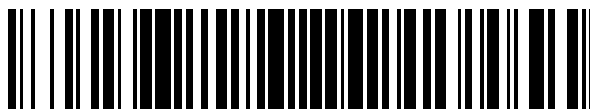


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 790 533**

51 Int. Cl.:

C12Q 1/68

(2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **20.05.2016 PCT/IB2016/000833**

87 Fecha y número de publicación internacional: **01.12.2016 WO16189388**

96 Fecha de presentación y número de la solicitud europea: **20.05.2016 E 16738215 (9)**

97 Fecha y número de publicación de la concesión europea: **12.02.2020 EP 3298166**

54 Título: **Análisis paralelo múltiplex de regiones genómicas blanco para análisis prenatales no invasivos**

30 Prioridad:

22.05.2015 US 201562165593 P
04.12.2015 US 201562263320 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
28.10.2020

73 Titular/es:

NIPD GENETICS PUBLIC COMPANY LIMITED
(100.0%)
Neas Engomis 31, Engomi
Nicosia 2409, CY

72 Inventor/es:

KOUMBARIS, GEORGE;
KYPRI, ELENA;
TSANGARAS, KYRIAKOS;
ACHILLEOS, ACHILLEAS;
MINA, PETROS;
PAPAGEORGIOU, ELISAVET, A. y
PATSALIS, PHILIPPOS, C.

74 Agente/Representante:

VEIGA SERRANO, Mikel

ES 2 790 533 T3

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

DESCRIPCION

Análisis paralelo múltiple de regiones genómicas blanco para análisis prenatales no invasivos

5 **Estado de la técnica**

El descubrimiento de ADN fetal libre (ADNfl) en la circulación materna (Lo, Y.M. *et al.* (1997) *Lancet* 350:485-487) marcó un hito hacia el desarrollo de análisis prenatales no invasivos para la detección de aneuploidías y ha abierto nuevas posibilidades en el entorno clínico. El ADNfl se ha usado exitosamente para la determinación del sexo fetal y el factor Rh fetal a partir de plasma materno (véase, por ejemplo, Bianchi, D. *et al.* (2005) *Obstet. Gynecol.* 106:841-844; Lo, Y.M. *et al.* (1998) *N. Engl. J. Med.* 339: 1734-1738; patente estadounidense n.º 6,258,540; publicación del PCT WO 91/07660). Estos métodos se han convertido en análisis de rutina en múltiples laboratorios de diagnóstico en todo el mundo. Sin embargo, el análisis directo de la cantidad reducida de ADNfl en presencia de un exceso de ADN materno es un gran desafío para la evaluación de las aneuploidías fetales mediante análisis prenatales no invasivos (NIPT)

Originalmente, se estimó que el ADNfl en la circulación materna representaba entre 3 y 6 % del ADN total (Lo, Y.M. *et al.* (1998) *Am. J. Hum. Genet.* 62:768-775). Sin embargo, estudios recientes sugieren que el ADN fetal puede alcanzar un porcentaje de 10 a 20 % del ADN total en la circulación materna (Lun, F.M. *et al.* (2008) *Clin. Chem.* 54:1664-1672). En las aneuploidías, uno de los cromosomas está presente en más o menos copias de lo normal. Por ejemplo, en los casos de trisomía 21, el cromosoma 21 está presente en tres copias en lugar de dos. Por ende, la capacidad de distinguir los casos normales de los casos de trisomía 21 dependen de la capacidad de detectar la copia adicional del cromosoma 21. Sin embargo, los niveles elevados de ADN materno en la circulación materna en comparación con la cantidad reducida de ADN fetal complica aún más la cuantificación.

A lo largo de la última década, se han aplicado una serie de métodos diferentes para discriminar el ADNfl del ADN materno circulante o para enriquecer el ADNfl (Chan, K.C. *et al.* (2004) *Clin. Chem.* 50:88-92; Papageorgiou, E.A. *et al.* (2009) *Am. J. Pathol.* 174: 1609-1618). Estos incluyen enfoques basados en el ADN, como enfoques de secuenciación (Chiu, R. W. *et al.* (2008) *Proc. Natl. Acad. Sci. USA* 105:20458-20463; Fan, H.C. *et al.* (2008) *Proc. Natl. Acad. Sci. USA* 105:16266-16271) o enfoques epigenéticos que se centran en la investigación de la metilación del ADN fetal mediante un tratamiento del ADN con bisulfato de sodio (Chim, S.S. *et al.* (2005) *Proc. Natl. Acad. Sci. USA* 102:14753-14758; publicación del PCT WO 2003/020974; publicación del PCT WO 2005/028674), enzimas de restricción sensibles a la metilación (Old, R.W. *et al.* (2007) *Reprod. Biomed. Online* 15:227-235; publicación del PCT WO 2005/035725) o anticuerpos específicos para los residuos de 5-metilcitosina de los dinucleótidos CpG a lo largo del genoma (Papageorgiou, E.A. *et al.* (2009) *Am. J. Pathol.* 174:1609-1618, Papageorgiou, E.A. *et al.* (2011) *Nature Medicine* 17:510-513; Tsaliki, E. *et al.* (2012) *Prenat. Diagn.* 32:996-1001; publicación del PCT WO 2011/092592). Otros enfoques se han dirigido al ARNm específico del feto (Ng, E.K. *et al.* (2003) *Proc. Natl. Acad. Sci. USA* 100:4 748-4 753) o bien a la investigación de proteínas específicas del feto (Avent, N.D. *et al.* (2008) *Semin. Fetal Neonatal Med.* 13:91-98).

La implementación de tecnologías de secuenciación de última generación (NGS) en el desarrollo de NIPT para la detección de aneuploidías ha revolucionado el campo. En 2008, dos grupos independientes demostraron que podía realizarse un NIPT para detectar la trisomía 21 mediante el uso de secuenciación masiva paralela *shotgun* (MPSS) de última generación (Chiu, R. W. *et al.* (2008) *Proc. Natl. Acad. Sci. USA* 105:20458-20463; Fan, H.C. *et al.* (2008) *Proc. Natl. Acad. Sci. USA* 105:16266-162710). La nueva era de los NIPT para la detección de aneuploidías ha abierto nuevas posibilidades para la implementación de estas tecnologías en la práctica clínica. Las empresas de biotecnología dedicadas total o parcialmente al desarrollo de NIPT han iniciado ensayos clínicos a gran escala tendientes a su implementación (Palomaki, G.E. *et al.* (2011) *Genet. Med.* 13:913-920; Ehrlich, M. *et al.* (2011) *Am. J. Obstet. Gynecol.* 204:205e1-11; Chen, E.Z. *et al.* (2011) *PLoS One* 6:e21791; Sehnert, A.J. *et al.* (2011) *Clin. Chem.* 57:1042-1049; Palomaki, G.E. *et al.* (2012) *Genet. Med.* 14:296-305; Bianchi, D.W. *et al.* (2012) *Obstet. Gynecol.* 119:890-901; Zimmerman, B. *et al.* (2012) *Prenat. Diagn.* 32:1233-1241; Nicolaides, K.H. *et al.* (2013) *Prenat. Diagn.* 33:575-579; Sparks, A.B. *et al.* (2012) *Prenat. Diagn.* 32:3-9). Actualmente, cuatro empresas en los Estados Unidos (SEQUENOM Inc., Verinata Health, Inc., Natera y Ariosa) ofrecen NIPT basados en enfoques de secuenciación de última generación.

Los primeros enfoques hacia los NIPT empleaban metodologías de secuenciación masiva paralela *shotgun* (MPSS) basada en tecnologías de NGS (véase, por ejemplo, la patente estadounidense n.º 7,888,017; la patente estadounidense n.º 8,008,018; la patente estadounidense n.º 8,195,415; la patente estadounidense n.º 8,296,076; la patente estadounidense n.º 8,682,594; la publicación de

patente estadounidense n.º 20110201507; y la patente estadounidense n.º 20120270739). Estos enfoques son de genoma completo, donde toda la muestra materna —muestra que contiene tanto ADN materno como ADN fetal libre— se somete a amplificación, secuenciación y análisis.

Más recientemente, se han desarrollado enfoques de NIPT basados en NGS dirigida, en los que se secuencian únicamente secuencias específicas de interés. Por ejemplo, se ha descrito un enfoque de NGS basado en polimorfismos de nucleótidos individuales (SNP) que involucra la amplificación y el análisis dirigidos de SNP en los cromosomas 13, 18, 21 X e Y en una única reacción (Zimmerman, B. et al. (2012) Prenat. Diag. 32:1233-1241; Nicolaides, K.H. et al. (2013) Prenat. Diagn. 33:575-579; publicación del PCT WO 2011/041485; patente estadounidense n.º 8,825,412). También se ha desarrollado un enfoque basado en NGS en el que se secuencian únicamente regiones específicas de interés, en el que se hibridan al molde complementario tres sondas por cada *locus* blanco. Una vez que las tres sondas hibridan, se ligan entre sí para formar una única sonda de mayor tamaño que, luego, se amplifica y secuencia (Sparks, A.B. et al. (2012) Prenat. Diagn. 32:3-9; publicación de patente estadounidense n.º 20120034603). Las muestras se analizan con una determinación paralela altamente multiplexada denominada análisis digital de regiones seleccionadas (DANSR). Tales enfoques dirigidos requieren un volumen de secuenciación considerablemente menor que los enfoques de MPSS, dado que únicamente se secuencian *loci* específicos del cromosoma de interés en lugar de regiones de todo el genoma.

Aún se requieren otras metodologías para enfoques de NIPT basados en NGS, particularmente enfoques que puedan dirigirse a secuencias específicas de interés y, así, permitan reducir considerablemente el volumen de secuenciación necesario en comparación con los enfoques de genoma completo.

Objeto de la invención

La invención provee métodos para análisis prenatales no invasivos que permiten detectar el riesgo de anomalías cromosómicas y utilizan un enfoque dirigido para enriquecer las secuencias de interés como paso previo a la secuenciación paralela en masa, junto con un análisis estadístico que permite un conteo y una evaluación altamente precisos de los constituyentes cromosómicos del plasma materno en regiones de interés. Así, los métodos de la invención reducen el volumen de secuenciación necesario para la secuenciación paralela en masa y hacen posible una aplicación de alta capacidad a menor costo y con un nivel muy alto de exactitud. Los métodos de la invención utilizan una mezcla de secuencias de captura de blancos (TACS) para enriquecer secuencias de interés en una muestra combinada que contiene tanto ADN materno como ADN fetal. En particular, la mezcla de TACS está diseñada de tal manera que las secuencias dentro de la mezcla tengan características que optimizan la eficiencia, la especificidad y la exactitud de la evaluación de las anomalías cromosómicas. Más específicamente, se ha optimizado el tamaño de las TACS, la cantidad de TACS, su ubicación en el/los cromosoma(s) de interés y su contenido de GC. La hibridación de las TACS a una biblioteca de secuenciación preparada a partir de una muestra combinada de ADN materno y fetal (p. ej., una muestra de plasma materno que contiene ADNf), seguida por el aislamiento de aquellas secuencias dentro de la biblioteca que se unen a las TACS, permite enriquecer únicamente las regiones cromosómicas de interés como paso previo a la secuenciación paralela en masa y el correspondiente análisis.

Así, en un aspecto, la invención provee un método para evaluar el riesgo de una anomalía cromosómica en un cromosoma de interés en ADN fetal en una mezcla combinada de ADN materno y fetal, donde el método comprende:

- (a) preparar la biblioteca de secuenciación a partir de la muestra combinada;
- (b) hibridar la biblioteca de secuenciación a una mezcla de secuencias de captura de blancos (TACS), donde la mezcla de TACS comprende secuencias que se unen a uno o más cromosomas de interés y donde:

- (i) cada secuencia de la mezcla tiene entre 100 y 260 pares de bases (bp) de longitud y/o entre 100 y 300 bp de longitud, y/o entre 100 y 350 bp de longitud, donde cada secuencia tiene un extremo 5' y un extremo 3';
- (ii) cada secuencia de la mezcla se une al/a los cromosoma(s) de interés a una distancia de al menos 150 pares de bases, tanto en el extremo 5' como en el extremo 3', de regiones que presentan variaciones en el número de copias, duplicaciones segmentales o elementos de ADN repetitivo; y
- (iii) el contenido de GC de las TACS se encuentra entre 19 % y 50 %, y/o entre 19 % y 60 %, y/o entre 19 % y 70 % y/o entre 19 % y 80 %;

- (c) aislar aquellos miembros de la biblioteca de secuenciación que se unen a las TACS para obtener una biblioteca enriquecida;
- (d) amplificar y secuenciar la biblioteca enriquecida; y
- (e) realizar análisis estadísticos sobre el resultado de la secuenciación de las secuencias de la biblioteca enriquecida para determinar así un riesgo asociado con la anomalía cromosómica y/o anomalía genética de otro tipo en el ADN fetal.

En una realización, la anomalía cromosómica es una aneuploidía, como una trisomía. El cromosoma de interés puede ser cualquier cromosoma, aunque los cromosomas preferentes incluyen el 13, el 18, el 21, el X y el Y. Una aneuploidía preferente para la detección es la trisomía 21 (T21). Además de las anomalías numéricas como las aneuploidías, la invención permite detectar otros tipos de anomalías cromosómicas, como anomalías estructurales, incluidas, sin carácter taxativo, las variaciones en el número de copias, incluidas, sin carácter taxativo, las microdeleciones y las microduplicaciones, las inserciones, las traslocaciones, las inversiones y las mutaciones pequeñas, incluidas las mutaciones puntuales y los perfiles de mutaciones.

En una realización, la mezcla de TACS está fijada a un sustrato sólido. Por ejemplo, las TACS pueden estar biotiniladas y fijadas a microesferas magnéticas recubiertas con estreptavidina. En otra realización, la mezcla de TACS puede estar libre en solución.

En una realización, las TACS están diseñadas para unirse a un cromosoma de interés y a una o más secuencias de referencia para detectar el riesgo de una anomalía cromosómica en el cromosoma de interés. Como alternativa, la mezcla de TACS puede estar diseñada para unirse a múltiples cromosomas de interés de manera que pueda detectarse el riesgo de múltiples anomalías cromosómicas, así como, por ejemplo, el género fetal, todo ello en un único análisis de la muestra. Por ejemplo, en una realización, la mezcla de TACS comprende diferentes secuencias que se unen a los cromosomas 13, 18, 21 y X, o a los cromosomas 13, 18, 21, X e Y.

En varias realizaciones, el contenido de GC de las TACS se encuentra entre 19 % y 80 %, entre 19 % y 70 %, entre 19 % y 60 %, entre 19 % y 50 %, entre 19 % y 49 %, entre 19 % y 48 %, entre 19 % y 47 %, entre 19 % y 46 %, entre 19 % y 45 %, entre 19 % y 44 %, entre 19 % y 43 %, entre 19 % y 42 %, entre 19 % y 41 % o entre 19 % y 40 %.

En varias realizaciones, cada secuencia dentro de la mezcla de TACS tiene entre 100 y 350 pares de bases, entre 150 y 260 pares de bases, entre 100 y 200 pares de bases o entre 200 y 260 pares de bases de longitud. En una realización, cada secuencia de la mezcla de TACS tiene 250 pares de bases de longitud.

En varias realizaciones, la mezcla de TACS puede comprender 800 o más secuencias distintas, 1500 o más secuencias distintas, 2000 o más secuencias distintas, 2500 o más secuencias distintas, o 3000 o más secuencias distintas. En una realización, la mezcla de TACS comprende 1600 secuencias distintas.

En una realización, la secuenciación de la biblioteca enriquecida provee una profundidad de lecturas correspondiente a los *loci* ubicados en el cromosoma de interés y profundidades de lecturas correspondientes a los *loci* de referencia, y el análisis estadístico se realiza aplicando un algoritmo que contrasta secuencialmente la profundidad de lecturas de los *loci* ubicados en el cromosoma de interés contra la profundidad de lecturas de los *loci* de referencia, donde las diferencias detectadas pueden indicar la presencia de variantes genéticas. Los pasos del algoritmo pueden incluir, sin carácter taxativo, los siguientes: (a) eliminar los *loci* secuenciados de forma inadecuada; (b) mitigar el sesgo inducido por el contenido de GC; y (c) determinar la ploidía. En una realización, el sesgo inducido por el contenido de GC se mitiga agrupando los *loci* con contenidos de GC equiparables.

En otra realización, la secuenciación de la biblioteca enriquecida provee el tamaño de los fragmentos del material genético acelular capturado por los TACS y el análisis estadístico comprende el uso de un algoritmo que compara y contrasta la distribución de tamaños de fragmentos entre los *loci* bajo prueba y los *loci* de referencia, donde las diferencias en la distribución indican la presencia de variantes genéticas. Los pasos del algoritmo pueden incluir, sin carácter taxativo, los siguientes: (a) eliminar los valores atípicos (*outliers*) de tamaños de fragmento; (b) crear una distribución binaria de tamaños de fragmentos; y (c) contrastar la distribución binaria de tamaños de fragmentos obtenida a partir de la región de interés con la distribución correspondiente a los *loci* de referencia para determinar la ploidía.

En general, la clasificación de la ploidía se logra mediante la aplicación de uno o más métodos estadísticos. Por ejemplo, el método estadístico puede seleccionarse de entre el grupo que consiste en una prueba *t*, una prueba de remuestreo (*bootstrap*) bivariada no paramétrica, una prueba de

permutación estratificada, una prueba binomial de proporciones y/o combinaciones de las anteriores. En una realización, se aplican a la muestra los cuatro métodos estadísticos mencionados arriba. En general, el método estadístico resulta en una puntuación asociada a la muestra combinada y se detecta el riesgo de la anomalía cromosómica en cuestión en el ADN fetal cuando la puntuación de la muestra combinada se encuentra por encima de un valor umbral de referencia. El método de la invención puede comprender, además, la estimación de la fracción de ADN fetal en la muestra combinada.

En otra realización, el método estadístico puede seleccionarse de entre el grupo que comprende una prueba t , una prueba de remuestreo bivariada no paramétrica y una prueba de permutación estratificada. En una realización, se aplican a la muestra todos los métodos estadísticos mencionados arriba. En general, el método estadístico resulta en una puntuación asociada a la muestra combinada y se detecta el riesgo de la anomalía cromosómica en cuestión en el ADN fetal cuando la puntuación de la muestra combinada se encuentra por encima de un valor umbral de referencia. El método de la invención puede comprender, además, la estimación de la fracción de ADN fetal en la muestra combinada.

También se divulgan kits para aplicar el método de la invención. En una realización, el kit comprende un contenedor que comprende la mezcla de TACS e instrucciones para aplicar el método. En varias otras realizaciones, el kit comprende otros componentes para realizar los otros pasos del método.

Descripción de las figuras

La Figura 1 es un diagrama esquemático del análisis paralelo múltiplex de regiones genómicas blanco para análisis prenatales no invasivos mediante secuencias de captura de blancos (TACS).

La Figura 2 es una lista de regiones cromosómicas ejemplares para amplificar TACS que se unan a los cromosomas 13, 18, 21 o X.

La Figura 3 es un gráfico de la asignación de un valor de puntuación a 98 muestras de sangre materna sujetas a un análisis paralelo múltiplex de regiones genómicas blanco mediante TACS para la detección del riesgo de trisomía 21, donde se evalúa la significación estadística de las diferencias entre las medianas de las profundidades de lecturas de los grupos condicionalmente apareados mediante una fórmula de una prueba t (denominado, en la presente, método estadístico 1).

La Figura 4 es un gráfico de la asignación de un valor de puntuación a 98 muestras de sangre materna sujetas a un análisis paralelo múltiplex de regiones genómicas blanco mediante TACS para la detección del riesgo de trisomía 21, donde el análisis se realizó mediante un método de remuestreo bivariado no paramétrico (denominado, en la presente, método estadístico 2).

La Figura 5 es un gráfico de la asignación de un valor de puntuación a 98 muestras de sangre materna sujetas a un análisis paralelo múltiplex de regiones genómicas blanco mediante TACS para la detección del riesgo de trisomía 21, donde el análisis se realizó mediante una prueba de permutación estratificada (denominado en la presente método estadístico 3).

La Figura 6 es un gráfico de la asignación de un valor de puntuación a 98 muestras de sangre materna sujetas a un análisis paralelo múltiplex de regiones genómicas blanco para la detección del riesgo de trisomía 21, donde el análisis se realizó mediante una prueba binomial de las proporciones de tamaños de fragmentos (denominado, en la presente, método estadístico 4).

La Figura 7 es una gráfica de los valores de puntuación ponderados de las 98 muestras de sangre materna obtenidos de un análisis mediante los métodos estadísticos 1, 2, 3 y el método de puntuación ponderada 1 que se muestran en las Figuras 3 a 5.

La Figura 8 es una gráfica de los valores de puntuación ponderados de las 98 muestras de sangre materna obtenidos de un análisis mediante los métodos estadísticos 1, 2, 3, 4 y el método de puntuación ponderada 1 que se muestran en las Figuras 3 a 6.

La Figura 9 es una gráfica un método ponderado alternativo, el método de valores de puntuación ponderados 2, de los valores de puntuación de las 98 muestras de sangre materna obtenidos de un análisis mediante los métodos estadísticos 1, 2, 3 que se muestran en las Figuras 3 a 5.

La Figura 10 es un gráfico de la asignación de un valor de puntuación a 9 muestras sintéticas sujetas a un análisis paralelo múltiplex de regiones genómicas blanco mediante TACS dirigidas al locus 7q11.23, donde el análisis se realizó mediante una realización del método estadístico 1 para la detección de microdeleciones.

La Figura 11 es un gráfico de la asignación de un valor de puntuación a 9 muestras sintéticas sujetas a un análisis paralelo múltiplex de regiones genómicas blanco mediante TACS dirigidas al locus 7q11.23, donde el análisis se realizó mediante una realización del método estadístico 1 para la detección de microduplicaciones.

Descripción detallada de la invención

La invención refiere a un método de NIPT que involucra un enriquecimiento basado en hibridación de determinadas regiones blanco a lo largo del genoma humano en un panel de determinaciones, seguida por la cuantificación, combinada con un nuevo *pipeline* bioinformático y matemático. El enriquecimiento por hibridación en solución se ha usado en el pasado para enriquecer regiones específicas de interés antes de la secuenciación (véase, por ejemplo, Meyer, M y Kirchner, M. (2010) Cold Spring Harb. Protoc. 2010(6):pdbprot5448; Liao, G.J. et al. (2012) PLoS One 7:e38154; Maricic, T. et al. (2010) PLoS One 5:e14004; Tewhey, R. et al. (2009) Genome Biol. 10:R116; Tsangaras, K. et al. (2014) PLoS One 9:e109101). Sin embargo, en los métodos de NIPT de la invención, las secuencias blanco empleadas para enriquecer regiones específicas de interés, relevantes para la detección del riesgo de una anomalía cromosómica, se han optimizado para maximizar la eficiencia, la especificidad y la exactitud. El genoma humano está repleto de elementos que pueden confundir y desconcertar todo tipo de análisis genéticos, lo que pone en evidencia el beneficio de un enfoque hacia los NIPT que sea de naturaleza dirigido. En vista de ello, la complejidad del genoma humano y la presencia de estos elementos de confusión exigen un diseño cuidadoso de las secuencias de captura de blancos empleadas para el enriquecimiento. De acuerdo con lo que se describe en el presente, se han desarrollado secuencias de captura de blancos (TACS) óptimas que permiten NIPT más simples y robustos a la vez que minimizan el riesgo de resultados falsos positivos y falsos negativos que suelen ocurrir en los NIPT de genomas completos a causa de la inevitable secuenciación de elementos de confusión.

El método de la invención para evaluar el riesgo de una anomalía cromosómica en un cromosoma de interés en ADN fetal en una mezcla combinada de ADN materno y fetal comprende:

- (a) preparar la biblioteca de secuenciación a partir de la muestra combinada;
- (b) hibridar la biblioteca de secuenciación a una mezcla de secuencias de captura de blancos (TACS), donde la mezcla de TACS comprende secuencias que se unen a uno o más cromosomas de interés y donde:
 - (i) cada secuencia de la mezcla tiene entre 100 y 260 pares de bases (bp) de longitud, donde cada secuencia tiene un extremo 5' y un extremo 3';
 - (ii) cada secuencia de la mezcla se une al/a los cromosoma(s) de interés a una distancia de al menos 150 pares de bases, tanto en el extremo 5' como en el extremo 3', de regiones que presentan variaciones en el número de copias (CNV), duplicaciones segmentales o elementos de ADN repetitivo; y
 - (iii) el contenido de GC de las TACS se encuentra entre 19 % y 50 %;
- (c) aislar aquellos miembros de la biblioteca de secuenciación que se unen a las TACS para obtener una biblioteca enriquecida;
- (d) amplificar y secuenciar la biblioteca enriquecida; y
- (e) realizar análisis estadísticos sobre el resultado de la secuenciación de las secuencias de la biblioteca enriquecida para determinar así un riesgo asociado con la anomalía cromosómica en el ADN fetal.

También se incluyen kits para aplicar el método de la invención.

Diversos aspectos de esta divulgación se describen en mayor detalle en las subsecciones siguientes.

