

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 739 430**

51 Int. Cl.:

C12Q 1/6837 (2008.01)

C12Q 1/6825 (2008.01)

C12Q 1/6827 (2008.01)

C12Q 1/6809 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **19.08.2014 PCT/US2014/051763**

87 Fecha y número de publicación internacional: **26.02.2015 WO15026873**

96 Fecha de presentación y número de la solicitud europea: **19.08.2014 E 14838189 (0)**

97 Fecha y número de publicación de la concesión europea: **03.07.2019 EP 3036545**

54 Título: **Ensayos para la detección de una sola molécula y uso de los mismos**

30 Prioridad:

19.08.2013 US 201361867559 P
19.08.2013 US 201361867554 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
31.01.2020

73 Titular/es:

SINGULAR BIO INC. (100.0%)
455 Mission Bay Boulevard South
San Francisco, CA 94158, US

72 Inventor/es:

FEHR, ADRIAN NIELSEN;
COLLINS, PATRICK JAMES;
HERSCHLEB, JILL LYNDON y
JONES, HYWEL BOWDEN

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 739 430 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

Ensayos para la detección de una sola molécula y uso de los mismos

Antecedentes de la invención

5 La invención se refiere a métodos para detectar una variación genética en una muestra genética de un sujeto. La detección de una variación genética es importante en muchos aspectos de la biología humana.

Compendio

La invención se refiere a métodos para detectar una variación genética en una muestra genética de un sujeto. La invención se refiere adicionalmente a métodos para detectar una variación genética en una muestra genética de un sujeto utilizando sondas marcadas y contando el número de marcas en las sondas.

10 **Breve descripción de los dibujos**

La Figura 1 representa miembros de una matriz ilustrativa que comprenden compañeros de unión, etiquetas, etiquetas de afinidad, sondas de etiquetado, conjuntos de sondas y/o conjuntos de sondas ligadas descritos en la presente memoria sobre un sustrato.

15 La Figura 2 representa un histograma normalizado de la intensidad de la señal medida tanto de muestras con una sola marca como en anticuerpos con múltiples marcas.

La Figura 3 representa los perfiles de blanqueo promedio de varias marcas.

Las Figuras 4-13 muestran los gráficos de intensidad de marcas integradas a lo largo del tiempo para varias marcas Alexa 488.

20 La Figura 14 representa el espectro de excitación y el espectro de emisión a través de una operación convencional cuando se logra la excitación de un fluoróforo iluminando con una banda espectral estrecha alineada con los máximos de absorción de esa especie.

La Figura 15 representa el espectro de excitación y el espectro de emisión a través de la interrogación con varios colores de excitación y bandas de emisión recolectadas diferentes de (o además de) el caso para la operación convencional.

25 La Figura 16 muestra los resultados cuando se recoge la luz de estas diversas configuraciones para la generación de imágenes, p. ej., varios filtros de emisión, y se compara con los valores de calibración para los fluoróforos de interés.

La Figura 17 muestra los resultados recopilados con varias referencias, incluidos aquellos con un perfil de emisión plano (Contaminante 1; triángulos) o un perfil ponderado en azul (Contaminante 2; estrellas).

30 La Figura 18 representa bandas de excitación significativamente diferentes de dos fluoróforos.

La Figura 19 representa un diagrama de flujo del sistema ilustrativo.

La Figura 20 representa un diagrama de flujo del sistema ilustrativo que incluye varios métodos para analizar datos.

Las Figuras 21-46 representan conjuntos de sondas ilustrativos descritos en la presente memoria.

35 Las Figuras 47 y 48 muestran los patrones de fluorescencia resultantes cuando los productos contienen secuencias de etiquetas de afinidad únicas y el sustrato subyacente contiene complementos para cada una de las etiquetas de afinidad únicas dentro de la misma ubicación (p. ej., como el mismo miembro) sobre un sustrato.

Las Figuras 49 y 51 muestran los patrones de fluorescencia resultantes cuando diferentes productos contienen secuencias de etiquetas de afinidad idénticas y el sustrato subyacente contiene el complemento de la etiqueta de afinidad.

40 Las Figuras 50 y 52 muestran las ubicaciones ampliadas de las Figuras 49 y 51, respectivamente.

Las Figuras 53 y 54 muestran los patrones de fluorescencia resultantes cuando los productos contienen secuencias de etiquetas de afinidad únicas y el sustrato subyacente tiene una ubicación (p. ej., como un miembro) que contiene el complemento de un complemento de etiqueta de afinidad, y otra ubicación separada (p. ej., como otro miembro) que contiene el complemento de la otra etiqueta de afinidad.

45 La Figura 55 representa dos conjuntos de sondas; un conjunto de sondas para el Locus 1 y un conjunto de sondas para el Locus 2 - aunque como se mencionó anteriormente, se pueden diseñar múltiples conjuntos de sondas para cada locus genómico.

La Figura 56 representa el flujo de trabajo de procedimiento que se aplicaría a la colección de conjuntos de sondas.

La Figura 57 representa una versión modificada del flujo de trabajo de procedimiento ilustrado en la Figura 56.

La Figura 58 proporciona un ejemplo de cómo las sondas producto para el Locus 1 y el Locus 2 pueden etiquetarse con diferentes moléculas marcadoras.

5 La Figura 59 proporciona evidencia de que las sondas producto que representan una multitud de ubicaciones genómicas para un locus se pueden generar de una manera específica de enzima ligasa utilizando el procedimiento de hibridación-ligación.

La Figura 60 proporciona datos que indican que los conjuntos de sondas se pueden utilizar para detectar cambios relativos en el estado del número de copias.

10 La Figura 61 proporciona evidencia de que se pueden utilizar mezclas de sondas producto para generar datos cuantitativos de micromatrices.

Las Figuras 62-64 ilustran modificaciones del procedimiento general descrito en las Figuras 55 a 58.

La Figura 65 representa una realización adicional del procedimiento modificado descrito en la Figura 62.

La Figura 66 representa otra realización más del procedimiento representado en la Figura 65.

15 La Figura 67 representa conjuntos de sondas ilustrativos utilizados en los métodos descritos en la presente memoria.

La Figura 68 representa conjuntos de sondas ilustrativos utilizados en los métodos descritos en la presente memoria cuando se someten a ensayo translocaciones que tienen puntos de ruptura conocidos.

La Figura 69 representa conjuntos de sondas ilustrativos utilizados en los métodos descritos en la presente memoria cuando se eligen como diana mutaciones en los SNP.

20 **Descripción detallada de la invención**

Los métodos descritos en la presente memoria pueden emplear, a menos que se indique lo contrario, técnicas convencionales y descripciones de biología molecular (incluidas las técnicas recombinantes), biología celular, bioquímica y tecnología de micromatrices y secuenciación, que están dentro del conocimiento práctico de los expertos en la técnica. Tales técnicas convencionales incluyen síntesis de matriz polimérica, hibridación y ligación de oligonucleótidos, secuenciación de oligonucleótidos y detección de hibridación utilizando una marca. Se pueden tener ilustraciones específicas de técnicas adecuadas mediante la referencia a los ejemplos en la presente memoria. Sin embargo, también se pueden utilizar, por supuesto, procedimientos convencionales equivalentes. Tales técnicas y descripciones convencionales se pueden encontrar, por ejemplo, en Kimmel y Oliver, DNA Microarrays (2006) Elsevier; Campbell, DNA Microarray, Synthesis and Synthetic DNA (2012) Nova Science; Bowtell y Sambrook, DNA Microarrays: Molecular Cloning Manual (2003) Cold Spring Harbor Laboratory Press. Antes de que se describan las presentes composiciones, herramientas y métodos de investigación, se debe entender que esta invención no está limitada a los métodos, composiciones, objetivos y usos específicos descritos, ya que, por supuesto, estos pueden variar. También se debe entender que la terminología utilizada en la presente memoria tiene el propósito de describir solo aspectos concretos y no se pretende que limite el alcance de la presente invención, que estará limitada solamente por las reivindicaciones adjuntas.

La presente invención se define en la reivindicación independiente 1 y ciertas características opcionales de la misma se definen en las reivindicaciones dependientes. En la medida en que los términos "invención", "ejemplo", "aspecto" y "realización" se utilizan en la presente memoria, esto se interpretará de tal manera que la única protección buscada es para la invención reivindicada.

40 La invención se refiere a métodos para detectar una variación genética en una muestra genética de un sujeto. La variación genética en la presente memoria puede incluir, pero no se limita a, una o más sustituciones, inversiones, inserciones, deleciones o mutaciones en secuencias de nucleótidos (p. ej., ADN y ARN) y proteínas (p. ej., péptido y proteína), uno o más alelos raros, polimorfismos, polimorfismos de un solo nucleótido (SNP), polimorfismos genéticos a gran escala, tales como inversiones y traslocaciones, diferencias en la abundancia y/o el número de copias (p. ej., variantes del número de copias, CNV) de una o más moléculas de nucleótidos (p. ej., ADN), trisomía, monosomía y reordenamientos genómicos. En algunas realizaciones, la variación genética puede estar relacionada con metástasis, presencia, ausencia y/o riesgo de una enfermedad, tal como cáncer, variabilidad farmacocinética, toxicidad del fármaco, eventos adversos, recurrencia y/o presencia, ausencia o riesgo de rechazo del trasplante de órganos en el sujeto. Por ejemplo, los cambios de número de copias en el gen *HER2* influyen en que un paciente con cáncer de mama responda o no al tratamiento con Herceptin. De manera similar, la detección de un aumento en el número de copias del cromosoma 21 (o 18, o 13, o cromosomas sexuales) en la sangre de una mujer embarazada puede ser utilizada como un diagnóstico no invasivo para el síndrome de Down en un feto. Un ejemplo adicional es la detección de alelos de un órgano trasplantado que no están presentes en el genoma del receptor - la verificación de la frecuencia, o el número de copias, de estos alelos puede identificar signos de rechazo potencial del órgano. Se

pueden utilizar varios métodos para detectar tales cambios (p. ej., rtPCR, secuenciación y micromatrices). Uno de los métodos consiste en contar las moléculas individuales marcadas para detectar la presencia de una mutación (p. ej., mutación EGFR en el cáncer) o un exceso de una secuencia o región genómica específica (p. ej., el cromosoma 21 en el síndrome de Down). El recuento de moléculas individuales se puede realizar de varias maneras, siendo una lectura común el depósito de las moléculas sobre una superficie y la generación de imágenes.

Además, la variación genética puede consistir en mutaciones genéticas de novo, tales como mutaciones de base única o múltiple, translocaciones, amplificaciones y deleciones subcromosómicas y aneuploidía. En algunas realizaciones, la variación genética puede significar una secuencia de nucleótidos alternativa en un locus genético que puede estar presente en una población de individuos y que incluye sustituciones, inserciones y deleciones de nucleótidos con respecto a otros miembros de la población. En realizaciones adicionales, la variación genética puede ser aneuploidía. En otras realizaciones adicionales, la variación genética puede ser trisomía 13, trisomía 18, trisomía 21, aneuploidía de X (p. ej., trisomía XXX y trisomía XXY), o aneuploidía de Y (p. ej., trisomía XYY). En realizaciones adicionales, la variación genética puede estar en la región 22q11.2, 1q21.1, 9q34, 1p36, 4p, 5p, 7q11.23, 11q24.1, 17p, 11p15, 18q o 22q13. En realizaciones adicionales, la variación genética puede ser una microdelección o una microamplificación.

En algunas realizaciones, la detección, descubrimiento, determinación, medición, evaluación, recuento y valoración de la variación genética se utilizan de manera indistinta e incluyen determinaciones cuantitativas y/o cualitativas, que incluyen, por ejemplo, la identificación de la variación genética, la determinación de la presencia y/o la ausencia de la variación genética, y la cuantificación de la variación genética. En realizaciones adicionales, los métodos de la presente descripción pueden detectar múltiples variaciones genéticas. El término "y/o" utilizado en la presente memoria se define para indicar cualquier combinación de los componentes. Además, las formas singulares "un", "uno", "una", "el" y "la" pueden incluir, adicionalmente, referentes plurales, a menos que el contexto indique claramente lo contrario. Así, por ejemplo, la referencia a "una región nucleotídica" se refiere a una, más de una, o mezclas de tales regiones, y la referencia a "un ensayo" puede incluir referencias a etapas y métodos equivalentes conocidos por los expertos en la técnica, etcétera.

"Muestra" significa una cantidad de material de una fuente biológica, ambiental, médica o del paciente en la que se busca la detección, medición o marcaje de los ácidos nucleicos, péptidos y/o proteínas diana. Por un lado, se pretende que incluya un espécimen o cultivo (p. ej., cultivos microbiológicos). Por otro lado, se pretende que incluya muestras biológicas y ambientales. Una muestra puede incluir un espécimen de origen sintético. Las muestras ambientales incluyen material ambiental, tal como muestras de materia superficial, suelo, agua e industriales, así como muestras obtenidas de instrumentos, aparatos, equipos, utensilios, elementos desechables y no desechables de procesamiento de alimentos y productos lácteos. "Muestra genética" puede ser cualquier muestra líquida o sólida con información biológica hereditaria y/o no hereditaria codificada en las secuencias de nucleótidos de los ácidos nucleicos. La muestra se puede obtener de una fuente, que incluye, pero no se limitan a, sangre completa, suero, plasma, orina, saliva, sudor, materia fecal, lágrimas, líquido intestinal, muestras de membrana mucosa, tejido pulmonar, tumores, órganos transplantados, feto, y/u otras fuentes. Las muestras genéticas pueden ser de un animal, incluido un ser humano, líquidas, sólidas (p. ej., heces) o tejido. Las muestras genéticas pueden incluir materiales tomados de un paciente que incluyen, pero no se limitan a, cultivos, sangre, saliva, líquido cefalorraquídeo, líquido pleural, leche, linfa, esputo, semen, productos aspirados con agujas y similares. Además, la muestra genética puede ser un material genético fetal de una muestra de sangre materna. El material genético fetal se puede aislar y separar de la muestra de sangre materna. La muestra genética puede ser una mezcla de material genético fetal y materno. Además, la muestra genética puede incluir secuencias genéticas aberrantes derivadas de la formación de tumores o metástasis, y/o firmas de ADN de donantes presentes en un receptor de trasplante. En realizaciones adicionales, cuando la muestra genética es plasma, el método puede comprender aislar el plasma de una muestra de sangre del sujeto. En realizaciones adicionales, cuando la muestra genética es suero, el método puede comprender aislar el suero de una muestra de sangre del sujeto. En otras realizaciones adicionales, cuando la muestra genética es una muestra de ADN sin células (ADNsc), el método comprende adicionalmente aislar la muestra de ADN sin células de una muestra obtenida de la fuente descrita en la presente memoria. La muestra de ADN sin células en la presente memoria significa una población de moléculas de ADN que circulan libremente en el torrente sanguíneo, fuera de cualquier célula u órgano. En el caso de un embarazo, el ADN sin células de la madre contiene una mezcla de ADN materno y ADN fetal. No se debe considerar que estos ejemplos sean limitantes de los tipos de muestra aplicables a la presente invención.

En algunas realizaciones, el método de la presente descripción puede comprender seleccionar y/o aislar locus o loci genéticos de interés, y cuantificar la cantidad de cada locus presente (por ejemplo para determinar el número de copias) y/o las cantidades relativas de diferentes variantes de locus (p. ej., dos alelos de una secuencia de ADN dada). La región, región de interés, locus, o locus de interés en referencia a un genoma o polinucleótido diana utilizado en la presente memoria significan una sub-región o segmento contiguo del genoma o polinucleótido diana. Como se emplea en la presente memoria, región, regiones de interés, locus, o locus de interés en una molécula de nucleótido se pueden referir a la posición de un nucleótido, un gen o una porción de un gen en un genoma, incluido el ADN mitocondrial u otro ADN no cromosómico, o se pueden referir a cualquier porción contigua de la secuencia genómica, esté o no dentro de un gen o esté o no asociada al mismo. Una región, región de interés, locus, o locus de interés en una molécula de nucleótido pueden tener desde un solo nucleótido hasta un segmento de unos pocos cientos o unos pocos miles de nucleótidos de longitud o más. En algunas realizaciones, una región o locus de interés

pueden tener una secuencia de referencia asociada a ellos. La "secuencia de referencia" utilizada en la presente memoria denota una secuencia con la que se compara un locus de interés en un ácido nucleico. En ciertas realizaciones, una secuencia de referencia se considera una secuencia de "tipo salvaje" para un locus de interés. Un ácido nucleico que contiene un locus de interés que tiene una secuencia que varía de una secuencia de referencia para el locus de interés a veces se denomina variación "polimórfica" o "mutante" o "genética". Un ácido nucleico que contiene un locus de interés que tiene una secuencia que no varía de una secuencia de referencia para el locus de interés a veces se denomina "no polimórfico" o de "tipo salvaje" o "variación no genética". En ciertas realizaciones, un locus de interés puede tener más de una secuencia de referencia distinta asociada con él (p. ej., cuando se sabe que un locus de interés tiene un polimorfismo que debe considerarse un tipo normal o salvaje). En algunas realizaciones, el método de la presente descripción también puede comprender elegir y/o aislar péptidos o péptidos de interés, y calificar la cantidad de cada péptido presente y/o cantidades relativas de diferentes péptidos.

En realizaciones adicionales, la región de interés descrita en la presente memoria puede incluir una "secuencia variante genética consenso", que se refiere a la secuencia de ácido nucleico o proteína, cuyos nucleótidos o aminoácidos se sabe que aparecen con alta frecuencia en una población de individuos que portan el gen que codifica una proteína que no funciona normalmente, o en la que el ácido nucleico en sí no funciona normalmente. Además, la región de interés descrita en la presente memoria puede incluir una "secuencia génica normal consenso" que se refiere a una secuencia de ácido nucleico, cuyo ácido nucleico se produce en sus respectivas posiciones con alta frecuencia en una población de individuos que portan el gen que codifica una proteína que no funciona normalmente, o que por sí misma no funciona normalmente. En realizaciones adicionales, la región de control que no es la región de interés o la secuencia de referencia descrita en la presente memoria puede incluir una "secuencia normal consenso" que se refiere a la secuencia de ácido nucleico o proteína, cuyos nucleótidos o aminoácidos se sabe que aparecen con una alta frecuencia en una población de individuos que portan el gen que codifica una proteína que funciona normalmente o en la que el ácido nucleico en sí tiene una función normal.

Los métodos descritos en la presente memoria pueden producir mediciones altamente precisas de variación genética. Un tipo de variación descrito en la presente memoria incluye la abundancia relativa de dos o más loci genómicos distintos. En este caso, los loci pueden ser pequeños (p. ej., tan pequeños como aproximadamente 300, 250, 200, 150, 100 o 50 nucleótidos o menos), de tamaño moderado (p. ej., de 1.000, 10.000, 100.000 o un millón de nucleótidos), y tan grande como una porción de un brazo cromosómico o el cromosoma completo o conjuntos de cromosomas. Los resultados de este método pueden determinar la abundancia de un locus con respecto a otro. La precisión y la exactitud de los métodos de la presente descripción pueden permitir la detección de cambios muy pequeños en el número de copias (tan bajo como aproximadamente 25, 10, 5, 4, 3, 2, 1, 0,5, 0,1, 0,05, 0,02 o 0,01% o menos), que permite la identificación de una firma muy diluida de variación genética. Por ejemplo, se puede encontrar una firma de aneuploidía fetal en una muestra de sangre materna donde la aberración genética fetal se diluye con la sangre materna, y un número de copias observable de un cambio de aproximadamente 2% es indicativo de trisomía fetal.

Como se emplea en la presente memoria, el término "aproximadamente" significa modificar, por ejemplo, las longitudes de las secuencias de nucleótidos, los grados de error, las dimensiones, la cantidad de un ingrediente en una composición, las concentraciones, los volúmenes, la temperatura del procedimiento, el tiempo del procedimiento, los rendimientos, las velocidades de flujo, las presiones, y valores similares, y sus intervalos, se refiere a la variación en la cantidad numérica que puede ocurrir, por ejemplo, a través de procedimientos típicos de medición y manejo utilizados para la preparación de compuestos, composiciones, productos concentrados o formulaciones de uso; por error inadvertido en estos procedimientos; por diferencias en la fabricación, fuente o pureza de los materiales de partida o ingredientes utilizados para llevar a cabo los métodos; y consideraciones similares. El término "aproximadamente" también abarca cantidades que difieren debido al envejecimiento, por ejemplo, de una composición, formulación o cultivo celular con una concentración inicial o mezcla concretas, y cantidades que difieren debido a la mezcla o procesamiento de una composición o formulación con una concentración inicial o mezcla concretas. Si se modifican por el término "aproximadamente", las reivindicaciones adjuntas incluyen equivalentes a estas cantidades. El término "aproximadamente" se puede referir adicionalmente a un intervalo de valores que son similares al valor de referencia establecido. En ciertas realizaciones, el término "aproximadamente" se refiere a un intervalo de valores que se encuentran dentro de 50, 25, 10, 9, 8,7, 6, 5,4, 3, 2, 1 por ciento o menos del valor de referencia establecido.

