

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 738 176**

21 Número de solicitud: 201830731

51 Int. Cl.:

G16B 20/20 (2009.01)

C12Q 1/6827 (2008.01)

12

SOLICITUD DE PATENTE

A1

22 Fecha de presentación:

20.07.2018

43 Fecha de publicación de la solicitud:

20.01.2020

71 Solicitantes:

BIOARRAY, S.L. (100.0%)
Avda. de la Universidad s/n, Fundación UMH.
Edificio Quorum III
03202 Elche (Alicante) ES

72 Inventor/es:

ALCARAZ MAS, Luis A. y
CASTEJÓN FERNÁNDEZ, Natalia

74 Agente/Representante:

PONS ARIÑO, Ángel

54 Título: **MÉTODO PARA EL ESTUDIO DE MUTACIONES EN EMBRIONES EN PROCESOS DE REPRODUCCIÓN IN VITRO**

57 Resumen:

Método para el estudio de mutaciones en embriones en procesos de reproducción in vitro con la particularidad de que combina las técnicas de detección de Aneuploidía (DPG-A) y el estudio de enfermedades monogénicas en embriones (DGP-M) y que se comprende, un proceso de selección de SNPs donde se toma como entrada los valores de unos n SNPs candidatos ($t_1 \dots t_k$) de cada individuo x, en una región cromosómica de interés y específicamente extraídos para una población objeto de estudio; un proceso de selección se SNPs se evalúan todas las combinaciones de SNPs para obtener un set t mínimo de tagSNPs a partir de la matriz M obtenida en el primer proceso de selección se SNPs; y un proceso de validación in-sílico del panel de tagSNPs obtenido en el segundo proceso.

ES 2 738 176 A1

DESCRIPCIÓN

MÉTODO PARA EL ESTUDIO DE MUTACIONES EN EMBRIONES EN PROCESOS DE REPRODUCCIÓN IN VITRO

5

El objeto de la presente invención es un método para el estudio de mutaciones en embriones en procesos de reproducción *in vitro* con la particularidad de que combina las técnicas de detección de Aneuploidía (DPG-A) y el estudio de enfermedades monogénicas en embriones (DGP-M) de acuerdo con la reivindicación 1.

10

CAMPO DE LA TÉCNICA

La presente invención está referida a un método para el estudio de mutaciones en embriones de parejas sometidas a ciclos de reproducción *in vitro* mediante análisis de SNP (polimorfismos de nucleótido único) por secuenciación masiva que combine la detección de Aneuploidías (DGP-A Diagnóstico Genético Preimplantacional de Aneuploidías) y el estudio de enfermedades monogénicas en embriones (DGP-M, Diagnóstico Genético Preimplantacional para enfermedades Monogénicas) con una única biopsia.

20

ESTADO DE LA TÉCNICA ANTERIOR

Un polimorfismo de nucleótido único o SNP es una variación en la secuencia de ADN que afecta a una sola base (adenina (A), timina (T), citosina (C) o guanina (G)) de una secuencia del genoma.

25

El diagnóstico genético preimplantacional (DGP) se desarrolló en la década de los 80 del siglo XX como una opción para aquellas parejas que estaban en riesgo de tener un hijo afectado por una enfermedad monogénica o una alteración cromosómica determinada y que querían evitar la posibilidad de requerir una terminación voluntaria del embarazo.

30

El DGP consiste en la realización de una biopsia de una o más células de los embriones generados durante un ciclo de fecundación *in vitro* cuando los embriones tienen entre 3 y 5 días, para utilizar dicho material en la elaboración de un diagnóstico genético. Así pues, sólo aquellos embriones diagnosticados como no afectados por una determinada alteración genética son transferidos para poder engendrar un niño sano.

35

Cabe destacar que el DGP tiene dos características que lo diferencian del diagnóstico genético aplicado a cualquier otro ámbito. En primer lugar, el tiempo de respuesta debe ser mucho menor, dado que, en muchas ocasiones, es necesario obtener resultados en menos de 24 horas para permitir la transferencia de embriones en el mismo ciclo. Así, por ejemplo, un embrión biopsiado en el día 3 debe ser transferido o vitrificado en el día 5 ó 6. Por otro lado, normalmente, cada pareja produce una media de 6 a 10 embriones, por lo que el coste del análisis debe ser bajo para poder analizar todos los embriones y que esto no suponga un incremento sustancial de lo que ya cuesta de por sí un ciclo de reproducción *in vitro*.

Debido a lo anterior, el DGP se realiza utilizando múltiples y variadas técnicas, dependiendo de la naturaleza de la alteración estudiada. Tradicionalmente, para las enfermedades monogénicas se utilizaban métodos basados en la reacción en cadena de la polimerasa (PCR) y análisis de fragmentos y, más recientemente, sistemas como el *Karyomapping* basado en la utilización de *microarrays* aplicados a la detección de SNP (por ejemplo, como se describe en el documento ES2360085T3). Aunque se trata de métodos muy diversos, todos ellos tienen características en común:

En primer lugar, requieren de una amplificación previa. En este caso se trabaja con muy poco material genético de partida -en muchas ocasiones, procedente de una única célula- por lo que es necesario amplificarlo. Para ello, tradicionalmente, se utiliza un método conocido como MDA (*Multiple Displacement Amplification*) que amplifica el material genético gracias a la polimerasa Phi29 a temperatura constante. Este método produce fragmentos de ADN largos con una baja tasa de error.

En segundo lugar, tienen una alta tasa de *Allele-Drop-Out* (ADO). El ADO es un artefacto frecuente en este tipo de análisis -sobre un 5% de los análisis lo tienen- que consiste en la amplificación preferencial de uno de los alelos. Por este motivo, si al realizar un análisis no se detecta mutación, cabe la posibilidad de que se deba a que el alelo mutado no se ha amplificado.

En tercer lugar, se realiza un análisis indirecto. Como consecuencia de lo anterior, siempre se hace un análisis indirecto de la mutación. Este análisis consiste en estudiar una serie de polimorfismos alrededor de la mutación -generalmente STR o *Short Tandem Repeats*- y determinar si los polimorfismos presentes en el embrión son los asociados al alelo patológico o al alelo sano, es decir, se realiza un análisis de ligamiento.

En cuarto lugar, requieren de un estudio de informatividad previo. Para poder determinar qué polimorfismos segregan con el alelo sano y cuáles con el patológico, es necesario realizar un estudio de informatividad previo. Este estudio consiste en analizar de 5 a 10 STR cercanos a la mutación -situados a ambos lados- en la pareja y en otros miembros de la familia. Idealmente, un hijo en común con la mutación, o bien padres y madres de la pareja. Así, por ejemplo, si se puede asociar que un número determinado de repeticiones de un STR concreto esté siempre presente en familiares con la mutación patogénica, se podrá asociar dicho polimorfismo al alelo patológico y descartar aquellos embriones que lo posean.

Finalmente, el análisis directo de la mutación no siempre es posible. Siempre que sea posible, se realiza el análisis directo de la mutación de los embriones, generalmente, mediante mini secuenciación. Sin embargo, esto sólo es posible si se trata de una mutación puntual. Con este tipo de técnicas, no es posible detectar otro tipo de alteraciones, como deleciones. Por su parte, *Karyomapping* no es capaz de detectar la mutación directa en ningún caso.

Recientemente, la secuenciación masiva (NGS) se ha incorporado como técnica al DGP - como se describe en los documentos US20140274741, WO2014082032, US20150038337 o EP2947156-. No obstante, en el estado de la técnica sólo se describe su aplicación a la detección de aneuploidías (DGP-A).

Si bien la tecnología ha ido evolucionando en el perfeccionamiento de estas técnicas de DGP, existen actualmente ciertas limitaciones en la técnica de análisis. Las principales técnicas y sus limitaciones en DGP son:

Microarray CGH (Hibridación Genómica Comparativa). Esta técnica se utiliza para la detección de aneuploidías (DGP-A) y de alteraciones desbalanceadas cuando uno de los padres es portador de una alteración balanceada. Las principales limitaciones de la técnica son:

Se trata de una técnica cara, de forma que muchas de las parejas que realmente se beneficiarían de ella -por ejemplo, madres de edad avanzada, dado que un número alto de sus ovocitos serán aneuploides- no pueden optar a ellas.

Se utiliza un método de amplificación basado en PCR, que tiene un porcentaje de ADO mucho mayor que los métodos basados en MDA, lo que dificulta la integración con técnicas para el estudio de enfermedades monogénicas, como se ha indicado anteriormente.