Diseño de las secuencias de captura de blancos

En la presente, los términos "secuencias de captura de blancos" o "TACS" hacen referencia a secuencias de ADN cortas que son complementarias a la(s) región/ones de interés de uno o más cromosomas de interés y que se usan como "carnada" para capturar y enriquecer la región de interés a partir de una biblioteca amplia de secuencias, como una biblioteca de secuenciación de genoma completo preparada a partir de una muestra de plasma materno. Para el enriquecimiento, se usa una mezcla de TACS, donde las secuencias dentro de la mezcla se han optimizado en relación con: (i) la longitud de las secuencias; (ii) la distribución de las TACS a lo largo de la(s) región/ones de interés; y (iii) el contenido de GC de las TACS. Asimismo, se ha optimizado la cantidad de secuencias dentro de la mezcla de TACS (tamaño de la mezcla).

Se ha descubierto que las TACS con una longitud de 100 a 260 pares de bases son óptimas para maximizar la eficiencia de enriquecimiento. En varias otras realizaciones, cada secuencia dentro de la mezcla de TACS tiene entre 150 y 260 pares de bases, entre 100 y 200 pares de bases, entre 200 y 260 pares de bases o entre 100 y 350 pares de bases de longitud. En realizaciones preferentes, la longitud de las TACS de la mezcla es de 250 o 260 pares de bases. La persona razonablemente versada en la técnica apreciará que, por lo general, puede usarse una pequeña variación en el tamaño de las TACS sin afectar los resultados (p. ej., la eliminación de unos pocos pares de bases a cada extremo de las TACS). Así, la cantidad de pares de bases que se indican en la presente deben considerarse como "aproximadas", admitiendo una pequeña variación (p. ej., de 1 a 5 %) en la longitud. Así, por ejemplo, una longitud de "250 pares de bases" hace referencia a "aproximadamente 250 pares de bases", de manera que también abarca, por ejemplo, 248 pares de bases o 252 pares de bases.

La distribución de las TACS a lo largo de cada región o cromosoma de interés se ha optimizado para evitar las repeticiones con alto número de copias, las repeticiones con bajo número de copias y las variantes en el número de copias, y, a la vez, poder direccionar la captura a polimorfismos de nucleótidos individuales (SNP) informativos para permitir la detección de las aneuploidías, la detección de variaciones en el número de copias y la estimación de la fracción fetal (ff). Así, cada secuencia dentro de la mezcla de TACS está diseñada de manera tal que el extremo 5' y el extremo 3' se encuentran, cada uno, a una distancia de al menos 150 pares de bases de regiones en el genoma de las que se sabe que contienen uno o más de los siguientes elementos: variaciones en el número de copias (CNV), duplicaciones segmentales y/o elementos de ADN repetitivo (como transposones o zonas de repeticiones en tándem). En varias otras realizaciones, cada secuencia dentro de la mezcla de TACS está diseñada de manera tal que el extremo 5' y el extremo 3' se encuentran, cada uno, a una distancia de al menos 200, 250, 300, 400 o 500 pares de bases de regiones en el genoma de las que se sabe que contienen uno o más de los elementos anteriores.

El término "variaciones en el número de copias" (CNV) es un término de la técnica que hace referencia a una forma de variante estructural del genoma humano en la cual existen alteraciones en el ADN del genoma de ciertos individuos que resultan en una cantidad menor o mayor a la normal de una o más secciones del genoma. Las CNV corresponden a regiones relativamente extensas del genoma que pueden haberse borrado (por ejemplo, una sección que normalmente es A-B-C-D puede ser A-B-D) o duplicado (por ejemplo, una sección que normalmente es A-B-C-D puede ser A-B-C-C-D). Las CNV representan aproximadamente 13 % del genoma humano y el tamaño de cada variación varía entre aproximadamente 1 kilobase a varias megabases.

El término "duplicaciones segmentales" (también "repeticiones con bajo número de copias") también es un término de la técnica que refiere a bloques de ADN cuya longitud varía entre 1 y 400 kilobases que ocurren en más de una posición del genoma y que, por lo general, tienen un alto grado de identidad a nivel de secuencia (mayor de 90 %). Las duplicaciones segmentales se abordan, por ejemplo, en Eichler. E.E. (2001) Trends Genet. 17:661-669.

El término "elementos de ADN repetitivo" (también "ADN repetitivo" o "ADN repetido") también es un término de la técnica que refiere a patrones de ADN que ocurren en múltiples copias a lo largo del genoma. El término "elemento de ADN repetitivo" abarca las repeticiones terminales, las repeticiones en tándem y las repeticiones intercaladas, incluidos los transposones. Los elementos de ADN repetitivo en el marco de las tecnologías de NGS se analiza en más detalle, por ejemplo, en Todd, J. et al. (2012) Nature Reviews Genet. 13:36-46.

Las TACS están diseñadas con características específicas de contenido de GC con el fin de minimizar el sesgo inducido por el contenido de GC en los datos y permitir el uso de un *pipeline* de análisis de datos personalizado e innovador. Se ha determinado que las TACS con un contenido de GC de 19 a 50 % logran un enriquecimiento óptimo y son las que muestran mayor rendimiento con el ADN fetal acelular. Dentro de una mezcla de TACS, diferentes secuencias pueden tener diferentes contenido de GC porcentuales, pero, para ser incluida en la mezcla, el contenido de GC porcentual de cada secuencia debe encontrarse en el rango de 19 a 50 %. En algunas instancias, la mezcla de TACS puede seleccionarse de manera de definir otro rango de contenido de GC porcentual que se considere más adecuado para la evaluación de anomalías genéticas específicas. Algunos ejemplos de rangos de contenidos de GC porcentuales pueden ser, sin carácter taxativo, entre 19 % y 75 %, entre 19 % y 65 %, entre 19 % y 55 %, entre 19 % y 50 %, entre 19 % y 49 %, entre 19 % y 48 %, entre 19 % y 47 %, entre 19 % y 46 %, entre 19 % y 45 %, entre 19 % y 44 %, entre 19 % y 43 %, entre 19 % y 42 %, entre 19 % y 41 % o entre 19 % y 40 %.

Según se describen en mayor detalle más abajo a propósito de una realización del análisis de datos, luego de la amplificación y secuenciación de las secuencias enriquecidas, los *loci* de prueba y los

loci de referencia pueden aparearse o agruparse en función de su contenido de GC porcentual (p. ej., los *loci* bajo análisis con un contenido de GC porcentual de 40 % se aparean con *loci* de referencia con un contenido de GC porcentual de 40 %). Se apreciará que el procedimiento de apareamiento en función del contenido de GC porcentual podría permitir una variación leve en el rango de contenido de GC porcentual admitido para un apareamiento. Como ejemplo no taxativo y con referencia al ejemplo descrito anteriormente, un *locus* de prueba con un contenido de GC porcentual de 40 % podría aparearse con *loci* de referencia con un rango de contenido de GC porcentual entre 39 y 41 %, abarcando el contenido de GC porcentual del *locus* de prueba dentro de un rango adecuado.

Para preparar una mezcla de TACS con los criterios optimizados mencionados arriba en relación con el tamaño, la ubicación en el genoma humano y el contenido de GC porcentual, pueden aplicarse métodos manuales o computarizados conocidos en la técnica para el análisis del genoma humano de referencia. En una realización, se implementa un método semiautomático en el que primeramente se designan manualmente regiones basadas en la versión 19 del genoma humano de referencia (hg19) de manera de evitar las regiones repetitivas mencionadas y, posteriormente, las regiones designadas se curan en relación con el contenido de GC con ayuda de software que calcula el contenido de GC de cada región en función de sus coordenadas en la versión 19 del genoma humano de referencia (hg19). En otra realización, se emplea software hecho a medida para analizar el genoma humano de referencia e identificar regiones adecuadas para TACS que cumplan ciertos criterios, incluidos, sin carácter taxativo, criterio referidos al contenido de GC porcentual, la proximidad a regiones repetitivas y/o la proximidad a otras TACS.

La cantidad de TACS en la mezcla se ha examinado y ajustado cuidadosamente para lograr el mejor equilibrio entre la robustez de los resultados y el costo/la capacidad de la determinación. La mezcla generalmente contiene 800 o más TACS, pero puede incluir más; por ejemplo, 1500 o más TACS, 2000 o más TACS, o 2500 o más TACS. Se ha detectado que un número óptimo de TACS en la mezcla es de 1600. La persona razonablemente versada en la técnica apreciará que, por lo general, puede usarse una pequeña variación en el tamaño de la mezcla sin afectar los resultados (p. ej., la eliminación o el agregado de una pequeña cantidad de TACS). Así, las cantidades de TACS de la mezcla que se indican en la presente deben considerarse como "aproximadas", admitiendo una pequeña variación (p. ej., de 1 a 5 %) en su tamaño. Así, por ejemplo, un tamaño de mezcla de "1600 secuencias" hace referencia a "aproximadamente 1600 secuencias", de manera que también abarca, por ejemplo, 1590 secuencias o 1610 secuencias.

En vista de lo anterior, en otro aspecto, la invención provee un método para preparar una mezcla de TACS para usar en el método de la invención para detectar el riesgo de una anomalía cromosómica y/u otra anomalía genética, donde el método para preparar la mezcla de TACS comprende: seleccionar regiones en uno o más cromosomas de interés que cumplan los criterios definidos arriba (p. ej., que se encuentren a una distancia de al menos 150 pares de bases a cada extremo de las secuencias repetitivas mencionadas y que tengan un contenido de GC entre 19 % y 50 %); preparar cebadores que amplifiquen las secuencias que hibridan con las regiones seleccionadas; y amplificar las secuencias, donde cada secuencia tiene entre 100 y 260 pares de bases de longitud.

Obtención y preparación de las muestras

Los métodos de la invención se realizan sobre una muestra combinada que contiene tanto ADN materno como ADN fetal. En general, la muestra es una muestra de plasma materno, aunque pueden usarse otras fuentes de tejido que contengan tanto ADN materno como ADN fetal. El plasma materno puede obtenerse de una muestra de sangre periférica entera de una mujer embarazada y el plasma puede obtenerse mediante métodos estándar. Apenas 2 a 4 ml de plasma son suficientes para obtener material de ADN suficiente para el análisis de acuerdo con el método de la invención. Luego, el ADN acelular total puede extraerse de la muestra mediante técnicas estándar, que incluyen, sin carácter taxativo, el protocolo Qiasymphony (Qiagen) adecuado para aislar ADN fetal libre, así como cualquier otro método de extracción manual o automático adecuado para aislar ADN acelular.

Luego de aislarse, el ADN acelular de la muestra combinada se usa para construir una biblioteca de secuenciación cuyo fin es hacer la muestra compatible con una tecnología de secuenciación ulterior (por ejemplo, sin carácter taxativo, la secuenciación de última generación Illumina). En general, esto involucra la ligación de adaptadores a los extremos de los fragmentos de ADN acelular, seguida por una amplificación. Hay kits para la preparación de bibliotecas de secuenciación disponibles comercialmente. En el Ejemplo 1 se describe en detalle un protocolo ejemplar —sin carácter taxativo— para la preparación de una biblioteca de secuenciación.

Enriquecimiento mediante hibridación de TACS

Para enriquecer la(s) región/ones de interés del/de los cromosoma(s) de interés, se hace hibridar la mezcla de TACS con la biblioteca de secuenciación y, a continuación, se aíslan las secuencias de la biblioteca de secuenciación que hibrida a las TACS. Para facilitar el aislamiento de las secuencias deseadas enriquecidas, las TACS generalmente se modifican de tal manera que las secuencias que hibridan con las TACS se pueden separar de aquellas que no hibridan con las TACS. En general, esto se logra fijando las TACS a un sustrato sólido. Esto permite la separación física de las secuencias que hibridan con las TACS de aquellas que no hibridan con las TACS. Por ejemplo, cada secuencia dentro de la mezcla de TACS puede marcarse con biotina y la mezcla puede fijarse a microesferas recubiertas con una sustancia que se una a la biotina, como estreptavidina o avidina. En una realización preferente, las TACS se marcan con biotina y se unen a microesferas magnéticas recubiertas con estreptavidina. Sin embargo, la persona razonablemente versada en la técnica apreciará que existen otros sistemas de unión por afinidad conocidos en la técnica y que pueden usarse en lugar del sistema de biotina-estreptavidina/avidina. Por ejemplo, puede usarse un sistema basado en anticuerpos en el que las TACS se marquen con un antígeno y, luego, se unan a microesferas recubiertas con el anticuerpo correspondiente. Además, las TACS pueden incorporar en un extremo una secuencia marca y pueden fijarse a un sustrato sólido mediante una secuencia complementaria en el sustrato sólido que hibride con la secuencia marca. Por otra parte, además de microesferas magnéticas, pueden usarse otros tipos de sustratos sólidos, como microesferas poliméricas y otros similares.

Luego del enriquecimiento de la(s) secuencia(s) de interés con las TACS y la formación de la biblioteca enriquecida, los miembros de la biblioteca enriquecida se eluyen del sustrato sólido y se amplifican y secuencian mediante métodos estándar conocidos en la técnica. En general, se emplea la tecnología estándar de secuenciación de última generación Illumina, aunque pueden emplearse también otras tecnologías de secuenciación que provean no solo información de secuencia sino también recuentos muy exactos. Para detectar anomalías genéticas —incluidas, sin carácter taxativo, las aneuploidías o las variaciones estructurales en el número de copias— exige un recuento muy exacto y la NGS es un tipo de tecnología que provee tal recuento muy exacto. Así, para la detección de anomalías genéticas, incluidas, sin carácter taxativo, las aneuploidías o las variaciones estructurales en el número de copias, pueden usarse otros métodos de recuento exactos, como la PCR digital y los microarreglos, en lugar de la NGS. En el Ejemplo 3 se describen en detalle protocolos ejemplares —sin carácter taxativo— para la amplificación y secuenciación de la biblioteca enriquecida.

Análisis de los datos

La información obtenida de la secuenciación de la biblioteca enriquecida se analiza mediante un innovador *pipeline* de análisis biomatemático/bioestadístico. Este *pipeline* de análisis explota las características de las TACS, en tanto que la alta eficiencia de la captura de blancos permite una detección eficiente de aneuploidías o variaciones estructurales en el número de copias, así como otros tipos de anomalías genéticas. En el Ejemplo 4 se describe en detalle un análisis ejemplar. En el análisis, los fragmentos de ADN secuenciados de la muestra se alinean primeramente al genoma humano de referencia. Se emplean métricas de control de calidad para inspeccionar las propiedades de la muestra alineada y decidir si la muestra es apta para clasificación. Estas métricas de calidad pueden incluir, sin carácter taxativo, el análisis de patrones de enriquecimiento de los *loci* de interés, como, por ejemplo, la profundidad de secuenciación general de la muestra, el volumen de secuenciación de la muestra que sea específico ("en el blanco"), el rendimiento de las TACS, el nivel esperado de sesgo inducido por el contenido de GC y la cuantificación de la fracción fetal. Para determinar el riesgo de una anomalía cromosómica en el ADN fetal de la muestra, se emplea un algoritmo innovador. Los pasos del algoritmo incluyen, sin carácter taxativo, eliminar los loci secuenciados de forma inadecuada; extraer información relativa a la profundidad de lecturas y el tamaño de fragmentos en coordenadas específicas de las TACS; mitigar el sesgo inducido por el contenido de GC; y determinar la ploidía.

La determinación de la ploidía se logra mediante uno o más métodos estadísticos, algunos de los cuales son, sin carácter taxativo, una prueba *t*, una prueba de remuestreo (*bootstrap*), una prueba de permutación, una prueba binomial de proporciones y/o combinaciones de las anteriores. La persona razonablemente versada en la técnica apreciará que la selección y aplicación de pruebas estadísticas para incluir en una determinación de la ploidía se basa en la cantidad de puntos de datos disponibles. Así, la idoneidad de cada prueba viene dada por distintos factores como —sin carácter taxativo— la cantidad de TACS utilizadas y la correspondiente aplicación para mitigar el sesgo inducido por el contenido de GC, si corresponde. Así, los métodos mencionados deben considerarse como ejemplos de los tipos de análisis estadísticos que pueden emplearse y no son los únicos métodos adecuados para la determinación de la ploidía. En general, el método estadístico resulta en una puntuación asociada a la muestra combinada y se detecta el riesgo de la anomalía cromosómica en cuestión en

el ADN fetal cuando la puntuación de la muestra combinada se encuentra por encima de un valor umbral de referencia.

En particular, un aspecto de los análisis estadísticos involucra cuantificar y mitigar el sesgo inducido por el contenido de GC. Además del desafío de detectar pequeños cambios de señal en el ADN fetal en la muestra combinada (por ejemplo, sin carácter taxativo, la presencia de más o menos material genético de ciertas regiones cromosómicas fetales), el propio proceso de secuenciación introduce ciertos sesgos que pueden oscurecer la detección de la señal. Uno de tales sesgos es la secuenciación/amplificación preferente de las regiones genéticas en función de su contenido de GC. Por ello, ciertos métodos de detección —incluidos, sin carácter taxativo, los métodos basados en la profundidad de lecturas— deben tener en cuenta tal sesgo al momento de analizar los datos de secuenciación. Así, se debe cuantificar el sesgo en los datos y, posteriormente, se deben aplicar métodos adecuados que contemplen dicho sesgo de manera tal que las dependencias del contexto genético no puedan afectar los métodos estadísticos que pudieran usarse para cuantificar el riesgo de anomalías genéticas fetales.

Por ejemplo, un método para cuantificar el sesgo inducido por el contenido de GC es aplicar la técnica de suavizado local ponderado de diagramas de dispersión (LOESS) a los datos de secuenciación. Cada *locus* blanco puede definirse por su profundidad de lecturas obtenidas en la secuenciación y su contenido de GC. Una recta de ajuste óptimo por estas dos variables sobre una gran cantidad de *loci* provee una estimación de la profundidad de lecturas de secuenciación esperada dado el contenido de GC. Una vez que se completa este paso de cuantificación del sesgo inducido por el contenido de GC, el siguiente paso es usar esta información para considerar posibles sesgos en los datos. Un método es normalizar la profundidad de lecturas de todos los *loci* respecto de su profundidad de lecturas esperada (en función del contenido de GC de cada *locus*). En principio, eso desvincula los datos de profundidad de lecturas de su contexto genético y torna todos los datos comparables entre sí. Así, los datos recuperados de regiones con distintos contenidos de GC, como, por ejemplo —sin carácter taxativo—, distintos cromosomas, se pueden usar en pruebas estadísticas posteriores para la detección de anomalías. Así, mediante el procedimiento LOESS, el sesgo inducido por el contenido de GC se desvincula de los datos antes de la aplicación de las pruebas estadísticas. En una realización, el análisis estadístico de las secuencias de la biblioteca enriquecida comprende mitigar el sesgo de GC con un procedimiento LOESS.

En una realización preferente alternativa, el sesgo inducido por el contenido de GC se cuantifica y mitiga agrupando los *loci* con contenidos de GC similares (equiparables). Así, conceptualmente, este método para mitigar el sesgo inducido por el contenido de GC comprende los tres pasos siguientes:

- 1) identificar y calcular el contenido de GC en las TACS;
- 2) mitigar/considerar el sesgo inducido por el contenido de GC mediante diversos procedimientos de apareamiento/agrupación de las TACS; y
- 3) calcular el riesgo de que distintas anomalías genéticas estén presentes en el feto aplicando métodos estadísticos y matemáticos a los conjuntos de datos resultantes del paso 2.

En el caso de la prueba *t*, el conjunto de datos se divide en dos grupos: los *loci* de prueba y los *loci* de referencia. Por cada grupo, se crean subconjuntos de grupos en los que los *loci* se categorizan de acuerdo con su contenido de GC como se ilustra en el ejemplo no taxativo de la Tabla 1, a continuación:

Tabla 1

GC	Profundidad de lecturas en los <i>loci</i> de referencia	Profundidad de lecturas en los <i>loci</i> de prueba
40 %	$x_1^{40}, x_2^{40}, \dots, x_{nx40}^{40}$	$y_1^{40}, y_2^{40}, \dots, y_{ny40}^{40}$
41 %	$x_1^{41}, x_2^{41}, \dots, x_{nx41}^{41}$	$y_1^{41}, y_2^{41}, \dots, y_{ny41}^{41}$
42 %	$x_1^{42}, x_2^{42}, \dots, x_{nx42}^{42}$	$y_1^{42}, y_2^{42}, \dots, y_{ny42}^{42}$
...	...	...

La persona razonablemente versada en la técnica apreciará que la creación de subgrupos puede involucrar abarcar un rango de contenidos de GC adecuados y/o un subconjunto de *loci* definidos por un contenido de GC y/o un rango de contenido de GC dados. Así, el contenido de GC porcentual dado en el ejemplo no taxativo de la Tabla 1 deben considerarse "aproximados", admitiendo una leve variación (p. ej., 1 o 2 %). Así, por ejemplo, un contenido de GC porcentual de "40 %" pretende hacer referencia a "aproximadamente 40 %", de tal manera que, por ejemplo, también podrían estar comprendidos los *loci* con contenidos de GC en el rango de 39 % a 41 % si se considera pertinente.

Así, cuando se hace referencia a un contenido de GC particular, se entiende que la referencia y los subgrupos de *loci* de prueba pueden comprender cualquier cantidad de *loci* relacionados con un contenido de GC porcentual y/o rango particulares.

Posteriormente, por cada subgrupo de contenido de GC, se calcula una profundidad de lecturas representativa. Para ello, se pueden usar distintos métodos, incluidos, sin carácter taxativo, la media, la mediana o la moda de caja conjunto. Así, se crean dos vectores de profundidades de lecturas representativas, donde uno corresponde a los *loci* de referencia y el otro a los *loci* de prueba (p. ej., X_m e Y_m). En una realización, ambos vectores pueden contrastarse entre sí para identificar diferencias significativas en la profundidad de lecturas. En otra realización, la diferencia entre ambos vectores puede usarse para evaluar si hay discrepancias considerables entre los *loci* de referencia y los *loci* de prueba. Se atribuye a la muestra la puntuación de la prueba. En el caso de los análisis estadísticos que emplean un enfoque de remuestreo (*bootstrap*), el conjunto de datos se divide en dos grupos: los *loci* de referencia y los *loci* de prueba. Entonces, se calcula el contenido de GC de cada *locus*. Luego, se realiza el siguiente procedimiento:

Se selecciona un *locus* al azar de entre los *loci* de referencia y se registran su profundidad de lecturas y su contenido de GC. Posteriormente, se selecciona un *locus* al azar de entre los *loci* de prueba, con la única condición de que su contenido de GC sea similar al del *locus* de referencia. Se registra su profundidad de lecturas. La persona razonablemente versada en la técnica apreciará que la similitud entre los contenidos de GC puede abarcar un rango de contenidos de GC adecuados. Así, la referencia a un contenido de GC porcentual específico puede considerarse "aproximada", "próxima" o "dentro de un rango adecuado" (p. ej., 1 a 2 %) que incluya el contenido de GC porcentual específico investigado. De este modo, se crea un par de *loci* de referencia y de prueba con contenidos de GC similares. Se registra la diferencia del par de *loci* de referencia y de prueba, notada E1. Luego, los *loci* se regresan a sus grupos respectivos. Este proceso se repite hasta crear una muestra de remuestreo del mismo tamaño que la cantidad de TACS de prueba presentes. Entonces, se estima y se registra una profundidad de lecturas representativa de la muestra de remuestreo, notada E_mu. Para ello, se pueden usar distintos métodos, incluidos, sin carácter taxativo, la media, la moda o la mediana del vector, y/o múltiplos de dichos valores.

El proceso descrito anteriormente se repite tantas veces como sea necesario y se crea una distribución de E_mu. Luego, se atribuye a la muestra una puntuación que corresponde a un percentil de esta distribución.

En el caso de los análisis estadísticos con una prueba de permutación, el conjunto de datos se ordena primero en dos grupos: los *loci* de prueba y los *loci* de referencia. Por cada grupo, se crean subconjuntos de grupos en los que los *loci* se categorizan de acuerdo con la similitud en sus contenidos de GC (véanse las columnas 2 y 3 del ejemplo no taxativo de la Tabla 2, a continuación). También se registra la cantidad de *loci* presentes en cada subgrupo de prueba. Los *loci* del grupo de prueba se usan para estimar la profundidad de lecturas del grupo de prueba, notada Yobs. Para hacerlo, puede seleccionarse un número representativo de cada subgrupo de contenido de GC. Para obtener una estimación de la profundidad de lecturas se pueden usar distintos métodos, incluidos, sin carácter taxativo, la media, la mediana o la moda de los *loci* seleccionados.

Tabla 2

GC	Profundidad de lecturas en los <i>loci</i> de referencia	Profundidad de lecturas en los <i>loci</i> de prueba	Cant. de <i>loci</i> de prueba	Unión de los <i>loci</i>
40 %	$x_1^{40}, x_2^{40}, \dots, x_{nx40}^{40}$	$y_1^{40}, y_2^{40}, \dots, y_{ny40}^{40}$	ny40	$x_1^{40}, \dots, x_{nx40}^{40}, y_1^{40}, \dots, y_{ny40}^{40}$
41 %	$x_1^{41}, x_2^{41}, \dots, x_{nx41}^{41}$	$y_1^{41}, y_2^{41}, \dots, y_{ny41}^{41}$	ny41	$x_1^{41}, \dots, x_{nx41}^{41}, y_1^{41}, \dots, y_{ny41}^{41}$
42 %	$x_1^{42}, x_2^{42}, \dots, x_{nx42}^{42}$	$y_1^{42}, y_2^{42}, \dots, y_{ny42}^{42}$	ny42	$x_1^{42}, \dots, x_{nx42}^{42}, y_1^{42}, \dots, y_{ny42}^{42}$
...	...	...	...	...

