

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 820 866**

51 Int. Cl.:

**G16B 20/20** (2009.01)

**C12Q 1/6827** (2008.01)

**C12Q 1/6883** (2008.01)

12

## TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **23.07.2008** **E 12180133 (6)**

97 Fecha y número de publicación de la concesión europea: **19.08.2020** **EP 2557519**

54 Título: **Determinación de un genotipo de un feto**

30 Prioridad:

**23.07.2007 US 951438 P**

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

**22.04.2021**

73 Titular/es:

**THE CHINESE UNIVERSITY OF HONG KONG  
(100.0%)  
Technology Licensing Office, Room 226 Pi Ch'iu  
Building Shatin  
New Territories Hong Kong , CN**

72 Inventor/es:

**LO, YUK-MING DENNIS;  
CHIU, ROSSA WAI KWUN;  
CHAN, KWAN CHEE;  
ZEE, BENNY CHUNG YING y  
CHONG, KA CHUN**

74 Agente/Representante:

**PONS ARIÑO, Ángel**

ES 2 820 866 T3

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

## DESCRIPCIÓN

Determinación de un genotipo de un feto

## 5 Antecedentes

Las enfermedades genéticas, cánceres y otras afecciones a menudo son el resultado o producen un desequilibrio en dos cromosomas o alelos correspondientes u otras secuencias de ácidos nucleicos. Es decir, una cantidad de una secuencia respecto a otra secuencia es mayor o menor de lo normal. Habitualmente, la proporción normal es una proporción 50/50 uniforme. El síndrome de Down (trisomía 21) es una enfermedad que tiene un desequilibrio de un cromosoma 21 de más.

Los métodos de diagnóstico prenatal convencionales de trisomía 21 implican el muestreo de materiales fetales mediante procedimientos invasivos tales como la amniocentesis o el muestreo de vellosidades coriónicas, que suponen un riesgo determinado de pérdida fetal. Los procedimientos no invasivos, tales como el cribado mediante ecografía y marcadores bioquímicos, se han utilizado para estratificar el riesgo de mujeres gestantes antes de los procedimientos de diagnóstico invasivos definitivos. Sin embargo, estos métodos de selección típicamente miden epifenómenos que se asocian con la trisomía 21 en lugar de la anomalía cromosómica del núcleo y por tanto, tienen una exactitud diagnóstica subóptima y otras desventajas, tal como estar influenciada altamente por la edad gestacional.

El descubrimiento de ADN fetal libre de células circulante en el plasma materno en 1997 abrió nuevas posibilidades para el diagnóstico prenatal no invasivo (Lo, YMD y Chiu, RWK 2007 Nat Rev Genet 8, 71-77). Si bien este método se ha aplicado fácilmente al diagnóstico prenatal ligado al sexo (Costa, JM et al. 2002 N Engl J Med 346, 1502) y a ciertos trastornos genéticos únicos (Lo, YMD et al. 1998 N Engl J Med 339, 1734-1738), su aplicación a la detección prenatal de aneuploidías cromosómicas fetales ha representado un desafío considerable (Lo, YMD y Chiu, RWK 2007, anteriormente citado). Primero, los ácidos nucleicos fetales coexisten en el plasma materno con un elevado fondo de ácidos nucleicos de origen materno que pueden interferir con el análisis (Lo, YMD et al. 1998 Am J Hum Genet 62, 768-775). Segundo, los ácidos nucleicos circulan en el plasma materno predominantemente en una forma libre de células, dificultando la derivación de la información de dosis de genes o cromosomas dentro del genoma fetal.

Se han realizado desarrollos considerables para superar estos desafíos recientemente (Benachi, A y Costa, JM 2007 Lancet 369, 440-442). Un enfoque detecta los ácidos nucleicos específicos fetales en el plasma materno, superando por tanto el problema de la interferencia del fondo materno (Lo, YMD y Chiu, RWK 2007, anteriormente citado). La dosis del cromosoma 21 se dedujo a partir de las proporciones de alelos polimórficos en las moléculas de ADN/ARN derivadas de placenta. Sin embargo, este método es menos exacto cuando las muestras contienen menos cantidad de ácido nucleico diana y solo pueden aplicarse a fetos que son heterocigóticos para los polimorfismos diana, lo cual es solo un subconjunto de la población si se usa un polimorfismo.

Dhallan et al. (Dhallan, R, et al. 2007, anteriormente citado Dhallan, R, et al. 2007 Lancet 369, 474-481) describen una estrategia alternativa de enriquecimiento de la proporción de ADN fetal circulante añadiendo formaldehído al plasma materno. La proporción de secuencias del cromosoma 21 aportadas por el feto en el plasma materno se determinó evaluando la proporción de alelos específicos fetales heredados por vía paterna respecto a alelos específicos no fetales para polimorfismos de un solo nucleótido, (SNP; del inglés, Single Nucleotide Polymorphisms) en el cromosoma 21. De forma análoga, se calcularon las proporciones de SNP para un cromosoma de referencia. A continuación, se dedujo un desequilibrio del cromosoma fetal 21 mediante la detección de una diferencia estadísticamente significativa entre las proporciones de SNP para el cromosoma 21 y aquellas del cromosoma de referencia, donde el nivel de significación se definió utilizando un valor fijo de  $p \leq 0,05$ . Para garantizar una alta cobertura de la población, se seleccionaron más de 500 SNP por cromosoma. Sin embargo, ha habido controversias con respecto a la eficacia del formaldehído para enriquecer a una elevada proporción (Chung, GTY, et al. 2005 Clin Chem 51, 655-658) y por tanto la reproducibilidad del método necesita evaluarse de forma adicional. Además, como cada feto y madre serían informativos para un número diferente de SNP para cada cromosoma, la potencia del ensayo estadístico para la comparación de proporciones de SNP variaría en cada caso (Lo, YMD y Chiu, RWK. 2007 Lancet 369, 1997). Asimismo, dado que estos enfoques dependen de la detección de polimorfismos genéticos, se limitan a fetos heterocigóticos para estos polimorfismos.

La utilización de la reacción en cadena de la polimerasa (PCR) y la cuantificación de ADN de un locus del cromosoma 21 y de un locus de referencia en cultivos de amniocitos, obtenidos de fetos con trisomía 21 y euploides, Zimmermann et al. (2002 Clin Chem. 48, 362-363) pudo diferenciar los dos grupos de fetos basándose en el aumento de 1,5 veces en las secuencias de ADN del cromosoma 21 en el primer grupo. Dado que una diferencia de 2 veces en la concentración del molde de ADN constituye una diferencia de un solo ciclo umbral (Ct), la discriminación de una diferencia de 1,5 veces ha sido el límite de la PCR en tiempo real convencional. Para conseguir grados de discriminación cuantitativa más sutiles, se requieren estrategias alternativas. Por consiguiente, algunas realizaciones de la presente invención utilizan PCR digital (Vogelstein, B et al. 1999 Proc Natl Acad Sci USA 96, 9236-9241) para este propósito.

La PCR digital se ha desarrollado para la detección del sesgo de la proporción alélica en muestras de ácidos nucleicos (Chang, HW et al. 2002 J Natl Cancer Inst. 94, 1697-1703). Clínicamente, se ha mostrado su utilidad para la detección de la pérdida de heterocigosidad (LOH; del inglés, loss of heterozygosity) en muestras de ADN tumorales (Zhou, W. et al. 2002 Lancet 359, 219-225). Para el análisis de resultados de PCR digital, se ha adoptado en estudios previos una prueba de la proporción de probabilidades secuencial (SPRT; del inglés, Sequential Probability Ratio Testing) para clasificar los resultados experimentales sugerentes de la presencia de LOH en una muestra o no (El Karoui et al. 2006 Stat Med 25, 3124-3133). En algunos métodos utilizados en los estudios previos, el valor de límite para determinar el LOH utilizó una proporción de referencia fija de los dos alelos en el ADN de 2/3. Como la cantidad, proporción y concentración de ácidos nucleicos fetales en plasma materno son variables, estos métodos no son adecuados para detectar la trisomía 21 utilizando ácidos nucleicos fetales en un fondo de ácidos nucleicos maternos en plasma materno.

Es deseable tener una prueba no invasiva para la detección de trisomía 21 fetal (y otros desequilibrios) basada en el análisis de ácidos nucleicos fetales circulantes, especialmente uno que sea independiente del uso de polimorfismos genéticos y/o marcadores específicos fetales. Es deseable también tener una determinación exacta de valores límite y de recuento de secuencias, que puedan reducir el número de pocillos de datos y/o la cantidad de moléculas de ácidos nucleicos de plasma materno necesarias para la exactitud, proporcionando por tanto una eficacia aumentada y una eficacia del coste. Es deseable también que las pruebas no invasivas tengan elevada sensibilidad y especificidad para minimizar los diagnósticos falsos.

Otra aplicación de la detección de ADN fetal en plasma materno es para el diagnóstico prenatal de trastornos genéticos únicos, tales como la beta-talasemia. Sin embargo, como el ADN fetal solo constituye una fracción minoritaria del ADN en el plasma materno, se cree que este enfoque solo puede detectar una mutación que un feto ha heredado de su padre, pero que está ausente de la madre. Entre los ejemplos de esto se incluye la delección de 4 pb en los codones 41/42 del gen de la beta-globina humana que provoca la beta-talasemia (Chiu RWK et al. 2002 Lancet, 360, 998-1000) y la mutación de Q890X del gen regulador de la conductancia transmembrana de la fibrosis quística (Gonzalez-Gonzalez et al. 2002 Prenat Diagn, 22, 946-8). Sin embargo, dado que tanto la beta-talasemia como la fibrosis quística son afecciones recesivas autosómicas, en las que el feto necesitaría heredar una mutación de cada parental antes de que la enfermedad se manifestara, la detección de la mutación heredada meramente por vía paterna solo incrementaría el riesgo de que el feto tenga la enfermedad de un 25 % a un 50 %. Diagnósticamente, esto no es lo ideal. Por tanto, la principal aplicación diagnóstica del enfoque existente sería para el escenario cuando se puede detectar la mutación fetal no heredada por vía paterna en el plasma materno, donde el feto a continuación, puede estar excluido de tener el estado de enfermedad homocigótico. Sin embargo, diagnósticamente, este enfoque tiene la desventaja de que la conclusión se realiza basada en la detección negativa de la mutación paterna. Por tanto, un enfoque que permitiera que se determine el genotipo fetal completo (sea homocigótico normal, homocigótico mutante o heterocigótico) a partir de plasma materno, sin la limitación anterior, sería muy deseable.

### Breve resumen

Los ejemplos desvelados en el presente documento proporcionan métodos, sistemas y aparatos para determinar si un desequilibrio de secuencias de ácido nucleico (por ejemplo, desequilibrio alélico, desequilibrio mutacional o desequilibrio cromosómico) existe dentro de una muestra biológica. Se eligen uno o más valores límite para determinar un desequilibrio de, por ejemplo, una proporción de cantidades de las dos secuencias (o conjuntos de secuencias).

En un ejemplo, el valor límite se determina basándose, al menos en parte, en el porcentaje de secuencias fetales (ácido nucleico clínicamente relevante) en una muestra biológica, tal como plasma materno o suero u orina, que contiene un fondo de secuencias de ácidos nucleicos maternos. En otra realización, el valor límite se determina basándose en una concentración promedio de una secuencia en una pluralidad de reacciones. En un aspecto, el valor límite se determina a partir de una proporción de pocillos informativos que se estima que contienen una secuencia de ácido nucleico particular, donde la proporción se determina basándose en el porcentaje mencionado anteriormente y/o la concentración promedio.

El valor límite se puede determinar utilizando muchos tipos diferentes de métodos, tales como SPRT, descubrimiento falso, intervalo de confianza, característica operativa del receptor (ROC; del inglés, receiver operating characteristic). Esta estrategia minimiza adicionalmente la cantidad de pruebas necesarias antes de que se pudiera realizar una clasificación segura. Esto es de una relevancia particular para los análisis de ácidos nucleicos plasmáticos en donde la cantidad de molde a menudo es limitante.

De acuerdo con la invención, se proporciona un método implementado en un ordenador para determinar un genotipo de un feto de una mujer gestante en un primer locus, el feto que tiene una madre que es la mujer gestante, en donde la madre es heterocigótica en el primer locus para un primer alelo y un segundo alelo que es diferente del primer alelo, comprendiendo el método: recibir datos a partir de una pluralidad de reacciones que implican moléculas de ADN libre de células de una muestra biológica de la madre, en donde la muestra biológica es plasma o suero materno y que incluye moléculas de ADN libre de células del sujeto femenino y del feto, en donde cada reacción indica una presencia o ausencia del primer alelo y del segundo alelo y en donde los datos incluyen: (1) un primer conjunto de datos cuantitativos que indican una primera cantidad de reacciones que son positivas para la presencia del primer alelo del

sujeto femenino y del feto en la pluralidad de reacciones; y (2) un segundo conjunto de datos cuantitativos que indican una segunda cantidad de reacciones que son positivas para la presencia del segundo alelo del sujeto femenino y del feto en la pluralidad de reacciones; determinar un parámetro a partir de los dos conjuntos de datos, en donde el parámetro incluye una proporción entre la primera cantidad y la segunda cantidad; comparar el parámetro con uno o más valores límite para determinar el genotipo del feto, en donde el feto se determina que tiene solo el primer alelo en el primer locus cuando la primera cantidad es mayor que la segunda cantidad mediante un grado especificado por uno de dichos uno o más valores límite, en donde el feto se determina que tiene solo el segundo alelo en el primer locus cuando la segunda cantidad es mayor que la primera cantidad por un grado especificado por uno de dichos uno o más valores límite y en donde el feto se determina que es heterocigótico para el primer alelo y el segundo alelo cuando la primera cantidad y la segunda cantidad son estadísticamente iguales, como se especifica por dichos uno o más valores límite.

Otro aspecto de la invención se dirige a un medio legible por ordenador asociado con el método de la invención.

- 15 Se puede obtener una mejor comprensión de la naturaleza y ventajas de la presente invención con referencia a la siguiente descripción detallada y los dibujos adjuntos.

### Breve descripción de los dibujos

20 La Fig. 1 es un diagrama de flujo que ilustra un experimento de PCR digital.

La Fig. 2A ilustra un método de análisis digital de ARN-SNP y RCD de acuerdo con una realización de la presente invención.

25 La Fig. 2B muestra una tabla de ejemplos de anomalías cromosómicas detectables frecuentemente en cánceres.

La Fig. 3 ilustra un gráfico que tiene las curvas SPRT utilizadas para determinar el síndrome de Down.

30 La Fig. 4 muestra un método para determinar un estado de enfermedad utilizando un porcentaje de células fetales.

La Fig. 5 muestra un método para determinar un estado de enfermedad utilizando una concentración promedio.

35 La Fig. 6 muestra una tabla que tabula la proporción alélica de ARN-SNP digital esperada y la  $P_r$  de muestras de trisomía 21 para un intervalo de concentraciones de moldes expresado como la concentración de moldes de referencia promedio por pocillo ( $m_r$ ).

40 La Fig. 7 muestra una tabla que tabula la  $P_r$  esperada para las concentraciones fraccionarias de ADN fetal de un 10 %, 25 %, 50 % y 100 % en muestras de trisomía 21 en un intervalo de concentraciones de moldes expresado como la concentración de moldes de referencia promedio por pocillo ( $m_r$ ).

La Fig. 8 muestra un gráfico que ilustra el grado de diferencias en las curvas de SPRT para los valores de  $m_r$  de 0,1, 0,5 y 1,0 para los análisis de ARN-SNP digital.

45 La Fig. 9A muestra una tabla de una comparación de la eficacia de los algoritmos de SPRT nuevo y viejo para clasificar casos de euploidía y trisomía 21 en los análisis de ARN-SNP digital de 96 pocillos.

La Fig. 9B muestra una tabla de una comparación de la eficacia de los algoritmos de SPRT nuevo y viejo para clasificar casos de euploidía y trisomía 21 en los análisis de ARN-SNP digital de 384 pocillos.

50 La Fig. 10 es una tabla que muestra los porcentajes de fetos clasificados correcta e incorrectamente como euploides y aneuploides y aquellos no clasificables para los recuentos informativos dados.

55 La Fig. 11 es una tabla 1100 que muestra simulaciones informáticas para un análisis de RCD digital para una muestra de ADN fetal puro (100 %).

La Fig. 12 es una tabla 1200 que muestra los resultados de la simulación informática de exactitudes de análisis de RCD digital para  $m_r=0,5$  para la clasificación de muestras de fetos euploides o con trisomía 21 con concentraciones fraccionarias diferentes de ADN fetal.

60 La Fig. 13A muestra una tabla 1300 de análisis de ARN-SNP digital en tejidos placentarios de gestaciones de euploides y trisómicos 21.

La Fig. 13B muestra una tabla 1350 de análisis de ARN-SNP digital de plasma materno de gestaciones de euploides y trisómicos 21.

65 Las Fig. 14A-14C muestran gráficos que ilustran una curva límite resultado de un análisis RCD.



La Fig. 15A muestra una tabla de análisis de ARN-SNP digital en tejidos placentarios de gestaciones de euploides o de trisómicos 21.

5 La Fig. 15B muestra una tabla de datos de ARN-SNP digital de los 12 paneles de reacción a partir de una muestra de tejido materno.

La Fig. 15C muestra una tabla de análisis de ARN-SNP digital de plasma materno de gestaciones de euploides o de trisómicos 21.

10 La Fig. 16A muestra una tabla para un análisis de ARN-SNP digital de placentas de euploides y trisómicos 18.

La Fig. 16B muestra una interpretación de SPRT de datos de ARN-SNP digital para placentas de euploides o de trisómicos 18.

15 La Fig. 17 muestra una tabla de un análisis de RCD digital de mezclas de ADN celular de placenta/sangre materna al 50 % de gestaciones de euploides y trisómicos 21.

La Fig. 18 muestra una curva de SPRT que ilustra los límites de decisión para la clasificación correcta.

20 La Fig. 19 muestra una tabla de análisis de RCD digital de muestras de líquido amniótico de gestaciones euploides y de trisómicos 21.

25 La Fig. 20 muestra una tabla de análisis de RCD digital de muestras de ADN placentario de gestaciones de euploides y trisómicos 18 (E= euploide; T18= trisomía 18).

La Fig. 21 muestra una tabla de un análisis de RCD digital multiplexado de mezclas de ADN celular de placenta/sangre materna al 50 % de gestaciones de euploides y trisómicos 21.(E= euploide; T21= trisomía 21; U= sin clasificar).

30 Las Figs. 22A y 22B muestran una tabla de análisis de RCD digital multiplexado de un 50 % de ADN genómico placentario de euploide o trisómico 21/50 % mezcla de ADN de capa leucocitaria materna. Sin clase indica inclasificable y T21 indica trisomía 21.

35 La Fig. 23 muestra un escenario en donde tanto las parejas masculinas como femeninas portan la misma mutación.

La Fig. 24A muestra una tabla de un análisis de RMD digital de mezclas de ADN mujer/hombre y hombre/hombre de acuerdo con una realización de la presente invención.

40 La Fig. 24B muestra una tabla de un análisis de RMD digital de mezclas con ADN un 25 % femenino y un 75 % masculino, de acuerdo con una realización de la presente invención.

La Fig. 25 muestra una tabla de un análisis de RMD digital de mezclas de un 15 %-50 % de ADN que imitan muestras de plasma materno para la mutación HbE de acuerdo con una realización de la presente invención.

45 La Fig. 26A muestra una tabla de un análisis de RMD digital de mezclas de un 5 %-50 % de ADN que imitan muestras de plasma materno para la mutación CD41/42 de acuerdo con una realización de la presente invención.

50 La Fig. 26B muestra una tabla de un análisis de RMD digital de mezclas de un 20 % de ADN que imitan muestras de plasma materno para la mutación CD41/42 de acuerdo con una realización de la presente invención.

La Fig. 27 muestra un diagrama de bloques de un aparato informático ejemplar utilizable con sistemas y métodos de acuerdo con las realizaciones de la presente invención.

## 55 Definiciones

La expresión "*muestra biológica*", como se utiliza en el presente documento, se refiere una cualquier muestra que se toma de un sujeto (por ejemplo, un ser humano, tal como una mujer gestante) y contiene una o más moléculas de ácido nucleico de interés.

60 La expresión "*ácido nucleico*" o "*polinucleótido*" se refiere a un ácido desoxirribonucleico (ADN) o ácido ribonucleico (ARN) y a un polímero de los mismos en forma mono- o bicatenaria. A menos que esté específicamente limitado, la expresión abarca los ácidos nucleicos que contienen análogos conocidos de nucleótidos naturales que tienen propiedades de unión similares a las del ácido nucleico de referencia y que se metabolizan de manera similar a la de los nucleótidos de origen natural. Salvo que se indique lo contrario, una secuencia particular del ácido nucleico también abarca implícitamente variantes modificadas conservativamente de la misma (por ejemplo, sustituciones

de codones degenerados), alelos, ortólogos, SNP y secuencias complementarias así como la secuencia explícitamente indicada. Específicamente, las sustituciones de codones degenerados se pueden lograr mediante la generación de secuencias en las que la tercera posición de uno o más codones seleccionados (o todos) se sustituye con restos de base mixta y/o de desoxiinosina (Batzer *et al.*, Nucleic Acid Res. 19:5081 (1991); Ohtsuka *et al.*, J. Biol. Chem. 260: 2605-2608 (1985); y Rossolini *et al.*, Mol. Cell. Probes 8: 91-98 (1994)). La expresión ácido nucleico se utiliza indistintamente con gen, ADNc, ARNm, ARN pequeño no codificante, microARN (miARN), ARN asociado a Piwi y ARN corto en horquilla (ARNhc) codificado por un gen o locus.

El término "*gen*" significa el segmento de ADN implicado en la producción de una cadena polipeptídica. Puede incluir regiones anteriores y posteriores a la región codificante (líder y remolque) así como las secuencias intervinientes (intrones) entre los segmentos de codificación individuales (exones).

El término "*reacción*", como se utiliza en el presente documento, se refiere a cualquier proceso que implica una acción química, enzimática o física que sea indicativa de la presencia o ausencia de una secuencia particular de polinucleótidos de interés. Un ejemplo de una "*reacción*" es una reacción de amplificación, tal como una reacción en cadena de la polimerasa (PCR). Otro ejemplo de una "*reacción*" es una reacción de secuenciación, por síntesis o por ligadura. Una "*reacción informativa*" es una que indica la presencia de una o más secuencias particulares de polinucleótidos de interés y en un caso en donde solo una secuencia de interés está presente. El término "pocillo", como se utiliza en el presente documento, se refiere a una reacción en un lugar predeterminado dentro de una estructura confinada, por ejemplo, un vial, célula o cámara bien conformados en una matriz de PCR.

La expresión "*secuencia de ácido nucleico clínicamente relevante*", como se utiliza en el presente documento, puede referirse a una secuencia de polinucleótidos correspondiente a un segmento de una secuencia genómica más grande cuyo desequilibrio potencial se va a someter a prueba o a la propia secuencia genómica más grande. Un ejemplo es la secuencia del cromosoma 21. Otros ejemplos incluyen el cromosoma 18, 13, X e Y. Todavía otros ejemplos incluyen secuencias genéticas mutadas o polimorfismos genéticos o variaciones del número de copias que un feto puede heredar de uno de sus padres o de ambos. Todavía otros ejemplos incluyen secuencias que están mutadas, delecionadas o amplificadas en un tumor maligno, por ejemplo, secuencias en las que se produce pérdida de heterocigosidad o duplicación génica. En algunas realizaciones, pueden usarse secuencias múltiples de ácido nucleico clínicamente relevantes o marcadores equivalentemente múltiples de la secuencia de ácido nucleico clínicamente relevante para proporcionar datos para detectar el desequilibrio. Por ejemplo, pueden usarse datos de cinco secuencias no consecutivas en el cromosoma 21 de una manera aditiva para la determinación de un posible desequilibrio del cromosoma 21, reduciendo eficazmente la necesidad de volumen de muestra a 1/5.

La expresión "*secuencia de ácido nucleico de fondo*", como se utiliza en el presente documento, se refiere a una secuencia de ácido nucleico cuya proporción normal se conoce respecto a la secuencia de ácido nucleico clínicamente relevante, por ejemplo, una proporción de 1 a 1. Como un ejemplo, la secuencia de ácido nucleico de fondo y la secuencia de ácido nucleico clínicamente relevante son dos alelos del mismo cromosoma que se diferencian debido a la heterocigosidad. En otro ejemplo, la secuencia de ácido nucleico de fondo es un alelo que es heterocigótico respecto a otro alelo que es la secuencia de ácido nucleico clínicamente relevante. Asimismo, algo de cada secuencia de ácido nucleico de fondo y de la secuencia de ácido nucleico clínicamente relevante puede proceder de diferentes individuos.

La expresión "*secuencia de ácido nucleico de referencia*", como se utiliza en el presente documento, se refiere a una secuencia de ácido nucleico cuya concentración promedio por reacción se conoce o se ha medido de forma equivalente.

La expresión "*secuencia de ácido nucleico sobrerrepresentada*" como se utiliza en el presente documento se refiere a la secuencia de ácido nucleico entre dos secuencias de interés (por ejemplo, una secuencia clínicamente relevante y una secuencia de fondo) que es más abundante que la otra secuencia en una muestra biológica.

La expresión "*basándose en*" como se utiliza en el presente documento significa "*basándose en al menos en parte*" y se refiere a un valor (o resultado) que se utiliza en la determinación de otro valor, tal como sucede en la relación de una entrada de un método y una salida de ese método. El término "*deriva*" como se utiliza en el presente documento se refiere también a la relación de una entrada de un método y a la salida de ese método, tal como sucede cuando la derivación es el cálculo de una fórmula.

La expresión "*datos cuantitativos*" como se utiliza en el presente documento significa que los datos que se obtienen de una o más reacciones y que proporcionan uno o más valores numéricos. Por ejemplo, el número de pocillos que muestran un marcador fluorescente para una secuencia particular sería un dato cuantitativo.

El término "*parámetro*" como se utiliza en el presente documento significa un valor numérico que caracteriza un conjunto de datos cuantitativo y/o una relación numérica entre conjuntos de datos cuantitativos. Por ejemplo, un parámetro es una proporción (o función de una proporción) entre una primera cantidad de una primera secuencia de ácido nucleico y una segunda cantidad de una segunda secuencia de ácido nucleico.

La expresión "*valor límite*" (en inglés, cutoff value), como se utiliza en el presente documento, significa un valor numérico cuyo valor se utiliza para arbitrar entre dos o más estados de clasificación (por ejemplo, enfermo y no enfermo) de una muestra biológica. Por ejemplo, si un parámetro es mayor que el valor límite, se realiza una primera clasificación del dato cuantitativo (por ejemplo, estado enfermo); o si el parámetro es menor que el valor límite, se realiza una clasificación diferente de los datos cuantitativos (por ejemplo, estado no enfermo).

La expresión "*desequilibrio*" como se utiliza en el presente documento significa cualquier desviación significativa como se define por al menos un valor límite en una cantidad de la secuencia de ácido nucleico clínicamente relevante a partir de una cantidad de referencia. Por ejemplo, la cantidad de referencia podría ser una proporción de 3/5 y por tanto podría producirse un desequilibrio si la proporción medida es de 1:1.

### Descripción detallada de la invención

Se desvelan los métodos, sistemas y aparatos para determinar si un aumento o una disminución en comparación con una cantidad de referencia (por ejemplo, no enfermo) de una secuencia de ácido nucleico clínicamente relevante en relación con otras secuencias de ácido nucleico clínicamente no relevantes (por ejemplo, un desequilibrio cromosómico o alélico) existe dentro de una muestra biológica. Se seleccionan uno o más valores límite para determinar si existe un cambio en comparación con la cantidad de referencia (por ejemplo un desequilibrio), por ejemplo, respecto a la proporción de cantidades de dos secuencias (o conjuntos de regiones). El cambio detectado en la cantidad de referencia puede ser cualquier desviación (ascendente o descendente) en la relación de la secuencia de ácido nucleico clínicamente relevante con las otras secuencias clínicamente no relevantes. Por tanto, el estado de referencia puede ser cualquier proporción u otra cantidad (por ejemplo, distinta de una correspondencia 1-1) y un estado medido que significa que un cambio puede ser cualquier proporción u otra cantidad que difiera de la cantidad de referencia como se determina por dichos uno o más valores límite.

La secuencia de ácido nucleico clínicamente relevante y la secuencia de ácido nucleico de fondo pueden proceder de un primer tipo de células y de uno o más segundos tipos de células. Por ejemplo, en una muestra biológica están presentes secuencias de ácido nucleico fetal originadas de células fetales/placentarias, tal como plasma materno, que contiene un fondo de secuencias de ácido nucleico materno que se originan de células maternas. Por tanto, en una realización, el valor límite se determina basándose, al menos en parte, en el porcentaje del primer tipo de células en una muestra biológica. Obsérvese que el porcentaje de secuencias fetales en una muestra puede determinarse mediante cualquier loci derivado del feto y no limitarse a medir las secuencias de ácido nucleico clínicamente relevantes. En otra realización, el valor límite se determina basándose, al menos en parte, en el porcentaje de secuencias tumorales en una muestra biológica, tal como plasma, suero, saliva u orina, que contiene un fondo de secuencias de ácido nucleico derivadas de las células no malignas dentro del cuerpo.

En todavía otra realización, el valor límite se determina basándose en una concentración promedio de una secuencia en una pluralidad de reacciones. En un aspecto, el valor límite se determina a partir de una proporción de pocillos informativos que se estima que contienen una secuencia de ácido nucleico particular, donde la proporción se determina basándose en el porcentaje mencionado anteriormente y/o la concentración promedio. El valor límite se puede determinar utilizando muchos tipos diferentes de métodos, tales como SPRT, descubrimiento falso, intervalo de confianza, característica operativa del receptor (ROC; del inglés, receiver operating characteristic). Esta estrategia minimiza adicionalmente la cantidad de pruebas necesarias antes de que se pueda realizar una clasificación segura. Esto es de una relevancia particular para los análisis de ácidos nucleicos plasmáticos en donde la cantidad de molde a menudo es limitante. Aunque se presenta con respecto a la PCR digital, se pueden utilizar otros métodos.

La PCR digital implica múltiples análisis de PCR sobre ácidos nucleicos extremadamente diluidos, de manera que la mayoría de amplificaciones positivas reflejan la señal a partir de una molécula molde individual. La PCR digital de este modo permite el recuento de moléculas molde individuales. La proporción de amplificaciones positivas entre el número total de PCR analizadas, permite una estimación de la concentración de molde en la muestra original sin diluir. Esta técnica se ha propuesto para permitir la detección de diversos fenómenos genéticos (Vogelstein, B *et al.* 1999, anteriormente citado) y se ha utilizado previamente para la detección de la pérdida de heterocigosidad en muestras de tumor (Zhou, W. *et al.* 2002, anteriormente citado) y en el plasma de pacientes con cáncer (Chang, HW *et al.* 2002, anteriormente citado). Dado que la cuantificación de moléculas molde mediante PCR digital no depende de las relaciones dependientes de la dosis entre los colorantes indicadores y las concentraciones de ácido nucleico, su precisión analítica teóricamente debe ser superior a la de la PCR en tiempo real. Por tanto, la PCR digital podría permitir potencialmente la discriminación de grados más sutiles de diferencias cuantitativas entre los loci diana y de referencia.

Para probar esto, los presentes investigadores primero calcularon si la PCR digital podría determinar la proporción alélica de ARNm de *PLAC4* (Lo, YMD, *et al.* 2007 Nat Med 13, 218-223), un transcrito placentario del cromosoma 21, en plasma materno y de este modo distinguir fetos trisómicos 21 y euploides. Se hace referencia a esta aproximación como el método de ARN-SNP digital. A continuación, los presentes inventores evaluaron si la precisión aumentada de la PCR digital permitiría la detección de aneuploidías cromosómicas fetales sin depender de polimorfismos genéticos. Los presentes inventores denominan a esto un análisis de dosis cromosómica relativa (RCD; del inglés, relative chromosome dosage) digital. El enfoque anterior es dependiente del polimorfismo, pero requiere menos precisión en

la discriminación cuantitativa, si bien el último enfoque es independiente del polimorfismo, pero requiere de una precisión más elevada para la discriminación cuantitativa.

## I. RNA-SNP DIGITAL

5

### **A. Visión general**

La PCR digital es capaz de detectar la presencia de sesgo de la proporción alélica de dos alelos en una muestra de ADN. Por ejemplo, se ha utilizado para detectar la pérdida de heterocigosidad (LOH) en una muestra de ADN de tumor. Asumiendo que hay dos alelos en la muestra de ADN, en concreto A y G y que el alelo A se perdería en las células con la LOH. Cuando la LOH está presente en un 50 % de células en la muestra de tumor, la proporción alélica G:A en la muestra de ADN sería 2:1. Sin embargo, si la LOH no está presente en la muestra de tumor, la proporción alélica de G:A sería 1:1.

10

15

La Fig. 1 es un diagrama de flujo 100 que ilustra un experimento de PCR digital. En la etapa 110, la muestra de ADN se diluyó y a continuación, se distribuyó en pocillos separados. Nótese que los inventores han determinado que algunas especies de ácidos nucleicos plasmáticos están ya bastante diluidos en la muestra original. Por consiguiente, no hay necesidad para la dilución de algunos moldes, si ya están presentes en las concentraciones necesarias. En los estudios previos (por ejemplo, Zhou *et al.* 2002, anteriormente citado), una muestra de ADN se diluyó hasta un punto de manera que la concentración promedio de un "ADN molde" específico es aproximadamente 0,5 molécula de uno de los dos moldes por pocillo. Nótese que el término "ADN molde" aparece para referirse a los alelos A o G y que no se proporciona ningún fundamento para esta concentración específica.

20

25

En la etapa 120, en cada pocillo, se lleva a cabo un proceso de PCR para detectar el(los) alelo(s) A y/o G simultáneamente. En la etapa 130, se identifican los marcadores en cada pocillo (por ejemplo, a través de fluorescencia), por ejemplo, A, G, A y G o ninguno. En ausencia de LOH, la abundancia de los alelos A y G en la muestra de ADN sería la misma (una copia por célula). Por tanto, las probabilidades de que un pocillo sea positivo para el alelo A y para el alelo G serían las mismas. Esto se reflejaría por cifras similares de pocillos que sean positivos para los alelos A o G. Sin embargo, cuando la LOH está presente en un 50 % o más de células en una muestra de tumor, la proporción alélica de los alelos G y A sería al menos 2:1. Los métodos previos simplemente asumieron que la muestra era al menos un 50 % cancerosa. Por tanto, la probabilidad de que un pocillo sea positivo para el alelo G sería superior que para el alelo A. Como resultado, el número de pocillos que son positivos para el alelo G sería superior que para el alelo A.

30

35

En la etapa 140, para clasificar los resultados de la PCR digital, se contaría el número de pocillos que son positivos para cada alelo pero no el otro. En el ejemplo anterior, se cuenta el número de pocillos que son positivos para el alelo A pero negativos para el alelo G y el número de pocillos positivos para el alelo G pero negativos para el alelo A. En una realización, el alelo q muestra menos pocillos positivos se considera como el alelo de referencia.

40

En la etapa 150, el número total de pocillos informativos se determina como la suma de las cifras de pocillos positivos para cualquiera de los dos alelos. En la etapa 160, se calcula la proporción ( $P_r$ ) de pocillos informativos (un ejemplo de un parámetro) aportados por el alelo con más pocillos positivos.

45

$P_r = N.^{\circ}$  de pocillos positivos solo para el alelo con más pocillos positivos /  $N.^{\circ}$  total de pocillos positivos solo para un alelo (A o G).

Otras realizaciones pueden utilizar todos los pocillos con uno de los alelos divididos por todos los pocillos con al menos un alelo.