En algunas realizaciones, el sujeto puede ser un sujeto gestante, ser humano, un sujeto con un alto riesgo de una enfermedad genética (p. ej., cáncer), todas las diversas familias de animales domésticos, así como animales silvestres o salvajes. En algunas realizaciones, la variación genética puede ser una variación genética en el feto de un sujeto gestante (p. ej., variantes del número de copias y aneuploidía en el feto). En algunas realizaciones, el sujeto es un sujeto gestante, y la variación genética es una variación en el feto del sujeto gestante en una región seleccionada del grupo que consiste en 22q11.2, 1q21.1, 9q34, 1p36, 4p, 5p 7q11.23, 11q24.1, 17p, 11p15, 18q y 22q13, (p. ej., una mutación y/o un cambio en el número de copias en cualquiera de las regiones 22q11.2, 1q21.1, 9q34, 1p36, 4p, 5p, 7q11.23, 11q24.1, 17p, 11p15, 18q y 22q13). El feto descrito en la presente memoria significa un vástago no nacido de un humano u otro animal. En algunas realizaciones, el feto puede ser el vástago de más de 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 semanas después de la concepción. En realizaciones adicionales, el feto puede ser un vástago concebido por implantes, fertilización in vitro, embarazos múltiples o gemelación. En realizaciones adicionales, el feto puede ser parte de un par de gemelos (idénticos o no idénticos), o

un trío de trillizos (idénticos o no idénticos).

Las invenciones de acuerdo con algunas realizaciones abarcan al menos dos componentes principales: un ensayo para la identificación selectiva de loci genómicos y una tecnología para cuantificar estos loci con alta precisión. El ensayo puede incluir métodos para marcar y/o aislar selectivamente una o más secuencias de ácido nucleico, de tal manera que la etapa de marcaje en sí sea suficiente para producir moléculas (definidas como "sondas producto", "conjunto de sondas ligadas", "conjuntos de sondas conjugadas", "sondas ligadas", "sondas conjugadas", o "moléculas marcadas" en esta invención) que contienen toda la información necesaria para la identificación de una secuencia concreta en el contexto de un ensayo concreto. Por ejemplo, el ensayo puede comprender poner en contacto, unir y/o hibridar sondas con una muestra, ligar y/o conjugar las sondas, amplificar opcionalmente las sondas ligadas/conjugadas, e inmovilizar las sondas en un sustrato. En algunas realizaciones, los ensayos y métodos descritos en la presente memoria se pueden realizar en una muestra de una sola entrada en paralelo como un ensayo múltiplex como se describe en la presente memoria.

La sonda producto, el conjunto de sondas ligadas, el conjunto de sondas conjugadas, las sondas ligadas, las sondas conjugadas y las moléculas marcadas pueden ser moléculas únicas, contiguas que resultan del rendimiento de la acción enzimática sobre un conjunto de sondas, tal como un ensayo. En una sonda producto o una molécula marcada, una o más sondas individuales de un conjunto de sondas se pueden modificar covalentemente de manera que formen una especie molecular distinta, singular en comparación con las sondas o los conjuntos de sondas. Como resultado, las sondas producto o una molécula marcada pueden ser químicamente diferentes y, por lo tanto, se pueden identificar, contar, aislar o manipular adicionalmente de las sondas o conjuntos de sondas.

Por ejemplo, las sondas producto pueden contener una o más marcas de identificación y una o más etiquetas de afinidad para su aislamiento y/o inmovilización. En algunas realizaciones, no es necesario realizar modificaciones adicionales de las sondas producto (p. ej., determinación de la secuencia de ADN). En algunas realizaciones, no se requieren interrogaciones adicionales de la secuencia de ADN. Los sondas producto que contienen las marcas se pueden contar directamente, típicamente después de una etapa de inmovilización sobre un sustrato sólido. Por ejemplo, las marcas de fluoróforos orgánicos se utilizan para marcar sondas producto, y las sondas producto se cuentan directamente inmovilizando las sondas producto en un sustrato de vidrio y generando a continuación imágenes a través de un microscopio fluorescente y una cámara digital. En otras realizaciones, la marca se puede extinguir o eliminarse selectivamente dependiendo de si la molécula marcada ha interactuado con su locus genómico complementario. En realizaciones adicionales, dos marcas en porciones opuestas de la sonda producto pueden trabajar en concierto para suministrar una señal de transferencia de energía de resonancia de fluorescencia (FRET) dependiendo de si la molécula marcada ha interactuado con su locus genómico complementario. Para un locus genómico dado, las sondas de marcaje que contienen las marcas se diseñan para cualquier región de secuencia dentro de ese locus. También se puede diseñar un conjunto de múltiples sondas de marcaje con la misma o diferentes marcas para un solo locus genómico. En este caso, una sonda puede aislar y etiquetar selectivamente una región diferente dentro de un locus particular, o regiones solapantes dentro de un locus. En algunas realizaciones, las sondas producto que contienen etiquetas de afinidad se inmovilizan sobre el sustrato a través de las etiquetas de afinidad. Por ejemplo, las etiquetas de afinidad se utilizan para inmovilizar sondas producto sobre el sustrato, y las sondas producto que contienen las etiquetas de afinidad se cuentan directamente. Para un locus genómico dado, las sondas de etiquetado que contienen las etiquetas de afinidad se diseñan para cualquier región de secuencia dentro de ese locus. También se puede diseñar un conjunto de múltiples sondas de etiquetado con etiquetas de afinidad iguales o diferentes para un solo locus genómico. En este caso, una sonda puede aislar y etiquetar selectivamente una región diferente dentro de un locus particular, o regiones solapantes dentro de un locus.

En un aspecto, los métodos de la presente descripción pueden comprender poner en contacto conjuntos de sondas descritos en la presente memoria con la muestra genética descrita en la presente memoria. En algunas realizaciones, los métodos de la presente descripción pueden comprender poner en contacto múltiples conjuntos de sondas, tales como conjuntos de sondas primero y segundo, con la muestra genética. En realizaciones adicionales, cada uno de los conjuntos de sondas comprende una sonda de marcaje y una sonda de etiquetado. Por ejemplo, el primer conjunto de sondas comprende una primera sonda de marcaje y una primera sonda de etiquetado, y el segundo conjunto de sondas comprende una segunda sonda de marcaje y una segunda sonda de etiquetado.

El contacto de los conjuntos de sondas con la muestra genética se puede realizar simultáneamente o después de hibridar, ligar, amplificar y/o inmovilizar las sondas. Además, el contacto de los conjuntos de sondas con la muestra genética se puede realizar simultáneamente o antes de hibridar, ligar, amplificar y/o inmovilizar las sondas.

Para un locus o región genómica dados de una molécula de nucleótido en la muestra genética, una única secuencia de ácido nucleico dentro de ese locus, o múltiples secuencias de ácido nucleico dentro de ese locus pueden ser interrogadas y/o cuantificadas a través de la creación de sondas producto. Las secuencias interrogadas dentro de un locus genómico pueden ser distintas y/o solapantes, y pueden contener o no polimorfismos genéticos. Una sonda producto se forma mediante el diseño de uno o más oligonucleótidos llamados "conjunto de sondas". Por ejemplo, la sonda producto se puede formar ligando el conjunto de sondas mediante ligación de las sondas en el conjunto de sondas. Un conjunto de sondas comprende al menos una sonda que hibrida con, se conjuga con, se une a, o se inmoviliza sobre una molécula diana, incluidos ácidos nucleicos (p. ej., ADN y ARN), péptidos y proteínas. En

algunas realizaciones, una sonda puede comprender un material aislado, purificado, de origen natural, de origen no natural y/o artificial, por ejemplo, incluyendo oligonucleótidos de cualquier longitud (p. ej., 5, 10, 20, 30, 40, 50, 100 o 150 nucleótidos o menos), en los que al menos una porción (p. ej., 50, 60, 70, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, o 100%) de las secuencias de oligonucleótidos es complementaria a un motivo de secuencia y/o dominio de hibridación presente en una o más moléculas diana, de manera que la sonda está configurada para hibridar (o interactuar de manera similar) en parte o en su totalidad con una o más moléculas diana o región de ácido nucleico de interés. La porción de la molécula diana o la región de ácido nucleico de interés con la que se hibrida una sonda se denomina "dominio de hibridación" de la sonda, que puede ser parte o la totalidad de la molécula diana o la región de ácido nucleico de interés como se describe en la presente memoria.

Una sonda puede ser de hebra sencilla o de doble hebra. En algunas realizaciones, la sonda se puede preparar a partir de un producto digerido con enzimas de restricción purificado o se puede producir sintéticamente, de forma recombinante o mediante amplificación por PCR. En realizaciones adicionales, la sonda puede comprender un material que se une a una secuencia peptídica concreta. Un conjunto de sondas descrito en la presente memoria puede comprender un conjunto de una o más sondas diseñadas para corresponder a una única ubicación genómica o un péptido en una secuencia proteica.

"Nucleótido" utilizado en la presente memoria significa un desoxirribonucleótido o un ribonucleótido o cualquier análogo de nucleótido (p. ej., ADN y ARN). Los análogos de nucleótidos incluyen nucleótidos que tienen modificaciones en la estructura química de la base, el azúcar y/o el fosfato, incluyendo, pero no limitadas a, modificaciones de la pirimidina de la posición 5', modificaciones de la purina de la posición 8, modificaciones en las aminas exocíclicas de la citosina, sustitución de 5-bromo-uracilo, y similares; y modificaciones del azúcar de la posición 2', incluyendo, pero no limitadas a, ribonucleótidos con azúcar modificado en los que el 2'-OH se reemplaza por un grupo seleccionado entre H, OR, R, halo, SH, SR, NH₂, NHR, NR₂, o CN. Los ARNhc también pueden comprender elementos no naturales tales como nucleótidos no naturales, p. ej., ionosina y xantina, azúcares no naturales, p. ej., 2'-metoxi ribosa, o enlaces fosfodiéster no naturales, p. ej., metilfosfonatos, fosforotioatos y péptidos. En una realización, el ARNhc comprende adicionalmente un elemento o una modificación que hace que el ARNhc sea resistente a la digestión con nucleasa. "Polinucleótido" u "oligonucleótido" se utilizan indistintamente y cada uno significa un polímero lineal de monómeros de nucleótidos. Los monómeros que forman polinucleótidos y oligonucleótidos son capaces de unirse específicamente a un polinucleótido natural y/o artificial por medio de un patrón regular de interacciones monómero a monómero, tales como el tipo de emparejamiento de bases de Watson-Crick, apilamiento de bases, tipos de emparejamiento de bases de Hoogsteen o de Hoogsteen inversos, o similares. Tales monómeros y sus enlaces internucleosídicos pueden ser naturales o pueden ser análogos de los mismos, p. ej., análogos de origen natural o de origen no natural. Los análogos de origen no natural pueden incluir PNA, LNA, enlaces internucleosídicos fosforotioato, nucleótidos que contienen grupos de enlace que permiten la unión de marcas, tales como fluoróforos o haptenos, y similares. Cuando el uso de un oligonucleótido o polinucleótido requiera un procesamiento enzimático, tal como la extensión por una polimerasa, la ligación por una ligasa o similar, un experto en la técnica entendería que los oligonucleótidos o polinucleótidos en esos casos no contendrían ciertos análogos de enlaces internucleosídicos, radicales de azúcar, o nucleótidos en cualquiera o algunas posiciones. Los polinucleótidos típicamente varían en tamaño de unas pocas unidades monoméricas, cuando se los denomina "oligonucleótidos" a varios miles de unidades monoméricas. Cuando un polinucleótido u oligonucleótido esté representado por una secuencia de letras (mayúsculas o minúsculas), tal como "ATGCCCTG", se entenderá que los nucleótidos están en el orden 5'→ 3' de izquierda a derecha. Normalmente, los polinucleótidos comprenden los cuatro nucleósidos naturales (p. ej., desoxiadenosina, desoxicitidina, desoxiguanosina, desoxitimidina para ADN o sus contrapartes de ribosa para ARN) unidos por enlaces fosfodiéster; sin embargo, también pueden comprender análogos de nucleótidos no naturales, p. ej., que incluyen nucleótidos modificados, azúcares o enlaces internucleosídicos. Está claro para los expertos en la técnica que cuando una enzima tiene requisitos específicos de sustrato de oligonucleótidos o polinucleótidos para su actividad, p. ej., ADN de hebra sencilla, ARN, dúplex ARN/ADN o similares, la selección de la composición apropiada para los sustratos oligonucleotídicos o polinucleotídicos se encuentra dentro del conocimiento de un experto en la técnica.

En otro aspecto, los métodos de la presente descripción pueden comprender hibridar al menos partes de los conjuntos de sondas primero y segundo con la primera y segunda regiones de ácido nucleico de interés en moléculas de nucleótidos de la muestra genética, respectivamente. La hibridación de las sondas con el ácido nucleico de interés se puede realizar simultáneamente o después de poner en contacto las sondas con la muestra genética, ligando, amplificando y/o inmovilizando las sondas. Además, la hibridación de las sondas con el ácido nucleico de interés se puede realizar simultáneamente o antes de ligar, amplificar y/o inmovilizar las sondas. Una parte o parte completa de la sonda puede hibridar con una parte o parte completa de la región de interés en moléculas de nucleótidos de hebra sencilla o doble, proteína o anticuerpo en una muestra. La región de interés hibridada a la sonda puede ser de 1 a 50 nucleótidos, 50 a 1000 nucleótidos, 100 a 500 nucleótidos, 5, 10, 50, 100, 200 nucleótidos o menos, o 2, 5, 10, 50, 100, 200, 500, 1000 nucleótidos o más. Las sondas se pueden diseñar o configurar para hibridar perfectamente con una región o molécula diana, o se pueden diseñar de tal manera que un emparejamiento erróneo de una sola base (p. ej., en un polimorfismo de un solo nucleótido, o sitio SNP), o un pequeño número de tales emparejamientos erróneos, no produzca un híbrido de sonda y molécula diana.

En realizaciones adicionales, la primera sonda de marcaje y/o la primera sonda de etiquetado hibridan con la primera región de ácido nucleico de interés, y la segunda sonda de marcaje y/o las segundas sondas de etiquetado se

5 hibridan con la segunda región de ácido nucleico de interés. En realizaciones adicionales, múltiples o todas las sondas y/u otros componentes (p. ej., sondas de marcaje, sondas de etiquetado y sondas de huecos) de un conjunto de sondas que hibridan con una región de ácido nucleico de interés son adyacentes entre sí. Cuando dos de las sondas y/o componentes hibridados con la región de ácido nucleico de interés son "adyacentes" o "inmediatamente adyacentes", no hay nucleótido entre los dominios de hibridación de las dos sondas en la región de ácido nucleico de interés. En esta realización, las diferentes sondas dentro de un conjunto de sondas se pueden unir covalentemente para formar una molécula de oligonucleótido más grande. En otra realización, un conjunto de sondas se puede diseñar para que hibride con una porción no contigua, pero proximal, de la región de ácido nucleico de interés, de tal manera que haya un "hueco" de uno o más nucleótidos en la región de ácido nucleico de interés, entre las sondas hibridadas de un conjunto de sondas, que no está ocupado por una sonda. En esta realización, se puede utilizar una ADN polimerasa u otra enzima para sintetizar una nueva secuencia de polinucleótidos, en algunos casos uniendo covalentemente dos sondas de un solo conjunto de sondas. Dentro de un conjunto de sondas, cualquier sonda puede llevar una o más marcas, o etiquetas de afinidad utilizadas para la identificación o el aislamiento del locus. En un aspecto, la primera y segunda sondas de marcaje hibridan con la primera y la segunda región de ácido nucleico de interés en moléculas de nucleótidos de la muestra genética, respectivamente; la primera y la segunda sondas de etiquetado hibridan con la primera y la segunda región de ácido nucleico de interés en las moléculas de nucleótidos de la muestra genética, respectivamente; la primera sonda de marcaje hibrida con una región adyacente a donde se hibrida la primera sonda de etiquetado; y la segunda sonda de marcaje hibrida a una región adyacente a donde hibrida la segunda sonda de etiquetado.

20 La hibridación ocurre de tal manera que las sondas dentro de un conjunto de sondas se pueden modificar para formar una nueva entidad molecular más grande (p. ej., una sonda producto). Las sondas en la presente memoria pueden hibridar con las regiones de ácido nucleico de interés en condiciones rigurosas. Como se emplea en la presente memoria, el término "rigurosidad" se utiliza en referencia a las condiciones de temperatura, fuerza iónica y la presencia de otros compuestos tales como disolventes orgánicos, en las cuales se realizan las hibridaciones de ácidos nucleicos. La "rigurosidad" se produce típicamente en un intervalo de aproximadamente T_m °C a aproximadamente 20°C a 25°C por debajo de T_m . Se puede utilizar una hibridación rigurosa para aislar y detectar secuencias de polinucleótidos idénticas o para aislar y detectar secuencias de polinucleótidos similares o relacionadas. En "condiciones rigurosas", la secuencia de nucleótidos, en su totalidad o partes de la misma, hibridará con su complemento exacto y secuencias estrechamente relacionadas. Las condiciones de baja rigurosidad comprenden condiciones equivalentes a la unión o hibridación a 68°C en una solución que consiste en 5 × SSPE (43,8 g/l de NaCl, 6,9 g/l de $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$ y 1,85 g/l de EDTA, pH ajustado a 7,4 con NaOH), SDS al 0,1%, 5x reactivo de Denhardt (50x Denhardt contiene por 500 mL: 5 g de Ficoll (Tipo 400), 5 g de BSA) y 100 µg/mL de ADN de esperma de salmón desnaturalizado seguido de lavado en una solución que comprende 2,0 + SSPE, SDS al 0,1% a temperatura ambiente cuando se emplea una sonda de aproximadamente 100 a aproximadamente 1000 nucleótidos de longitud. Es bien conocido en la técnica que se pueden emplear numerosas condiciones equivalentes para comprender condiciones de baja rigurosidad; factores tales como la longitud y la naturaleza (ADN, ARN, composición de bases) de la sonda y la naturaleza de la diana (ADN, ARN, composición de bases, presente en solución o inmovilizado, etc.) y la concentración de las sales y otros componentes (p. ej., la presencia o ausencia de formamida, sulfato de dextrano, polietilenglicol, así como los componentes de la solución de hibridación pueden variar para generar condiciones de hibridación de baja rigurosidad diferentes de, pero equivalentes a, las condiciones enumeradas anteriormente. Además, las condiciones que promueven la hibridación en condiciones de alta rigurosidad (p. ej., el aumento de la temperatura de las etapas de hibridación y/o lavado, el uso de formamida en la solución de hibridación, etc.) son bien conocidas en la técnica. Las condiciones de alta rigurosidad, cuando se utilizan en referencia a la hibridación de ácidos nucleicos, comprenden condiciones equivalentes a la unión o hibridación a 68°C en una solución que consiste en 5 + SSPE, SDS al 1%, 5x reactivo de Denhardt y 100 µg/mL de ADN de esperma de salmón desnaturalizado seguido de lavado en una solución que comprende 0,1 + SSPE y SDS al 0,1% a 68°C cuando se emplea una sonda de aproximadamente 100 a aproximadamente 1000 nucleótidos de longitud.

50 En algunas realizaciones, la sonda producto se puede formar solo si las sondas dentro de un conjunto de sondas están correctamente hibridadas. Por lo tanto, las sondas producto se pueden formar con alta rigurosidad y alta precisión. De nuevo, las sondas producto pueden contener información suficiente para identificar mediante interrogación la secuencia genómica para la cual fue diseñada la sonda producto. Por lo tanto, la generación y la cuantificación directa de una sonda producto concreta (en este caso, por recuento molecular) puede reflejar la abundancia de una secuencia genética concreta en la muestra de origen.

55 En realizaciones adicionales, las regiones de ácido nucleico de interés, para las cuales se configuran las sondas para su hibridación, están ubicadas en cromosomas diferentes. Por ejemplo, la primera región de ácido nucleico de interés está ubicada en el cromosoma 21, y la segunda región de ácido nucleico de interés no está ubicada en el cromosoma 21 (p. ej., ubicada en el cromosoma 18).