Luego, se crea una distribución de los valores Yobs de prueba usando *loci* indistintamente de que provengan del grupo de prueba o del grupo de referencia, tal como se explica a continuación. Los *loci* de prueba y de referencia de cada subgrupo de contenido de GC (véase la última columna del ejemplo de la Tabla 2) se combinan para permitir calcular una nueva estimación de la profundidad de lecturas. De cada subgrupo combinado se escoge una cantidad de *loci* al azar, cantidad acotada superiormente por la cantidad de *loci* de prueba usados en el cálculo original de Yobs (p. ej., en el caso de un contenido de GC de 40 % y en el contexto del ejemplo no taxativo de la Tabla 2, esta cantidad de *loci* puede estar en el rango [1,ny40]). La nueva estimación de la profundidad de lecturas se calcula a partir de todos los *loci* seleccionados. El procedimiento se repite tantas veces como sea

necesario para construir una distribución de las medias observadas. Luego, se le atribuye a una muestra una puntuación que corresponde a la posición de Yobs en esta distribución, usando una transformación adecuada que tiene en cuenta los momentos de la distribución construida. Al igual que con los métodos descritos anteriormente, se apreciará que se admite una ligera variación en el contenido de GC porcentual (p. ej., 1 a 2 %) si se considera adecuado. Así, la referencia a un contenido de GC específico puede interpretarse como "aproximada", de manera que, por ejemplo, al hacer referencia a un contenido de GC de 40 %, el método puede utilizar los *loci* que tengan un contenido de GC de "aproximadamente" 40 % (p. ej., entre 39 y 41 %).

En el caso del análisis estadístico con una prueba binomial de proporciones, se emplean los tamaños de fragmentos alineados a coordenadas genómicas específicas de los TACS. Se ha demostrado que los fragmentos de material genético acelular provenientes de la placenta son de menor tamaño que los de otros tipos de material genético acelular (Chan, K.C. (2004) Clin. Chem. 50:88-92). Así, el estadístico de interés es si la proporción de fragmentos pequeños alineados a una región de prueba específica de una TACS se aparta significativamente de lo que se espera en comparación con la proporción respectiva de otras regiones de referencia específicas de TACS, lo que indicaría una anomalía genética fetal.

De este modo, los tamaños de fragmentos se dividen en dos grupos. Los tamaños relacionados con los *loci* de prueba se asignan a un grupo y los tamaños de fragmentos relacionados con los *loci* de referencia se asignan al otro grupo. Posteriormente, en cada grupo, los tamaños de fragmentos se distribuyen en dos subgrupos, donde los fragmentos pequeños se asignan a un subgrupo y todos los fragmentos restantes se asignan al otro subgrupo. El último paso calcula la proporción de fragmentos pequeños en cada grupo y usa estas cantidades en una prueba binomial de proporciones. La puntuación de la prueba se atribuye a la muestra investigada.

El resultado final de una muestra puede asignarse combinado una o más puntuaciones derivadas de los diferentes métodos estadísticos, ejemplos no taxativos de lo cual se muestran en el Ejemplo 4.

Kits

En otro aspecto, la divulgación provee kits para aplicar los métodos de la invención. En una realización, el kit comprende un contenedor que consiste de la mezcla de TACS e instrucciones para aplicar el método. En una realización, las TACS se proveen en un formato que permite unirlos a un sustrato sólido, por ejemplo, pueden ser TACS biotiniladas. En otra realización, las TACS se proveen junto con un sustrato sólido, por ejemplo, TACS biotiniladas provistas junto con microesferas magnéticas recubiertas con estreptavidina. En varias otras realizaciones, el kit puede comprender otros componentes para realizar otros aspectos del método. Por ejemplo, además de la mezcla de TACS, el kit puede comprender uno o más de los siguientes: (i) uno o más componentes para aislar ADN acelular de una muestra de plasma materno (p. ej., los que se describen en el Ejemplo 1); (ii) uno o más componentes para preparar la biblioteca de secuenciación (p. ej., cebadores, adaptadores, ligadores, enzimas de restricción, enzimas de ligación, polimerasas y similares, como los que se describen en el Ejemplo 1); (iii) uno o más componentes para amplificar y/o secuenciar la biblioteca enriquecida (p. ej., los que se describen en el Ejemplo 3); y/o (iv) software para realizar análisis estadístico (p. ej., el que se describe en el Ejemplo 4).

IV. Ejemplos

La presente invención se ilustra en mayor detalle mediante los siguientes ejemplos, pero no se debe interpretar que estos ejemplos limiten el alcance de la invención.

Ejemplo 1: Obtención de las muestras y preparación de las bibliotecas

Obtención de las muestras

Se obtuvieron muestras de plasma anónimas de mujeres embarazadas luego de la 10.^a semana de gestación. Los protocolos empleados para la obtención de muestras para nuestro estudio fueron aprobados por el Comité de Bioética Nacional de Chipre y se obtuvo el consentimiento informado de todas las participantes.

Extracción de las muestras

Se extrajo ADN acelular de 2 a 4 ml de plasma de cada individuo mediante un método de extracción manual o automático adecuado para aislar ADN acelular, como, por ejemplo —sin carácter taxativo— el protocolo Qiasymphony adecuado para aislar ADN fetal libre (Qiagen).

Preparación de las bibliotecas

El ADN extraído de las muestras de plasma materno se usaron para la construcción de bibliotecas de secuenciación. Se usaron métodos estándar de preparación de bibliotecas con las siguientes modificaciones (Meyer, M. and Kircher, M. (2010) Cold Spring Harb. Protoc. 2010(6):pdb prot5448). Se preparó una biblioteca de extracción independiente como control negativo para evaluar toda contaminación introducida durante el experimento. Durante este paso, las salientes 5' y 3' se completaron agregando 12 unidades de polimerasa T4 (NEB) y se incorporaron fosfatos 5' usando 40 unidades de polinucleótido quinasa T4 (NEB) en una reacción de 100 µl con posterior incubación a 25 °C durante 15 minutos y, luego, 12 °C durante 15 minutos. Los productos de reacción se purificaron con el kit MinElute (Qiagen). Posteriormente, se ligaron los adaptadores PS y P7 (ver la sección preparación de los adaptadores) diluidos 1: 10 a ambos extremos del ADN con 5 unidades de ligasa de ADN T4 (NEB) en una reacción de 40 µl durante 20 minutos a temperatura ambiente, seguido por una purificación con el kit MinElute (Qiagen). Se eliminaron las hendiduras de hebra simple (*nicks*) en una reacción de relleno con 16 unidades de polimerasa Bst (NEB) en una reacción de 40 µl con posterior incubación a 65 °C durante 25 minutos y, luego, 12 °C durante 20 minutos. Los productos se purificaron con el kit MinElute (Qiagen). La amplificación de las bibliotecas se realizó con una polimerasa Fusion (Herculase II Fusion DNA polymerase de Agilent Technologies o Pfusion High Fidelity Polymerase de NEB) en reacciones de 50 µl con las siguientes condiciones de termociclado: 95 °C durante 3 min; seguido por 10 ciclos a 95 °C durante 30 s, 60 °C durante 30 s, 72 °C durante 30 s y, por último, 72 °C durante 3 min. Los productos finales de las bibliotecas se purificaron con el kit de purificación MinElute (Qiagen) y se midieron por espectrofotometría.

Preparación de los adaptadores

Se prepararon mezclas de hibridación de los adaptadores P5 y P7 (26) por separado y se las incubó durante 10 s a 95 °C, seguido por una rampa de 95 °C a 12 °C a razón de 0,1° C/s. Las reacciones de P5 y P7 se combinaron para obtener una mezcla de adaptadores lista para usar (100 µM de cada adaptador). Las mezclas de hibridación se prepararon de la siguiente manera. La mezcla de reacción P5 contenía el adaptador P5_F (500 µM) a una concentración final de 200 µM y el adaptador P5+P7_R (500 µM) a una concentración final de 200 µM, con *buffer* de hibridación de oligonucleótidos en concentración 1 X. Por su parte, la mezcla de reacción P7 contenía el adaptador P7_F (500 µM) a una concentración final de 200 µM y el adaptador P5+P7_R (500 µM) a una concentración final de 200 µM, con *buffer* de hibridación de oligonucleótidos en concentración 1 X (30). Las secuencias fueron las siguientes, donde * representa un enlace fosforotioato (PTO) (Integrated DNA Technologies) (Meyer, M. y Kircher, M. (2010) Cold Spring Harb. Protoc. 2010(6):pdb prot5448):

adaptador P5_F:

A*C*A*C*TCTTTCCCTACACGACGCTCTTCCG*A*T*C*T (ID. DE SEC. N.º 1).

adaptador P7_F:

G*T*G*A*CTGGAGTTCAGACGTGTGCTCTTCCG*A*T*C*T (ID. DE SEC. N.º 2),

adaptador P5+P7_R:

A*G*A*T*CGGAA*G*A*G*C (ID. DE SEC. N.º 3).

Ejemplo 2: Diseño y preparación de secuencias de captura de blancos (TACS)

Se prepararon TACS personalizadas para la detección de anomalías cromosómicas totales o parciales en los cromosomas 13, 18, 21, X Y o cualquier otro, así como otras anomalías genéticas, incluidas, sin carácter taxativo, síndromes asociados con microdeleciones o microduplicaciones, traslocaciones, inversiones, inserciones y otras mutaciones puntuales o pequeñas. Los *loci* genómicos blancos usados para el diseño de las TACS se seleccionaron en función de su contenido de GC y su distancia a elementos de ADN repetitivo (distancia mínima de 150 bp). El tamaño de las TACS puede ser variable. En una realización del método, el tamaño de las TACS varía entre 100 y 260 bp y las TACS se generan mediante un método basado en la reacción en cadena de la polimerasa (PCR), que se describe a continuación. Las TACS se prepararon mediante una PCR símples con polimerasa Taq estándar, cebadores diseñados para amplificar los *loci* blanco y ADN normal como molde. En la Figura 2 se muestran las regiones cromosómicas usadas para diseñar cebadores para amplificar *loci* adecuados en los cromosomas 13, 18, 21 y X, con el fin de preparar la mezcla de TACS para el análisis de los cromosomas 13, 18, 21 y X. Todas las TACS personalizadas se generaron empleando las siguientes condiciones de termociclado: 95 °C durante 3 min; 40 ciclos de 95 °C durante 15 s, 60 °C durante 15 s, 72 °C durante 12 s; y 72 °C durante 12 s, seguido por la verificación mediante electroforesis en gel de agarosa y purificación mediante kits estándar de limpieza de productos de PCR como el Qiaquick PCR Purification Kit (Qiagen), el NucleoSpin 96

PCR Clean-up Kit (Macherey-Nagel) o el kit Agencourt AMPure XP for PCR Purification (Beckman Coulter). La concentración se midió con un instrumento NanoDrop (Thermo Scientific).

Ejemplo 3: Hibridación de las TACS y amplificación

Biotinilación de las TACS

Las TACS se prepararon para su hibridación de acuerdo con un protocolo descrito anteriormente (Maricic, T. et al. (2010) PLoS One 5:e14004) con pequeñas modificaciones. Se comenzó por obtener extremos romos con el kit Quick Blunting Kit (NEB) y una incubación a temperatura ambiente durante 30 minutos. Posteriormente, los productos de la reacción se purificaron con el kit MinElute (Qiagen) y se ligaron a un adaptador de biotina mediante el kit Quick Ligation Kit (NEB) en una reacción de 40 µl a temperatura ambiente durante 15 minutos. Los productos de la reacción se purificaron con el kit MinElute (Qiagen) y se desnaturalizaron para obtener ADN de hebra simple antes de su inmovilización en microesferas magnéticas recubiertas con estreptavidina (Invitrogen).

Hibridación de TACS

Las bibliotecas amplificadas se mezclaron con oligonucleótidos bloqueadores (Maricic, T. supra) (200 µM), 5 µg de ADN Cot-1 (Cot-1 DNA, Invitrogen), 50 µg de ADN de esperma de salmón (Salmon Sperm DNA, Invitrogen), *buffer* de hibridación Agilent en concentración 2 X y agente bloqueador Agilent en concentración 10 X, y se calentaron a 95 °C durante 3 min para desnaturalizar las hebras de ADN. Luego de la desnaturalización, se realizó una incubación durante 30 minutos a 37 °C para bloquear los elementos repetitivos y las secuencias de los adaptadores. La mezcla resultante se agregó a las TACS biotiniladas. Todas las muestras se incubaron en un incubador giratorio durante 12 a 48 h a 66 °C. Luego de la incubación, las microesferas se lavaron como se describió anteriormente y el ADN se eluyó por calentamiento (Maricic, T. supra). Los productos eluidos se amplificaron con cebadores de adaptadores dirigidos hacia fuera. Los productos amplificados enriquecidos se combinaron de forma equimolar y se secuenciaron en una plataforma Illumina o cualquier otra plataforma adecuada.

Ejemplo 4: Análisis bioinformático de las muestras

Alineación con el genoma humano

Por cada muestra, se aplicó la rutina bioinformática (*pipeline*) que se describe a continuación para alinear los fragmentos de ADN secuenciados de cada muestra con el genoma humano de referencia. Los fragmentos de lecturas apareadas dirigidas obtenidos de los resultados de la NGS se procesaron para eliminar las secuencias de adaptadores y las lecturas de baja calidad (puntuación Q < 25) con el software cutadapt (Martin, M. et al. (2011) EMB.netJournal 17.1). La calidad de las lecturas crudas y/o procesadas, junto con toda estadística descriptiva que ayudara a evaluar la calidad del resultado de la secuenciación de la muestra, se obtuvieron con el software FastQC (Babraham Institute (2015) FastQC) y/u otras herramientas de software propias. Las lecturas procesadas de al menos 25 bases de longitud se alinearon al genoma humano de referencia, versión hg19 (UCSC Genome Bioinformatics) con un algoritmo de alineación basado en la transformación de Burrows-Wheeler (Li, H. and Durbin, R. (2009) Bioinformatics 25: 1754-1760). Si correspondía, las lecturas duplicadas se eliminaron luego de la alineación. En los casos en que correspondía, el resultado de la secuenciación obtenido de la misma muestra, pero procesado en distintas calles (*lanes*) de secuenciación se combinaron en un único archivo de salida. Los procedimientos de eliminación de lecturas duplicadas y combinación se realizaron con el paquete de herramientas de software Picard (Broad Institute (2015) Picard) y/o el paquete de herramientas de software Sambamba (Tarasov, Artern, et al. "Sambamba: fast processing of NGS alignment formats." Bioinformatics 31.12 (2015): 2032-2034).

El análisis de software anterior resultó en una versión final alineada de una muestra secuenciada contra el genoma humano de referencia y todos los pasos ulteriores se basaron en esta versión alineada. La información relativa a los polimorfismos de nucleótidos individuales (SNP) en los *loci* de interés se obtuvo con la herramienta bcftools del paquete de software SAMtools (Li, H. et al. (2009) Bioinformatics 25:2078-2079). La profundidad de lecturas por base en los *loci* de interés —denominada, en lo sucesivo, archivo mpileup— se obtuvo con la opción mpileup del paquete de software SAMtools. La información relativa al tamaño de los fragmentos alineados —denominada, en lo sucesivo, archivo de tamaños de fragmentos— se obtuvo con la opción view del paquete de software SAMtools.

El archivo mpileup y el archivo de tamaños de fragmentos se procesaron con interfaces de programación de aplicación (API) personalizadas escritas en los lenguajes de programación Python y R (Python Software Foundation (2015) Python; The R Foundation (2015) The R Project for Statistical Computing). Las API se usaron para determinar la ploidía de los cromosomas de interés mediante una serie de pasos (denominados colectivamente, en lo sucesivo, "el algoritmo") y también para recopilar otras estadísticas descriptivas previstas para usar como métricas de control de calidad, tales como —sin carácter taxativo— la cuantificación de la fracción fetal (denominadas colectivamente, en lo sucesivo, "métricas de control de calidad"). Las API también pueden usarse para la evaluación de anomalías genéticas a partir de datos generados al aplicar el método descrito en casos de embarazos múltiples, así como otras anomalías genéticas como —sin carácter taxativo— las microdeleciones, las microduplicaciones, las variaciones en el número de copias, las traslocaciones, las inversiones, las inserciones, las mutaciones puntuales y los perfiles de mutaciones.

Métricas de control de calidad

Se emplearon métricas de control de calidad para inspeccionar las propiedades de una muestra alineada y decidir si la muestra era apta para clasificación. Estas métricas fueron, sin carácter taxativo, las siguientes:

(a) El enriquecimiento de una muestra. Los patrones de enriquecimiento son indicativos de si una muestra tiene un enriquecimiento adecuado entre los *loci* de interés en un experimento de secuenciación dado (denominado, en lo sucesivo, una "corrida"). Para evaluar esto, se consideran distintas métricas, que incluyen, sin carácter taxativo:

- (i) la profundidad de lecturas de la muestra en su conjunto que sean específicas ("en el blanco");
- (ii) el volumen de secuenciación de la muestra que sean específico ("en el blanco") con relación a la cantidad total de lecturas mapeadas;
- (iii) el rendimiento de TACS individuales en términos de la profundidad de lecturas alcanzada;
- (iv) la curtosis y la asimetría estadística del enriquecimiento de TACS individuales;
- y
- (v) los momentos de curtosis y asimetría estadística que surgen de todas las TACS.

Las comprobaciones anteriores también se tienen en consideración en relación con el enriquecimiento del sesgo inducido por el contenido de GC. Las muestras que no cumplen con uno o más de los criterios mencionados anteriormente se identifican para su inspección ulterior antes de la clasificación.

(b) La fracción fetal de una muestra. Las muestras con una fracción fetal estimada por debajo de un umbral específico no se clasifican.

El algoritmo

El algoritmo es una colección de rutinas de procesamiento de datos y modelos matemáticos y estadísticos organizados como una serie de pasos. Los pasos del algoritmo buscan determinar la ploidía individual de un cromosoma de interés en relación con todos los demás cromosomas de la muestra secuenciada y se usan para la detección de anomalías cromosómicas totales o parciales en los cromosomas 13, 18, 21, X, Y o cualquier otro, así como otras anomalías genéticas como, por ejemplo —sin carácter taxativo—, los síndromes de microdeleciones/microduplicaciones y otras mutaciones puntuales o pequeñas. Así, el algoritmo puede usarse, sin carácter taxativo, para la detección de anomalías cromosómicas totales o parciales en los cromosomas 13, 18, 21, X Y o cualquier otro, así como otras anomalías genéticas, incluidas, sin carácter taxativo, microdeleciones, microduplicaciones, variaciones en el número de copias, traslocaciones, inversiones, inserciones, mutaciones puntuales y otros perfiles de mutaciones. El algoritmo realiza, sin carácter taxativo, dos tipos de evaluaciones, una relativa a la información sobre la profundidad de lecturas en cada muestra y otra relativa a la distribución de tamaños de fragmentos en regiones específicas de las TACS. A cada tipo de evaluación pueden asociarse una o más pruebas estadísticas, ejemplos no taxativos de los cuales se proveen en los métodos estadísticos descritos en la presente.

En caso de pruebas asociadas con la profundidad de lecturas, el algoritmo compara secuencialmente la profundidad de lecturas de los *loci* de cada cromosoma de interés (denominado, en lo sucesivo, cromosoma de prueba) contra la profundidad de lecturas de todos los demás *loci* (denominados, en lo sucesivo, *loci* de referencia) para determinar su ploidía. Por cada muestra, los pasos fueron, sin carácter taxativo, los siguientes:

(a) Eliminar los loci secuenciados de forma inadecuada. Se obtuvo la profundidad de lecturas de cada *locus*. Los *loci* que no alcanzaron una cantidad mínima de lecturas se consideraron insuficientemente enriquecidos y se eliminaron antes de los pasos ulteriores.

(b) Mitigar el sesgo inducido por el contenido de GC. El procedimiento de secuenciación introduce discrepancias en la profundidad de lecturas entre los *loci* de interés en función de su contenido de GC. Para considerar tal sesgo, se empleó un nuevo enfoque de apareamiento de secuencias que aumenta tanto la sensibilidad y la especificidad a la hora de detectar aneuploidías cromosómicas. Se identificó el contenido de GC de cada *locus* del cromosoma de prueba y se agruparon los *loci* genéticos similares para formar grupos genéticamente apareados. El procedimiento se repitió con los *loci* de referencia. Luego, los grupos genéticamente apareados del cromosoma de prueba se aparearon condicionalmente con sus grupos genéticamente apareados contrapartes en el/los cromosoma(s) de referencia. Los grupos pueden tener cualquier cantidad de miembros. Luego, los grupos condicionalmente apareados se usaron para determinar la ploidía de los cromosomas de prueba.

(c) Determinar la ploidía. La determinación de la ploidía se realizó empleando un único método estadístico y/o una puntuación ponderada del resultado de los siguientes métodos estadísticos (sin carácter taxativo):

Método estadístico 1: Las diferencias en las profundidades de lecturas entre los grupos condicionalmente apareados se evaluaron para determinar su significación estadística con la fórmula de la prueba t :

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

donde t es el resultado de la prueba t , $\bar{x}$ es el promedio de las diferencias entre los grupos condicionalmente apareados, μ es la profundidad de lecturas esperada y se fija en un valor que representa diferencias de profundidad de lecturas insignificantes entre ambos grupos, s es la desviación estándar de las diferencias entre los grupos condicionalmente apareados y n es la longitud del vector de las diferencias condicionalmente apareadas. La magnitud del estadístico t se usó para identificar la evidencia, si correspondía, contra la hipótesis nula de igualdad de ploidía entre los cromosomas de referencia y de prueba. Específicamente, $t \geq c_1$ (donde c_1 es un umbral predefinido que pertenece al conjunto de todos los números positivos) indica evidencia en contra de la hipótesis nula. Los resultados del análisis de 98 muestras maternas correspondientes al cromosoma 21 con este método se muestran en la Figura 3.

Método estadístico 2: Remuestreo bivariado no paramétrico. El método de remuestreo (*bootstrap*) depende de la relación entre las variables aleatorias X (profundidad de lecturas de los *loci* de referencia) e Y (profundidad de lecturas de los *loci* de prueba). Aquí, tratamos la profundidad de lecturas de los señuelos en el grupo de referencia (variable aleatoria denotada por X) como la covariable independiente. El primer paso del procedimiento iterativo involucró un muestreo al azar con remuestreo (*bootstrapping*) de las profundidades de lecturas de los *loci* en los cromosomas de referencia, es decir, (x_1, g_1) , ..., (x_n, g_n) , donde el parámetro g es conocido y representa el contenido de GC del señuelo escogido. Luego, por cada señuelo de referencia seleccionado aleatoriamente (x_i, g_i) , se generó una profundidad de lecturas correspondiente para un *locus* genéticamente apareado, a saber, (y_1, g_1) , ..., (y_n, g_n) . Así, se obtuvieron los datos bivariados (x_1, y_1) , (x_2, y_2) , ..., (x_n, y_n) , condicionalmente apareados en función de su contenido de GC (parámetro g_i). Las diferencias entre las profundidades de lecturas de los valores remuestreados genéticamente apareados x_i e y_i se usaron para calcular el estadístico de interés en cada iteración. En una realización, esta medida estadística puede ser, sin carácter taxativo, la moda, la media o la mediana de las diferencias registradas, y/o múltiplos de dichos valores. El procedimiento se repitió la cantidad de veces necesarias para construir la distribución del estadístico de interés a partir de estas diferencias. Se asignó a la muestra una puntuación que corresponde a un percentil específico de la distribución construida (p. ej., el 5.º percentil). Bajo la hipótesis nula, la ploidía es la misma entre los cromosomas de los grupos de referencia y de prueba. Así, las muestras en las que la puntuación de un cromosoma en particular era mayor a un umbral predefinido, notado c_2 , se clasificaron como muestras en las que era estadísticamente improbable que la ploidía fuera la misma que en la referencia. También pueden emplearse otras medidas estadísticas. Los resultados del análisis de 98 muestras maternas correspondientes al cromosoma 21 con este método se muestran en la Figura 4.

Método estadístico 3: Prueba de permutaciones estratificadas. El estadístico de interés es la profundidad de lecturas estimada del cromosoma de prueba, denotada por $\widehat{Y}_{obs}$, que se calcula a partir de todos los *loci* de los grupos genéticamente apareados de los cromosomas de prueba de la siguiente manera:

$$\hat{Y}_{obs} = \frac{\sum_{j=1}^{\tau} \sum_{i=1}^{N_j} y_{ij}}{\sum_{j=1}^{\tau} N_j}$$

donde y_{ij} es la profundidad de lecturas del *loci* i que forma parte del grupo genéticamente apareado j (es decir, los *loci* que pertenecen a un grupo específico en virtud de su contenido de GC), N_j es la cantidad de *loci* de prueba que son parte del grupo genéticamente apareado j y τ es la cantidad de grupos genéticamente apareados.