50

En la etapa 170, se determina si el valor de  $P_r$  muestra un desequilibrio alélico. Como se desea exactitud y eficacia, esta tarea no es sencilla. Un método para determinar un desequilibrio utiliza un método de probabilidad de tipo Bayesiano, la prueba de la proporción de probabilidades secuencial (SPRT). La SPRT es un método que permite comparar dos hipótesis probabilísticas según se acumulan datos. En otras palabras, es un método estadístico para clasificar los resultados de la PCR digital sugerentes de la presencia o ausencia de sesgo alélico. Tiene la ventaja de minimizar el número de pocillos a analizar para alcanzar una potencia estadística y una exactitud dadas.

55

En un análisis de SPRT ejemplar, los resultados experimentales se ensayaron frente a la hipótesis nula y a la alternativa. La hipótesis alternativa se acepta cuando hay un sesgo de la proporción alélica en la muestra. La hipótesis nula se acepta cuando no hay un sesgo de la proporción alélica en la muestra. Se comparó el valor de  $P_r$  con dos valores límite para aceptar la hipótesis nula o la alternativa. Si ninguna de las hipótesis se acepta, la muestra se marcó como sin clasificar, que significa que el resultado de la PCR digital observado no es suficiente para clasificar la muestra con la confianza estadística deseada.

60

Los valores límite para aceptar la hipótesis nula o la alternativa se han calculado típicamente basándose en un valor fijo de  $P_r$  bajo las asunciones establecidas en la hipótesis. En la hipótesis nula, se asume que la muestra no exhibe sesgo de la proporción alélica. Por tanto, las probabilidades de cada pocillo de ser positivo para los alelos A y G serían

65

las mismas y por tanto, el valor esperado de  $P_r$  sería de 1/2. En la hipótesis alternativa, el valor esperado de  $P_r$  se toma como de 2/3 o aproximadamente a intermedio entre 0,5 y 2/3, por ejemplo, 0,585. Además, debido al un número limitado de experimentos, se puede elegir un corte superior ( $,585+3/N$ ) y un corte inferior ( $,585-3/N$ ).

## 5 B. Detección del síndrome de Down

En un ejemplo se utiliza la SNP digital para detectar el síndrome de Down fetal a partir del plasma de una mujer gestante. Utilizando marcadores específicos para células fetales/placentarias, se puede medir la proporción de alelos en el cromosoma 21. Por ejemplo, para determinar si un grado observado de sobrerepresentación de un alelo de *PLAC4* es estadísticamente significativo, se utiliza la SPRT.

De acuerdo con un ejemplo, el ARN-SNP digital determina un desequilibrio en la proporción de alelos polimórficos de un SNP A/G, rs8130833, localizado en el ARNm de *PLAC4* que se transcribe del cromosoma 21 y se expresa por la placenta. Para un feto euploide heterocigótico, los alelos A y G deben estar igualmente representados en el genoma fetal (proporción genómica de 1:1); si bien en la trisomía 21, el cromosoma 21 trisómico se asociaría con una copia adicional de uno de los alelos SNP en el genoma fetal, dado una proporción de 2:1. El objetivo del análisis de PCR digital es determinar si las cantidades de los dos alelos de *PLAC4* en la muestra analizada son iguales o de otro modo. Por tanto, los alelos de *PLAC4* tanto A como G son los moldes diana. Se diseñó un ensayo de PCR en tiempo real para amplificar ARNm de *PLAC4* y se discriminaron los dos alelos SNP mediante sondas fluorescentes TaqMan. En la Fig. 2A se muestra una ilustración esquemática de las etapas analíticas.

La Fig. 2A ilustra un método de ARN-SNP digital 200. En la etapa 210, se recibe la muestra. En la etapa 220, la secuencia de ácido nucleico, por ejemplo, ARNm de *PLAC4*, se cuantifica en las muestras de ARN extraídas. En un ejemplo, esto se realiza mediante PCR en tiempo real para el ARNm de *PLAC4*. En un aspecto, esta etapa proporciona al operador una idea acerca de cuánta dilución se necesita antes de que la diana alcance la "esfera" del análisis de la PCR digital.

En la etapa 230, se diluye la muestra. En la etapa 240, se mide una concentración de la muestra diluida. La concentración de la muestra diluida se puede confirmar que es  $\sim 1$  molde/pocillo (es decir, secuencia de referencia o no de referencia o cualquier alelo). Algunos ejemplos utilizan técnicas descritas en la sección IV para esta medición. Por ejemplo, los presentes inventores distribuyeron la muestra diluida en 96 pocillos para el análisis de RCP en tiempo real para confirmar que se había logrado una dilución utilizable. La concentración de dilución se puede dejar también como desconocida, eliminando por tanto esta etapa, como se explicará más adelante.

En la etapa 250, se realiza una PCR digital sobre cada pocillo de la matriz. Por ejemplo, la misma muestra diluida se distribuyó en 384 pocillos para el análisis de PCR en tiempo real. A partir de los resultados de la PCR, se identificó una cantidad de marcadores para cada secuencia de ácido nucleico y el número de pocillos informativos. Un pocillo informativo se define como uno que es positivo solo para el alelo A o el G, pero no para ambos. En la etapa 260, se realizó un cálculo de un valor esperado de  $P_r$ . Estas etapas se describen en más detalle más adelante. El cálculo incluye determinar un parámetro a partir de valores determinados en la etapa 250. Por ejemplo, se puede calcular la concentración de molde promedio real por pocillo.

En la etapa 270, se puede realizar un SPRT u otra prueba de la proporción de probabilidades para determinar si existe o no desequilibrio. Para un caso euploide, los presentes investigadores esperan un número igual de pocillos A-positivos y G-positivos. Sin embargo, cuando se analizan las moléculas molde de un feto con trisomía 21, el número de pocillos que contienen solo un alelo debe ser más elevado que el que contiene solo el otro alelo. En resumen, se espera un desequilibrio alélico para la trisomía 21.

Como se ha mencionado anteriormente, la SPRT es un método de probabilidad de tipo Bayesiano, que permite comparar dos hipótesis probabilísticas según se acumulan datos. En el análisis de PCR digital para la detección de trisomía 21, se acepta la hipótesis alternativa cuando existe desequilibrio alélico (es decir, trisomía 21 detectada); y se acepta la hipótesis nula cuando no existe desequilibrio alélico (es decir, trisomía 21 no detectada). El alelo con el número más elevado de recuentos se denomina como el alelo sobrerepresentado potencialmente y se calcularía su proporción entre todos los pocillos informativos ( $P_r$ ). La SPRT se aplica a determinar si la  $P_r$  indica el suficiente grado de desequilibrio alélico que se esperaría para una muestra de trisomía 21.

Operacionalmente, la SPRT se puede aplicar e interpretar a través de la utilización de gráficos con un par de curvas de SPRT que se construyen para definir los límites probabilísticos para aceptar o rechazar cualquiera de las hipótesis. La Fig. 3 ilustra un gráfico que tiene las curvas de SPRT para determinar el síndrome de Down de acuerdo con una realización de la presente invención. Las curvas de SPRT representan la proporción requerida de pocillos informativos positivos para el alelo sobrerepresentado  $P_r$  potencialmente (eje y), para un número total dado de pocillos informativos (eje x) cuando se pudiera hacer una clasificación segura. Como se muestra en la Fig. 3, la curva superior establece los límites probabilísticos para aceptar de la hipótesis alternativa; si bien la curva inferior establece los límites probabilísticos para aceptar de la hipótesis nula.

El valor de  $P_r$  derivado experimentalmente se compararía con el valor esperado de  $P_r$  para aceptar o rechazar

cualquiera de las hipótesis. Si se aceptó la hipótesis nula, las muestras se clasificaron como que se han obtenido de mujeres gestantes con fetos euploides. Si se acepta la hipótesis alternativa, las muestras se clasificaron como que se han obtenido de mujeres gestantes con fetos trisómicos 21. Como alternativa, cualquiera de las hipótesis podría no aceptarse si la  $P_r$  para el número dado de recuentos informativos no ha alcanzado el nivel requerido de confianza estadística para la clasificación de enfermedades. Estos casos se consideraron inclasificables hasta que se disponga de más casos. Si la clasificación de enfermedad no es posible, se pueden realizar placas de 384 pocillos adicionales hasta que los datos agregados se conviertan en clasificables mediante la SPRT.

La SPRT ofrece por tanto la ventaja de que se requiere una cantidad más pequeña de prueba para un nivel de confianza dado respecto a otros métodos estadísticos. En términos prácticos, la SPRT permite la aceptación o rechazo de cualquiera de las hipótesis tan pronto como la cantidad de datos requeridos se haya acumulado y por tanto minimiza los análisis adicionales innecesarios. Esta característica es de relevancia particular para el análisis de ácidos nucleicos plasmáticos que están presentes habitualmente en bajas concentraciones, en donde el número de moléculas molde disponibles es limitado. Además de una clasificación estricta, la clasificación puede incluir también un porcentaje de exactitud. Por ejemplo, una clasificación que resulta de una comparación con un valor límite puede proporcionar que una muestra exhiba una probabilidad de un desequilibrio de secuencias de ácido nucleico con un cierto porcentaje o de forma equivalente, que un desequilibrio determinado sea exacto hasta cierto porcentaje u otro valor.

Un enfoque similar podría aplicarse para determinar el genotipo de un feto respecto a una mutación o un polimorfismo genético, utilizando ácidos nucleicos fetales en plasma materno o suero. Debe recordarse que un feto heredaría la mitad de su genoma de su madre. Como una ilustración, considérese un locus genético particular con dos alelos, A y B. Si la madre es una heterocigótica con un genotipo de AB, el feto podría tener teóricamente un genotipo de AA, BB o AB. Si el feto tiene un genotipo de AB, es decir, el mismo que el de la madre, a continuación, solo habrá ácidos nucleicos del genotipo AB (tanto de la madre como del feto) en el plasma materno. Por tanto, en el plasma materno se observa el equilibrio de ácidos nucleicos o alélico. Por otro lado, si el feto tiene un genotipo de AA o BB, a continuación, habría un desequilibrio alélico con una sobrerrepresentación del alelo A o el B, respectivamente, en el plasma materno. Esta consideración también es aplicable a las mutaciones que causan enfermedades (por ejemplo, aquellas que causan fibrosis quística o beta-talasemia o atrofia muscular espinal), en cuyo caso se podría considerar A como el alelo silvestre y el B se podría considerar como el alelo mutante.

## II. RCD DIGITAL

Una desventaja del ARN-SNP digital es que solo puede aplicarse a casos de heterocigosis para el SNP analizado. Una mejora que sería ideal si una prueba no invasiva para detectar la trisomía 21 fetal u otra aneuploidía cromosómica fetal (por ejemplo, la trisomía 18, 13 y las aneuploidías cromosómicas sexuales) basada en el análisis de ácidos nucleicos fetales circulantes fuera independiente del uso de polimorfismos genéticos. Por tanto, en un ejemplo, la dosis de cromosomas se determina mediante análisis de PCR digital de un locus del cromosoma 21 no polimórfico respecto a uno situado en un cromosoma de referencia, en concreto, el cromosoma 1 en este estudio. Un cambio de la proporción del cromosoma 21 respecto al cromosoma 1 de 2:2 en el genoma de un feto euploide se diferencia de un caso de trisomía 21. En el análisis de PCR digital para la detección de trisomía 21, las dos hipótesis a comparar serían la hipótesis nula en la que no hay desequilibrio cromosómico (es decir, trisomía 21 no detectada) y la hipótesis alternativa en la que existe un desequilibrio cromosómico (es decir, trisomía 21 detectada).

Este enfoque se puede generalizar a los otros cromosomas implicados en otras aneuploidías cromosómicas, por ejemplo, el cromosoma 18 en la trisomía 18, cromosoma 13 en la trisomía 13, cromosoma X en el síndrome de Turner. Además, aparte del cromosoma 1, pueden utilizarse otros cromosomas no implicados en las aneuploidías afectadas como un cromosoma de referencia. Un enfoque similar se puede aplicar también a la detección del cáncer, analizando el cambio de proporción de un cromosoma suprimido comúnmente, en parte, en el cáncer, respecto a un cromosoma de referencia. Entre los ejemplos de lo primero se incluyen el cromosoma 5q en el cáncer colorrectal, el cromosoma 3p en el cáncer de pulmón y el cromosoma 9p en el carcinoma nasofaríngeo. La Fig. 2B enumera algunas de las anomalías cromosómicas relacionadas con cánceres comunes que dan como resultado un desequilibrio de secuencias.

La Fig. 2A ilustra también un método RCD digital 205. En las etapas 220-230, se cuantifica el ADN extraído, por ejemplo, a través de técnicas Nanodrop y se diluyó hasta una concentración de aproximadamente un molde diana de cromosomas 21 o del cromosoma que normaliza (tal como el cromosoma 1) por pocillo. En la etapa 240, la confirmación se puede realizar analizando la muestra de ADN diluida mediante el ensayo utilizando la sonda del cromosoma 1 solo en un formato de 96 pocillos para confirmar si el nivel ~37 % de los pocillos fueron negativos antes de que se procediera al análisis de RCD digital utilizando ambas sondas TaqMan en placas de 384 pocillos. La significación del 37 % se describirá posteriormente en la Sección IV.

La prueba de la etapa 240 y los resultados de la etapa 250 pueden hacerse con un ensayo de PCR en tiempo real diseñado para amplificar una secuencia paróloga (Deutsch, S. *et al.* 2004 J Med Genet 41, 908-915) presente en ambos cromosomas que se distinguen por variaciones de secuencias parálogas que se discriminan mediante un par de sondas TaqMan. En este contexto, un pocillo informativo se define como uno que es positivo solo para el locus del cromosoma 21 o el locus del cromosoma 1, pero no para ambos. Para un feto euploide, el número de pocillos

informativos positivos para cualquier locus debe ser aproximadamente igual. Para un feto con trisomía 21, debe haber una sobrerrepresentación de pocillos positivos para el cromosoma 21 respecto al cromosoma 1. La proporción exacta de la sobrerrepresentación se describe en las siguientes secciones.

### 5 III. INCORPORACIÓN DEL PORCENTAJE DE SECUENCIAS FETALES

Una desventaja de los métodos 200 y 205 descritos anteriormente es que se necesitan marcadores específicos fetales. Por consiguiente, se utilizan marcadores específicos no fetales. Para utilizar tales marcadores específicos no fetales, se mide la concentración fraccionaria de ADN fetal en el plasma materno (es decir, la muestra biológica). Con tal información, se puede calcular un valor más útil de  $P_r$  como sigue.

Incluso con el pequeño porcentaje fraccionario de ADN fetal en el plasma materno, un feto con trisomía 21 aportaría una dosis adicional de secuencias del cromosoma 21 por equivalentes de genoma (GE; del inglés, genome-equivalent) de ADN fetal liberado en el plasma materno. Por ejemplo, una muestra de plasma materno de una gestante euploide que contiene 50 GE/ml de ADN total con 5 GE/ml de ADN aportado por el feto (es decir, un 10 % de concentración fraccionaria de ADN fetal) debe contener un total de 100 copias (90 copias maternas + 10 copias fetales) de secuencias del cromosoma 21 por mililitro de plasma materno. Para una gestante de trisómico 21, cada GE fetal aportaría tres copias del cromosoma 21, dando como resultado un total de 105 copias/ml (90 copias maternas + 15 copias fetales) de secuencias del cromosoma 21 en el plasma materno. En un 10 % de concentración de ADN fetal, la cantidad de secuencias derivadas del cromosoma 21 en el plasma materno de una gestación de trisómico sería por tanto 1,05 veces la de un caso euploide. Por tanto, si se pudiera desarrollar un enfoque analítico para determinar este pequeño grado de diferencia cuantitativa, se conseguiría una prueba independiente del polimorfismo para el diagnóstico prenatal no invasivo de la trisomía 21 fetal.

Por consiguiente, el grado de sobrerrepresentación sería dependiente de la concentración de ADN fetal fraccionaria en la muestra de ADN analizada. Por ejemplo, cuando el ADN placentario se analiza, la proporción de RCD teórica en el genoma debe ser 3:2, es decir, una diferencia de 1,5 veces. Sin embargo, como se ha descrito anteriormente, la proporción de RCD teórica disminuiría a 1,05 cuando se analiza una muestra de plasma materno que contiene un 10 % de ADN fetal. La  $P_r$  derivada experimentalmente se calcula dividiendo el número de pocillos positivos solo para el locus del cromosoma 21 locus por el número total de pocillos informativos. La  $P_r$  derivada experimentalmente está sujeta al análisis de SPRT con la  $P_r$  calculada y la proporción de RCD teórica.

La Fig. 4 muestra un método 400 para determinar un estado de enfermedad utilizando un porcentaje de ácidos nucleicos fetales. En la etapa 410, se mide el porcentaje fraccionario de material fetal. En una realización, el porcentaje fraccionario se determina midiendo la cantidad de un marcador específico fetal (por ejemplo, el cromosoma Y, marcadores de polimorfismo genético (por ejemplo, SNP), firmas epigenéticas placentarias) en relación con un marcador específico no fetal (es decir, secuencia génica presente tanto en la madre como en el feto). La medición real se podría hacer mediante PCR en tiempo real, PCR digital, reacciones de secuenciación (incluyendo la secuenciación paralela masiva) o cualesquiera otros métodos cuantitativos. En un aspecto, es preferible no utilizar el gen diana que podría estar en desequilibrio alélico potencialmente para esta medición.

En la etapa 420, se realiza una PCR digital u otro método de medición, incluyendo la dilución de la muestra, situando la muestra diluida en los pocillos y midiendo las reacciones en cada pocillo. En la etapa 430, los resultados de la PCR se utilizan para identificar los marcadores de diferentes secuencias de ácidos nucleicos de referencia (tales como cromosomas o alelos). En la etapa 440, se calcula la proporción real ( $P_r$ ) de la secuencia sobrerrepresentada. En la etapa 450, el valor límite para determinar un estado de enfermedad se calcula utilizando el porcentaje de material fetal en la muestra. En la etapa 460, a partir de la  $P_r$  real y el valor límite, se determina si un existe un desequilibrio.

En un ejemplo, el porcentaje fraccionario de las secuencias de ácidos nucleicos de referencia se incorpora en un método de ARN-SNP digital. Por tanto, cuando se investiga una LOH debido a células cancerosas, esto se puede realizar con muestras de tumor con menos de un 50 % de células cancerosas. Se puede utilizar también en muestras con más de un 50 % de células cancerosas para obtener una  $P_r$  más exacta y por tanto reducir el número de falsos positivos, que conducirían a diagnósticos incorrectos. En otra realización, el porcentaje de ácido nucleico fetal se incorpora en un método de PCR digital para determinar si un feto ha heredado una mutación génica parental (por ejemplo, que causa fibrosis quística o beta-talasemia o atrofia muscular espinal) o un polimorfismo, a partir del análisis de ácidos nucleicos plasmáticos maternos.

### IV. INCORPORACIÓN DE UNA CONCENTRACIÓN PROMEDIO POR POCILLO

Otra desventaja de los métodos previos (por ejemplo, Zhou, W. *et al.* 2002, anteriormente citado) es que la concentración promedio de moldes por pocillo ( $m$ ) es necesario que sea de 1 por pocillo. Dado que es difícil obtener una concentración exacta, esto puede conducir a inexactitudes. Asimismo, incluso con una concentración exacta de 1 molde por pocillo, los métodos previos han ignorado la distribución estadística de los moldes en un pocillo. En los métodos previos, es decir, el algoritmo viejo, el valor esperado de  $P_r$  para aceptar la hipótesis alternativa se asume que es la proporción alélica y por tanto, es independiente de la concentración promedio del ADN molde por pocillo.

Sin embargo, debido a una variación estadística natural de los moldes en la muestra diluida, no habrá exactamente 1 molde por pocillo. Se puede medir la concentración promedio de al menos una de las secuencias, que se utiliza a continuación para calcular el valor límite, es decir, la  $P_r$  esperada. En un aspecto, este cálculo implica una distribución estadística para determinar una probabilidad de que un pocillo contenga las secuencias de ácidos nucleicos diferentes, que se utiliza a continuación para determinar la  $P_r$  esperada.

En un ejemplo, la concentración promedio se toma de una secuencia de ácido nucleico de referencia, que en un caso es una secuencia de ácido nucleico con la concentración más baja en la muestra de ADN. En el caso de una muestra sin un desequilibrio, las concentraciones de las dos secuencias en la muestra serían las mismas y cualquiera de las dos se podría considerar como el alelo de referencia. En el caso de una muestra con, por ejemplo, LOH, el alelo que se suprime en las células cancerosas se podría considerar como el alelo de referencia. La concentración promedio del alelo de referencia se indicaría como  $m_r$ . En otro ejemplo, la secuencia con la concentración más elevada se podría tomar como la secuencia de referencia.

#### A. SNP digital. Ejemplo utilizando SPRT y PCR digital

La Fig. 5 muestra un método 500 para determinar un estado de enfermedad utilizando una concentración de molde promedio. En la etapa 510, se mide una cantidad de las diferentes secuencias. Esto se puede hacer, por ejemplo, contando los marcadores en un experimento de PCR digital, como se explicó anteriormente. Sin embargo, se puede hacer mediante otros métodos que no implican una etapa de amplificación o que no utilizan un marcador fluorescente, pero podrían utilizar otras propiedades, tales como propiedades físicas como la masa, propiedades ópticas específicas o propiedades de emparejamiento de bases.

En la etapa 520, se determina la proporción real de la secuencia sobrerrepresentada. Esto puede realizarse como se ha descrito anteriormente tomando el número de pocillos que muestran solo esa secuencia y dividiendo por el número de pocillos informativos. En la etapa 530, se mide la concentración promedio de al menos una de las secuencias (la secuencia de referencia). En un ejemplo, la secuencia de referencia es la secuencia sobrerrepresentada. En otro ejemplo, la secuencia de referencia es la secuencia infrarrepresentada. La medición se puede realizar mediante el recuento del número de pocillos negativos para la secuencia de referencia en el experimento de PCR digital. La relación entre la proporción de pocillos negativos y la concentración de molde promedio se describe por la distribución de Poisson, como se describe en la siguiente subsección.

En la etapa 540, se calcula una cantidad esperada de pocillos positivos para las diferentes secuencias, por ejemplo, utilizando la distribución de Poisson. La cantidad esperada puede ser como una probabilidad de la secuencia por pocillo, una secuencia promedio por pocillo, el número de pocillos que contienen la secuencia o cualquier otra cantidad adecuada. En la etapa 550, la  $P_r$  esperada se calcula a partir de las cantidades esperadas. En la etapa 560, se calcula un valor límite a partir de la  $P_r$  esperada, por ejemplo, mediante una SPRT. En la etapa 570, se determina una clasificación del desequilibrio de las secuencias de ácido nucleico. A continuación, se describen aspectos específicos del método 500.

##### 1. Determinación de la cantidad esperada de secuencias

Una vez se conoce la concentración promedio por pocillo (reacción o mezcla de reacción) a partir de la etapa 530, el número esperado de pocillos que muestra esa secuencia se puede calcular en la etapa 540. Esta cantidad se puede expresar como un %, un valor fraccionario o un valor entero. Utilizando un ejemplo específico con fines ilustrativos, se supone que la concentración promedio del molde de referencia por pocillo ( $m_r$ ) es 0,5 y que el genotipo del feto con trisomía 21 en el SNP de *PLAC4*, rs8130833, es AGG. Por tanto, el molde de referencia sería el alelo A y el molde sobrerrepresentado sería el alelo G.

En un ejemplo, se asume una distribución de Poisson para la distribución del alelo A entre las mezclas de reacción de los pocillos del procedimiento de medición, tal como la PCR digital. En otros ejemplos, se utilizan otras funciones de distribución, tales como la distribución binomial.

$$P(n) = \frac{m^n e^{-m}}{n!}$$

La ecuación de Poisson es: donde,  $n$ =número de moléculas molde por pocillo;  $P(n)$ =probabilidad de  $n$  moléculas molde en un pocillo particular; y  $m$ = número promedio de moléculas molde en un pocillo en un experimento de PCR digital.

Por consiguiente, la probabilidad de que cualquier pocillo no contenga ninguna molécula del alelo A en una concentración promedio del alelo A de 0,5 sería:

$$P(0) = \frac{0,5^0 e^{-0,5}}{0!} = e^{-0,5} = 0,6065.$$



Por tanto, la probabilidad de que cualquier pocillo contenga al menos una molécula del alelo A sería:  $1 - 0,6065 = 0,3935$ . Por tanto, se esperaría que ~39 % de los pocillos contuviera al menos una molécula del alelo A.

En cuanto a la secuencia de ácido nucleico no de referencia, para cada célula de un feto con trisomía 21, la proporción genómica de A respecto a G sería de 1:2. Asumiendo que la proporción de A respecto a G en el ARN o ADN extraído de la muestra permanecería sin cambiar, la concentración promedio del alelo G por pocillo sería dos veces la del alelo A, es decir,  $2 \times 0,5 = 1$ .

Por consiguiente, la probabilidad de que cualquier pocillo no contenga ninguna molécula del alelo G en una concentración promedio del alelo G de 1 sería:

$$P(0) = \frac{1^0 e^{-1}}{0!} = e^{-1} = 0,3679$$

Por tanto, la probabilidad de que cualquier pocillo contenga al menos una molécula del alelo G sería:  $1 - 0,3679 = 0,6321$ . Por tanto, se esperaría que ~63 % de los pocillos contuviera al menos una molécula del alelo G.

## 2. Determinación de la proporción de la secuencia sobrerrepresentada

Después de calcularse las cantidades esperadas, se puede determinar la proporción de la secuencia de ácido nucleico sobrerrepresentada. Asumiendo que el llenado de los pocillos con el alelo A y el alelo G es independiente, la probabilidad de que un pocillo contenga ambos alelos sería de  $0,3935 \times 0,6321 = 0,2487$ . Por tanto, se esperaría que ~25 % de los pocillos contuviera ambos alelos.

La proporción de pocillos esperada que contienen el alelo A pero no el alelo G sería el número de pocillos que contienen al menos un alelo A al que se le deduce el número de pocillos que contienen tanto el alelo A como el alelo G:  $0,3935 - 0,2487 = 0,1448$ . De forma análoga, la proporción de pocillos esperada que contienen el alelo G pero no el alelo A sería:  $0,6321 - 0,2487 = 0,3834$ . Un pocillo informativo se define como un pocillo que es positivo para el alelo A o alelo G, pero no para ambos.

Por tanto, la proporción esperada de pocillos que contienen el alelo A respecto al alelo G en un análisis de ARN-SNP digital es  $0,1448/0,3834$ . En otras palabras, la proporción de pocillos positivos solo para el alelo G es 2,65 veces la de pocillos positivos solo para el alelo A. Esto está en contraste con la proporción genómica fetal en donde el alelo sobrerrepresentado es 2 veces la del otro alelo.

Para el análisis de SPRT, la proporción de los pocillos informativos positivos para el alelo sobrerrepresentado ( $P_r$ ) se calcula e interpreta utilizando las curvas de SPRT. En el ejemplo actual, la proporción de pocillos informativos sería:  $0,1448 + 0,3834 = 0,5282$ . Por tanto, la  $P_r$  esperada de un caso de trisomía 21 en  $m_r$  0,5 es:  $0,3834/0,5282 = 0,73$ .

Como la concentración de molde promedio ( $m$ ) es un parámetro clave en la ecuación de Poisson, la  $P_r$  variará con  $m$ . La Fig. 6 muestra una tabla 600 que tabula la proporción alélica de ARN-SNP digital esperada y la  $P_r$  de muestras de trisomía 21 para un intervalo de concentraciones de molde expresado como la concentración de molde de referencia promedio por pocillo ( $m_r$ ), de acuerdo con una realización de la presente invención. La Tabla 600 muestra la proporción alélica esperada y la proporción de pocillos informativos positivos para el alelo sobrerrepresentado ( $P_r$ ) para una serie de concentraciones de molde de referencia promedio por pocillo ( $m_r$ ).

El valor esperado de  $P_r$  varía con la concentración promedio del alelo de referencia por pocillo ( $m_r$ ) en un modo no lineal. Como se muestra en la Tabla 600, el valor esperado de  $P_r$  para aceptar la hipótesis alternativa aumentaría con  $m_r$ . Como el valor esperado de  $P_r$  para aceptar la hipótesis nula está fijo en 0,5, las muestras con y sin desequilibrio alélico se separarían más ampliamente en términos del valor de  $P_r$  cuando aumenta  $m_r$ . Nótese que en otras realizaciones el valor para aceptar la hipótesis nula puede ser diferente de 0,5. Esto podría ocurrir cuando la proporción normal es diferente de 1:1, por ejemplo, 5:3 y por tanto, ocurriría un desequilibrio cuando la proporción se desviara de 5:3. La diferencia en las cantidades de las dos secuencias de ácido nucleico diferentes a continuación se determinaría individualmente.

Sin embargo, como en los métodos previos (por ejemplo, Zhou, W. *et al.* 2002, anteriormente citado) utilizaron un valor esperado de  $P_r$  fijo para las muestras de LOH, subestimaron el valor de  $P_r$  para aquellas muestras con LOH (hipótesis alternativa aceptada). El grado de subestimación aumentaría con  $m_r$ . En otras palabras, cuanto mayor es la concentración promedio del alelo de referencia en la muestra de ADN, más inexactos son los viejos métodos. Esta subestimación de  $P_r$  para aceptar la hipótesis alternativa conduciría al cálculo inexacto de los valores límite para aceptar tanto la hipótesis nula como la alternativa.

## 3. Cálculo de los valores límite basados en la $P_r$ esperada

Para los ejemplos que utilizan una SPRT, se pueden utilizar las ecuaciones para calcular los límites superiores e

inferiores de las curvas de SPRT a partir de El Karoui *et al.* (2006). Asimismo, el nivel de confianza estadística preferido para aceptar la hipótesis nula o alternativa podría variarse ajustando el umbral de probabilidad en las ecuaciones. En este estudio, se utiliza una proporción del umbral de probabilidad de 8 debido a que este valor se ha demostrado que proporciona un rendimiento satisfactorio para discriminar muestras con y sin desequilibrio alélico en el contexto de la detección del cáncer. Por tanto, las ecuaciones para calcular los límites superiores e inferiores de las curvas de SPRT son:

$$\text{Límite superior} = [(\ln 8)/N - \ln \delta]/\ln \gamma$$

$$\text{Límite inferior} = [(\ln 1/8)/N - \ln \delta]/\ln \gamma$$

donde,

$$\delta = (1 - \theta_1)/(1 - \theta_0)$$

$$\gamma = -(\theta_1(1 - \theta_0)/\theta_0(1 - \theta_1))$$

$\theta_0$  = proporción de pocillos informativos que contienen el alelo no de referencia si la hipótesis nula es verdadera = 0,5 (véase más adelante)

$\theta_1$  = proporción de pocillos informativos que contienen el alelo no de referencia (es decir, sobrerrepresentado) si la hipótesis alternativa es verdadera

N = número de pocillos informativos = número de pocillos para cualquier alelo solo

(ln es un símbolo matemático que representa el logaritmo natural, es decir,  $\log\theta$ .)

Para la determinación de  $\theta_0$  para aceptar la hipótesis nula, se asume que la muestra se ha obtenido de una mujer gestante que porta un feto euploide. Bajo esta asunción, el número esperado de pocillos positivos para cualquier molde sería 1:1 y por tanto, la proporción esperada de pocillos informativos que contienen el alelo no de referencia sería de 0,5.

Para la determinación de  $\theta_1$  para aceptar la hipótesis alternativa, se asume que la muestra se ha obtenido de una mujer gestante que porta un feto con trisomía 21. Los cálculos para la  $P_r$  esperada de casos de trisomía 21 para el análisis de ARN-SNP digital se detallan en la Tabla 600. Por tanto,  $\theta_1$  para el análisis de ARN-SNP digital se refiere a los datos mostrados en la última columna de la tabla 600.

La medición de  $m_r$  se puede realizar a través de diferentes mecanismos como conocen o conocerán los expertos en la materia. En un ejemplo, se determina el valor de  $m_r$  durante el proceso experimental del análisis de PCR digital. Como la relación entre el valor de  $m_r$  y el número total de pocillos que son positivos para el alelo de referencia se puede regir por una distribución (por ejemplo, la distribución de Poisson, se puede calcular la  $m_r$  a partir del número de pocillos que son positivos para el alelo de referencia utilizando esta fórmula:

$$m_r = -\ln(1 - \text{proporción de pocillos que son positivos para el alelo de referencia})$$

Nótese que ln es el logaritmo natural, es decir,  $\log_e$ . Este enfoque proporciona una estimación directa y precisa de  $m_r$  en la muestra de ADN utilizada para el experimento de PCR digital.

Este método puede utilizarse para lograr una concentración deseada. Por ejemplo, los ácidos nucleicos extraídos de una muestra se pueden diluir hasta una concentración específica, tal como una molécula molde por pocillo de reacción, como se hizo en la etapa 240 del método 200. En un ejemplo utilizando la distribución de Poisson, la proporción esperada de pocillos sin molde se puede calcular como  $e^{-m}$ , donde  $m$  es la concentración promedio de moléculas molde por pocillo. Por ejemplo, a una concentración promedio de una molécula molde por pocillo, la proporción esperada de pocillo sin molécula molde viene dada por  $e^{-1}$ , es decir, 0,37 (37 %). El 63 % de pocillos restantes contendrán una o más moléculas molde. Normalmente, a continuación se contaría el número de pocillos positivos e informativos en una ejecución de PCR digital. La definición de pocillos informativos y la manera en la que los datos de PCR digital se interpretan dependen de la aplicación.

En otros ejemplos, la concentración promedio por pocillo,  $m_r$ , se mide mediante otro método de cuantificación, por ejemplo, PCR cuantitativa en tiempo real, PCR competitiva semicuantitativa, PCR competitiva real utilizando métodos de espectrometría de masas, etc.

## B. RCD digital

Se puede realizar una RCD digital utilizando la concentración promedio de una forma similar al método de SNP digital descrito anteriormente. Las cifras de pocillos positivos para el marcador del cromosoma de referencia (no el cromosoma 21), el marcador del cromosoma 21 y ambos marcadores, se pueden determinar mediante PCR digital. La concentración promedio del marcador de referencia por pocillo ( $m_r$ ) se puede calcular a partir del número total de

pocillos negativos para el marcador de referencia, independientemente del positivo del marcador del cromosoma 21, de acuerdo con la función de probabilidad de Poisson, como en el cálculo de  $m_r$  para el análisis de SNP digital.

A continuación, el análisis de SPRT puede utilizarse para clasificar una muestra de plasma como proveniente de una mujer gestante que porta un feto euploide o trisómico 21. La hipótesis nula se aceptaría cuando el feto fuese euploide. En este escenario, la proporción esperada para los pocillos positivos para el marcador de referencia y el marcador del cromosoma 21 sería 1:1 y por tanto, la proporción esperada de pocillos informativos con señal positiva para el marcador del cromosoma 21 sería 0,5. La hipótesis alternativa se aceptaría cuando el feto fuese trisómico para el cromosoma 21. En este escenario, si la muestra de ADN se derivara solamente del feto, la concentración promedio del marcador del cromosoma 21 en cada pocillo sería 3/2 veces la concentración promedio del marcador de referencia ( $m_r$ ).

Si bien el RCD digital puede utilizarse para determinar la dosis de cromosomas a través de la detección de marcadores específicos fetales, por ejemplo, las firmas epigenéticas de la placenta (Chim, SSC. *et al.* 2005 Proc Natl Acad Sci USA 102, 14753-14758), un ejemplo del análisis de RCD digital utiliza marcadores específicos no fetales. Por tanto, una etapa adicional de medición del porcentaje de material fetal tendría lugar cuando se utilizan marcadores específicos no fetales. Por tanto, la concentración promedio del marcador del cromosoma 21 por pocillo sería dependiente de la proporción del ADN fetal en la muestra y podría calcularse utilizando:  $m_r [(200 \% + \text{porcentaje de ADN fetal})/200 \%]$ .