5

Requiere un equipamiento específico, tal como un escáner de microarrays.

No permite el análisis de mutaciones, por lo que no puede ser utilizado para seleccionar embriones libres de patología genética, cuando los padres son portadores.

10

No permite distinguir entre embriones normales y aquellos que portan una translocación balanceada.

15

No permite identificar aquellos embriones que portan anomalías numéricas en mosaico.

Mini secuenciación. Se utiliza para la detección directa de una mutación específica en el estudio de enfermedades monogénicas en embriones (DGP-M). Las principales limitaciones son:

20

Requiere el diseño de oligos específicos en una región muy concreta, lo que puede dificultar el análisis.

25

Requiere amplificación previa mediante MDA, por lo que no es fácilmente combinable con técnicas para el cribado de aneuploidías.

Se necesita conocimiento previo de la mutación a analizar, y de un proceso largo de puesta a punto.

30

Requiere de un secuenciador capilar.

No permite detectar anomalías estructurales de ningún tipo.

35

Sólo es útil en el caso de mutaciones puntuales.

Análisis de fragmentos. Se usa sólo o en conjunto con el anterior, para hacer el estudio indirecto de las patologías, por lo que las limitaciones son las mismas que en el caso de la mini-secuenciación, además de:

5 Requiere el diseño de oligos específicos marcados fluorescentemente para diseñar el estudio de informatividad previo.

En ocasiones, el estudio de informatividad se puede alargar debido a la dificultad de encontrar STRs informativos.

10

Karyomapping. Es una técnica que analiza mediante microarray miles de polimorfismos, aunando el análisis de aneuploidías y el estudio de enfermedades monogénicas. La gran ventaja, es que el mismo microarray sirve para analizar distintas patologías. Sus limitaciones son las siguientes:

15

Aunque detecta aneuploidías, no es capaz de detectar errores mitóticos, lo que supone que no es capaz de determinar la presencia de mosaicismo.

20

Requiere de muestras en trío (padre, madre y niño previo afecto) para determinar la segregación de los alelos. Esto es debido a que únicamente hace estudio indirecto.

No es posible detectar la mutación en sí misma en ningún caso, con lo que el riesgo de recombinación nunca queda excluido.

25

Es cara.

El protocolo es largo, no pudiéndose utilizar para ciclos con transferencia en fresco.

30

El documento ES2360085T3 describe que el análisis cromosómico por cariotipado molecular (por ejemplo, para la detección de trisomía) se puede llevar a cabo usando un análisis de marcadores bialélicos del genoma completo (p. ej., polimorfismos de nucleótido único (SNP) bialélicos) que están distribuidos por el genoma, y que se pueden detectar fácilmente usando tecnologías existentes. Este descubrimiento es inesperado por varias razones, principalmente porque a priori se supondría que un marcador bialélico (que proporciona solo

35

información binaria en una posición dada en el cromosoma) no podría identificar positivamente la presencia de tres o más cromosomas diferentes.

No obstante, este documento realiza un análisis de alta densidad de SNP adyacentes
5 cercanos es capaz de identificar positivamente, entre otros, la presencia de dos cromosomas derivados de un progenitor y que basado en suposiciones bien establecidas sobre la frecuencia y espaciado de sucesos de recombinación entre cromosomas parentales durante la meiosis, esto permitirá la detección precisa de trisomía. Además, se identifica en cada caso el origen parental del error, lo cual no es posible mediante algunos
10 procedimientos de cariotipado.

Sin embargo, todavía no se ha podido establecer con éxito un método rápido, eficaz y económico que permita combinar DGP-A y DGP-M con una única biopsia. Por tanto, el avance de la técnica va en el sentido del perfeccionamiento de las herramientas DGP-A y
15 DGP-M con una única biopsia mediante secuenciación masiva, línea principal en el desarrollo del presente proyecto.

EXPLICACIÓN DE LA INVENCION

Es un objeto de la presente invención un método que combine técnicas DGP-A y DGP-M
20 con una única biopsia mediante secuenciación masiva. Este objeto se alcanza con el método de la reivindicación 1. En las reivindicaciones dependientes se muestran realizaciones particulares del método de la invención. En otros aspectos se reivindica el kit que incluye un dispositivo electrónico que ejecuta el método de la invención así como el
25 producto de software que contiene las instrucciones ejecutables para llevar a cabo el método de la presente invención.

Para la combinación de ambas técnicas, primero se realiza la preparación de una librería para DGP-A. En la preparación de esta librería se puede utilizar cualquier kit comercial
30 disponible. como, por ejemplo, *Ion Reproseq (ThermoFisher)*, *PicoPlex (Rubicon Genomics)*, *Veriseq (Illumina)* y *Repli-G (Qiagen)*.

Posteriormente, se toma una alícuota de ese ADN amplificado, y se aplica un método para enriquecer los SNPs que sean interesantes. Preferentemente, este método está basado en
35 amplificación multiplex, pero también se pueden utilizar métodos de captura, PCR simple, o cualquier otro. Tras la amplificación de las regiones de interés, se procede a la preparación

de librerías, añadiendo los adaptadores y códigos de barras necesarios. Así, se tendrá la librería para DGP-M.

Finalmente, tras la cuantificación de ambas librerías, se combinan en unas proporciones
5 determinadas, y se prosigue con el protocolo estándar de secuenciación para la plataforma
elegida. Como se utilizan métodos de preparación de librerías diferentes para ambos
procesos, cada uno de ellos con su código de barras, el secuenciador arroja los resultados
de DGP-A y DGP-M por separado. Así, el análisis se realiza de manera independiente para
DGP-A, utilizando la solución apropiada según el método de preparación de librería y
10 secuenciador seleccionado, mientras que el análisis de DGP-M se realiza haciendo un
fasado de SNPs, como se describe posteriormente.

Este protocolo, es más rápido que otras tecnologías como Karyomapping (ES2360085T3)
porque sólo añade cuatro horas (tiempo de preparación de la librería de DGP-M) al proceso
15 global de DGP-A mediante secuenciación masiva. De esta manera, si se emplea, por
ejemplo, la combinación de *Ion Rerproseq* para DGP-A junto con el método de la invención,
el proceso completo se puede realizar en menos de doce horas, siendo, por tanto, un
método ideal para realizar biopsia en D-3 y transferencia en D-5, pero también para biopsia
en D-5 y transferencia en D-6 y, obviamente, para ciclos con transferencia en diferido debido
20 a que en el protocolo existen múltiples pasos donde se puede parar.

Este método es también más económico que los métodos basados en cariotipado o
cariomapeado (ES2360085T3) porque el precio depende del coste de la preparación de
librería para DGP-A. Además, es adaptable al número de muestras que se tenga para
25 analizar, mientras que en el cariomapeado se utiliza un slide de 12 arrays, por lo que el
número de muestras debe ser múltiplo de 12 para maximizar resultados.

Por ejemplo, se puede preparar una librería para DGP-A utilizando Ion Reproseq. Si la
pareja tiene, por ejemplo, 8 embriones, se habrán preparado 8 librerías diferentes con los
30 códigos de barras correspondientes, del 1 al 8. Tras esta librería, se toma una alícuota y se
amplifican los polimorfismos correspondientes. Si la pareja es portadora, por ejemplo, de
mutaciones en el gen SMN1, se utilizará un kit que amplifica determinados polimorfismos
alrededor de esa mutación. Tras la amplificación, se preparará la librería utilizando el
método correspondiente. Un método podría ser amplificación mediante PCR multiplex
35 utilizando Ampliseq, con su correspondiente kit de preparación de librería, y añadiendo unos
códigos de barras distintos a los anteriores (por ejemplo, del 9 al 16). Se cuantifican ambas

librerías, y se mezclan en una proporción 3:1, es decir, tres veces más librería DGP-A que librería DGP-M. Tras la cuantificación, si el método de secuenciación seleccionado es, por ejemplo, Ion PGM, se procede a la preparación de la secuenciación y a la secuenciación en sí misma, utilizando el protocolo estándar del fabricante del equipo. Una vez finalizada la
 5 secuenciación, el secuenciador nos arrojará 16 ficheros, 8 para la librería DGP-A y 8 para DGP-M. Los ficheros para DGP-A se analizarán utilizando el software del fabricante del equipo, o cualquier otra solución. Para el análisis DGP-M se realizará el fasado de SNPs utilizando el software que se propone en la presente invención. La FIG. 1 muestra un esquema de todo el proceso.