A continuación, se construyó una distribución nula para evaluar $\hat{Y}_{obs}$. Para hacerlo, por cada grupo j , los *loci* de prueba y de referencia se combinaron (intercambiabilidad bajo la hipótesis nula) y cada grupo j se muestreó aleatoriamente hasta N_j sin reposición (permutación estratificada). Esto creó un vector de valores, notado y_i , del cual se calculó el valor promedio, notado $\hat{y}_i$. El procedimiento se repitió la cantidad de veces necesarias para construir la distribución nula. Por último, $\hat{Y}_{obs}$ se studentizó respecto de la distribución nula con la siguiente fórmula:

$$Z_{Y_{obs}} = \frac{\hat{Y}_{obs} - \hat{Y}}{\sigma_Y}$$

donde $\hat{Y}$ y σ_Y son el primer momento y la raíz cuadrada del segundo momento de todos los valores estadísticos $\hat{y}_i$ permutados. En las muestras cuyos valores de $Z_{Y_{obs}}$ eran mayores de un umbral predefinido, notado $c3$, era estadísticamente menos probable que la ploidía fuera la misma entre los grupos de referencia y de prueba. Los resultados del análisis de 98 muestras maternas correspondientes al cromosoma 21 con este método se muestran en la Figura 5.

En el caso de las pruebas asociadas con los tamaños de fragmentos, el algoritmo calcula la proporción de fragmentos pequeños en los *loci* de prueba y la compara con la correspondiente proporción en los *loci* de referencia, como se describe en el método estadístico 4, a continuación.

Método estadístico 4: Proporciones de tamaños de fragmentos. Por cada muestra, se extrae la cantidad y el tamaño de los fragmentos alineados al genoma humano de referencia en las coordenadas correspondientes a las TACS. Posteriormente, los datos se filtran para eliminar los tamaños de fragmentos que se consideran valores estadísticamente atípicos usando el método de detección de valores atípicos basado en la mediana. Específicamente, se definen como valores atípicos aquellos fragmentos cuyo tamaño es mayor o menor a los umbrales F_{thr} , dados por la siguiente ecuación:

$$F_{thr} = F_{median} \pm (X \times IQR)$$

donde F_{median} es la mediana del tamaño de todos los fragmentos de una muestra, X es una variable que puede asumir valores en el conjunto $\mathbb{R}^+$, e IQR es el rango intercuartílico de los tamaños de fragmentos. Posteriormente, se realiza una prueba binomial de proporciones para buscar evidencia en contra de la hipótesis nula, H_0 , definida de la siguiente manera:

H_0 : La proporción de fragmentos pequeños en la región de prueba no difiere de la proporción de fragmentos pequeños en la región de referencia.

En varias realizaciones de la invención, los fragmentos pequeños se definen como aquellos fragmentos cuyo tamaño es menor o igual a un subconjunto de $\mathbb{Z}^+$ acotado superiormente por 160 bp. Si definimos como T el conjunto de todas las TACS, la región de prueba puede ser cualquier subconjunto propio S que defina la región investigada, y la región de referencia es el complemento relativo de S en T . Por ejemplo, en una realización de la invención, el conjunto S se define como el conjunto de todas las secuencias capturadas por TACS del cromosoma 21 y, por consiguiente, el conjunto de referencia se define como el conjunto de todos los fragmentos capturados por TACS en los cromosomas de referencia, y/u otros *loci* de referencia.

La hipótesis alternativa, H_1 , se define de la siguiente manera:

H_1 : La proporción de fragmentos pequeños en la región de prueba difiere de la proporción de fragmentos pequeños en la región de referencia.

Así, teniendo en cuenta la corrección por continuidad, se calcula la siguiente puntuación (Brown et al, Harrel):

$$W_{test} = (\bar{p} - p_{ref}) / \sqrt{\frac{\bar{p}(1 - \bar{p})}{N_{test}}}$$

donde

$$\bar{p} = \frac{(\bar{F} + 0.5)}{(N_{test} + 1)}$$

$$p_{ref} = \frac{(F_{ref} + 0.5)}{(N_{ref} + 1)}$$

5 $\bar{F}$ es la cantidad de fragmentos pequeños en la región de prueba, F_{ref} es la cantidad de fragmentos pequeños en la región de referencia, N_{test} es la cantidad total de fragmentos en la región de prueba y N_{ref} es la cantidad total de fragmentos en la región de referencia.

10 Por cada muestra, el algoritmo contrasta secuencialmente la proporción de tamaños de fragmentos de las regiones investigadas (por ejemplo, sin carácter taxativo, el cromosoma 18, el cromosoma 21 y el cromosoma 13) contra las regiones de referencia, es decir, aquellas no investigadas en la prueba. Por cada muestra, se asigna una puntuación por cada prueba. Las puntuaciones superiores a un umbral, notado c4, proveen evidencia en contra de la hipótesis nula. Los resultados del análisis de 15 98 muestras maternas correspondientes al cromosoma 21 con este método se muestran en la Figura 6.

Método de puntuación ponderada 1: En una realización del método, se atribuyó a cada una de las muestras una puntuación ponderada, calculada como la suma ponderada de todos los métodos estadísticos de acuerdo con la siguiente fórmula:

20
$$V_s(R, F) = z_1 \max\{R_s, F_s\} + (1 - z_1) \min\{R_s, F_s\}$$

donde R_s es la puntuación corregida específica de la corrida obtenida de una contribución ponderada de cada método estadístico relacionado con la profundidad de lecturas de las muestras, que se define como:

$$R_s = \frac{(\sum_i w_i S_{is} - \bar{R}_r)}{\sigma_r}$$

25 y $\bar{R}_r$ es la mediana específica de la corrida, calculada a partir del vector de todas las puntuaciones no ajustadas ponderadas relacionadas con la profundidad de lecturas obtenidas de una única corrida de secuenciación, y σ_r es un múltiplo del desvío estándar de las puntuaciones R calculado a partir de un conjunto de referencia de 100 muestra euploides. Los términos $\max\{R_s, F_s\}$ y $\min\{R_s, F_s\}$ denotan los valores máximo y mínimo del conjunto entre llaves, respectivamente.

30 F_s es la puntuación corregida específica de la corrida obtenida del método estadístico relacionado con los tamaños de fragmentos y se define como:

$$F_s = \frac{(W_{test} - \bar{R}_f)}{\sigma_f}$$

35 donde W_{test} tiene la misma definición que antes, $\bar{R}_f$ es la mediana específica de la corrida, calculada a partir del vector de todas las puntuaciones estadísticas no ajustadas relacionadas con los fragmentos obtenidas de una única corrida de secuenciación, y σ_f es un múltiplo del desvío estándar de las puntuaciones F calculado a partir de un conjunto de referencia de 100 muestra euploides.

Una puntuación de clasificación única inferior a un valor predefinido indica que no hay evidencia en los datos observados de que una muestra tenga un riesgo significativo de aneuploidía. En las Figuras 7 y 8 se muestran los resultados correspondientes a las 98 muestras maternas usando el enfoque de puntuación ponderada sobre un subconjunto de los métodos o sobre todos los métodos, respectivamente.

Método de puntuación ponderada 2: En otra realización del método, la puntuación ponderada obtenida de los métodos estadísticos descritos anteriormente se usó para asignar a cada muestra una puntuación única de riesgo de aneuploidía de acuerdo con la siguiente fórmula:

$$R(t, c) = \sum_{j=0}^{j=N} w_j \frac{t_j}{c_j}$$

donde R es el resultado de la puntuación ponderada, w_j es el peso asignado al método j , t_j la puntuación observada resultante del método j , y c_j el umbral del método j .

Una puntuación de clasificación única inferior a un valor predefinido indica que no hay evidencia en los datos observados de que una muestra tenga un riesgo significativo de aneuploidía. En la Figura 9 se muestran los resultados de las 98 muestras maternas usando los métodos estadísticos 1 a 3 con el método de puntuación ponderada 2.

Dado que se asumió que todas las profundidades de lecturas de los señuelos del grupo de referencia provenían de la misma población y para tener un umbral universal, se emplearon ajustes específicos de cada corrida para aliviar los sesgos específicos de cada corrida.

El/los método(s) mencionado(s) también es/son adecuado(s) para la detección de otras anomalías genéticas, incluidas, sin carácter taxativo, las anomalías subcromosómicas. Un ejemplo no taxativo es la pérdida parcial contigua de material cromosómico que da lugar a una microdelección, o la incorporación parcial contigua de material cromosómico que da lugar a una microduplicación. Un *locus* genético conocido sujeto a ambas anomalías es el 7q11.23. En una realización del método estadístico 1, se evaluaron muestras de plasma sintéticas con 5 %, 10 % y 20 % de material fetal para evaluar el aumento del riesgo de microdelecciones y/o microduplicaciones en el *locus* genético 7q11.23. Los resultados se ilustran en la Figura 10 en los casos de microdelecciones y en la Figura 11 en los casos de microduplicaciones.

En el caso de las mutaciones puntuales, se realizan varias pruebas binomiales que consideran la estimación de la fracción fetal de la muestra, notada f , la profundidad de lecturas del alelo menor, notada r , y la profundidad de lecturas total de la base secuenciada, notada n . Dos ejemplos frecuentes, pero no taxativos involucran la evaluación del riesgo cuando la anomalía genética es una mutación puntual recesiva o una mutación puntual dominante.

En el ejemplo no taxativo de una mutación puntual recesiva, la hipótesis nula evaluada es que tanto la madre como el feto son heterocigotas (es decir, la frecuencia del alelo menor es de 0,5), en tanto que la hipótesis alternativa es que el feto es homocigoto (la frecuencia del alelo menor es 0,5-f/2). Un valor p pequeño en la correspondiente prueba de cocientes de verosimilitudes (*likelihood ratio*) indicaría evidencia contraria a la hipótesis nula. En el ejemplo no taxativo de una mutación puntual dominante, la hipótesis nula evaluada es que tanto la madre como el feto son homocigotas en la posición dada, en tanto que la hipótesis alternativa es que solo el feto es heterocigoto en la posición dada. Un valor p pequeño en la correspondiente prueba de cocientes de verosimilitudes (*likelihood ratio*) indicaría evidencia contraria a la hipótesis nula.

Además de los anteriores, se desarrollaron métodos de determinación del sexo, ejemplos no taxativos de los cuales se describen a continuación. En una realización de la invención, se asignó el sexo fetal a una muestra usando una prueba de Poisson dada por la siguiente fórmula:

$$Pr(r_y \leq k) = e^{-\lambda} \sum_{i=0}^{i=k} \frac{\lambda^i}{i!}$$

donde $\lambda = (fB\mu)/2$ y f es la fracción fetal estimada de la muestra, B es la cantidad de secuencias blanco en el cromosoma Y, μ es la profundidad de lecturas de la muestra y k es la suma de las

lecturas obtenidas de todos los blancos B. La hipótesis nula de la prueba de Poisson fue que la muestra era masculina. Un valor de $\Pr(r_y)$ menor a un umbral c_y se consideró como evidencia suficiente para rechazar la hipótesis nula, es decir, concluir que la muestra no era masculina. En los casos en los que uno o más de los términos para calcular $\Pr(r_y)$ no estaban disponibles, el sexo de la muestra se clasificó como no disponible (NA).

En otra realización de la invención, el sexo fetal se asignó usando la profundidad de lecturas promedio de las secuencias blanco en el cromosoma Y. Si la profundidad de lecturas promedio de las secuencias blanco era superior a un umbral predefinido, donde tal umbral puede definirse en función de otras características específicas de la muestra como la profundidad de lecturas y la fracción fetal estimada, el sexo fetal se clasificó como masculino. Si la profundidad de lecturas promedio era inferior a dicho umbral, la muestra se clasificó como femenina.

Estimación de la fracción fetal

Se han desarrollado diversos métodos para estimar la fracción fetal que pueden aplicarse a embarazos simples y/o múltiples. Así, en función del tipo de embarazo, la fracción fetal estimada puede obtenerse de cualquiera de los métodos o como una estimación ponderada a partir de un subconjunto de los métodos desarrollados y/o de todos ellos. A continuación, se dan algunos ejemplos no taxativos.

En una realización, se desarrolló una técnica de aprendizaje computarizado basada en la inferencia bayesiana para calcular la distribución *a posteriori* de la fracción fetal de ADN empleando los recuentos alélicos en los *loci* heterocigotas del plasma materno de embarazos simples. Se utilizaron tres combinaciones informativas posibles de genotipos maternos/fetales dentro del modelo para identificar aquellos valores de fracción fetal de ADN con mayor respaldo de los datos observados.

Sea f la fracción fetal de ADN. Si la madre es heterocigota en un locus dado, el genotipo fetal puede ser heterocigota u homocigota, lo que resulta en frecuencias esperadas del alelo menor de 0,5 y 0,5- $f/2$, respectivamente. Si la madre es homocigota y el feto es heterocigota, la frecuencia esperada del alelo menor será de $f/2$. Se empleó un método de Monte Carlo basado en una cadena de Markov (algoritmo de Metropolis-Hastings) (The R Foundation (2015) The R Project for Statistical Computing) con una distribución *a priori* no informativa o informativa (es decir, que incorporara información adicional como la edad gestacional, el peso materno, etc.) para obtener una secuencia de muestras aleatorias de la distribución de probabilidad *a posteriori* de la fracción fetal de ADN basada en un modelo de mezclas finitas.

En otra realización, la fracción fetal estimada se calcula únicamente a partir del clúster de frecuencia alélica menor (MAF) específico del feto, es decir, el clúster formado cuando la madre es homocigota y el feto es heterocigota en un *locus* genómico dado. Se asume que la media de la fracción fetal estimada tiene una distribución normal $N(2\bar{x}, \sigma_{\bar{x}})$, donde $\bar{x}$ es la media de la MAF específica del feto y $\sigma_{\bar{x}}$ es el desvío estándar de la MAF específica del feto. Luego, la fracción fetal estimada se obtiene de los percentiles de la distribución calculada, $N(2\bar{x}, \sigma_{\bar{x}})$.

En los embarazos de gestación múltiple —ejemplos no taxativos de los cuales incluyen los embarazos de gemelos monocigóticos y de mellizos dicigóticos, los embarazos de trillizos y distintos casos de donantes de óvulos y/o esperma—, la fracción fetal puede estimarse empleando información obtenida a partir de *loci* genéticos cuyo valor de MAF sea menor que un umbral, notado M_{thresh} , y derivada de posibles SNP específicos del feto. La persona razonablemente versada en la técnica apreciará que los SNP específicos del feto pueden originarse de cualquiera de los fetos, de cualquier combinación posible de los fetos o de todos los fetos de la gestación. Así, se ha desarrollado un algoritmo que estima la fracción fetal del feto con la menor contribución al contenido fetal total teniendo en cuenta la contribución combinatoria de cada feto a los valores de MAF que definen los SNP específicos del feto y también permite la contribución no homogénea de material fetal al contenido fetal total del material obtenido del plasma. Con este fin, el algoritmo emplea un enfoque de dos pasos.

En una realización del algoritmo, el embarazo múltiple considerado es un embarazo de mellizos dicigóticos. Como primer paso, la implementación algorítmica del modelo utiliza todos los SNP informativos y permite una contribución fetal no homogénea que puede explicarse con una diferencia relativa en las fracciones fetales estimadas respecto de un umbral definido, notado cf . Específicamente, si f_1 y f_2 representan las fracciones fetales de los fetos uno y dos, respectivamente, y $f_1 \leq f_2$, la suposición es que $f_2 \leq cf f_1$, donde cf es una constante real positiva mayor o igual a 1. Bajo esta hipótesis, los datos observados D , definidos como los recuentos de los alelos alternativo y de referencia en los *loci* de los SNP informativos, se suponen generados por una distribución combinada de tres binomiales (definidas por los parámetros $f_1/2$, $f_2/2$ y $(f_1+f_2)/2$), siendo la distribución *a posteriori* $p(f_1, f_2 | D)$ proporcional al modelo de observación, que puede expresarse

como $p(f_1|f_2, D)$ $p(f_2|D)$. La distribución *a posteriori* $p(f_1, f_2|D)$ se muestra con un algoritmo de Metropolis-Hastings MCMC usando una distribución *a priori* uniforme. El enfoque de cuantiles empíricos se aplica al arreglo de datos generado para inferir las fracciones fetales.

Como segundo paso, el algoritmo ejecuta un algoritmo de agrupamiento (*clustering*) basado en un modelo (modelo de mezcla gaussiana finita ajustado mediante el algoritmo EM; paquete mclust en R) para identificar si existe un clúster independiente de SNP atípicos del que se crea que está centrado en torno de $f_1/2$. La existencia de tal clúster con una media que invalide la hipótesis $cf \geq f_2/f_1$ lleva a la estimación de f_1 usando únicamente SNP que son parte del clúster identificado.

Referencias

Chris Fraley and Adrian E. Raftery (2002). Model-based Clustering, Discriminant Analysis and Density Estimation. Journal of the American Statistical Association, 97:611-631

Chris Fraley, Adrian E. Raftery, T. Brendan Murphy, and Luca Scrucca (2012). mclust Version 4 for R: Normal Mixture Modeling for Model-Based Clustering, Classification, and Density Estimation. Technical Report No. 597, Department of Statistics, University of Washington

Ejemplo 5: Resultados del análisis de muestras maternas

Se analizaron noventa y ocho muestras maternas para evaluar el riesgo de trisomía del cromosoma 21 (T21) de acuerdo con las metodologías descritas en los ejemplos 1 a 4. Los valores de puntuación del riesgo detectado de T21 en cada una de las 98 muestras con los métodos estadísticos 1, 2, 3 y 4 se muestran en las gráficas de las Figuras 3, 4, 5 y 6, respectivamente. Cada punto representa el valor de puntuación de una muestra individual. La línea representa el umbral "c" ($c = 3,00$ en el caso del método 1, $c = 5,00$ en el caso del método 2, $c = 4,00$ en el caso del método 3 y $c = 0,91$ en el caso del método 4). A las puntuaciones que superaron el umbral en la dirección positiva (es decir, valor de puntuación $> c$, puntos oscuros) se les asignó un riesgo alto de aneuploidía en comparación con las puntuaciones que no superaron el umbral (es decir, valor de puntuación $< c$, puntos claros). Con los cuatro métodos estadísticos, a cuatro muestras se les asignó un riesgo alto de aneuploidía T21. Todos los métodos estadísticos identificaron un riesgo elevado en las mismas muestras, a pesar de que las puntuaciones tenían distribuciones diferentes en los distintos métodos.

Las puntuaciones ponderadas resultantes de las combinaciones de los métodos estadísticos 1, 2, 3 y 4 del algoritmo (descritas en el Ejemplo 4) se representan en las gráficas de las Figuras 7 a 9, donde, una vez, más cada punto representa el valor de puntuación de una muestra individual y la línea representa el umbral. Los valores de puntuación ponderados identificaron un riesgo elevado en las mismas cuatro muestras.

Las cuatro muestras en las que se identificó un riesgo elevado de T21 mediante los análisis anteriores se evaluaron de forma independiente y se verificó que se trataba de embarazos con un feto con trisomía 21, lo que confirma la exactitud del método para la evaluación del riesgo de T21.

Los resultados del análisis de microdeleciones y microduplicaciones en muestras sintéticas se ilustran en las Figuras 10 y 11.

REIVINDICACIONES

1. Un método para evaluar el riesgo de una anomalía cromosómica en ADN fetal en una mezcla combinada de ADN materno y fetal, donde el método comprende:

(a) preparar la biblioteca de secuenciación a partir de la muestra combinada;

5 (b) hibridar la biblioteca de secuenciación a una mezcla de secuencias de captura de blancos (TACS), donde la mezcla de TACS comprende secuencias que se unen a uno o más cromosomas de interés que comprenden una anomalía cromosómica y donde:

10 (i) cada secuencia de la mezcla tiene entre 100 y 260 nucleótidos de longitud, donde cada secuencia tiene un extremo 5' y un extremo 3';
(ii) cada secuencia de la mezcla se une al/a los cromosoma(s) de interés a una distancia de al menos 150 pares de bases, tanto en el extremo 5' como en el extremo 3', de regiones que presentan variaciones en el número de copias (CNV), duplicaciones segmentales o elementos de ADN repetitivo; y
(iii) el contenido de GC de las TACS se encuentra entre 19 % y 50 %;

15 (c) aislar aquellos miembros de la biblioteca de secuenciación que se unen a las TACS para obtener una biblioteca enriquecida;

(d) amplificar y secuenciar la biblioteca enriquecida; y

(e) realizar análisis estadísticos sobre las secuencias de la biblioteca enriquecida para determinar así un riesgo asociado con la anomalía cromosómica en el ADN fetal.

20 2. El método de la reivindicación 1, donde la anomalía cromosómica es una aneuploidía.

3. El método de la reivindicación 1, donde el/los cromosoma(s) de interés se selecciona(n) de entre el grupo que comprende los cromosomas 13, 18, 21, X e Y.

25 4. El método de la reivindicación 1, donde la anomalía cromosómica es la trisomía 21.

5. El método de la reivindicación 1, donde la anomalía cromosómica es una anomalía estructural, incluidas, sin carácter taxativo, las variaciones en el número de copias, incluidas las microdelecciones y las microduplicaciones, las inserciones, las traslocaciones, las inversiones y las mutaciones pequeñas, incluidas las mutaciones puntuales y los perfiles de mutaciones.

30 6. El método de la reivindicación 1, donde la muestra combinada es una muestra de plasma materno.

35 7. El método de la reivindicación 1, donde la mezcla de TACS:

(i) está fija a un sustrato sólido o puede estar libre en solución y, opcionalmente, las TACS están biotiniladas y unidas a microesferas magnéticas recubiertas con estreptavidina, o bien están libres en solución;
40 (ii) la mezcla de TACS se une a múltiples cromosomas de interés, de manera que pueden detectarse múltiples anomalías cromosómicas; o
(iii) la mezcla de TACS contiene diferentes secuencias que se unen a los cromosomas 13, 18, 21, X e Y.

8. El método de la reivindicación 1, donde:

45 (i) el contenido de GC de las TACS se encuentra entre 19 % y 46 %, por ejemplo, entre 19 % y 43 %;
(ii) la mezcla de TACS comprende 800 o más secuencias distintas, por ejemplo, 1500 o más secuencias distintas, 1600 secuencias distintas, 2000 o más secuencias distintas, 2500 o más secuencias distintas, o 20 000 o más secuencias distintas; o
50 (iii) cada secuencia de la mezcla de TACS tiene entre 100 y 260 pares de bases de longitud.

9. El método de la reivindicación 1, donde la secuenciación de la biblioteca enriquecida provee una profundidad de lecturas correspondiente al cromosoma de interés y profundidades de lecturas correspondientes a los *loci* de referencia, y el análisis estadístico comprende aplicar un algoritmo que contrasta secuencialmente la profundidad de lecturas de los *loci* del cromosoma

de interés contra la profundidad de lecturas de los *loci* de referencia, donde el algoritmo comprende pasos para: (a) eliminar los *loci* secuenciados de forma inadecuada; (b) mitigar el sesgo inducido por el contenido de GC; y (c) determinar la ploidía.

- 5 10. El método de la reivindicación 9, donde:

 (i) el sesgo inducido por el contenido de GC se mitiga agrupando los *loci* con contenidos de GC equiparables; o
10 (ii) el método estadístico resulta en una puntuación asociada a la muestra combinada y se detecta el riesgo de la anomalía cromosómica en cuestión en el ADN fetal cuando la puntuación de la muestra combinada se encuentra por encima de un valor umbral de referencia.
- 15 11. El método de la reivindicación 1, donde secuenciar la biblioteca enriquecida provee la cantidad y el tamaño de los fragmentos secuenciados correspondientes a coordenadas específicas de las TACS y el análisis estadístico comprende aplicar un algoritmo que contrasta secuencialmente la proporción de tamaños de fragmento del cromosoma de interés contra la proporción de tamaños de fragmentos de los *loci* de referencia, donde el algoritmo comprende pasos para: (a) eliminar los valores atípicos de tamaños de fragmento; (b) calcular la proporción de tamaños de fragmentos; y (c) determinar la ploidía.
- 20 12. El método de las reivindicaciones 9 u 11, donde la determinación de la ploidía se realiza mediante la aplicación de uno o más métodos estadísticos.
- 25 13. El método de las reivindicaciones 1, 9 u 11, donde el/los método(s) estadístico(s) se selecciona(n) de entre el grupo que consiste en una prueba *t*, una prueba de remuestreo bivariada no paramétrica, una prueba de permutación estratificada y una prueba binomial de proporciones.
14. El método de la reivindicación 1, que comprende, además, la estimación de la fracción de ADN fetal en la muestra combinada.