Utilizando, de nuevo, un ejemplo específico con fines ilustrativos, la concentración promedio del molde de referencia, cromosoma 1, por pocillo ( $m_r$ ) se asume que es 0,5 y se asume que un 50 % del ADN se deriva del feto y que un 50 % del ADN en la muestra se deriva de la madre.

Por consiguiente, utilizando la distribución de Poisson, la probabilidad de que cualquier pocillo no contenga ninguna molécula del locus del cromosoma 1 cuando su concentración promedio es 0,5 por pocillo sería:

$$P(0) = \frac{0,5^0 e^{-0,5}}{0!} = e^{-0,5} = 0,6065$$

Por tanto, la probabilidad de que cualquier pocillo contenga al menos una molécula del locus del cromosoma 1 sería:  $1 - 0,6065 = 0,3935$ . Por tanto, se esperaría que ~39 % de los pocillos contuviera al menos una molécula del locus.

Para cada célula de este feto con trisomía 21, la proporción genómica del cromosoma 21 frente al cromosoma 1 sería 3:2. La proporción entre el cromosoma 21 y el cromosoma 1 en la muestra de ADN dependería de la concentración de ADN fetal fraccionaria (% de ADN fetal) y sería:  $3 \text{ } 3 \% \text{ de ADN fetal} + 2 (1 - \% \text{ ADN fetal}) : 2 \text{ } 3 \% \text{ de ADN fetal} + 3 (1 - \% \text{ ADN fetal})$ . Por tanto, en este caso cuando la concentración de ADN fetal fraccionaria es un 50 %, la proporción sería:  $(3 \text{ } 3 50 \% + 2 \text{ } 3 50 \%) / (2 \text{ } 3 50 \% + 3 \text{ } 3 50 \%) = 1,25$ . Si el método de SNP digital no utilizó marcadores específicos fetales, tal cálculo se podría utilizar también para calcular la concentración promedio de la secuencia no de referencia.

Por tanto, cuando la concentración promedio del locus del cromosoma 1 por pocillo es 0,5, la concentración promedio del locus del cromosoma 21 por pocillo es:  $1,25 \text{ } 0,5 = 0,625$ . Por consiguiente, la probabilidad de que cualquier pocillo no contenga ninguna molécula del locus del cromosoma 21 cuando su concentración promedio es 0,625 por pocillo sería:

$$P(0) = \frac{0,625^0 e^{-0,625}}{0!} = e^{-0,625} = 0,5353$$

Por tanto, la probabilidad de que cualquier pocillo contenga al menos una molécula del locus del cromosoma 21 sería:  $1 - 0,5353 = 0,4647$ . Por tanto, se esperaría que ~46 % de los pocillos contenga al menos una molécula del locus.

Asumiendo que el llenado de los pocillos con cada loci es independiente, la probabilidad de que un pocillo contenga ambos loci sería de  $0,3935 \text{ } 0,4647 = 0,1829$ . Por tanto, se esperaría que ~18 % de los pocillos contuviera ambos loci.

La proporción de pocillos esperada que contienen el locus del cromosoma 1 pero no el locus del cromosoma 21 sería el número de pocillos que contienen al menos un locus del cromosoma 1 al que se le deduce el número de pocillos que contienen ambos loci:  $0,3935 - 0,1829 = 0,2106$ . De forma análoga, la proporción de pocillos esperada que contienen el locus del cromosoma 21 pero no ambos loci sería:  $0,4647 - 0,1829 = 0,2818$ . Un pocillo informativo se define como un pocillo positivo para el locus del cromosoma 1 o el locus del cromosoma 21, pero no para ambos.

Por tanto, la proporción esperada del cromosoma 21 respecto al cromosoma 1 en el análisis de RCD digital es  $0,2818/0,2106 = 1,34$ . En otras palabras, la proporción de pocillos positivos solo para el locus del cromosoma 21 es 1,34 veces la de pocillos positivos solo para el locus del cromosoma 1. Esto está en contraste con la proporción de

1,25 en la muestra de ADN.

Para el análisis de SPRT, se necesitaría que la proporción de los pocillos informativos positivos para el locus del cromosoma 21 ( $P_r$ ) se calculara e interpretara utilizando las curvas de SPRT. En el ejemplo actual, la proporción de pocillos informativos sería:  $0,2106 + 0,2818 = 0,4924$ . Por tanto, la  $P_r$  esperada de un caso de trisomía 21 con un 50 % de ADN fetal en  $m_r$  0,5 es:  $0,2818/0,4924 = 0,57$ .

Como la concentración de molde promedio ( $m$ ) es un parámetro clave en la ecuación de Poisson, la  $P_r$  variará con  $m$ . La Fig. 7 muestra una tabla 700 que tabula la  $P_r$  esperada para las concentraciones de ADN fetal fraccionarias de un 10 %, 25 %, 50 % y 100 % en muestras de trisomía 21 en un intervalo de concentraciones de molde expresado como la concentración de molde de referencia promedio por pocillo ( $m_r$ ) de acuerdo con una realización la presente invención. Los cálculos para la  $P_r$  esperada de casos de trisomía 21 para el análisis de RCD digital se detallan en la Tabla 700. Por tanto,  $\theta_1$  para el análisis de RCD digital de muestras con concentraciones de ADN fetal fraccionarias variables, se puede obtener a partir de las columnas que muestran los valores de  $P_r$  esperados correspondientes en la tabla 700.

### C. Resultados

#### 1. Comparación de diferentes $m_r$

La base para la diferencia entre el grado teórico (como en el genoma fetal) y el esperado experimentalmente de desequilibrio alélico o cromosómico y los cálculos para determinar el último para un intervalo de valores de  $m_r$  se muestran en las tablas 600 y 700. En el análisis de ARN-SNP digital de una muestra de trisomía 21, cuando  $m_r=0,5$ , los pocillos que contienen solo el alelo sobrerrepresentado con respecto a los pocillos que contienen solo el alelo de referencia, en concreto la proporción ARN-SNP digital, es 2,65 (tabla 600). En el análisis de RCD digital de un espécimen compuesto de un 100 % de ADN fetal, cuando  $m_r=0,5$ , los pocillos positivos solamente para el locus del cromosoma 21 con respecto a aquellos positivos solamente para el locus del cromosoma 1, en concreto la proporción RCD digital, es 1,7 (tabla 700) ( $P_r = 0,63$ , por tanto, la proporción de RCD digital es  $0,63/(1-0,63) = 1,7$ ). Según la concentración de ADN fetal fraccionaria disminuye, la proporción de RCD digital disminuye para la misma  $m_r$  (tabla 700).

Como se muestra en las Tabla 600 y 700, la extensión de sobrerrepresentación alélica o cromosómica se incrementa con  $m_r$ . Sin embargo, el porcentaje de pocillos informativos se aproxima a su máximo alrededor de  $m_r = 0,5$  y disminuye gradualmente con aumento adicional en  $m_r$ . En la práctica, la disminución en la proporción de pocillos informativos se podría compensar aumentando el número total de pocillos analizados si la cantidad de molécula molde para el espécimen no es limitante, pero los pocillos adicionales requerirían un aumento en costes de reactivos. Por tanto, el rendimiento de PCR digital óptimo es una compensación entre la concentración de molde y el número total de pocillos ensayados por muestra.

#### 2. Ejemplo utilizando las curvas de SPRT

Como se ha analizado anteriormente, el grado esperado de desequilibrio alélico o cromosómico para un experimento de PCR digital depende de la concentración de molde real por mezcla de reacción (por ejemplo, un pocillo). Los presentes inventores describen la concentración de molde basada en el alelo de referencia, es decir, la concentración de molde de referencia promedio por pocillo ( $m_r$ ). Como se muestra en la ecuación anterior, la  $P_r$  esperada se puede utilizar para determinar la representación de las curvas de SPRT superior e inferior. Dado que la  $P_r$  esperada es, a su vez, dependiente del valor de  $m_r$ , la representación de las curvas de SPRT dependerían esencialmente del valor de  $m_r$ . Por tanto, en la práctica, se necesitaría un conjunto de curvas de SPRT relevantes para la  $m_r$  real de un conjunto de datos de PCR digital para la interpretación de la  $P_r$  a partir de esa ejecución particular.

La Fig. 8 muestra un gráfico 800 que ilustra el grado de diferencias en las curvas de SPRT para los valores de  $m_r$  de 0,1, 0,5 y 1,0 para los análisis de ARN-SNP digital de acuerdo con una realización de la presente invención. Cada conjunto de datos de PCR digital debe interpretarse con las curvas específicas relevantes para el valor de  $m_r$  exacto de esa ejecución particular. Nótese que dado que grado de desequilibrio alélico o cromosómico esperado para los enfoques de ARN-SNP y RCD digitales son diferentes (2:1 para el primero y 3:2 para el último), se necesitan diferentes conjuntos de curvas de SPRT para los dos sistemas de PCR digital. La  $P_r$  derivada experimentalmente se interpreta con las curvas de SPRT relevantes seleccionadas mediante la  $m_r$  correspondiente de la ejecución de PCR digital. Esto está en contraste con el uso informado anterior de la SPRT para la detección molecular de LOH mediante PCR digital en donde se ha utilizado un conjunto fijo de curvas.

La forma práctica de interpretar los datos de PCR digital utilizando SPRT se ilustra más adelante utilizando una ejecución de ARN-SNP digital hipotética. Tras el análisis de ARN-SNP digital de cada caso, se cuenta el número de pocillos positivos para el alelo A solo, el alelo G solo o ambos alelos. El alelo de referencia se define como el alelo con el número más pequeño de pocillos positivos. El valor de  $m_r$  se calcula utilizando el número total de pocillos negativos para el alelo de referencia, independientemente de si el otro alelo es positivo, de acuerdo con la función de densidad de probabilidad de Poisson. Los datos del ejemplo hipotético de los presentes investigadores son como sigue:

En una reacción de 96 pocillos, 20 pocillos son positivos para el alelo A solo, 24 pocillos son positivos para el alelo G solo y 33 pocillos son positivos para ambos alelos. El alelo A se considera como el alelo de referencia debido a que existen menos pocillos positivos para A que positivos para G. El número de pocillos negativos para el alelo de referencia es  $96 - 20 - 33 = 43$ . Por tanto, se puede calcular  $m_r$  utilizando la ecuación de Poisson:  $-\ln(43/96) = 0,80$ . La  $P_r$  de este caso determinada experimentalmente es:  $24 / (20 + 24) = 0,55$ .

De acuerdo con la Tabla 600, la  $P_r$  esperada de una muestra de trisomía 21 en  $m_r = 0,8$  es 0,76. Por tanto,  $\theta_1$  es 0,76 para este caso. Las curvas de SPRT basadas en  $\theta_1 = 0,76$  se utilizarían para interpretar la  $P_r$  derivada experimentalmente de este caso que es 0,55. Cuando se ajusta la  $P_r = 0,55$  en las curvas de SPRT relevantes, los puntos de datos caen bajo la curva inferior. Por tanto, el caso se clasifica como euploide, véase la Fig. 3.

### 3. Comparación con el método viejo

La Fig. 9A muestra una tabla 900 de una comparación de la eficacia de los algoritmos de SPRT nuevo y viejo para clasificar casos de euploidía y trisomía 21 en los análisis de ARN-SNP digital de 96 pocillos. La Fig. 9B muestra una tabla 950 de una comparación de la eficacia de los algoritmos de SPRT nuevo y viejo para clasificar casos de euploidía y trisomía 21 en los análisis de ARN-SNP digital de 384 pocillos. El nuevo algoritmo se refiere a la selección de curvas de SPRT específicas para la  $m_r$  derivada de los datos de PCR digital. El viejo algoritmo se refiere al uso de un conjunto fijo de curvas de SPRT para todas las ejecuciones de PCR digital. El efecto de un cálculo incorrecto de los valores límite sobre la exactitud de la clasificación se revela mediante los análisis de simulación mostrados en la tabla 900.

En comparación con el uso de un conjunto fijo de curvas de SPRT en estudios previos, la proporción de datos inclasificables es mucho más baja con el enfoque de los presentes inventores, como se muestra en las tablas 900 y 950. Por ejemplo, utilizando el enfoque de los presentes inventores, con  $m_r = 0,5$ , un 14 % y un 0 % de muestras de trisomía 21 serían inclasificables para los análisis de ARN-SNP digital de 96 pocillos y de 384 pocillos, respectivamente, pero serían un 62 % y un 10 %, respectivamente, con el uso de curvas fijas (véase la Tabla 900). Por tanto, el enfoque de los presentes inventores permite la clasificación de enfermedades con menor número de pocillos informativos.

Como se muestra en la Tabla 900, el nuevo algoritmo es más exacto en la clasificación de las muestras como que tienen o no tienen sesgo en la proporción alélica para todos los valores de  $m_r$  de 0,1 a 2,0. Por ejemplo, cuando  $m_r$  es igual a 1,0 y se realiza una ejecución de ARN-SNP digital de 96 pocillos, el nuevo algoritmo clasifica correctamente un 88 % y un 92 % de muestras con y sin sesgo de la proporción alélica, respectivamente, si bien el porcentaje de clasificación correcta de muestras con y sin sesgo de la proporción alélica utilizando el algoritmo viejo es solo un 19 % y un 36 %, respectivamente.

Utilizando el algoritmo nuevo, la separación de muestras con y sin sesgo de la proporción alélica aumentaría con  $m_r$ . Como resultado, las exactitudes de la clasificación aumentarían con  $m_r$ . El efecto de aumentar la separación de los dos grupos de muestras sobre la exactitud de la clasificación disminuiría cuando aumentara  $m_r$  a más de 2,0 debido a la reducción en el porcentaje de pocillos informativos. Por el contrario, utilizando el algoritmo viejo, las exactitudes de la clasificación se reducen significativamente cuando aumenta  $m_r$  debido a la desviación aumentada del valor de  $P$  esperado respecto a su valor verdadero.

Los datos experimentales y de simulación de los presentes investigadores muestran que el ARN-SNP digital es un método eficaz y exacto para la detección de la trisomía 21. Dado que el ARNm de PLAC4 en el plasma materno se deriva puramente del feto, para 12 de las 13 muestras de plasma materno ensayadas, solo se necesitó uno de los experimentos de PCR digital de 384 pocillos para una clasificación correcta (Tabla 1350 de la Fig. 13B). Este enfoque homogéneo basado en PCR digital en tiempo real ofrece por tanto una alternativa al enfoque basado en espectrometría de masas para el análisis de ARN-SNP (Lo, YMD, *et al.* 2007 Nat Med, anteriormente citado). Además de los transcritos de ARNm específicos de placenta, los presentes inventores también conciben que se podrían utilizar otros tipos de especies de ácidos nucleicos específicos fetales en plasma materno para la detección basada en PCR digital de aneuploidías cromosómicas fetales. Un ejemplo son los marcadores epigenéticos fetales (Chim, SSC *et al.* (2005) Proc Natl Acad Sci USA 102, 14753-14758; Chan, KCA *et al.* (2006) Clin Chem 52, 2211-2218), que se han utilizado recientemente para la detección prenatal no invasiva de trisomía 18 utilizando el enfoque de la proporción alélica epigenética (EAR; del inglés, epigenetic allelic ratio) (Tong, YK *et al.* (2006) Clin Chem 52, 2194-2202). Por tanto, los presentes inventores predicen que la EAR digital sería una técnica analítica posible.

### V. AUMENTO DEL %, MARCADORES MÚLTIPLES Y ALTERNATIVAS DE PCR

Como se ha descrito anteriormente, la aplicación de las realizaciones de la presente invención al ADN extraído de plasma materno puede complicarse cuando el ADN fetal solo constituye una fracción menor del ADN del plasma materno, con una concentración fraccionaria promedio de alrededor de un 3% entre las semanas 11 y 17 de gestación. No obstante, como se muestra en el presente documento, el RCD digital permite la detección de aneuploidías incluso cuando el ADN aneuploide está presente como una población menor. Con una concentración fraccionaria decreciente de ADN fetal, tal como puede estar presente durante la gestación temprana, se necesita un mayor número de

recuentos informativos para el RCD digital. Un significado del presente trabajo, como se resume en la tabla 1200 de la Fig. 12, es que los presentes investigadores han proporcionado un conjunto de parámetros de referencia, por ejemplo, se requiere ADN fetal fraccionario y moléculas molde totales, sobre las que se pueden construir ensayos de diagnóstico. En opinión de los presentes investigadores, el número total de 7680 reacciones para una concentración de ADN fetal fraccionaria de un 25% es un conjunto particularmente atractivo de parámetros de referencia. Estos parámetros deberían permitir que las muestras de euploides y de trisomía 21 se clasifiquen correctamente un 97% de las veces, como se muestra en la tabla 1200.

El número de moléculas de ADN plasmático que están presentes por unidad de volumen de plasma materno es limitado (Lo, YMD. *et al.* 1998 Am J Hum Genet 62, 768-7758). Por ejemplo, en la gestación temprana, la concentración plasmática materna promedio de un locus autosómico, el gen  $\beta$ -globina, se ha demostrado que es de 986 copias/ml, con contribuciones tanto del feto como de la madre (Lo, YMD. *et al.* 1998 Am J Hum Genet 62, 768-7758). Para capturar 7.680 moléculas, se necesitaría ADN extraído de unos 8 ml de plasma materno. Este volumen de plasma, obtenible de ~15 ml de sangre materna, está en el límite de la práctica habitual. Sin embargo, los presentes investigadores prevén que se pueden combinar múltiples conjuntos diana de cr21 (cr, cromosoma) y cromosomas de referencia para el análisis digital de RCD. Para 5 pares de cr21 y cromosomas de referencia diana, solo se necesitarían 1,6 ml de plasma materno para proporcionar el número de moléculas molde necesarias para el análisis. Podría realizarse una PCR multiplexada de una sola molécula. La solidez de dicho análisis de una sola molécula multiplexado se ha demostrado previamente para el haplotipado de una sola molécula (Ding, C. y Cantor, CR. 2003 Proc Natl Acad Sci USA 100, 7449-7453).

Como alternativa, para lograr una concentración de ADN fetal fraccionaria del 25%, los métodos pueden permitir el enriquecimiento selectivo del ADN fetal (Li, Y. *et al.* 2004 Clin Chem 50, 1002-1011) o la supresión del ADN materno de fondo (Dhallan, R *et al.* 2004 JAMA 291, 1114-1119) o ambos, en el plasma materno. Además de tales métodos físicos para el enriquecimiento del ADN fetal y la supresión del ADN materno, también sería posible utilizar una estrategia de enriquecimiento molecular, tal como dirigirse a moléculas de ADN fetal que exhiban un patrón de metilación de ADN particular (Chim, SSC *et al.*, 2005 Proc Natl Acad Sci USA 102, 14753-14758, Chan, KCA *et al.* 2006 Clin Chem 52, 2211-2218; Chiu, RWK *et al.* 2007 Am J Pathol 170, 941-950.)

Adicionalmente, ahora hay varios enfoques alternativos a la configuración manual de análisis de PCR digital en tiempo real como se utiliza en el estudio actual para realizar la PCR digital. Estos enfoques alternativos incluyen chips de PCR digital microfluídicas (Warren *et al.* 2006 Proc Natl Acad Sci USA 103, 17807-17812; Ottesen, EA *et al.* 2006 Science 314, 1464-1467), PCR en emulsión (Dressman, D *et al.* 2003 Proc Natl Acad Sci USA 100, 8817-8822) y secuenciación genómica paralela masiva (Margulies, M. *et al.* 2005 Nature 437, 376-380) utilizando por ejemplo la plataforma Roche 454, la plataforma de Illumina Solexa y el sistema SOLiD™ de Applied Biosystems, etc. Con respecto a lo último, el método de los presentes inventores también es aplicable a métodos de secuenciación paralela masiva en moléculas de ADN individuales, que no requieren un paso de amplificación, tal como la tecnología de secuenciación de ADN Helicos True Single Molecule DNA (Harris TD *et al.* 2008 Science, 320, 106-109), la molécula individual, tecnología en tiempo real (SMRT™) de Pacific Biosciences y la secuenciación de nanoporos (Soni GV y Meller A. 2007 Clin Chem 53, 1996-2001). Con la utilización de estos métodos, el ARN-SNP digital y el RCD digital podrían realizarse rápidamente sobre un gran número de muestras, mejorando por tanto la viabilidad clínica de los métodos propuestos aquí para el diagnóstico prenatal no invasivo.

## Ejemplos

Los siguientes ejemplos se ofrecen para ilustrar, pero no para limitar la invención reivindicada.

### I. SIMULACIONES POR ORDENADOR

Se realizó una simulación por ordenador para estimar la exactitud del diagnóstico de trisomía 21 utilizando el enfoque de SPRT. La simulación por ordenador se realizó con el software Microsoft Excel 2003 (Microsoft Corp., EE.UU.) y el software SAS 9.1 para Windows (SAS Institute Inc., Carolina del Norte, EE.UU.). El rendimiento de la PCR digital es una interacción entre la concentración del molde de referencia ( $m_r$ ), el número de recuentos informativos y el grado proyectado de desequilibrio alélico o cromosómico ( $P_r$ ). Se realizaron simulaciones por separado para una gama de cada una de estas variables. Dado que los límites de decisión de las curvas SPRT para el ARN-SNP digital y el RCD digital eran diferentes, los análisis de simulación para los dos sistemas se realizaron por separado.

Para cada condición de PCR digital simulada (es decir,  $m_r$ , concentración de ADN fetal fraccionaria, número total de pocillos), se realizaron dos rondas de simulación. La primera ronda simuló el escenario de que las muestras analizadas se obtuvieron de mujeres gestantes portadoras de fetos euploides. La segunda ronda simuló el escenario donde las muestras analizadas se obtuvieron de mujeres gestantes con fetos con trisomía 21. Para cada ronda, se simuló la prueba de 5000 fetos.

### A. ARN-SNP

Para el ARN-SNP digital, se realizaron simulaciones de un experimento de 384 pocillos con  $m_r = 0,1$  a  $m_r = 2,0$ . En

cada valor de  $m_r$ , los presentes inventores simularon el escenario de modo que se ensayaron 5000 fetos euploides y 5000 fetos con trisomía 21. Las curvas de SPRT adecuadas para la  $m_r$  dada se utilizaron para clasificar los 10.000 fetos. La Fig. 10 es una tabla 1000 que muestra los porcentajes de fetos clasificados correcta e incorrectamente como euploides o aneuploides y aquellos no clasificables para los recuentos informativos dados. Las exactitudes para diagnosticar casos euploides y aneuploides son ambas de un 100%, para  $m_r$  entre 0,5 y 2,0. Cuando  $m_r=0,1$ , sólo un 57% y un 88% de los fetos euploides y con trisomía 21 pudieron clasificarse con precisión después del análisis de 384 pocillos.

Los datos de simulación se generaron como se describe en las siguientes etapas:

En la etapa 1, para cada pocillo, se generaron dos números aleatorios utilizando la función Random(Poisson) del programa informático SAS ([www.sas.com/technologies/analytics/statistics/index.html](http://www.sas.com/technologies/analytics/statistics/index.html)) para representar los alelos A y G, respectivamente. La función Random(Poisson) generaría números enteros positivos comenzando a partir del 0 (es decir, 0, 1, 2, 3,...) y la probabilidad de que se genere cada número entero dependió de la probabilidad de este número de acuerdo con la función de densidad de probabilidad de Poisson para un valor promedio dado que representó la concentración promedio de los alelos por pocillo. Un pocillo se consideró positivo para el alelo A si el número aleatorio que representa el alelo A fue mayor que cero, es decir, contenía 1 o más moléculas del alelo A. De forma análoga, el pocillo se consideró como positivo para el alelo G si el número aleatorio que representa el alelo G fue mayor que cero.

Para simular el escenario de una mujer gestante que porta un feto euploide, se utilizó el mismo valor promedio para generar los números aleatorios para el alelo A y el alelo G. Por ejemplo, en el análisis que simula el análisis de ARN-SNP digital con  $m_r = 0,5$ , el valor promedio para el alelo A o el alelo G se estableció de forma idéntica en 0,5, lo que significó una concentración promedio para cada alelo de 0,5 moléculas por pocillo. Utilizando la ecuación de Poisson, con una concentración promedio de 0,5, la proporción de pocillos positivos para los alelos A y G serían las mismas y fue de 0,3935, véase la tabla 600.

Cuando se simula el análisis de ARN-SNP digital de una mujer gestante con un feto con trisomía 21 con  $m_r = 0,5$ , la concentración promedio del alelo sobrerrepresentado por pocillo se esperaría que fuera 2 veces la del alelo de referencia, es decir, 1. En esta situación, la probabilidad de que un pocillo sea positivo para el alelo sobrerrepresentado fue de 0,6321, véase la tabla 600.

Después de generar un número aleatorio para un pocillo de PCR digital, el pocillo podría clasificarse en uno de los siguientes estados:

- a. negativo para los alelos A y G
- b. positivo para los alelos A y G
- c. positivo para el alelo A pero negativo para el alelo G
- d. positivo para el alelo G pero negativo para el alelo A

En la etapa 2, se repitió el paso 1 hasta que se generó el número deseado de pocillos, 384 pocillos para la simulación actual. Se contó el número de pocillos positivos solo para el alelo A y solo para el alelo G. El alelo con menos pocillos positivos se consideró como el alelo de referencia y el alelo con más pocillos positivos se consideró como el alelo potencialmente sobrerrepresentado. El número de pocillos informativos fue el número total de pocillos positivos para cualquiera de los alelos, pero no para ambos. A continuación, se calculó la proporción de pocillos informativos que contienen el alelo potencialmente sobrerrepresentado ( $P_r$ ). Se calcularon los límites superior e inferior para las curvas de SPRT relevantes para aceptar la hipótesis nula o alternativa.

En la etapa 3, se realizaron 5000 simulaciones para cada uno de los dos escenarios de la mujer gestante que porta un feto euploide o con trisomía 21. Cada simulación podría considerarse como una muestra biológica independiente obtenida de mujeres gestantes. En la Tabla 1000, la clasificación correcta de casos euploides se refiere a aquellos casos euploides en los que se aceptó la hipótesis nula y la clasificación incorrecta de casos euploides se refiere a aquellos casos euploides en los que se aceptó la hipótesis alternativa. De forma análoga, aquellos casos de trisomía 21 en los que se aceptó la hipótesis alternativa se consideraron como correctamente clasificados y aquellos casos de trisomía 21 en los que se aceptó la hipótesis nula se consideraron incorrectamente clasificados. Para ambos grupos, aquellos casos en los que no se aceptó la hipótesis nula o la alternativa, después de que se hubiera simulado el número total preespecificado de pocillos, se consideraron sin clasificar.

En la etapa 4, las etapas 1 a 3 se realizaron para  $m_r$  que varían de 0,1 a 2,0 en incrementos de 0,1.

## **B. RCD**

La Fig. 11 es una tabla 1100 que muestra simulaciones por ordenador para el análisis RCD digital para una muestra de ADN fetal puro (100%) para  $m_r$  que varía de 0,1 a 2,0, de acuerdo con una realización de la presente invención. Según disminuye la concentración de ADN fetal fraccionaria, decrece el grado de sobrerrepresentación del cromosoma 21 y por tanto, se requiere un mayor número de pocillos informativos para una clasificación precisa de la enfermedad. Por tanto, se realizaron simulaciones adicionales para concentraciones de ADN fetal de un 50%, 25% y 10% para un

número total de pocillos que varían de 384 a 7680 pocillos con  $m_r=0,5$ .

La Fig. 12 es una tabla 1200 que muestra los resultados de la simulación informática de exactitudes de análisis de RCD digital con  $m_r=0,5$  para la clasificación de muestras de fetos euploides o con trisomía 21 con diferentes concentraciones fraccionarias de ADN fetal de acuerdo con una realización de la presente invención. El rendimiento del RCD digital es mejor para los casos con una concentración fraccionaria de ADN fetal más elevada. Con una concentración de ADN fetal de un 25% y un número total de 7680 análisis de PCR, un 97% de casos tanto euploides como aneuploides serían clasificables sin una clasificación incorrecta. El 3% restante de los casos requiere análisis adicionales hasta que se pueda lograr la clasificación.

Los procedimientos para simular análisis digitales RCD fueron similares a aquellos descritos para el análisis digital ARN-SNP. Las etapas para las simulaciones se describen a continuación:

En la etapa 1, se generaron dos números aleatorios bajo la función de densidad de probabilidad de Poisson para representar el locus de referencia, el cromosoma 1 y el locus del cromosoma 21. Para los sujetos femeninos portadores de fetos euploides, las concentraciones promedio de ambos loci del cromosoma 1 y del cromosoma 21 fueron las mismas. En este análisis de simulación, se utilizó una concentración promedio de molde de 0,5 para cada locus por pocillo. Para los sujetos femeninos portadores de fetos con trisomía 21, el  $m_r$  en esta simulación fue de 0,5, pero la concentración promedio del locus del cromosoma 21 por pocillo dependería de la concentración de ADN fetal fraccionario en la muestra analizada, como se muestra en la Tabla 700. La distribución del locus de referencia y/o del cromosoma 21 para un pocillo se determinó mediante los números aleatorios que representan el locus respectivo que se generaron de acuerdo con la función de densidad de probabilidad de Poisson con la concentración promedio adecuada del locus por pocillo.

En la etapa 2, se repitió el paso 1 hasta que se generó el número deseado de pocillos, por ejemplo, 384 pocillos para un experimento de placas de 384 pocillos. Se contó el número de pocillos positivos para el cromosoma 1 solo y el cromosoma 21 solo. El número de pocillos informativos fue el número total de pocillos positivos para uno de los cromosomas, pero no para ambos. La proporción de pocillos informativos positivos para el cromosoma 21 ( $P_r$ ) se calculó, a continuación. Los límites superior e inferior de las curvas de SPRT relevantes para aceptar la hipótesis nula o alternativa se calcularon como se describe en la sección anterior sobre el análisis de SPRT.

En la etapa 3, se realizaron 5000 simulaciones para cada uno de los dos escenarios de la mujer gestante que porta un feto euploide o con trisomía 21. Cada simulación podría considerarse como una muestra biológica independiente obtenida de mujeres gestantes. En la Tabla 1100, la clasificación correcta de casos euploides se refiere a aquellos casos euploides en los que se aceptó la hipótesis nula y la clasificación incorrecta de casos euploides se refiere a aquellos casos euploides en los que se aceptó la hipótesis alternativa. De forma análoga, aquellos casos de trisomía 21 en los que se aceptó la hipótesis alternativa se consideraron como correctamente clasificados y aquellos casos de trisomía 21 en los que se aceptó la hipótesis nula se consideraron incorrectamente clasificados. Para ambos grupos, aquellos casos en los que no se aceptó la hipótesis nula o la alternativa, después de que se hubiera simulado el número total preespecificado de pocillos, se consideraron sin clasificar.

En la etapa 4, se repitieron las etapas 1 a 3 para las muestras con 10%, 25%, 50% y 100% de ADN fetal con cifras totales de pocillos que varían de 384 a 7680.

## II. VALIDACIÓN DE LA DETECCIÓN DE TRISOMÍA 21

### A. ARN-SNP para *PLAC4*

La viabilidad práctica del ARN-SNP digital se demostró utilizando el SNP rs8130833 sobre el gen *PLAC4* en el cromosoma 21 (Lo, YMD *et al.* 2007 Nat Med 13, 218-223). Se analizaron muestras de ADN y ARN placentario de placentas heterocigóticas, dos euploides y dos con trisomía 21. Las muestras de ADN placentario se analizaron con el protocolo de ARN-SNP digital, pero con la omisión de la etapa de transcripción inversa, por tanto, convirtiendo el procedimiento esencialmente en un análisis de ADN-SNP digital. Para encontrar el equilibrio entre la posibilidad de una clasificación correcta de los casos y la proporción de pocillos informativos, los presentes inventores diluyeron las muestras con el objetivo de un alelo de cualquier tipo por pocillo y lo confirmaron mediante un análisis de PCR digital de 96 pocillos. A esto le siguió un experimento de ARN-SNP digital de 384 pocillos. Se calcularon  $P_r$  y  $m_r$  y la curva de SPRT para este valor de  $m_r$  se utilizó para la interpretación de los datos.

La Fig. 13A muestra una tabla 1300 de análisis de ARN-SNP digital en tejidos placentarios de gestaciones euploides y de trisomía 21. Los genotipos se determinaron mediante un ensayo de espectrometría de masas. "Euploide" se asignó cuando el  $P_r$  obtenido experimentalmente estuvo por debajo de la región inclasificable; "T21", que representa la trisomía 21, se asignó cuando la  $P_r$  obtenida experimentalmente estuvo por encima de la región inclasificable. T21, trisomía 21. Cada uno de estos casos se clasificó correctamente, utilizando muestras tanto de ADN como de ARN, con un experimento de 384 pocillos.

Los presentes inventores analizaron adicionalmente muestras de ARN plasmático de nueve mujeres portadoras de



fetos euploides y cuatro mujeres portadoras de fetos con trisomía 21. La Fig. 13B muestra una tabla 1350 de análisis de ARN-SNP digital de plasma materno de gestaciones de euploides y trisómicos 21. Todos los casos se clasificaron correctamente. Los resultados iniciales para un caso de trisomía 21 (M2272P) cayeron dentro de la región inclasificable, entre las curvas de SPRT, después de un experimento de 384 pocillos. Por tanto, se realizó una ejecución adicional de 384 pocillos. Se calcularon valores nuevos de  $m_r$  y  $P_r$  a partir de los datos agregados del total de 768 pocillos y la clasificación se realizó utilizando un nuevo conjunto de curvas de SPRT seleccionadas basándose en este valor de  $m_r$ . A continuación, el caso se puntuó correctamente como aneuploide.

Los datos experimentales y de simulación de los presentes investigadores muestran que el ARN-SNP digital es un método eficaz y exacto para la detección de la trisomía 21. Dado que el ARNm de *PLAC4* en el plasma materno se deriva puramente del feto, para 12 de las 13 muestras de plasma materno ensayadas, solo se necesitó uno experimento de PCR digital de 384 pocillos para una clasificación correcta. Este enfoque homogéneo, basado en PCR digital en tiempo real ofrece por tanto una alternativa al enfoque basado en espectrometría de masas para el análisis de ARN-SNP. Además de los transcritos de ARNm específicos de placenta, los presentes inventores también conciben que se podrían utilizar otros tipos de especies de ácidos nucleicos específicos fetales en plasma materno para la detección basada en PCR digital de aneuploidías cromosómicas fetales. Un ejemplo son los marcadores epigenéticos fetales que se han utilizado recientemente para la detección prenatal no invasiva de la trisomía 18 utilizando el enfoque de la relación alélica epigenética (EAR) (Tong YK et al. 2006 Clin Chem, 52, 2194-2202). Por tanto, los presentes inventores predicen que la EAR digital sería una técnica analítica posible.

## B. RCD

La viabilidad práctica del RCD digital para la detección de trisomía 21 se investigó también utilizando un ensayo de PCR dirigido a secuencias parálogas en los cromosomas 21 y 1. Los loci parálogos se utilizaron aquí a modo de ejemplos. Las secuencias no parálogas en el cromosoma 21 y cualquier otro cromosoma de referencia pueden utilizarse también para el RCD. Las muestras de ADN placentario de dos placentas euploides y dos de trisomía 21 se diluyeron hasta aproximadamente un molde diana de cada cromosoma por pocillo y se confirmaron mediante un análisis de PCR digital de 96 pocillos. Cada muestra confirmada se analizó mediante un experimento de RCD digital de 384 pocillos y se calcularon los valores de  $P_r$  y  $M_r$ . Para el RCD digital, el parálogo del cromosoma 1 fue el molde de referencia. Este valor de  $m_r$  se utilizó para seleccionar un conjunto correspondiente de curvas de SPRT para la interpretación de datos. Todas las muestras de placenta se clasificaron correctamente como se muestra en la Fig. 14A.