60 En otro aspecto, los métodos de la presente descripción pueden comprender ligar la primera sonda de marcaje y la primera sonda de etiquetado, y ligar la segunda sonda de marcaje y la segunda sonda de etiquetado. La ligación de las sondas se puede realizar simultáneamente o después de poner en contacto las sondas con la muestra genética, amplificar y/o inmovilizar las sondas. Además, la ligación de las sondas se puede realizar simultáneamente o antes de poner en contacto las sondas con la muestra genética, amplificar y/o inmovilizar las sondas. La ligación en la

presente memoria significa el procedimiento de unión de dos sondas (p. ej., unión de dos moléculas de nucleótidos) entre sí. Por ejemplo, la ligación en la presente memoria puede implicar la formación de un enlace 3',5'-fosfodiéster que conecta dos nucleótidos, y un agente de unión que es un agente capaz de causar la ligación puede ser una enzima o un agente químico.

5 En otro aspecto, los métodos de la presente descripción pueden comprender amplificar las sondas ligadas y/o los conjuntos de sondas ligadas. La amplificación de las sondas ligadas se puede realizar simultáneamente o después de poner en contacto las sondas con la muestra genética, ligar, hibridar y/o inmovilizar las sondas. Además, la amplificación de las sondas ligadas se puede realizar simultáneamente o antes de inmovilizar las sondas. La amplificación en la presente memoria se define como la producción de copias adicionales de la sonda y/o la sonda producto y se puede llevar a cabo utilizando tecnologías de reacción en cadena de la polimerasa bien conocidas en la técnica. Como se emplea en la presente memoria, el término "reacción en cadena de la polimerasa" ("PCR") se refiere a un método para aumentar la concentración de un segmento de una secuencia diana (p. ej., en una mezcla de ADN genómico) sin clonación o purificación. La longitud del segmento amplificado de la secuencia diana deseada está determinada por las posiciones relativas de dos cebadores oligonucleotídicos entre sí, y por lo tanto, esta longitud es un parámetro controlable. En virtud del aspecto repetitivo del procedimiento, el método se denomina "reacción en cadena de la polimerasa" (en adelante, "PCR"). Debido a que los segmentos amplificados deseados de la secuencia diana se convierten en las secuencias predominantes (en términos de concentración) en la mezcla, se dice que están "amplificadas por PCR". Con la PCR, es posible amplificar una copia única de una secuencia diana específica en el ADN genómico a un nivel detectable mediante varias metodologías diferentes (p. ej., hibridación con una sonda marcada). Además del ADN genómico, cualquier secuencia de oligonucleótido se puede amplificar con el conjunto apropiado de moléculas cebadoras. En particular, los segmentos amplificados creados por el propio procedimiento de PCR son, en sí mismos, moldes eficaces para las posteriores amplificaciones por PCR. Una amplificación puede ser una amplificación en "tiempo real" si se dispone de una química de detección que permita medir un producto de reacción a medida que progresa la reacción de amplificación, p. ej., "PCR en tiempo real" o "NASBA en tiempo real" como describen Leone et al., en *Nucleic Acids Research*, 26: 2150-2155 (1998).

Los cebadores generalmente son de una sola hebra para una máxima eficacia en la amplificación, pero alternativamente pueden ser de doble hebra. Si es de doble hebra, el cebador generalmente se trata primero para separar sus hebras antes de utilizarse para preparar productos de extensión. Esta etapa de desnaturalización está típicamente influenciada por el calor, pero se puede llevar a cabo alternativamente con álcali, seguido de neutralización. Por lo tanto, un "cebador" es complementario a un molde, y forma complejos mediante puentes de hidrógeno o hibridación con el molde para proporcionar un complejo de cebador/molde para el inicio de la síntesis por una polimerasa, que se extiende mediante la adición de nucleótidos unidos covalentemente conectados a su extremo 3' complementario al molde en el proceso de síntesis de ADN.

Un "par de cebadores", como se emplea en la presente memoria, se refiere a un cebador directo y a un cebador inverso correspondiente, que tienen secuencias de ácido nucleico adecuadas para la amplificación basada en ácido nucleico de un ácido nucleico diana. Tales pares de cebadores generalmente incluyen un primer cebador que tiene una secuencia que es igual o similar a la de una primera porción de un ácido nucleico diana, y un segundo cebador que tiene una secuencia que es complementaria a una segunda porción de un ácido nucleico diana para proporcionar la amplificación del ácido nucleico diana o un fragmento del mismo. La referencia a los cebadores "primero" y "segundo" en la presente memoria es arbitraria, a menos que se indique específicamente lo contrario. Por ejemplo, el primer cebador se puede diseñar como un "cebador directo" (que inicia la síntesis de ácido nucleico desde un extremo 5' del ácido nucleico diana) o como un "cebador inverso" (que inicia la síntesis de ácido nucleico desde un extremo 5' del producto de extensión producido a partir de la síntesis iniciada desde el cebador directo). Del mismo modo, el segundo cebador se puede diseñar como un cebador directo o un cebador inverso.

En algunas realizaciones, la región de ácido nucleico de interés en la molécula de nucleótido de la presente memoria se puede amplificar mediante los métodos de amplificación descritos en la presente memoria. Los ácidos nucleicos en una muestra pueden o no amplificarse antes del análisis, utilizando un método de amplificación universal (p. ej., amplificación del genoma completo y PCR del genoma completo). La amplificación de la región de ácido nucleico de interés se puede realizar simultáneamente o después de poner en contacto las sondas con la muestra genética, ligar, amplificar y/o inmovilizar las sondas. Además, la amplificación de las sondas ligadas se puede realizar simultáneamente o antes de poner en contacto las sondas con la muestra genética, ligar las sondas, inmovilizar las sondas y/o contar las marcas.

En realizaciones adicionales, el método excluye la amplificación de las moléculas de nucleótidos de la muestra genética después de la hibridación o la ligación. En realizaciones adicionales, el método excluye la amplificación de las moléculas de nucleótidos de la muestra genética después de la hibridación y la ligación.

En otro aspecto, los métodos de la presente descripción pueden comprender inmovilizar las sondas de etiquetado en una ubicación predeterminada sobre un sustrato. La inmovilización de la sonda sobre un sustrato se puede realizar simultáneamente o después de poner en contacto las sondas con la muestra genética, hibridar las sondas con la región de interés del ácido nucleico, ligar y/o amplificar las sondas. Además, la inmovilización de la sonda sobre un sustrato se puede realizar simultáneamente o antes de poner en contacto las sondas con la muestra genética, hibridar las sondas con la región de ácido nucleico de interés, ligar, amplificar y/o contar las sondas. La

inmovilización en la presente memoria significa unir directa o indirectamente las sondas de etiquetado a la ubicación predeterminada en el sustrato mediante un enlace físico o químico. En algunas realizaciones, el sustrato en la presente memoria puede comprender un compañero de unión que está configurado para contactar y unirse a una parte o a la totalidad de una etiqueta en la sonda de etiquetado descrita en la presente memoria e inmovilizar la etiqueta y, por lo tanto, la sonda de etiquetado que comprende la etiqueta. La etiqueta de la sonda de etiquetado puede comprender un compañero de unión correspondiente del compañero de unión sobre el sustrato como se describe en la presente memoria.

La inmovilización se puede realizar mediante la hibridación de una parte o la totalidad de la sonda de etiquetado con una parte o la totalidad del compañero de unión sobre el sustrato. Por ejemplo, la etapa de inmovilización comprende hibridar al menos una parte de la etiqueta o secuencia de nucleótidos de etiquetado con una molécula de nucleótido correspondiente inmovilizada sobre el sustrato. Aquí, la molécula de nucleótido correspondiente es un compañero de unión de la etiqueta o secuencia de nucleótidos de etiquetado que está configurada para hibridar parcial o totalmente con etiqueta o secuencia de nucleótidos de etiquetado. En algunas realizaciones, los compañeros de unión oligonucleotídicos o polinucleotídicos pueden ser de hebra sencilla y pueden anclarse covalentemente al sustrato, por ejemplo, por el extremo 5' o el extremo 3'. La inmovilización también se puede realizar mediante los siguientes compañeros de unión y medios de unión ilustrativos: Biotina-oligonucleótido complejo con Avidina, Estreptavidina o Neutravidina; oligonucleótido SH unido covalentemente a través de un enlace disulfuro a una superficie SH; Amina-oligonucleótido unido covalentemente a un carboxilato activado o un grupo aldehído; Ácido fenilborónico (PBA)-oligonucleótido complejo con ácido salicilhidroxámico (SHA); acridita-oligonucleótido que reacciona con la superficie de tiol o silano o co-polimeriza con monómero de acrilamida para formar poli(acrilamida), o mediante otros métodos conocidos en la técnica. Para algunas aplicaciones donde es preferible tener una superficie cargada, las capas superficiales pueden estar compuestas por una estructura de múltiples capas de polielectrolito (PEM) como se muestra en la Publicación de la Solicitud de Patente de Estados Unidos Núm. 2002/025529. En algunas realizaciones, la inmovilización se puede realizar mediante procedimientos bien conocidos, por ejemplo, que comprenden poner en contacto las sondas con el soporte que tiene compañeros de unión anclados durante un cierto período de tiempo, y después de que las sondas se agoten para la extensión, el soporte con los productos de extensión inmovilizados se enjuaga opcionalmente utilizando un líquido adecuado. En realizaciones adicionales, la inmovilización de sondas producto sobre un sustrato puede permitir un lavado riguroso para eliminar los componentes de la muestra biológica y el ensayo, reduciendo así el ruido de fondo y mejorando la precisión.

"Soporte sólido", "soporte", "sustrato" y "soporte en fase sólida" se utilizan indistintamente y se refieren a un material o grupo de materiales que tienen una superficie o superficies rígidas o semirrígidas. En algunas realizaciones, al menos una superficie del sustrato será sustancialmente plana, aunque en algunas realizaciones puede ser deseable separar físicamente las regiones de síntesis para diferentes compuestos, por ejemplo, con pocillos, regiones elevadas, agujas, fosas de grabado al ácido o similares. En realizaciones adicionales, el sustrato puede comprender al menos un soporte plano en fase sólida (p. ej., un portaobjetos de microscopio de vidrio). De acuerdo con otras realizaciones, el sustrato o los sustratos adoptarán la forma de cuentas, resinas, geles, microesferas u otras configuraciones geométricas. En un aspecto, el sustrato de acuerdo con algunas realizaciones de la presente descripción excluye cuentas, resinas, geles y/o microesferas.

En algunas realizaciones, como se muestra en la Figura 1, los compañeros de unión, las etiquetas, las etiquetas de afinidad, las marcas, las sondas (p. ej., las sondas de etiquetado y las sondas de marcaje), y/o los conjuntos de sondas descritos en la presente memoria pueden inmovilizarse sobre un sustrato (1) como una matriz (2). La matriz en la presente memoria tiene múltiples miembros (3-10) que pueden o no tener un solapamiento (6) entre los miembros. Cada miembro puede tener al menos un área sin solapamiento con otro miembro (3-5 y 7-10). En realizaciones adicionales, cada miembro puede adoptar diferentes formas (p. ej., puntos circulares (3-8), triángulos (9) y cuadrados (10)) y dimensiones. Un miembro de una matriz puede tener un área de 1 a 10^7 micras², de 100 a 10^7 micras², de 10^3 a 10^8 micras², de 10^4 a 10^7 micras²; de 10^5 a 10^7 micras²; aproximadamente 0,0001, 0,001, 0,01, 0,1, 1, 10, 100, 10^3 , 10^4 , 10^5 , 10^6 , 10^7 , 10^8 o más micras²; y/o aproximadamente 0,001, 0,01, 0,1, 1, 10, 100, 10^3 , 10^4 , 10^5 , 10^6 , 10^7 , 10^8 o menos micras². Una imagen de un miembro (8) ilustrativo de acuerdo con algunas realizaciones de la presente invención se muestra como elemento 12. Además, dos o más miembros que comprenden los compañeros de unión, las etiquetas, las etiquetas de afinidad, las marcas, las sondas (p. ej., sondas de etiquetado y sondas de marcaje), y/o los conjuntos de sondas del mismo tipo pueden tener la misma forma y dimensión. Específicamente, los miembros de una matriz que comprende los compañeros de unión, las etiquetas, las etiquetas de afinidad, las marcas, las sondas de etiquetado y/o los conjuntos de sondas configurados o utilizados para detectar la misma variación genética o un control de acuerdo con los métodos descritos en la presente memoria pueden tener las mismas formas y dimensiones. Además, todos y cada uno de los miembros de las matrices sobre el sustrato pueden tener las mismas formas y dimensiones. En otras realizaciones, los miembros de una matriz que comprende los compañeros de unión, etiquetas, etiquetas de afinidad, marcas, sondas y/o conjuntos de sondas configurados o utilizados para detectar diferentes variaciones genéticas y/o controles de acuerdo con los métodos descritos en la presente memoria pueden tener las mismas formas y dimensiones. Además, cada miembro de la matriz puede comprender diferentes compañeros de unión, las etiquetas, las etiquetas de afinidad, las marcas, las sondas y/o los conjuntos de sondas.

En algunas realizaciones, dos miembros de la matriz pueden estar separados por (i) una distancia, en la que puede

no haber o haber solo muy pocos compañeros de unión, etiquetas, etiquetas de afinidad, marcas, sondas (p. ej., sondas de etiquetado y sondas de marcaje), y/o conjuntos de sondas inmovilizados, y/o (ii) cualquier separador que distinga un miembro de otro (p. ej., sustrato elevado, cualquier material que impida la unión de los compañeros de unión, etiquetas, etiquetas de afinidad, sondas (p. ej., sondas de etiquetado), y/o conjuntos de sondas con respecto al sustrato, y cualquier material que no sea sonda entre los miembros). En realizaciones adicionales, los miembros de la matriz se pueden distinguir entre sí al menos solo por sus ubicaciones. Los miembros de la matriz pueden estar separados por una distancia de 0 a 10^4 micras, de 0 a 10^3 micras, desde 10^2 a 10^4 micras, o de 10^2 a 10^3 micras aproximadamente 0, 0,001, 0,1, 1, 2, 3, 4, 5, 10, 50, 100, 10^3 , 10^4 , 10^5 , 10^6 , 10^7 , o 10^8 micras o más; y/o aproximadamente 0, 0,001, 0,1, 1, 2, 3, 4, 5, 10, 50, 100, 10^3 , 10^4 , 10^5 , 10^6 , 10^7 , o 10^8 micras o menos. Aquí, la distancia por la cual se separan dos miembros de la matriz se puede determinar por la distancia más corta entre los bordes de los miembros. Por ejemplo, en la Figura 1, la distancia por la que se separan dos miembros, los elementos 3 y 4, de una matriz (2) es la distancia indicada por el elemento n. Además, por ejemplo, la distancia más corta por la cual se separan los miembros de la matriz (2) sobre un sustrato (1) es 0, como la distancia por la cual se separan dos miembros, elementos 10 y 11, de la matriz. En otras realizaciones, dos miembros de la matriz pueden no estar separados y pueden estar solapados (6). En tales realizaciones, cada miembro puede tener al menos un área sin solapamiento con otro miembro (7).

En realizaciones adicionales, una matriz y los miembros de la matriz de los compañeros de unión, las etiquetas, las etiquetas de afinidad, las marcas, las sondas y/o los conjuntos de sondas descritos en la presente memoria se pueden ubicar en ubicaciones predeterminadas sobre el sustrato, y las formas y las dimensiones de cada miembro de la matriz y la distancia entre los miembros se pueden determinar antes de la inmovilización. La ubicación predeterminada en la presente memoria significa una ubicación que se determina o identifica antes de la inmovilización. Por ejemplo, la forma y dimensión de cada miembro de una matriz se determina o identifica antes de la inmovilización.

En realizaciones adicionales, el sustrato puede comprender una matriz de compañeros de unión, comprendiendo cada miembro de la matriz los compañeros de unión, tales como oligonucleótidos o polinucleótidos, que están inmovilizados (p. ej., por medio de un enlace químico que no se rompería durante la hibridación de las sondas a los compañeros de unión del sustrato descrito en la presente memoria) en una región o ubicación espacialmente definidas; es decir, las regiones o ubicaciones son espacialmente discretas o separadas por una región o ubicación definidas sobre el sustrato. En realizaciones adicionales, el sustrato puede comprender una matriz, cada miembro de la cual comprende compañeros de unión que se unen a una región o ubicación espacialmente definidas. Cada una de las ubicaciones espacialmente definidas configuradas para comprender los compañeros de unión puede ser adicionalmente "direccionables" en el sentido de que su ubicación y la identidad de sus compañeros de unión inmovilizados son conocidas o predeterminadas, por ejemplo, antes de su uso, análisis, o anclaje a sus compañeros de unión en las sondas de etiquetado y/o conjuntos de sondas. El término "direccionable" con respecto a los conjuntos de sondas inmovilizadas sobre el sustrato significa que la secuencia de nucleótidos u otras características físicas y/o químicas de una parte anclada al extremo (p. ej., un compañero de unión del compañero de unión del sustrato, etiqueta, etiqueta de afinidad y sonda de etiquetado) de un conjunto de sondas descrito en la presente memoria se pueden determinar a partir de su dirección, es decir, una correspondencia uno a uno entre la secuencia u otra propiedad de la parte anclada al extremo del conjunto de sondas y una ubicación espacial en, o característica del sustrato sobre el que se inmoviliza el conjunto de sondas. Por ejemplo, una dirección de una parte anclada al extremo de un conjunto de sondas es una ubicación espacial, p. ej., las coordenadas planas de una región particular que inmoviliza copias de la parte anclada al extremo del conjunto sondas. Sin embargo, las partes unidas al extremo de los conjuntos de sondas también se pueden dirigir de otras maneras, p. ej., por el color, la frecuencia del micro-transpondedor, o similares, p. ej., Chandler et al. Publicación PCT WO 97/14028. En realizaciones adicionales, los métodos descritos en la presente memoria excluyen "micromatrices aleatorias", que se refiere a una micromatriz cuyas regiones espacialmente discretas de los compañeros de unión (p. ej., oligonucleótidos o polinucleótidos) del sustrato y/o las partes ancladas al extremo de los conjuntos de sondas no están espacialmente dirigidas. Es decir, la identidad de los compañeros de unión, la etiqueta, la etiqueta de afinidad, la sonda de etiquetado y/o los conjuntos de sondas anclados no se puede discernir, al menos inicialmente, desde su ubicación. En un aspecto, los métodos descritos en la presente memoria excluyen micromatrices aleatorias que son matrices planas de microcuentas.

Una matriz de ácido nucleico de acuerdo con algunas realizaciones de la presente descripción se puede producir por cualquier método bien conocido en la técnica, incluyendo, pero sin limitarse a los descritos en la Publicación de la Solicitud de Patente de Estados Unidos Núm. 2013/0172216; Schena, Microarrays: A Practical Approach (IRL Press, Oxford, 2000). Por ejemplo, se puede utilizar una matriz de captura de ADN. La matriz de captura de ADN es un sustrato sólido (p. ej., un portaobjetos de vidrio) con oligonucleótidos localizados anclados covalentemente a la superficie. Estos oligonucleótidos pueden tener uno o más tipos en la superficie, y pueden adicionalmente ser segregados geográficamente a través del sustrato. En condiciones de hibridación, las matrices de captura de ADN se unirán preferentemente a dianas complementarias en comparación con otros radicales no específicos, actuando así para localizar las dianas en la superficie y separarlas de las especies no deseadas.

En algunas realizaciones, la primera y la segunda sondas de marcaje y/o las sondas de marcaje amplificadas de las mismas ligadas a las sondas de etiquetado inmovilizadas comprenden la primera y la segunda marcas, respectivamente.

La sonda de marcaje significa en la presente memoria una sonda que comprende o está configurada para unirse a una marca. La propia sonda de marcaje puede comprender una marca o se puede modificar para comprender o unirse a una marca. La sonda amplificada en la presente memoria se define como las copias adicionales de una sonda inicial producida después de la amplificación de la sonda inicial como se describe en la presente memoria. Por consiguiente, las sondas amplificadas pueden tener una secuencia que es la secuencia de nucleótidos de las sondas iniciales y/o la secuencia complementaria de las secuencias de nucleótidos de las sondas iniciales. Las sondas amplificadas pueden contener una secuencia que coincide parcial o completamente con las secuencias de nucleótidos de las sondas iniciales. Los términos "complementario" o "complementariedad" se utilizan en referencia a una secuencia de nucleótidos relacionada por las reglas de emparejamiento de bases. Por ejemplo, la secuencia "5'-CAGT-3'", es complementaria a la secuencia "5'-ACTG-3'". La complementariedad puede ser "parcial" o "total". La complementariedad "parcial" es aquella en la que uno o más nucleótidos de ácido nucleico en una sonda no coinciden de acuerdo con las reglas de emparejamiento de bases, mientras que otros coinciden. La complementariedad "total" o "completa" entre los ácidos nucleicos es aquella en la que todas y cada una de las bases del ácido nucleico de la sonda coinciden con otra base según las reglas de emparejamiento de bases.