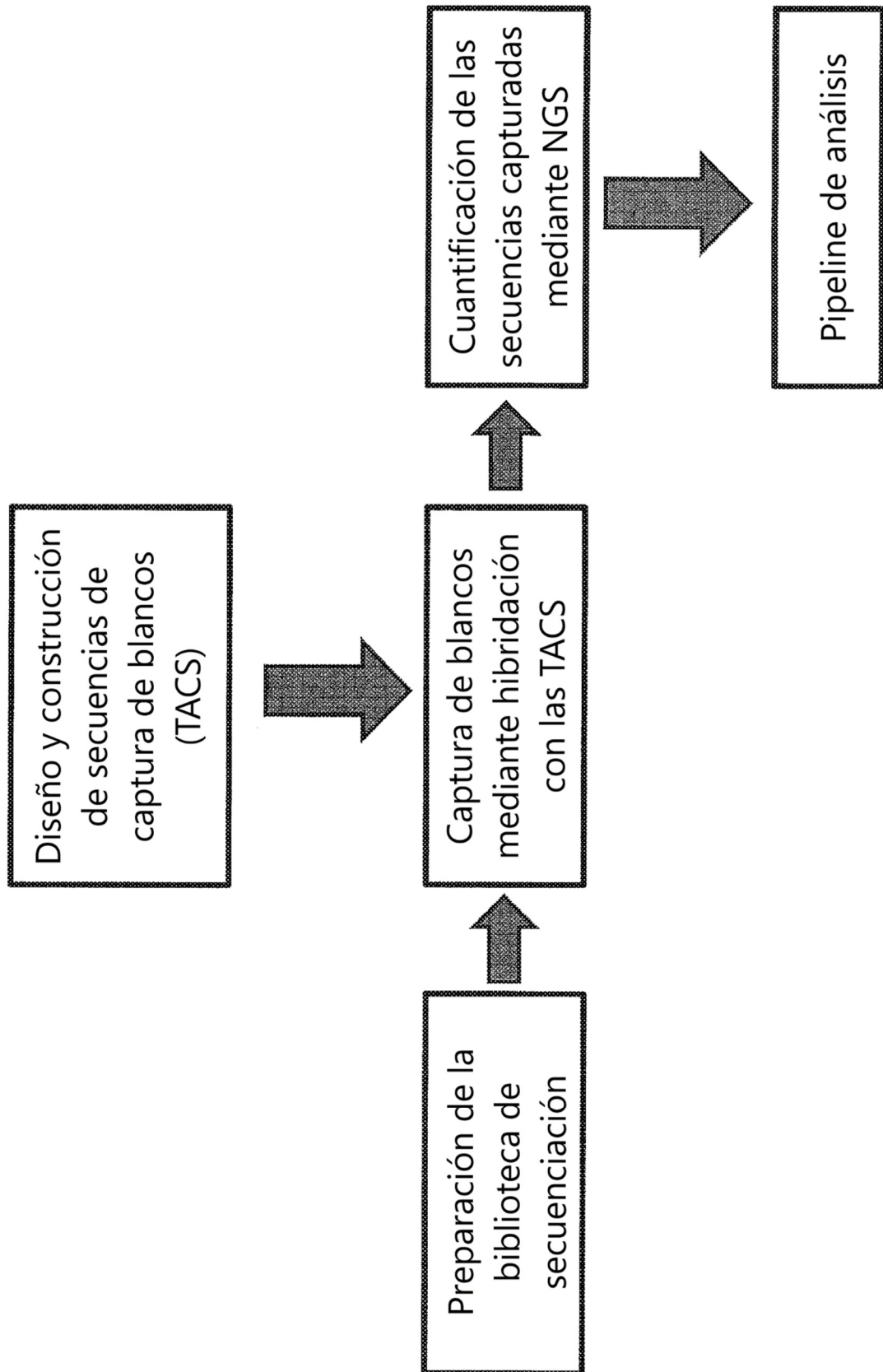


FIGURA 1

cromosoma	inicio	fin
chr13	91049752	91050001
chr13	92177139	92177388
chr13	93071049	93071298
chr13	943226973	943227222
chr13	94415507	94415756
chr13	95600034	95600283
chr13	112910529	112910778
chr13	53328405	53328654
chr13	20533985	20534234
chr13	20535728	20535977
chr13	20544227	20544476
chr13	20567359	20567607
chr13	20577159	20577407
chr13	20598772	20599021
chr13	20600556	20600804
chr13	20610779	20611027
chr13	20621336	20621585
chr13	20624392	20624641
chr13	20625461	20625710
chr13	20632609	20632857
chr13	20635196	20635444
chr13	21127405	21127653
chr13	21163998	21164247
chr13	21170950	21171199
chr13	21172515	21172764
chr13	21173774	21174022
chr13	21182580	21182828
chr13	21203385	21203634
chr13	21217173	21217422
chr13	21228403	21228652
chr13	21237418	21237666
chr13	21239719	21239968

cromosoma	inicio	fin
chr13	21245232	21245481
chr13	21251423	21251672
chr13	21276601	21276850
chr13	21296989	21297237
chr13	21321012	21321261
chr13	21330138	21330386
chr13	21348854	21349102
chr13	21351692	21351941
chr13	21356527	21356775
chr13	21361621	21361869
chr13	21396343	21396591
chr13	21399744	21399992
chr13	21417352	21417601
chr13	21422981	21423229
chr13	21607075	21607324
chr13	21631007	21631256
chr13	22114658	22114907
chr13	22130956	22131205
chr13	22189915	22190164
chr13	22270443	22270692
chr13	22396225	22396474
chr13	22609278	22609527
chr13	22691882	22692131
chr13	22765144	22765393
chr13	27175379	27175628
chr13	27405837	27406086
chr13	27646720	27646969
chr13	27674747	27674996
chr13	27969360	27969609
chr13	27978231	27978480
chr13	28003481	28003730
chr13	28073180	28073429

cromosoma	inicio	fin
chr13	28077809	28078058
chr13	28212786	28213035
chr13	28302961	28303210
chr13	28325874	28326123
chr13	28395893	28396142
chr13	28549591	28549840
chr13	28592602	28592851
chr13	28608293	28608542
chr13	28617043	28617292
chr13	28622471	28622720
chr13	28623594	28623843
chr13	28941450	28941699
chr13	28958898	28959147
chr13	28960375	28960624
chr13	29003985	29004234
chr13	29052686	29052935
chr13	52293092	52293341
chr13	52302656	52302905
chr13	53286404	53286653
chr13	53306303	53306552
chr13	53497224	53497473
chr13	53548644	53548893
chr13	53652715	53652964
chr13	53861872	53862121
chr13	53897429	53897678
chr13	53938205	53938454
chr13	54626806	54627055
chr13	54740174	54740423
chr13	54753834	54754083
chr13	54768238	54768487
chr13	54773333	54773582
chr13	54910969	54911218

FIGURA 2 (Página 1 de 31)

cromosoma	inicio	fin	cromosoma	inicio	fin	cromosoma	inicio	fin
chr13	54917347	54917596	chr13	111956688	111956937	chr13	60654975	60655224
chr13	54926251	54926500	chr13	112517966	112518215	chr13	61151386	61151635
chr13	55009732	55009981	chr13	112531519	112531768	chr13	62064603	62064852
chr13	59785342	59785591	chr13	112986709	112986958	chr13	71479359	71479608
chr13	60086302	60086551	chr13	113000054	113000303	chr13	39233534	39233783
chr13	60668020	60668269	chr13	21391429	21391678	chr13	39281667	39281916
chr13	60707747	60707996	chr13	21394616	21394865	chr13	39343524	39343773
chr13	60975949	60976198	chr13	21404268	21404516	chr13	39444897	39445146
chr13	61177164	61177413	chr13	21618233	21618482	chr13	39469458	39469707
chr13	61495655	61495904	chr13	21619521	21619769	chr13	39486237	39486486
chr13	67720914	67721163	chr13	21629602	21629851	chr13	39541262	39541511
chr13	67799557	67799806	chr13	21654127	21654376	chr13	39597307	39597556
chr13	70101783	70102032	chr13	22215795	22216044	chr13	39662550	39662799
chr13	71512697	71512946	chr13	22227261	22227510	chr13	39681445	39681694
chr13	71821654	71821903	chr13	22243011	22243260	chr13	39707053	39707302
chr13	72405844	72406093	chr13	22255153	22255402	chr13	39791099	39791348
chr13	74087455	74087704	chr13	22639387	22639636	chr13	39882804	39883053
chr13	74192477	74192726	chr13	27183101	27183350	chr13	39887397	39887646
chr13	74260893	74261142	chr13	27387690	27387939	chr13	40023278	40023527
chr13	78218064	78218313	chr13	30466189	30466438	chr13	40038277	40038526
chr13	88033439	88033688	chr13	30977198	30977447	chr13	40071926	40072175
chr13	90954378	90954627	chr13	31061235	31061484	chr13	40138914	40139163
chr13	91290109	91290358	chr13	31205037	31205286	chr13	40171414	40171663
chr13	91874767	91875016	chr13	32625308	32625557	chr13	40405751	40406000
chr13	92125914	92126163	chr13	34428102	34428351	chr13	40483291	40483540
chr13	92308079	92308328	chr13	36054899	36055148	chr13	40562330	40562579
chr13	92971372	92971621	chr13	36105686	36105935	chr13	40590631	40590880
chr13	94063232	94063481	chr13	40676702	40676951	chr13	40633944	40634193
chr13	94453836	94454085	chr13	40745383	40745632	chr13	40644042	40644291
chr13	94713643	94713892	chr13	49281723	49281972	chr13	40661582	40661831
chr13	95766126	95766375	chr13	53624527	53624776	chr13	40667939	40668188
chr13	95814599	95814848	chr13	60069277	60069526	chr13	40684306	40684555

FIGURA 2 (Página 2 de 31)

chromosoma	inicio	fin
chr13	40747946	40748195
chr13	40767281	40767530
chr13	41165017	41165266
chr13	41176869	41177118
chr13	41573543	41573792
chr13	41581284	41581533
chr13	41588844	41589093
chr13	41628316	41628565
chr13	41633944	41634193
chr13	41827812	41828061
chr13	42280338	42280587
chr13	42349979	42350228
chr13	42589616	42589865
chr13	42773088	42773337
chr13	42811933	42812182
chr13	44211957	44212206
chr13	44239246	44239495
chr13	44247433	44247682
chr13	44359516	44359765
chr13	44388992	44389241
chr13	44493443	44493692
chr13	44557886	44558135
chr13	44959162	44959411
chr13	45043950	45044199
chr13	45115235	45115484
chr13	45147386	45147635
chr13	47067063	47067312
chr13	47120595	47120844
chr13	47154937	47155186
chr13	47258834	47259083
chr13	47261849	47262098
chr13	47286641	47286890

chromosoma	inicio	fin
chr13	47314093	47314342
chr13	47623498	47623747
chr13	48562058	48562307
chr13	48698166	48698415
chr13	48762184	48762433
chr13	48806515	48806764
chr13	48811339	48811588
chr13	49119202	49119451
chr13	49479756	49480005
chr13	49505070	49505319
chr13	49530727	49530976
chr13	50387096	50387317
chr13	50960526	50960775
chr13	51783208	51783457
chr13	51816719	51816968
chr13	51825241	51825490
chr13	52025153	52025402
chr13	52249504	52249753
chr13	29111011	29111260
chr13	29192994	29193243
chr13	29233772	29234021
chr13	29310139	29310388
chr13	29563140	29563389
chr13	29580073	29580322
chr13	29929175	29929424
chr13	29934475	29934724
chr13	29939640	29939889
chr13	29957853	29958102
chr13	29995732	29995981
chr13	30003375	30003624
chr13	30014954	30015203
chr13	30114022	30114271

chromosoma	inicio	fin
chr13	30145894	30146143
chr13	30187253	30187502
chr13	30421518	30421767
chr13	30531947	30532196
chr13	30778718	30778967
chr13	30905760	30906009
chr13	31011560	31011809
chr13	31089936	31090185
chr13	31227324	31227573
chr13	31430437	31430686
chr13	31575060	31575309
chr13	31641685	31641934
chr13	31718167	31718416
chr13	31723797	31724044
chr13	32249897	32250146
chr13	32296888	32297137
chr13	32311823	32312072
chr13	32316395	32316644
chr13	32321163	32321412
chr13	32348262	32348511
chr13	32565108	32565357
chr13	32578964	32579213
chr13	32584041	32584290
chr13	32596887	32597136
chr13	32651705	32651954
chr13	32760468	32760717
chr13	32829312	32829561
chr13	32852622	32852871
chr13	32872003	32872252
chr13	33535420	33535669
chr13	33634729	33634978
chr13	33660366	33660615

FIGURA 2 (Página 3 de 31)

cromosoma	inicio	fin
chr13	33742309	33742558
chr13	33760763	33761012
chr13	33787480	33787729
chr13	33803219	33803468
chr13	33925858	33926107
chr13	34286418	34286667
chr13	34301780	34302029
chr13	34380783	34381032
chr13	35504924	35505173
chr13	35532835	35533084
chr13	35538772	35539021
chr13	36233165	36233414
chr13	36343466	36343715
chr13	36348003	36348252
chr13	36380590	36380839
chr13	36386774	36387023
chr13	36410707	36410956
chr13	36427833	36428082
chr13	36445018	36445267
chr13	36451093	36451342
chr13	36568336	36568585
chr13	36576882	36577131
chr13	36663599	36663848
chr13	37595672	37595921
chr13	38283224	38283473
chr13	38320894	38321143
chr13	38361974	38362223
chr13	38380696	38380945
chr13	38789453	38789702
chr13	39009512	39009761
chr13	39096011	39096260
chr13	39202882	39203131

cromosoma	inicio	fin
chr13	39264689	39264941
chr13	39265199	39265451
chr13	39265501	39265748
chr13	39266515	39266762
chr13	39266915	39267157
chr13	39322100	39322347
chr13	39338010	39338257

cromosoma	inicio	fin
-----------	--------	-----

FIGURA 2 (Página 4 de 31)

cromosoma	inicio	fin
chrX	44004660	44004909
chrX	44005572	44005821
chrX	80622102	80622351
chrX	81576088	81576337
chrX	83458870	83459119
chrX	83461989	83462238
chrX	83972195	83972444
chrX	83973091	83973340
chrX	84308155	84308404
chrX	84412933	84413182
chrX	84530826	84531075
chrX	84531356	84531605
chrX	84561384	84561633
chrX	85583338	85583587
chrX	85584098	85584347
chrX	85643367	85643616
chrX	85643847	85644096
chrX	85644765	85645014
chrX	85682212	85682461
chrX	85715710	85715959
chrX	85764768	85765017
chrX	85988749	85988998
chrX	85989214	85989463
chrX	85997574	85997823
chrX	87225841	87226090
chrX	91517109	91517358
chrX	91517578	91517827
chrX	94601042	94601291
chrX	95282210	95282459
chrX	95700693	95700942
chrX	96016317	96016566
chrX	96474722	96474971

cromosoma	inicio	fin
chrX	96475190	96475439
chrX	96522258	96522507
chrX	96685785	96686034
chrX	98226881	98227130
chrX	98227467	98227716
chrX	98314697	98314946
chrX	98331718	98331967
chrX	102986351	102986600
chrX	112117885	112118134
chrX	116461543	116461792
chrX	116746667	116746916
chrX	117121287	117121536
chrX	117494627	117494876
chrX	118400460	118400709
chrX	119760145	119760394
chrX	120436725	120436974
chrX	120437307	120437556
chrX	121854317	121854566
chrX	121855104	121855353
chrX	122863038	122863287
chrX	123203756	123204005
chrX	123989030	123989279
chrX	124394967	124395216
chrX	125455385	125455634
chrX	125500101	125500350
chrX	125500582	125500831
chrX	127973758	127974007
chrX	128274367	128274616
chrX	128274689	128274938
chrX	128275238	128275487
chrX	128694864	128695113
chrX	131201343	131201592

cromosoma	inicio	fin
chrX	132160888	132161137
chrX	133560531	133560780
chrX	133847304	133847553
chrX	133886077	133886326
chrX	133886778	133887027
chrX	135279738	135279987
chrX	135581031	135581280
chrX	136941543	136941792
chrX	138913313	138913562
chrX	138913920	138914169
chrX	139857034	139857283
chrX	139857318	139857567
chrX	141400062	141400311
chrX	141408144	141408393
chrX	141408548	141408797
chrX	142517923	142518172
chrX	143466384	143466633
chrX	147766073	147766322
chrX	147869477	147869726
chrX	147880766	147881015
chrX	147881811	147882060
chrX	151796091	151796340
chrX	104102171	104102420
chrX	104124449	104124698
chrX	104234605	104234854
chrX	104235195	104235444
chrX	104291380	104291629
chrX	104291786	104292035
chrX	104647680	104647929
chrX	104648691	104648940
chrX	106886942	106887191
chrX	3343487	3343736

FIGURA 2 (Página 5 de 31)

cromosoma	inicio	fin
chrX	4013011	4013260
chrX	5008870	5009119
chrX	5687782	5688031
chrX	5717003	5717252
chrX	5808998	5809247
chrX	5814187	5814436
chrX	5836574	5836823
chrX	5849776	5850025
chrX	6895474	6895723
chrX	6965226	6965475
chrX	6965573	6965822
chrX	7152743	7152992
chrX	9529739	9529988
chrX	9605323	9605572
chrX	9615158	9615407
chrX	9766329	9766578
chrX	9793372	9793621
chrX	9849298	9849547
chrX	9850182	9850431
chrX	10002581	10002830
chrX	10301512	10301761
chrX	10354762	10355011
chrX	10375361	10375610
chrX	10384260	10384509
chrX	10398645	10398894
chrX	10572719	10572968
chrX	10579800	10580049
chrX	10580482	10580731
chrX	10627502	10627751
chrX	10645064	10645313
chrX	10669664	10669913
chrX	10709674	10709923

cromosoma	inicio	fin
chrX	10710087	10710336
chrX	11139863	11140112
chrX	11658912	11659161
chrX	11681358	11681607
chrX	11686196	11686445
chrX	11686913	11687162
chrX	11687182	11687431
chrX	11917847	11918096
chrX	12016474	12016723
chrX	12158815	12159064
chrX	12485152	12485401
chrX	12608017	12608266
chrX	12724960	12725209
chrX	12841791	12842040
chrX	12904558	12904807
chrX	12937788	12938037
chrX	12938100	12938349
chrX	12992266	12992515
chrX	12995926	12996175
chrX	12997273	12997522
chrX	13293562	13293811
chrX	13338207	13338456
chrX	13340422	13340671
chrX	13351532	13351781
chrX	13624857	13625106
chrX	15525635	15525884
chrX	15543469	15543718
chrX	15543853	15544102
chrX	15844810	15845059
chrX	16860721	16860970
chrX	16861729	16861978
chrX	18307677	18307926

cromosoma	inicio	fin
chrX	18600682	18600931
chrX	19376384	19376633
chrX	20026004	20026253
chrX	20627108	20627357
chrX	21537202	21537451
chrX	21629967	21630216
chrX	21887062	21887311
chrX	22202443	22202692
chrX	23018625	23018874
chrX	23019006	23019255
chrX	24521525	24521774
chrX	24522047	24522296
chrX	25401534	25401783
chrX	25403300	25403549
chrX	28689766	28690015
chrX	28693600	28693849
chrX	28717026	28717275
chrX	28720131	28720380
chrX	28808128	28808377
chrX	28824813	28825062
chrX	28825844	28826093
chrX	29487025	29487274
chrX	30847194	30847443
chrX	32660941	32661190
chrX	32949975	32950224
chrX	35971182	35971431
chrX	35971544	35971793
chrX	43599991	43600240
chrX	43600262	43600511
chrX	43807492	43807741
chrX	63137429	63137678
chrX	63137928	63138177

FIGURA 2 (Página 6 de 31)

chromosoma	inicio	fin
chrX	136320360	136320609
chrX	139380500	139380749
chrX	139381074	139381323
chrX	144897923	144898172
chrX	144908837	144909086
chrX	144958353	144958602
chrX	144959377	144959626
chrX	145799354	145799603
chrX	145922717	145922966
chrX	146994477	146994726
chrX	146995311	146995560
chrX	147540977	147541226
chrX	149764787	149765036
chrX	150808513	150808762
chrX	150809248	150809497
chrX	150942801	150943050
chrX	79712826	79713075
chrX	80374192	80374441
chrX	80491243	80491492
chrX	80514786	80515035
chrX	80553311	80553560
chrX	104028154	104028403
chrX	104039199	104039448
chrX	104101675	104101924

chromosoma	inicio	fin
chrX	78401753	78402002
chrX	78403248	78403497
chrX	78405536	78405785
chrX	78411791	78412040
chrX	78427326	78427575
chrX	78430991	78431240
chrX	78431372	78431621
chrX	78437851	78438100
chrX	78526539	78526788
chrX	78531686	78531935
chrX	78996675	78996924
chrX	107682036	107682285
chrX	107683930	107684179
chrX	110366745	110366994
chrX	110367927	110368176
chrX	110368330	110368579
chrX	110443307	110443556
chrX	110466084	110466333
chrX	110537791	110538040
chrX	110538060	110538309
chrX	110538444	110538693
chrX	110575740	110575989
chrX	110576099	110576348
chrX	110950649	110950898
chrX	110971144	110971393
chrX	121597454	121597703
chrX	135558463	135558712
chrX	135570815	135571064
chrX	135760826	135761075
chrX	135831981	135832230
chrX	136044394	136044643
chrX	136318161	136318410

chromosoma	inicio	fin
chrX	74269728	74269977
chrX	74270926	74271175
chrX	74282022	74282271
chrX	74329664	74329913
chrX	74344769	74345018
chrX	74741441	74741690
chrX	76192017	76192266
chrX	76194203	76194452
chrX	76221949	76222198
chrX	76226569	76226818
chrX	76234682	76234931
chrX	76308591	76308840
chrX	76647025	76647274
chrX	76711476	76711725
chrX	76759387	76759636
chrX	76778190	76778439
chrX	76778847	76779096
chrX	76875971	76876220
chrX	76923678	76923927
chrX	76937441	76937690
chrX	76938848	76939097
chrX	76939481	76939730
chrX	76949897	76950146
chrX	76951314	76951563
chrX	77040328	77040577
chrX	77168344	77168593
chrX	77634270	77634519
chrX	78176076	78176325
chrX	78176555	78176804
chrX	78178901	78179150
chrX	78208245	78208494
chrX	78290291	78290540

FIGURA 2 (Página 7 de 31)

chromosoma	inicio	fin
chr21	17021878	17022127
chr21	17046718	17046967
chr21	17088576	17088825
chr21	17134784	17135033
chr21	17161709	17161958
chr21	17183683	17183932
chr21	17301983	17302232
chr21	17302814	17303063
chr21	17336649	17336898
chr21	17346228	17346477
chr21	17348050	17348299
chr21	17351309	17351558
chr21	17362239	17362488
chr21	17367713	17367962
chr21	17506670	17506919
chr21	17548792	17549041
chr21	17591325	17591574
chr21	17695933	17696182
chr21	19087043	19087292
chr21	19178365	19178614
chr21	22579319	22579568
chr21	22605212	22605461
chr21	22964667	22964916
chr21	22984832	22985081
chr21	22985386	22985635
chr21	36076596	36076845
chr21	36169368	36169617
chr21	36789088	36789337
chr21	37939587	37939836
chr21	37959766	37960015
chr21	37983316	37983565
chr21	38559539	38559788

chromosoma	inicio	fin
chr21	38751525	38751774
chr21	39173262	39173511
chr21	39285267	39285516
chr21	39330340	39330589
chr21	39442392	39442641
chr21	39499219	39499468
chr21	39541298	39541547
chr21	39563720	39563969
chr21	39584873	39585122
chr21	46530377	46530731
chr21	46534043	46534292
chr21	46535763	46536012
chr21	46537454	46537703
chr21	46542089	46542338
chr21	46556856	46557105
chr21	46564465	46564714
chr21	46567255	46567504
chr21	46572969	46573218
chr21	46605329	46605578
chr21	32134809	32135058
chr21	32148136	32148385
chr21	32165994	32166243
chr21	32174173	32174422
chr21	32179882	32180131
chr21	32181384	32181633
chr21	32184463	32184712
chr21	32201994	32202243
chr21	32208609	32208858
chr21	32319600	32319849
chr21	35034581	35034830
chr21	35060177	35060426
chr21	35070605	35070854

chromosoma	inicio	fin
chr21	35093305	35093554
chr21	35094930	35095179
chr21	35099418	35099667
chr21	35138457	35138706
chr21	35152877	35153126
chr21	35157173	35157422
chr21	35161795	35162044
chr21	35296082	35296331
chr21	35450331	35450580
chr21	35473022	35473271
chr21	35473971	35474220
chr21	35498164	35498413
chr21	35540943	35541192
chr21	35702965	35703214
chr21	28574819	28575068
chr21	28586377	28586626
chr21	28597009	28597258
chr21	28616121	28616370
chr21	28631408	28631657
chr21	28891355	28891604
chr21	28892223	28892472
chr21	28919446	28919695
chr21	29012106	29012355
chr21	29021007	29021256
chr21	29753618	29753867
chr21	29796479	29796728
chr21	29812165	29812414
chr21	29813966	29814215
chr21	29814951	29815200
chr21	29820028	29820277
chr21	29825992	29826241
chr21	29827616	29827865