10 Gracias al método así descrito es posible la combinación de técnicas DGP-A y DGP-M mediante secuenciación masiva. Además, es considerablemente más rápido que las tecnologías basadas en cariotipado o cariomapeado (*Karyomapping*) además de resultar más económico.

15 A lo largo de la descripción y de las reivindicaciones, la palabra «comprende» y sus variantes no pretenden excluir otras características técnicas, aditivos, componentes o pasos. Para los expertos en la materia, otros objetos, ventajas y características de la invención se desprenderán en parte de la invención y en parte de la práctica de la invención. Los
 20 siguientes ejemplos y dibujos se proporcionan a modo de ilustración y no se pretende que restrinjan la presente invención. Además, la invención cubre todas las posibles combinaciones de realizaciones particulares y preferidas aquí indicadas.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

25 A continuación, se pasa a describir de manera muy breve una serie de dibujos que ayudan a comprender mejor la invención y que se relacionan expresamente con una realización de dicha invención, que se ilustra como un ejemplo no limitativo de ésta.

30 La **FIG.1** muestra un esquema del método objeto de la presente invención. En primer lugar, se analiza el gen con la mutación a descartar en los embriones y se selecciona una región alrededor (típicamente, de 4 Mb). Se seleccionan aquellos polimorfismos que no cumplen ningún criterio de exclusión, y se buscan aquellos con mayor probabilidad de ser
 35 informativos. Por otro lado, se identifican aquellos polimorfismos que pueden ser *tag* mediante análisis de correlación y desequilibrio de ligamiento. Con ambas cosas, se realiza

un diseño final de panel, y se hace una validación *in silico* simulando entrecruzamientos entre multitud de individuos cuyos genomas se encuentran depositados en bases de datos públicas.

5 La **FIG.2** muestra un esquema de todo el proceso, incluyendo el método de la invención, para un primer ejemplo de uso, empezando por el diseño y síntesis del panel que luego se utilizará para la amplificación de la librería para DGP-M. Por un lado, los pacientes y otros familiares son analizados para determinar la distribución de los polimorfismos en los distintos alelos. En cuanto a las muestras de los embriones, tras la biopsia, se prepara la
10 librería para DGP-A, y con una alícuota de esta se prepara la librería para DGP-M. Se cuantifican y mezclan todas las librerías, y se procede a la secuenciación. Finalmente, se realiza el análisis bioinformático de aneuploidías y de la enfermedad monogénica, de manera independiente.

15 La **FIG.3** muestra un ejemplo de fasado de SNPs, en el caso de un trío formado por la pareja y un hijo afecto. Se muestran los SNPs antes y después de fasar, con la distribución de los mismos en los alelos. Además, se señala cuáles son informativos (dan información sobre la fase) y cuáles no.

20 La **FIG.4** muestra un ejemplo de fasado de SNPs con análisis de resultados en embriones. Para simplificar, se muestran sólo aquellos SNPs que son informativos. En este caso, la pareja es portadora de mutaciones en heterocigosis en el gen *CEP290*, que provoca síndrome de Meckel. Este síndrome es letal prenatal, por lo que para el fasado se utilizó ADN de un feto anterior. En este caso, el panel se diseñó incluyendo el análisis directo de la
25 mutación, que aparece sombreada.

La **FIG.5** muestra el resultado del fasado de SNPs en un embrión, donde aparecen distintos problemas y artefactos, junto con la descripción de estos. El algoritmo de fasado desarrollado permite detectar estos errores.

30

La **FIG.6** muestra un ejemplo de una pareja con una translocación balanceada, y los embriones que como consecuencia se pueden generar. El 50% de estos embriones pueden heredar una alteración desbalanceada con graves consecuencias (desde abortos de repetición hasta niños con retraso mental y rasgos dismórficos). El 25% de los embriones
35 serán normales, y el otro 25% tendrán la alteración balanceada como uno de los progenitores.

EXPLICACIÓN DE UN MODO DETALLADO DE REALIZACIÓN DE LA INVENCION

El método para el estudio de mutaciones en embriones en procesos de reproducción in vitro, objeto de la presente invención puede dividirse, a su vez, en tres procesos secuenciales y diferenciados. El método de la invención combina varias técnicas de selección de tagSNPs, pues calcula las correlaciones de desequilibrio de ligamiento para el bloque de interés (por defecto, 4 Mb alrededor de la mutación, 1 Mb es igual a un millón de nucleótidos). Los SNPs de esa región, a su vez, serán considerados como en un enfoque libre de bloques, de manera que todas las correlaciones entre SNPs son calculadas y tenidas en cuenta en la selección de los tagSNPs. De esta forma la presente invención selecciona polimorfismos con alta probabilidad de ser informativos, considerando las frecuencias alélicas de los mismos dentro de la población objetivo y si forman parte o no de un mismo haplobloque (conjunto de SNPs que se heredan juntos). Un tagSNP es aquel SNP que se considera representativo para todo el haplobloque, es decir, que si el tagSNP está en heterocigosis, por ejemplo, todo los SNPs pertenecientes a ese mismo haplobloque estarán en heterocigosis. Un tagSNP evita tener que analizar todos los polimorfismos porque conociendo cómo se comporta, se puede deducir cómo lo hace el resto de polimorfismos del haplobloque.

Por lo tanto, el objetivo del método de la invención (**FIG.1**) se centra en la obtención de un panel mínimo de tagSNPs con máxima capacidad informativa, simplificando el posterior análisis e interpretación de los resultados de un análisis de informatividad.

Más concretamente, el método de la invención comprende se divide en dos procesos básicos y un tercer proceso de validación. Los dos primeros procesos no tienen porqué ser, necesariamente, en este orden:

- (i) Un primer proceso de selección de SNPs:
 - a. En este proceso se toma como entrada los valores de los n SNPs candidatos ($t_1 \dots t_k$) de cada individuo x , en una región cromosómica de interés y específicamente extraídos para la población objeto de estudio. Se seleccionan aquellos SNPS que son bialélicos, por lo que los individuos pueden ser representados como haplotipos de longitud m formados por cadenas binarias $\{1,0\}$, siendo 1|0 y 0|1 los valores para SNPs heterocigotos y 0|0 y 1|1 los valores para los homocigotos. Esto se realiza en toda la región cromosómica de interés que queda definida como toda posición que se

encuentre 2 Mb (es un valor por defecto que es posible modificar) aguas arriba y 2 Mb aguas abajo del gen/mutación que se desee estudiar.

b. Tras la obtención, se analizan los n SNPs candidatos de la región y se excluyen aquellos que cumplan alguna de las siguientes condiciones:

- 5 i. SNPs con más de un alelo alternativo (SNPs no bialélicos)
- ii. SNPs cuyos alelos sean diferentes al cambio de un nucleótido simple (indels, cambios de patrón de polinucleótidos, entre otros)
- iii. SNPs que se encuentren en homocigosis en al menos el 99% de la población de interés
- 10 iv. SNPs no comunes, es decir, cuya frecuencia alélica menor sea inferior al 1%

c. El siguiente punto trata de maximizar la situación en que uno de los parentales presente el valor de un SNP en estado heterocigoto, mientras en el otro parental se presente como homocigoto, es decir, sea informativo. Esto se consigue a través de la maximización del valor de dos funciones por encima de cierto valor umbral:

$$\text{MaxP: } p - (3p^2) + (4p^3) - (2p^4)$$

$$\text{HET rate: } 2pq$$

20 siendo p y q respectivamente las frecuencias alélicas de los alelos referencia y alternativo para cada SNP. Estas son las ecuaciones del equilibrio de Hardy-Weinberg y su derivada.

d. La salida de este algoritmo será un panel de z SNPs optimizados para ambos valores en forma de *matriz* M cuyas columnas corresponden a los individuos de la población y las filas a los valores de cada SNP para cada individuo.

(ii) Un segundo proceso de selección de SNPs:

a. Mediante una búsqueda exhaustiva, se evalúan todas las combinaciones de SNPs para obtener un set t mínimo de tagSNPs a partir de la matriz M obtenida en el punto (i).

b. En primer lugar los SNPs de la matriz M del block-región se organizan en grupos de alta correlación en base al criterio del pairwise r^2 . Para ello se calcula el valor de pairwise r^2 a partir de la frecuencia alélica calculada para la matriz M . De esta manera SNPs de diferentes grupos presentarán baja correlación, por lo que dos SNPs pertenecerán a un mismo grupo tan solo cuando el r^2 pairwise entre ellos supere cierto valor umbral (fijado por el

usuario).