La sonda inmovilizada en la presente memoria se define como una sonda que se une directa o indirectamente al sustrato por un enlace físico o químico. En algunas realizaciones, una sonda de marcaje puede inmovilizarse sobre un sustrato indirectamente a través de la ligación a una sonda de etiquetado inmovilizada sobre el sustrato descrito en la presente memoria.

Una marca en la presente memoria significa una molécula, colorante o radical orgánico, natural, sintético, artificial o no natural que tiene una propiedad o característica que es susceptible de detección y, opcionalmente, de cuantificación. Una marca puede ser directamente detectable (p. ej., radioisótopos, fluoróforos, quimioluminóforos, enzimas, partículas coloidales, sustancias fluorescentes, puntos cuánticos u otras nanopartículas, nanoestructuras, compuestos metálicos, marcas organometálicas y aptámeros peptídicos); o una marca puede ser detectable indirectamente utilizando compañeros de unión específicos. Los ejemplos de sustancias fluorescentes incluyen colorantes fluorescentes tales como fluoresceína, fósforo, rodamina, derivados de colorantes de polimetina y similares. Los ejemplos de una sustancia fluorescente disponible en el mercado incluyen colorantes fluorescentes, como BODIPY FL (marca registrada, producido por Molecular Probes, Inc.), FluorePrime (nombre de producto, producido por Amersham Pharmacia Biotech, Inc.), Fluoredate (nombre de producto, producido por Millipore Corporation), FAM (producido por ABI Inc.), Cy 3 y Cy 5 (producido por Amersham Pharmacia), TAMRA (producido por Molecular Probes, Inc.), Pacific Blue, TAMRA, Alexa 488, Alexa 594, Alexa 647, Atto 488, Atto 590, Atto 647N y similares. "Punto cuántico" (QD) significa una estructura cristalina semiconductor a nanoescala, generalmente preparada a partir de seleniuro de cadmio, y absorbe la luz y a continuación la emite un par de nanosegundos más tarde en un color específico. Los QD con una variedad de superficies conjugadas o reactivas, p. ej., amino, carboxilo, estreptavidina, proteína A, biotina e inmunoglobulinas, también se incluyen en la presente descripción.

En realizaciones adicionales, las primera y segunda marcas son diferentes, de modo que las marcas se pueden distinguir entre sí. En realizaciones adicionales, la primera y segunda marcas son diferentes en sus propiedades físicas, ópticas y/o químicas.

En algunas realizaciones, las marcas inmovilizadas se pueden resolver ópticamente. El término "marca que se puede resolver ópticamente" o "marca que se puede resolver individualmente ópticamente en la presente memoria" significa un grupo de marcas que se pueden distinguir entre sí por su emisión fotónica u otras propiedades ópticas, por ejemplo, después de la inmovilización como se describe en la presente memoria. En realizaciones adicionales, aunque las marcas pueden tener las mismas propiedades de emisión óptica y/o espectral, las marcas inmovilizadas se pueden distinguir entre sí espacialmente. En algunas realizaciones, las marcas del mismo tipo, que se definen por ser marcas que tienen las mismas propiedades ópticas, se inmovilizan sobre el sustrato, por ejemplo como un miembro de una matriz descrita en la presente memoria, a una densidad y/o separación tal que las sondas producto individuales se pueden resolver como se muestra en el elemento **12** de la Figura 1. En esta descripción, las "mismas marcas" se definen por ser marcas que tienen composiciones químicas y físicas idénticas. Las "diferentes marcas" en la presente memoria significan marcas que tienen diferentes composiciones químicas y/o físicas, incluyendo "marcas de diferentes tipos" que tienen diferentes propiedades ópticas. Las "diferentes marcas del mismo tipo" en la presente memoria significan marcas que tienen diferentes composiciones químicas y/o físicas, pero las mismas propiedades ópticas.

El elemento **12** de la Figura 1 representa una imagen de un miembro ilustrativo de una matriz que comprende marcas inmovilizadas. En estas realizaciones, las marcas son direccionables espacialmente ya que la ubicación de una molécula específica su identidad (y en la síntesis combinatoria espacial, la identidad es una consecuencia de la ubicación). En realizaciones adicionales, un miembro de la matriz sobre el sustrato puede tener una o múltiples sondas marcadas inmovilizadas sobre el miembro. Cuando se inmovilizan múltiples sondas marcadas en un miembro de la matriz, las marcas del mismo tipo en las sondas marcadas inmovilizadas en un miembro de una matriz sobre el sustrato se pueden distinguir entre sí espacialmente como se muestra en el elemento **12** de la Figura 1. En algunas realizaciones, las marcas inmovilizadas del mismo tipo están separadas por una distancia de aproximadamente 1 a 1000 nm, de 5 a 100 nm, o de 10 a 100 nm; aproximadamente 100, 150, 200, 250, 300, 350 o 400 nm o más; y/o aproximadamente 50, 100, 150, 200, 250, 300, 350 o 400 nm o menos en todas las dimensiones. La densidad de las sondas producto y sus marcas en los sustratos puede ser de hasta muchos millones (y hasta mil

millones o más) de sondas producto que se contarán por sustrato. La capacidad de contar un gran número de sondas producto que contienen las marcas permite una cuantificación precisa de las secuencias de ácido nucleico.

En algunas realizaciones, las primera y segunda sondas de etiquetado inmovilizadas y/o las sondas de etiquetado amplificadas de las mismas comprenden una primera y una segunda etiquetas, respectivamente. La sonda de etiquetado en la presente memoria significa una sonda que está configurada para unirse directa o indirectamente al sustrato. La propia sonda de etiquetado puede unirse al sustrato o puede modificarse para unirse al sustrato. Una etiqueta o etiqueta de afinidad en la presente memoria significan un motivo para el aislamiento, enriquecimiento o inmovilización específicos de sondas producto. Los ejemplos de la etiqueta o etiqueta de afinidad incluyen un compañero de unión descrito en la presente memoria, secuencias de ADN únicas que permiten la captura específica de la secuencia, incluida la secuencia genómica natural y/o no genómica artificial, biotina-estreptavidina, etiquetas de His, octapéptido FLAG, química click (p. ej., pares de grupos funcionales que reaccionan rápida y selectivamente entre sí en condiciones acuosas suaves), y anticuerpos (p. ej., azida-ciclina). Por ejemplo, la etapa de inmovilización comprende hibridar al menos una parte de la etiqueta, la etiqueta de afinidad o la secuencia de nucleótidos de etiquetado con una molécula de nucleótido correspondiente inmovilizada sobre el sustrato. La etiqueta o etiqueta de afinidad se configuran para unirse a entidades que incluyen, pero no se limitan a, una cuenta, una cuenta magnética, un portaobjetos de microscopio, un cubreobjetos, una micromatriz o una molécula. En algunas realizaciones, la etapa de inmovilización se realiza inmovilizando las etiquetas sobre la ubicación predeterminada del sustrato.

En otro aspecto, se cuentan los números de diferentes marcas inmovilizadas sobre el sustrato y, por lo tanto, los números de diferentes sondas producto inmovilizadas que comprenden las marcas. Por ejemplo, las sondas producto de cada locus genético se agrupan, y se cuentan las marcas en las sondas producto inmovilizadas. En algunas realizaciones, se pueden interrogar múltiples secuencias dentro de un locus genómico a través de la creación de múltiples tipos de sondas producto. Para este ejemplo, se pueden combinar diferentes sondas producto para el mismo locus genómico (posiblemente mediante la inmovilización en una ubicación común de un sustrato, p. ej., como miembro de una matriz descrita en la presente memoria), y las marcas en estas sondas producto se pueden contar directamente. También se pueden separar diferentes sondas producto para el mismo locus genómico (posiblemente mediante la inmovilización en diferentes ubicaciones de un sustrato, p. ej., como miembros diferentes de una matriz descrita en la presente memoria), y las marcas en estas sondas producto se pueden contar directamente. En realizaciones adicionales, el sustrato puede tener una o más etiquetas de afinidad específicas en cada ubicación en un sustrato, p. ej., como miembro de una matriz en el sustrato. Por lo tanto, otro método para cuantificar las secuencias de ácido nucleico se produce mediante la inmovilización de sondas producto para un solo locus genómico (este puede ser un tipo de sonda producto, o puede ser un conjunto de más de una sonda producto para un locus genómico particular) sobre la misma ubicación de un sustrato (p. ej., como el mismo miembro de una matriz descrita en la presente memoria) como sondas producto correspondientes a un segundo locus genómico, que pueden o no servir como locus de referencia o de control. En este caso, las sondas producto del primer locus genómico serán distinguibles de las sondas producto del segundo locus genómico, en función de la presencia de diferentes marcas utilizadas en la generación de las sondas producto.

En un ejemplo, para detectar la trisomía 21 (aneuploidía) de un feto mediante el examen de una muestra de sangre materna, se generaría un conjunto de sondas producto correspondientes al cromosoma 21, por ejemplo, con una marca de fluoróforo de color rojo, y se contaría. También se generaría un segundo conjunto de sondas producto a partir de un locus de referencia o control, p. ej., el cromosoma 18, y se contaría. Este segundo conjunto de sondas producto se puede generar, por ejemplo, con una marca de fluoróforo de color verde.

En algunas realizaciones, estas sondas producto se pueden preparar de tal manera que se agrupen por locus (en este caso, cromosoma 21 o cromosoma 18) y se cuenten por separado sobre un sustrato. Es decir, las sondas producto correspondientes al cromosoma 21 se pueden aislar y contabilizar por separado, y las sondas producto correspondientes al cromosoma 18 se pueden aislar y contabilizar por separado. En realizaciones adicionales, estas sondas producto también se pueden preparar de tal manera que se agrupen en la misma ubicación de un sustrato (p. ej., como el mismo miembro de una matriz descrita en la presente memoria). En este caso, en la misma región de un sustrato, las sondas producto que portan un fluoróforo de color rojo responderán al cromosoma 21, y las sondas producto con un fluoróforo de color verde responderán al cromosoma 18. Por ejemplo, ya que todas estas sondas producto se pueden resolver individualmente y, por lo tanto, se pueden contar con mucha precisión, una mayor frecuencia de sondas producto del cromosoma 21 con respecto a las sondas producto del cromosoma 18 (incluso tan pequeña como 0,01, 0,1, uno o más por ciento o menos) significará la presencia de trisomía 21 en un feto. En este caso, las sondas producto para el cromosoma 18 pueden servir como control.

En otro aspecto, los métodos de la presente descripción pueden comprender el recuento de las marcas de los conjuntos de sondas inmovilizados sobre el sustrato. En algunas realizaciones, los métodos pueden comprender el recuento de (i) un primer número de la primera marca inmovilizada sobre el sustrato, y (ii) un segundo número de la segunda marca inmovilizada sobre el sustrato. La etapa de recuento se puede realizar después de inmovilizar el conjunto de sondas ligadas a un sustrato, y el sustrato con conjuntos de sondas ligadas inmovilizadas se puede almacenar en condiciones que eviten la degradación de los conjuntos de sondas ligadas (p. ej., a temperatura ambiente o una temperatura por debajo de la temperatura ambiente) antes de llevar a cabo la etapa de recuento.

Con el fin de cuantificar con precisión la abundancia relativa de diferentes secuencias genómicas, por ejemplo, para

la cuantificación del número de copias de ADN o para la cuantificación de la frecuencia de alelos, se puede contar una gran cantidad de sondas producto. Por ejemplo, una marca puede ser detectada y contada en base a la medición, por ejemplo, de propiedades físicoquímicas, electromagnéticas, eléctricas, optoelectrónicas o electroquímicas, o características de la marca inmovilizada.

5 En algunas realizaciones, la marca se puede detectar mediante microscopía de sonda de barrido (SPM), microscopía de efecto túnel (STM) y microscopía de fuerza atómica (AFM), microscopía electrónica, técnicas de interrogación/detección óptica que incluyen, pero no se limitan a, microscopía óptica de barrido de campo cercano (NSOM), microscopía confocal y excitación de ondas evanescentes. Las versiones más específicas de estas técnicas incluyen microscopía confocal de campo lejano, microscopía de dos fotones, epi-iluminación de campo
10 amplio y microscopía de reflexión interna total (TIR). Muchas de las técnicas anteriores también se pueden utilizar en un modo espectroscópico. La detección real es mediante cámaras de dispositivos de carga acoplada (CCD) y CCD intensificados, fotodiodos y/o tubos fotomultiplicadores. En algunas realizaciones, la etapa de recuento comprende un análisis óptico, que detecta una propiedad óptica de una etiqueta. En realizaciones adicionales, el análisis óptico comprende un análisis de imagen como se describe en la presente memoria.

15 En otro aspecto, la etapa de recuento comprende la lectura del sustrato en el primer y segundo canales de imagen que corresponden a la primera y segunda marcas, respectivamente, y la producción de una o más imágenes del sustrato, en donde la primera y la segunda sondas de marcaje se pueden resolver en la una o más imágenes. En algunas realizaciones, la etapa de recuento comprende el filtrado espacial para la segmentación de imágenes. En realizaciones adicionales, la etapa de recuento comprende el análisis de segmentación de imágenes "water
20 shedding", o un método híbrido para la segmentación de imágenes.

Los métodos descritos en la presente memoria también pueden observar la frecuencia de diferentes alelos en el mismo locus genético (p. ej., dos alelos de un polimorfismo de un solo nucleótido dado). La precisión de estos métodos puede detectar cambios muy pequeños en la frecuencia (p. ej., tan bajos como aproximadamente 10, 5, 4,
25 3, 2, 1, 0,5, 0,1 o 0,01% o menos). Como ejemplo, en el caso de un trasplante de órganos, una muestra de sangre contendrá una firma genética muy diluida del órgano donado. Esta firma puede ser la presencia de un alelo que no se encuentra en el receptor del genoma del órgano donado. Los métodos descritos en la presente memoria pueden detectar desviaciones muy pequeñas en la frecuencia de los alelos (p. ej., tan bajas como aproximadamente 10, 5, 4,
30 3, 2, 1, 0,5, 0,1 o 0,01% o menos) y pueden identificar la presencia de ADN del donante en una muestra del anfitrión (p. ej., muestra de sangre). Un órgano trasplantado no saludable puede producir niveles elevados de ADN de donante en la sangre del anfitrión, un aumento de solo un pequeño porcentaje (p. ej., tan bajo como aproximadamente 10, 5, 4, 3, 2, 1, 0,5, 0,1 o 0,01% o menos). Los métodos descritos en la presente memoria pueden ser lo suficientemente sensibles como para identificar cambios en la frecuencia de los alelos con la sensibilidad necesaria y, por lo tanto, pueden determinar con precisión la presencia y las cantidades cambiantes de ADN del donante en la sangre del anfitrión.

35 En otro aspecto, los métodos de la presente descripción pueden comprender comparar los números primero y segundo para determinar la variación genética en la muestra genética. En algunas realizaciones, la etapa de comparación comprende obtener una estimación de un número relativo de las moléculas de nucleótido que tienen la primera y la segunda regiones de ácido nucleico de interés.

40 En otro aspecto, los métodos de la presente descripción pueden comprender el marcaje de la primera y segunda sondas de marcaje con primera y segunda marcas, respectivamente, antes de la etapa de contacto (p. ej., durante la fabricación de las sondas). El marcaje de la sonda se puede realizar simultáneamente o después de poner en contacto las sondas con la muestra genética, hibridar, ligar, amplificar y/o inmovilizar las sondas. Además, el marcaje de la sonda se puede realizar simultáneamente o antes de poner en contacto las sondas con la muestra genética, hibridar, ligar, amplificar y/o inmovilizar las sondas. El marcaje de una sonda puede comprender agregar, inmovilizar o unir una marca a la sonda mediante un enlace físico o químico. Las marcas se pueden colocar en cualquier lugar dentro de la secuencia de una sonda, incluso en el extremo 5' o 3'.

45 En otro aspecto, los métodos de la presente descripción pueden comprender el etiquetado de la primera y la segunda sondas de etiquetado con la primera y la segunda etiquetas, respectivamente, antes de la etapa de contacto (p. ej., durante la fabricación de las sondas). El etiquetado de la sonda se puede realizar simultáneamente o después de poner en contacto las sondas con la muestra genética, hibridar, ligar, amplificar y/o marcar las sondas. Además, el etiquetado de la sonda se puede realizar simultáneamente o antes de poner en contacto las sondas con la muestra genética, hibridar, ligar, amplificar, inmovilizar y/o marcar las sondas. El etiquetado de una sonda puede comprender agregar, inmovilizar o unir una etiqueta a la sonda mediante un enlace físico o químico. Las etiquetas se pueden colocar en cualquier lugar dentro de la secuencia de una sonda, incluso en el extremo 5' o 3'.

55 En otro aspecto, los conjuntos de sondas en la presente memoria se pueden diseñar para que tengan etiquetas de acuerdo con las ubicaciones predeterminadas en las que se van a inmovilizar las etiquetas. En algunas realizaciones, las etiquetas en todos los conjuntos de sondas configuradas para detectar una variación genética son las mismas y están configuradas para ser inmovilizadas en las mismas ubicaciones en el sustrato directa o indirectamente. En realizaciones adicionales, la primera y segunda etiquetas son iguales, y cada una de las demás etiquetas es diferente de la primera o la segunda etiquetas. En realizaciones adicionales, cada uno o un grupo de
60

miembros de la matriz de múltiples ubicaciones predeterminadas en un sustrato pueden tener una etiqueta única que se debe inmovilizar.

5 En otro aspecto, los conjuntos de sondas de acuerdo con algunas realizaciones se pueden amplificar, y se pueden producir conjuntos de sondas marcadas durante el proceso de amplificación. En otro aspecto, cada una de las sondas de marcaje puede comprender una secuencia de cebado directo o inverso, y cada una de las sondas de etiquetado puede comprender una secuencia de cebado inverso o directo correspondiente y una secuencia de nucleótidos de etiquetado como etiqueta. Las secuencias de cebado directo e inverso son las secuencias que están configuradas para hibridar con los cebadores directo e inverso correspondientes, respectivamente. En algunas realizaciones, la etapa de amplificación comprende amplificar (i) las primeras sondas de marcaje y etiquetado ligadas con los primeros cebadores directo e inverso que hibridan con las secuencias de cebado directo e inverso, respectivamente, en donde el primer cebador directo o inverso que hibrida con la primera sonda de marcaje comprende la primera marca, y (ii) las segundas sondas de marcaje y etiquetado ligadas con los segundos cebadores directo e inverso que hibridan con las secuencias de cebado directo e inverso, respectivamente, en donde el segundo cebador directo o inverso que hibrida con la segunda sonda de marcaje comprende la segunda marca. En realizaciones adicionales, las secuencias de nucleótidos de etiquetado amplificadas de las sondas de etiquetado se inmovilizan en una ubicación predeterminada sobre un sustrato, en donde las secuencias de nucleótidos de etiquetado amplificadas de la primera y segunda sondas de etiquetado son la primera y segunda etiquetas. En algunas realizaciones, la primera y segunda etiquetas son iguales y/o están configuradas para unirse a la misma ubicación en el sustrato. En otra realización, la primera y segunda etiquetas son diferentes y/o están configuradas para unirse a diferentes ubicaciones en el sustrato. En realizaciones adicionales, cuando las sondas se amplifican, el método comprende contar el número de marcas en las sondas y/o los conjuntos de sondas amplificados inmovilizados sobre el sustrato. Por ejemplo, el primer número es el número de la primera marca en el primer conjunto de sondas amplificadas inmovilizadas sobre el sustrato, y el segundo número es el número de la segunda marca en el segundo conjunto de sondas amplificadas inmovilizadas sobre el sustrato.

25 En otro aspecto, los conjuntos de sondas de acuerdo con algunas realizaciones se pueden amplificar, y se pueden producir conjuntos de sondas marcadas utilizando cebadores inversos marcados sin utilizar un cebador directo. En otro aspecto, cada una de las sondas de marcaje puede comprender una secuencia de cebado inverso, y cada una de las sondas de etiquetado puede comprender una secuencia de nucleótidos de etiquetado como etiqueta. En algunas realizaciones, la etapa de amplificación puede comprender amplificar (i) las primeras sondas de marcaje y etiquetado ligadas con un primer cebador inverso que hibrida con una primera secuencia de cebado inverso de la primera sonda de marcaje, en donde el primer cebador inverso comprende la primera marca, y (ii) las segundas sondas de marcaje y etiquetado ligadas con un segundo cebador inverso que hibrida con una segunda secuencia de cebado inverso de la segunda sonda de marcaje, en donde el segundo cebador inverso comprende la segunda marca. En realizaciones adicionales, las secuencias de nucleótidos de etiquetado amplificadas de las sondas de etiquetado se inmovilizan en una ubicación predeterminada sobre un sustrato, en donde las secuencias de nucleótidos de etiquetado amplificadas de la primera y segunda sondas de etiquetado son la primera y segunda etiquetas. En realizaciones adicionales, el primer número es el número de la primera marca en el primer conjunto de sondas amplificadas inmovilizadas sobre el sustrato, y el segundo número es el número de la segunda marca en el segundo conjunto de sondas amplificadas inmovilizadas sobre el sustrato.

