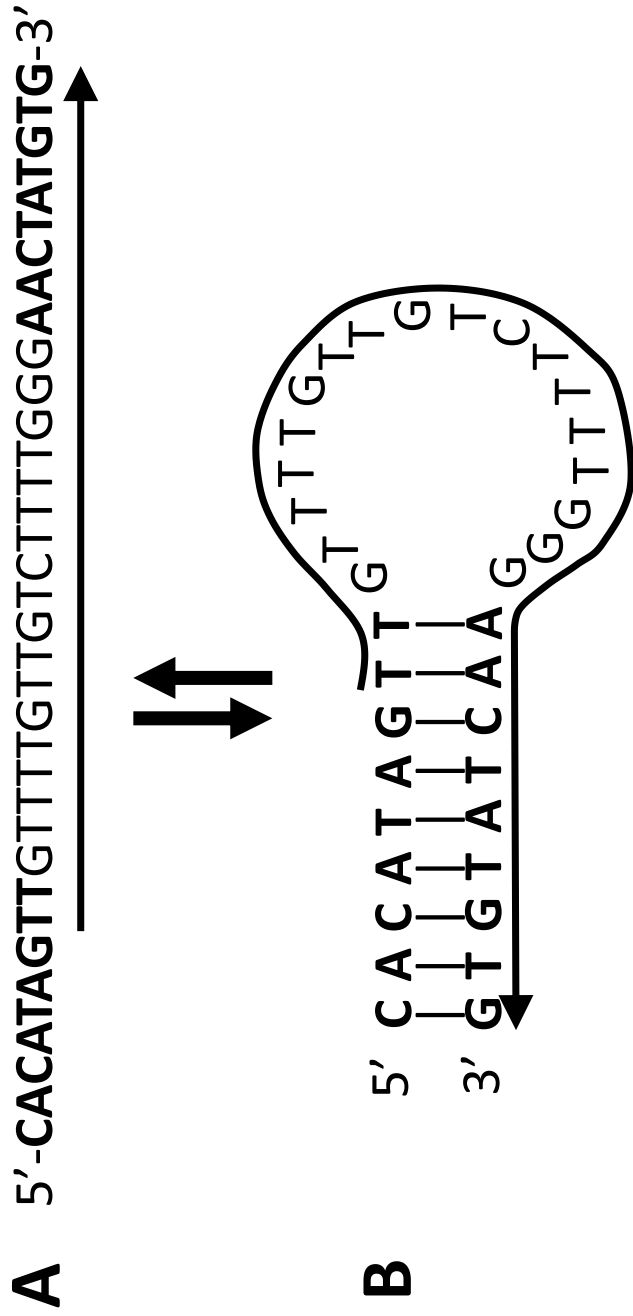


FIG 2

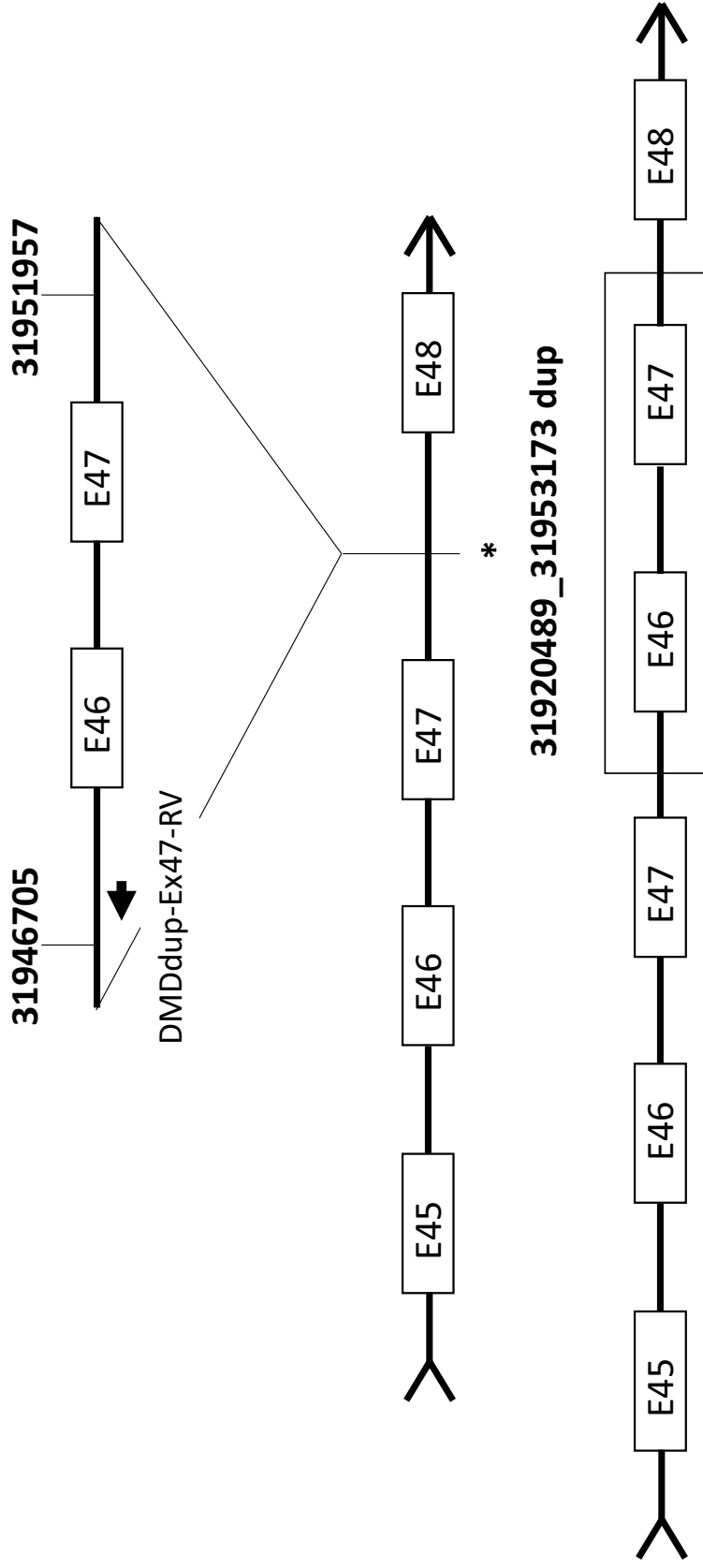


FIG 3

ES 2 641 690 A1

Listado de secuencias

<110> Health in Code
<120> Método de identificación de mutaciones
<130> 212/15
<160> 7
<170> PatentIn version 3.5
<210> 1
<211> 7
<212> DNA
<213> Artificial Sequence
<220>
<223> Cebador degenerado.

<220>
<221> misc_feature
<222> (1)..(7)
<223> n is a, c, g, or t

<400> 1
nnnnnnnn 7

<210> 2
<211> 22
<212> DNA
<213> Artificial Sequence
<220>
<223> Cebador FBN1-Ex58-FW

<400> 2
gctttcccct ctgcttctt ct 22

<210> 3
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> FBN1_de18-FW

<400> 3
tggaaaagca caagtcctt 20

<210> 4
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> Cebador FBN1_de18-RV

<400> 4
ccaggcaagt gtcagcatta 20

<210> 5
<211> 36

ES 2 641 690 A1

Listado de secuencias

<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador DMDdup-Ex47-RV

<400> 5
cacatagttg ttttgttgtc ttttggaac tatgtg 36

<210> 6
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador DMD_46-47dup-FW

<400> 6
cagttggcag agaaaacacg 20

<210> 7
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador DMD_46-47dup-RV

<400> 7
tatcgctttg cctacgctct 20



- ②① N.º solicitud: 201630599
②② Fecha de presentación de la solicitud: 09.05.2016
③② Fecha de prioridad:

INFORME SOBRE EL ESTADO DE LA TECNICA

⑤① Int. Cl.: Ver Hoja Adicional

DOCUMENTOS RELEVANTES

Categoría	⑤⑥ Documentos citados	Reivindicaciones afectadas
A	GB 2364054 A (SMITHKLINE BEECHAM CORP) 16/01/2002, página 3, último párrafo – página 4, primer párrafo; Página 7, segundo párrafo; reivindicación 3 y 4.	1-25
A	US 2011306036 A1 (DAUNER ALLISON <i>et al.</i>) 15/12/2011, Párrafos [0011] y [0012], reivindicación 1.	1-25
A	US 2007148637 A1 (CHANG YA-CHUN <i>et al.</i>) 28/06/2007, Párrafos [0004]-[0007].	1-25
A	US 2013149695 A1 (LEE SEUNG-TAE <i>et al.</i>) 13/06/2013, Párrafo [0010] y reivindicación 1 y 3.	1-25
A	LIBORIO STUPPIA <i>et al.</i> USE OF THE MLPA ASSAY IN THE MOLECULAR DIAGNOSIS OF GENE COPY NUMBER ALTERATIONS IN HUMAN GENETIC DISEASES. . Int. J. Mol. Sci 2012, Vol. 13, Páginas 3245-3276. Página 3249. Página 3249.	1-25

Categoría de los documentos citados

X: de particular relevancia

Y: de particular relevancia combinado con otro/s de la misma categoría

A: refleja el estado de la técnica

O: referido a divulgación no escrita

P: publicado entre la fecha de prioridad y la de presentación de la solicitud

E: documento anterior, pero publicado después de la fecha de presentación de la solicitud

El presente informe ha sido realizado

para todas las reivindicaciones

para las reivindicaciones nº:

Fecha de realización del informe
23.10.2017

Examinador
S. González Peñalba

Página
1/4

CLASIFICACIÓN OBJETO DE LA SOLICITUD

C12Q1/68 (2006.01)

C12N15/10 (2006.01)

G01N33/50 (2006.01)

Documentación mínima buscada (sistema de clasificación seguido de los símbolos de clasificación)

C12N, C12Q, G01N

Bases de datos electrónicas consultadas durante la búsqueda (nombre de la base de datos y, si es posible, términos de búsqueda utilizados)

INVENES, EPODOC, WPI, MEDLINE, BIOSIS, EMBASE, NPL, INTERNET

Fecha de Realización de la Opinión Escrita: 23.10.2017

Declaración

Novedad (Art. 6.1 LP 11/1986)	Reivindicaciones 1-25	SI
	Reivindicaciones	NO
Actividad inventiva (Art. 8.1 LP11/1986)	Reivindicaciones 1-25	SI
	Reivindicaciones	NO

Se considera que la solicitud cumple con el requisito de aplicación industrial. Este requisito fue evaluado durante la fase de examen formal y técnico de la solicitud (Artículo 31.2 Ley 11/1986).

Base de la Opinión.-

La presente opinión se ha realizado sobre la base de la solicitud de patente tal y como se publica.

1. Documentos considerados.-

A continuación se relacionan los documentos pertenecientes al estado de la técnica tomados en consideración para la realización de esta opinión.

Documento	Número Publicación o Identificación	Fecha Publicación
D01	GB 2364054 A (SMITHKLINE BEECHAM CORP)	16.01.2002
D02	US 2011306036 A1 (DAUNER ALLISON <i>et al.</i>)	15.12.2011
D03	US 2007148637 A1 (CHANG YA-CHUN <i>et al.</i>)	28.06.2007
D04	US 2013149695 A1 (LEE SEUNG-TAE <i>et al.</i>)	13.06.2013
D05	LIBORIO STUPPIA <i>et al.</i> USE OF THE MLPA ASSAY IN THE MOLECULAR DIAGNOSIS OF GENE COPY NUMBER ALTERATIONS IN HUMAN GENETIC DISEASES. . Int. J. Mol. Sci , Vol. 13, Páginas 3245-3276.	2012

2. Declaración motivada según los artículos 29.6 y 29.7 del Reglamento de ejecución de la Ley 11/1986, de 20 de marzo, de Patentes sobre la novedad y la actividad inventiva; citas y explicaciones en apoyo de esta declaración**NOVEDAD Y ACTIVIDAD INVENTIVA ARTS. 6 Y 8 DE LA LP.**

La presente solicitud de patente a la vista de los documentos citados del estado de la técnica y tal y como ha sido definida en las reivindicaciones 1 a 25 parece tener novedad y actividad inventiva por no estar incluida en el estado de la técnica analizado ni poder deducirse de éste de un modo evidente por un experto en la materia.