FIGURA 2 (Página 8 de 31)

cromosoma	inicio	fin
chr21	29841662	29841911
chr21	29884231	29884480
chr21	29898255	29898504
chr21	30368230	30368479
chr21	30413906	30414155
chr21	30415555	30415804
chr21	30428953	30429202
chr21	30477650	30477899
chr21	30528161	30528410
chr21	30550132	30550381
chr21	30580738	30580987
chr21	30600446	30600695
chr21	30628668	30628917
chr21	30660609	30660858
chr21	30680391	30680640
chr21	30720214	30720463
chr21	30782462	30782711
chr21	30810199	30810448
chr21	30849402	30849651
chr21	30868525	30868774
chr21	30988776	30989025
chr21	31012406	31012655
chr21	31032720	31032969
chr21	31085356	31085605
chr21	31494755	31495004
chr21	31560667	31560916
chr21	31605973	31606222
chr21	31621826	31622075
chr21	31653822	31654071
chr21	31670830	31671079
chr21	31691347	31691596
chr21	31708936	31709185

cromosoma	inicio	fin
chr21	31733443	31733692
chr21	31768996	31769245
chr21	31886189	31886438
chr21	31886808	31887057
chr21	32119833	32120082
chr21	32125263	32125512
chr21	25795207	25795456
chr21	25802273	25802522
chr21	26355452	26355702
chr21	26365770	26366019
chr21	26373440	26373689
chr21	26374774	26375023
chr21	26390279	26390528
chr21	26390951	26391200
chr21	26402802	26403051
chr21	26496769	26497018
chr21	26515400	26515649
chr21	26571662	26571911
chr21	26589786	26590035
chr21	26604824	26605073
chr21	26933670	26933919
chr21	26950703	26950952
chr21	27045134	27045383
chr21	27398390	27398639
chr21	27441346	27441595
chr21	27465490	27465739
chr21	27485182	27485431
chr21	27486006	27486255
chr21	27502429	27502678
chr21	27520192	27520441
chr21	27539803	27540052
chr21	27623130	27623379

cromosoma	inicio	fin
chr21	27623692	27623941
chr21	27763679	27763928
chr21	27894000	27894249
chr21	27926498	27926747
chr21	28077177	28077426
chr21	28122462	28122711
chr21	28161337	28161586
chr21	28180925	28181174
chr21	28275754	28276003
chr21	28291449	28291698
chr21	28294820	28295069
chr21	28295351	28295600
chr21	28295857	28296106
chr21	28311866	28312115
chr21	28334784	28335033
chr21	28354814	28355063
chr21	28415341	28415590
chr21	28433186	28433435
chr21	28450226	28450475
chr21	28474083	28474332
chr21	28491062	28491311
chr21	28509439	28509688
chr21	28534176	28534425
chr21	40042441	40042690
chr21	40427338	40427587
chr21	40509976	40510225
chr21	40531291	40531540
chr21	40578323	40578572
chr21	40601158	40601407
chr21	40630577	40630826
chr21	41280723	41280972
chr21	41607273	41607522

FIGURA 2 (Página 9 de 31)

cromosoma	inicio	fin
chr21	16265484	16265733
chr21	16273734	16273983
chr21	16275865	16276114
chr21	16290234	16290483
chr21	16295877	16296126
chr21	16296637	16296936
chr21	16298842	16299091
chr21	16316213	16316462
chr21	16319836	16320085

cromosoma	inicio	fin
chr21	15843963	15844212
chr21	15858450	15858699
chr21	15859015	15859264
chr21	16002625	16002874
chr21	16030414	16030663
chr21	16063119	16063368
chr21	16068358	16068607
chr21	16071495	16071744
chr21	16081282	16081531
chr21	16115078	16115327
chr21	16122085	16122334
chr21	16130688	16130937
chr21	16133865	16134114
chr21	16139969	16140218
chr21	16141592	16141841
chr21	16160517	16160766
chr21	16165861	16166110
chr21	16189746	16189995
chr21	16190806	16191055
chr21	16192618	16192867
chr21	16195809	16196058
chr21	16198337	16198586
chr21	16203731	16203980
chr21	16207035	16207284
chr21	16209341	16209590
chr21	16211270	16211519
chr21	16219907	16220156
chr21	16230785	16231034
chr21	16236026	16236275
chr21	16241139	16241388
chr21	16242812	16243061
chr21	16246944	16247193

cromosoma	inicio	fin
chr21	41629820	41630069
chr21	41703462	41703711
chr21	41729089	41729338
chr21	41761636	41761885
chr21	41839021	41839270
chr21	41908902	41909151
chr21	41933321	41933570
chr21	41981868	41982117
chr21	42029142	42029391
chr21	42030642	42030891
chr21	42080190	42080439
chr21	42251219	42251468
chr21	42559559	42559808
chr21	42601828	42602077
chr21	42851340	42851589
chr21	42872713	42872962
chr21	43032728	43032977
chr21	43406661	43406910
chr21	46521300	46521549
chr21	16323548	16323797
chr21	16330249	16330498
chr21	16332622	16332871
chr21	16333315	16333564
chr21	16335355	16335604
chr21	16344361	16344610
chr21	16349606	16349855
chr21	16358650	16358899
chr21	15713251	15713500
chr21	15714357	15714606
chr21	15725977	15726226
chr21	15731366	15731615
chr21	15816232	15816481

FIGURA 2 (Página 10 de 31)

chromosoma	inicio	fin
chr18	64634331	64634580
chr18	64645925	64646174
chr18	64975938	64976187
chr18	64996573	64996822
chr18	65066425	65066674
chr18	65073729	65073978
chr18	65074170	65074419
chr18	66343957	66344206
chr18	66344406	66344655
chr18	66585247	66585496
chr18	66596489	66596738
chr18	66601231	66601480
chr18	69966208	69966457
chr18	70088939	70089188
chr18	70093364	70093613
chr18	70107935	70108184
chr18	70198868	70199117
chr18	70200128	70200377
chr18	70200907	70201156
chr18	70674976	70675225
chr18	70682967	70683216
chr18	73541191	73541440
chr18	73548315	73548564
chr18	73605288	73605537
chr18	73704485	73704734
chr18	73752546	73752795
chr18	73764755	73765004
chr18	73768420	73768669
chr18	73775027	73775276
chr18	73788751	73789000
chr18	73851108	73851357
chr18	73854713	73854962

chromosoma	inicio	fin
chr18	73871422	73871671
chr18	73876103	73876352
chr18	73879805	73880054
chr18	73884248	73884497
chr18	73931865	73932114
chr18	73938014	73938263
chr18	73943212	73943461
chr18	73947628	73947877
chr18	73964168	73964417
chr18	73966321	73966570
chr18	73974071	73974320
chr18	73985549	73985798
chr18	73987174	73987423
chr18	48466086	48466335
chr18	48928528	48928777
chr18	48929078	48929327
chr18	48951524	48951773
chr18	48951856	48952105
chr18	48959882	48960131
chr18	48960304	48960553
chr18	49004516	49004765
chr18	49006425	49006674
chr18	49017695	49017944
chr18	49148010	49148259
chr18	49149627	49149876
chr18	49149933	49150182
chr18	49150459	49150708
chr18	50239403	50239652
chr18	50242789	50243038
chr18	50243069	50243318
chr18	50243367	50243616
chr18	50247229	50247478

chromosoma	inicio	fin
chr18	50308727	50308976
chr18	50309031	50309280
chr18	50338747	50338996
chr18	50423284	50423533
chr18	50424422	50424671
chr18	50424789	50425038
chr18	50448837	50449086
chr18	50449492	50449741
chr18	50514079	50514328
chr18	50519889	50520138
chr18	56064244	56064493
chr18	56148666	56148915
chr18	56149037	56149286
chr18	56244171	56244420
chr18	56245242	56245491
chr18	56247063	56247312
chr18	56527049	56527298
chr18	56528011	56528260
chr18	56532190	56532439
chr18	56532882	56533131
chr18	56534208	56534457
chr18	56601487	56601736
chr18	56633131	56633380
chr18	56635259	56635508
chr18	56719296	56719545
chr18	58037828	58038077
chr18	58038328	58038577
chr18	58039111	58039360
chr18	58040290	58040539
chr18	59424085	59424334
chr18	59424765	59425014
chr18	59441324	59441573

FIGURA 2 (Página 11 de 31)

cromosoma	inicio	fin	cromosoma	inicio	fin	cromosoma	inicio	fin
chr18	59499718	59499967	chr18	60401084	60401333	chr18	60974839	60975088
chr18	59500226	59500475	chr18	60412199	60412448	chr18	60984378	60984627
chr18	59507920	59508169	chr18	60420053	60420302	chr18	61055997	61056246
chr18	59543044	59543293	chr18	60434561	60434810	chr18	61088669	61088918
chr18	59556453	59556702	chr18	60449047	60449296	chr18	61090462	61090711
chr18	59614976	59615225	chr18	60452471	60452720	chr18	61100325	61100574
chr18	59625553	59625802	chr18	60456757	60457006	chr18	61154459	61154708
chr18	59631090	59631339	chr18	60467147	60467396	chr18	61162901	61163150
chr18	59631779	59632028	chr18	60469757	60470006	chr18	61170629	61170878
chr18	59632388	59632637	chr18	60514168	60514417	chr18	61197492	61197741
chr18	59695718	59695967	chr18	60537888	60538137	chr18	61213452	61213701
chr18	59696087	59696336	chr18	60544380	60544629	chr18	61213900	61214149
chr18	59696563	59696812	chr18	60544742	60544991	chr18	61216749	61216998
chr18	59712419	59712668	chr18	60545407	60545656	chr18	61232928	61233177
chr18	59756979	59757228	chr18	60611934	60612183	chr18	61236079	61236328
chr18	59770789	59771038	chr18	60617081	60617330	chr18	61253277	61253526
chr18	59805993	59806242	chr18	60617641	60617890	chr18	61259478	61259727
chr18	59813876	59814125	chr18	60630438	60630687	chr18	61264533	61264782
chr18	59855411	59855660	chr18	60630849	60631098	chr18	61289440	61289689
chr18	59909054	59909303	chr18	60639840	60640089	chr18	61368889	61369138
chr18	59909336	59909585	chr18	60640300	60640549	chr18	61370663	61370912
chr18	59910140	59910389	chr18	60776944	60777193	chr18	63685629	63685878
chr18	59910602	59910851	chr18	60813738	60813987	chr18	63711583	63711832
chr18	59973803	59974052	chr18	60814585	60814834	chr18	63721183	63721432
chr18	59996728	59996977	chr18	60938666	60938915	chr18	64379625	64379874
chr18	60010541	60010790	chr18	60939202	60939451	chr18	64398368	64398617
chr18	60011381	60011630	chr18	60950541	60950790	chr18	64410092	64410341
chr18	60240694	60240943	chr18	60953225	60953474	chr18	64412606	64412855
chr18	60378353	60378602	chr18	60958474	60958723	chr18	64417267	64417516
chr18	60384593	60384842	chr18	60960879	60961128	chr18	64421562	64421811
chr18	60384979	60385228	chr18	60965583	60965832	chr18	64426871	64427120
chr18	60389544	60389793	chr18	60970051	60970300	chr18	64436287	64436536

FIGURA 2 (Página 12 de 31)

chromosoma	inicio	fin
chr18	64593727	64593976
chr18	24962594	24962843
chr18	24964748	24964997
chr18	25003553	25003802
chr18	25033345	25033594
chr18	25037434	25037683
chr18	25052992	25053241
chr18	25080042	25080291
chr18	25124915	25125164
chr18	25128057	25128306
chr18	25140257	25140506
chr18	25140722	25140971
chr18	25147759	25148008
chr18	25157828	25158077
chr18	25160244	25160493
chr18	25166354	25166603
chr18	25172986	25173235
chr18	25174972	25175221
chr18	25175640	25175889
chr18	25182855	25183104
chr18	25198531	25198780
chr18	25201886	25202135
chr18	25209395	25209644
chr18	25213880	25214129
chr18	25216888	25217137
chr18	25228617	25228866
chr18	25231535	25231784
chr18	25245662	25245911
chr18	25246300	25246549
chr18	25254843	25255092
chr18	25265476	25265725
chr18	25278532	25278781

chromosoma	inicio	fin
chr18	25279010	25279259
chr18	25283817	25284066
chr18	25292042	25292291
chr18	25295429	25295678
chr18	25330345	25330594
chr18	25332963	25333212
chr18	25333284	25333533
chr18	25345252	25345501
chr18	25384702	25384951
chr18	25510767	25511016
chr18	25519108	25519357
chr18	25531724	25531973
chr18	25543168	25543417
chr18	25547425	25547674
chr18	25557343	25557592
chr18	25564483	25564732
chr18	25583087	25583336
chr18	25585083	25585332
chr18	25589902	25590151
chr18	25601966	25602215
chr18	25611986	25612235
chr18	25617498	25617747
chr18	25629105	25629354
chr18	25631180	25631429
chr18	25643639	25643888
chr18	25645371	25645620
chr18	25665025	25665274
chr18	25666614	25666863
chr18	25667057	25667306
chr18	25670741	25670990
chr18	25673208	25673457
chr18	25681018	25681267

chromosoma	inicio	fin
chr18	25682079	25682328
chr18	25689906	25690155
chr18	25700540	25700789
chr18	25703648	25703897
chr18	25707188	25707437
chr18	25716100	25716349
chr18	25722951	25723200
chr18	25725055	25725304
chr18	25748105	25748354
chr18	26214441	26214690
chr18	26236788	26237037
chr18	26261336	26261585
chr18	26265937	26266186
chr18	26285286	26285535
chr18	26289006	26289255
chr18	26308405	26308654
chr18	26308962	26309211
chr18	26312054	26312303
chr18	26316149	26316398
chr18	47100905	47101154
chr18	47130278	47130527
chr18	47366936	47367185
chr18	47371131	47371380
chr18	47689868	47690117
chr18	47690267	47690516
chr18	47703689	47703938
chr18	47708139	47708388
chr18	47717445	47717694
chr18	47738531	47738780
chr18	47777438	47777687
chr18	47794212	47794461
chr18	47799400	47799649

FIGURA 2 (Página 13 de 31)

chromosoma	inicio	fin	chromosoma	inicio	fin	chromosoma	inicio	fin
chr18	47802662	47802911	chr18	43918519	43918768	chr18	46599847	46600096
chr18	39874986	39875235	chr18	43918998	43919247	chr18	46601866	46602115
chr18	39891490	39891739	chr18	43919544	43919793	chr18	46983278	46983527
chr18	42204240	42204489	chr18	44086123	44086372	chr18	47013103	47013352
chr18	42313146	42313395	chr18	44094299	44094548	chr18	47016802	47017051
chr18	42313939	42314188	chr18	44167214	44167463	chr18	286722	286971
chr18	42314302	42314551	chr18	44167791	44168040	chr18	298393	298642
chr18	42349923	42350172	chr18	44168296	44168545	chr18	338127	338376
chr18	42407165	42407414	chr18	44173512	44173761	chr18	340427	340676
chr18	42418033	42418282	chr18	44174626	44174875	chr18	387135	387384
chr18	42449702	42449951	chr18	44857765	44858014	chr18	391175	391424
chr18	42450400	42450649	chr18	45110525	45110774	chr18	434440	434689
chr18	42461427	42461676	chr18	45197621	45197870	chr18	457527	457776
chr18	42462834	42463083	chr18	45197871	45198120	chr18	460384	460633
chr18	42468595	42468844	chr18	45198760	45199009	chr18	469281	469530
chr18	42468938	42469187	chr18	45234886	45235135	chr18	572716	572965
chr18	42644003	42644252	chr18	45235523	45235772	chr18	598352	598601
chr18	42644494	42644743	chr18	45238370	45238619	chr18	615413	615662
chr18	42645387	42645636	chr18	45335896	45336145	chr18	618280	618529
chr18	42741059	42741308	chr18	45336353	45336602	chr18	901111	901360
chr18	42745733	42745982	chr18	45357461	45357710	chr18	939383	939632
chr18	42746460	42746709	chr18	45357745	45357994	chr18	26321283	26321532
chr18	42746911	42747160	chr18	45358502	45358751	chr18	26337865	26338114
chr18	42747473	42747722	chr18	45362379	45362628	chr18	26345717	26345966
chr18	43796659	43796908	chr18	45397588	45397837	chr18	26365387	26365636
chr18	43820471	43820720	chr18	45864183	45864432	chr18	26367767	26367767
chr18	43841913	43842162	chr18	46464382	46464631	chr18	26392500	26392749
chr18	43842202	43842451	chr18	46473428	46473677	chr18	26406499	26406748
chr18	43842647	43842896	chr18	46474545	46474794	chr18	26711590	26711839
chr18	43845701	43845950	chr18	46481797	46482046	chr18	26729342	26729591
chr18	43846353	43846602	chr18	46483016	46483265	chr18	26733981	26734230
chr18	43917652	43917901	chr18	46589014	46589263	chr18	26739336	26739585

FIGURA 2 (Página 14 de 31)

cromosoma	inicio	fin
chr18	26750685	26750934
chr18	26752487	26752736
chr18	26757573	26757822
chr18	27244230	27244479
chr18	35076347	35076596
chr18	35147912	35148161
chr18	35282655	35282904
chr18	35292141	35292390
chr18	35304247	35304496
chr18	35311302	35311551
chr18	36463709	36463958
chr18	36473033	36473282
chr18	36475403	36475652
chr18	36479145	36479394
chr18	36482038	36482287
chr18	36483959	36484208
chr18	36485761	36486010
chr18	36487627	36487876
chr18	36489135	36489384
chr18	36531244	36531493
chr18	36554006	36554255
chr18	36558050	36558299
chr18	36569619	36569868
chr18	1017928	1018177
chr18	241538	241787
chr18	242831	243080

FIGURA 2 (Página 15 de 31)

chromosoma	inicio	fin	chromosoma	inicio	fin	chromosoma	inicio	fin
chr13	19520254	19520499	chr13	26939424	26939672	chr13	32759010	32759261
chr13	19523790	19524042	chr13	26939942	26940196	chr13	32784101	32784351
chr13	20531155	20531403	chr13	26952790	26953038	chr13	32784570	32784820
chr13	20542998	20543246	chr13	26956549	26956797	chr13	32786822	32787069
chr13	21182580	21182828	chr13	26957204	26957453	chr13	32787129	32787381
chr13	22255353	22255601	chr13	26957736	26957984	chr13	32811213	32811463
chr13	26795002	26795250	chr13	26971501	26971749	chr13	32931779	32932029
chr13	26795380	26795628	chr13	26974853	26975100	chr13	33016510	33016761
chr13	26809864	26810112	chr13	26976810	26977059	chr13	33016929	33017182
chr13	26810422	26810669	chr13	27117460	27117707	chr13	33017432	33017681
chr13	26810951	26811198	chr13	27216274	27216526	chr13	36342494	36342743
chr13	26811424	26811674	chr13	29111011	29111260	chr13	36343935	36344179
chr13	26820545	26820794	chr13	29233772	29234021	chr13	363444330	36344577
chr13	26824800	26825048	chr13	29310139	29310388	chr13	36345572	36345822
chr13	26825155	26825405	chr13	29563140	29563389	chr13	36346001	36346253
chr13	26827150	26827400	chr13	31499212	31499458	chr13	36349204	36349453
chr13	26827690	26827939	chr13	31499519	31499772	chr13	36354263	36354513
chr13	26829149	26829396	chr13	31609109	31609361	chr13	36358322	36358572
chr13	26829663	26829911	chr13	31653571	31653822	chr13	36361603	36361850
chr13	26830241	26830491	chr13	32149129	32149380	chr13	36362050	36362297
chr13	26830739	26830987	chr13	32169142	32169392	chr13	36362531	36362778
chr13	26834633	26834880	chr13	32173373	32173622	chr13	36363439	36363689
chr13	26836514	26836761	chr13	32271292	32271539	chr13	36366103	36366353
chr13	26843066	26843315	chr13	32316109	32316362	chr13	36387184	36387431
chr13	26851259	26851507	chr13	32316651	32316898	chr13	39584383	39584630
chr13	26864367	26864615	chr13	32339360	32339608	chr13	39585052	39585299
chr13	26888243	26888485	chr13	32359633	32359883	chr13	39585675	39585922
chr13	26888660	26888907	chr13	32433382	32433629	chr13	40090054	40090298
chr13	26898691	26898936	chr13	32434990	32435237	chr13	40093695	40093942
chr13	26925436	26925684	chr13	32600409	32600659	chr13	40098460	40098707
chr13	26926976	26927224	chr13	32653613	32653857	chr13	40098913	40099157
chr13	26938928	26939175	chr13	32758635	32758877	chr13	40101053	40101303

FIGURA 2 (Página 16 de 31)

chromosoma	inicio	fin
chr13	40101582	40101832
chr13	40384942	40385189
chr13	40391440	40391687
chr13	40392221	40392468
chr13	40398309	40398557
chr13	40398589	40398840
chr13	40406460	40406711
chr13	40420794	40421037
chr13	40613790	40614037
chr13	42846940	42847187
chr13	42876485	42876732
chr13	42896653	42896897
chr13	43149529	43149781
chr13	43164772	43165022
chr13	43177841	43178091
chr13	52163649	52163899
chr13	52541402	52541652
chr13	52549279	52549529
chr13	53323273	53323520
chr13	53350485	53350727
chr13	53505942	53506191
chr13	53555869	53556121
chr13	53606740	53606987
chr13	53610142	53610389
chr13	53619813	53620063
chr13	53622959	53623209
chr13	59239399	59239649
chr13	62064603	62064852
chr13	68236363	68236610
chr13	68238595	68238842
chr13	68241140	68241393
chr13	68310701	68310948

chromosoma	inicio	fin
chr13	68311152	68311402
chr13	68316131	68316378
chr13	68428699	68428952
chr13	68429013	68429258
chr13	68570544	68570791
chr13	71269531	71269778
chr13	71454173	71454420
chr13	71456180	71456429
chr13	71463069	71463321
chr13	71464644	71464887
chr13	71465060	71465307
chr13	71466315	71466560
chr13	71466724	71466967
chr13	71479089	71479339
chr13	71479359	71479608
chr13	71619104	71619349
chr13	71750165	71750409
chr13	71750461	71750708
chr13	71762243	71762493
chr13	71762640	71762893
chr13	71781780	71782031
chr13	71783224	71783469
chr13	71784572	71784822
chr13	73226435	73226679
chr13	73233924	73234168
chr13	73251601	73251851
chr13	73261064	73261309
chr13	73284370	73284617
chr13	73296842	73297089
chr13	73305479	73305731
chr13	73305843	73306093
chr13	73309111	73309357

chromosoma	inicio	fin
chr13	73320095	73320339
chr13	77198180	77198433
chr13	77422635	77422880
chr13	77449372	77449622
chr13	77454212	77454458
chr13	77454985	77455228
chr13	77455721	77455971
chr13	77456394	77456643
chr13	77476090	77476339
chr13	77477580	77477833
chr13	77513439	77513692
chr13	77591495	77591742
chr13	77595913	77596160
chr13	77598010	77598257
chr13	77630724	77630968
chr13	77633431	77633680
chr13	77636364	77636611
chr13	77639621	77639873
chr13	77642789	77643041
chr13	77652265	77652509
chr13	77652705	77652952
chr13	77675169	77675423
chr13	77699605	77699849
chr13	77713099	77713352
chr13	77714235	77714481
chr13	77742675	77742920
chr13	77759281	77759528
chr13	78540689	78540943
chr13	89222075	89222318
chr13	94063232	94063481

FIGURA 2 (Página 17 de 31)

chromosoma	inicio	fin	chromosoma	inicio	fin	chromosoma	inicio	fin
chr18	11701840	11702094	chr18	13040411	13040657	chr18	29705097	29705347
chr18	11742882	11743124	chr18	13049051	13049304	chr18	29707000	29707248
chr18	11743713	11743964	chr18	13053013	13053260	chr18	29711514	29711758
chr18	11743974	11744216	chr18	13057128	13057370	chr18	29772771	29773020
chr18	11745797	11746044	chr18	13725324	13725566	chr18	29773038	29773291
chr18	11766505	11766752	chr18	13731669	13731919	chr18	29790110	29790357
chr18	11784594	11784841	chr18	13930816	13931068	chr18	29794196	29794446
chr18	11784857	11785107	chr18	13936510	13936761	chr18	29814455	29814703
chr18	11789414	11789659	chr18	13939030	13939277	chr18	29816677	29816924
chr18	11791405	11791649	chr18	13941500	13941742	chr18	29824628	29824872
chr18	11831898	11832147	chr18	18709626	18709868	chr18	29941738	29941986
chr18	11840071	11840318	chr18	19345013	19345263	chr18	29942404	29942657
chr18	11853018	11853271	chr18	22314388	22314639	chr18	29947043	29947286
chr18	12786330	12786572	chr18	22317204	22317445	chr18	29963605	29963859
chr18	12793790	12794034	chr18	22318584	22318826	chr18	29967321	29967565
chr18	12813342	12813590	chr18	22474179	22474425	chr18	29988241	29988491
chr18	12813796	12814046	chr18	22503094	22503336	chr18	29989060	29989304
chr18	12836508	12836756	chr18	22514422	22514675	chr18	33702205	33702458
chr18	12837826	12838071	chr18	22527670	22527920	chr18	33703293	33703541
chr18	12839370	12839619	chr18	22713619	22713863	chr18	33703805	33704052
chr18	12839810	12840054	chr18	22715692	22715939	chr18	33710269	33710511
chr18	12840789	12841033	chr18	22746015	22746258	chr18	33881399	33881647
chr18	12955546	12955788	chr18	22751334	22751581	chr18	33884883	33885133
chr18	12968393	12968643	chr18	26362055	26362304	chr18	33965911	33966161
chr18	12984202	12984449	chr18	27204825	27205070	chr18	33967910	33968160
chr18	12984953	12985198	chr18	27683185	27683429	chr18	35703018	35703269
chr18	12986018	12986263	chr18	28529366	28529610	chr18	35705739	35705986
chr18	13007427	13007681	chr18	28538550	28538803	chr18	36487627	36487876
chr18	13007815	13008064	chr18	28550611	28550855	chr18	52428207	52428460
chr18	13019314	13019562	chr18	28551683	28551928	chr18	52472800	52473050
chr18	13029804	13030046	chr18	28584849	28585102	chr18	52702246	52702490
chr18	13036834	13037081	chr18	29704799	29705048	chr18	52703207	52703449