- c. Tras esto se realiza la selección de tagSNPs dentro de cada grupo en base al criterio de LD, empezando por $k=1$ SNPs y estudiando todas las k -combinaciones posibles, organizando los SNPs dentro de cada grupo.
- d. Suponiendo dos SNPs cuyas frecuencias sean p (alelo de mayor frecuencia) y $q=1-p$. Se emplean las ecuaciones:

$$D = pAB - pApB \qquad D' = D/D_{max} \qquad \left\{ \begin{array}{l} D > 0 \rightarrow D_{max} = \min(pApb, papB) \\ D < 0 \rightarrow D_{max} = \max(-pApB, -papb) \end{array} \right.$$

Siendo pA y pB la probabilidad observada del alelo p para el SNP1 y el SNP2 respectivamente y pa y pb la probabilidad para el alelo menor. Finalmente, pAB es la probabilidad combinada del par pq . Mediante el uso de estas ecuaciones se considerarán tagSNPs aquellos SNPs cuyos rangos funcionales superen cierto valor umbral.

La organización de los SNPs dentro de cada grupo en base al criterio de LD permite realizar una selección en base al rango funcional, asegurando que la primera solución encontrada es una solución óptima y permitiendo reducir considerablemente el tiempo de computación.

Finalmente, si un SNP no supera los umbrales de r^2 o LD será considerado en un grupo sólo y tomado como tagSNP por sí mismo.

- (iii) Un tercer proceso de validación que consiste en una validación in-silico del panel de tagSNPs obtenido Para ello y empleando la base de datos *1000Genomes db* [1], se escogen aleatoriamente individuos para realizar múltiples cruces. Tras esto se contabiliza el número de tagSNPS que resultaron informativos de cada cruce y se aporta el promedio a modo de dato informativo de la potencia informativa del panel.

Como se ha indicado, la **FIG.2** muestra un esquema del método de la invención, en donde en primer lugar, se analiza el gen con la mutación a descartar en los embriones y se selecciona una región alrededor (típicamente, de 4 Mb). Se seleccionan aquellos polimorfismos que no cumplen ningún criterio de exclusión, y se buscan aquellos con mayor

probabilidad de ser informativos. Por otro lado, se identifican aquellos polimorfismos que pueden ser *tag* mediante análisis de correlación y desequilibrio de ligamiento. Con ambas cosas, se realiza un diseño final de panel, y se hace una validación *in silico* simulando entrecruzamientos entre multitud de individuos cuyos genomas se encuentran depositados en bases de datos públicas.

Ejemplo 1. Diagnóstico de aneuploidías

En la **FIG.2** se muestra el esquema completo para el uso en el diagnóstico de aneuploidías. Se considera una pareja de población europea donde un miembro de la pareja es portador de la variante patogénica autosómica dominante *VHL:c.233A>G p.(Asn78Ser)* (posición cromosómica *chr13:10183764*) causante de una afección conocida como síndrome de Von Hippel-Lindau, que tiene un modo de herencia autosómico dominante.

El input del software serán los 69473 SNPs que contiene el *block-region chr13:9181319-11681319*. El output de este algoritmo será una matriz M de 1625 SNPs candidatos, los cuales actuarán como input para el algoritmo de selección de SNPs, cuyo output estará formado por un panel de 283 tagSNPs. En la fase de validación se obtuvo que el 49% de tagSNPs del panel resultaba informativo en promedio.

Parte de laboratorio húmedo. Una vez seleccionados los polimorfismos que queremos secuenciar, se introducen las posiciones en la plataforma correspondiente de enriquecimiento. Preferiblemente, Ion Ampliseq. Esta plataforma diseña los oligos necesarios para la captura de las regiones. En el laboratorio de IVF se generan los embriones que se biopsian cuando llegan a la etapa de blastocisto. La biopsia se mete en un tubo de PCR y se envía al laboratorio. En el laboratorio, se amplifica el ADN utilizando, por ejemplo, Ion Reproseq, de manera que, además de la amplificación, se hace la librería para DGP-A. Siguiendo el protocolo adecuado, se amplifican todas las regiones diseñadas con Ampliseq en el material previamente amplificado, y se produce la librería para DGP-M. Posteriormente, hacemos la secuenciación masiva. Hay que tener en cuenta que para poder secuenciar de manera simultánea múltiples muestras, es necesario marcar las muestras con un código de barras molecular. Se debe tener especial precaución en que los códigos de barras no coincidan entre las muestras.

Análisis de datos. Una vez finalizada la secuenciación, se obtienen una serie de ficheros con lecturas para todo el embrión (DGP-A) y otros ficheros con los polimorfismos detectados

(DGP-M). Con estos ficheros, se debe hacer un análisis bioinformático. Con los primeros, se determinan las aneuploidías utilizando el software más apropiado según la plataforma. Con los segundos, se determina el patrón de segregación de cada polimorfismo para el análisis de DGP-M. En primer lugar, las lecturas obtenidas se alinean al genoma de referencia, y se identifican los polimorfismos en todas y cada una de las muestras, incluidas pacientes, familiares utilizados como referencia y embriones. En el caso de los familiares, la situación más sencilla es aquella en la que disponemos de la pareja y de un hijo afecto. Para el fasado de SNPs, debemos determinar qué polimorfismos se comparten en el trío y de esta manera averiguar cuáles segregan con el alelo sano y cuáles con el patogénico. Como se han seleccionado SNPs bialélicos, cada una de las muestras puede ser 0/0, 0/1, 1/1 si son homocigotas para el SNP silvestre, heterocigotas u homocigotas para el SNP alternativo, respectivamente. El número 0 está indicando que se trata del SNP de referencia (sea cual fuere) mientras que el 1 indica que se trata del SNP alternativo. Esto es así para todos los cromosomas, excepto los sexuales, en los cuales las mujeres pueden ser homocigotas u heterocigotas, mientras que los hombres son siempre hemizogotos (0 ó 1). Con estos datos, se procede a realizar el fasado de SNPs. Para ello, se analizan aquellos SNPs que son informativos en la pareja. Los SNPs informativos son aquellos en los cuales uno de los pacientes es heterocigoto (0/1) y el otro homocigoto (0/0 o 1/1). El polimorfismo que se utiliza para casar es aquel que es distinto en el individuo heterocigoto. Por ejemplo, si tenemos un individuo 0/1 y el otro 1/1, el polimorfismo que utilizaremos para fasar es el 0. Por comparación con uno o más individuos en la familia, se determina arbitrariamente a qué alelo pertenece cada uno. Un ejemplo de fasado de SNPs sería el siguiente:

Se considera que la situación en la que tenemos una pareja con un hijo, cada uno con sus alelos. Consideramos que el padre tiene los alelos P1 y P2, la madre M1 y M2. Lógicamente, el hijo habrá heredado un alelo de cada uno, por ejemplo, P1 del padre y M1 de la madre. Si al analizar unos polimorfismos, se obtienen que el padre es 0/1, la madre 1/1 y el hijo 0/1, determinaremos que el polimorfismo 0 pertenece al alelo P1, que es el compartido entre el padre y el hijo, dado que ambos presentan dicho polimorfismo. En otro caso, podemos tener el padre 0/1, la madre 1/1 y el hijo 1/1. En ese caso, el polimorfismo 0 debe necesariamente pertenecer al alelo P2 del padre, puesto que no es compartido entre padre e hijo. De manera análoga, se procede con todos los polimorfismos que sean informativos para la madre.

Una vez determinados los haplotipos de los padres, se procede al análisis de los embriones. Para ello, se trata de identificar el SNP informativo en cada uno de los

embriones. Por ejemplo, si el polimorfismo 0 pertenece al alelo P1, si un embrión presenta dicho polimorfismo, significa que tiene dicho haplotipo. En definitiva, se trata de identificar los SNPs informativos en los distintos embriones y así determinar el haplotipo de cada uno de ellos.

5

Un ejemplo del proceso de fasado de SNPs se muestra en la **FIG.3**, donde se especifica, además, si los SNPs son informativos, no informativos o semiinformativo. Un SNP informativo es aquel que cumple lo anteriormente dicho, es decir, que sea heterocigoto en uno de los progenitores, y homocigoto en el otro; un SNP no informativo puede darse
10 cuando ambos progenitores sean homocigotos, o que ambos sean heterocigotos y el hijo también lo sea; un SNP semiinformativo se da cuando ambos progenitores son heterocigotos y el hijo es homocigoto. Esta clasificación se puede extender a otras combinaciones familiares.