Para demostrar que el enfoque de RCD digital se puede utilizar para detectar ADN de trisomía 21 que se mezcla con un exceso de ADN euploide, tal como el escenario de ADN fetal en plasma materno, se analizaron las mezclas que contienen un 50 % y un 25 % de ADN placentario de trisomía 21 en un fondo de ADN de células sanguíneas maternas euploides. El ADN de la placenta de 10 casos de trisomía 21 y de 10 casos euploides se mezcló cada uno con una cantidad igual de ADN de células sanguíneas euploides maternas, produciendo por tanto veinte mezclas de ADN al 50 %. La Fig. 14B muestra un gráfico 1440 que ilustra la interpretación de SPRT para el análisis RCD de las mezclas de un 50 % de ADN fetal. De forma análoga, el ADN placentario de 5 casos de trisomía 21 y de 5 casos euploides se mezcló cada uno con un exceso 3 veces mayor de ADN de células sanguíneas euploides maternas, produciendo por tanto diez mezclas de ADN al 25 %. La Fig. 14C muestra un gráfico 1470 que ilustra la interpretación de SPRT para el análisis RCD de las mezclas de un 25 % de ADN fetal. Todas las mezclas de ADN euploide y aneuploide se clasificaron correctamente, como se muestra en las Figs. 14B y 14C.

Cada muestra alcanzó el punto de ser clasificable después de varios análisis de PCR digital de 384 pocillos, como se marca en las Figs. 14B y 14C. Para las mezclas de ADN al 50 %, el número de placas de 384 pocillos necesarios varió de uno a cinco. Para las mezclas de ADN al 25 %, el número de placas de 384 pocillos necesarios varió de uno a siete. La proporción acumulada de casos que se clasificaron correctamente con la adición de cada análisis de PCR digital de 384 fue como la predicha por la simulación por ordenador presentada en la Tabla 1200.

## III. MÉTODO CON PCR DIGITAL

### I. ARN-SNP DIGITAL

Todas las muestras de ARN se sometieron primero a transcripción inversa con un cebador de transcripción inversa específico de gen utilizando la transcriptasa inversa de ThermoScript (Invitrogen). La secuencia del cebador de transcripción inversa fue 5'-AGTATATAGAACCATGTTTAGGCCAGA-3' (Integrated DNATechnologies, Coralville, Iowa). El tratamiento posterior de las muestras de ARN transcrito de forma inversa (es decir, el ADNc) para el ARN-SNP digital y las muestras de ADN (por ejemplo, ADN placentario) fue el mismo, esencialmente. Antes del análisis de PCR digital, las muestras de ADN y de ADNc primero se cuantificaron, utilizando un ensayo de PCR en tiempo real para *PLAC4*, que consta de los cebadores 5'-CCGCTAGGGTGTCTTTTAAGC-3', 5'-GTGTTGCAATACAAATGAGTTTCT-3' y de la sonda fluorescente 5'-(FAM)ATTGGAGCAAATTC(MGBNFQ)-3' (Applied Biosystems, Foster City, California), en donde FAM es la 6-carboxifluoresceína y MGBNFQ es un desactivador no fluorescente de unión a surco menor.

Se preparó una curva de calibración mediante diluciones en serie de oligonucleótidos de ADN sintético monocatenario purificados por HPLC (Proligo, Singapur) especificando el amplicón. La secuencia fue 5'-CGCCGCTAGGGTGTCTTTTAAAGCTATTGGAGCAAATTCAAATTTGGCTTAAAGAA

AAAGAAACTCATTTTGTATTGCAACACCAGGAGTATCCCAAGGGACTCG-3'. La reacción se estableció utilizando 2X TaqMan Universal PCR Master Mix (Applied Biosystems) en un volumen de reacción de 25 µl. Se utilizaron 400 nM de cada cebador y 80 nM de la sonda en cada reacción. La reacción se inició a 50 °C durante 2 min, seguido de 95 °C durante 10 min y 45 ciclos de 95 °C durante 15 seg y 60 °C durante 1 min en un sistema de detección de secuencias ABI PRISM 7900HT (Applied Biosystems). A continuación, se llevaron a cabo diluciones en serie de las muestras de ADN o de ADNc de manera que la amplificación posterior por PCR digital se pudiera realizar con aproximadamente una molécula molde por pocillo. A tal concentración, se esperaba que aproximadamente el 37 % de los pocillos de reacción mostraran amplificación negativa y se confirmó primero realizando un análisis de PCR en tiempo real digital de 96 pocillos. A esto le siguió un análisis de ARN-SNP digital realizado en placas de 384 pocillos utilizando un conjunto de cebadores que no abarcan intrones: el cebador directo 5'-TTTGTATTGCAACACCATTGG-3' y el cebador de transcripción inversa específico del gen descrito anteriormente.

Se diseñaron dos sondas TaqMan específicas de alelo dirigidas a cada uno de los dos alelos del SNP rs8130833 en el *PLAC4*. Sus secuencias fueron 5'-(FAM)TCGTCGTCTAACTTG(MGBNFQ)-3' y 5'-(VIC)ATTCGTCATCTAACTTG(MGBNFQ) para los alelos G y A, respectivamente. La reacción se estableció utilizando 2X TaqMan Universal PCR Master Mix en un volumen de reacción de 5 µl. Cada reacción contiene 1X TaqMan Universal PCR Master Mix, 572 nM de cada cebador, 107 nM de la sonda específica del alelo G y 357 nM de la sonda específica del alelo A. La reacción se llevó a cabo en el sistema de detección de secuencias ABI PRISM 7900HT. La reacción se inició a 50 °C durante 2 min, seguido de 95 °C durante 10 min y 45 ciclos de 95 °C durante 15 seg y 57 °C durante 1 min. Durante la reacción, los datos de fluorescencia se recogieron mediante la aplicación "Absolute Quantification" del software SDS 2.2.2 (Applied Biosystems). El software calculó automáticamente los valores basales y umbrales. El número de pocillos que fueron positivos para los alelos A o G se registró y se sometió a análisis de SPRT.

## B. Análisis de RCD digital

Todas las muestras de ADN placentario y de capa leucocitaria materna utilizadas en este estudio se cuantificaron primero mediante el espectrofotómetro de NanoDrop (NanoDrop Technology, Wilmington, Delaware). La concentración de ADN se convierte en copias/ml utilizando una conversión de 6,6 pg/célula. La cantidad de ADN correspondiente a aproximadamente un molde por pocillo se determinó diluyendo en serie las muestras de ADN y se confirmó con el ensayo de PCR en tiempo real en un formato de 96 pocillos en donde los presentes investigadores esperan que aproximadamente el 37 % de los pocillos muestre una amplificación negativa. La configuración de la PCR para la placa confirmatoria fue la misma, como se describe a continuación, excepto que solo se agregó la sonda para el cromosoma de referencia. En el análisis de RCD digital, los loci parálogos en el cromosoma 21 y 1 (Deutsch, S. *et al.* 2004 J Med Genet 41, 908-915) se coamplificaron primero mediante el cebador directo 5'-GTTGTTCTGCAAAAAACCTTCGA-3' y el cebador inverso 5'-CTTGGCCAGAAATACTTCATTACCATAT-3'. Se diseñaron dos sondas TaqMan específicas de cromosomas para dirigirse a los parálogos del cromosoma 21 y 1 y sus secuencias fueron 5'-(FAM)TACCTCCAT AATGAGTAAA(MGBNFQ)-3' y 5'-(VIC)CGTACCTCTGTAATGTGTAA(MGBNFQ)-3', respectivamente. Cada reacción contiene 1X TaqMan Universal PCR Master Mix (Applied Biosystems), 450 nM de cada cebador y 125 nM de cada sonda. El volumen total de reacción fue de 5 µl/pocillo. La reacción se inició a 50 °C durante 2 min, seguido de 95 °C durante 10 min y 50 ciclos de 95 °C durante 15 seg y 60 °C durante 1 min. Todos los experimentos de PCR en tiempo real se llevaron a cabo en un sistema de detección de secuencias ABI PRISM 7900HT (Applied Biosystems) y los datos de fluorescencia se recogieron mediante la aplicación "Absolute Quantification" del software SDS 2.2.2 (Applied Biosystems). Se utilizaron los valores basales por defecto y los valores umbral manuales. Se registró el número de pocillos que fueron positivos para el cromosoma 21 o el cromosoma 1 y se sometió a análisis de SPRT. Se analizarían una o más placas de 384 pocillos hasta que se pudiera clasificar la enfermedad mediante la SPRT.

## IV. UTILIZACIÓN DE PCR DIGITAL BASADA EN MICROFLUIDOS

### I. ARN-SNP DIGITAL

Este ejemplo demuestra el rendimiento del análisis de PCR digital utilizando PCR digital basada en microfluidos. Una variante de este enfoque se ilustra aquí, a modo de ejemplo, aunque no de forma limitativa, utilizando un sistema Fluidigm BioMark™. Este sistema puede realizar más de 9000 PCR digitales por ejecución.

Se obtuvieron muestras de tejido placentario y de sangre periférica materna de mujeres gestantes portadoras de fetos euploides o trisómicos 21. El genotipado del SNP rs8130833 en el gen *PLAC4* en muestras de ADN placentario mediante la prolongación del cebador seguido de espectrometría de masas. Se extrajo ARN de las muestras de placenta y de plasma materno.

Todas las muestras de ARN se sometieron a transcripción inversa con un cebador de transcripción inversa específico de gen (5-AGTATATAGAACCATGTTTAGGCCAGA-3') utilizando la transcriptasa inversa ThermoScript (Invitrogen).

Para las muestras de ADNc placentario, se realizaron diluciones en serie de manera que la amplificación posterior por PCR digital podría realizarse en aproximadamente una molécula molde por pocillo.

La PCR digital se realizó en el sistema BioMark™ (Fluidigm) con una matriz digital 12.765 (Fluidigm). Cada matriz digital consta de 12 paneles para alojar 12 mezclas de ensayo de muestras. Cada panel se divide adicionalmente en 765 pocillos para llevar a cabo una reacción de 7-nL/pocillo. La región de SNP rs8130833 en el gen *PLAC4* se amplificó mediante el cebador directo (5'-TTTGTATTGCAACACCATTTGG-3') y el cebador de transcripción inversa específico del gen descrito anteriormente. Se diseñaron dos sondas TaqMan específicas de alelo dirigidas a cada uno de los dos alelos del SNP rs8130833. Sus secuencias fueron 5'-(FAM)TCGTCGTCTAACTTG(MGBNFQ)-3' y 5'-(VIC)ATTCGTCATCTAACTTG(MGBNFQ) para los alelos G y A, respectivamente. La reacción para un panel de matriz se estableció utilizando 2X TaqMan Universal PCR Master Mix en un volumen de reacción de 10 µl. Cada reacción contiene 1X TaqMan Universal PCR Master Mix, 572 nM de cada cebador, 53,5 nM de la sonda específica del alelo G, 178,5 nM de la sonda específica del alelo A y 3,5 µl de la muestra de ADNc. Se utilizó un panel de reacción para cada muestra de ADNc placentario, si bien se utilizaron 12 paneles para cada muestra de plasma materno. Las mezclas de ensayo de muestra se cargaron en el Digital Array mediante un Controlador IFC de NanoFlex™ (Fluidigm). La reacción se llevó a cabo en el sistema BioMark™. La reacción se inició a 50 °C durante 2 min, seguido de 95 °C durante 10 min y 40 ciclos de 95 °C durante 15 seg y 57 °C durante 1 min.

Se analizaron muestras de ARN placentario de una placenta euploide y dos placentas heterocigóticas T21 en un panel de reacción de 765 pocillos. Para cada muestra, se contó el número de pocillos informativos, que comprenden los positivos para el alelo A o para el G (pero no para ambos). Se determinó la proporción del alelo sobrerrepresentado entre todos los pocillos informativos (Pr). Las curvas de SPRT adecuadas para la concentración de molde de referencia promedio exacta por pocillo ( $m_r$ ) de estas ejecuciones se aplicaron para determinar si el Pr obtenido experimentalmente indicaba una muestra euploide o T21. Como se muestra en la Fig. 15A, todas las muestras de ARN se clasificaron correctamente utilizando este enfoque.

Los presentes inventores analizaron adicionalmente muestras de ARN plasmático de cuatro mujeres portadoras de fetos euploides y una mujer portadora de un feto con trisomía 21. Cada muestra se analizó en doce paneles de reacción de 765 pocillos, es decir, 9180 reacciones por muestra de ARN de plasma. La Fig. 15B muestra el número de pocillos informativos para cada uno de los 12 paneles para esta muestra de ARN de plasma. Como se muestra en la tabla, la concentración de molde en la muestra de plasma estaba tan diluida que el número de pocillos informativos en cualquier panel de reacción no era suficiente para la clasificación SPRT. Los pocillos informativos de tres paneles de reacción tuvieron que combinarse antes de que esta muestra se clasificara como muestra euploide (Fig. 15C). La Fig. 15C muestra que utilizando los datos agregados de dos a doce paneles, podrían clasificarse correctamente todos los casos de plasma.

En comparación con el método manual para realizar la PCR digital, este método basado en microfluidos es mucho más rápido y requiere menos mano de obra. El proceso completo podría completarse en dos horas y media.

## B. Análisis de ARN-SNP digital para la detección prenatal de trisomía 18

En este ejemplo, Los presentes inventores utilizaron un ensayo de discriminación alélica digital basado en PCR sobre el inhibidor del ARNm de *serpina peptidasa clado B (ovalbúmina) miembro 2 (SERPINB2)*, un transcrito expresado por la placenta en el cromosoma 18, para detectar un desequilibrio en la proporción de alelos polimórficos para fetos con trisomía 18. La extracción de ADN y ARN de muestras de tejido placentario se realizó utilizando el mini Kit de ADN QIAamp (Qiagen, Hilden, Alemania) y el reactivo TRIzol (Invitrogen, Carlsbad, California), respectivamente, como se describe en los protocolos de los fabricantes. Las muestras de ARN placentario extraídas se sometieron a tratamiento con ADNasa I (Invitrogen) para eliminar el ADN genómico contaminante. La genotipado del SNP rs6098 en el gen *SERPINB2* se llevó a cabo en muestras de ADN de tejido placentario con un ensayo Homogenous MassEXTEND (hME) utilizando MassARRAY Compact (Sequenom, San Diego) como se describió anteriormente.

La transcripción inversa para el transcrito de *SERPINB2* se realizó en las muestras de ARN de tejido placentario con un cebador específico de gen 5'-CGCAGACTTCTACCAAACA-3' (Integrated DNA Technologies, Coralville, Iowa) utilizando la transcriptasa inversa ThermoScript (Invitrogen). Todas las muestras de ADNc se diluyeron hasta una concentración de manera que la amplificación posterior por PCR digital podría realizarse en una concentración promedio de una molécula molde por pocillo. La PCR digital se estableció utilizando la TaqMan Universal PCR Master Mix (Applied Biosystems, Foster City, California) y los reactivos Biomark™ PCR (Fluidigm, San Francisco). El cebador directo 5'-CTCAGCTCTGCAATCAATGC-3' (Integrated DNA Technologies) y el cebador inverso (idéntico al cebador específico del gen utilizado para la transcripción inversa) se utilizaron a una concentración de 600 nM. Las dos sondas TaqMan dirigidas al alelo A o G del SNP rs6098 en la secuencia *SERPINB2* fueron 5'-(FAM)CCACAGGGAATTATTT(MGBNFQ)-3' y 5'-(FAM)CCACAGGGGATTATTT(MGBNFQ)-3' (Applied Biosystems). FAM es la 6-carboxifluoresceína y MGBNFQ es un desactivador no fluorescente de unión al surco menor y se utilizaron en concentraciones de 300 nM y 500 nM, respectivamente. Cada mezcla de muestra-reactivo se dividió en 765 pocillos de reacción en una matriz digital 12.765 de BioMark™ utilizando el controlador IFC de NanoFlex™ (Fluidigm). Tras la partición, la matriz se colocó en el sistema de PCR en tiempo real de BioMark™ (Fluidigm) para la amplificación térmica. La reacción se inició a 50 °C durante 2 min y continuó a 95 °C durante 5 min seguido de 45 ciclos de 95 °C

durante 15 seg y 59 °C durante 1 min. Después de la amplificación, se contó el número de pocillos informativos (uno que fue positivo solo para el alelo A o para el G) y el número de pocillos positivos para ambos alelos y se sometió a un análisis de prueba de la proporción de probabilidades secuencial (SPRT).

- 5 Para un feto euploide heterocigótico, los alelos A y G deben estar igualmente representados (1: 1) en el genoma fetal, si bien para la trisomía 18, habría una copia adicional de un alelo, dando por tanto una proporción de 2:1 en el genoma fetal. Se generó una serie de curvas de SPRT para la interpretación de diferentes muestras. Estas curvas ilustran la proporción esperada de pocillos informativos positivos para el alelo P<sub>r</sub> sobrerrepresentado (eje y) para un número total  
10 dado de pocillos informativos (eje x) necesarios para la clasificación. Para cada muestra, la P<sub>r</sub> derivada experimentalmente se comparó con el valor de P<sub>r</sub> esperado. Las muestras por encima de la curva superior se clasificaron como trisomía 18, si bien aquellas por debajo de la curva inferior se clasificaron como euploides. El área entre las dos curvas es la región inclasificable.

- 15 La viabilidad del análisis digital de ARN-SNP para la detección de trisomía 18 fetal se demostró utilizando el SNP rs6098 en el gen SERPINB2. Las muestras de ADN de tejido placentario de sujetos con fetos euploides y trisomía 18 primero se genotiparon mediante espectrometría de masas para identificar los casos heterocigóticos. Se encontraron nueve placentas heterocigóticas euploides y tres con trisomía 18 y se sometieron a análisis de ARN-SNP digital. Para cada muestra, se calcularon P<sub>r</sub> y m<sub>r</sub> y la curva de SPRT para este valor de m<sub>r</sub> se utilizó para la clasificación de enfermedades. Como se muestra en la Fig. 16A, todas las muestras se clasificaron correctamente. Los valores de P<sub>r</sub>  
20 de las placentas con trisomía 18 estuvieron por encima de la región inclasificable, si bien las placentas euploides cayeron por debajo de esta región.

- 25 Las muestras con curvas de SPRT basadas en m<sub>r</sub>= 0,1, 0,2 y 0,3 se ilustran en la Fig. 16B. Estos datos sugieren que el método ARN-SNP digital es una herramienta de diagnóstico valiosa para gestaciones con trisomía 18. Las dos curvas representan los límites de la región inclasificable. Las muestras con puntos de datos por encima de la curva superior se clasificaron como aneuploides, si bien aquellas con puntos de datos por debajo de la curva inferior se clasificaron como euploides.

### C. Análisis de RCD digital

- 30 Este ejemplo demuestra el rendimiento del análisis de RCD digital utilizando PCR digital basada en microfluidos. Una variante de este enfoque se ilustra aquí, a modo de ejemplo, aunque no de forma limitativa, utilizando un sistema Fluidigm BioMark™. Este sistema puede realizar más de 9000 PCR digitales por ejecución.

- 35 Se obtuvieron muestras de tejido placentario y de sangre periférica materna de mujeres gestantes portadoras de fetos euploides o trisómicos 21 (T21). El ADN de la placenta de 10 casos de T21 y de 10 casos euploides se mezcló cada uno con una cantidad igual de ADN de células sanguíneas euploides maternas, produciendo por tanto veinte mezclas de ADN al 50 %. Para asegurar una proporción fetal precisa en las muestras de mezcla, el ADN extraído se cuantificó primero mediante la medición de la densidad óptica (DO) a 260 nm. A continuación, se cuantificaron digitalmente  
40 mediante el sistema BioMark™ (Fluidigm) utilizando las matrices digitales 12.765 (Fluidigm). El ensayo para cuantificar las muestras fue el mismo que se describe a continuación, con la excepción de que solo se utilizó la sonda para el cromosoma de referencia.

- 45 La dosis de cromosomas en las mezclas de ADN al 50 % y las muestras de líquido amniótico se determinaron mediante análisis de PCR digital de un locus del cromosoma 21 no polimórfico relacionado con uno situado en el cromosoma 1. Un amplicón de 101 pb de un par de loci parálogos en el cromosoma 21 y 1 se coamplificó primero mediante el cebador directo 5'-GTTGTTCTGCAAAAACCTTCGA-3' y el cebador inverso 5'-CTTGGCCAGAAATACTTCATTACCATAT-3'. Se diseñaron dos sondas TaqMan específicas de cromosomas para distinguir a los parálogos del cromosoma 21 y 1 y sus secuencias fueron 5'-(FAM)TACCTCCAT AATGAGTAAA(MGBNFQ)-3' y 5'-(VIC)CGTACCTCTGTAATGTGTAA(MGBNFQ)-3', respectivamente. Los loci parálogos se utilizaron aquí solo a modo de ejemplo. En otras palabras, los loci no parálogos también podrían utilizarse para tal análisis.

- 50 Con el fin de demostrar el uso del enfoque de RCD digital para detectar trisomía 18 (T18), se diseñó otro ensayo dirigido a secuencias parálogas en los cromosomas 21 y 18. Se coamplificó primero un amplicón de 128 pb de los loci parálogos en los cromosomas 21 y 18 mediante el cebador directo 5'-GTTGTTCTGCAAAAACCTTCGA-3' y el cebador inverso 5'-CTTGGCCAGAAATACTTCATTACCATAT-3'. Se diseñaron dos sondas TaqMan específicas de cromosomas para distinguir a los parálogos del cromosoma 21 y 18 y sus secuencias fueron 5'-(FAM)TACCTCCAT AATGAGTAAA(MGBNFQ)-3' y 5'-(VIC)CGTACCTCTGTAATGTGTAA(MGBNFQ)-3', respectivamente. Los loci parálogos se utilizaron aquí solo a modo de ejemplo. En otras palabras, los loci no parálogos también podrían utilizarse para tal análisis.  
60

- 65 Todos los experimentos se llevaron a cabo en el sistema BioMark™ (Fluidigm) utilizando las matrices digitales 12.765 (Fluidigm). La reacción para un panel se estableció utilizando 2X TaqMan Universal PCR Master Mix en un volumen de reacción de 10 ml. Cada reacción contenía un 1X TaqMan Universal PCR Master Mix, 900 nM de cada cebador, 125 nM de cada sonda y 3,5 ml de una muestra de ADN de células de sangre materna/placentaria al 50 %. La mezcla de muestra/ensayo se cargó en la matriz digital (Digital Array) mediante un Controlador IFC de NanoFlex™ (Fluidigm).

La reacción se llevó a cabo en el sistema BioMark™ para la detección. La reacción se inició a 50 °C durante 2 min, seguido de 95 °C durante 10 min y 40 ciclos de 95 °C durante 15 seg y 57 °C durante 1 min.

Las muestras de un 50 % de ADN de células sanguíneas placentarias/maternas de euploides y de T21 se analizaron en las matrices digitales en el ensayo de cr21/cr1 (cr, cromosoma). Para cada muestra, se contó el número de pocillos informativos, que comprenden los positivos para los marcadores de cr21 o cr1 (pero no ambos). Se determinó la proporción del marcador sobrerrepresentado entre todos los pocillos informativos ( $P_r$ ). Las curvas de SPRT adecuadas para la concentración de molde de referencia promedio exacta por pocillo ( $m_r$ ) para cualquiera de los paneles de PCR digital se aplicaron para determinar si el  $P_r$  obtenido experimentalmente indicaba una muestra euploide o T21. Los datos se agregaron a partir de paneles extra para muestras que permanecieron sin clasificar hasta que se pudiera tomar una decisión. Como se muestra en la Fig. 17, todas las muestras de un 50 % de ADN de células sanguíneas placentarias/maternas se clasificaron correctamente utilizando este enfoque con datos que varían de uno a cuatro paneles necesarios. También se representó una curva de SPRT para mostrar los límites de decisión para una clasificación correcta, tal como se muestra en la FIG. 18.

Los presentes inventores aplicamos adicionalmente el análisis RCD en muestras de líquido amniótico obtenidas de 23 mujeres que portan de fetos euploides y 6 mujeres que portan fetos T21. Cada muestra se analizó en un solo panel de reacción de 765 pocillos con el ensayo cr21/cr1. La Fig. 19 muestra el sumario de la clasificación de SPRT. Como se muestra en la Fig. 19, las 29 muestras se clasificaron correctamente. El método de RCD digital es por tanto un enfoque alternativo para la detección de trisomías utilizando microsatélites (Levett LJ, *et al.* A large-scale evaluation of amnio-PCR for the rapid prenatal diagnosis of fetal trisomy. *Ultrasound Obstet Gynecol* 2001; 17: 115-8) o de polimorfismos de un solo nucleótido (SNP) (Tsui NB, *et al.* Detection of trisomy 21 by quantitative mass spectrometric analysis of single-nucleotide polymorphisms. *Clin Chem* 2005; 51: 2358-62) marcadores o PCR en tiempo real no digital (Zimmermann B, *et al.* Novel real-time quantitative PCR test for trisomy 21. *Clin Chem* 2002; 48: 362-3) en diversos tipos de muestra utilizadas para el diagnóstico prenatal, tal como biopsias de líquido amniótico y vellosidades coriónicas.

En un intento por detectar casos de T18, los presentes investigadores aplicamos el ensayo cr21/cr18 en 3 muestras de ADN placentario euploide y 5 de T18. Se calculó la proporción del marcador sobrerrepresentado entre todos los pocillos informativos ( $P_r$ ). Todos se clasificaron correctamente, con la excepción de un caso de T18 que se clasificó erróneamente como euploide. Los resultados se resumen en la Fig. 20.

#### V. UTILIZACIÓN DE ENSAYOS DE RCD DIGITAL MULTIPLEXADOS EN PLATAFORMA DE ESPECTROMÉTRICA DE MASAS

El número de moléculas de ADN plasmático que están presentes por unidad de volumen de plasma materno es limitado (Lo, YMD. *et al.* 1998 *Am J Hum Genet* 62, 768-7758). Por ejemplo, en la gestación temprana, la concentración plasmática materna promedio de un locus autosómico, el gen  $\beta$ -globina, se ha demostrado que es de 986 copias/ml, con contribuciones tanto del feto como de la madre (Lo, YMD. *et al.* 1998 *Am J Hum Genet* 62, 768-7758). Para capturar 7.680 moléculas, se necesitaría ADN extraído de unos 8 ml de plasma materno. Este volumen de plasma, obtenible de ~15 ml de sangre materna, está en el límite de la práctica habitual. Sin embargo, los presentes investigadores prevén que se pueden combinar múltiples conjuntos diana de Cr21 y cromosomas de referencia para el análisis digital de RCD. Para 5 pares de cr21 y cromosomas de referencia diana, solo se necesitarían 1,6 ml de plasma materno para proporcionar el número de moléculas molde necesarias para el análisis. Podría realizarse una PCR multiplexada de una sola molécula. La solidez de dicho análisis de una sola molécula multiplexado se ha demostrado previamente para el haplotipado de una sola molécula (Ding, C. y Cantor, CR. 2003 *Proc Natl Acad Sci USA* 100, 7449-7453).

En un ejemplo, se obtuvieron muestras de tejido placentario y de sangre periférica materna de mujeres gestantes que portan fetos euploides o trisómicos 21 (T21). Se mezclaron 5 muestras de ADN placentario euploide y 5 de T21 cada una con proporciones iguales de ADN de células de sangre materna para producir 10 mezclas de ADN que imitan muestras de plasma con un 50 % de ADN fetal. Para asegurar una proporción fetal precisa en las muestras de mezcla, el ADN extraído se cuantificó primero mediante la medición de la densidad óptica (DO) a 260 nm. A continuación, se cuantificaron digitalmente mediante PCR en tiempo real en el formato de 384 pocillos. El ensayo para cuantificar las muestras fue el mismo que se ha descrito en el ejemplo previo de análisis de RCD digital.

La dosis de cromosomas en las mezclas al 50 % se determinó mediante análisis de PCR digital de un locus del cromosoma 21 no polimórfico relacionado con uno situado en el cromosoma 1. El método se denomina análisis de dosis cromosómica relativa (RCD) digital. Un amplicón de 121 pb (incluido un 10-mer en cada cebador) de un par de loci parálogos en el cromosoma 21 y el 1 se coamplificó mediante el cebador directo 5'-GTTGTTCTGCAAAAACCTTCGA-3' y el cebador inverso 5'-CTTGGCCAGAAATACTTCATTACCATAT-3'. Se diseñó un cebador de extensión que se dirige a las diferencias de bases entre el cromosoma 21 y el cromosoma 1 y su secuencia es 5'-CTCATCCTCACTTCGTACCTC-3'.

Con el fin de demostrar la utilidad de los ensayos de PCR digital multiplexadas para detectar casos de T21, se diseñó otro ensayo de RCD digital dirigido a secuencias parálogas en los cromosomas 21 y 18. Un amplicón de 148 pb

(incluido un 10-mer en cada cebador) de los loci parálogos en los cromosomas 21 y 18 fue coamplificado por el cebador directo 5'-ACGTTGGATGGTACAGAAACCACAACTGATCGG-3' y el cebador inverso 5'-ACGTTGGATGGTCCAGGCTGTGGGCCT-3'. Se diseñó un cebador de extensión que se dirige a las diferencias de bases entre el cromosoma 21 y el cromosoma 18 y su secuencia es 5'-ACAAAAGGGGGAAGAGG-3'.

El análisis RCD digital multiplexado se realizó utilizando el protocolo de prolongación del cebador. La reacción de PCR se estableció utilizando un kit de reactivos GeneAmp PCR Core Reagent Kit (Applied Biosystems) en un volumen de reacción de 5 ml. Cada reacción contuvo tampón 1X, MgCl<sub>2</sub> 2 mM, mezcla de dNTP 200 mM, AmpliTaq Gold 0,2 U, 200 nM de cada uno de los 4 cebadores y la mezcla de ADN al 50 %. La mezcla de ensayo/muestra se dispensó en una placa de PCR de 384 pocillos y la reacción se inició a 50 °C durante 2 min, seguido de 95 °C durante 10 min y 40 ciclos de 95 °C durante 15 seg y 57 °C durante 1 min.

Los productos de PCR se sometieron a un tratamiento de fosfatasa alcalina de camarón (SAP; del inglés, shrimp alkaline phosphatase) para eliminar los dNTP no incorporados. La mezcla se incubó a 37 °C durante 40 min seguido de 85 °C durante 5 min. A continuación, se llevó a cabo la reacción de prolongación del cebador. En síntesis, se añadieron 771 nM del cebador de extensión del ensayo cr21/cr1, 1,54 mM de cebador de extensión del ensayo cr21/cr18, 0,67 U de Thermosequenase (Sequenom) y 64 mM de cada uno de ddCTP, ddGTP, dATP y dTTP en un cóctel de extensión a los productos de PCR tratados con SAP. Las condiciones de reacción fueron 94 °C durante 2 min, seguido de 94 °C durante 5 min, 50 °C durante 5 seg y 72 °C durante 5 seg durante 80 ciclos. Se agregaron 16 ml de agua y 3 mg de la resina limpia Clean Resin (Sequenom) a los productos de extensión para una limpieza final. Las mezclas se mezclaron en un rotador durante 20 a 30 min, seguido de una etapa de centrifugación a 361 g durante 5 min. Se dispensaron de quince a 25 nL de los productos finales en un SpectroCHIP (Sequenom) mediante un nanodispensador MassARRAY Nanodispenser S (Sequenom). La adquisición de datos del SpectroCHIP se realizó en el espectrómetro de masas compacto MassARRAY Analyzer (Sequenom). Los datos de masas se importaron al software MassARRAY Typer (Sequenom) para su análisis.

Las muestras de un 50 % de ADN placentario/materno, cinco euploides y cinco T21, se analizaron con el ensayo dúplex de RCD. Para cada muestra, se contó el número de pocillos informativos del ensayo individual, que comprenden los positivos para los marcadores solo de cr21 o de cr1 o de cr18. Se calculó separadamente para cada ensayo de RCD, la proporción del marcador de cr21 entre todos los pocillos informativos ( $P_r$ ). A continuación, se aplicó la prueba de la proporción de probabilidades secuencial (SPRT) para determinar si la  $P_r$  indicó una muestra euploide o T21. De este modo, se redujo el número de pocillos necesarios dado que cada placa se contó dos veces.

El ensayo cr21/cr1 normalmente se aplicó primero. Si la muestra permaneció sin clasificar, a continuación, los valores del ensayo cr21/cr18 se añadirían para cálculos adicionales. Se utilizaron paneles extra para muestras que permanecieron sin clasificar hasta que se pudiera tomar una decisión. Como se muestra en la Fig. 21, todas las muestras de mezcla de un 50 % de euploides se clasificaron correctamente utilizando una única placa de 384 pocillos. Varios casos de T21 requirieron 2 o más placas para una clasificación correcta. Si solo se utilizara un ensayo, se necesitaría un mayor número de placas para lograr el número de pocillos informativos requeridos cuando se alcanzó la clasificación. Por ejemplo, los datos para el caso N0230 fueron inclasificables cuando se usó individualmente cualquiera de los ensayos RCD. Sin embargo, se logró una clasificación correcta cuando se combinaron los datos de los dos ensayos. Si no se utilizaran los ensayos dúplex de RCD, se necesitarían placas de análisis adicionales. Los presentes inventores esperarían una reducción adicional del número de pocillos con un nivel más elevado de multiplexación de los ensayos.

En otro ejemplo, los presentes inventores desarrollaron un ensayo 4-plex dirigido a 4 amplicones diferentes en el cromosoma 21 y sus correspondientes parejas parálogas situadas en otros autosomas no del cromosoma 21. Este ensayo 4-plex se utilizó en el análisis de RCD digital, seguido de la clasificación de SPRT de muestras de gestaciones de euploides y de trisómicos 21. Se realizaron extracciones de ADN de muestras de placenta utilizando el kit de tejido QIAamp (Qiagen, Hilden, Alemania).