40 En otro aspecto, los conjuntos de sondas ligadas de acuerdo con algunas realizaciones se pueden producir utilizando una reacción en cadena de la ligasa. En otro aspecto, el método descrito en la presente memoria comprende poner en contacto el tercer y cuarto conjuntos de sondas con la muestra genética, en donde el tercer conjunto de sondas comprende una tercera sonda de marcaje y una tercera sonda de etiquetado, y el cuarto conjunto de sondas comprende una cuarta sonda de marcaje y una cuarta sonda de etiquetado. El método puede comprender adicionalmente hibridar el primer y segundo conjuntos de sondas con la primera y segunda hebras de ácido nucleico efectoras de interés en moléculas de nucleótidos de hebra sencilla a partir de moléculas de nucleótidos de doble hebra de la muestra genética, respectivamente; e hibridar el tercer y cuarto conjuntos de sondas con las hebras de ácido nucleico antisentido de la primera y segunda hebras de ácido nucleico efectoras de interés, respectivamente. El método puede comprender adicionalmente producir un primer, segundo, tercero y cuarto conjuntos de sondas ligadas al menos ligando (i) la primera sonda de marcaje y la primera sonda de etiquetado, (ii) la segunda sonda de marcaje y la segunda sonda de etiquetado, (iii) la tercera sonda de marcaje y la tercera sonda de etiquetado, y (iv) la cuarta sonda de marcaje y la cuarta sonda de etiquetado. El método puede comprender adicionalmente realizar una reacción en cadena de la ligasa conocida en la técnica para amplificar la sonda ligada y/o los conjuntos de sondas ligadas. En algunas realizaciones, la reacción en cadena de la ligasa puede comprender hibridar un primer, segundo, tercero y cuarto conjuntos de sondas no ligadas con el tercer, cuarto, primer y segundo los conjuntos de sondas ligadas, respectivamente, y ligar al menos (i) la primera sonda de marcaje y la primera sonda de etiquetado, (ii) la segunda sonda de marcaje y la segunda sonda de etiquetado, (iii) la tercera sonda de marcaje y la tercera sonda de etiquetado, y (iv) la cuarta sonda de marcaje y la cuarta sonda de etiquetado de los conjuntos de sondas no ligadas. El método puede comprender adicionalmente inmovilizar las sondas de etiquetado en la ubicación predeterminada sobre un sustrato, en donde la primera, segunda, tercera y cuarta sondas de marcaje ligadas a la primera, segunda, tercera y cuarta sondas de etiquetado inmovilizadas, respectivamente, comprenden una primera, segunda, tercera y cuarta marcas, respectivamente; las marcas inmovilizadas se pueden resolver ópticamente; la primera, segunda, tercera y cuarta sondas de etiquetado inmovilizadas comprenden una

primera, segunda, tercera y cuarta etiquetas, respectivamente, y la etapa de inmovilización se realiza inmovilizando las etiquetas en la ubicación predeterminada. El método puede comprender adicionalmente contar (i) la primera suma de la primera y tercera marcas inmovilizadas sobre el sustrato, y (ii) la segunda suma de las segunda y cuarta marcas inmovilizadas sobre el sustrato, y comparar las primeras y segundas sumas para determinar la variación genética en la muestra genética. En otras realizaciones adicionales, el método comprende adicionalmente marcar la primera, segunda, tercera y cuarta sondas de marcaje con la primera, segunda, tercera y cuarta marcas, respectivamente, antes de la etapa de contacto. En otras realizaciones adicionales, la primera y tercera marcas son iguales, y la segunda y cuarta marcas son iguales.

En otro aspecto, el método descrito en la presente memoria comprende poner en contacto el tercer y cuarto conjuntos de sondas con la muestra genética, en donde el tercer conjunto de sondas comprende una tercera sonda de marcaje y una tercera sonda de etiquetado, y el cuarto conjunto de sondas comprende una cuarta sonda de marcaje y una cuarta sonda de etiquetado, la primera y la tercera sondas de marcaje comprenden una primera secuencia de cebado inverso, la segunda y cuarta sondas de marcaje comprenden una segunda secuencia de cebado inverso, y cada una de las sondas de etiquetado comprende una secuencia de nucleótidos de etiquetado como etiqueta. El método puede comprender adicionalmente hibridar el primer y segundo conjuntos de sondas con la primera y segunda hebras de ácido nucleico efectoras de interés, respectivamente, en moléculas de nucleótidos de hebra sencilla a partir de moléculas de nucleótidos de doble hebra de la muestra genética; e hibridar al menos partes del tercer y cuarto conjuntos de sondas con hebras de ácido nucleico antisentido de la primera y segunda hebras de ácido nucleico efectoras de interés, respectivamente; producir el primer, segundo, tercer y cuarto conjuntos de sondas ligadas ligando (i) la primera sonda de marcaje y la primera sonda de etiquetado, (ii) la segunda sonda de marcaje y la segunda sonda de etiquetado, (iii) la tercera sonda de marcaje y la tercera sonda de etiquetado, y (iv) la cuarta sonda de marcaje y la cuarta sonda de etiquetado. El método puede comprender adicionalmente realizar una reacción en cadena de la ligasa. En algunas realizaciones, la reacción en cadena de la ligasa comprende hibridar al menos partes del primer, segundo, tercer y cuarto conjuntos de sondas no ligadas con el tercero, cuarto, primer y segundo conjuntos de sondas ligadas, respectivamente, y ligar (i) la primera sonda de marcaje y la primera sonda de etiquetado, (ii) la segunda sonda de marcaje y la segunda sonda de etiquetado, (iii) la tercera sonda de marcaje y la tercera sonda de etiquetado, y (iv) la cuarta sonda de marcaje y la cuarta sonda de etiquetado del conjunto de sondas no ligadas. El método puede comprender adicionalmente amplificar (i) el primer y tercer conjuntos de sondas ligadas con un primer cebador inverso que hibrida con la primera secuencia de cebado inverso, en donde el primer cebador inverso comprende la primera marca, y (ii) el segundo y cuarto conjuntos de sondas ligadas con un segundo cebador inverso que hibrida con la segunda secuencia de cebado inverso, en donde el segundo cebador inverso comprende la segunda marca, las secuencias de nucleótidos de etiquetado amplificadas de las sondas de etiquetado se inmovilizan en una ubicación predeterminada sobre un sustrato, en donde las secuencias de nucleótidos de etiquetado amplificadas de la primera, segunda, tercera y cuarta sondas de etiquetado son la primera, segunda, tercera y cuarta etiquetas, el primer número es el número de la primera marca en el primer y tercer conjuntos de sondas amplificadas inmovilizados sobre sustrato, y el segundo número es el número de la segunda marca en el segundo y cuarto conjuntos de sondas amplificadas inmovilizado sobre el sustrato.

En otro aspecto, la primera y segunda sondas de marcaje ligadas están en el extremo 3' del primer y segundo conjuntos de sondas ligadas y comprenden la primera y la segunda secuencias de cebado inverso que hibridan con el primer y el segundo cebadores inversos, respectivamente. En algunas realizaciones, el primer y segundo cebadores inversos comprenden la primera y segunda marcas. En realizaciones adicionales, la primera y segunda sondas de etiquetado ligadas están en el extremo 5' del primer y segundo conjunto de sondas ligadas. En realizaciones adicionales, la primera y segunda sondas de etiquetado ligadas están en el extremo 5' del primer y segundo conjunto de sondas ligadas y comprenden la primera y la segunda secuencias de cebado directo correspondientes que hibridan con el primero y segundo cebadores directos, respectivamente.

En otro aspecto, el método en la presente memoria comprende digerir moléculas de doble hebras en la muestra para producir moléculas de hebra sencillas. En algunas realizaciones, la etapa de amplificación comprende poner en contacto una exonucleasa con la sonda y/o el conjunto de sondas amplificadas, y digerir la sonda y/o conjunto de sondas amplificadas desde el extremo 5' de una hebra de la sonda y/o conjunto de sondas amplificadas de doble hebra. Por ejemplo, la etapa de amplificación comprende poner en contacto una exonucleasa con la sonda amplificadas en un conjunto de sondas, y digerir el conjunto de sondas amplificadas a partir del extremo 5' de una hebra del conjunto de sondas amplificadas de doble hebra. En realizaciones adicionales, la única hebra de la sonda y conjunto de sondas amplificadas en contacto con la exonucleasa no tiene ninguna marca en el extremo 5'. El contacto de la exonucleasa con las sondas de doble hebras no marcadas puede digerir la hebra no marcada del extremo 5' produciendo sondas de hebra sencilla. En otro aspecto, el extremo 5' del conjunto de sondas amplificadas que comprende la marca en el extremo 5' se puede proteger de la digestión con exonucleasa.

En otro aspecto, el método puede detectar de 1 a 100, de 1 a 50, de 2 a 40 o de 5 a 10 variaciones genéticas; 2, 3, 4, 5, 6, 7, 8, 9, 10 o más variaciones genéticas; y 100, 50, 30, 20, 10 o menos variaciones genéticas. En algunas realizaciones, el método descrito en la presente memoria puede detectar x número de variaciones genéticas utilizando al menos (x + 1) número de diferentes conjuntos de sondas. En estas realizaciones, se pueden comparar un número de marcas de un tipo de conjuntos de sondas con uno o más números de marcas del resto de los diferentes tipos de conjuntos de sondas. En algunas realizaciones, el método descrito en la presente memoria puede detectar variaciones genéticas de manera continua en todo el genoma a varias resoluciones, por ejemplo, a una

resolución de 300.000 bases, de manera que 100 variaciones distribuidas en todos los cromosomas se interrogan y cuantifican por separado. En realizaciones adicionales, la resolución de bases está en el intervalo de uno o diez a cien mil nucleótidos hasta un millón, diez millones o 100 millones o más nucleótidos.

5 En otro aspecto, el método de acuerdo con algunas realizaciones puede detectar al menos dos variaciones genéticas. En algunas realizaciones, el método descrito en la presente memoria puede comprender adicionalmente poner en contacto un quinto conjunto de sondas con la muestra genética, en donde el quinto conjunto de sondas comprende una quinta sonda de marcaje y una quinta sonda de etiquetado. El método puede comprender adicionalmente hibridar al menos una parte del quinto conjunto de sondas con la tercera región de ácido nucleico de interés en moléculas de nucleótidos de la muestra genética, en donde la tercera región de ácido nucleico de interés es diferente de la primera y segunda regiones de ácido nucleico de interés. El método puede comprender adicionalmente ligar el quinto conjunto de sondas ligando al menos la quinta sonda de marcaje y la quinta sonda de etiquetado. El método puede comprender adicionalmente amplificar los conjuntos de sondas ligadas. El método puede comprender adicionalmente inmovilizar cada una de las sondas de etiquetado en una ubicación predeterminada sobre un sustrato, en donde la quinta sonda de marcaje y/o la sonda de marcaje amplificada de la misma ligada a la sonda de etiquetado inmovilizada comprenden una quinta marca, la quinta etiqueta es diferente de la primera y segunda marcas, las marcas inmovilizadas se pueden resolver ópticamente, la quinta sonda de etiquetado inmovilizada y/o la sonda de etiquetado amplificada de la misma comprenden una quinta etiqueta, y la etapa de inmovilización se realiza inmovilizando las etiquetas en la ubicación predeterminada. El método puede comprender contar un tercer número de la quinta marca inmovilizada sobre sustrato, y comparar el tercer número con el primer y/o segundo números para determinar la segunda variación genética en la muestra genética. En algunas realizaciones, el sujeto puede ser un sujeto gestante, la primera variación genética es la trisomía 21 en el feto del sujeto gestante y la segunda variación genética se selecciona del grupo que consiste en la trisomía 13, trisomía 18, aneuploidía de X y aneuploidía de Y en el feto del sujeto gestante.

25 En otro aspecto, el método de acuerdo con algunas realizaciones puede detectar al menos tres variaciones genéticas. En algunas realizaciones, el método descrito en la presente memoria comprende adicionalmente poner en contacto un sexto conjunto de sondas con la muestra genética, en donde el sexto conjunto de sondas comprende una sexta sonda de marcaje y una sexta sonda de etiquetado. El método puede comprender adicionalmente hibridar al menos una parte del sexto conjunto de sondas con la cuarta región de ácido nucleico de interés en moléculas de nucleótidos de la muestra genética, en donde la cuarta región de ácido nucleico de interés es diferente de la primera, segunda y tercera regiones de ácido nucleico de interés. El método puede comprender adicionalmente ligar el sexto conjunto de sondas ligando al menos la sexta sonda de marcaje y la sexta sonda de etiquetado. El método puede comprender adicionalmente amplificar los conjuntos de sondas ligadas. El método puede comprender adicionalmente la inmovilización de cada una de las sondas de etiquetado en una ubicación predeterminada sobre un sustrato, en donde la sexta sonda de marcaje y/o la sonda de marcaje amplificada de la misma ligadas a la sonda de etiquetado inmovilizada comprenden una sexta marca, la sexta marca es diferente de la primera y segunda marcas, las marcas inmovilizadas se pueden resolver ópticamente, la sexta sonda de etiquetado inmovilizada y/o la sonda de etiquetado amplificada de la misma comprenden una sexta etiqueta, y la etapa de inmovilización se realiza inmovilizando las etiquetas en la ubicación predeterminada. El método puede comprender adicionalmente contar un cuarto número de la sexta marca inmovilizada sobre sustrato, y comparar el cuarto número con el primer, segundo y/o tercer números para determinar la tercera variación genética en la muestra genética.

40 En otro aspecto, el método puede, según algunas realizaciones, detectar al menos cuatro variaciones genéticas. En algunas realizaciones, el método descrito en la presente memoria comprende adicionalmente poner en contacto un séptimo conjunto de sondas con la muestra genética, en donde el séptimo conjunto de sondas comprende una séptima sonda de marcaje y una séptima sonda de etiquetado. El método puede comprender adicionalmente hibridar al menos una parte del séptimo conjunto de sondas con la quinta región de ácido nucleico de interés en moléculas de nucleótidos de la muestra genética, en donde la quinta región de ácido nucleico de interés es diferente de la primera, segunda, tercera y cuarta regiones de ácido nucleico de interés. El método puede comprender adicionalmente ligar el séptimo conjunto de sondas al menos ligando la séptima sonda de marcaje y la séptima sonda de etiquetado. El método puede comprender adicionalmente amplificar opcionalmente los conjuntos de sondas ligadas. El método puede comprender adicionalmente inmovilizar cada una de las sondas de etiquetado en una ubicación predeterminada sobre un sustrato, en donde la séptima sonda de marcaje y/o la sonda de marcaje amplificada de la misma ligadas a la sonda de etiquetado inmovilizada comprenden una séptima marca, la séptima marca es diferente de la primera y segunda marcas, las marcas inmovilizadas se pueden resolver ópticamente, la séptima sonda de etiquetado inmovilizada y/o la sonda de etiquetado amplificada de la misma comprenden una séptima etiqueta, y la etapa de inmovilización se realiza inmovilizando las etiquetas en la ubicación predeterminada. El método puede comprender adicionalmente contar un quinto número de la séptima etiqueta inmovilizada sobre el sustrato, y comparar el quinto número con el primer, segundo, tercer y/o cuarto números para determinar la cuarta variación genética en la muestra genética.

50 En otro aspecto, el método de acuerdo con algunas realizaciones puede detectar al menos cinco variaciones genéticas. En algunas realizaciones, el método descrito en la presente memoria comprende adicionalmente poner en contacto un octavo conjunto de sondas con la muestra genética, en donde el octavo conjunto de sondas comprende una octava sonda de marcaje y una octava sonda de etiquetado. El método puede comprender adicionalmente hibridar al menos una parte del octavo conjunto de sondas con la sexta región de ácido nucleico de interés en

moléculas de nucleótidos de la muestra genética, en donde la sexta región de ácido nucleico de interés es diferente de la primera, segunda, tercera, cuarta, y quinta regiones de ácido nucleico de interés. El método puede comprender adicionalmente ligar el octavo conjunto de sondas ligando al menos la octava sonda de marcaje y la octava sonda de etiquetado. El método puede comprender adicionalmente amplificar los conjuntos de sondas ligadas. El método puede comprender adicionalmente inmovilizar cada una de las sondas de etiquetado en una ubicación predeterminada sobre un sustrato, en donde la octava sonda de marcaje y/o la sonda de marcaje amplificada de la misma ligadas a la sonda de etiquetado inmovilizada comprenden una octava etiqueta, la octava marca es diferente de la primera y segunda marcas, las marcas inmovilizadas se pueden resolver ópticamente, la octava sonda de etiquetado inmovilizada y/o la sonda de etiquetado amplificada de la misma comprenden una octava etiqueta, y la etapa de inmovilización se realiza inmovilizando las etiquetas en la ubicación predeterminada. El método puede comprender adicionalmente contar un sexto número de la octava etiqueta inmovilizada sobre sustrato, y comparar el sexto número con el primer, segundo, tercer, cuarto y/o quinto números para determinar la quinta variación genética en la muestra genética. En algunas realizaciones, el sujeto es un sujeto gestante, y la primera, segunda, tercera, cuarta y quinta variaciones genéticas son la trisomía 13, la trisomía 18, la trisomía 21, aneuploidía X y aneuploidía Y en el feto del sujeto gestante.

En otro aspecto, el sujeto es un sujeto gestante, la variación genética es la trisomía 21 en el feto del sujeto gestante, la primera región de ácido nucleico de interés se ubica en el cromosoma 21 y la segunda región de ácido nucleico de interés no se ubica en el cromosoma 21.

En otro aspecto, el sujeto es un sujeto gestante, la variación genética es la trisomía 21 en el feto del sujeto gestante, la primera región de ácido nucleico de interés se ubica en el cromosoma 21 y la segunda región de ácido nucleico de interés se ubica en el cromosoma 18.

En un aspecto, el conjunto de sondas de la presente memoria puede comprender dos, tres, cuatro, cinco o más sondas de marcaje, y/o dos, tres, cuatro, cinco o más marcas. En algunas realizaciones, el método descrito en la presente memoria puede comprender adicionalmente el primer y segundo conjuntos de sondas, comprender adicionalmente la tercera y cuarta sondas de marcaje, respectivamente; el primer conjunto de sondas inmovilizadas y/o el primer conjunto de sondas amplificadas comprenden adicionalmente una novena marca en la tercera sonda de marcaje y/o su producto amplificado; y el segundo conjunto de sondas inmovilizadas y/o el segundo conjunto de sondas amplificadas comprenden adicionalmente una décima marca en la cuarta sonda de marcaje y/o su producto amplificado. En estas realizaciones, si la novena y décima marcas son diferentes de la primera y segunda marcas, este método se puede utilizar para confirmar el número contado para la primera y segunda marcas. Si la novena y décima marcas son iguales que la primera y segunda marcas, respectivamente, este método se puede utilizar para mejorar la precisión de las marcas de detección inmovilizadas en cada una de las regiones de ácido nucleico de interés. Por ejemplo, el uso de múltiples marcas sería más brillante que el uso de una marca y, por lo tanto, se pueden detectar más fácilmente múltiples marcas que una sola marca.

En realizaciones adicionales, (i) el primer conjunto de sondas inmovilizadas y/o el primer conjunto de sondas amplificadas comprenden adicionalmente una undécima marca en la sonda de marcaje, y (ii) el segundo conjunto de sondas inmovilizadas y/o el segundo conjunto de sondas amplificadas comprenden adicionalmente una duodécima marca que es diferente de la undécima marca en la sonda de marcaje. En realizaciones adicionales, en donde la primera, segunda, undécima y duodécima marcas son diferentes entre sí, y la etapa de recuento comprende adicionalmente los números de recuento de la undécima y duodécima marcas inmovilizadas sobre el sustrato.

En otro aspecto, el método descrito en la presente memoria se puede realizar con una muestra de control. En algunas realizaciones, el método puede comprender adicionalmente repetir las etapas con una muestra de control diferente de la muestra genética del sujeto. El método puede comprender adicionalmente contar los números de control de las marcas inmovilizadas sobre el sustrato, y comparar los números de control con el primer, segundo, tercer, cuarto, quinto y/o sexto números para confirmar la variación genética en la muestra genética.