15 Una vez hecho el fasado de SNP con familiares, se debe comparar el patrón de polimorfismos con los embriones secuenciados, y de esta manera determinaremos que embriones son portadores y cuáles son normales. En la **FIG.4** se muestra un ejemplo de fasado de SNPs que incluye el análisis de los embriones.

20 El algoritmo de fasado de SNPs es, además, capaz de identificar las distintas posibles fuentes de error y poner sobre alerta al analista para que las pueda sopesar y analizar. Estas fuentes de error son diversas. En la **FIG.5** se muestra el resultado para un embrión similar a la **FIG.4**, pero donde se identifican y describen las distintas fuentes de error:

25 Por ejemplo, podemos tener un fenómeno de allele-drop-out. Este fenómeno implica que, en un embrión, sólo se amplifica uno de los alelos. De manera que, al secuenciar, podemos interpretar que un embrión es homocigoto para un polimorfismo, cuando en realidad es heterocigoto. Por ejemplo, un embrión puede ser 0/1 pero haber sufrido un fenómeno de ADO durante la amplificación y mostrarse
30 como 1/1 en la secuenciación. Este fallo puede dar lugar a una mala interpretación de los resultados, asignando equívocamente el alelo de los embriones. A fin de evitarlo, es necesario distinguir aquellos polimorfismos que, además de ser informativos, son *clave*. Los polimorfismos clave son aquellos que además de ser informativos son heterocigotas en el embrión. Por ejemplo, si tenemos un padre 0/1 y
35 una madre 1/1, y el embrión en 0/1, el polimorfismo 0 será clave porque es informativo y heterocigoto. Por el contrario, si el embrión es 1/1, el polimorfismo sólo

será informativo. En este último caso no podemos determinar si el embrión es realmente 1/1, o si hemos sufrido un allele-drop-out y sólo estamos viendo uno de los alelos 1.

5 Otra fuente de error se debe a lo que se llama *No Call*. Se produce cuando no se obtienen señal para ninguno de los alelos para un polimorfismo determinado, así que no podemos saber ni siquiera uno de ellos.

10 Finalmente, otra fuente que puede dar lugar a confusión es se produzca una recombinación. Una recombinación es cuando los alelos en uno de los padres se intercambian, y esto se ve reflejado en el embrión. Por ejemplo, si analizamos 100 polimorfismos en un embrión, puede suceder que parte de ellos (por ejemplo, 60) pertenezcan al alelo P1 y el resto al alelo P2. Para poder identificar la recombinación, debe suceder que el cambio de alelo P1 al P2 sea de manera secuencial. Es decir,
15 por ejemplo, que los primeros 60 polimorfismos pertenezcan al alelo P1 y los siguientes al P2. Si el intercambio de polimorfismos se produjera de manera más o menos aleatoria, eso significaría que se trata de un error de la secuenciación, porque estadísticamente no puede suceder que un embrión tenga más de dos recombinaciones en un espacio tan pequeño como el fragmento analizado (4 Mb).

20 También puede suceder que haya errores de secuenciación o artefactos. Estos artefactos se pueden identificar fácilmente porque parecen recombinaciones o allele-drop-outs, pero suceden de manera esporádica en una o muy pocas posiciones.

25 Finalmente, para poder realizar la asignación de alelos de manera inequívoca, se requiere que haya al menos tres polimorfismos informativos y que sean además clave a cada lado de la mutación, de manera consecutiva, además de al menos otros 3 polimorfismos más no clave.

30 **Ejemplo 2. Identificación de embriones triploides**

Los embriones triploides son un problema importante en cualquier ciclo de IVF. Estos suponen un 15% de los abortos espontáneos por anomalías cromosómicas. Los embriones triploides deben descartarse siempre de cualquier ciclo de fecundación *in vitro*, pero es difícil
35 identificarlos porque no se observan diferencias en cuanto a calidad embrionaria con respecto a embriones normales. En ocasiones, es posible distinguirlos porque en D+1 se

observan tres pronúcleos, pero no siempre es posible. Los embriones triploides pueden ser de origen dispérmico (en casos de FIV) o bien originarse por un fallo del ovocito al no realizar la extrusión del segundo corpúsculo polar.

5 Los embriones triploides no se pueden identificar mediante técnicas corrientes de DGP-A, a pesar de tratarse de una anomalía numérica. En ocasiones, mediante inspección visual, es posible detectar embriones 46,XXY al observar una distribución anormal de las lecturas de los cromosomas sexuales, pero no siempre es posible y requiere de personal entrenado.

10 El método aquí descrito se puede utilizar para identificar este tipo de embriones. Se pueden seleccionar polimorfismos informativos a lo largo del genoma y determinar si son triploides analizando los polimorfismos presentes y la frecuencia de los mismos. Normalmente, un polimorfismo en heterocigosis se debe encontrar en una proporción alrededor de 0,5, debido a que la mitad de las lecturas se corresponderán a un alelo y la mitad a otro. Un embrión
15 triploide tiene tres alelos, por lo que esta proporción se verá desviada. Así, el resultado puede ser de tres polimorfismos para la misma posición (si son multialélicos) o bien dos polimorfismos pero uno de ellos con frecuencia sobre el 33% y el otro sobre el 66%. Si todos los polimorfismos con suficientes lecturas siguen este patrón a lo largo de todo el genoma, eso significa que el embrión es triploide.

20

Ejemplo 3. Identificación de embriones con traslocaciones balanceadas

En ocasiones, algunas parejas deciden acudir a ciclos de fecundación *in vitro* porque uno de ellos es portador de una translocación balanceada. En estos casos, estos padres tienen
25 un riesgo reproductivo elevado, porque el 50% de sus embriones tendrán una translocación desbalanceada como consecuencia de heredar uno de los cromosomas alterados. Además, tendrá un 25% de posibilidades de generar embriones completamente normales, y un 25% de generar embriones con la alteración balanceada. En la **FIG.6** se muestra un esquema de los posibles embriones generados. Las técnicas actuales permiten distinguir aquellos
30 embriones con alteraciones desbalanceadas, la mayoría de las veces utilizando simplemente DGP-A. Sin embargo, no es posible diferenciar aquellos embriones con la alteración balanceada de aquellos completamente normales, dado que no se producen cambios de número de copia. Mediante el presente desarrollo, es posible mapear todo el cromosoma mediante distintos polimorfismos y, estudiando la distribución de esos
35 polimorfismos en los embriones desbalanceados, determinar cuáles están presentes en el cromosoma alterado y cuáles en el normal. De esta manera, podremos saber si los

embriones sin cambios de número de copia han recibido de su progenitor el cromosoma normal o el alterado. Este estudio es posible gracias a la combinación de DGP-A y DGP-M.

REIVINDICACIONES

1.- Un método para el estudio de mutaciones en embriones en procesos de reproducción in vitro con la particularidad de que combina las técnicas de detección de Aneuploidía (DPG-A) y el estudio de enfermedades monogénicas en embriones (DGP-M) y que se **caracteriza por que** comprende, a su vez los procesos de:

un proceso de selección de SNPs donde se toma como entrada los valores de unos n SNPs candidatos ($t_1 \dots t_k$) de cada individuo x , en una región cromosómica de interés y específicamente extraídos para una población objeto de estudio; y donde este proceso está configurado para maximizar la situación en que uno de los parentales presente el valor de un SNP en estado heterocigoto, mientras en el otro parental se presente como homocigoto y obtener un panel de z SNPs optimizados para ambos valores maximizados en forma de matriz M cuyas columnas corresponden a los individuos de la población y las filas a los valores de cada SNP para cada individuo;

un proceso de selección se SNPs se evalúan todas las combinaciones de SNPs para obtener un set t mínimo de tagSNPs a partir de la matriz M obtenida en el primer proceso de selección se SNPs; y

un proceso de validación in-sílico del panel de tagSNPs obtenido en el segundo proceso.

2.- El método de acuerdo con la reivindicación 1 donde el primer proceso de selección de SNPs comprende la selección de aquellos SNPs que son bialélicos, por lo que los individuos pueden ser representados como haplotipos de longitud m formados por cadenas binarias $\{1,0\}$, siendo 1|0 y 0|1 los valores para SNPs heterocigotos y 0|0 y 1|1 los valores para los homocigotos; realizándose esta selección en toda la región cromosómica de interés.

3.- El método de acuerdo con la reivindicación 2 donde la región cromosómica de interés queda definida como toda posición que se encuentre dos megabases por encima y dos megabases por debajo del gen o mutación objeto de estudio.