Todas las muestras de ADN placentario y de capa leucocitaria materna utilizadas en este estudio se cuantificaron primero mediante el espectrofotómetro de NanoDrop (NanoDrop Technology, Wilmington, Delaware). La concentración de ADN se convirtió en equivalentes de genoma (GE)/ml utilizando una conversión de 6,6 pg/célula. La cantidad de ADN correspondiente a aproximadamente un molde por pocillo se determinó diluyendo en serie las muestras de ADN. Bajo tal condición, los presentes investigadores esperaríamos que aproximadamente un 37 % de los pocillos mostraran amplificación negativa. En el análisis de RCD digital multiplexado, se seleccionaron 4 conjuntos de secuencias parálogas diana: los loci parálogos en el cromosoma 21 y 1 se coamplificaron por el cebador directo 5'-ACGTTGGATGTTGATGAAGTCTCATCTCTACTTCG -3' y el cebador inverso 5'-ACGTTGGATGCAATAAGCTTGGCCAGAAATACT -3', dando como resultado un amplicón de 81 pb. Los loci parálogos en el cromosoma 21 y 7 se coamplificaron por el cebador directo 5'-ACGTTGGATGGAATTTAAGCTAAATCAGCCTGAACTG-3' y el cebador inverso 5'-ACGTTGGATGGTTTCTCATAGTTCATCGTAGGCTTAT-3', dando como resultado un amplicón de 82 pb. Los loci parálogos en el cromosoma 21 y 2 se coamplificaron por el cebador directo 5'-ACGTTGGATGTCAGGCAGGGTTCTATGCAG -3' y el cebador inverso 5'-ACGTTGGATGAGGCGGCTTCCTGGCTCTTG -3', dando como resultado un amplicón de 101 pb. Los loci parálogos

en el cromosoma 21 y 6 se coamplificaron por el cebador directo 5'-ACGTTGGATGGCTCGTCTCAGGCTCGTAGTT-3' y el cebador inverso 5'-ACGTTGGATGTTTCTTCGAGCCCTTCTTGG-3', dando como resultado un amplicón de 102 pb. Cada reacción contenía 10X tampón II (Applied Biosystems),  $MgCl_2$  y 100 nM de cada cebador. El volumen total de reacción fue de 5 ml/pocillo. La reacción se inició a 95 °C durante 5 min, seguido de 45 ciclos a 95 °C durante 30 seg, 62 °C durante 30 seg y 72 °C durante 30 seg y una extensión final a 72 °C durante 7 min. Todas las amplificaciones de PCR convencionales se llevaron a cabo en un sistema GeneAmp PCR System 9700 (Applied Biosystems). Los nucleótidos no incorporados se desactivaron mediante tratamiento con fosfatasa alcalina de camarón (SAP). Cada reacción contuvo tampón 10X SAP (Sequenom) y enzima SAP (Sequenom). Se agregaron 2 ml de mezcla SAP a cada PCR. La reacción con SAP se incubó a 37 °C durante 40 min y 85 °C durante 5 min. Tras el tratamiento de SAP, se llevó a cabo la reacción de prolongación del cebador en los productos de PCR utilizando el kit iPLEX Gold (Sequenom). Los desajustes de secuencia parálogos (PSM) en los loci parálogos en los cromosomas 21 y 1 se examinaron por el cebador de extensión 5'-GTCTCATCTCTACTTCGTACCTC-3'. Los PSM en los loci parálogos en los cromosomas 21 y 7 se examinaron por el cebador de extensión 5'-TTTTACGCTGTCCCCATTT-3'. Los PSM en los loci parálogos en los cromosomas 21 y 2 se examinaron por el cebador de extensión 5'-GGTCTATGCAGGAGCCGAC-3'. Los PSM en los loci parálogos en los cromosomas 21 y 6 se examinaron por el cebador de extensión 5'-TGGGCGCGGGAGCGGACTTCGCTGG-3'. Cada reacción contuvo 10X tampón de reacción iPLEX (Sequenom), mezcla de terminación iPLEX (Sequenom), enzima iPLEX (Sequenom) y 343 nM de cada cebador de extensión, excepto para el cebador de extensión para PSM en los cromosomas 21 y 6, que se utilizaron en 1,03 mM. Se agregaron 2 ml de mezcla iPLEX a 5 ml de producto de la PCR. La reacción iPLEX se cicló de acuerdo con un programa de 200 ciclos cortos. Brevemente, las muestras se desnaturalizaron primero a 94 °C durante 35 seg, seguido de hibridación a 52 °C durante 5 seg y prolongación a 80 °C durante 5 seg. El ciclo de hibridación y prolongación se repitió cuatro veces más durante un total de cinco ciclos y a continuación, se volvió a un ciclo de desnaturalización a 94 °C durante 5 seg, tras lo cual se volvió a realizar el ciclo cerrado de hibridación y prolongación de 5 ciclos. Los cinco ciclos de hibridación y prolongación con la única etapa de desnaturalización se repitieron 39 veces para un total de 40. Se realizó una prolongación final a 72 °C durante 3 min. Los productos de la reacción iPLEX se diluyeron con 16 ml de agua y se desalaron con 6 mg de resina para cada PCR. La placa de 384 pocillos se centrifugó a 1600 g durante 3 min antes de dispensarla en el SpectroCHIP (Sequenom) y el análisis de MS de espectrometría de masas (Sequenom) de desorción e ionización por láser asistida por matriz-tiempo de vuelo (MALDI-TOF).

El número de pocillos que fueron positivos solo para el cromosoma 21 o solo para el cromosoma de referencia para cada uno de los cuatro ensayos se registró independientemente. Para cada ensayo, se calcularon los números de moléculas corregidos por Poisson para el cromosoma 21 y el cromosoma de referencia. La suma del número corregido por Poisson de moléculas para el cromosoma 21, así como la suma del número corregido por Poisson de cromosomas de referencia de los cuatro ensayos, se calculó y se consideró como los recuentos informativos para el ensayo de 4-plex. El valor de la  $P_r$  fue el recuento del cromosoma 21 para el ensayo de 4-plex dividido por la suma del cromosoma 21 y los recuentos de cromosomas de referencia para el ensayo 4-plex. Los valores de  $P_r$  derivados experimentalmente se sometieron a análisis de SPRT. Se analizarían una o más placas de 384 pocillos hasta que se pudiera clasificar la enfermedad mediante la SPRT. Se analizaron un total de dos mezclas de un 50 % de ADN genómico placentario euploide/50 % de ADN de capa leucocitaria materna y dos mezclas de un 50 % de ADN genómico placentario de trisomía 21/50 % de ADN de capa leucocitaria materna.

El valor de  $P_r$  derivado experimentalmente se compararía con el valor esperado de  $P_r$  para probar las hipótesis nula o alternativa. Como alternativa, ninguna de las hipótesis podría aceptarse si la  $P_r$  para el número dado de recuentos informativos no ha alcanzado el nivel requerido de confianza estadística para la clasificación de enfermedades. Estos casos se consideraron inclasificables hasta que se disponga de más casos.

Los resultados y la clasificación de SPRT de cada muestra están tabulados en las Figs. 22A y 22B. Las dos muestras euploides requirieron 2 y 5 análisis de RCD digital multiplexado de 384 pocillos antes de poder alcanzar la clasificación de SPRT. Los datos de ninguno de los miembros individuales del ensayo 4-plex permitieron la clasificación de enfermedades mediante SPRT. Ambas muestras de trisomía 21 se clasificaron correctamente cada una con solo un análisis RCD digital multiplexado de 384 pocillos. De forma análoga, los datos de ninguno de los miembros individuales del ensayo 4-plex permitieron la clasificación de enfermedades mediante SPRT. Sin embargo, los recuentos compuestos del ensayo 4-plex permitieron una clasificación de SPRT correcta. Estos datos ilustraron que mediante el uso de RCD digital multiplexado, el número efectivo de recuentos informativos aumentó sustancialmente para un número dado de análisis de PCR digital realizados, en comparación con la utilización de un ensayo de RCD digital de un solo plex.

## VI. UTILIZACIÓN DE DOSIFICACIÓN CROMOSÓMICA RELATIVA EPIGENÉTICA DIGITAL

Los presentes inventores perfilan aquí un enfoque llamado dosificación cromosómica relativa epigenética digital (ERCD digital; del inglés, digital epigenetic relative chromosome dosage) en el que los marcadores epigenéticos muestran un patrón de metilación del ADN específico del feto u otros cambios epigenéticos, en un cromosoma implicado en una aneuploidía cromosómica (por ejemplo, el cromosoma 21 en la trisomía 21) y en un cromosoma de referencia, se someten a análisis de PCR digital. La proporción entre el número de pocillos positivos para el marcador epigenético del cromosoma 21 y el positivo para el marcador epigenético del cromosoma de referencia en el ADN

plasmático extraído de mujeres gestantes que portan fetos normales nos dará el intervalo de referencia. Se espera que la proporción aumente si el feto tiene trisomía 21. Es obvio para los expertos en la materia que podría utilizarse en este análisis más de un marcador del cromosoma 21 y más de un marcador cromosómico de referencia.

Un ejemplo de un gen en el cromosoma 21 que presenta un patrón de metilación específico del feto (placenta) es el gen de la *Holocarboxilasa sintetasa* (*HLCS*). La *HLCS* se hipermetila en la placenta, pero está hipometilado en las células sanguíneas maternas; y se cubre en el documento US 2007/275402. Un ejemplo de un gen en un cromosoma de referencia que muestra un patrón de metilación específico del feto (placenta) es el gen *RASSF1A* en el cromosoma 3 [10]. *RASSF1A* está hipermetilado en la placenta pero está hipometilado en las células sanguíneas de la madre, véase el documento US 2009/155776.

En la aplicación de *HLCS* hipermetilado e *RASSF1A* hipermetilado para la detección por PCR digital de la trisomía 21 en un feto utilizando plasma materno, primero se recoge sangre periférica materna. A continuación, la sangre se somete a centrifugación y se recoge el plasma. A continuación, se extrae el ADN del plasma utilizando técnicas bien conocidas por los expertos en la materia, tales como la utilización de un kit QIAamp Blood (Qiagen). A continuación, el ADN plasmático se somete a digestión utilizando una o más enzimas de restricción sensibles a la metilación, tales como *HpaII* y *BstUI*. Esta(s) enzima(s) de restricción sensible(s) a la metilación cortará(n) las versiones maternas no metiladas de estos genes, si bien dejan intactas las secuencias hipermetiladas fetales. A continuación, la muestra de ADN de plasma digerida se diluye hasta un punto en que de promedio se detectarán aproximadamente de 0,2 a 1 molécula de las secuencias de *HLCS* o *RASSF1A* tratadas por enzimas de restricción pero intactas por pocillo de reacción. Se utilizarán dos sistemas de PCR en tiempo real para amplificar el ADN diluido, uno con dos cebadores y una sonda TaqMan específica para el gen *HLCS*, que abarca la región que se cortará por la(s) enzima(s) de restricción si la secuencia no está metilada; y el otro hacia el gen *RASSF1A*, de manera similar con dos cebadores y una sonda TaqMan. Respecto al último conjunto cebador/sonda de *RASSF1A*, se ha descrito un ejemplo por Chan *et al.* 2006, Clin Chem 52, 2211-2218. Las sondas TaqMan hacia las dianas *HLCS* y *RASSF1A* tendrán distintos indicadores fluorescentes, tales como FAM y VIC, respectivamente. A continuación, se utiliza una placa de 384 pocillos para realizar el experimento de PCR digital. El número de pocillos puntuados positivos solo para *HLCS* y aquellos puntuados positivos solo para *RASSF1A* se contará y se tomará una proporción de estos recuentos. La proporción *HLCS*:*RASSF1A* se espera que sea mayor para el plasma materno extraído de una mujer gestante portadora de un feto con trisomía 21, en comparación con una portadora de un feto euploide normal. El grado de sobrerepresentación dependerá de la concentración promedio del molde de referencia por pocillo en la ejecución de la PCR digital.

Otros métodos para puntuar estos resultados serán posibles, por ejemplo, el recuento del número de pocillos positivos para *HLCS*, independientemente de la positividad concurrente para *RASSF1A*; y viceversa para *RASSF1A*, independientemente de la positividad concurrente para *HLCS*. Asimismo, en sustitución del cálculo de la proporción, el número total o la diferencia en los recuentos de *HLCS* y *RASSF1A* podría utilizarse para indicar el estado de trisomía 21 de un feto.

Aparte de realizar la PCR digital en placas, también será obvio para los expertos en la materia que se pueden utilizar otras variantes de PCR digital, por ejemplo, chips de microfluidos, sistemas de microplacas de PCR de nanolitros, PCR en emulsión, PCR de colonia y amplificación de círculo rodante, extensión del cebador y espectrometría de masas, etc. Estas variantes de PCR digital se nombran a modo de ejemplos y no como limitaciones.

Aparte de la PCR en tiempo real, también resultará obvio para los expertos en la materia que se pueden utilizar métodos tales como la espectrometría de masas para puntuar los resultados de la PCR digital.

Aparte de utilizar enzimas de restricción sensibles a la metilación para diferenciar las versiones fetal y materna de *HLCS* y *RASSF1A*, será obvio para los expertos en la materia que serían aplicables también otros métodos para determinar el estado de metilación, por ejemplo, la modificación con bisulfito, PCR específica de metilación, inmunoprecipitación utilizando un anticuerpo frente a la citosina metilada, espectrometría de masas, etc.

Será obvio también para los expertos en la materia que el enfoque ilustrado en este ejemplo y otros ejemplos en esta solicitud de patente se puede utilizar en los otros fluidos corporales en los que se puede encontrar ADN fetal, incluyendo orina materna, líquido amniótico, lavados transcervicales, vellosidades coriónicas, saliva materna, etc.

## VII. SECUENCIACIÓN GENÓMICA PARALELA MASIVA UTILIZANDO PCR EN EMULSIÓN Y OTRAS ESTRATEGIAS

Los presentes inventores describirán aquí otro ejemplo por el cual una lectura digital de moléculas de ácido nucleico se puede utilizar para la detección de aneuploidías cromosómicas fetales, por ejemplo, trisomía 21, en el plasma materno. La aneuploidía cromosómica fetal es el resultado de (una) dosis anormal(es) de un cromosoma o región cromosómica. Es deseable que las pruebas no invasivas tengan una elevada sensibilidad y especificidad para minimizar los diagnósticos falsos. Sin embargo, el ADN fetal está presente en una concentración absoluta baja y representa una porción menor de todas las secuencias de ADN en el plasma y suero maternos. Por tanto, el número de muestras de PCR digital dirigidas a loci de genes específicos no se puede aumentar infinitamente dentro del mismo espécimen biológico. Por tanto, el análisis de múltiples conjuntos de loci diana específicos se puede utilizar para



aumentar la cantidad de datos que podrían obtenerse de un espécimen sin aumentar el número de muestreos de PCR digital realizados.

Por consiguiente, los ejemplos permiten la detección no invasiva de aneuploidía cromosómica fetal maximizando la cantidad de información genética que se podría inferir de la cantidad limitada de ácidos nucleicos fetales que existen como una población menor en una muestra biológica que contiene ácidos nucleicos maternos de fondo. En un aspecto, la cantidad de información genética obtenida es suficiente para un diagnóstico exacto, pero no demasiado excesiva para contener los costes y el aporte de muestra biológica requerida.

La secuenciación paralela masiva, tal como la alcanzable en la plataforma 454 (Roche) (Margulies, M. *et al.* 2005 Nature 437, 376-380), analizador genómico Illumina Genome Analyzer (o plataforma Solexa) o el sistema SOLiD System (Applied Biosystems) o la tecnología de secuenciación de ADN de una sola molécula verdadera de Helicos True Single Molecule DNA sequencing technology (Harris TD *et al.* 2008 Science, 320, 106-109), la molécula individual, la tecnología en tiempo real (SMRT™) de Pacific Biosciences y la secuenciación de nanoporos (Soni GV y Meller A. 2007 Clin Chem 53: 1996-2001), permiten la secuenciación de muchas moléculas de ácido nucleico aisladas de un espécimen en órdenes elevados de multiplexación de forma paralela (Dear Brief Funct Genomic Proteomic 2003; 1: 397-416). Cada una de estas plataformas secuencia moléculas individuales expandidas clonalmente o incluso no amplificadas de fragmentos de ácido nucleico.

A medida que se genera un gran número de lecturas de secuenciación, del orden de cientos de miles a millones o incluso posiblemente cientos de millones o miles de millones, a partir de cada muestra en cada ejecución, las lecturas secuenciadas resultantes forman un perfil representativo de la mezcla de especies de ácidos nucleicos en el espécimen original. Por ejemplo, los perfiles de haplotipo, transcrito y metilación de las lecturas secuenciadas se asemejan a los del espécimen original (Brenner *et al.* Nat Biotech 2000; 18: 630-634; Taylor *et al.* Cancer Res 2007; 67: 8511-8518). Debido a la gran cantidad de muestras de secuencias de cada espécimen, el número de secuencias idénticas, tal como las generadas a partir de la secuenciación de un conjunto de ácidos nucleicos en varios dobleces de cobertura o elevada redundancia, también es una buena representación cuantitativa del recuento de una determinada especie de ácido nucleico o locus en la muestra original.

En una realización, la secuenciación aleatoria se realiza sobre fragmentos de ADN que están presentes en el plasma de una mujer gestante y se obtienen secuencias genómicas que originalmente habrían provenido del feto o de la madre. La secuenciación aleatoria implica muestrear (secuenciar) una porción aleatoria de las moléculas de ácido nucleico presentes en la muestra biológica. Como la secuenciación es aleatoria, en cada análisis se puede secuenciar un subconjunto (fracción) diferente de las moléculas de ácido nucleico (y por tanto el genoma). Las realizaciones funcionarán incluso cuando este subconjunto varíe de una muestra a otra y de un análisis a otro, lo que puede ocurrir incluso utilizando la misma muestra. Los ejemplos de la fracción son aproximadamente un 0,1 %, 0,5 % o 1 % del genoma. En otras realizaciones, la fracción es al menos cualquiera de estos valores.

A continuación, puede utilizarse un procedimiento bioinformático para localizar cada una de estas secuencias de ADN en el genoma humano. Es posible que una proporción de tales secuencias se descarte del análisis posterior porque están presentes en las regiones repetidas del genoma humano o en regiones sujetas a variaciones interindividuales, por ejemplo, variación del número de copias. se puede determinar una cantidad del cromosoma de interés y de otro u otros cromosomas.

En un ejemplo, un parámetro (por ejemplo, una representación fraccionada) de un cromosoma potencialmente implicado en una aneuploidía cromosómica, por ejemplo, el cromosoma 21 o el cromosoma 18 o el cromosoma 13, se puede calcular a partir de los resultados del procedimiento bioinformático. La representación fraccionaria se puede obtener basándose en una cantidad de todas las secuencias (por ejemplo, alguna medida de todos los cromosomas) o un subconjunto particular de cromosomas (por ejemplo, solo un cromosoma respecto al que se está analizando)

A continuación, esta representación fraccionaria se compara con un intervalo de referencia establecido en gestaciones que implican fetos normales (es decir, euploides). Es posible que en algunas variantes del procedimiento, el intervalo de referencia se ajustaría de acuerdo con la concentración fraccionaria de ADN fetal (f) en una muestra de plasma materna particular. El valor de f puede determinarse a partir del conjunto de datos de secuenciación, por ejemplo, utilizando secuencias asignables al cromosoma Y si el feto es varón. El valor de f se puede determinar también en un análisis por separado, por ejemplo, utilizando marcadores epigenéticos fetales (Chan KCA *et al.* 2006 Clin Chem 52, 2211-8) o a partir del análisis de polimorfismos de un solo nucleótido.

En un aspecto, incluso cuando un grupo de ácidos nucleicos en un espécimen se secuencia con una cobertura genómica <100 % y entre la proporción de moléculas de ácido nucleico capturadas, la mayoría de cada especie de ácido nucleico solo se secuencia una vez, el desequilibrio de dosis de un locus de gen en particular o del cromosoma también se pueden determinar cuantitativamente. En otras palabras, el desequilibrio de dosis del locus génico o de cromosoma se infiere a partir de la representación porcentual de dicho locus entre todas las etiquetas secuenciadas cartografiadas del espécimen.

En un aspecto para el enfoque de secuenciación genómica paralela masiva, se pueden generar datos representativos

de todos los cromosomas al mismo tiempo. El origen de un fragmento en particular no se selecciona con anticipación. La secuenciación se realiza al azar y a continuación, se puede realizar una búsqueda de base de datos para ver de dónde proviene un fragmento en particular. Esto contrasta con las situaciones en las que se amplifica un fragmento específico del cromosoma 21 y otro del cromosoma 1.

5 En un ejemplo, una proporción de tales secuencias procedería del cromosoma implicado en una aneuploidía tal como el cromosoma 21 en este ejemplo ilustrativo. Otras secuencias de tal ejercicio de secuenciación todavía se derivarían de los otros cromosomas. Teniendo en cuenta el tamaño relativo del cromosoma 21 en comparación con los otros cromosomas, se podría obtener una frecuencia normalizada, dentro de un intervalo de referencia, de secuencias  
10 específicas del cromosoma 21 a partir de tal ejercicio de secuenciación. Si el feto tiene trisomía 21, la frecuencia normalizada de secuencias derivadas del cromosoma 21 de dicho ejercicio de secuenciación aumentará, permitiendo por tanto la detección de trisomía 21. El grado de cambio en la frecuencia normalizada dependerá de la concentración fraccionaria de ácidos nucleicos fetales en la muestra analizada.

15 En una realización, los presentes investigadores utilizaron el analizador de genoma Illumina para la secuenciación de un solo extremo de ADN genómico humano y muestras de ADN de plasma humano. El analizador de genoma de Illumina secuencia las moléculas de ADN individuales expandidas clonalmente capturadas en una superficie sólida denominada celda de flujo. Cada celda de flujo tiene 8 carriles para la secuenciación de 8 especímenes individuales o grupos de especímenes. Cada carril es capaz de generar ~200Mb de secuencia, que es solo una fracción de los 3  
20 mil millones de pares de bases de secuencias en el genoma humano. Cada muestra de ADN genómico o de ADN plasmático se secuenció utilizando un carril de una celda de flujo. Las etiquetas de secuencia corta generadas se alinearon con el genoma humano de referencia y se anotó el origen cromosómico. Se tabuló el número total de etiquetas secuenciadas individuales alineadas con cada cromosoma y se comparó con el tamaño relativo de cada cromosoma como se esperaba a partir del genoma humano de referencia o de los especímenes no representativos de la enfermedad. A continuación, se identificaron las ganancias o pérdidas de cromosomas.

El enfoque descrito es sólo una ejemplificación de la estrategia de dosificación de genes/cromosomas actualmente descrita. Como alternativa, se podría realizar una secuenciación de extremos emparejados. En lugar de comparar la longitud de los fragmentos secuenciados respecto a la esperada en el genoma de referencia como describen Campbell  
30 *et al.* (Nat Genet 2008; 40: 722-729), se contó y clasificó el número de etiquetas secuenciadas alineadas de acuerdo con la ubicación cromosómica. Las ganancias o pérdidas de regiones cromosómicas o cromosomas completos se determinaron comparando el recuento de etiquetas con el tamaño de cromosoma esperado en el genoma de referencia o el de un espécimen no representativo de la enfermedad.

35 En otra realización, la fracción del conjunto de ácidos nucleicos que se secuencia en una ejecución se subselecciona adicionalmente antes de secuenciar. Por ejemplo, las técnicas basadas en hibridación, tal como la matriz de oligonucleótidos, se podrían utilizar para subseleccionar primero las secuencias de ácido nucleico de ciertos cromosomas, por ejemplo, un cromosoma potencialmente aneuploide y otro(s) cromosoma(s) no implicado(s) en la aneuploidía analizada. Otro ejemplo es que se subselecciona o se enriquece antes de la secuenciación una  
40 determinada subpoblación de secuencias de ácido nucleico del conjunto de muestras. Por ejemplo, se ha informado de que las moléculas de ADN fetal en el plasma materno están compuestas por fragmentos más cortos que las moléculas de ADN de fondo materno (Chan *et al.* Clin Chem 2004; 50: 88-92). Por tanto, se pueden utilizar uno o más métodos conocidos por los expertos en la materia para fraccionar las secuencias de ácido nucleico en la muestra de acuerdo con el tamaño de la molécula, por ejemplo, por electroforesis en gel o por columnas de exclusión por tamaño  
45 o mediante un enfoque basado en microfluidos. Sin embargo, como alternativa, en el ejemplo del análisis de ADN fetal libre de células en el plasma materno, la porción de ácido nucleico fetal se podría enriquecer mediante un método que suprima el fondo materno, como mediante la adición de formaldehído (Dhallan *et al.* JAMA 2004; 291: 1114-9).

Otras estrategias de secuenciación de una molécula individual tal como la de la plataforma Roche 454, la plataforma SOLiD de Applied Biosystems, la tecnología de secuenciación de ADN Helicos True Single Molecule, la molécula individual, la tecnología en tiempo real (SMRT™) de Pacific Biosciences y la secuenciación de nanoporos podría utilizarse de manera similar en esta aplicación.

Se pueden encontrar ejemplos de resultados y una discusión adicional (por ejemplo, para secuenciar y calcular parámetros) en la solicitud presentada simultáneamente "DIAGNOSING FETAL CHROMOSOMAL ANEUPLOIDY USING GENOMIC SEQUENCING", US 2009/029377. Nótese que los métodos descritos en el presente documento para determinar un valor límite pueden implementarse cuando la reacción es de secuenciación, por ejemplo, como se describe en esta sección.

60 La determinación de la concentración fraccionaria de ADN fetal en el plasma materno también se puede realizar por separado de la ejecución de la secuenciación. Por ejemplo, la concentración de ADN del cromosoma Y podría determinarse utilizando PCR en tiempo real, PCR de microfluidos o espectrometría de masas. De hecho, la concentración de ADN fetal podría determinarse utilizando loci distintos del cromosoma Y y podría ser aplicable a fetos femeninos. Por ejemplo, Chan *et al.* demostraron que las secuencias de RASSF1A metiladas derivadas de origen fetal se detectarían en el plasma de mujeres gestantes en el fondo de secuencias de RASSF1A no metiladas derivadas maternalmente (Chan *et al.*, Clin Chem 2006;52:2211-8). La concentración de ADN fetal fraccionario se puede  
65

determinar por tanto dividiendo la cantidad de secuencias de *RASSF1A* metiladas por la cantidad total de secuencias de *RASSF1A* (metiladas y no metiladas).

Se espera que se prefiera el plasma materno al suero materno para la práctica de la invención de los presentes investigadores, porque el ADN se libera de las células sanguíneas de la madre durante la coagulación de la sangre. Por tanto, si se utiliza el suero, se espera que la concentración fraccionaria de ADN fetal sea menor en el plasma materno que en el suero materno. En otras palabras, si se utiliza suero materno, se espera que fuera necesario generar más secuencias para diagnosticar la aneuploidía cromosómica fetal, en comparación con una muestra de plasma obtenida de la misma mujer gestante al mismo tiempo.

Todavía otra forma alternativa de determinar la concentración fraccionaria de ADN fetal sería a través de la cuantificación de diferencias polimórficas entre las mujeres gestantes y el feto (Dhallan R, *et al.* 2007 Lancet, 369, 474-481). Un ejemplo de este método sería dirigirse a sitios polimórficos en los que la mujer gestante es homocigótica y el feto es heterocigótico. La cantidad de alelo específico del feto se puede comparar con la cantidad del alelo común para determinar la concentración fraccionaria de ADN fetal.

A diferencia de las técnicas existentes para detectar anomalías cromosómicas, incluida la hibridación genómica comparativa, hibridación genómica comparativa de micromatrices, la reacción en cadena de la polimerasa cuantitativa en tiempo real, que detectan y cuantifican una o más secuencias específicas, la secuenciación paralela masiva no depende de la detección o análisis de un conjunto predeterminado o predefinido de secuencias de ADN. Se secuencian una fracción aleatoria representativa de moléculas de ADN del grupo de especímenes. El número de etiquetas de secuencia diferentes alineadas con varias regiones cromosómicas se compara entre muestras que contienen o no ADN tumoral. Las anomalías cromosómicas se revelarían por las diferencias en el número (o porcentaje) de secuencias alineadas con cualquier región cromosómica dada en los especímenes.

En otro ejemplo, la técnica de secuenciación en ADN libre de células de plasma puede utilizarse para detectar las anomalías cromosómicas en el ADN plasmático para la detección de un cáncer específico. Los diferentes cánceres tienen un conjunto de anomalías cromosómicas típicas. Pueden utilizarse cambios (amplificaciones y deleciones) en múltiples regiones cromosómicas. Por tanto, habría una proporción aumentada de secuencias alineadas respecto a las regiones amplificadas y una proporción disminuida de secuencias alineadas respecto a las regiones disminuidas. La representación porcentual por cromosoma podría compararse con el tamaño de cada cromosoma correspondiente en un genoma de referencia expresado como el porcentaje de representación genómica de cualquier cromosoma dado con respecto al genoma completo. Se pueden utilizar también comparaciones directas o comparaciones con un cromosoma de referencia.

## VIII. DETECCIÓN DE MUTACIONES

El ADN fetal en el plasma materno existe como una población menor, con un promedio de un 3 % a un 6 % del ADN del plasma materno aportado por el feto. Debido a esta razón, la mayor parte del trabajo previo en la materia se ha centrado en la detección de dianas de ADN que el feto ha heredado del padre y que son distinguibles de la mayoría del ADN materno de fondo en el plasma materno. Entre los ejemplos de tales dianas detectadas previamente se incluyen el gen *SRY* en el cromosoma Y (Lo YMD *et al.* 1998 Am J Hum Genet, 62, 768-775) y el gen *RHD* cuando la madre es RhD negativa (Lo YMD *et al.* 1998 N Engl J Med, 339, 1734-1738).

Para la detección de mutaciones fetales, las estrategias previas que utilizan plasma materno se limitan a condiciones autosómicas dominantes en las que el padre es un portador, la exclusión de enfermedades autosómicas recesivas mediante la detección de mutaciones directa cuando el padre y la madre portan mutaciones diferentes o mediante análisis de enlace (Ding C *et al.* 2004 Proc Natl Acad Sci USA 101, 10762-10767). Estas estrategias anteriores tienen limitaciones considerables. Por ejemplo, para una pareja en donde tanto el hombre como la mujer de la pareja portan la misma mutación, a continuación, sería imposible llevar a cabo un diagnóstico prenatal significativo mediante la detección directa de la mutación en el plasma materno.

Tal escenario se ilustra en la Fig. 23. En este escenario, habrá tres genotipos fetales posibles, NN, NM y MM, donde N representa el alelo normal y M representa el alelo mutante. Entre los ejemplos de alelos mutantes se incluyen aquellos que causan fibrosis quística, beta-talasemia, alfa-talasemia, anemia de células falciformes, atrofia muscular espinal, hiperplasia suprarrenal congénita, etc. Se pueden encontrar otros ejemplos de estos trastornos en la herencia mendeliana en el hombre en línea (OMIM; del inglés, Online Mendelian Inheritance in Man) [www.ncbi.nlm.nih.gov/sites/entrez?db=OMIM&itool=toolbar](http://www.ncbi.nlm.nih.gov/sites/entrez?db=OMIM&itool=toolbar). En el plasma materno, la mayor parte del ADN vendrá de la madre y sería NM. Para cualquiera de los tres genotipos fetales, no habrá ningún alelo fetal único que permita su detección única en el plasma materno. Por tanto, la estrategia convencional no se puede aplicar aquí.

La invención permite manejar tales escenarios. En el escenario donde la madre y el feto son ambos NM, a continuación el alelo N y el alelo M estarán en equilibrio alélico. Sin embargo, si la madre es NM y el feto es NN, a continuación, habrá un desequilibrio alélico en el plasma materno, con el alelo N estando sobrerrepresentado. Por otro lado, si la madre es NM y el feto es MM, a continuación, habrá un desequilibrio alélico en el plasma materno, con el alelo M estando sobrerrepresentado. Por tanto, para la detección de mutaciones fetales, la hipótesis nula se refiere a la

ausencia de desequilibrio alélico cuando el feto es del genotipo NM. La hipótesis alternativa se refiere a la presencia de desequilibrio alélico y el genotipo fetal podría ser NN o MM dependiendo de si el alelo N o M está sobrerrepresentado.

5 La presencia o ausencia de desequilibrio alélico se puede determinar mediante PCR digital utilizando las realizaciones descritas en el presente documento. En un primer escenario, un volumen particular de plasma materno contiene el ADN liberado de 100 células, de las cuales 50 son de la madre y 50 son del feto. Por tanto, la concentración fraccionaria de ADN fetal en este volumen de plasma es de un 50 %. Cuando la madre es del genotipo NM, a continuación, habrá 50 alelos N y 50 alelos M aportados por la madre. Si el feto es del genotipo NM, a continuación, habrá 50 alelos N y 10  
10 50 alelos M aportados por el feto. Por tanto, no habrá desequilibrio alélico entre el alelo N y el alelo M, un total de 100 copias cada uno. Por otro lado, Si el feto es del genotipo NN, habrá 100 alelos N derivados del feto en este volumen de plasma. Por tanto, habrá un total de 150 alelos N por 50 alelos M. En otras palabras, habrá un desequilibrio alélico entre N y M, con N estando sobrerrepresentado en una proporción de 3:1 respecto a M.

15 En la situación inversa, si el feto es del genotipo MM, a continuación, habrá 100 alelos M derivados del feto en este volumen de plasma. Por tanto, habrá un total de 150 alelos M por 50 alelos N. En otras palabras, habrá un desequilibrio alélico entre N y M, estando M sobrerrepresentado en una proporción de 3:1 respecto a N. Tal desequilibrio alélico puede medirse mediante PCR digital. El alelo con el número menor de pocillos positivos se considera como el molde de referencia. Igual que los análisis de ARN-SNP digital y RCD digital, la distribución real de los alelos en el experimento de PCR digital estaría gobernada por la función de densidad de probabilidad de Poisson. Por tanto, si bien el grado teórico de desequilibrio alélico en el escenario presente es 3:1, el grado esperado de desequilibrio alélico dependería de la concentración molde promedio por pocillo durante el análisis de PCR digital. Por tanto, los límites de interpretación, tal como para el análisis de SPRT, adecuados para la concentración del molde de referencia promedio por pocillo ( $m_r$ ) deberían utilizarse para la clasificación de casos.

25 Asimismo, el grado de desequilibrio alélico que necesita medirse depende de la concentración de ADN fetal fraccionaria. En contraste con el ejemplo anterior, se puede considerar un volumen particular de plasma materno que contiene el ADN liberado de 100 células, de las cuales 90 son de la madre y 10 son del feto. Por tanto, la concentración fraccionaria de ADN fetal en este volumen de plasma es de un 10 %. Cuando la madre es del genotipo NM, a continuación, habrá 90 alelos N y 90 alelos M aportados por la madre. Si el feto es del genotipo NM, a continuación, habrá 10 alelos N y 10 alelos M aportados por el feto. Por tanto, no habrá desequilibrio alélico entre el alelo N y el alelo M, un total de 100 copias cada uno. Por otro lado, Si el feto es del genotipo NN, habrá 20 alelos N derivados del feto en este volumen de plasma. Por tanto, habrá un total de 110 alelos N por 90 alelos M.

35 En otras palabras, habrá un desequilibrio alélico entre N y M, con N estando sobrerrepresentado. En la situación inversa, si el feto es del genotipo MM, a continuación, habrá 20 alelos M derivados del feto en este volumen de plasma. Por tanto, habrá un 110 alelos M por 90 alelos N. En otras palabras, habrá un desequilibrio alélico entre N y M, con M estando sobrerrepresentado. El grado teórico de desequilibrio alélico cuando la concentración fraccionaria de ADN fetal es de un 10 % sería de 110:90, que es diferente a la proporción 3:1 cuando hay un 50 % de ADN fetal como se muestra en el ejemplo anterior. Por tanto, los límites de interpretación, tal como para el análisis de SPRT, adecuados para la concentración de ADN fetal fraccionaria deberían utilizarse para la clasificación de casos.

Por tanto, se extraerá el ADN plasmático. La cantidad de ADN materno y fetal en la muestra de plasma se cuantificará, por ejemplo, mediante los ensayos de PCR en tiempo real previamente establecidos (Lo, *et al.* 1998 Am J Hum Genet 62, 768-775) u otros tipos de cuantificadores bien conocidos por los expertos en la materia, por ejemplo, marcadores SNP (Dhallan R *et al.* 2007 Lancet, 369, 474-481) y marcadores epigenéticos fetales (Chan KCA *et al.* 2006 Clin Chem, 52, 2211-2218). Se calculará el porcentaje de ADN fetal. A continuación, la muestra de ADN plasmática cuantificada se prepara (por ejemplo, diluida o concentrada) de manera que durante el análisis de PCR digital, cada pocillo de reacción contendrá un promedio de una molécula molde (puede ser el alelo N o M). El análisis de PCR digital se llevará a cabo utilizando un par de cebadores, más dos sondas TaqMan, una específica para el alelo N, si bien la otra es específica para el alelo M. Se contará el número de pocillos que son positivos solo para M y el número de pocillos que son positivos solo para N. La proporción de estos pocillos se utilizará para determinar si hay evidencia de desequilibrio alélico. La evidencia estadística del desequilibrio alélico puede buscarse mediante métodos bien conocidos por los expertos en la materia, tales como el uso de SPRT. En una variante de este análisis, también es posible contar el número de pocillos que son positivos para M solamente o para M y N; así como contar el número de pocillos que son positivos solo para N o para M y N; y derivar una proporción de estos recuentos. Una vez más, la evidencia estadística del desequilibrio alélico puede buscarse mediante métodos bien conocidos por los expertos en la materia, tales como el uso de SPRT.