En otro aspecto, el sujeto puede ser un sujeto gestante, y la variación genética es una variación genética en el feto del sujeto gestante. En tales realizaciones, el método puede utilizar un sitio de polimorfismo de un solo nucleótido (SNP) para determinar si la proporción (p. ej., concentración y porcentaje en número en función del número de moléculas de nucleótido en la muestra) de material fetal (p. ej., fracción fetal) es suficiente para que la variación genética del feto se pueda detectar en una muestra del sujeto gestante con una significación estadística razonable. En realizaciones adicionales, el método puede comprender adicionalmente poner en contacto conjuntos de sondas maternas y paternas con la muestra genética, en donde el conjunto de sondas maternas comprende una sonda de marcaje materna y una sonda de etiquetado materna, y el conjunto de sondas paternas comprende una sonda de marcaje paterna y una sonda de etiquetado paterna. El método puede comprender adicionalmente hibridar al menos una parte de cada uno de los conjuntos de sondas maternas y paternas en una región de ácido nucleico de interés en moléculas de nucleótidos de la muestra genética, comprendiendo la región de ácido nucleico de interés un sitio SNP predeterminado, en donde al menos una parte del conjunto de sondas maternas hibrida con un primer alelo en el sitio del SNP, la al menos una parte del conjunto de sondas paternas hibrida con un segundo alelo en el sitio del SNP, y el primer y segundo alelos son diferentes entre sí. El método puede comprender adicionalmente ligar el material y los conjuntos de sondas paternas ligando al menos (i) las sondas de marcaje y etiquetado maternas, y (ii) las sondas de marcaje y etiquetado paternas. El método puede comprender adicionalmente amplificar las sondas

ligadas. El método puede comprender adicionalmente inmovilizar las sondas de etiquetado en una ubicación predeterminada sobre un sustrato, en donde las sondas de marcaje maternas y paternas y/o las sondas de marcaje amplificadas de las mismas ligadas a las sondas de etiquetado inmovilizadas comprenden marcas maternas y paternas, respectivamente; Las marcas maternas y paternas son diferentes, y las marcas inmovilizadas se pueden resolver ópticamente. El método puede comprender adicionalmente contar los números de las marcas maternas y paternas, y determinar si una proporción de material fetal en la muestra genética es suficiente para detectar la variación genética en el feto basándose en los números de marcas maternas y paternas. El método puede comprender adicionalmente determinar la proporción de material fetal en la muestra genética.

En algunas realizaciones, cuando el sujeto es un sujeto gestante y la variación genética es una variación genética en el feto del sujeto gestante, el método puede comprender adicionalmente poner en contacto conjuntos de sondas del alelo A y alelo B que son específicas de cada alelo con la muestra genética, en donde el conjunto de sondas del alelo A comprende una sonda de marcaje del alelo A y una sonda de etiquetado del alelo A, y el conjunto de sondas del alelo B comprende una sonda de marcaje del alelo B y una sonda de etiquetado del alelo B. El método puede comprender adicionalmente hibridar al menos una parte de cada uno de los conjuntos de sondas del alelo A y alelo B con una región de ácido nucleico de interés en moléculas de nucleótidos de la muestra genética, comprendiendo la región de ácido nucleico de interés un sitio de polimorfismo de un solo nucleótido predeterminado (SNP) para el que un perfil alélico materno (es decir, genotipo) difiere de un perfil alélico fetal en el sitio SNP (p. ej., la composición alélica materna puede ser AA y la composición alélica fetal puede ser AB, o BB. En otro ejemplo, la composición alélica materna puede ser AB y la composición alélica fetal puede ser AA o BB), en donde la al menos una parte del conjunto de sondas del alelo A hibrida con un primer alelo en el sitio SNP, la al menos una parte del conjunto de sondas del alelo B hibrida a un segundo alelo en el sitio SNP, y el primer y segundo alelos son diferentes entre sí. El método puede comprender adicionalmente ligar los conjuntos de sondas del alelo A y alelo B ligando al menos (i) las sondas de marcaje y etiquetado del alelo A, y (ii) las sondas de marcaje y etiquetado del alelo B. El método puede comprender adicionalmente amplificar los conjuntos de sondas ligadas. El método puede comprender adicionalmente inmovilizar las sondas de etiquetado en una ubicación predeterminada sobre un sustrato, en donde las sondas de marcaje del alelo A y el alelo B y/o las sondas de marcaje amplificadas de las mismas ligadas a las sondas de etiquetado inmovilizadas comprenden las marcas del alelo A y del alelo B, respectivamente, las marcas del alelo A y del alelo B son diferentes, y las marcas inmovilizadas se pueden resolver ópticamente. El método puede comprender adicionalmente contar los números de las marcas del alelo A y del alelo B, y determinar si una proporción de un material fetal en la muestra genética es suficiente para detectar la variación genética en el feto basándose en los números de las marcas del alelo A y del alelo B. El método puede comprender adicionalmente determinar la proporción del material fetal en la muestra genética.

En algunas realizaciones, cuando el sujeto es un sujeto gestante, la variación genética es una variación genética en el feto del sujeto gestante y la muestra genética comprende un cromosoma Y, el método puede comprender adicionalmente el contacto de conjuntos de sondas maternas y paternas con la muestra genética, en donde el conjunto de sondas maternas comprende una sonda de marcaje materna y una sonda de etiquetado materna, y el conjunto de sondas paternas comprende una sonda de marcaje paterna y una sonda de etiquetado paterna. El método puede comprender adicionalmente hibridar al menos partes de los conjuntos de sondas maternas y paternas con las regiones de ácido nucleico maternas y paternas de interés en las moléculas de nucleótidos de la muestra genética, respectivamente, en donde la región de ácido nucleico paterna de interés está ubicada en el cromosoma Y, y la región de ácido nucleico materno de interés no está localizada en el cromosoma Y. El método puede comprender adicionalmente ligar los conjuntos de sondas maternas y paternas ligando al menos (i) las sondas de marcaje y etiquetado maternas, y (ii) las sondas de marcaje y etiquetado paternas. El método puede comprender adicionalmente amplificar las sondas ligadas. El método puede comprender adicionalmente una región de ácido nucleico de interés que comprende un sitio de polimorfismo de un solo nucleótido (SNP) predeterminado que contiene más de un SNP, por ejemplo, dos o tres SNP. Adicionalmente, el sitio SNP puede contener SNP con un alto desequilibrio de enlace, de modo que las sondas de marcaje y etiquetado están configuradas para aprovechar la mejora energética de múltiples coincidencias o discrepancias de SNP en comparación con solo una. El método puede comprender adicionalmente inmovilizar las sondas de etiquetado en una ubicación predeterminada sobre un sustrato, en donde las sondas de marcaje maternas y paternas y/o las sondas de marcaje amplificadas de las mismas ligadas a las sondas de etiquetado inmovilizadas comprenden marcas maternas y paternas, respectivamente, las marcas maternas y paternas son diferentes, y las marcas inmovilizadas se pueden resolver ópticamente. El método puede comprender adicionalmente contar los números de las marcas maternas y paternas, y determinar si una proporción de material fetal en la muestra genética es suficiente para detectar la variación genética en el feto basándose en los números de las marcas maternas y paternas. El método puede comprender adicionalmente determinar la proporción del material fetal en la muestra genética.

En realizaciones adicionales, se pueden utilizar otras variaciones genéticas (p. ej., eliminación de una sola base, microsatélite y pequeñas inserciones) en lugar de la variación genética en el sitio SNP descrito en la presente memoria.

En un aspecto, el conjunto de sondas descrito en la presente memoria puede comprender tres o más sondas, que incluyen al menos una sonda entre las sondas de marcaje y etiquetado. En algunas realizaciones, el primer y segundo conjuntos de sondas comprenden adicionalmente la primera y segunda sondas de huecos, respectivamente; la primera sonda de huecos hibrida con una región entre las regiones donde se hibridan la primera

sonda de marcaje y la primera sonda de etiquetado; la segunda sonda de huecos hibrida con una región entre las regiones donde hibridan la segunda sonda de marcaje y la segunda sonda de etiquetado. El método puede comprender adicionalmente la etapa de ligación que comprende ligar al menos (i) la primera sonda de marcaje, la primera sonda de etiquetado y la primera sonda de huecos, y (ii) la segunda sonda de marcaje, la segunda sonda de etiquetado y la segunda sonda de huecos. En realizaciones adicionales, la sonda de huecos puede comprender una marca. Por ejemplo, la primera y la segunda sondas de huecos y/o sus productos amplificados están marcados con marcas (p. ej., decimotercera y decimocuarta marcas, respectivamente), y cada una de las marcas puede ser diferente del resto de las marcas (p. ej., la primera y segunda marcas). Las marcas en las sondas de huecos (p. ej., decimotercera y decimocuarta marcas) pueden ser iguales o diferentes entre sí. En otro aspecto, la primera y segunda sondas de marcaje hibridan con la primera y la segunda regiones de ácido nucleico de interés en moléculas de nucleótidos de la muestra genética, respectivamente; la primera y segunda sondas de etiquetado hibridan con la primera y la segunda regiones de ácido nucleico de interés en las moléculas de nucleótidos de la muestra genética, respectivamente; la primera y la segunda sondas de huecos hibridan con la primera y la segunda regiones de ácido nucleico de interés en las moléculas de nucleótidos de la muestra genética, respectivamente. En algunas realizaciones, hay de 0 a 100 nucleótidos, de 1 a 100 nucleótidos, de 2 a 50 nucleótidos; de 3 a 30 nucleótidos, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 100, 150 o 200 o más; o 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 25, 35, 45, 55, 110, 160 o 300 o menos entre las regiones donde hibridan la primera sonda de marcaje y las sondas de etiquetado y hay de 0 a 100 nucleótidos, 1 a 100 nucleótidos, 2 a 50 nucleótidos; 3 a 30 nucleótidos, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 100, 150 o 200 nucleótidos o más; o 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 25, 35, 45, 55, 110, 160 o 300 nucleótidos o menos entre las regiones donde hibridan la segunda sonda de marcaje y las sondas de etiquetado. En realizaciones adicionales, la sonda de huecos entre una sonda de marcaje y una sonda de etiquetado puede tener una longitud de 0 a 100 nucleótidos, 1 a 100 nucleótidos, 2 a 50 nucleótidos; 3 a 30 nucleótidos, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 100, 150 o 200 o más; o 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 25, 35, 45, 55, 110, 160 o 300 o menos.

En otro aspecto, el conjunto de sondas descrito en la presente memoria puede comprender un espaciador ligado y/o conjugado con la sonda de marcaje y la sonda de etiquetado. El espaciador puede comprender o no oligonucleótidos. El espaciador puede comprender un material aislado, purificado, natural o no natural, que incluye un oligonucleótido de cualquier longitud (p. ej., 5, 10, 20, 30, 40, 50, 100 o 150 nucleótidos o menos). En algunas realizaciones, la sonda puede estar en un producto digerido con enzimas de restricción purificado o se puede producir mediante síntesis, de forma recombinante o mediante amplificación por PCR. Por ejemplo, las primeras sondas de marcaje y etiquetado se conjugan mediante un primer espaciador, las segundas sondas de marcaje y etiquetado se conjugan mediante un segundo espaciador, y el primer y segundo espaciadores no hibridan con las moléculas de nucleótidos de la muestra genética. En algunas realizaciones, el método comprende adicionalmente digerir la muestra genética hibridada con una enzima y romper un enlace en el primer y segundo espaciadores después de la digestión.

En otro aspecto, el método descrito en la presente memoria excluye la identificación de una secuencia en las moléculas de nucleótidos de la muestra genética, y/o la secuenciación de la región o regiones de ácido nucleico de interés y/o las sondas. En algunas realizaciones, el método que excluye la secuenciación de las sondas incluye la exclusión de la secuenciación de un código de barras y/o una etiqueta de afinidad en una sonda de etiquetado. En realizaciones adicionales, no se necesita que los conjuntos de sondas inmovilizadas para detectar diferentes variaciones genéticas, regiones de nucleótidos de interés y/o péptidos de interés sean detectadas o escaneadas por separado ya que no se requiere la secuenciación en los métodos descritos en la presente memoria. En realizaciones adicionales, los números de diferentes marcas inmovilizadas sobre el sustrato se contaron simultáneamente (p. ej., mediante un solo escaneo y/o generación de imagen), y así los números de las diferentes marcas no se contaron por separado. En otro aspecto, el método descrito en la presente memoria excluye la lectura de la matriz global o la cuantificación analógica. La lectura de la matriz global significa en la presente memoria una única medición que mide la señal combinada acumulativa de múltiples marcas de un solo tipo, opcionalmente combinada con una segunda medición de la señal combinada acumulativa de numerosas marcas de un segundo tipo, sin resolver una señal de cada marca. Se obtiene un resultado de la combinación de una o más de tales mediciones en las que no se resuelven las marcas individuales. En otro aspecto, el método descrito en la presente memoria puede incluir una única medición que mide las mismas marcas, diferentes marcas del mismo tipo y/o marcas del mismo tipo en las que se resuelven las marcas individuales. El método descrito en la presente memoria puede excluir la cuantificación analógica y puede emplear una cuantificación digital, en la que solo se determina el número de marcas (que se determina a través de mediciones de la intensidad y la forma de las marcas individuales), y no la intensidad óptica acumulada o combinada de las marcas.

En otro aspecto, el conjunto de sondas descrito en la presente memoria puede comprender un aglutinante. Un aglutinante es el mismo material que la etiqueta o etiqueta de afinidad que se describe en la presente memoria. En algunas realizaciones, el método comprende adicionalmente inmovilizar el aglutinante sobre una fase sólida después de las etapas de ligación. El método puede comprender adicionalmente aislar los conjuntos de sondas ligadas de las sondas no ligadas. En realizaciones adicionales, el aglutinante comprende biotina, y la fase sólida comprende una cuenta magnética.

En otro aspecto, la etapa de recuento descrita en la presente memoria puede comprender adicionalmente calibrar, verificar y/o confirmar los números contados. Calibrar significa en la presente memoria verificar y/o ajustar la

precisión del número contado. Verificación y confirmación significan en la presente memoria determinar si el número contado es exacto o no, y/o cuánto es el error, si existe.

En otro aspecto, la intensidad y/o señal respecto a ruido se utilizan como un método para identificar marcas individuales. Cuando las moléculas de colorante u otras marcas ópticas están muy cerca, a menudo son imposibles de discriminar con imágenes basadas en fluorescencia debido al límite intrínseco de la difracción de la luz. Es decir, dos marcas que están juntas no serán distinguibles sin un hueco visible entre ellas. Un método ilustrativo para determinar el número de marcas en una ubicación dada consiste en examinar la señal relativa y/o la señal respecto a ruido en comparación con las ubicaciones que se sabe que tienen un solo flúor. Generalmente, dos o más marcas emitirán una señal más brillante (y una que se puede diferenciar más claramente del fondo) que un solo flúor. La Figura 2 muestra el histograma normalizado de la intensidad de la señal medida tanto de muestras de marcas individuales como de anticuerpos con múltiples marcas (ambos Alexa 546; verificados a través de perfiles de blanqueamiento). Las dos poblaciones fueron claramente separables, y las marcas múltiples se pueden distinguir claramente de las marcas individuales.

En algunas realizaciones, la etapa de recuento puede comprender medir señales ópticas de las marcas inmovilizadas, y calibrar los números contados distinguiendo una señal óptica de una marca única del resto de las señales ópticas del fondo y/o marcas múltiples. En algunas realizaciones, la distinción comprende calcular una intensidad de señal relativa y/o la señal respecto a ruido de la señal óptica en comparación con la intensidad de una señal óptica de una sola marca. La distinción puede comprender adicionalmente determinar si la señal óptica es de una sola marca. En realizaciones adicionales, la señal óptica es de una sola marca si la intensidad de señal relativa y/o de la señal respecto a ruido de una señal óptica difiere de la intensidad de una señal óptica de una sola etiqueta en una cantidad predeterminada o menos. En realizaciones adicionales, la cantidad predeterminada es de 0% a 100%, de 0% a 150%, de 10% a 200%, 0, 1, 2, 3, 4, 5, 10, 20, 30 o 40% o más, y/o 300, 200, 100, 50, 30, 10 o 5% o menos de la intensidad de la señal óptica de una sola marca.

En otro aspecto, diferentes marcas pueden tener diferentes propiedades de parpadeo y blanqueo. También pueden tener diferentes propiedades de excitación. Con el fin de comparar el número de moléculas de colorante para dos marcas diferentes, es necesario asegurarse de que los dos colorantes se comporten de manera similar y tengan características de emisión similares. Por ejemplo, si un colorante es mucho más tenue que otro, el número de moléculas puede ser subestimado en este canal. Se pueden titular varios factores para proporcionar la equivalencia óptima entre los colorantes. Por ejemplo, la etapa de recuento y/o la etapa de calibración pueden comprender optimizar (i) las potencias de las fuentes de luz para excitar las marcas, (ii) los tipos de fuentes de luz, (iii) los tiempos de exposición de las marcas y/o (iv) los conjuntos de filtros para que las marcas coincidan con las señales ópticas de las marcas y midan las señales ópticas de las marcas. Estos factores se pueden variar individualmente o combinados. Adicionalmente, la métrica que se optimiza puede variar. Por ejemplo, puede ser intensidad global, señal con respecto a ruido, menor fondo, varianza más baja en intensidad o cualquier otra característica.

Los perfiles de blanqueo son específicos para cada marca y se pueden utilizar para añadir información para distinguir los tipos de marcas. La Figura 3 muestra los perfiles de blanqueo promedio de varias marcas. La gráfica muestra los recuentos normalizados por tipo de marca en función de las imágenes sucesivas que se recopilaron en un intervalo de 60 segundos. El elemento c1 es Cy3 flúor, el elemento c2 es Atto647 flúor y el elemento c3 es Alexa488 flúor.

En otro aspecto, el comportamiento de parpadeo se puede utilizar como un método para identificar marcas individuales. Se sabe que muchas moléculas de colorante pasan temporalmente a un estado oscuro (p. ej., Burnette et al., Proc. Natl Acad Sci. USA (2011) 108: 21081-21086). Esto produce un efecto de parpadeo, donde una marca pasará por una o más etapas de brillante-oscuro-brillante. La duración y el número de estos períodos oscuros pueden variar. La presente invención utiliza este comportamiento de parpadeo para discriminar una marca de dos o más marcas que pueden parecer similares en generación de imágenes limitada por difracción. Si hay varias marcas presentes, es poco probable que la señal desaparezca por completo durante el parpadeo. Lo más probable es que la intensidad caiga a medida que una de las marcas se oscurezca, pero las otras no. La probabilidad de que todas las marcas parpadeen simultáneamente (y, por lo tanto, se vea como un solo flúor) puede calcularse en función de las características específicas de parpadeo de un colorante.

En algunas realizaciones, las señales ópticas de las marcas se miden durante al menos dos puntos de tiempo, y una señal óptica proviene de una sola marca si la intensidad de la señal óptica se reduce en una función de una sola etapa. En algunas realizaciones, los dos puntos de tiempo pueden estar separados de 0,1 a 30 minutos, de 1 segundo a 20 minutos, de 10 segundos a 10 minutos; 0,01, 0,1, 1, 2, 3, 4, 5, 10, 20, 30, 40, 50, 60 segundos o más; y/o 1, 2, 3, 4, 5, 10, 20, 30, 40, 50, 60 segundos o menos. En realizaciones adicionales, una intensidad de la señal óptica de una sola marca tiene una disminución de una sola etapa con el tiempo, y una intensidad de la señal óptica de dos o más marcas tiene disminuciones de múltiples etapas con el tiempo. En realizaciones adicionales, las señales ópticas de las marcas se miden durante al menos dos puntos de tiempo y se normalizan a los perfiles de blanqueo de las marcas. En otro aspecto, el método descrito en la presente memoria y/o la etapa de recuento pueden comprender adicionalmente medir una señal óptica de una marca de control durante al menos dos puntos de tiempo, y comparar la señal óptica de la marca de control con las señales ópticas de las marcas para determinar un aumento o disminución de la señal óptica de las marcas.