FIGURA 2 (Página 18 de 31)

cromosoma	inicio	fin
chr18	52719979	52720226
chr18	53843255	53843507
chr18	53843510	53843764
chr18	53843879	53844121
chr18	53870090	53870332
chr18	53873286	53873528
chr18	53913040	53913287
chr18	54141132	54141379
chr18	54154870	54155117
chr18	54183137	54183381
chr18	54271758	54272011
chr18	54352988	54353238
chr18	60378353	60378602
chr18	61164144	61164393

FIGURA 2 (Página 19 de 31)

cromosoma	inicio	fin
chr21	15774354	15774598
chr21	15774701	15774951
chr21	15941894	15942144
chr21	15942369	15942621
chr21	16330536	16330783
chr21	16331481	16331728
chr21	16332238	16332485
chr21	16333694	16333941
chr21	16335023	16335271
chr21	16336651	16336901
chr21	16337562	16337816
chr21	16338381	16338633
chr21	16338902	16339154
chr21	16339772	16340019
chr21	16340409	16340656
chr21	16344080	16344330
chr21	16349101	16349354
chr21	16362700	16362954
chr21	16375294	16375539
chr21	16376059	16376308
chr21	16376695	16376942
chr21	16519760	16520004
chr21	16529308	16529552
chr21	16529940	16530189
chr21	16531755	16532002
chr21	16533111	16533355
chr21	16537385	16537631
chr21	16545527	16545774
chr21	16788480	16788722
chr21	16792457	16792708
chr21	17149630	17149877
chr21	45114956	45115203

cromosoma	inicio	fin
chr21	17160410	17160656
chr21	17160972	17161216
chr21	19300496	19300748
chr21	19396568	19396818
chr21	25701930	25702183
chr21	25702313	25702560
chr21	27985131	27985378
chr21	27989556	27989803
chr21	28000089	28000336
chr21	28007209	28007456
chr21	28007550	28007793
chr21	28008308	28008555
chr21	28054591	28054835
chr21	28292245	28292493
chr21	28294400	28294644
chr21	28308980	28309224
chr21	28311566	28311813
chr21	28316801	28317045
chr21	29796060	29796309
chr21	30112410	30112660
chr21	30112772	30113024
chr21	30113106	30113350
chr21	30114870	30115119
chr21	30301073	30301322
chr21	30301466	30301713
chr21	30317970	30318220
chr21	30358601	30358851
chr21	30397555	30397801
chr21	30419470	30419720
chr21	30527461	30527714
chr21	30528513	30528763
chr21	30554566	30554808

cromosoma	inicio	fin
chr21	30568450	30568700
chr21	30901196	30901442
chr21	30950156	30950403
chr21	30950626	30950878
chr21	30952977	30953227
chr21	30953281	30953523
chr21	30953836	30954083
chr21	30966825	30967077
chr21	30968330	30968574
chr21	30969659	30969903
chr21	31015490	31015733
chr21	31022236	31022488
chr21	31091818	31092065
chr21	31109787	31110039
chr21	31110149	31110396
chr21	31113996	31114249
chr21	31120815	31121064
chr21	31144588	31144835
chr21	35154438	35154687
chr21	40549354	40549598
chr21	40759240	40759493
chr21	40839325	40839569
chr21	40846844	40847093
chr21	41160090	41160343
chr21	41171755	41171999
chr21	41298214	41298467
chr21	41592339	41592583
chr21	41596188	41596432
chr21	41741900	41742144
chr21	43219112	43219356
chr21	46468654	46468898
chr21	46495701	46495954

FIGURA 2 (Página 20 de 31)

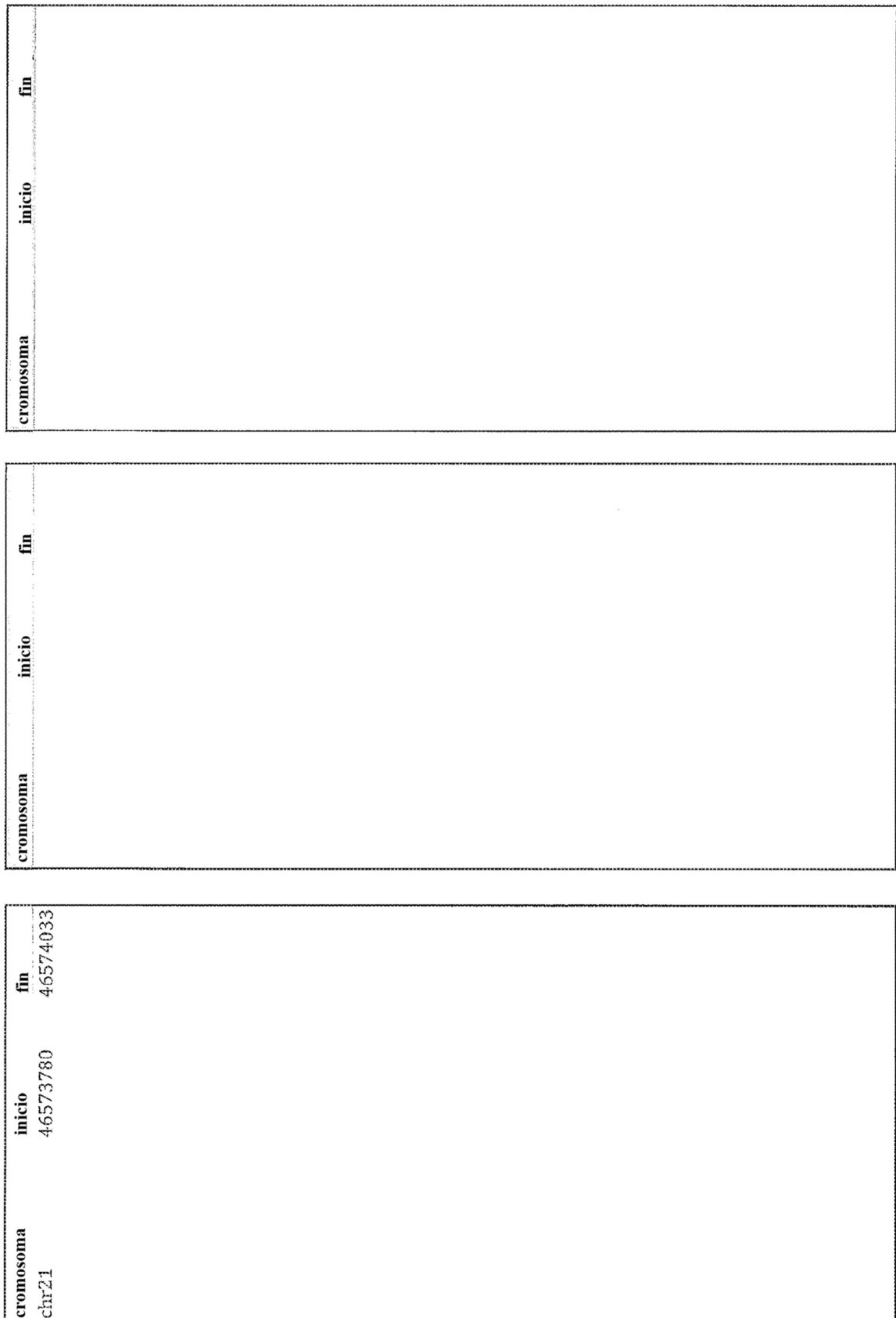


FIGURA 2 (Página 21 de 31)

cromosoma	inicio	fin
chr2	50176039	50176288
chr2	63596895	63597144
chr2	103538910	103539159

FIGURA 2 (Página 22 de 31)

cromosoma	inicio	fin
chr3	23866353	23866602

FIGURA 2 (Página 23 de 31)

cromosoma	inicio	fin
-----------	--------	-----

[illegible]

chromosoma	inicio	fin
chr4	21947614	21947863
chr4	62556697	62556946
chr4	72418734	72418983
chr4	122570216	122570465
chr4	130264593	130264842

FIGURA 2 (Página 24 de 31)

cromosoma	inicio	fin
chr5	59283789	59284038
chr5	88597843	88598092
chr5	163891445	163891694
chr5	163896919	163897168

FIGURA 2 (Página 25 de 31)

cromosoma	inicio	fin
chr6	8529895	8530144
chr6	96221217	96221466
chr6	112441026	112441275
chr6	118481182	118481431
chr6	129800139	129800388

FIGURA 2 (Página 26 de 31)

cromosoma	inicio	fin
chr7	22725339	22725588
chr7	42250786	42251035

FIGURA 2 (Página 27 de 31)

cromosoma	inicio	fin
chr9	2600491	2600740
chr9	14299731	14299980
chr9	72782230	72782479
chr9	82880926	82881175

FIGURA 2 (Página 28 de 31)

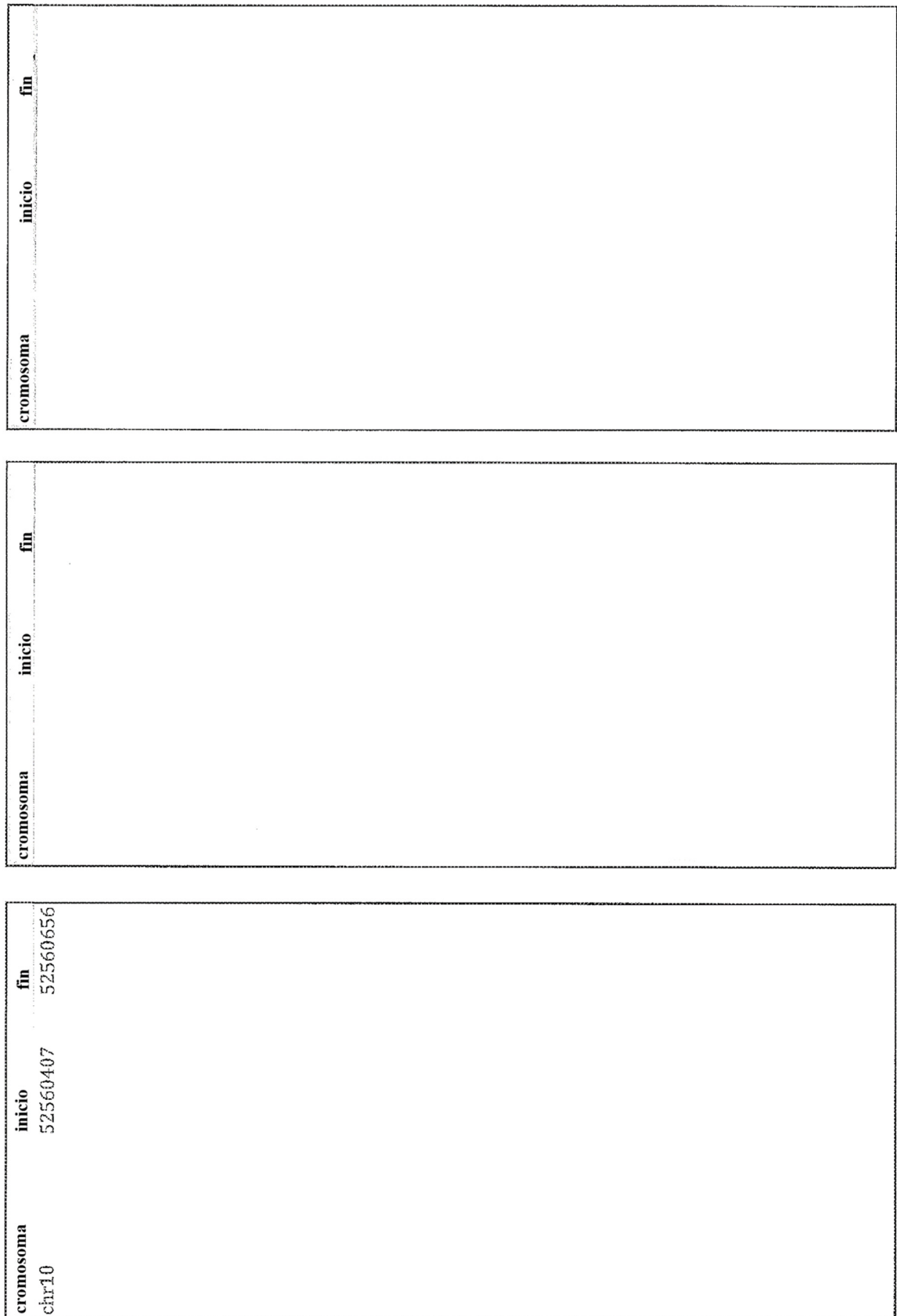


FIGURA 2 (Página 29 de 31)

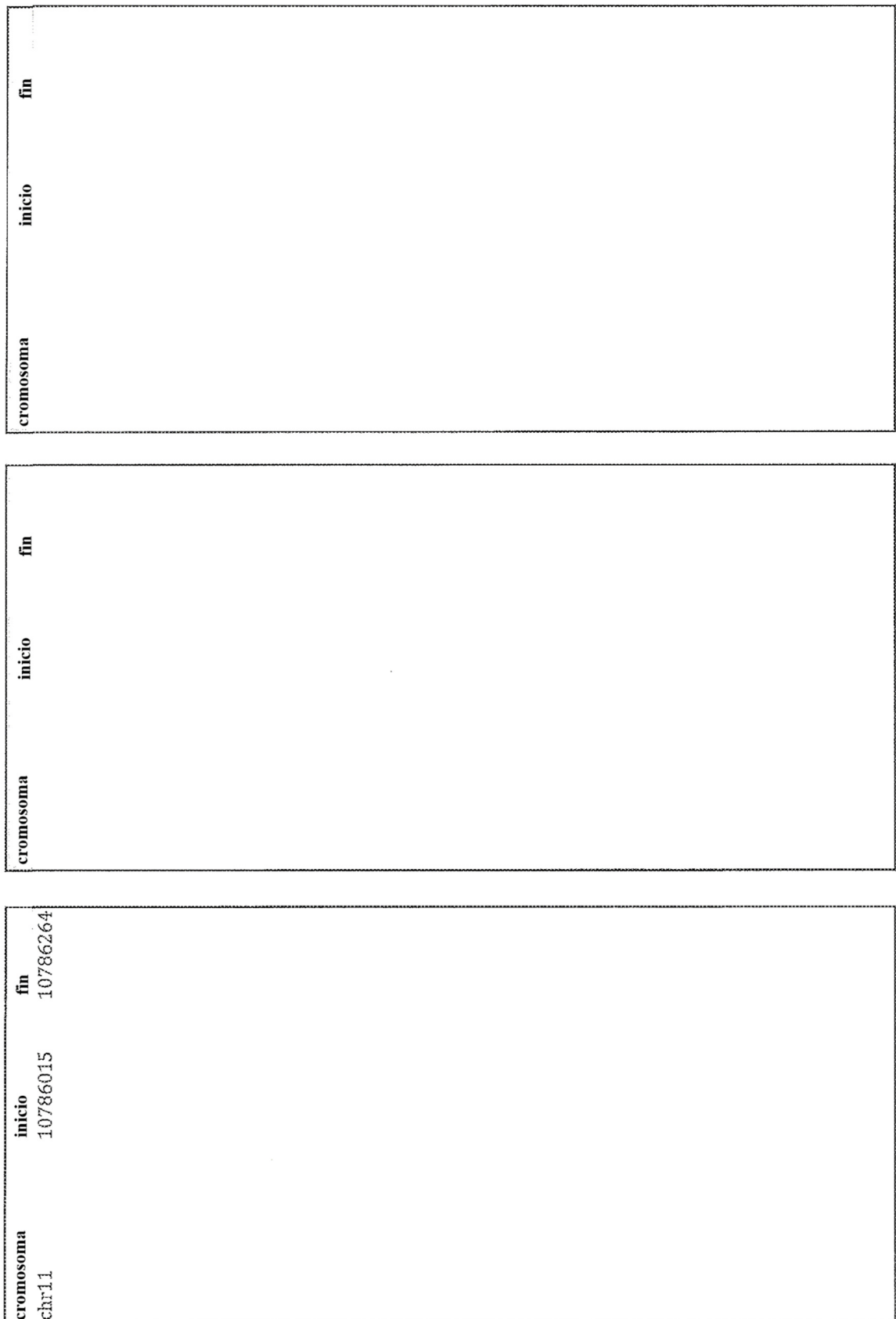


FIGURA 2 (Página 30 de 31)