4.- El método de acuerdo con una cualquiera de las reivindicaciones 1 a 3 donde el primer proceso comprende una etapa de analizar los n SNPs candidatos de la región y excluir aquellos que cumplan alguna de las siguientes condiciones: SNPs con más de un alelo alternativo (SNPs no bialélicos); SNPs cuyos alelos sean diferentes al cambio de un nucleótido

simple; SNPs que se encuentren en homocigosis en al menos el 99% de la población de interés; y SNPs no comunes, es decir, cuya frecuencia alélica menor sea inferior al 1%.

5 5.- El método de acuerdo con una cualquiera de las reivindicaciones 1 a 4 donde el primer proceso comprende una etapa de maximizar la situación en que uno de los parentales presente el valor de un SNP en estado heterocigoto, mientras en el otro parental se presente como homocigoto, es decir, que sea informativo a través de la maximización del valor de dos funciones por encima de cierto valor umbral:

10 **MaxP:** $p-(3p^2) + (4p^3) - (2p^4)$

HET rate: $2pq$

siendo p y q respectivamente las frecuencias alélicas de los alelos referencia y alternativo para cada SNP.

15 6.- El método de acuerdo con una cualquiera de las reivindicaciones 1 a 5 donde el segundo proceso de selección de SNPs que comprende, en primer lugar, que los SNPs de la matriz M del block-región se organicen en grupos de alta correlación en base al criterio del pairwise r^2 ; de tal forma que se calcula el valor de pairwise r^2 a partir de la frecuencia alélica calculada para la matriz M .

20

7.- El método de acuerdo con la reivindicación 6 donde los SNPs de diferentes grupos presentarán baja correlación, por lo que dos SNPs pertenecerán a un mismo grupo tan solo cuando el r^2 pairwise entre ellos supere cierto valor umbral fijado por el usuario.

25

8.- El método de acuerdo con una cualquiera de las reivindicaciones 1 a 7 donde la selección de tagSNPs dentro de cada grupo se realiza en base al criterio de límite de la detección (LD), empezando por $k=1$ SNPs y estudiando todas las k -combinaciones posibles, organizando los SNPs dentro de cada grupo.

30

9.- El método de acuerdo con una cualquiera de las reivindicaciones 6 a 8 donde si un SNP no supera los umbrales de r^2 o LD será considerado en un grupo sólo y tomado como tagSNP por sí mismo.

10.- El método de acuerdo con una cualquiera de las reivindicaciones 1 a 9 donde en el tercer proceso de validación se emplea una base de datos genómica donde se escogen aleatoriamente individuos para realizar 300 cruces, tras lo cual se contabiliza el número de tagSNPs que resultaron informativos de cada cruce y se aporta el promedio a modo de dato informativo de la potencia informativa.

11.- Un kit para el estudio de mutaciones en embriones en procesos de reproducción in vitro que se caracteriza porque comprende, al menos, un dispositivo electrónico con un procesador o procesadores y una memoria, donde la memoria almacena instrucciones que cuando son ejecutadas por el procesador o procesadores hacen que el dispositivo electrónico ejecute el método de acuerdo con una cualquiera de las reivindicaciones 1 a 10.

12.- Un producto de programa informático con instrucciones configuradas para ser ejecutadas por uno o más procesadores que hacen que el dispositivo electrónico del kit de la reivindicación 11 lleve a cabo el método de acuerdo con una cualquiera de las reivindicaciones 1 a 10.