En otro aspecto, la etapa de recuento comprende adicionalmente confirmar el recuento utilizando una molécula de control. Se puede utilizar una molécula de control para determinar el cambio en la frecuencia de un tipo de molécula. A menudo, el objetivo experimental es determinar la abundancia de dos o más tipos de moléculas, ya sea en forma absoluta o en relación entre sí. Considérese el ejemplo de dos moléculas marcadas con dos colorantes diferentes. Si la hipótesis nula es que tienen la misma frecuencia, se pueden enumerar en una matriz de una sola molécula y la proporción de los recuentos se puede comparar con la hipótesis nula. La "matriz de una sola molécula" en la presente memoria se define como una matriz configurada para detectar una sola molécula, que incluye, por ejemplo, las matrices descritas en la Publicación de la Solicitud de Patente de Estados Unidos Núm. 2013/0172216. Si la razón varía de 1:1, esto implica que dos moléculas están en diferentes frecuencias. Sin embargo, puede no estar claro a priori si ha aumentado la abundancia de una o ha disminuido la abundancia de la otra. Si se utiliza un tercer colorante como una molécula de control que también debería tener la misma frecuencia, esto debería tener una razón de 1:1 con los otros dos colorantes. Considérese el ejemplo de dos moléculas marcadas con los colorantes A y B, cuyo objetivo es ver si la molécula marcada con el colorante B tiene un aumento o disminución de frecuencia en comparación con la molécula marcada con el colorante A. Una tercera molécula marcada con el colorante C se incluye en el experimento de manera que debería estar en la misma abundancia que las otras dos moléculas. Si la razón de moléculas marcadas A y B respectivamente es 1:2, la primera molécula tiene una disminución de la frecuencia o la segunda tiene un aumento de la frecuencia. Si la razón de las moléculas marcadas A y C es 1:1 y la razón de las moléculas marcadas B y C es 1:2, es probable que la molécula marcada con el colorante B haya aumentado con la frecuencia con respecto a la molécula marcada con el colorante A. Un ejemplo de esto sería al determinar los cambios en el número de copias de ADN en un genoma diploide. Es importante saber si una secuencia se amplifica o si la otra se elimina y el uso de una molécula de control permite esta determinación. Téngase en cuenta que el control puede ser otra región del genoma o una secuencia de control artificial.

En algunas realizaciones, los resultados del método descrito en la presente memoria (p. ej., números contados de marcas) pueden confirmarse utilizando marcas diferentes pero las mismas etiquetas utilizadas en el método inicial. Tal confirmación se puede realizar simultáneamente con el método inicial o después de realizar el método inicial. En realizaciones adicionales, la confirmación descrita en la presente memoria comprende poner en contacto el primer y el segundo conjunto de sondas de control con la muestra genética, en donde el primer conjunto de sondas de control comprende una primera sonda de marcaje de control y la primera sonda de etiquetado, que es la misma etiqueta del primer conjunto de sondas descrito en la presente memoria, y el segundo conjunto de sondas de control comprende una segunda sonda de marcaje de control y la segunda sonda de etiquetado, que es la misma etiqueta del segundo conjunto de sondas descrito en la presente memoria. La confirmación puede comprender adicionalmente hibridar al menos una parte del primer y segundo conjuntos de sondas de control con la primera y segunda regiones de ácido nucleico de interés en moléculas de nucleótidos de la muestra genética, respectivamente. La confirmación puede comprender adicionalmente ligar el primer conjunto de sondas de control ligando al menos la primera sonda de marcaje de control y la primera sonda de etiquetado. La confirmación puede comprender adicionalmente ligar el segundo conjunto de sondas de control ligando al menos la segunda sonda de marcaje de control y la segunda sonda de etiquetado. La confirmación puede comprender adicionalmente amplificar los conjuntos de sondas ligadas. La confirmación puede comprender adicionalmente inmovilizar cada una de las sondas de etiquetado en una ubicación predeterminada sobre un sustrato, en donde la primera y segunda sondas de marcaje de control y/o las sondas de marcaje amplificadas de las mismas ligadas a las sondas de etiquetado inmovilizadas comprenden la primera y segunda marcas de control, respectivamente, la primera y la segunda marcas de control son diferentes, y las marcas inmovilizadas se pueden resolver ópticamente. La confirmación puede comprender adicionalmente medir las señales ópticas de las marcas de control inmovilizadas sobre el sustrato. La confirmación puede comprender adicionalmente comparar las señales ópticas de la primera y segunda marcas de control inmovilizadas con las señales ópticas de la primera y segunda marcas inmovilizadas para determinar si existe un error basado en las marcas. El "error basado en una marca" utilizado en la presente memoria significa cualquier error causado por la marca que podría no haberse producido si se utilizara una marca diferente en el método. En algunas realizaciones, la primera marca y la segunda marca de control son las mismas, y la segunda marca y la primera marca de control son las mismas.

El blanqueo se puede utilizar como un método para identificar marcas individuales. Un elemento clave de la lectura es que las marcas individuales se pueden "resolver", es decir, son distintas. Esto es trivial a bajas densidades en una superficie cuando la probabilidad de marcas en proximidad es muy baja. Para densidades más altas, suponiendo que las marcas se encuentran en ubicaciones aleatorias (es decir, Poissonianas), las posibilidades de vecinos próximos aumentan hasta el punto en que números significativos de marcas tienen vecinos cuyas emisiones fluorescentes se solapan parcialmente (o totalmente) con su propia emisión. En este punto, las marcas ya no se pueden "resolver" y, en un régimen de transición, existe entre la detección de una sola marca (es decir, la detección digital) y la detección clásica de tipo matriz de marca múltiple (p. ej., detección analógica), donde se mide la señal promedio de muchas moléculas. Dicho de otra manera, un régimen de recuento digital de moléculas individuales se cambia a un régimen analógico de intensidad fluorescente promedio de muchas moléculas.

Una solución para aumentar el intervalo de carga mientras se mantiene la capacidad de resolución individual es aprovechar el blanqueo con fluoróforo. La exposición prolongada a la luz puede hacer que las marcas se blanqueen, es decir, pierdan su propiedad de fluorescencia. Es decir, con el tiempo, una marca puede extinguirse. Esto usualmente ocurre como una función escalonada, pareciendo que la marca se "apaga". La presente invención puede

utilizar este comportamiento de blanqueo para discriminar una marca de dos o más marcas que pueden parecer similares en la generación de imágenes limitada por difracción. Para marcas múltiples, se esperaría que la extinción se produjera a través de una serie de disminuciones paso a paso en la intensidad de la señal. Por ejemplo, las Figuras 4-13 muestran los gráficos de intensidad de marca integrada en función del tiempo (que muestran los eventos de blanqueamiento como cambios en la intensidad) que se obtuvieron para varias marcas de Alexa 488. Las especies de una única marca frente a múltiples marcas se pueden diferenciar fácilmente (p. ej., dependiendo de si la intensidad de la señal óptica se reduce mediante etapas únicas frente a múltiples, como se muestra en los gráficos).

En otro aspecto, el método de la presente memoria puede comprender calibrar y/o confirmar los números contados mediante intercambio de marcas o intercambio de colorante. En algunas realizaciones donde las sondas producto 1 y 2 están marcadas con las marcas 1 y 2, respectivamente, varios modos de error pueden imitar la frecuencia diferencial de las sondas producto. Por ejemplo, si se observa una razón 1:2 entre la marca 1 y la marca 2, esto puede deberse a diferencias genuinas en la frecuencia (la sonda producto 2 es dos veces más común que la sonda producto 1), las diferencias en la eficacia de hibridación (las sondas producto) tienen la misma abundancia, pero la sonda producto 2 hibrida más eficazmente que la sonda producto 1) o las diferencias en las propiedades de las marcas (por ejemplo, si las marcas son colorantes fluorescentes, la marca 1 puede blanquearse más rápido, parpadear más frecuentemente, proporcionar una señal más baja o menor señal con respecto a ruido que la marca 2). Si se repite el mismo experimento con las marcas cambiadas, la razón se debe revertir, si se trata de una observación genuina de diferentes frecuencias de las moléculas, con la marca 1 ahora dos veces más común que la marca 2. Sin embargo, si esto se debe a una eficacia de hibridación diferencial la razón será $\leq 2:1$. Si la razón 1:2 se debió a las propiedades de las marcas, la razón cambiará a 2:1 de la marca 1 con respecto a la marca 2 si en realidad están en la misma frecuencia. Este enfoque puede extenderse a cualquier número de conjuntos de sondas marcadas.

En algunas realizaciones, la primera región de ácido nucleico de interés se ubica en un primer cromosoma, y la segunda región de ácido nucleico de interés se ubica en un segundo cromosoma, diferente del primer cromosoma. La etapa de recuento puede comprender adicionalmente confirmar el recuento, en donde la etapa de confirmación comprende poner en contacto el primer y segundo conjuntos de sondas de control con la muestra genética, en donde el primer conjunto de sondas de control comprende una primera sonda de marcaje de control y una primera sonda de etiquetado de control, y el segundo conjunto de sondas de control comprende una segunda sonda de marcaje de control y la segunda sonda de etiquetado de control. La etapa de confirmación puede comprender adicionalmente hibridar al menos una parte del primer y segundo conjuntos de sondas de control con la primera y segunda regiones de control ubicadas en el primer y segundo cromosomas, respectivamente, en donde la primera y segunda regiones de control son diferentes de la primera y la segunda regiones de ácido nucleico de interés. La etapa de confirmación puede comprender adicionalmente ligar el primer y el segundo conjunto de sondas de control ligando al menos (i) las primeras sondas de marcaje y etiquetado de control, y (ii) las segundas sondas de marcaje y etiquetado de control. La etapa de confirmación puede comprender adicionalmente amplificar los conjuntos de sondas ligadas. La etapa de confirmación puede comprender adicionalmente inmovilizar (i) el primer conjunto de sondas y el segundo conjunto de sondas de control en una primera ubicación predeterminada, y (ii) el segundo conjunto de sondas y el primer conjunto de sondas de control en una segunda ubicación predeterminada. En algunas realizaciones, la primera y segunda sondas de marcaje de control y/o las sondas de marcaje amplificadas de las mismas ligadas a las sondas de etiquetado inmovilizadas comprenden una primera y segunda marcas de control, respectivamente, la primera marca y la segunda marca de control son diferentes, la segunda marca y las primeras marcas de control son diferentes, las marcas inmovilizadas se pueden resolver ópticamente, la primera y segunda sondas de etiquetado de control inmovilizadas y/o las sondas de etiquetado amplificadas de las mismas comprenden la primera y segunda etiquetas de control, respectivamente, y la etapa de inmovilización se realiza inmovilizando las etiquetas en las ubicaciones predeterminadas. La etapa de confirmación puede comprender adicionalmente medir las señales ópticas de las marcas de control inmovilizadas sobre sustrato. La etapa de confirmación puede comprender adicionalmente comparar las señales ópticas de las marcas de control inmovilizadas con las señales ópticas de la primera y segunda marcas inmovilizadas para determinar si existe un error basado en la región de ácido nucleico de interés. En realizaciones adicionales, la primera etiqueta y la segunda etiqueta de control son las mismas, y la segunda etiqueta y la primera etiqueta de control son las mismas.

En otro aspecto, la etapa de recuento del método descrito en la presente memoria puede comprender adicionalmente calibrar y/o confirmar los números contados (i) repitiendo algunas o todas las etapas de los métodos (p. ej., etapas que incluyen el contacto, la unión, la hibridación, la ligación, amplificación y/o inmovilización) descritos en la presente memoria con un conjunto de sondas diferentes configuradas para unirse y/o hibridar con la misma región o regiones de nucleótidos y/o péptidos de interés o una región o regiones diferentes en el mismo cromosoma de interés, y (ii) promediando los números contados de las marcas en los conjuntos de sondas unidas y/o hibridadas con la misma región de nucleótido y/o péptido de interés o con el mismo cromosoma de interés. En algunas realizaciones, la etapa de promediado se puede realizar antes de la etapa de comparación de manera que se comparen los números contados promediados de marcas en un grupo de diferentes conjuntos de sondas que se unen y/o hibridan con la misma región de nucleótidos y/o péptidos de interés, en lugar de los números contados de las marcas en los conjuntos de sondas individuales. En otro aspecto, el método descrito en la presente memoria puede comprender adicionalmente calibrar y/o confirmar la detección de la variación genética (i) repitiendo algunas o todas las etapas de los métodos (p. ej., etapas que incluyen el contacto, la unión, la hibridación, la ligación, la

amplificación), la inmovilización y/o el recuento) descritos en la presente memoria con diferentes conjuntos de sondas configurados para unirse y/o hibridar con regiones de control que no tienen ninguna variación genética conocida, y (ii) promediar los números contados de marcas en los conjuntos de sondas unidas y/o hibridadas con las regiones de control. En algunas realizaciones, los números promediados de las marcas en los conjuntos de sondas que se unen y/o hibridan con las regiones de control se comparan con los números de las marcas en los conjuntos de sondas que se unen y/o hibridan con las regiones de interés descritas en la presente memoria para confirmar la variación genética en la muestra genética. En otro aspecto, las etapas de calibrado y/o confirmación se pueden repetir simultáneamente con las etapas iniciales, o después de realizar las etapas iniciales.

En otro aspecto, las marcas (p. ej., colorantes fluorescentes) de una o más poblaciones se pueden medir y/o identificar en función de sus características espectrales subyacentes. La mayoría de los sistemas de generación de imágenes fluorescentes incluye la opción de recolectar imágenes en múltiples canales espectrales, controlados por la combinación de la fuente de luz y filtros de excitación/emisión/dicroicos espectrales. Esto permite interrogar las mismas especies fluorescentes en una muestra dada con múltiples bandas de color de luz de entrada diferentes, así como capturar las bandas de color de luz de salida deseadas. En el funcionamiento normal, la excitación de un fluoróforo se logra iluminando con una banda espectral estrecha alineada con los máximos de absorción de esa especie (p. ej., con un LED de banda ancha o un arco y un filtro de excitación para dar forma espectral a la salida, o un láser espectralmente homogéneo), y la mayor parte de la emisión del fluoróforo se recolecta con un filtro de emisión adaptado y uno dicroico de paso largo para diferenciar la excitación y la emisión (Figura 14). En operaciones alternativas, la identidad única de un radical fluorescente se puede confirmar a través de la interrogación con varios colores de excitación y bandas de emisión recolectadas diferentes (o además de) el caso para la operación convencional (Figura 15). La luz de estas diversas configuraciones de generación de imágenes, p. ej., varios filtros de emisión, se recolecta y se compara con los valores de calibración para los fluoróforos de interés (Figura 16). En el caso del ejemplo, la medición experimental (puntos) coincide con los datos de calibración/referencia esperados para ese fluoróforo (triángulos) pero no concuerda con una hipótesis alternativa (cuadrados). Dados los datos de prueba y calibración para uno o más canales, se puede calcular una bondad de ajuste o chi-cuadrado para cada espectro de calibración de hipótesis, y el mejor ajuste seleccionado, de forma automatizada y sólida. Pueden ser interesantes varias referencias, incluidos los fluoróforos utilizados en el sistema, así como los contaminantes fluorescentes comunes, p. ej., aquellos con un perfil de emisión plano (Contaminante 1; triángulo), o un perfil ponderado en azul (Contaminante 2; estrellas) (Figura 17).

Las restricciones de diseño para la selección del filtro pueden ser diferentes de los diseños convencionales para los cuales el objetivo es simplemente maximizar la luz recolectada en un solo canal, evitando contribuciones significativas de otros canales. En la invención de los autores de la presente invención, el objetivo es la selectividad espectral en lugar de la colección de luz únicamente. Por ejemplo, considérense dos fluoróforos con bandas de excitación significativamente diferentes, que se muestran en la Figura 18 (obsérvese que solo se muestran las regiones de excitación y no hay espectros de excitación). Un diseño convencional maximizaría la captura de la emisión de Flúor 1 (con filtro Em1, línea continua) y minimizaría la captura del borde de ataque de Flúor 2, y el Flúor 2 sería óptimamente capturado por Em2 (que está ligeramente desplazado hacia el rojo) para evitar una recolección significativa de la luz de Flúor 1). En el diseño de los autores de la presente invención, se desea verificar la presencia de Flúor 2 con el filtro Em1, lo que conduce a la ampliación de la banda que se va a capturar ("Em1+", línea discontinua fina). Esto crea información adicional para verificar la identidad de Flúor 2. Del mismo modo, Em2 puede ampliarse o desplazarse hacia Flúor 1 para capturar más luz de ese flúor (Em2+, línea discontinua fina). Este aumento en la información espectral también debe equilibrarse con la luz total disponible de un fluoróforo dado para mantener la detectabilidad. En otras palabras, la contribución de un fluoróforo dado en un canal dado solo es significativa si la señal correspondiente está por encima del ruido de fondo, y por lo tanto es informativa, a menos que se pretenda un control negativo. De esta manera, la firma espectral de una entidad fluorescente se puede utilizar para una identificación robusta y la captura de más luz puede ser una segunda prioridad si las características de especies únicas pueden cuantificarse de manera más eficaz.

Los sondas producto proporcionadas pueden estar marcadas con más de un tipo de fluoróforo, de manera que la firma espectral es más compleja. Por ejemplo, las sondas producto siempre pueden llevar un flúor universal, p. ej., Alexa647, y un fluoróforo específico del locus, p. ej., Alexa 555 para el locus 1 y Alexa 594 para el locus 2. Dado que los contaminantes rara vez se transportan, se obtiene la firma de dos fluoros. Esto puede aumentar adicionalmente la confianza del rechazo de la contaminación. La implementación involucraría la obtención de imágenes en tres o más canales en este ejemplo, de modo que la presencia o ausencia de cada flúor se pueda determinar, mediante el método de bondad de ajuste mencionado anteriormente, comparando la prueba con la referencia, produciendo llamadas de locus 1, locus 2 o ningún locus producto. Agregar fluoros adicionales ayuda a la identificación de flúor ya que hay más luz disponible para la recolección, pero a expensas del rendimiento de los productos de ensayo formados correctamente y del tiempo total de generación imágenes (pueden requerirse canales adicionales). También se pueden utilizar otros modificadores espectrales para aumentar la información espectral y la singularidad, incluidos los pares de FRET que cambian de color cuando están cerca u otros radicales.

En otro aspecto, como se describe en la presente memoria, el método de la presente descripción se puede utilizar para detectar una variación genética en péptidos o proteínas. En tal caso, los métodos pueden comprender poner en contacto el primer y el segundo conjunto de sondas con la muestra genética, en donde el primer conjunto de sondas comprende una primera sonda de marcaje y una primera sonda de etiquetado, y el segundo conjunto de sondas

comprende una segunda sonda de marcaje y una segunda sonda de etiquetado. Los métodos pueden comprender adicionalmente unir los conjuntos de sondas a regiones peptídicas de interés mediante un enlace físico o químico, en lugar de la etapa de hibridación descrita en la presente memoria en el caso de la detección de la variación genética en las moléculas de ácido nucleico. Específicamente, los métodos pueden comprender adicionalmente unir al menos partes del primer y segundo conjuntos de sondas a la primera y segunda regiones peptídicas de interés en un péptido o proteína de la muestra genética, respectivamente. Por ejemplo, la unión se puede realizar teniendo un aglutinante en al menos una sonda en el conjunto de sondas que se une específicamente a la región peptídica de interés.

En algunas realizaciones, los métodos para detectar una variación genética en péptidos o proteínas pueden comprender adicionalmente conjugar el primer conjunto de sondas mediante un enlace químico al menos conjugando la primera sonda de marcaje y la primera sonda de etiquetado, y conjugando el segundo conjunto de sondas conjugando al menos la segunda sonda de marcaje y la segunda sonda de etiquetado, en lugar de la etapa de ligación descrita en la presente memoria en caso de detectar la variación genética en las moléculas de ácido nucleico. El método puede comprender adicionalmente inmovilizar las sondas de etiquetado en una ubicación predeterminada sobre un sustrato como se describe en la presente memoria. En realizaciones adicionales, las primera y segunda sondas de marcaje conjugadas con las sondas de etiquetado inmovilizadas comprenden una primera y segunda marcas, respectivamente; la primera y segunda marcas son diferentes; las marcas inmovilizadas se pueden resolver ópticamente; la primera y segunda sondas de etiquetado inmovilizadas y/o las sondas de etiquetado amplificadas de las mismas comprenden una primera y segunda etiquetas, respectivamente; y la etapa de inmovilización se realiza inmovilizando las etiquetas en la ubicación predeterminada. Los métodos pueden comprender adicionalmente, como se describe en la presente memoria, contar (i) un primer número de la primera marca inmovilizada sobre el sustrato, y (ii) un segundo número de la segunda marca inmovilizada sobre el sustrato; y comparar el primer y segundo números para determinar la variación genética en la muestra genética.