60 La determinación de la dosis de la mutación genética fetal, denominada dosis de mutante relativo (RMD; del inglés, relative mutant dosage) digital, se validó utilizando mezclas de ADN femenino/masculino (XX/XY). El ADN de células sanguíneas de un hombre y una mujer se mezcló cada uno con ADN masculino, produciendo muestras con genotipos XX o XY en un fondo de XY a concentraciones fraccionarias de un 25 % y un 50 %, respectivamente, como se muestra en la Fig. 24A.

65 Además, también se obtuvieron muestras de células sanguíneas de 12 hombres y 12 mujeres. El ADN de células

sanguíneas femeninas (genotipo XX) se mezcló cada uno con un exceso de 3 veces el ADN de células sanguíneas masculinas (genotipo XY), por tanto, produciendo 12 mezclas de ADN con un 25 % de ADN con el genotipo XX en un fondo de un 75 % de ADN con genotipo XY, con los resultados mostrados en la Fig. 24B.

- 5 Un objetivo del análisis SPRT fue determinar el genotipo menor presente en el ADN de fondo. En mezclas de ADN con un 25 % de ADN XX en un fondo de un 75 % de ADN XY, el alelo menor sería el Y derivado del 75 % de ADN. Dado que el 25 % del ADN en la muestra era de genotipo XX, si hubiera un total de 200 moléculas de ADN en la muestra, a continuación, 150 moléculas se habrían originado a partir del individuo XY. Por tanto, se esperaría que el número de alelos Y fuera 75. El número de alelos X aportados por la proporción masculina de ADN (genotipo XY) también es 75. El número de alelos X aportados por la hembra (genotipo XX) es 50 (2 veces 25). Por tanto, la proporción de X respecto a Y es  $125/75 = (1 + 25\%)/(1 - 25\%) = 5/3$ .

Para la segunda parte de este estudio, se obtuvieron muestras de células sanguíneas de sujetos masculinos y femeninos que portaban mutaciones HbE (G → A) y CD41/42 (CTTT/-) en el gen *beta-globina*, es decir, el gen *hemoglobina beta* (*HBB*). Para imitar las muestras de plasma materno obtenidas de madres heterocigóticas (MN, donde M = mutante y N = tipo salvaje) que portan fetos masculinos con todos los genotipos posibles (MM, MN o NN), el ADN de células sanguíneas de hombres que eran homocigóticos para los alelos de tipo silvestre (NN) o heterocigótico (MN) para una de las dos mutaciones, se mezcló cada uno con una muestra de ADN de células sanguíneas obtenida de hembras heterocigóticas para la misma mutación (MN). Por tanto, se produjeron mezclas de ADN a diversas concentraciones fraccionarias de ADN masculino/mutante. También se utilizó una muestra de ADN de células sanguíneas de un homocigótico femenino para la delección CD41/42 (MM) para preparar las mezclas de ADN. Para asegurar una proporción masculina exacta para la clasificación de SPRT, la concentración fraccionaria de ADN masculino de cada mezcla de ADN se determinó utilizando el ensayo ZFY/X.

- 25 El ensayo ZFY/X digital se utilizó para validar el SPRT, así como para determinar la concentración fraccionaria de ADN masculino en las mezclas de ADN. La dosis de secuencias de proteínas de *Zinc Finger* en el cromosoma X (ZFX) y Y (ZFY) se determinó mediante análisis de PCR digital. Un amplicón de 87 pb de los loci ZFX y ZFY se coamplificó primero mediante el cebador directo 5'-CAAGTGCTGGACTCAGATGTAAGT-3' y el cebador inverso 5'-TGAAGTAATGTCAGAAGCTAAAACATCA-3'. Se diseñaron dos sondas TaqMan específicas de cromosomas para distinguir los parálogos del cromosoma X e Y y sus secuencias fueron 5'-(VIC)TCTTTAGCACATTGCA(MGBNFQ)-3' y 5'-(FAM)TCTTTACCACACTGCAC(MGBNFQ)-3', respectivamente.

La dosis de mutante en las mezclas de ADN se determinó mediante análisis de PCR digital del alelo normal respecto al alelo mutante. Para la mutación HbE, un amplicón de 87 pb de los alelos normal y mutante se coamplificó primero con el cebador directo 5'-GGGCAAGGTGAACGTGGAT-3' y el cebador inverso 5'-CTATTGGTCTCCTTAAACCTGTCTTGTA-3'. Se diseñaron dos sondas TaqMan específicas de alelo para distinguir entre los alelos normal (G) y mutante (A) y sus secuencias fueron 5'-(VIC)TTGGTGGTGAGGCC(MGBNFQ)-3' y 5'-(FAM)TTGGTGGTAAGGCC(MGBNFQ)-3', respectivamente. Los resultados de la mutación de HbE se muestran en la Fig. 25.

40 Para la mutación por delección CD41/42, un amplicón de 87 y 83 pb de los alelos normal y mutante se coamplificó primero con el cebador directo 5'-TTTTCCCACCCTTAGGCTGC-3' y el cebador inverso 5'-ACAGCATCAGGAGTGACAGATC-3', respectivamente. Se diseñaron dos sondas TaqMan específicas de alelo para distinguir entre los alelos normal (sin delección) y mutante (con delección) y sus secuencias fueron 5'-(VIC)CAGAGGTTCTTTGAGTCCT(MGBNFQ)-3' y 5'-(FAM)AGAGGTTGAGTCCT(MGBNFQ)-3', respectivamente. Los resultados de la mutación de HbE se muestran en las Figs. 26A y 26B.

Estos experimentos se llevaron a cabo en un sistema BioMark™ (Fluidigm) utilizando las matrices digitales 12.765 (Fluidigm). La reacción para un panel se estableció utilizando 2X TaqMan Universal PCR Master Mix en un volumen de reacción de 10 ml. Para los ensayos CD41/42 y ZFY/X, cada reacción contenía un 1X TaqMan Universal PCR Master Mix, 900 nM de cada cebador, 125 nM de cada sonda y 3,5 ml de mezcla de ADN a 1 ng/ml. Para el ensayo de HbE, se agregaron 250 nM y 125 nM de sondas dirigidas a los alelos normal (G) y mutante (A), respectivamente. La mezcla de muestra/ensayo se cargó en la matriz digital mediante un Controlador IFC de NanoFlex™ (Fluidigm). La reacción se llevó a cabo en el sistema BioMark™ para la detección de señales. La reacción se inició a 50 °C durante 2 min, seguido de 95 °C durante 10 min y 50 ciclos de 95 °C durante 15 seg y 57 °C (para ZFY/X y CD41/42) o 56 °C (para HbE) durante 1 min. Se utilizó al menos un panel de reacción para cada caso y los datos se añadieron de paneles adicionales para las muestras que permanecieron sin clasificar hasta que se pudo tomar una decisión.

También será obvio para los expertos en la materia que la PCR digital se puede realizar utilizando métodos bien conocidos por los expertos en la materia, por ejemplo, chips de microfluidos, sistemas de microplacas de PCR de nanolitros, PCR en emulsión, PCR de colonia, amplificación de círculo rodante, extensión del cebador y espectrometría de masas.

En conclusión, los presentes inventores perfilan un enfoque para detectar el desequilibrio de secuencias en una muestra. En un ejemplo, este enfoque puede utilizarse para la detección no invasiva de aneuploidia cromosómica fetal, tal como la trisomía 21, mediante análisis de ácidos nucleicos fetales en plasma materno. Este enfoque también

se puede aplicar a otros materiales biológicos que contienen ácidos nucleicos fetales, incluyendo líquido amniótico, muestras de vellosidades coriónicas, orina materna, muestras endocervicales, saliva materna, etc. Primero, los presentes investigadores demostraron la utilización de este enfoque para determinar el desequilibrio alélico de un SNP en ARNm de *PLAC4*, un transcrito expresado por la placenta en el cromosoma 21, en el plasma materno de mujeres portadoras de fetos con trisomía 21. Segundo, los presentes investigadores demostraron que este enfoque se puede utilizar como un método no basado en polimorfismo, a través del análisis de dosis cromosómica relativa (RCD), para la detección prenatal no invasiva de trisomía 21. Tal enfoque basado en RCD digital implica la evaluación directa de si el número total de copias del cromosoma 21 en una muestra que contiene ADN fetal está sobrerrepresentado con respecto a un cromosoma de referencia. Incluso sin una instrumentación elaborada, el RCD digital permite la detección de trisomía 21 en muestras que contienen un 25 % de ADN fetal. Aplicamos la prueba de prueba de la proporción de probabilidades secuencial (SPRT) para interpretar los datos de la PCR digital. Los análisis de simulación por ordenador confirmaron la elevada exactitud del algoritmo de clasificación de enfermedades.

Los presentes inventores perfilaron adicionalmente que el enfoque se puede aplicar a la determinación de otras formas de desequilibrios de secuencias de ácido nucleico más allá de la aneuploidía cromosómica, tal como la detección de mutaciones fetales o la detección de polimorfismos en el plasma materno.

Cualquiera de los componentes de software o funciones descritos en esta aplicación, se puede implementar como código de software para ejecutarse por un procesador utilizando cualquier lenguaje de ordenador adecuado tal como, por ejemplo, Java, C++ o Perl utilizando, por ejemplo, técnicas convencionales u orientadas por objetivos. El código de software puede almacenarse como una serie de instrucciones o comandos en un medio legible por ordenador para almacenamiento y/o transmisión, los medios adecuados incluyen la memoria de acceso aleatorio (RAM), una memoria de solo lectura (ROM), un medio magnético como un disco duro o disquete o un medio óptico tal como un disco compacto (CD) o un DVD (disco digital versátil), memoria flash y similares. El medio legible por ordenador puede ser cualquier combinación de tales dispositivos de almacenamiento o transmisión.

Tales programas también pueden codificarse y transmitirse utilizando señales portadoras adaptadas para la transmisión a través de redes cableadas, ópticas y/o inalámbricas que se ajusten a una variedad de protocolos, incluyendo Internet. Como tal, se puede crear un medio legible por ordenador según una realización de la presente invención utilizando una señal de datos codificada con tales programas. Los medios legibles por ordenador codificados con el código del programa pueden empaquetarse con un dispositivo compatible o proporcionarse por separado de otros dispositivos (por ejemplo, a través de descarga de Internet). Cualquier medio legible por ordenador puede residir en o dentro de un solo producto de programa de ordenador (por ejemplo, un disco duro o un sistema informático completo) y puede estar presente en o dentro de diferentes productos de programa de ordenador dentro de un sistema o red. Un sistema informático puede incluir un monitor, una impresora u otra pantalla adecuada para proporcionar a un usuario cualquiera de los resultados mencionados en el presente documento.

Un ejemplo de un sistema informático se muestra en la Figura 27. Los subsistemas mostrados en la Fig. 27 están interconectados a través de un bus de sistema 2775. Se muestran subsistemas adicionales tales como una impresora 2774, teclado 2778, disco fijo 2779, monitor 2776, que está acoplado al adaptador de pantalla 2782 y otros. Los dispositivos periféricos y de entrada/salida (E/S), que se acoplan al controlador de I/O 2771, se pueden conectar al sistema informático mediante cualquier número de medios conocidos en la técnica, tal como el puerto serie 2777. Por ejemplo, el puerto serie 2777 o la interfaz externa 2781 se pueden utilizar para conectar el aparato informático a una red de área amplia, tal como Internet, un dispositivo de entrada de ratón o un escáner. La interconexión a través del bus del sistema permite que el procesador central 2773 se comuniquen con cada subsistema y controle la ejecución de instrucciones desde la memoria del sistema 2772 o el disco fijo 2779, así como el intercambio de información entre subsistemas. La memoria del sistema 2772 y/o el disco fijo 2779 pueden incorporar un medio legible por ordenador.

La descripción anterior de las realizaciones ejemplares de la invención se ha presentado con fines ilustrativos y descriptivos. No pretenden ser exhaustivos ni limitar la invención a la forma precisa descrita y son posibles muchas modificaciones y variaciones a la luz de las enseñanzas anteriores. Las realizaciones se han elegido y descrito para explicar los principios de la invención y sus aplicaciones prácticas para permitir a otros expertos en la materia utilizar la invención en diversas realizaciones y con diversas modificaciones adecuadas para el uso particular contemplado.

## REIVINDICACIONES

1. Un método implementado en un ordenador para determinar un genotipo de un feto de una mujer gestante en un primer locus, teniendo el feto una madre que es la mujer gestante, en donde la madre es heterocigótica en el primer locus para un primer alelo y un segundo alelo que es diferente del primer alelo, comprendiendo el método:
  - recibir datos a partir de una pluralidad de reacciones que implican moléculas de ADN libre de células de una muestra biológica de la madre, en donde la muestra biológica es plasma o suero materno y que incluye moléculas de ADN libre de células del sujeto femenino y del feto, en donde cada reacción indica una presencia o ausencia del primer alelo y del segundo alelo y en donde los datos incluyen:
    - (1) un primer conjunto de datos cuantitativos que indican una primera cantidad de reacciones que son positivas para la presencia del primer alelo del sujeto femenino y del feto en la pluralidad de reacciones; y
    - (2) un segundo conjunto de datos cuantitativos que indican una segunda cantidad de reacciones que son positivas para la presencia del segundo alelo del sujeto femenino y del feto en la pluralidad de reacciones;
  - determinar un parámetro a partir de los dos conjuntos de datos, en donde el parámetro incluye una proporción entre la primera cantidad y la segunda cantidad;
  - comparar el parámetro con uno o más valores límite para determinar el genotipo del feto, en donde se determina que el feto tiene solo el primer alelo en el primer locus cuando la primera cantidad es mayor que la segunda cantidad en un grado especificado por uno de dichos uno o más valores límite, en donde se determina que el feto tiene solo el segundo alelo en el primer locus cuando la segunda cantidad es mayor que la primera cantidad en un grado especificado por uno de dichos uno o más valores límite y en donde se determina que el feto es heterocigótico para el primer alelo y el segundo alelo cuando la primera cantidad y la segunda cantidad son estadísticamente iguales como se especifica por dichos uno o más valores límite.
2. El método de la reivindicación 1, en donde el primer y segundo alelos son alelos de un polimorfismo de un solo nucleótido.
3. El método de la reivindicación 1, en donde el primer alelo es un alelo mutante.
4. El método de la reivindicación 3, en donde el alelo mutante provoca fibrosis quística, beta-talasemia, alfa-talasemia, anemia de células falciformes, atrofia muscular espinal o hiperplasia suprarrenal congénita.
5. El método de la reivindicación 3, en donde el primer alelo incluye una delección o una amplificación.
6. El método de la reivindicación 1, que comprende adicionalmente: determinar dichos uno o más valores límite basados en una concentración de ADN fetal fraccionaria en la muestra biológica.
7. El método de la reivindicación 1, en donde una reacción es una reacción de secuenciación o una reacción de amplificación.
8. El método de la reivindicación 7, en donde la pluralidad de reacciones son reacciones en cadena de la polimerasa, en donde la pluralidad de reacciones contiene un promedio de un primer alelo por reacción o en donde la pluralidad de reacciones contiene un promedio de un segundo alelo por reacción y en donde la primera cantidad son varias reacciones que indican una presencia del primer alelo y en donde la segunda cantidad son varias reacciones que indican una presencia del segundo alelo.
9. El método de la reivindicación 1, que comprende adicionalmente: determinar que la madre es heterocigótica en el primer locus para el primer alelo y el segundo alelo.
10. Un producto de programa informático que comprende un medio legible por ordenador codificado con una pluralidad de instrucciones que, cuando se ejecutan en un sistema informático, provocan que el sistema informático realice una operación para determinar un genotipo de un feto de una mujer gestante en un primer locus de acuerdo con el método de cualquiera de las reivindicaciones 2 a 8, teniendo el feto una madre que es la mujer gestante, en donde la madre es heterocigótica en el primer locus para un primer alelo y un segundo alelo que es diferente del primer alelo.

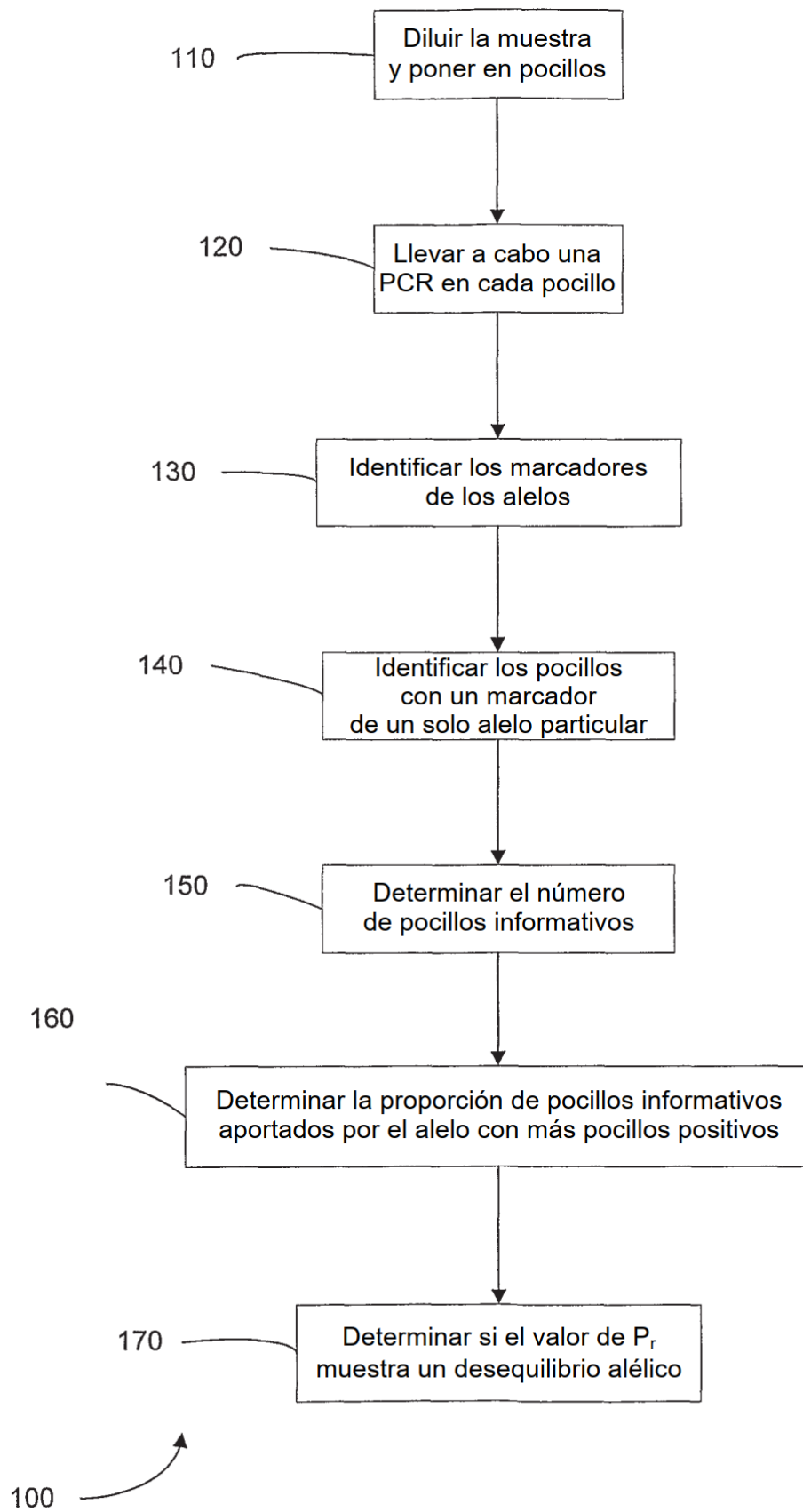


FIG. 1



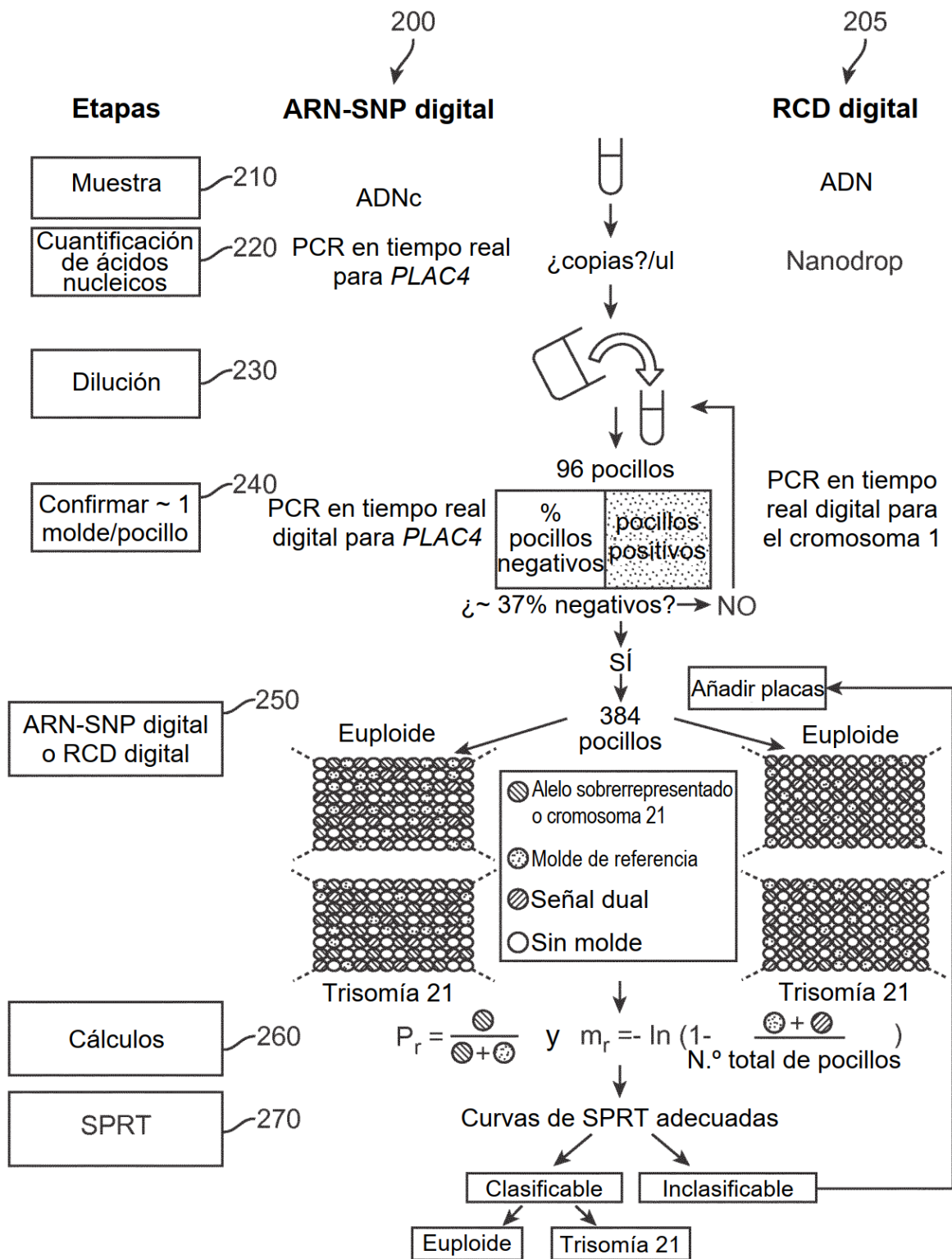


FIG. 2A

Cáncer	Anomalías cromosómicas comunes	
	Ganancia	Pérdida
Cáncer de vejiga	7	9p; 11q; 17p
Cáncer de mama	1q; 8q; 17q; 20q	8p; 13q; 16q; 18q
Cáncer colorrectal	7; 8q; 13q; 20q	1p; 5q; 8p; 15q; 17p; 18q
Carcinoma hepatocelular	1q; 8q; 20q	4q; 8p; 13q; 16q; 17p
Carcinoma de pulmón de células no pequeñas	1q; 3q; 5p; 8q	3p; 8p; 9p; 13q; 17p
Carcinoma de pulmón de células pequeñas	3q; 5p; 8q; 19q	3p; 4q; 5q; 10q; 13q; 16q
Carcinoma nasofaríngeo	1q; 3q; 8; 12; 19	1p; 3p; 9; 11q; 13q; 14q; 16q
Cáncer de próstata	7; 8q; Xq	8p; 13q; 16q

FIG. 2B

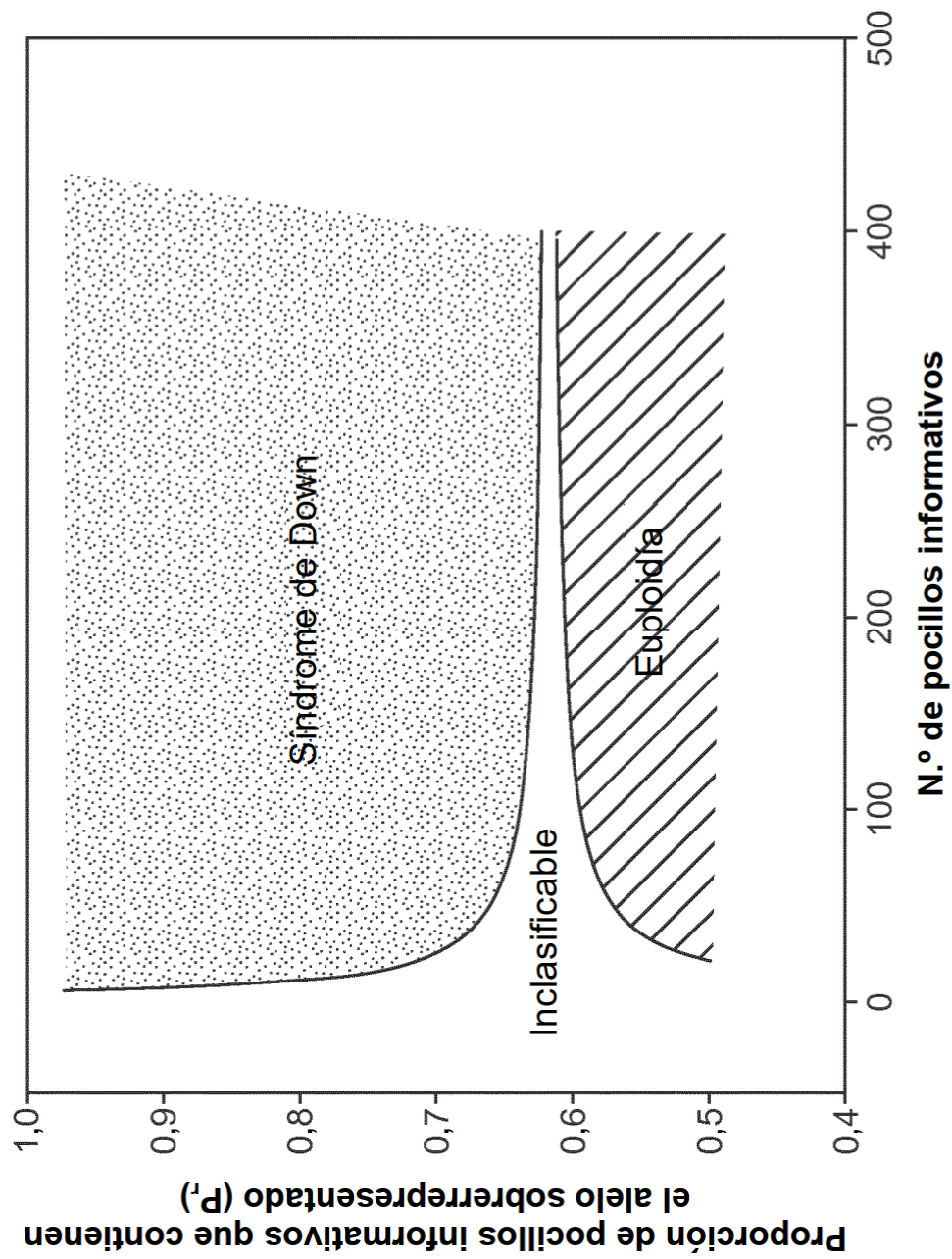


FIG. 3

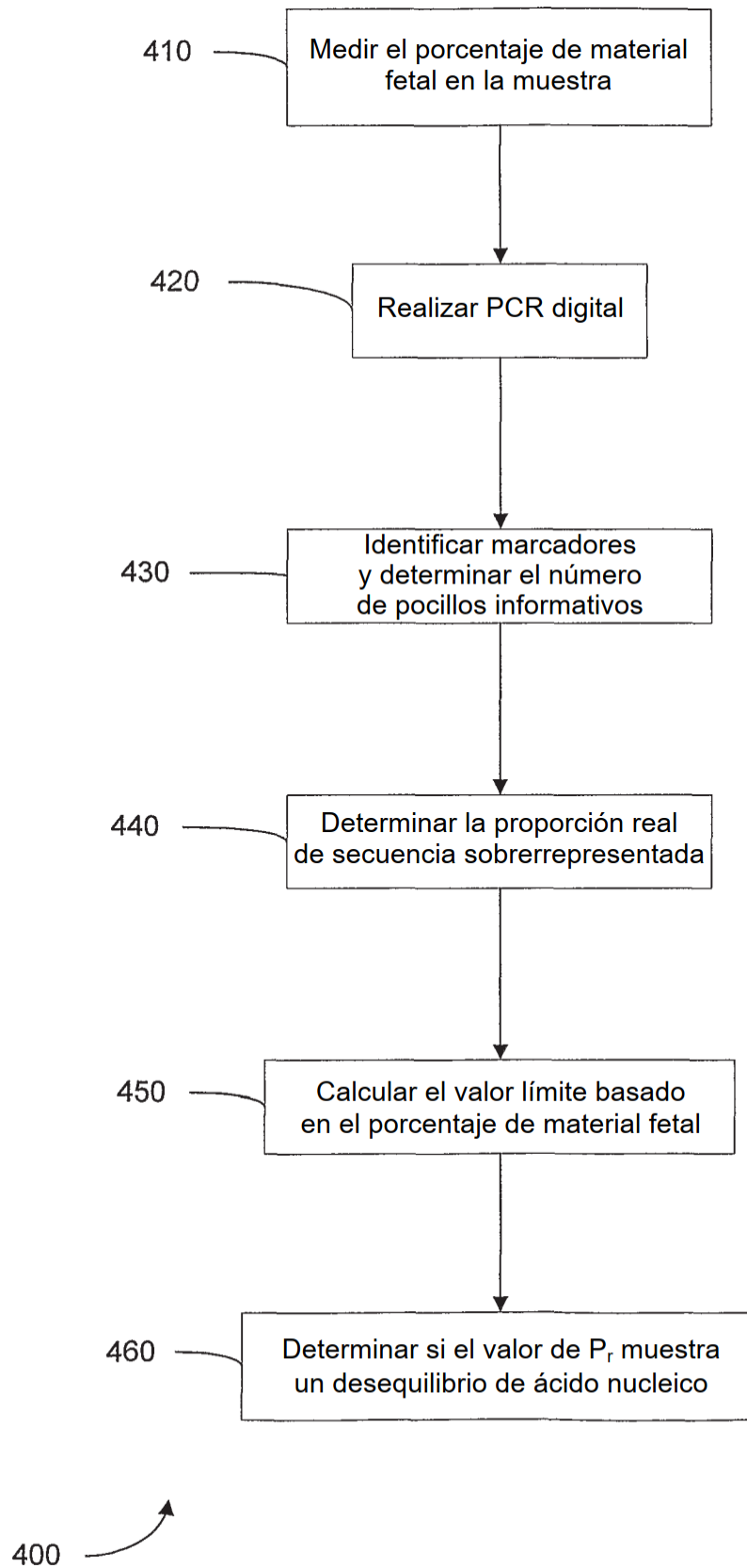


FIG. 4

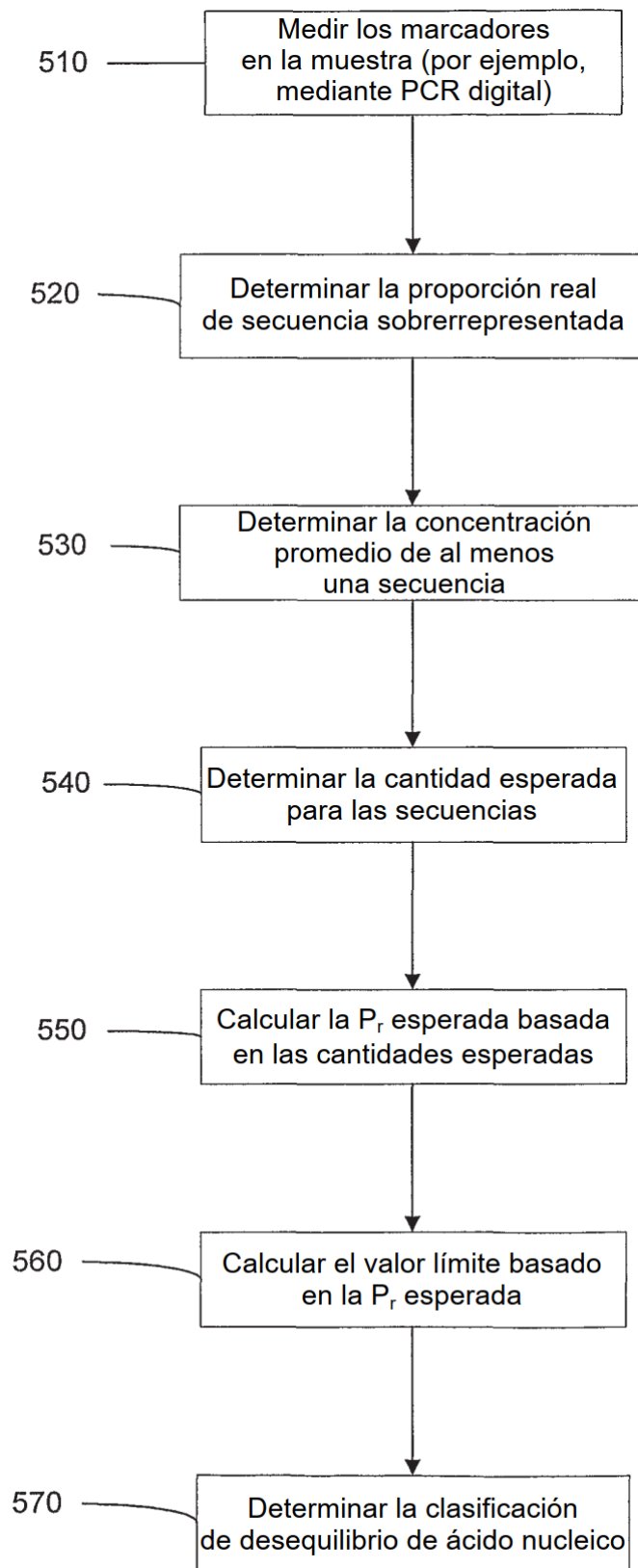


FIG. 5

500

600

$m_r$	Proporción de pocillos positivos para el alelo de referencia	Proporción de pocillos positivos para el alelo sobrerrepresentado	Proporción de pocillos positivos para ambos alelos	Proporción de pocillos positivos para el alelo de referencia SOLO	Proporción de pocillos positivos para el alelo sobrerrepresentado SOLO	Proporción alélica de ARN-SNP digital	Proporción de pocillos de informativos	$P_r$
0,1	0,0952	0,1813	0,0173	0,0779	0,164	2,11	0,2419	0,68
0,2	0,1813	0,3297	0,0598	0,1215	0,2699	2,22	0,3914	0,69
0,3	0,2592	0,4512	0,1169	0,1422	0,3342	2,35	0,4765	0,7
0,4	0,3297	0,5507	0,1815	0,1481	0,3691	2,49	0,5173	0,71
0,5	0,3935	0,6321	0,2487	0,1448	0,3834	2,65	0,5282	0,73
0,6	0,4512	0,6988	0,3153	0,1359	0,3835	2,82	0,5194	0,74
0,7	0,5034	0,7534	0,3793	0,1241	0,3741	3,01	0,4983	0,75
0,8	0,5507	0,7981	0,4395	0,1112	0,3586	3,23	0,4698	0,76
0,9	0,5934	0,8347	0,4953	0,0981	0,3394	3,46	0,4375	0,78
1	0,6321	0,8647	0,5466	0,0855	0,3181	3,72	0,4036	0,79
1,1	0,6671	0,8892	0,5932	0,0739	0,296	4	0,3699	0,8
1,2	0,6988	0,9093	0,6354	0,0634	0,2739	4,32	0,3373	0,81
1,3	0,7275	0,9257	0,6734	0,054	0,2523	4,67	0,3063	0,82
1,4	0,7534	0,9392	0,7076	0,0458	0,2316	5,06	0,2774	0,83
1,5	0,7769	0,9502	0,7382	0,0387	0,212	5,48	0,2507	0,84
1,6	0,7981	0,9592	0,7656	0,0325	0,1937	5,95	0,2262	0,85
1,7	0,8173	0,9666	0,79	0,0273	0,1766	6,47	0,2039	0,86
1,8	0,8347	0,9727	0,8119	0,0228	0,1608	7,05	0,1836	0,87
1,9	0,8504	0,9776	0,8314	0,019	0,1462	7,69	0,1652	0,88
2	0,8647	0,9817	0,8488	0,0158	0,1329	8,39	0,1487	0,89