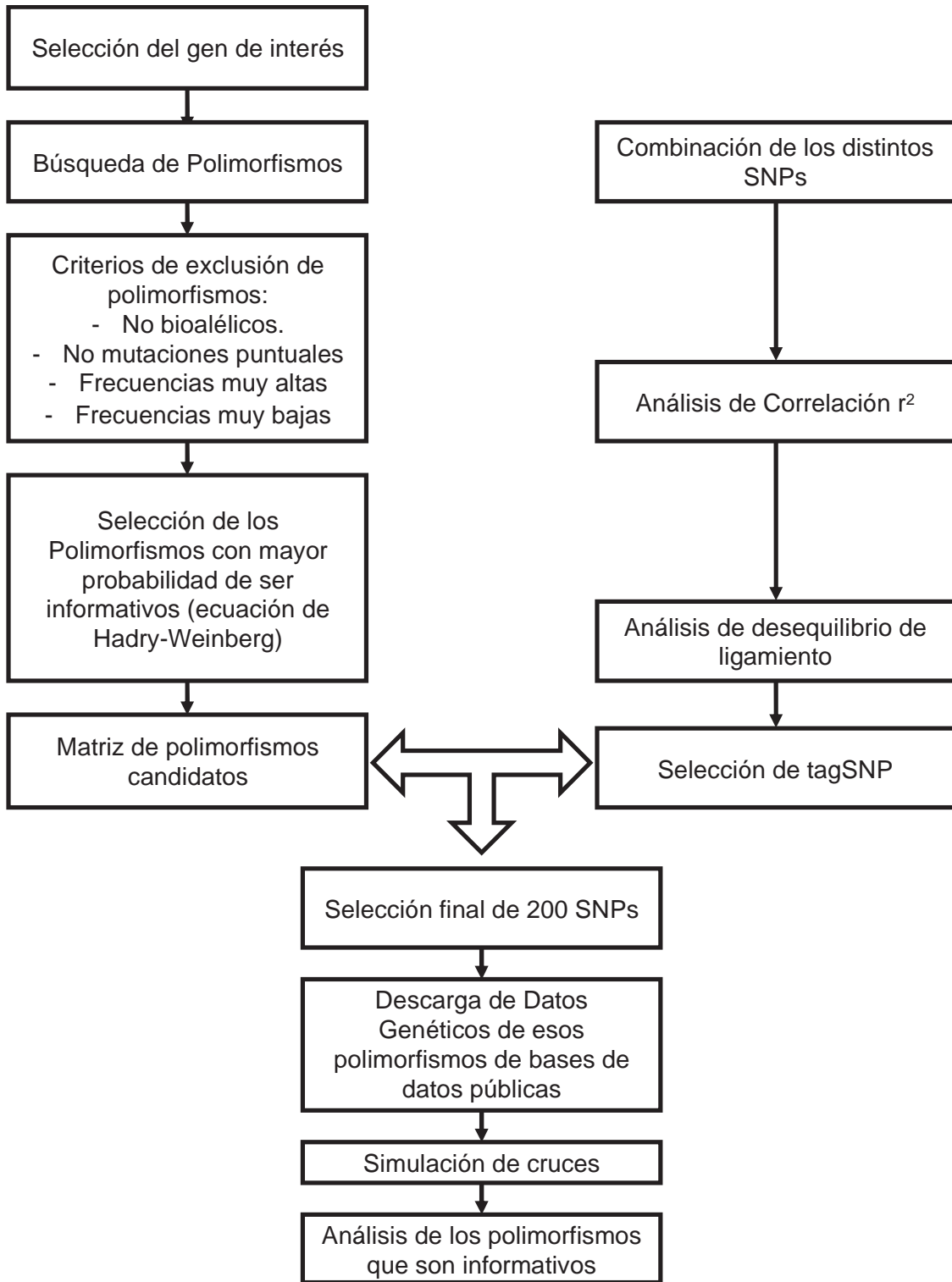


FIG.1

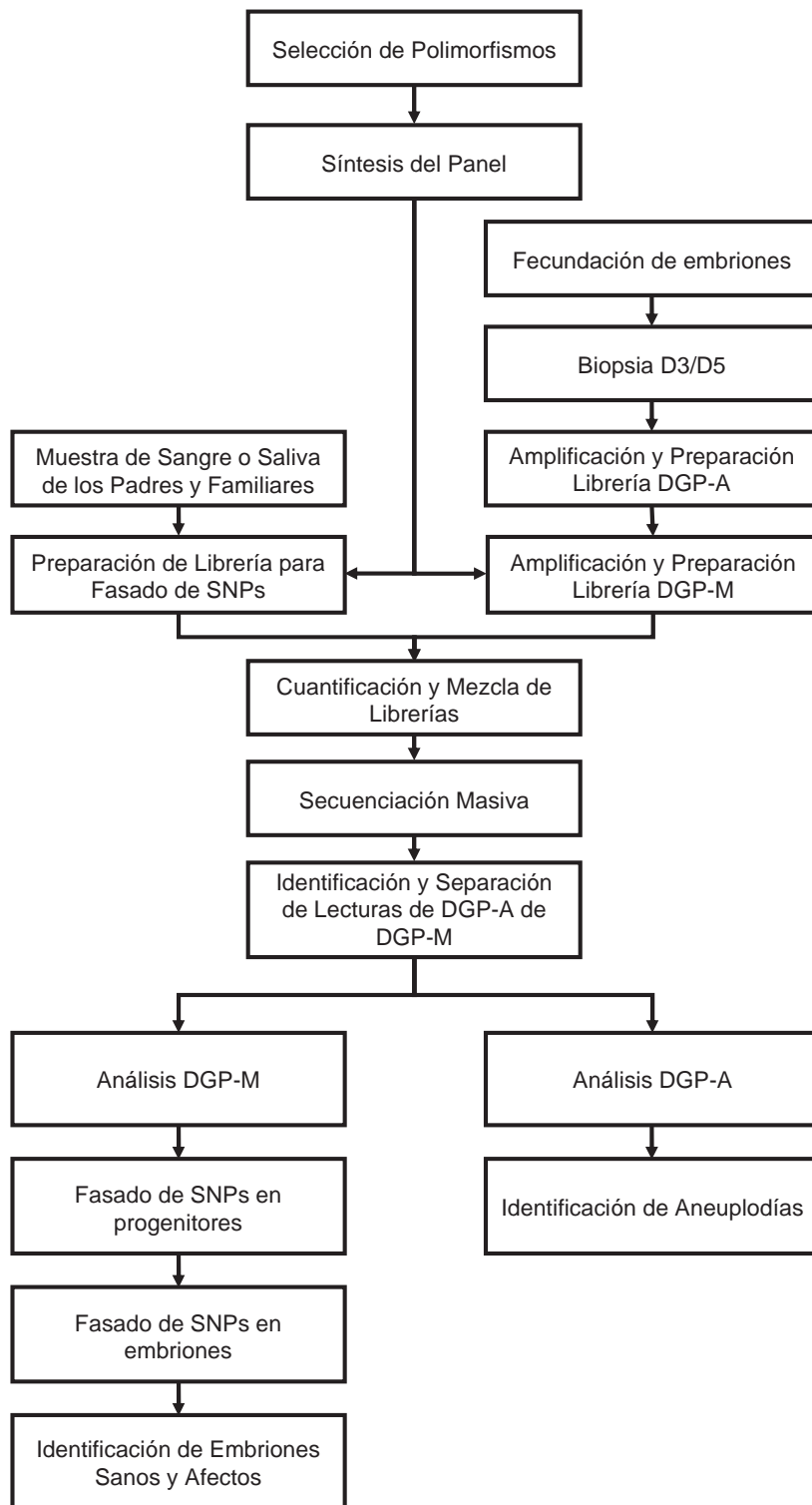


FIG.2

SNP	Sin Fasar			Fasado						Tipo de SNP
	P	M	H	P		M		H		
				P1	P2	M1	M2	P1	M1	
1	0/0	1/1	0/1	0	0	1	1	0	1	No informativo
2	0/1	0/0	1/0	1	0	0	0	1	0	Informativo
3	0/1	0/0	0/0	0	1	0	0	0	0	Informativo
4	1/1	0/1	0/1	1	1	0	1	1	0	Informativo
5	1/1	0/1	1/1	1	1	1	0	1	1	Infomativo
6	0/1	0/1	0/0	0	1	0	1	0	0	Semiinformativo
7	0/1	0/1	0/1	0	1	0	1	-	-	No informativo
9	0/0	0/0	0/0	0	0	0	0	0	0	No informativo
10	1/1	1/1	1/1	1	1	1	1	1	1	No informativo

FIG.3

ES 2 738 176 A1

Posición	Padre		Madre		Feto		Emb. 1		Emb. 2		Emb. 3		Emb. 4	
	Portador		Portadora		Afecto		Afecto		Sano		Portador		Portador	
	P1	P2	M1	M2	P1	M1	P1	M1	P2	M2	P1	M2	P2	M1
chr12:87254962	A	G	A	A	A	A	A	A	G	A	A	A	G	A
chr12:87260006	C	G	C	C	C	C	C	C	G	C	C	C	G	C
chr12:87302691	A	C	A	A	A	A	A	A	C	A	A	A	C	A
chr12:87302747	A	C	A	A	A	A	A	A	C	A	A	A	C	A
chr12:87391646	T	A	T	T	T	T	T	T	A	T	T	T	A	T
chr12:87640519	T	T	T	C	T	T	T	T	T	C	T	C	T	T
chr12:87674929	C	C	C	T	C	C	C	C	C	T	C	T	C	C
chr12:87714749	A	G	A	A	A	A	A	A	G	A	A	A	G	A
chr12:87714817	G	G	G	C	G	G	G	G	G	C	G	C	G	G
chr12:87901028	A	G	A	A	A	A	A	A	G	A	A	A	G	A
chr12:88238333	C	T	C	T	C	C	C	C	T	T	C	T	T	C
chr12:88346589	T	C	T	T	T	T	T	T	C	T	T	T	C	T
chr12:88450462	T	T	T	A	T	T	T	T	T	A	T	A	T	T
chr12:88471567	-	T	-	T	-	-	-	-	T	T	-	T	T	-
chr12:89400654	T	T	T	C	T	T	T	T	T	C	T	C	T	T
chr12:89447754	T	A	T	T	T	T	T	T	A	T	T	T	A	T
chr12:89449487	A	G	A	A	A	A	A	A	G	A	A	A	G	A
chr12:89456904	A	G	A	A	A	A	A	A	G	A	A	A	G	A
chr12:89459695	G	T	G	G	G	G	G	G	T	G	G	G	T	G
chr12:89460093	A	C	A	A	A	A	A	A	C	A	A	A	C	A
chr12:89462738	C	T	T	T	T	T	T	T	T	T	C	T	T	T
chr12:89463041	C	A	C	C	C	C	C	C	A	C	C	C	A	C
chr12:89468283	G	A	G	G	G	G	G	G	A	G	G	G	A	G
chr12:89469954	C	T	C	C	C	C	C	C	T	C	C	C	T	C
chr12:89470751	C	G	C	C	C	C	C	C	G	C	C	C	G	C
chr12:89471403	G	A	G	G	G	G	G	G	A	G	G	G	A	G
chr12:89473626	A	C	A	A	A	A	A	A	C	A	A	A	C	A
chr12:89474352	C	T	C	C	C	C	C	C	T	C	C	C	T	C
chr12:89475779	G	A	G	G	G	G	G	G	A	G	G	G	A	G
chr12:89490962	C	T	C	C	C	C	C	C	T	C	C	C	T	C
chr12:89491074	G	A	G	G	G	G	G	G	A	G	G	G	A	G
chr12:89492878	C	T	C	C	C	C	C	C	T	C	C	C	T	C
chr12:89496725	C	C	C	T	C	C	C	C	C	T	C	T	C	C
chr12:89510796	A	A	A	G	A	A	A	A	A	G	A	G	A	A
chr12:89510867	T	T	T	G	T	T	T	T	T	G	T	G	T	T
chr12:89526316	A	A	A	G	A	A	A	A	A	G	A	G	A	A
chr12:89528992	G	G	G	A	G	G	G	G	G	A	G	A	G	G
chr12:89555278	T	T	T	C	T	T	T	T	T	C	T	C	T	T
chr12:89562599	A	A	A	G	A	A	A	A	A	G	A	G	A	A
chr12:89566879	A	A	A	G	A	A	A	A	A	G	A	G	A	A
chr12:89568689	A	A	A	G	A	A	A	A	A	G	A	G	A	A
chr12:89577715	T	T	T	G	T	T	T	T	T	G	T	G	T	T
chr12:89585187	T	T	T	A	T	T	T	T	T	A	T	A	T	T
chr12:89588699	G	G	G	A	G	G	G	G	G	A	G	A	G	G
chr12:89620540	G	G	G	T	G	G	G	G	G	T	G	T	G	G
chr12:89632748	C	C	C	T	C	C	C	C	C	T	C	T	C	C
chr12:89669785	A	A	A	G	A	A	A	A	A	G	A	G	A	A
chr12:89675376	A	A	A	G	A	A	A	A	A	G	A	G	A	A
chr12:89680524	T	C	T	C	T	T	T	T	C	C	T	C	C	T
chr12:89687490	T	C	T	C	T	T	T	T	C	C	T	C	C	T
chr12:89690312	G	A	G	A	G	G	G	G	A	A	G	A	A	G
chr12:89692857	G	A	G	A	G	G	G	G	A	A	G	A	A	G
chr12:89693983	G	T	G	T	G	G	G	G	T	T	G	T	T	G
chr12:89696297	A	A	A	T	A	A	A	A	A	T	A	T	A	A
chr12:89697513	T	C	T	C	T	T	T	T	C	C	T	C	C	T
chr12:89698922	A	G	A	G	A	A	A	A	G	G	A	G	G	A
chr12:89706102	C	C	C	A	C	C	C	C	C	A	C	A	C	C
chr12:89708616	T	C	C	C	C	C	C	C	C	C	T	C	C	C
chr12:89708714	A	C	A	A	A	A	A	A	C	A	A	A	C	A
chr12:89709057	C	T	C	C	C	C	C	C	T	C	C	C	T	C
chr12:89712445	C	T	C	C	C	C	C	C	T	C	C	C	T	C
chr12:89712461	T	G	T	T	T	T	T	T	G	T	T	T	G	T
chr12:89712504	A	G	A	A	A	A	A	A	G	A	A	A	G	A
chr12:89717805	A	T	A	T	A	A	A	A	T	T	A	T	T	A