Se han encontrado en el estado de la técnica documentos que hacen referencia al uso de cebadores degenerados para identificar regiones de secuencias de polinucleótidos bacterianos. Así, el documento D01 se refiere a un método para identificar mutaciones, concretamente en la región QRDR. Para ello utiliza un método en el que emplea un cebador directo degenerado y un cebador reverso degenerado. El método amplifica la región QRDR mediante PCR, secuencia dicho producto amplificado para obtener una primera secuencia de polinucleótidos y compara dicha primera secuencia de polinucleótidos con una segunda secuencia de polinucleótidos de un producto amplificado hecho utilizando dicho primer cebador directo degenerado y dicho cebador inverso degenerado, para identificar, de este modo, las diferencias entre dichas secuencias (véase página 3, último párrafo - página 4, primer párrafo). Los cebadores presentan una longitud comprendida entre 10 y 30 nucleótidos (véase reivindicación 3) y la etapa de amplificación se lleva a cabo mediante PCR (véase reivindicación 4). La temperatura de hibridación se encuentra comprendida entre 25 y 75 grados centígrados (véase página 7, segundo párrafo).

Otros documentos, como el documento D02, emplean métodos en los que se utilizan cebadores degenerados y cebadores específicos para detectar simultáneamente serotipos de virus, como por ejemplo, cuatro serotipos del virus del dengue, para ello se utilizan los reactivos de LAMP (loop-mediated isothermal amplification) (véase párrafos [0011] y [0012] y reivindicación 1).

Y otros, como el documento D03 divulgan métodos para identificar un virus que utiliza un par de cebadores degenerados que corresponden a regiones altamente conservadas de virus y un par de cebadores específicos de acuerdo con secuencias altamente variables dentro de las regiones conservadas de dichos virus (véase párrafos [0004]-[0007]).

Aunque en dichos documentos se emplean cebadores degenerados y cebadores específicos, los métodos no utilizan un cebador degenerado y un cebador específico que hibride específicamente dentro de la región de la mutación del estudio a una distancia de 0,2-100 kb de cualquiera de sus extremos si la mutación comprende una ganancia de material genético, o fuera de dicha región a una distancia de 0,2-100 Kb de cualquiera de sus extremos si la mutación comprende una pérdida de material genético y además, no se emplean para detectar mutaciones.

Por otro lado, se han encontrado documentos, como el documento D04 que hace referencia a la detección de genes mutantes con una alta sensibilidad y especificidad (véase párrafo [0010]) y que describe que la distancia entre el sitio de mutación y el cebador más cercano a dicho sitio es de 1 a 9 pares de bases (véase reivindicación 3). Pero dicho método (véase reivindicación 1) no se emplea cebadores degenerados.

Por último, se han encontrado documentos para detectar mutaciones en el gen DMD, como por ejemplo, el documento D05 en el que se muestra que la técnica MLPA (Multiplex Ligation-dependent Probe Amplification) es adecuada para la identificación de deleciones y duplicaciones del gen DMD (véase página 3249). En la presente solicitud de patente, los genes estudiados para identificar las mutaciones por el procedimiento reivindicado han sido DMD y FBN1. Pero el método al que hace referencia el documento D02 es diferente al empleado en la presente solicitud de patente.

Por lo tanto, la presente solicitud de patente parece cumplir los requisitos de novedad y actividad inventiva en su reivindicación 1.

Las restantes reivindicaciones 2-25 dependen directa o indirectamente de la primera y han de interpretarse como añadidas a ésta, por lo que parecen tener, también, novedad y actividad inventiva. Por consiguiente las reivindicaciones 1-25 parecen cumplir los requisitos de novedad y actividad inventiva según los artículos 6 y 8 de la LP.