FIG. 6



700

m <sub>r</sub>	Porcentaje de ADN fetal en la muestra							
	10%		25%		50%		100%	
	Porcentaje de pocillos informativos	P <sub>r</sub>	Porcentaje de pocillos informativos	P <sub>r</sub>	Porcentaje de pocillos informativos	P <sub>r</sub>	Porcentaje de pocillos informativos	P <sub>r</sub>
0,1	18%	0,51	18%	0,53	19%	0,56	21%	0,61
0,2	30%	0,51	31%	0,53	32%	0,56	35%	0,61
0,3	39%	0,51	40%	0,53	41%	0,57	43%	0,62
0,4	45%	0,51	45%	0,54	46%	0,57	48%	0,63
0,5	48%	0,52	49%	0,54	49%	0,57	51%	0,63
0,6	50%	0,52	50%	0,54	50%	0,58	51%	0,64
0,7	50%	0,52	50%	0,54	50%	0,58	50%	0,65
0,8	49%	0,52	49%	0,54	49%	0,58	48%	0,65
0,9	48%	0,52	47%	0,55	47%	0,59	46%	0,66
1	46%	0,52	45%	0,55	44%	0,59	43%	0,67
1,1	44%	0,52	43%	0,55	42%	0,6	40%	0,68
1,2	41%	0,52	40%	0,55	39%	0,6	37%	0,69
1,3	39%	0,52	38%	0,55	36%	0,6	34%	0,69
1,4	36%	0,52	35%	0,56	33%	0,61	31%	0,7
1,5	34%	0,52	33%	0,56	31%	0,61	28%	0,71
1,6	31%	0,52	30%	0,56	28%	0,62	26%	0,72
1,7	29%	0,53	28%	0,56	26%	0,62	23%	0,73
1,8	27%	0,53	25%	0,57	24%	0,63	21%	0,73
1,9	24%	0,53	23%	0,57	21%	0,63	19%	0,74
2	22%	0,53	21%	0,57	20%	0,64	17%	0,75

FIG. 7

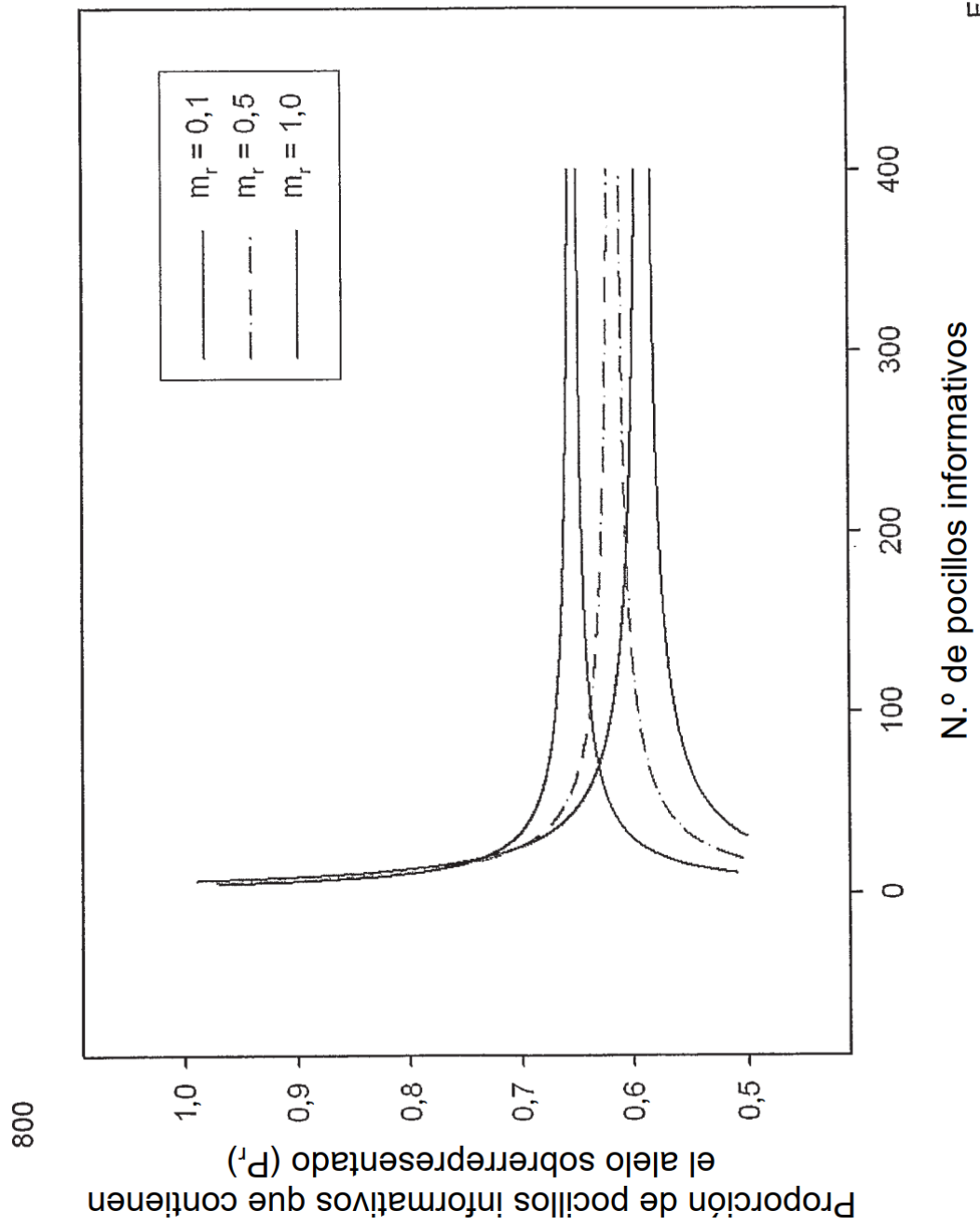


FIG. 8



900

m <sub>r</sub>	Clasificación	Algoritmo nuevo		Algoritmo viejo		m <sub>r</sub>	Clasificación	Algoritmo nuevo		Algoritmo viejo	
		Euploide	Trisomia 21	Euploide	Trisomia 21			Euploide	Trisomia 21	Euploide	Trisomia 21
0,1	Correcta	0%	37%	0%	35%	1,1	Correcta	89%	92%	14%	35%
	Incorrecta	3%	2%	3%	1%		Incorrecta	1%	0%	4%	0%
	Sin clasificar	97%	61%	97%	63%		Sin clasificar	10%	7%	81%	65%
	Correcta	6%	64%	0%	40%		Correcta	90%	92%	11%	35%
0,2	Incorrecta	4%	2%	4%	0%	1,2	Incorrecta	1%	0%	4%	0%
	Sin clasificar	90%	34%	96%	60%		Sin clasificar	9%	8%	84%	64%
	Correcta	34%	75%	6%	39%		Correcta	91%	92%	7%	34%
	Incorrecta	4%	2%	4%	0%		Incorrecta	1%	1%	4%	0%
0,3	Sin clasificar	62%	24%	90%	61%	1,3	Sin clasificar	8%	7%	89%	66%
	Correcta	54%	81%	14%	39%		Correcta	92%	91%	5%	34%
	Incorrecta	3%	1%	4%	0%		Incorrecta	1%	1%	4%	0%
	Sin clasificar	43%	19%	82%	61%		Sin clasificar	8%	7%	89%	66%
0,4	Correcta	64%	85%	20%	38%	1,4	Correcta	92%	90%	2%	32%
	Incorrecta	3%	1%	4%	0%		Incorrecta	1%	1%	4%	0%
	Sin clasificar	33%	14%	76%	62%		Sin clasificar	8%	9%	93%	68%
	Correcta	74%	88%	23%	38%		Correcta	92%	89%	1%	31%
0,5	Incorrecta	2%	1%	4%	0%	1,5	Incorrecta	1%	1%	4%	0%
	Sin clasificar	24%	11%	72%	62%		Sin clasificar	8%	10%	94%	69%
	Correcta	80%	89%	25%	37%		Correcta	92%	89%	1%	30%
	Incorrecta	2%	1%	4%	0%		Incorrecta	1%	1%	4%	0%
0,6	Sin clasificar	18%	10%	71%	63%	1,6	Sin clasificar	7%	10%	96%	70%
	Correcta	83%	91%	23%	37%		Correcta	92%	88%	0%	28%
	Incorrecta	2%	1%	4%	0%		Incorrecta	1%	1%	4%	0%
	Sin clasificar	15%	9%	73%	63%		Sin clasificar	7%	11%	96%	72%
0,7	Correcta	86%	91%	21%	36%	1,7	Correcta	92%	87%	0%	26%
	Incorrecta	1%	1%	4%	0%		Incorrecta	1%	1%	3%	0%
	Sin clasificar	12%	8%	75%	64%		Sin clasificar	8%	12%	97%	74%
	Correcta	88%	92%	19%	36%		Correcta	91%	86%	0%	24%
0,8	Incorrecta	1%	0%	4%	0%	1,8	Incorrecta	1%	1%	4%	0%
	Sin clasificar	11%	8%	77%	64%		Sin clasificar	9%	13%	96%	76%
	Correcta	88%	92%	19%	36%		Correcta	91%	86%	0%	24%
	Incorrecta	1%	0%	4%	0%		Incorrecta	1%	1%	4%	0%
0,9	Sin clasificar	12%	8%	75%	64%	1,9	Sin clasificar	8%	12%	97%	74%
	Correcta	88%	92%	19%	36%		Correcta	91%	86%	0%	24%
	Incorrecta	1%	0%	4%	0%		Incorrecta	1%	1%	4%	0%
	Sin clasificar	11%	8%	77%	64%		Sin clasificar	9%	13%	96%	76%
1	Correcta	86%	91%	21%	36%	2	Correcta	92%	87%	0%	26%
	Incorrecta	1%	1%	4%	0%		Incorrecta	1%	1%	3%	0%
	Sin clasificar	12%	8%	75%	64%		Sin clasificar	8%	12%	97%	74%
	Correcta	88%	92%	19%	36%		Correcta	91%	86%	0%	24%
0,9	Incorrecta	1%	0%	4%	0%	2	Incorrecta	1%	1%	4%	0%
	Sin clasificar	11%	8%	77%	64%		Sin clasificar	9%	13%	96%	76%
	Correcta	88%	92%	19%	36%		Correcta	91%	86%	0%	24%
	Incorrecta	1%	0%	4%	0%		Incorrecta	1%	1%	4%	0%

FIG. 9A

950

$m_r$	Clasificación	Algoritmo nuevo		Algoritmo viejo		$m_r$	Clasificación	Algoritmo nuevo		Algoritmo viejo	
		Euploide	Trisomia 21	Euploide	Trisomia 21			Euploide	Trisomia 21	Euploide	Trisomia 21
0,1	Correcta	57%	88%	48%	89%	1,1	Correcta	100%	100%	92%	90%
	Incorrecta	3%	1%	3%	1%		Incorrecta	0%	0%	1%	0%
	Sin clasificar	40%	11%	49%	10%		Sin clasificar	0%	0%	7%	10%
0,2	Correcta	89%	98%	79%	90%	1,2	Correcta	100%	100%	90%	90%
	Incorrecta	1%	0%	2%	0%		Incorrecta	0%	0%	1%	0%
	Sin clasificar	10%	2%	19%	10%		Sin clasificar	0%	0%	9%	10%
0,3	Correcta	96%	99%	88%	90%	1,3	Correcta	100%	100%	89%	90%
	Incorrecta	0%	0%	1%	0%		Incorrecta	0%	0%	10%	10%
	Sin clasificar	4%	1%	11%	10%		Sin clasificar	0%	0%	10%	10%
0,4	Correcta	99%	100%	92%	90%	1,4	Correcta	100%	100%	87%	90%
	Incorrecta	0%	0%	1%	0%		Incorrecta	0%	0%	1%	10%
	Sin clasificar	1%	0%	7%	10%		Sin clasificar	0%	0%	12%	10%
0,5	Correcta	100%	100%	94%	90%	1,5	Correcta	100%	100%	85%	90%
	Incorrecta	0%	0%	1%	0%		Incorrecta	0%	0%	1%	0%
	Sin clasificar	0%	0%	6%	10%		Sin clasificar	0%	0%	13%	10%
0,6	Correcta	100%	100%	94%	90%	1,6	Correcta	100%	100%	82%	90%
	Incorrecta	0%	0%	1%	0%		Incorrecta	0%	0%	1%	0%
	Sin clasificar	0%	0%	5%	10%		Sin clasificar	0%	0%	17%	10%
0,7	Correcta	100%	100%	95%	90%	1,7	Correcta	100%	100%	79%	90%
	Incorrecta	0%	0%	1%	0%		Incorrecta	0%	0%	2%	0%
	Sin clasificar	0%	0%	5%	10%		Sin clasificar	0%	0%	19%	10%
0,8	Correcta	100%	100%	95%	90%	1,8	Correcta	100%	100%	74%	90%
	Incorrecta	0%	0%	1%	0%		Incorrecta	0%	0%	2%	0%
	Sin clasificar	0%	0%	5%	10%		Sin clasificar	0%	0%	24%	10%
0,9	Correcta	100%	100%	94%	90%	1,9	Correcta	100%	100%	70%	90%
	Incorrecta	0%	0%	1%	0%		Incorrecta	0%	0%	2%	0%
	Sin clasificar	0%	0%	6%	10%		Sin clasificar	0%	0%	28%	10%
1	Correcta	100%	100%	94%	90%	2	Correcta	100%	100%	67%	90%
	Incorrecta	0%	0%	1%	0%		Incorrecta	0%	0%	3%	0%
	Sin clasificar	0%	0%	6%	10%		Sin clasificar	0%	0%	30%	10%

FIG. 9B



$m_r$	Resultado de la clasificación	Euploide	T21	$m_r$	Resultado de la clasificación	Euploide	T21
0,1	Correcta	57%	88%	0,9	Correcta	100%	100%
	Incorrecta	3%	1%		Incorrecta	0%	0%
	Sin clasificar	40%	11%		Sin clasificar	0%	0%
0,2	Correcta	89%	98%	1	Correcta	100%	100%
	Incorrecta	1%	0%		Incorrecta	0%	0%
	Sin clasificar	10%	2%		Sin clasificar	0%	0%
0,3	Correcta	96%	99%	1,1	Correcta	100%	100%
	Incorrecta	0%	0%		Incorrecta	0%	0%
	Sin clasificar	4%	1%		Sin clasificar	0%	0%
0,4	Correcta	99%	100%	1,2	Correcta	100%	100%
	Incorrecta	0%	0%		Incorrecta	0%	0%
	Sin clasificar	1%	0%		Sin clasificar	0%	0%
0,5	Correcta	100%	100%	1,3	Correcta	100%	100%
	Incorrecta	0%	0%		Incorrecta	0%	0%
	Sin clasificar	0%	0%		Sin clasificar	0%	0%
0,6	Correcta	100%	100%	1,4	Correcta	100%	100%
	Incorrecta	0%	0%		Incorrecta	0%	0%
	Sin clasificar	0%	0%		Sin clasificar	0%	0%
0,7	Correcta	100%	100%	1,5	Correcta	100%	100%
	Incorrecta	0%	0%		Incorrecta	0%	0%
	Sin clasificar	0%	0%		Sin clasificar	0%	0%
0,8	Correcta	100%	100%	2	Correcta	100%	100%
	Incorrecta	0%	0%		Incorrecta	0%	0%
	Sin clasificar	0%	0%		Sin clasificar	0%	0%

FIG. 10

1000

$m_r$	Resultado de la clasificación	Euploide	Trisomía 21	$m_r$	Resultado de la clasificación	Euploide	Trisomía 21
0,1	Correcta	38%	44%	0,7	Correcta	95%	94%
	Incorrecta	2%	2%		Incorrecta	0%	1%
	Sin clasificar	60%	54%		Sin clasificar	5%	5%
0,2	Correcta	64%	69%	0,8	Correcta	95%	95%
	Incorrecta	2%	2%		Incorrecta	1%	0%
	Sin clasificar	34%	29%		Sin clasificar	4%	5%
0,3	Correcta	78%	82%	0,9	Correcta	96%	96%
	Incorrecta	2%	1%		Incorrecta	0%	0%
	Sin clasificar	20%	17%		Sin clasificar	4%	4%
0,4	Correcta	85%	87%	1	Correcta	97%	96%
	Incorrecta	1%	1%		Incorrecta	0%	0%
	Sin clasificar	14%	12%		Sin clasificar	3%	4%
0,5	Correcta	89%	90%	1,5	Correcta	97%	96%
	Incorrecta	1%	1%		Incorrecta	0%	0%
	Sin clasificar	10%	9%		Sin clasificar	3%	4%
0,6	Correcta	92%	93%	2	Correcta	96%	95%
	Incorrecta	1%	1%		Incorrecta	0%	0%
	Sin clasificar	7%	7%		Sin clasificar	4%	5%

1100

FIG. 11

Número de pocillos	Resultados de la clasificación	Porcentaje de ADN fetal en la muestra					
		10%		25%		50%	
		Euploide	Trisomía 21	Euploide	Trisomía 21	Euploide	Trisomía 21
384	Correcta	0%	0%	6%	7%	47%	49%
	Incorrecta	0%	0%	0%	1%	2%	2%
	Sin clasificar	100%	100%	93%	92%	51%	49%
768	Correcta	0%	0%	23%	24%	73%	76%
	Incorrecta	0%	0%	2%	2%	2%	1%
	Sin clasificar	100%	100%	75%	74%	25%	23%
1152	Correcta	1%	1%	39%	40%	86%	87%
	Incorrecta	0%	0%	2%	2%	1%	1%
	Sin clasificar	99%	99%	59%	58%	13%	12%
1536	Correcta	2%	2%	50%	51%	93%	93%
	Incorrecta	0%	0%	2%	2%	1%	1%
	Sin clasificar	98%	97%	48%	47%	7%	6%
1920	Correcta	4%	4%	60%	60%	96%	96%
	Incorrecta	0%	1%	2%	2%	0%	0%
	Sin clasificar	96%	95%	38%	37%	4%	3%
3840	Correcta	18%	19%	84%	83%	100%	100%
	Incorrecta	1%	1%	1%	1%	0%	0%
	Sin clasificar	81%	80%	15%	15%	18%	22%
7680	Correcta	43%	45%	97%	97%	100%	100%
	Incorrecta	2%	2%	0%	0%	0%	0%
	Sin clasificar	55%	53%	3%	3%	0%	0%

1200

FIG. 12

1300

muestra	genotipo <sup>a</sup>	n.º de pocillos	n.º de pocillos positivos para alelos individuales				m <sub>r</sub> <sup>b</sup>	Resultado de SPRT		
			solo A	solo G	AG	todos negativos		P <sub>r</sub> <sup>c</sup>	región inclasificable <sup>d</sup>	clasificación <sup>e</sup>
<b><u>ADN placentario</u></b>										
N677	AG	384	85	83	126	90	0,79	0,51	0,63 - 0,65	euploide
N710	AG	384	102	83	73	126	0,52	0,55	0,61 - 0,63	euploide
N435	AGG	384	49	157	130	48	0,63	0,76	0,62 - 0,64	T21
N981	AAG	384	135	69	82	98	0,5	0,66	0,61 - 0,63	T21
<b><u>ADN placentario</u></b>										
V533	AG	384	103	93	71	117	0,56	0,53	0,61 - 0,63	euploide
N943	AG	384	89	100	74	121	0,55	0,53	0,61 - 0,63	euploide
N435	AGG	384	52	138	95	99	0,48	0,73	0,61 - 0,63	T21
T215	AAG	384	146	58	138	42	0,71	0,72	0,62 - 0,64	T21

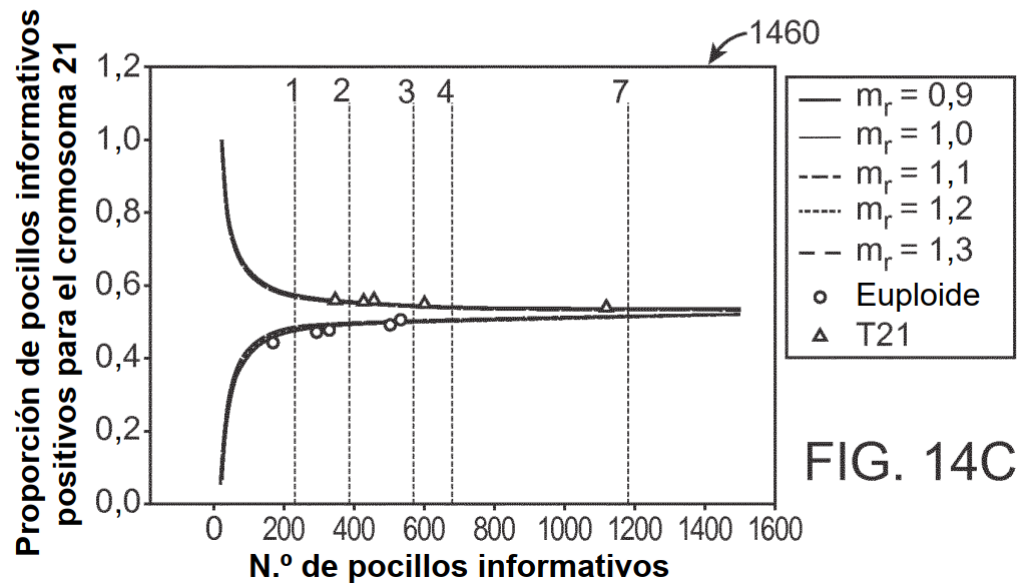
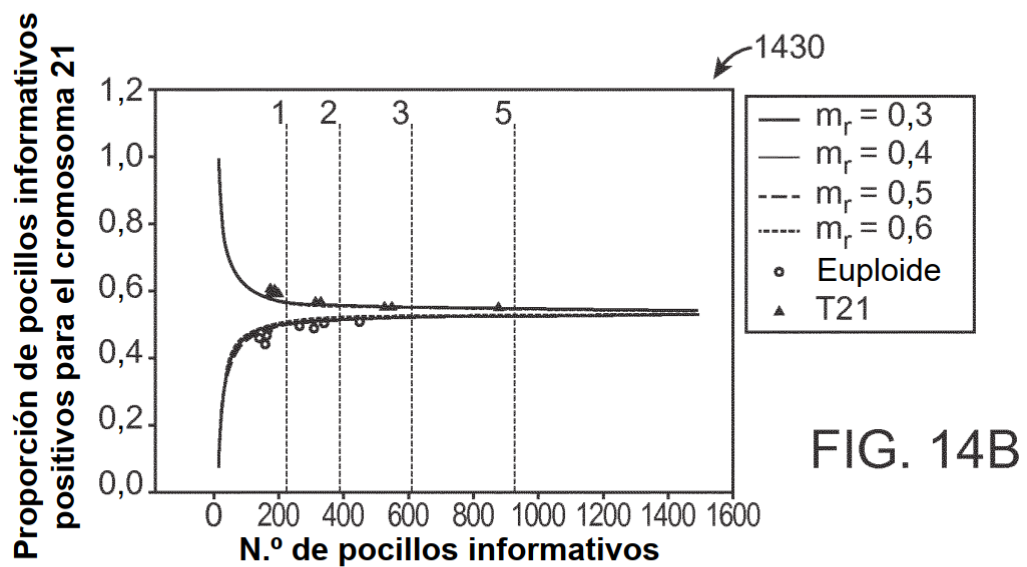
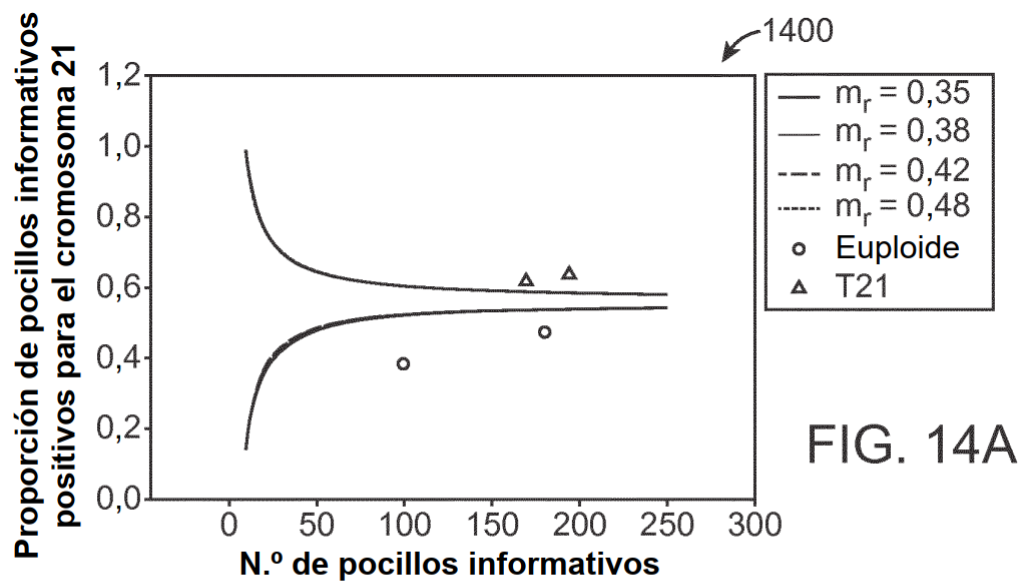
FIG. 13A

1350

muestra	genotipo	n.º de pocillos	n.º de pocillos positivos para alelos individuales				m	proporción	Resultado de SPRT	
			solo A	solo G	AG	todos negativos			región inclasificable	clasificación
M2390P	AG	384	90	100	97	97	0,67	0,526	0,62 - 0,64	euploide
M2391P	AG	384	97	105	65	117	0,55	0,52	0,61 - 0,63	euploide
M2473P	AG	384	66	92	34	192	0,3	0,582	0,59 - 0,62	euploide
M2524P	AG	384	29	28	3	324	0,08	0,509	0,54 - 0,64	euploide
M2528P	AG	384	112	85	44	143	0,41	0,569	0,60 - 0,62	euploide
M2601P	AG	384	90	101	72	121	0,55	0,529	0,61 - 0,63	euploide
M2607P	AG	384	73	91	57	163	0,41	0,555	0,60 - 0,63	euploide
M2638P	AG	384	66	90	52	176	0,37	0,577	0,59 - 0,62	euploide
M2639P	AG	384	71	56	17	240	0,21	0,559	0,58 - 0,62	euploide
M2525P	AAG	384	110	53	21	200	0,21	0,675	0,58 - 0,61	T21
M2272P	AAG	768	246	127	112	283	0,37	0,66	0,60 - 0,61	T21
M2718P	AGG	384	66	114	66	138	0,42	0,633	0,60 - 0,62	T21
M1519P	AGG	384	58	130	54	142	0,34	0,691	0,59 - 0,62	T21

FIG. 13B





muestra	genotipo	n.º de pocillos	n.º de pocillos positivos para alelos individuales				$m_r^a$	$P_r^b$	Análisis de SPRT	
			solo A	solo G	GA	todos negativos			Región inclasificable	Clasificación <sup>c</sup>
V943	AG	765	168	209	117	271	0,47	0,554	0,609 - 0,621	euploide
N435	AGG	765	92	281	105	287	0,30	0,753	0,597 - 0,610	T21
T21-5	AAG	765	267	126	141	231	0,43	0,679	0,607 - 0,618	T21

<sup>a</sup> n.º de moléculas de referencia por pocillo

<sup>b</sup> Los valores de  $P_r$  se calcularon mediante la ecuación:

$$P_r = \frac{\text{n.º de pocillos positivos para el alelo sobrerrepresentado}}{\text{n.º de pocillos positivos solo para A} + \text{n.º de pocillos positivos solo para G}}$$

<sup>c</sup> "Euploide" se asignó cuando la  $P_r$  estuvo por debajo de la región inclasificable; "T21" se asignó cuando la  $P_r$  estuvo por encima de la región inclasificable. T21, trisomía 21

FIG. 15A

n.º panel	n.º de pocillos	n.º de pocillos positivos para alelos individuales			
		solo A	solo G	GA	todos negativos
S1	765	14	15	0	736
S2	765	7	10	0	748
S3	765	13	9	1	744
S4	765	7	8	0	750
S5	765	12	8	0	745
S6	765	8	14	0	743
S7	765	13	9	0	743
S8	765	9	8	0	748
S9	765	14	15	0	736
S10	765	9	12	0	744
S11	765	8	17	0	740
S12	765	11	9	0	745

FIG. 15B

muestra	geno-tipo	paneles <sup>a</sup>	n.º de pocillos positivos para alelos individuales				$P_r$	Análisis de SPRT	
			n.º de pocillos	solo A	solo G	GA		Región inclasificable	Clasificación
M2879P	AG	5	3825	82	103	5	0,557	0,570 - 0,602	Euploide
M2875P	AG	3	2295	35	40	0	0,533	0,546 - 0,625	Euploide
M3078P	AG	3	2295	33	33	1	0,500	0,541 - 0,631	Euploide
M2976P	AG	12	9180	56	46	0	0,549	0,556 - 0,615	Euploide
M2831P	AAG	2	1530	42	13	0	0,764	0,531 - 0,640	T21

<sup>a</sup> n.º de paneles necesarios para la clasificación

FIG. 15C



N.º de pocillos positivos para el alelo A o G					Análisis de SPRT			
Muestra	Genotipo <sup>a</sup>	solo A	solo G	Ambos	$m_r^b$	$P_r^c$	Región inclasificable <sup>d</sup>	Clasificación <sup>e</sup>
N3505	AG	157	131	48	0,27	0,545	0,59- 0,61	Euploide
N3530	AG	174	132	52	0,28	0,569	0,59- 0,61	Euploide
N3531	AG	146	129	28	0,23	0,531	0,58- 0,60	Euploide
N3533	AG	149	135	46	0,27	0,525	0,59- 0,61	Euploide
N2772	AG	159	149	79	0,35	0,516	0,60- 0,61	Euploide
N3506	AG	154	145	66	0,32	0,515	0,60- 0,61	Euploide
N3555	AG	94	73	7	0,11	0,563	0,58- 0,61	Euploide
N3560	AG	101	109	28	0,18	0,519	0,58- 0,61	Euploide
N3561	AG	91	70	17	0,12	0,565	0,58- 0,61	Euploide
N828	AAG	153	89	22	0,16	0,632	0,58- 0,60	T18
CA2	AGG	12	28	0	0,016	0,700	0,51- 0,66	T18
CA7	AGG	97	171	56	0,22	0,638	0,59- 0,61	T18

<sup>a</sup> Los genotipos se determinaron mediante espectrometría de masas.

<sup>b</sup> El valor  $m_r$  indica el número promedio de moléculas molde de referencia por pocillo.

<sup>c</sup> El valor  $P_r$  indica la proporción de pocillos informativos positivos para el alelo sobrerrepresentado.

<sup>d</sup> La región inclasificable varía con el valor de  $m_r$ .

<sup>e</sup> "Euploide" se asignó cuando  $P_r$  estuvo por debajo de la región inclasificable; "T18" se asignó cuando  $P_r$  estuvo por encima de la región inclasificable.

FIG. 16A

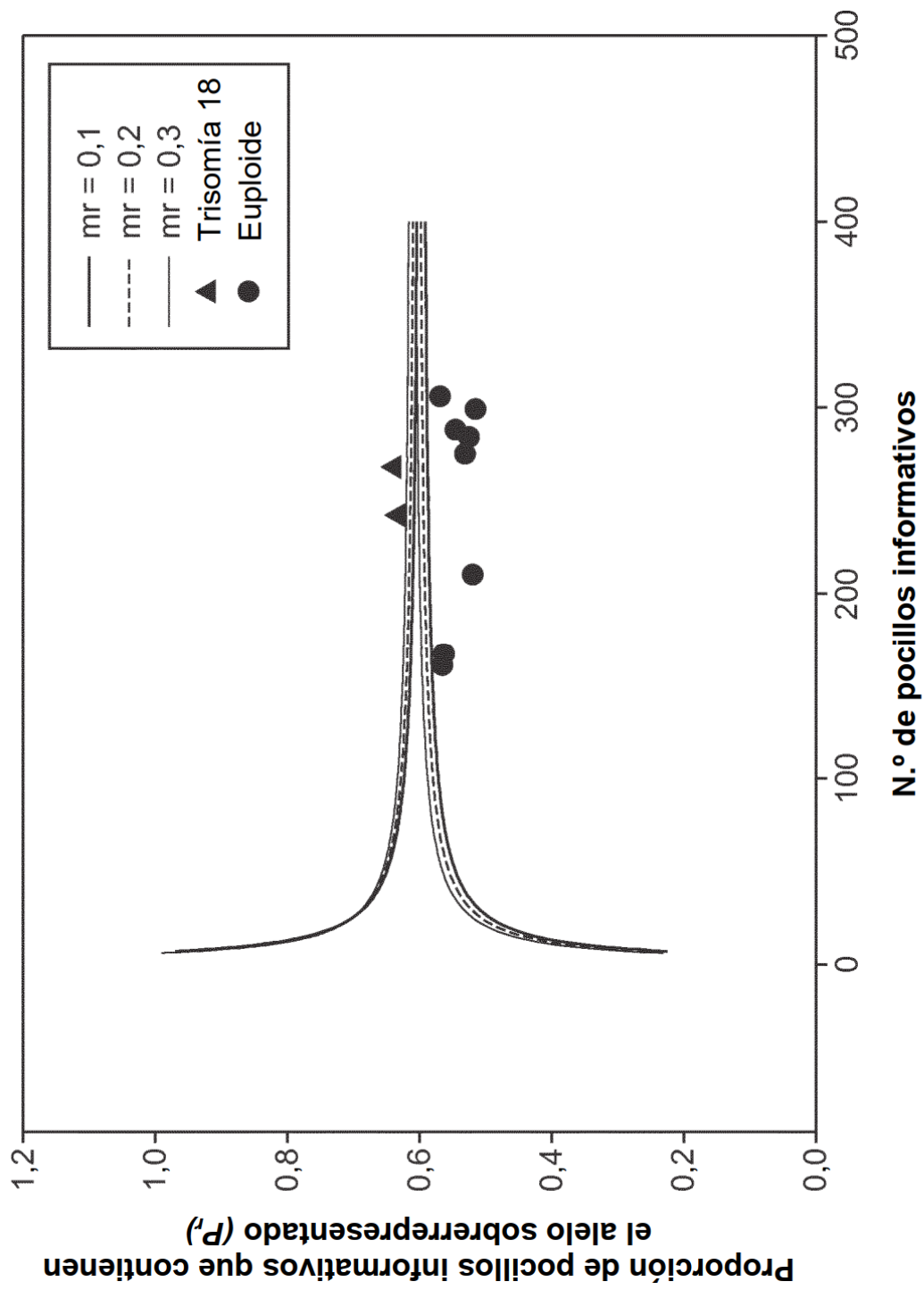


FIG. 16B

Tipo	Casos	N.º total de paneles utilizados	N.º total de pocillos informativos	N.º de recuento de pocillos			Análisis de SPRT			
				Ambos positivos	Solo Cr 21	Solo Cr 1	m <sub>r</sub> <sup>a</sup>	P <sub>r</sub> <sup>a</sup>	Región inclasificable	Clasificación <sup>c</sup>
T21	N0234	1	303	68	172	131	0,30	0,568	0,507 - 0,559	T21
	N0274	2	558	118	307	251	0,28	0,550	0,518 - 0,547	T21
	N0332	1	293	73	168	125	0,30	0,573	0,506 - 0,560	T21
	N0349	2	574	107	322	252	0,27	0,561	0,518 - 0,546	T21
	N0435	1	299	67	176	123	0,29	0,589	0,506 - 0,559	T21
	N0456	1	318	68	189	129	0,30	0,594	0,508 - 0,558	T21
	N0667	2	602	134	345	257	0,30	0,573	0,519 - 0,556	T21
	N0764	1	289	62	167	122	0,28	0,578	0,505 - 0,560	T21
	N0890	4	1206	272	657	549	0,31	0,545	0,526 - 0,539	T21
	N0891	1	317	103	180	137	0,38	0,568	0,510 - 0,558	T21
Normal	N0196	1	262	69	111	151	0,34	0,424	0,504 - 0,563	Euploide
	N0198	1	285	68	132	153	0,34	0,463	0,506 - 0,561	Euploide
	N0230	2	527	109	272	255	0,27	0,516	0,517 - 0,548	Euploide
	N0231	2	542	98	275	267	0,27	0,507	0,517 - 0,547	Euploide
	N0262	1	295	57	144	151	0,32	0,488	0,506 - 0,560	Euploide
	N0463	1	263	70	120	143	0,33	0,456	0,504 - 0,563	Euploide
	N0527	1	292	56	137	155	0,32	0,469	0,506 - 0,560	Euploide
	N0619	1	281	60	138	143	0,31	0,491	0,505 - 0,561	Euploide
	N0625	1	291	60	134	157	0,33	0,460	0,507 - 0,560	Euploide
	N0651	2	566	118	287	279	0,30	0,507	0,519 - 0,547	Euploide

<sup>a</sup> n.º de moléculas de referencia por pocillo

<sup>b</sup> Los valores de P<sub>r</sub> se calcularon mediante la ecuación:

$$n.º \text{ de pocillos positivos para el alelo sobrerrepresentado (cr21)} = \frac{n.º \text{ de pocillos positivos solo para cr21} + n.º \text{ de pocillos positivos solo para Ref}}{2}$$

<sup>c</sup> "Euploide" se asignó cuando la Pr estuvo por debajo de la región inclasificable; "T21" se asignó cuando la Pr estuvo por encima de la región inclasificable.