55

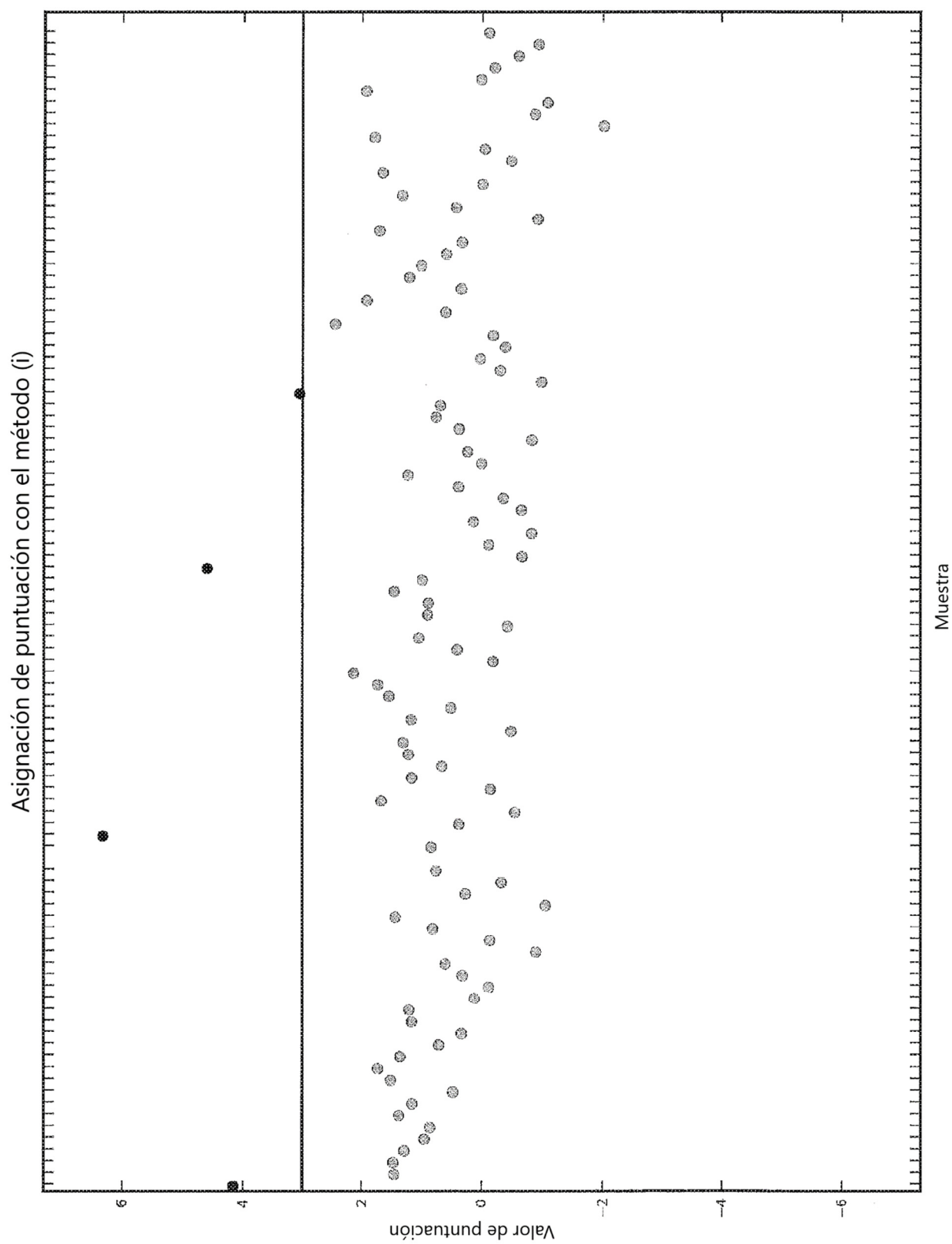


FIGURA 3

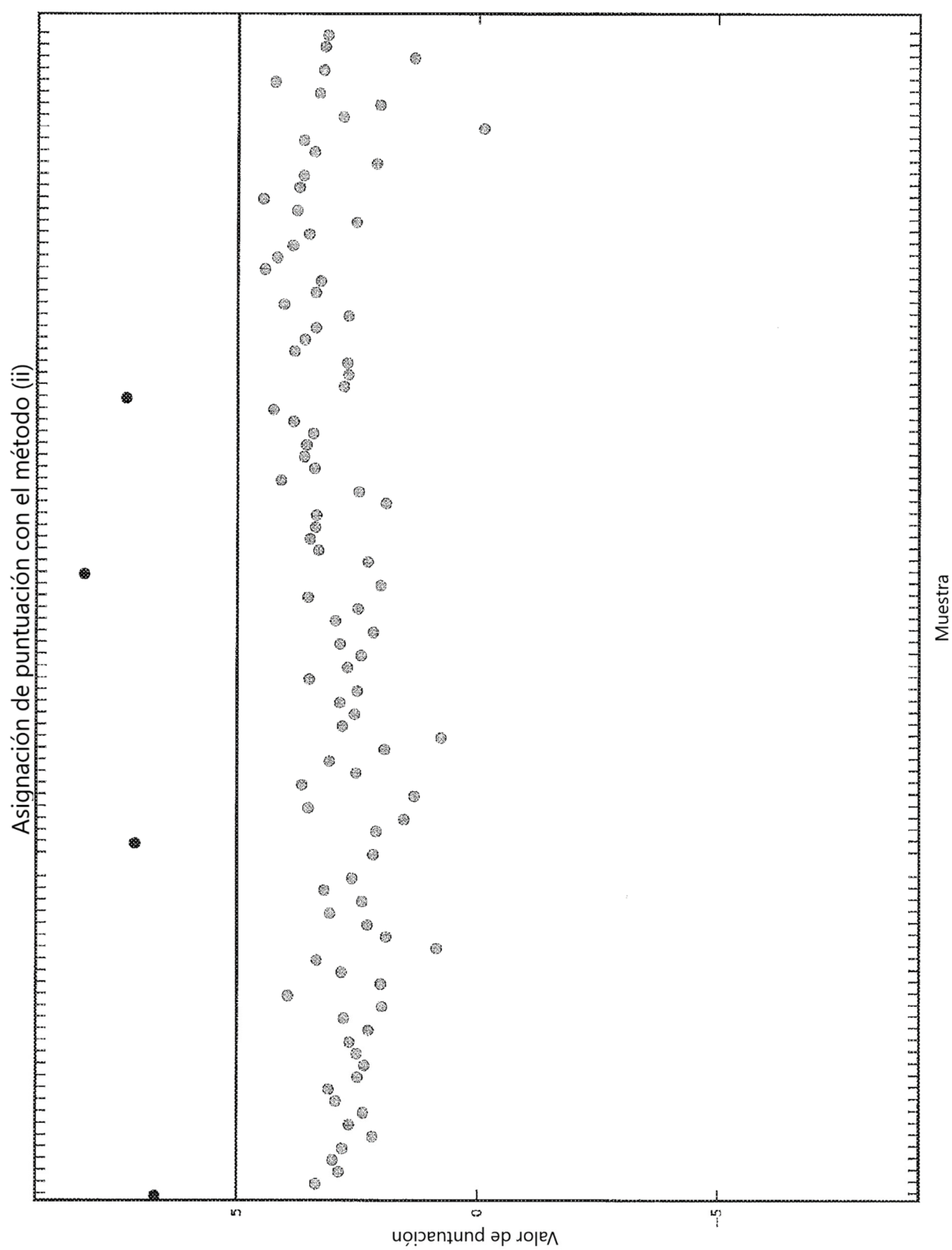


FIGURA 4

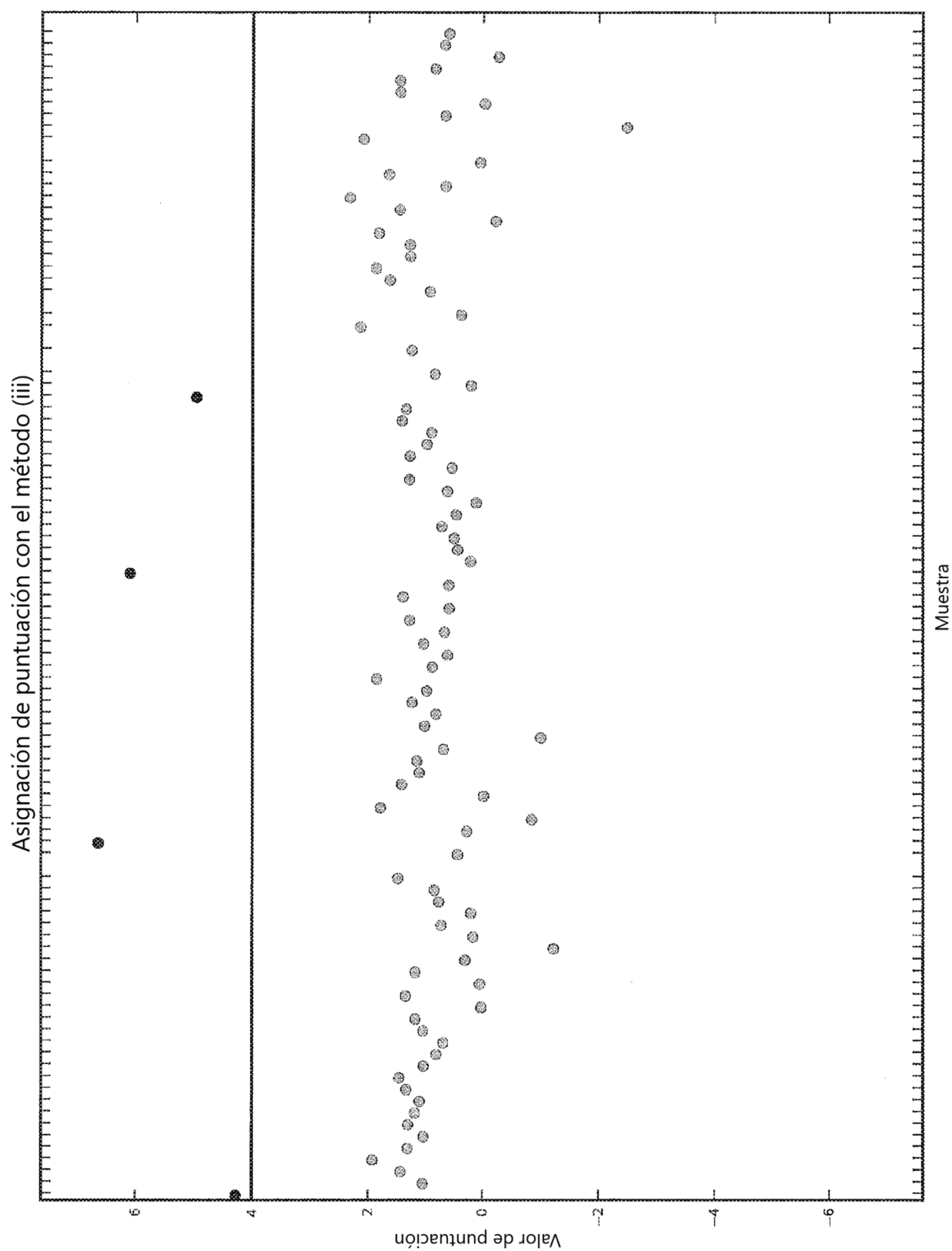


FIGURA 5

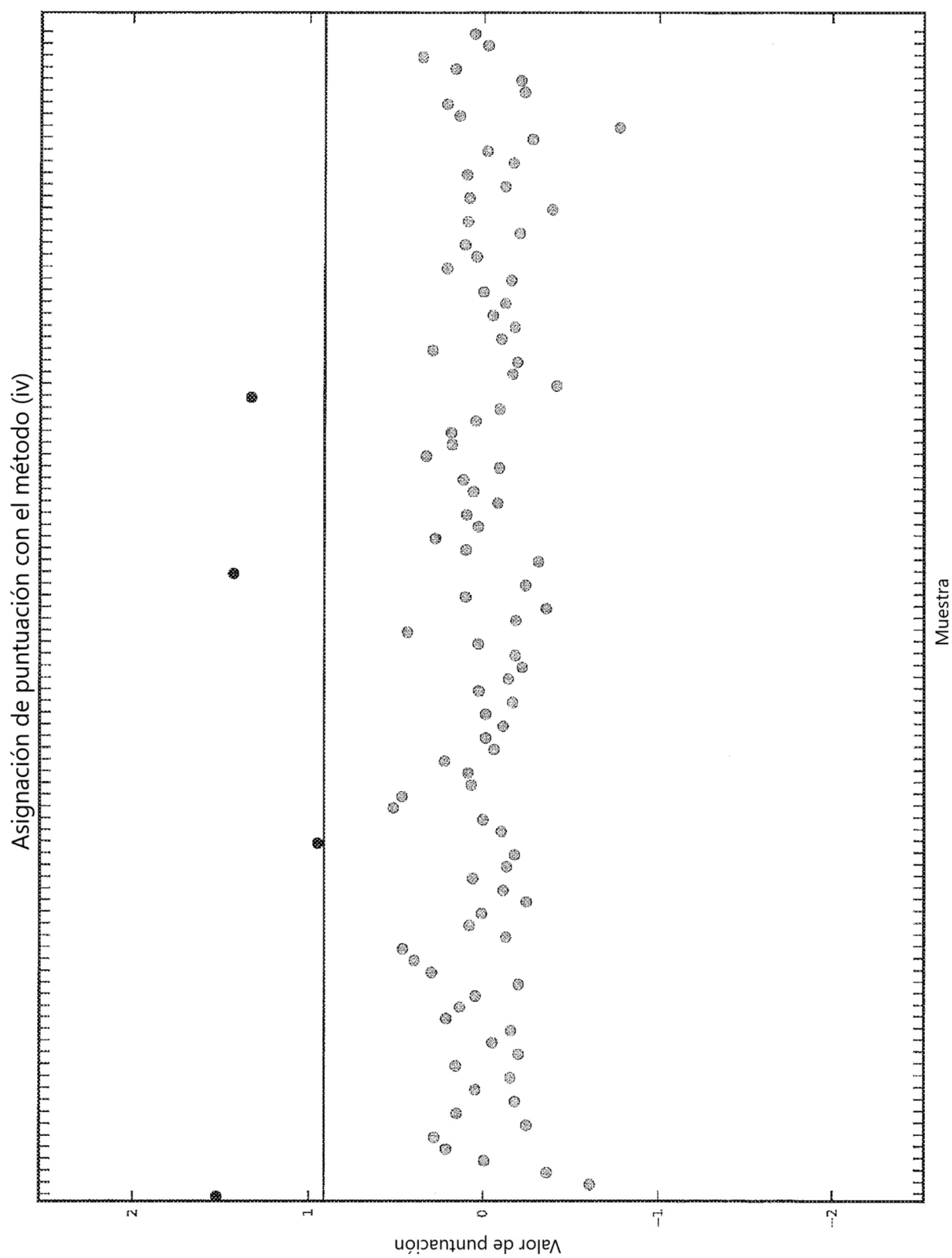


FIGURA 6

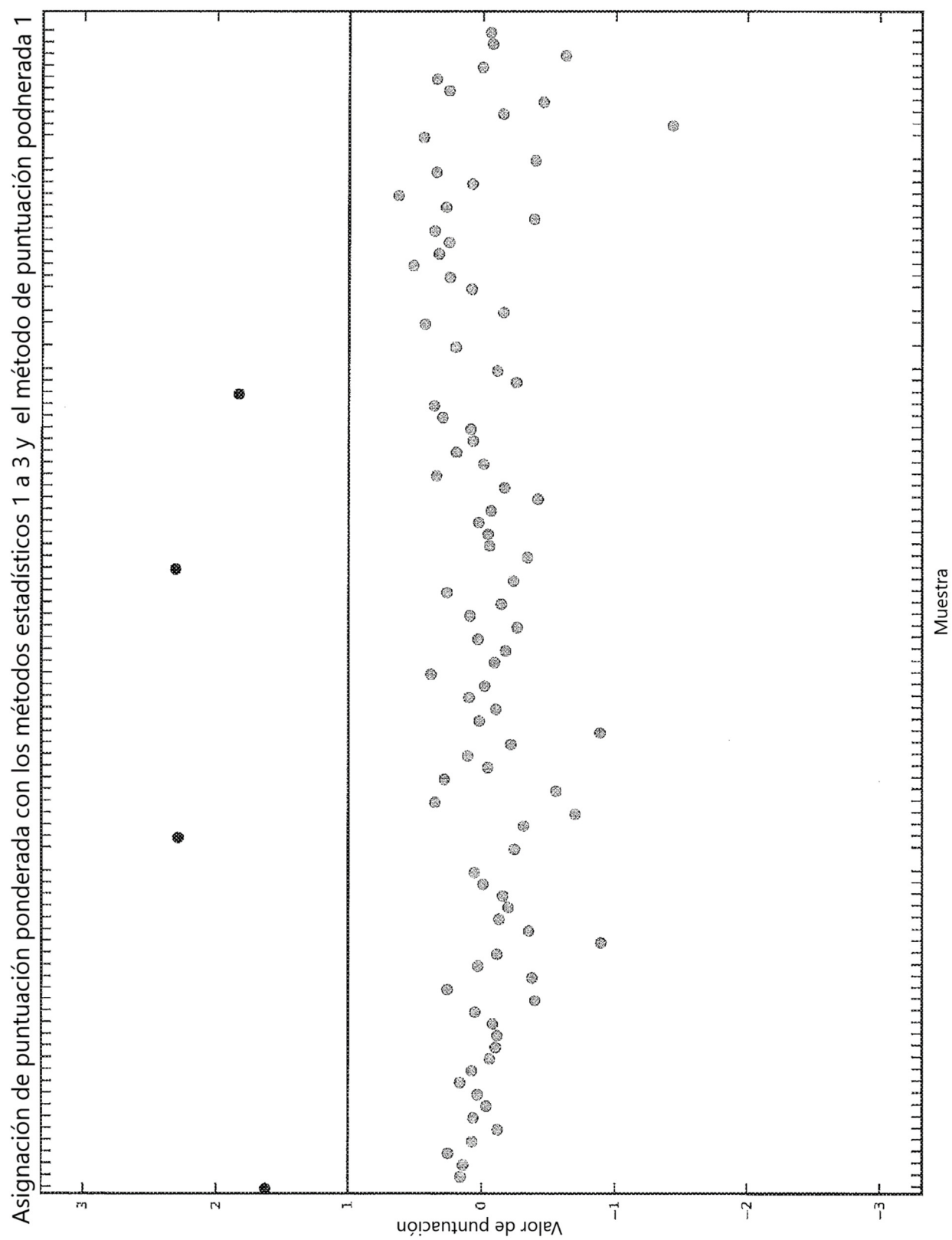


FIGURA 7

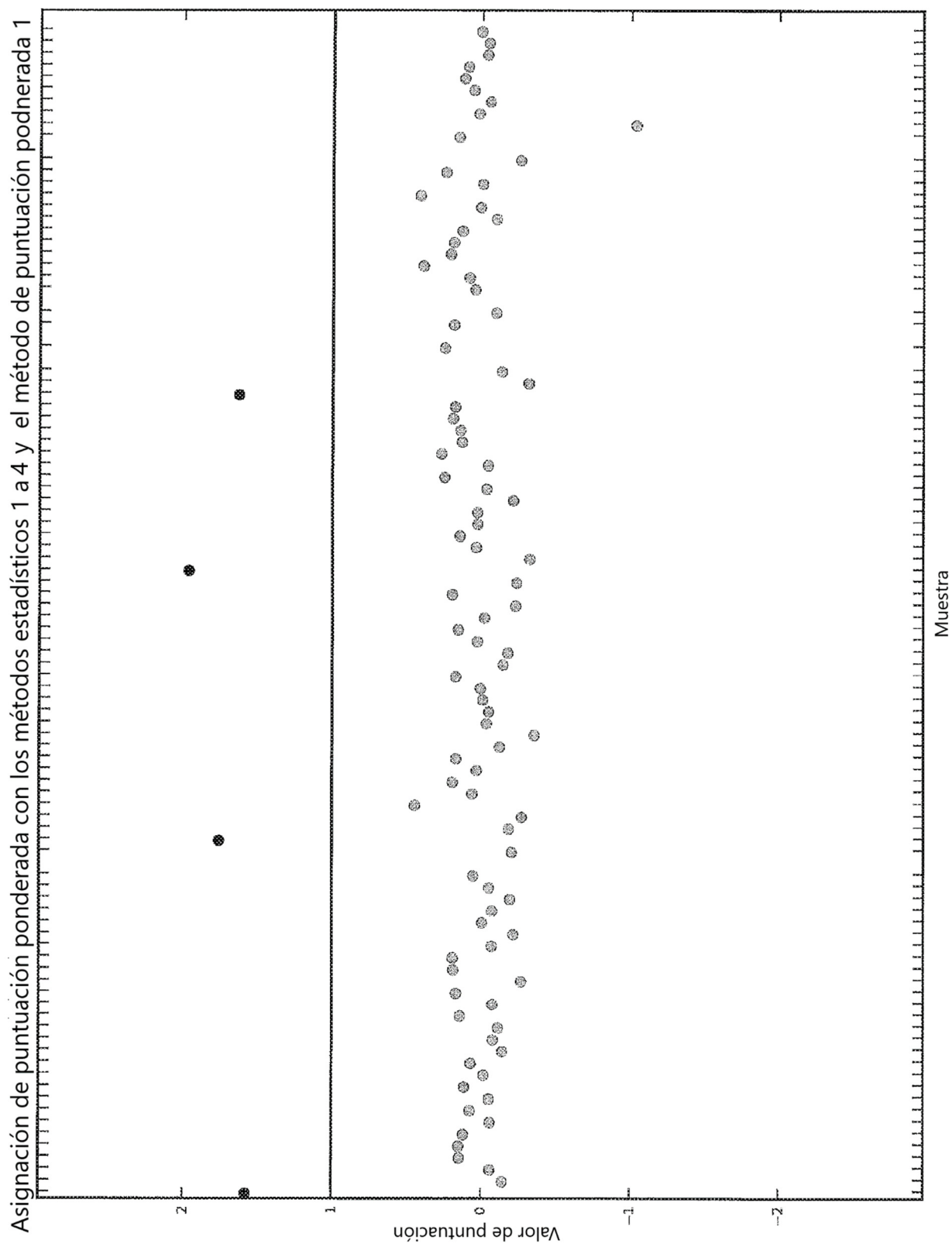


FIGURA 8

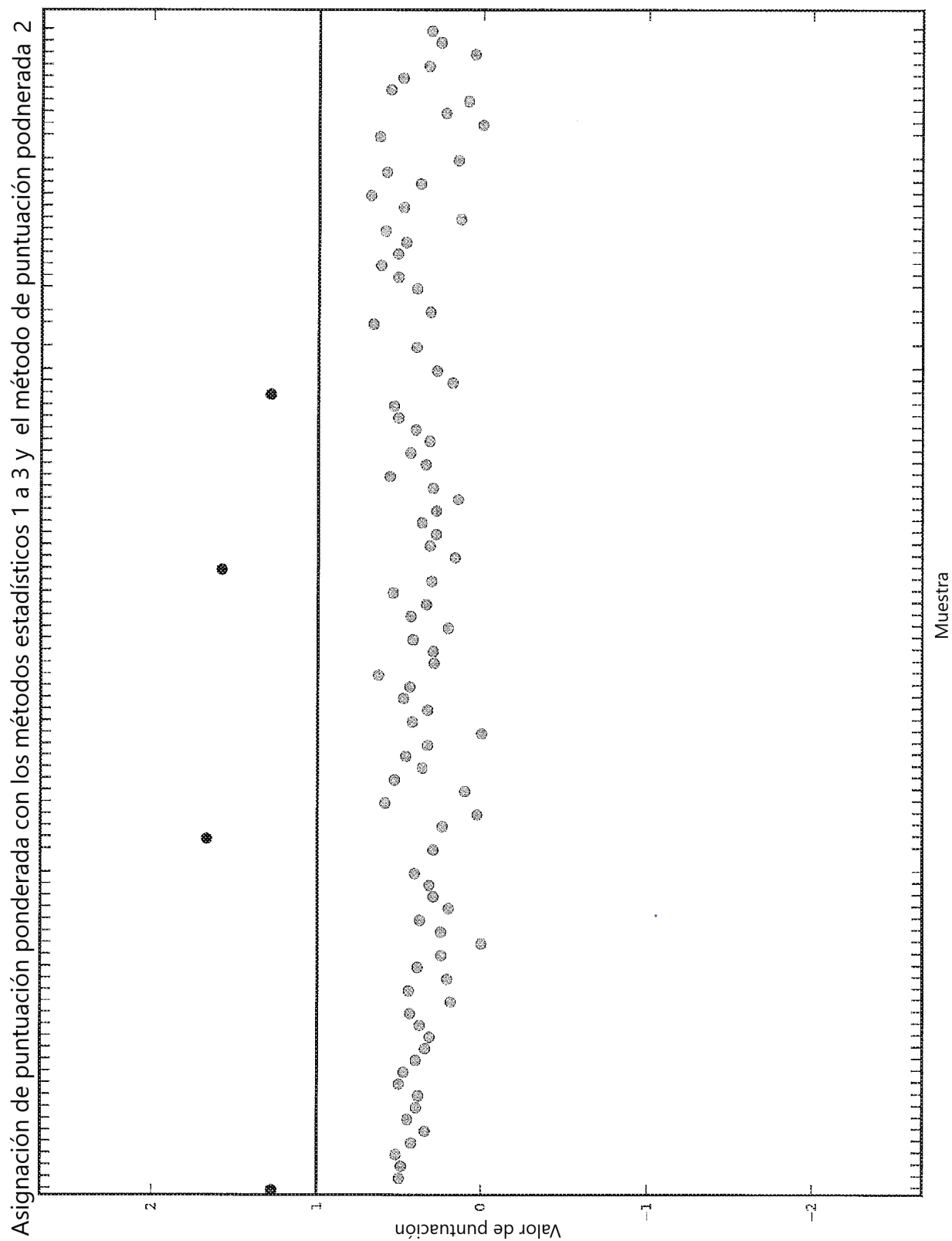


FIGURA 9

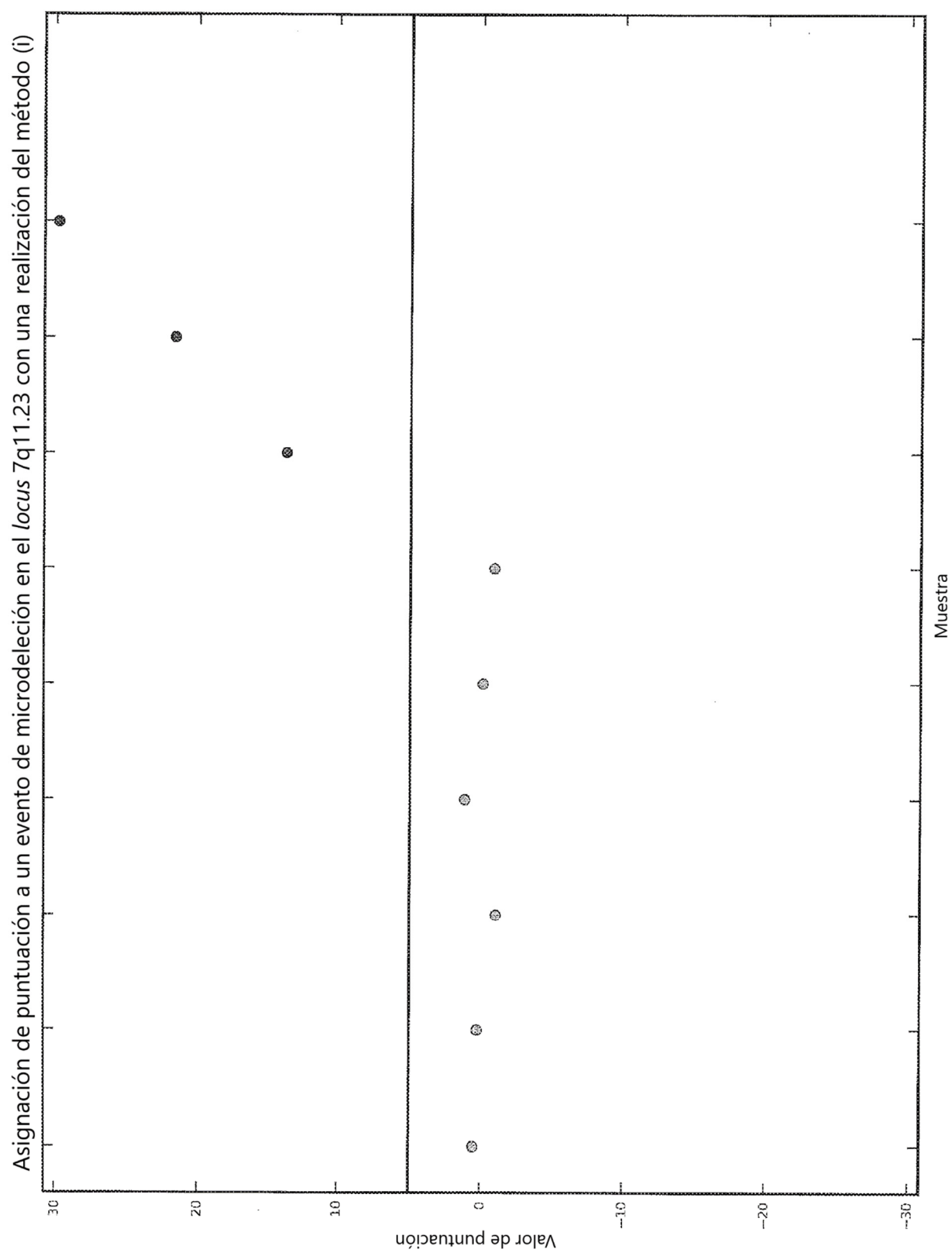


FIGURA 10

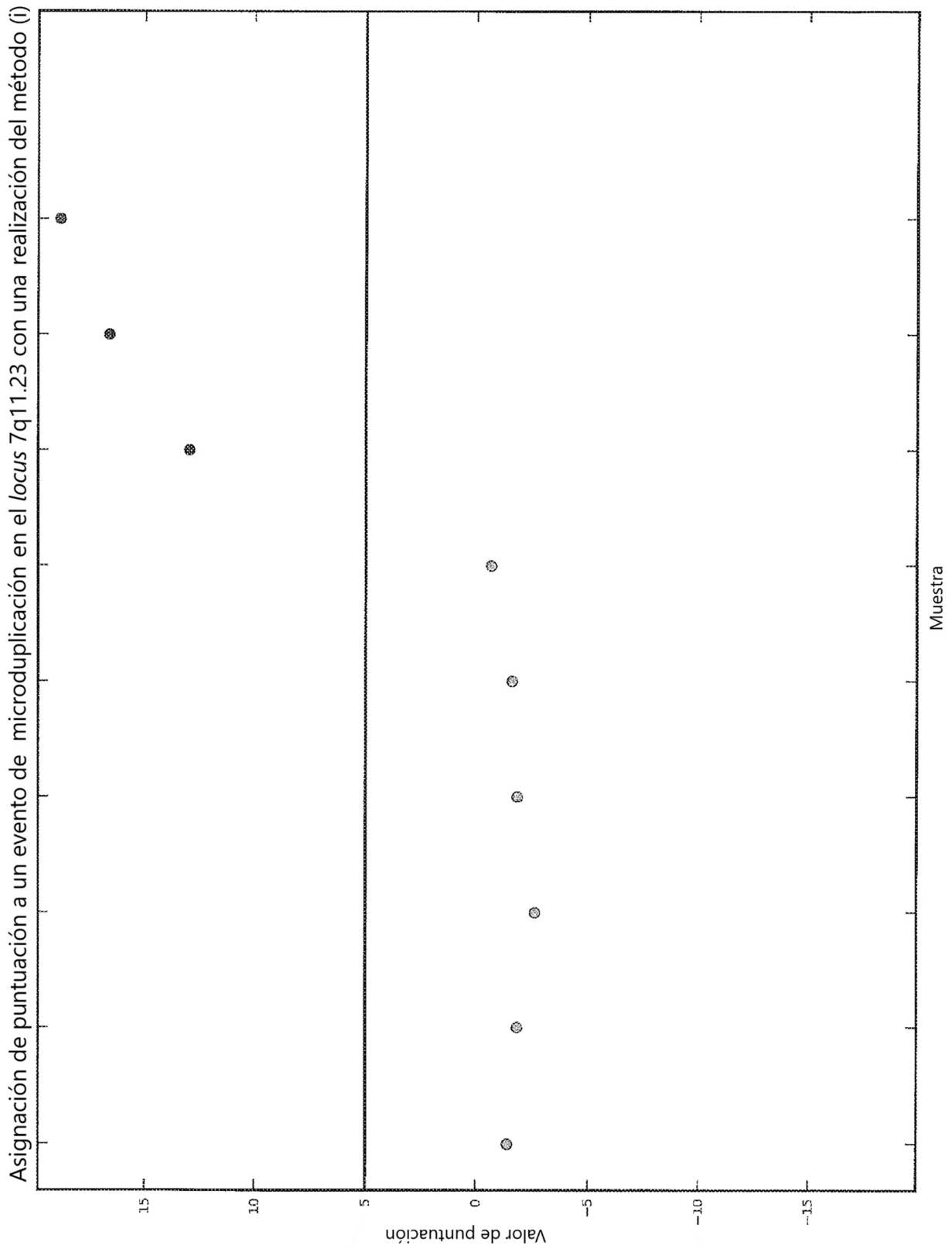


FIGURA 11