FIG.4

Posición	Padre		Madre		Feto		Emb. 2		Tipo de Error
	Portador		Portadora		Afecto		Sano		
	P1	P2	M1	M2	P1	M1	P2	M2	
chr12:87254962	A	G	A	A	A	A	G	A	<p>ADO. Esta posición se muestra como homocigota, pero no es posible porque debería ser C según la asignación del alelo. En este caso, sólo se ha amplificado el alelo M2 del embrión.</p> <p>Error de secuenciación. Ninguno de los padres presenta una T, por lo que no es posible que el embrión la tenga.</p> <p>Punto de recombinación. A partir de este punto todos los SNPs cambian del alelo M2 al M1</p> <p>No Call. En esta posición no se ha obtenido ninguna lectura, por lo que no es posible determinar el alelo.</p>
chr12:87260006	C	G	C	C	C	C	G	C	
chr12:87302691	A	C	A	A	A	A	A	A	
chr12:87302747	A	C	A	A	A	A	C	A	
chr12:87391646	T	A	T	T	T	T	A	T	
chr12:87640519	T	T	T	C	T	T	T	C	
chr12:87674929	C	C	C	T	C	C	C	T	
chr12:87714749	A	G	A	A	A	A	G	A	
chr12:87714817	G	G	G	C	G	G	G	C	
chr12:87901028	A	G	A	A	A	A	G	T	
chr12:88238333	C	T	C	T	C	C	T	T	
chr12:88346589	T	C	T	T	T	T	C	T	
chr12:88450462	T	T	T	A	T	T	T	A	
chr12:88471567	-	T	-	T	-	-	T	T	
chr12:89400654	T	T	T	C	T	T	T	T	
chr12:89447754	T	A	T	T	T	T	A	T	
chr12:89449487	A	G	A	A	A	A	G	A	
chr12:89456904	A	G	A	A	A	A	G	A	
chr12:89459695	G	T	G	G	G	G	T	G	
chr12:89460093	A	C	A	A	A	A	C	A	
chr12:89462738	C	T	T	T	T	T	T	T	
chr12:89463041	C	A	C	C	C	C	A	C	
chr12:89468283	G	A	G	G	G	G	A	G	
chr12:89469954	C	T	C	C	C	C	-	-	
chr12:89470751	C	G	C	C	C	C	G	C	
chr12:89471403	G	A	G	G	G	G	A	G	
chr12:89473626	A	C	A	A	A	A	C	A	
chr12:89474352	C	T	C	C	C	C	T	C	

FIG.5

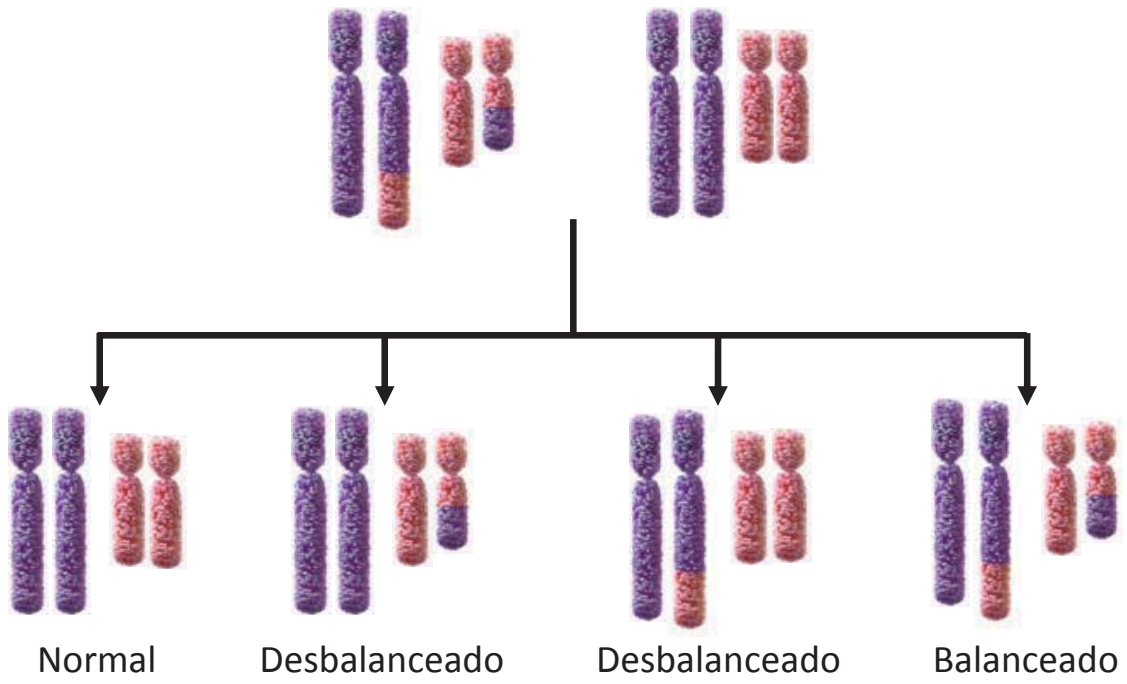


FIG.6



21 N.º solicitud: 201830731

22 Fecha de presentación de la solicitud: 20.07.2018

32 Fecha de prioridad:

INFORME SOBRE EL ESTADO DE LA TECNICA

51 Int. Cl.: **G16B20/20** (2019.01)
C12Q1/6827 (2018.01)

DOCUMENTOS RELEVANTES

Categoría	56 Documentos citados	Reivindicaciones afectadas
A	YAN LIYING <i>et al.</i> Live births after simultaneous avoidance of monogenic diseases and chromosome abnormality by next-generation sequencing with linkage analyses. Proceedings of the National Academy of Sciences of the United States of America DEC 29 2015. , 29/12/2015, Vol. 112, Páginas 15964-15969, ISSN 0027-8424, <DOI: doi:10.1073/pnas.1523297113>, página 15964, columna izquierda, primer párrafo; página 15965, columna izquierda, primer y segundo párrafos; página 15965, columna derecha, segundo y tercer párrafos; página 15968, columna derecha, segundo párrafo.	1-12
A	GONZALEZ-REIG S <i>et al.</i> New all-in-one protocol for 24-chromosome aneuploidies and monogenic diseases detection by next- generation sequencing: First-year experience. Reproductive BioMedicine Online 20180301 Elsevier Ltd nld. , 01/03/2018, Vol. 36, Páginas e34 - e35, ISSN 1472-6491 (print), página e34, columna derecha.	1-12
A	TREFF NATHAN R <i>et al.</i> Evaluation of targeted next-generation sequencing-based preimplantation genetic diagnosis of monogenic disease. FERTILITY AND STERILITY, 20130109 ELSEVIER SCIENCE INC, NEW YORK, NY, USA. , 09/01/2013, Vol. 99, Páginas 1377, ISSN 0015-0282, <DOI: doi:10.1016/j.fertnstert.2012.12.018>. Página 1377; página 1382, columna izquierda; segundo párrafo; página 1383.	1-12
A	US 2013210644 A1 (STOUGHTON ROLAND <i>et al.</i>) 15/08/2013, Párrafo [0007] y reivindicación 1.	1-12

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
04.06.2019

Examinador
S. González Peñalba

Página
1/2

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

G16B, C12Q

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, BIOSIS, MEDLINE, NPL, EMBASE, INTERNET