FIG. 17

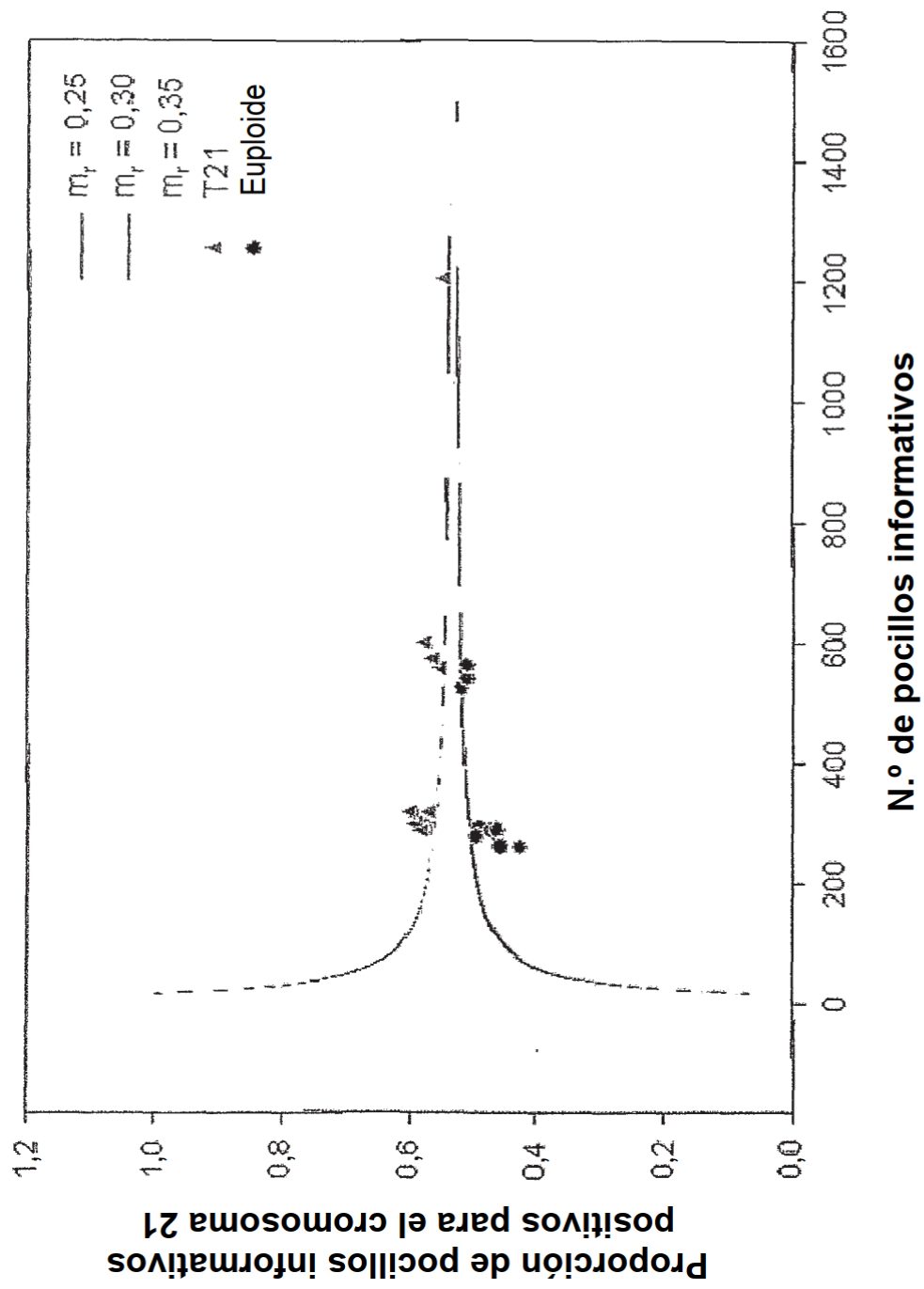


FIG. 18

Tipo	Casos	N.º total de paneles utilizados	N.º de pocillos informativos	N.º de recuento de pocillos			Análisis de SPRT		
				Ampos positivos	Solo Cr 21	Solo Cr 1	m <sup>a</sup> r	P <sup>a</sup> r	Región inclasificable Clasificación <sup>c</sup>
T21	A0274	1	380	269	249	131	0,74	0,655	0,567 - 0,585
	A2006	1	348	373	233	115	1,02	0,670	0,579 - 0,596
	A2178	1	384	288	239	145	0,83	0,622	0,571 - 0,588
	AS363A	1	364	109	244	120	0,36	0,670	0,551 - 0,573
	AS979A	1	256	492	188	68	0,32	0,734	0,590 - 0,610
	AS3457A	1	376	83	256	120	0,31	0,681	0,540 - 0,572
Normal	A0236	1	387	275	195	192	0,94	0,504	0,576 - 0,592
	A0237	1	392	228	184	208	0,84	0,469	0,572 - 0,588
	A0255	1	378	222	195	183	0,75	0,516	0,568 - 0,585
	A0263	1	254	43	127	127	0,25	0,500	0,541 - 0,576
	AS365A	1	372	135	192	178	0,53	0,519	0,558 - 0,578
	AS463A	1	413	64	186	168	0,36	0,525	0,550 - 0,574
	AS484A	1	316	271	194	178	0,88	0,522	0,573 - 0,590
	AS760A	1	190	100	200	213	0,74	0,484	0,560 - 0,584
	AS780A	1	152	67	159	157	0,35	0,503	0,548 - 0,575
	AS897A	1	394	6	87	111	0,17	0,439	0,532 - 0,579
	AS821A	1	334	8	69	83	0,13	0,454	0,523 - 0,586
	AS980A	1	120	137	191	203	0,59	0,485	0,561 - 0,580
	AS2282A	1	379	342	177	157	1,06	0,530	0,580 - 0,597
	AS2425A	1	354	6	64	64	0,10	0,500	0,516 - 0,591
	AS2434A	1	369	97	203	176	0,44	0,536	0,555 - 0,575
	A1119	1	256	123	185	169	0,48	0,523	0,555 - 0,577
	A1122	1	246	100	195	174	0,44	0,528	0,554 - 0,576
	A1152	1	373	40	120	128	0,26	0,500	0,554 - 0,576
	A1180	1	359	499	133	113	1,61	0,541	0,604 - 0,622
	A1275	1	359	95	201	172	0,43	0,539	0,554 - 0,575
	A1375	1	359	55	190	169	0,35	0,529	0,550 - 0,573
	A1390	1	357	97	165	192	0,47	0,462	0,555 - 0,577
	A1399	1	371	114	197	174	0,47	0,531	0,556 - 0,577

<sup>a</sup> n.º de moléculas de referencia por pocillo

<sup>b</sup> Los valores de P<sub>r</sub> se calcularon mediante la ecuación:

n.º de pocillos positivos para el alelo sobrerrepresentado (cr21)

n.º de pocillos positivos solo para cr21 + n.º de pocillos positivos solo para Ref

<sup>c</sup> "Euploide" se asignó cuando la Pr estuvo por debajo de la región inclasificable; "T21" se asignó cuando la Pr estuvo por encima de la región inclasificable.

**FIG. 19**

Tipo	Casos	N.º total de paneles utilizados	N.º total de pocillos informativos	N.º de recuento de pocillos			m <sup>a</sup> <sub>r</sub>	Análisis de SPRT		
				Ambos positivos	Solo Cr18	Solo Cr21		P <sup>b</sup> <sub>r</sub>	Región inclasificable	Clasificación <sup>c</sup>
Euploide	N0230	1	380	243	181	199	0,86	0,476	0 ,573 - 0,589	E
	N0625	1	370	225	179	191	0,78	0,484	0 ,569 - 0,586	E
	N0651	1	358	206	192	166	0,67	0,536	0,564 - 0,583	E
T18	N0828	1	543	256	329	214	0,95	0,606	0,579 - 0,590	T18
	N0830	1	346	353	240	106	0,92	0,694	0,574 - 0,592	T18
	N0936	1	372	285	209	163	0,88	0,562	0,573 - 0,590	E
	N1678	1	343	115	221	122	0,37	0,644	0,551 - 0,574	T18
	N1701	1	303	94	191	112	0,31	0,630	0,546 - 0,575	T18

<sup>a</sup> n.º de moléculas de referencia por pocillo

<sup>b</sup> Los valores de P<sub>r</sub> se calcularon mediante la ecuación:

$$n.º \text{ de pocillos positivos para el alelo sobrerrepresentado (cr18)}$$

$$n.º \text{ de pocillos positivos solo para cr21} + n.º \text{ de pocillos positivos solo para cr18}$$

<sup>c</sup> "Euploide" se asignó cuando la Pr estuvo por debajo de la región inclasificable; "T18" se asignó cuando la Pr estuvo por encima de la región inclasificable.

FIG. 20



Tipo	Casos	N.º de placa	Ensayo	N.º total de pocillos informativos	N.º de recuento de pocillos			Análisis de SPRT				
					Ambos positivos	Solo Cr 21	Solo Ref	m <sub>r</sub> <sup>a</sup>	P <sub>r</sub> <sup>a</sup>	Región inclasificable	Clasificación <sup>c</sup>	
Normal	N0196	1	cr21/cr1	148	78	71	77	0,52	0,480	0,489 - 0,584	E	
		1	cr21/cr18	152	63	79	73	0,44	0,520	0,487 - 0,584	U	
	N0198		SUMA	300	141	150	150	0,48	0,500	0,512 - 0,560	E	
		1	cr21/cr1	162	73	70	92	0,56	0,432	0,490 - 0,581	E	
		1	cr21/cr18	183	68	93	90	0,53	0,508	0,498 - 0,575	U	
			SUMA	345	141	163	182	0,55	0,472	0,517 - 0,557	E	
	N0230	1	cr21/cr1	162	58	81	81	0,45	0,500	0,490 - 0,581	U	
		1	cr21/cr18	165	45	81	81	0,41	0,491	0,489 - 0,580	U	
	N0231		SUMA	327	141	103	165	0,43	0,495	0,517 - 0,558	E	
		1	cr21/cr1	173	53	81	92	0,47	0,468	0,494 - 0,578	E	
		1	cr21/cr18	166	51	85	81	0,42	0,512	0,490 - 0,580	U	
			SUMA	339	104	166	173	0,45	0,490	0,514 - 0,557	E	
	N0262	1	cr21/cr1	186	97	101	85	0,64	0,543	0,503 - 0,575	U	
		1	cr21/cr18	176	111	91	85	0,71	0,517	0,504 - 0,557	U	
	T21			SUMA	362	208	192	170	0,68	0,530	0,522 - 0,557	U
			2	cr21/cr1	174	103	85	89	0,69	0,489	0,503 - 0,577	E
				SUMA	536	311	277	259	0,68	0,517	0,523 - 0,552	E
N0234			1	cr21/cr1	193	68	108	85	0,51	0,560	0,500 - 0,573	U
		1	cr21/cr18	178	65	92	86	0,50	0,517	0,496 - 0,576	U	
			SUMA	371	133	200	171	0,50	0,539	0,517 - 0,556	U	
		2	cr21/cr1	178	73	92	86	0,53	0,517	0,498 - 0,576	U	
			SUMA	549	206	292	257	0,51	0,532	0,524 - 0,549	U	
		2	cr21/cr18	178	62	95	83	0,47	0,534	0,495 - 0,576	U	
			SUMA	727	268	387	340	0,50	0,532	0,527 - 0,546	U	
		3	cr21/cr1	184	108	103	81	0,68	0,560	0,504 - 0,575	U	
		SUMA	911	376	490	421	0,54	0,538	0,529 - 0,545	U		
	3	cr21/cr18	198	102	120	78	0,63	0,606	0,505 - 0,572	T21		
	SUMA	1109	478	610	499	0,55	0,550	0,531 - 0,544	T21			

FIG. 21

N0232	1	cr21/cr1	173	109	103	70	0,63	0,595	0,500 - 0,577	T21
	1	cr21/cr18	191	82	110	81	0,55	0,576	0,501 - 0,574	T21
		<b>SUMA</b>	364	191	213	151	0,59	0,585	0,519 - 0,557	T21
N0349	1	cr21/cr1	158	62	83	75	0,44	0,525	0,489 - 0,582	U
	1	cr21/cr18	166	46	87	79	0,39	0,524	0,489 - 0,580	U
		<b>SUMA</b>	324	108	170	154	0,42	0,525	0,512 - 0,558	U
2	2	cr21/cr1	149	66	86	63	0,41	0,577	0,485 - 0,585	U
		<b>SUMA</b>	473	174	256	217	0,41	0,541	0,519 - 0,551	U
	2	cr21/cr18	168	43	100	68	0,34	0,595	0,487 - 0,580	T21
N0764		<b>SUMA</b>	641	217	356	285	0,40	0,555	0,523 - 0,546	T21
	1	cr21/cr1	182	82	106	76	0,53	0,582	0,498 - 0,575	T21
	1	cr21/cr18	178	83	97	81	0,56	0,545	0,498 - 0,576	U
N0890		<b>SUMA</b>	360	165	203	157	0,54	0,564	0,518 - 0,556	T21
	1	cr21/cr1	180	103	108	72	0,61	0,600	0,501 - 0,576	T21
	1	cr21/cr18	188	68	104	84	0,50	0,553	0,499 - 0,574	U
		<b>SUMA</b>	368	171	212	156	0,55	0,576	0,518 - 0,556	T21

a n.º de moléculas de referencia por pocillo

b Los valores de P<sub>r</sub> se calcularon mediante la ecuación:

$$n.º \text{ de pocillos positivos para el alelo sobrerrepresentado (cr21)} = \frac{n.º \text{ de pocillos positivos solo para cr21} + n.º \text{ de pocillos positivos solo para Ref}}{n.º \text{ de pocillos positivos solo para cr21} + n.º \text{ de pocillos positivos solo para Ref}}$$

c "Euploide" se asignó cuando la Pr estuvo por debajo de la región inclasificable; "T21" se asignó cuando la Pr estuvo por encima de la región inclasificable.

FIG. 21 (Cont.)



			N.º total corregido por Poisson		P <sub>r</sub>	Clasificación
			Cromosoma 21	Cr de referencia		
Euploide	Caso 1	Ensayo 1	37	33	0,52345751	Sin clase
		Ensayo 2	31	42	0,42506276	Sin clase
		Ensayo 3	29	37	0,44198785	Sin clase
		Ensayo 4	39	26	0,6008691	Sin clase
		Placa 1 subtotal	136	138	0,495821	Sin clase
		Ensayo 1	32	25	0,56605316	Sin clase
		Ensayo 2	27	24	0,53165424	Sin clase
		Ensayo 3	10	17	0,3681535	Sin clase
		Ensayo 4	24	18	0,5626764	Sin clase
		Placa 2 subtotal	93	84	0,52474	Sin clase
	Total		229	223	0,507182	Euploide
	Caso 2	Ensayo 1	45	35	0,56321585	Sin clase
		Ensayo 2	31	45	0,41245525	Sin clase
		Ensayo 3	32	22	0,59954763	Sin clase
		Ensayo 4	38	35	0,5228175	Sin clase
		Placa 1 subtotal	146	135	0,519153	Sin clase
		Ensayo 1	28	23	0,55272135	Sin clase
		Ensayo 2	23	22	0,51196923	Sin clase
		Ensayo 3	23	15	0,59688187	Sin clase
		Ensayo 4	29	26	0,5293349	Sin clase
		Placa 2 subtotal	102	85	0,545204	Sin clase
	Placa 1+2 subtotal		248	221	0,529583	Sin clase
		Ensayo 1	17	25	0,4122149	Sin clase
		Ensayo 2	25	16	0,6027299	Sin clase
		Ensayo 3	13	26	0,3382802	Sin clase
		Ensayo 4	27	24	0,5316872	Sin clase
		Placa 3 subtotal	82	91	0,475775	Sin clase
	Placa 1-3 subtotal		331	311	0,515069	Sin clase
		Ensayo 1	15	9	0,6269193	Sin clase
		Ensayo 2	14	15	0,4824155	Sin clase
		Ensayo 3	18	17	0,5146313	Sin clase
		Ensayo 4	21	18	0,5270080	Sin clase
		Placa 4 subtotal	69	60	0,532264	Sin clase
	Placa 1-4 subtotal		399	372	0,517942	Sin clase

FIG. 22A

# ES 2 820 866 T3

	Ensayo 1	10	13	0,4337957	<b>Sin clase</b>
	Ensayo 2	11	14	0,4390082	<b>Sin clase</b>
	Ensayo 3	7	16	0,3018022	<b>Sin clase</b>
	Ensayo 4	19	10	0,6579085	<b>Sin clase</b>
	Placa 5 subtotal	48	54	0,469951	<b>Sin clase</b>
	<b>Total</b>	<b>447</b>	<b>426</b>	<b>0,5123444</b>	<b>Euploide</b>
<b>T21</b>					
Caso 3	Ensayo 1	50	33	0,60012253	<b>Sin clase</b>
	Ensayo 2	56	47	0,54458055	<b>Sin clase</b>
	Ensayo 3	46	28	0,6196626	<b>Sin clase</b>
	Ensayo 4	56	36	0,6108731	<b>Sin clase</b>
	<b>Total</b>	<b>208</b>	<b>144</b>	<b>0,59079</b>	<b>T21</b>
Caso 4	Ensayo 1	56	36	0,61085513	<b>Sin clase</b>
	Ensayo 2	38	32	0,53906912	<b>Sin clase</b>
	Ensayo 3	40	24	0,62788068	<b>Sin clase</b>
	Ensayo 4	47	27	0,634511	<b>Sin clase</b>
	<b>Total</b>	<b>180</b>	<b>119</b>	<b>0,603476</b>	<b>T21</b>

FIG. 22B

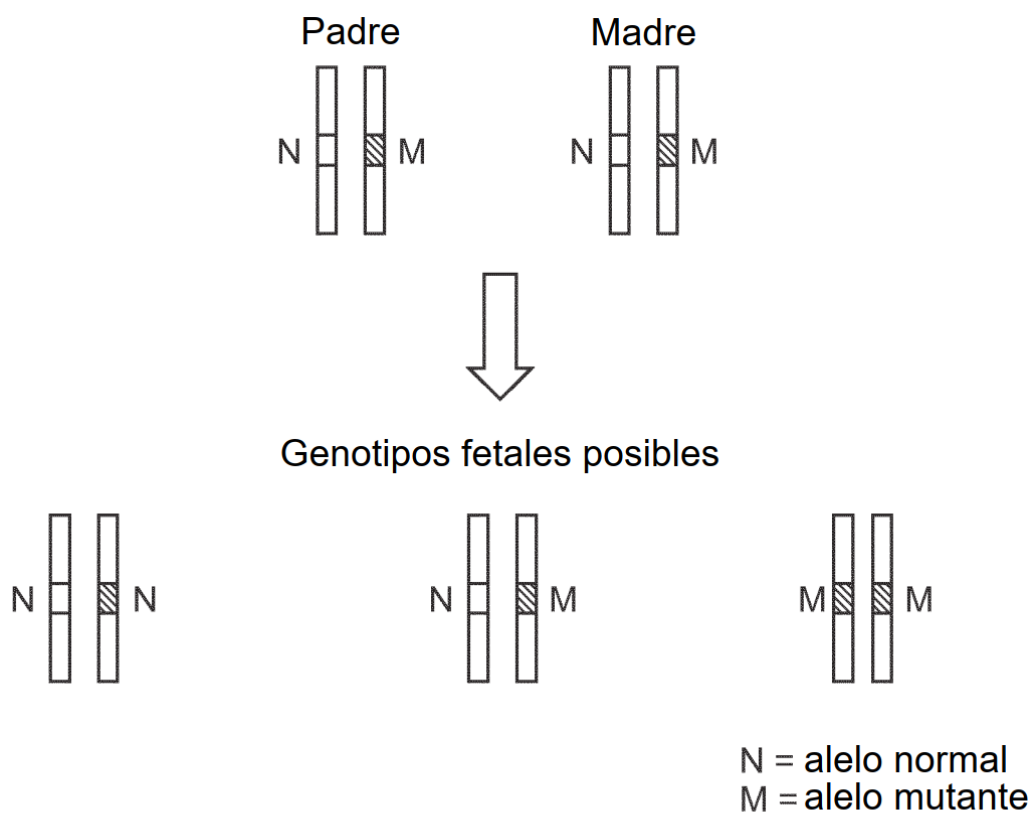


FIG. 23

Casos	Muestras	Esperados %	N.º de paneles	ZFY	ZFX	TODOS POS	ZFY NETO	ZFX NETO	Alelo Ref	m <sub>r</sub>	SPRT <sup>a</sup>
1	XY + XX (XX %)	50%	1	239	465	155	84	310	Y	0,37	XX
		25%	1	328	447	205	123	242	Y	0,56	XX
2	XY <sub>1</sub> + XY <sub>2</sub> (XY <sub>2</sub> %)	50%	1	414	405	220	194	185	X	0,85	XY
		25%	1	432	426	245	187	181	X	0,81	XY

FIG. 24A

Casos	N.º de paneles	ZFY	ZFX	TODOS POS	ZFY NETO	ZFX NETO	Alelo Ref	m <sub>r</sub>	SPRT
1	1	81	171	18	63	153	Y	0,11	XX
2	1	103	175	23	80	152	Y	0,14	XX
3	1	110	175	28	82	147	Y	0,16	XX
4	1	111	156	24	87	132	Y	0,16	XX
5	2	236	322	50	186	272	Y	0,17	XX
6	1	116	184	29	87	155	Y	0,16	XX
7	1	94	169	19	75	150	Y	0,13	XX
8	1	91	155	24	67	131	Y	0,13	XX
9	3	333	447	74	259	373	Y	0,16	XX
10	1	110	154	24	86	130	Y	0,16	XX
11	1	83	170	26	57	144	Y	0,11	XX
12	1	85	159	16	69	143	Y	0,12	XX

FIG. 24B

Casos	N.º de paneles	Genotipo materno <sup>a</sup>	Genotipo fetal <sup>a</sup>	% Fetal esperado	A	G	TODOS POS	A NETO	G NETO	% Fetal real <sup>b</sup>	Alelo Ref	m <sub>r</sub>	SPRT <sup>d</sup>
1	1	AG	GG	15%	264	318	116	148	202	9,6	A	0,42	GG
2	1	AG	GG	15%	208	276	81	127	195	14,3	A	0,32	GG
3	1	AG	AG	15%	248	251	87	161	164	15,4	A	0,39	AG
4	1	AG	GG	15%	158	240	51	107	189	13,7	A	0,23	GG
5	1	AG	AG	15%	226	245	70	156	175	16,8	A	0,35	AG
6	1	AG	AG	15%	319	319	130	189	189	18,1	A	0,54	AG
7	1	AG	AG	15%	254	279	94	160	185	19,8	A	0,40	AG
8	1	AG	GG	15%	296	363	144	152	219	15,7	A	0,49	GG
9	1	AG	AG	15%	279	285	100	179	185	13,0	A	0,45	AG
10	1	AG	AA	15%	337	278	120	217	158	NA <sup>c</sup>	G	0,45	AA
1	1	AG	GG	20%	248	295	96	152	199	13,0	A	0,39	GG
2	1	AG	GG	20%	191	289	79	112	210	19,5	A	0,29	GG
3	1	AG	AG	20%	248	218	75	173	143	18,2	G	0,34	U
4	1	AG	GG	20%	158	257	59	99	198	25,9	A	0,23	GG
5	1	AG	AG	20%	231	215	71	160	144	20,3	G	0,33	AG
6	1	AG	AG	20%	281	283	98	183	185	26,3	A	0,46	AG
7	1	AG	AG	20%	256	231	88	168	143	22,7	G	0,36	AG
8	1	AG	GG	20%	246	311	104	142	207	23,0	A	0,39	GG
9	1	AG	AG	20%	267	249	83	184	166	21,9	G	0,39	AG
10	1	AG	AA	20%	303	203	77	226	126	NA <sup>c</sup>	G	0,31	AA
1	1	AG	GG	25%	244	304	99	145	205	15,7	A	0,38	GG
2	1	AG	GG	25%	209	314	88	121	226	24,9	A	0,32	GG
3	1	AG	AG	25%	204	212	55	149	157	28,0	A	0,31	AG
4	1	AG	GG	25%	185	290	76	109	214	25,9	A	0,28	GG
5	1	AG	AG	25%	233	243	77	156	166	24,5	A	0,36	AG
6	1	AG	AG	25%	292	305	110	182	195	27,3	A	0,48	AG
7	1	AG	AG	25%	261	275	95	166	180	22,5	A	0,42	AG
8	1	AG	GG	25%	236	327	96	140	231	26,7	A	0,37	GG
9	1	AG	AG	25%	271	280	104	167	176	25,2	A	0,44	AG
10	1	AG	AA	25%	243	154	46	197	108	NA <sup>c</sup>	G	0,22	AA
1	1	AG	GG	50%	136	293	57	79	236	33,6	A	0,20	GG
2	1	AG	GG	50%	163	358	78	85	280	45,8	A	0,24	GG
3	1	AG	AG	50%	200	208	53	147	155	55,7	A	0,30	AG
4	1	AG	GG	50%	137	356	66	71	290	53,0	A	0,18	GG
5	1	AG	AG	50%	247	246	91	156	155	57,7	G	0,39	AG
6	1	AG	AG	50%	218	245	68	150	177	53,8	A	0,34	AG
7	1	AG	AG	50%	255	282	103	152	179	51,3	A	0,41	AG
8	1	AG	GG	50%	158	398	83	75	315	57,6	A	0,23	GG
9	1	AG	AG	50%	235	225	74	161	151	55,7	G	0,35	AG
10	1	AG	AA	50%	317	106	35	282	71	NA <sup>c</sup>	G	0,15	AA

<sup>a</sup> A = Alelo mutado, G = Alelo silvestre.

<sup>b</sup> El % de ADN fetal real se determina mediante el ensayo ZFY/X digital.

<sup>c</sup> El feto es femenino y el ensayo ZFY/X no se puede aplicar.

<sup>d</sup> La SPRT se realizó con el % fetal determinado mediante el ensayo ZFY/X excepto en el caso 10; U, sin clasificar.

FIG. 25

Caso	Genotipo materno <sup>a</sup>	Genotipo fetal <sup>a</sup>	N.º de paneles	M	N	TODOS POS	M NETO	N NETO	% Fetal esperado	1 <sup>b</sup>			2 <sup>c</sup>			Alelo Ref	m <sub>r</sub>
1	MN	MM	3	846	763	273	573	490	5%				MM	MM	7,8%	N	0,40
	MN	MM	3	863	748	289	574	459	10%				MM	MM	9,6%	N	0,39
	MN	MM	1	323	219	79	244	140	25%				MM	MM	24,1%	N	0,34
	MN	MM	1	418	160	92	326	68	50%				MM	MM	57,2%	N	0,23
	MN	MN	9	2559	2476	887	1672	1589	5%				U	U	3,4%	N	0,45
2	MN	MN	3	895	907	360	535	547	10%				MN	MN	9,7%	N	0,49
	MN	MN	1	277	263	97	180	166	25%				MN	MN	24,2%	N	0,42
	MN	MN	1	273	295	101	172	194	50%				MN	MN	58,0%	N	0,44

<sup>a</sup> M = con delección de CTTT, N = sin delección de CTTT.

<sup>b</sup> La SPRT se realizó con el % esperado.

<sup>c</sup> La SPRT se realizó con el % fetal determinado mediante el ensayo ZFY/X; U, sin clasificar.

FIG. 26A

Caso	N.º de paneles	Genotipo materno <sup>a</sup>	Genotipo fetal <sup>a</sup>	% Fetal esperado	M	N	TODOS POS	M NETO	N NETO	% fetal real <sup>b</sup>	Alelo Ref	m <sub>r</sub>	SPRT <sup>c</sup>
1	1	MN	MN	20%	203	200	58	145	142	29,0	N	0,30	MN
2	1	MN	MN	20%	149	204	39	110	165	29,7	M	0,22	NN
3	1	MN	MN	20%	128	212	31	97	181	30,7	M	0,18	NN
4	1	MN	MN	20%	173	196	45	128	151	27,0	M	0,26	MN
5	1	MN	MN	20%	147	230	47	100	183	24,3	M	0,21	NN
6	1	MN	MN	20%	199	206	45	154	161	19,2	M	0,30	MN
7	1	MN	MN	20%	165	237	58	107	179	18,7	M	0,24	NN
8	1	MN	MN	20%	203	180	43	160	137	23,4	N	0,27	MN
9	1	MN	MN	20%	173	241	61	112	180	16,9	M	0,26	NN
10	1	MN	MN	20%	174	190	52	122	138	16,8	M	0,26	U
11	1	MN	MN	20%	167	242	49	118	193	17,7	M	0,25	NN
12	1	MN	MN	20%	213	199	47	166	152	21,6	N	0,30	MN

<sup>a</sup> M = con delección de CTTT, N = sin delección de CTTT.

<sup>b</sup> El % de ADN fetal real se determinó mediante el ensayo ZFXY digital.

<sup>c</sup> La SPRT se realizó con el % fetal determinado mediante el ensayo ZFY/X; U, sin clasificar.

FIG. 26B

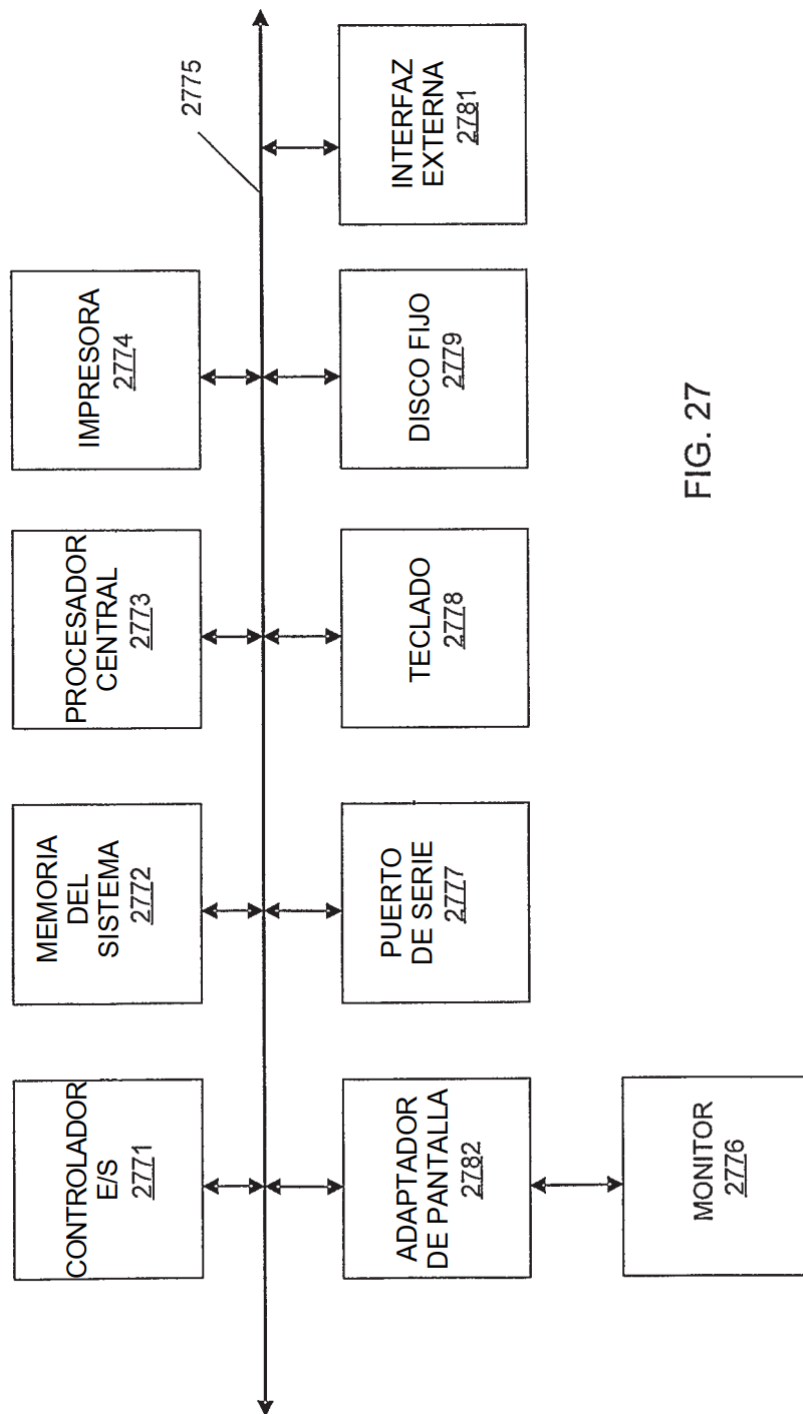


FIG. 27