Un sistema para detectar una variación genética de acuerdo con los métodos descritos en la presente memoria incluye varios elementos. Algunos elementos incluyen la transformación de una muestra biológica bruta en un analito útil. Este analito se detecta a continuación, generando datos que se procesan después en un informe. En la Figura 19 se muestran varios módulos que se pueden incluir en el sistema. En la Figura 20 se muestran más detalles de varios métodos para analizar datos, incluido el procesamiento de imágenes. El análisis se puede realizar en una computadora e implica una red conectada al dispositivo que genera los datos y un servidor de datos para el almacenamiento de datos y el informe. Opcionalmente, se puede incorporar información adicional más allá de los datos del analito en el informe final, p. ej., la edad materna o los riesgos conocidos anteriores. En algunas realizaciones, el sistema de prueba incluye una serie de módulos, algunos de los cuales son opcionales o pueden repetirse dependiendo de los resultados de los módulos anteriores. La prueba puede comprender: (1) recibir una solicitud, p. ej., de un médico clínico o facultativo encargado de la orden, (2) recibir una muestra del paciente, (3) realizar un ensayo que incluya controles de calidad en esa muestra que dé como resultado un producto de ensayo en un sustrato de generación de imágenes apropiado (p. ej., contacto, unión y/o hibridación de sondas a una muestra, ligación de las sondas, opcionalmente amplificación de las sondas ligadas e inmovilización de las sondas sobre un sustrato como se describe en la presente memoria), (4) generación de imágenes del sustrato en uno o más canales espectrales, (5) analizar datos de las imágenes, (6) realizar cálculos estadísticos (p. ej., comparar el primer y segundo números para determinar la variación genética en la muestra genética), (7) crear y aprobar el informe clínico, y (8) devolver el informe al médico clínico o facultativo que realizó la orden. El sistema de prueba puede comprender un módulo configurado para recibir una solicitud, p. ej., de un médico clínico o facultativo encargado de la orden, un módulo configurado para recibir una muestra del paciente, (3) un módulo configurado para realizar un ensayo que incluye controles de calidad en esa muestra, lo que da como resultado un producto de ensayo en un sustrato de generación de imágenes apropiado, (4) un módulo configurado para visualizar el sustrato en uno o más canales espectrales, (5) un módulo configurado para analizar los datos de la imagen, (6) un módulo configurado para realizar cálculos estadísticos, (7) un módulo configurado para crear y confirmar el informe clínico, y/o (8) un módulo configurado para devolver el informe al médico clínico o facultativo que realizó la orden.

En un aspecto, los ensayos y métodos descritos en la presente memoria se pueden realizar en una única muestra de entrada simultáneamente. Por ejemplo, el método puede comprender verificar la presencia de moléculas genómicas fetales en o por encima de un umbral mínimo como se describe en la presente memoria, seguido de una etapa para estimar el estado del número de copias diana si y solo si se cumple ese umbral mínimo. Por lo tanto, se puede ejecutar por separado un ensayo específico del alelo en la muestra de entrada para realizar el cálculo de la fracción fetal, y un ensayo diana genómico para calcular el estado del número de copias. En otras realizaciones, tanto los ensayos como los métodos descritos en la presente memoria se pueden llevar a cabo en paralelo en la misma muestra al mismo tiempo en el mismo volumen de fluido. Otros ensayos de control de calidad también se pueden llevar a cabo en paralelo con las mismas etapas de procesamiento del ensayo universal. Dado que las etiquetas, las etiquetas de afinidad y/o las sondas de etiquetado en las sondas producto, el conjunto de sondas ligadas o la molécula marcada que se van a inmovilizar sobre el sustrato se pueden diseñar de forma única para cada ensayo y cada producto de ensayo, se pueden localizar, generar imágenes y cuantificar todos los productos de ensayo paralelos en diferentes ubicaciones físicas en el sustrato de generación de imágenes. En otro aspecto, el mismo ensayo o método (o algunas de sus etapas) descritos en la presente memoria que utilizan las mismas sondas y/o la detección de la misma variación genética o control pueden realizarse en múltiples muestras simultáneamente en los

mismos o diferentes módulos (p. ej., tubo de ensayo) descritos en la presente memoria. En otro aspecto, los ensayos y métodos (o algunas de sus etapas) descritos en la presente memoria utilizando diferentes sondas y/o detectando diferentes variaciones genéticas o controles se pueden realizar en una o varias muestras simultáneamente en el mismo módulo o en diferentes módulos (p. ej., tubo de ensayo).

5 En otro aspecto, el análisis de imágenes puede incluir el pre-procesamiento de imágenes, la segmentación de imágenes para identificar las marcas, la caracterización de la calidad de la marca, el filtrado de la población de las marcas detectadas basándose en la calidad y la realización de cálculos estadísticos en función de la naturaleza de los datos de las imágenes. En algunos casos, por ejemplo cuando se realiza y se generan imágenes de un ensayo específico de un alelo, se puede calcular la fracción fetal. En otros, tales como el ensayo y la generación de
10 imágenes de dianas genómicas, se calcula el estado del número de copias relativo entre dos regiones genómicas diana. El análisis de los datos de la imagen se puede producir en tiempo real en el mismo ordenador que controla la adquisición de imágenes, o en un ordenador en red, de manera que los resultados del análisis pueden incorporarse al árbol de decisiones del flujo de trabajo de la prueba casi en tiempo real.

15 En otro aspecto, las etapas (4) y (5) de la prueba anterior se pueden repetir varias veces para diferentes porciones del sustrato de generación de imágenes, de manera que los resultados dictan las siguientes etapas. Por ejemplo, las pruebas y los métodos descritos en la presente memoria comprenden confirmar la presencia y el nivel preciso de una muestra fetal en una muestra genética obtenida de un sujeto antes de someter a prueba el estado del número de copias relativo de las dianas genómicas. Como se describe en la presente memoria, se puede utilizar un ensayo sensible a los alelos para cuantificar los niveles de ADN fetal con respecto al ADN materno. Las sondas producto
20 resultantes se pueden bajar a una región de fracción fetal 1 en el sustrato, y generar imágenes. En algunas realizaciones, si y solo si la fracción fetal calculada está por encima del requisito mínimo del sistema, la prueba puede continuar y proporcionar un resultado válido. De esta manera, la prueba de muestras que no confirman al menos la fracción fetal de entrada mínima se puede finalizar antes de que tenga lugar la generación de imágenes y los análisis adicionales. Por el contrario, si la fracción fetal está por encima del umbral mínimo, se pueden generar
25 imágenes adicionales (etapa 4 de la prueba) de las dianas genómicas (p. ej., cromosomas 21, 18 o 13) seguido de un análisis adicional (etapa 5 de la prueba). También se pueden utilizar y probar otros criterios.

En otro aspecto, no todos los SNP analizados en el ensayo específico de alelos pueden dar como resultado información útil. Por ejemplo, el material genómico materno puede tener alelos heterocigotos para un SNP dado (p. ej., par de alelos AB), y el material fetal también puede ser heterocigoto en ese sitio (p. ej., AB), por lo que el material fetal es indistinguible y el cálculo de la fracción fetal falla. Sin embargo, otro sitio SNP para la misma muestra de entrada puede mostrar nuevamente que el material materno es heterocigoto (p. ej., AB) mientras que el material fetal es homocigoto (p. ej., AA). En este ejemplo, el ensayo específico de alelos puede producir un poco más de recuentos de A que de B debido a la presencia del ADN fetal, a partir del cual se puede calcular la fracción fetal. Debido a que el perfil de SNP (es decir, el genotipo) no se puede conocer a priori para una muestra determinada, se
30 deben diseñar múltiples o numerosos sitios de SNP de manera que casi todas las muestras posibles produzcan un sitio SNP informativo. Cada sitio SNP se puede localizar en una ubicación física diferente en el sustrato de generación de imágenes, por ejemplo, utilizando una etiqueta diferente para cada SNP. Sin embargo, para una prueba dada, la fracción fetal solo se puede calcular con éxito una vez. Por lo tanto, se pueden generar imágenes y analizar una o varias ubicaciones en el sustrato utilizado para interrogar los SNP (p. ej., en grupos de uno, dos, tres, cuatro, cinco, diez, veinte, cincuenta o menos y/o uno, dos, tres, cuatro, cinco, diez, veinte, cincuenta o más) hasta
40 que se detecte un SNP informativo. Al alternar la generación de imágenes y el análisis, se pueden pasar por alto la generación de imágenes de todas las posibles manchas de SNP y reducir significativamente la duración promedio de la prueba, manteniendo la precisión y la robustez.

45 En otro aspecto, la determinación de la fracción fetal de una muestra puede ayudar a otros aspectos del sistema más allá de las pruebas finales para las cuales la porción de la fracción fetal en una muestra es inadecuada. Por ejemplo, si la fracción fetal es alta (p. ej., 20%) en ese caso para una potencia estadística dada, el número de recuentos requeridos por diana genética (p. ej., chr21) será menor; si la fracción fetal es baja (p. ej., 1%), en ese caso para la misma potencia estadística, se requiere un número muy alto de recuentos por diana genómica para alcanzar la misma significación estadística. Por lo tanto, después de la generación de imágenes (4-1) de la región de la fracción fetal 1, el análisis (5-1) de los datos que dan como resultado un rendimiento de recuento requerido por
50 diana genómica, la generación de imágenes (4-2) de la región de la diana genómica 2 comienza en el rendimiento requerido, seguido del análisis (5-2) de esos datos de imágenes y el resultado de la prueba para la variación genómica de las dianas de entrada.

55 En otro aspecto, las etapas (4) y (5) de la prueba anterior se pueden repetir adicionalmente con fines de control de calidad, incluida la evaluación de los niveles de fondo de los fluoros en el sustrato de generación de imágenes, los radicales contaminantes, los controles positivos u otras causas de variación del número de copias más allá de la prueba inmediata (p. ej., cáncer en la madre o el feto, quimerismo fetal, gemelación). Debido a que el análisis de imágenes se puede realizar en tiempo real y no requiere completar la ejecución completa de la generación de imágenes antes de generar los resultados (a diferencia de los métodos de secuenciación de ADN), los resultados intermedios pueden dictar las siguientes etapas pasos de un árbol de decisión, y adaptar la prueba para un funcionamiento ideal en una muestra individual. El control de calidad también puede abarcar la verificación de que la muestra es de calidad aceptable y está presente, el sustrato de generación de imágenes está configurado
60

correctamente, que el producto de ensayo está presente y/o a la concentración o densidad correctas, que hay niveles aceptables de contaminación, que el aparato de generación de imágenes es funcional y que el análisis está proporcionando los resultados adecuados, introduciéndose todo en un informe de prueba final para que lo revise el equipo clínico.

5 En otro aspecto, la prueba anterior comprende una o más de las siguientes etapas: (1) recibir una solicitud (de, por ejemplo, un médico clínico o facultativo encargado de la orden), (2) recibir una muestra del paciente, (3) realizar un ensayo (incluyendo una porción específica del alelo, una porción diana genómica y controles de calidad) en esa muestra que da como resultado un sustrato de generación de imágenes que contiene un producto de ensayo, (4-1) generar imágenes de la región específica del alelo del sustrato en uno o más canales espectrales, (5-1) analizar los datos de imágenes específicas del alelo para calcular la fracción fetal, (pendiente de suficiente fracción fetal) (4-2) generar imágenes de la región diana genómica del sustrato en uno o más canales espectrales, (5-2) analizar los datos de imagen de la región la diana genómica para calcular el estado del número de copia de las dianas genómicas, (4-3) generar imágenes de la región de control de calidad del sustrato en uno o más canales espectrales, (5-3) analizar los datos de imágenes de control de calidad para calcular, validar y verificar la prueba, (6) realizar cálculos estadísticos, (7) crear y aprobar el informe clínico, y (8) devolver el informe al médico clínico o médico que realizó la orden.

En la siguiente descripción, se exponen diversas realizaciones ilustrativas en vista de las figuras.

La Figura 21 es una implementación de un ensayo para cuantificar el número de copias genómicas en dos loci genómicos. En esta realización del ensayo, 105 y 106 son moléculas diana. 105 contiene la secuencia correspondiente al primer locus genómico "Locus 1" interrogado para determinar el número de copias (ejemplo, cromosoma 21), y 106 contiene la secuencia correspondiente al segundo locus genómico "Locus 2" interrogado para determinar el número de copias (ejemplo, cromosoma 18). La Figura 21 contiene un ejemplo de un conjunto de sondas por locus genómico, pero en algunas realizaciones de este ensayo, se diseñarán múltiples conjuntos de sondas para interrogar múltiples regiones dentro de un locus genómico. Por ejemplo, se pueden diseñar más de 10, o más de 100, o más de 500 conjuntos de sondas que corresponden al cromosoma 21. La Figura 21 ilustra solo un único conjunto de sondas para cada locus genómico, pero de manera importante que el alcance de esta invención permite múltiples conjuntos de sondas para cada locus genómico. La Figura 21 también ilustra un solo evento de hibridación entre una molécula diana y un conjunto de sondas. En la práctica, habrá múltiples moléculas diana presentes en una muestra de ensayo. Muchas moléculas diana contendrán las secuencias necesarias para la hibridación con un conjunto de sondas y la formación de una sonda producto. Diferentes moléculas diana pueden hibridar con conjuntos de sondas, ya que ciertas moléculas diana portarán polimorfismos genéticos. Además, las moléculas diana que surgen del ADN genómico pueden tener un surtido aleatorio de tamaños de moléculas, así como varias secuencias iniciales y finales. En esencia, existen múltiples moléculas diana que pueden hibridar con un conjunto de sondas dado. En un solo ensayo, se agregan múltiples copias de un conjunto de sondas dado. Por lo tanto, en un solo ensayo se pueden formar hasta miles, o cientos de miles, o millones de sondas producto específicas.

La Figura 21 muestra dos conjuntos de sondas, un conjunto de sondas para el Locus 1 y un conjunto de sondas para el Locus 2, aunque como se mencionó anteriormente, se pueden diseñar múltiples conjuntos de sondas para cada locus genómico. Un primer conjunto de sondas contiene las sondas miembro 101, 102, 103. El elemento 101 contiene la marca (100) de tipo "A". El elemento 103 contiene una etiqueta de afinidad (104) que se puede utilizar para el aislamiento y la identificación de la sonda producto. 102 puede no contener modificaciones, tales como una marca o un código de barras. Un segundo conjunto de sondas con las sondas miembro 108, 109, 110 porta características respectivas como en el primer conjunto de sondas. Sin embargo, 108 contiene una marca (107) de tipo "B", distinguible del tipo "A". El elemento 110 contiene una etiqueta de afinidad (111) que puede ser idéntica o única con respecto a 104. Se pueden diseñar muchos conjuntos de sondas que tienen como diana el "Locus 1", que contienen secuencias de sondas únicas pero el mismo tipo de marca "A". De manera similar, se pueden diseñar muchos conjuntos de sondas que tienen como diana el "Locus 2", que contienen secuencias de sondas únicas pero el mismo tipo de marca "B". En esta realización, las etiquetas de afinidad para los muchos conjuntos de sondas para el Locus 1 pueden ser idénticas o únicas, y las etiquetas de afinidad para los muchos conjuntos de sondas para el Locus 2 pueden ser idénticas o únicas.

Se añaden uno o más conjuntos de sondas a las moléculas diana en un solo recipiente y se exponen a condiciones de hibridación específicas de la secuencia.

Para cada conjunto de sondas, las tres sondas (p. ej., 101, 102, 103) hibridan (o se anclan mediante una interacción sonda-diana similar) a la molécula diana (105), por lo que no hay huecos entre las sondas en la molécula diana. Es decir, las sondas del conjunto de sondas son adyacentes entre sí y son competentes para la ligación.

Se añade la ligasa a las sondas hibridadas y se expone a condiciones de ligasa convencionales. Las sondas ligadas forman una sonda producto. Todas las sondas producto (o la mayoría de ellas) del Locus 1 tienen una marca de tipo "A". Todas las sondas producto del Locus 2 tienen una marca de tipo "B". La cuantificación de las sondas producto correspondientes a los loci genómicos 1 y 2 se realiza utilizando las marcas "A" y "B".

En algunas realizaciones, las sondas producto se inmovilizan sobre un sustrato utilizando sus etiquetas de afinidad. Por ejemplo, si la etiqueta de afinidad es una secuencia de ADN, las sondas producto se pueden hibridar con las regiones de una matriz de captura de ADN a la densidad apropiada para la posterior generación de imágenes.

5 En algunas realizaciones, las etiquetas de afinidad 104 y 111 contienen secuencias únicas y ortogonales que permiten el posicionamiento basado en la superficie en una o más ubicaciones, que pueden compartirse entre productos de hibridación o no. Las Figuras 47 y 48 muestran los patrones de fluorescencia resultantes cuando los productos contienen secuencias de etiquetas de afinidad únicas y el sustrato subyacente contiene complementos para cada una de las etiquetas de afinidad únicas dentro de la misma región (p. ej., como el mismo miembro de una matriz) sobre un sustrato. Las imágenes son de la misma región de un sustrato, pero la Figura 47 muestra las marcas Cy3 (unidas covalentemente al producto del cromosoma 18), y la Figura 48 muestra las marcas Alexa Fluor 647 (unidas covalentemente al producto del cromosoma 21). Se pueden generar patrones similares para otras realizaciones de ensayo que siguen.

15 En otra realización, las etiquetas de afinidad 104 y 111 contienen secuencias idénticas que permiten el posicionamiento basado en la superficie en la misma región (p. ej., como el mismo miembro de una matriz) sobre un sustrato. Es decir, diferentes productos compiten por los mismos sitios de unión. Las Figuras 49 y 51 muestran los patrones de fluorescencia resultantes cuando diferentes productos contienen secuencias de etiquetas de afinidad idénticas y el sustrato subyacente contiene el complemento de la etiqueta de afinidad. Las imágenes son de la misma ubicación en un sustrato, pero la Figura 49 muestra las marcas Cy3 (unidas covalentemente al producto del cromosoma 18) y la Figura 51 muestra las marcas Alexa Fluor 647 (unidas covalentemente al producto del cromosoma 21). Las Figuras 50 y 52 muestran regiones ampliadas de las Figuras 49 y 51, respectivamente, demostrando claramente la resolución de una sola molécula y las marcas distinguibles individualmente. Se pueden generar patrones similares para otras realizaciones de ensayo que siguen.

25 En otra realización, las etiquetas de afinidad 104 y 111 contienen secuencias únicas y ortogonales que permiten el posicionamiento basado en la superficie en más de una ubicación sobre un sustrato. Las Figuras 53 y 54 muestran los patrones de fluorescencia resultantes cuando los productos contienen secuencias de etiquetas de afinidad únicas y el sustrato subyacente tiene una región que contiene el complemento de una etiqueta de afinidad, y otra región separada que contiene el complemento de la otra etiqueta de afinidad. Las imágenes son de dos regiones separadas de un sustrato, conteniendo cada región un único complemento de etiqueta de afinidad como se describió anteriormente. La Figura 53 muestra las marcas Cy3 (unidas covalentemente al producto del cromosoma 21), y la Figura 54 muestra las marcas Alexa Fluor 647 (unidas covalentemente al producto del cromosoma 18). Se pueden generar patrones similares para otras realizaciones de ensayo que siguen.

30 Una característica de esta invención según algunas realizaciones es que la especificidad se logra a través de la combinación de múltiples sondas adyacentes que deben ligarse con éxito para que la sonda producto se forme, capture y detecte satisfactoriamente. Si una sonda producto no se forma satisfactoriamente por algún motivo, no se puede aislar o enriquecer para el uso de una etiqueta de afinidad y la detección. Por ejemplo, si la sonda 101 no se liga satisfactoriamente a la sonda 102, el producto resultante no puede ser detectado. De manera similar, si la sonda 103 no se liga satisfactoriamente a la sonda 102, el producto resultante no se puede aislar o enriquecer utilizando una etiqueta de afinidad.

40 El requerimiento de que todas las sondas del conjunto de sondas hibriden satisfactoriamente con la molécula diana y se ligen satisfactoriamente proporciona una alta especificidad y reduce en gran medida los problemas de hibridación cruzada y, por lo tanto, señales de falsos positivos.

45 En este ensayo, la especificidad se logra a través de la hibridación y la ligación específicas de la secuencia. En una realización preferida, la especificidad para formar sondas producto se produce en el recipiente de reacción, antes de aislar o enriquecer las sondas producto, por ejemplo, inmovilización sobre una superficie u otro sustrato sólido. Esto evita el reto de la hibridación basada en la superficie convencional (p. ej., micromatriz genómica) en la que la especificidad debe lograrse completamente mediante la hibridación solo con secuencias largas de oligonucleótidos (> 40 pb) (p. ej., matrices de Agilent y Affymetrix).

50 El uso de etiquetas de afinidad permite que las sondas producto se inmovilicen sobre un sustrato y, por lo tanto, el exceso de sondas no unidas se puedan lavar con métodos convencionales o se eliminen con métodos convencionales. Por lo tanto, todas o la mayoría de las marcas en la superficie son parte de una sonda producto formada específicamente que está inmovilizada en la superficie.

55 Una característica de esta invención según algunas realizaciones es que la captura sobre la superficie no afecta a la precisión. Es decir, no introduce ningún sesgo. En un ejemplo, si se utiliza la misma etiqueta de afinidad para conjuntos de sondas de diferentes loci genómicos, los conjuntos de sondas que se dirigen a cada locus tienen una marca diferente. Las sondas producto de ambos loci genómicos se pueden inmovilizar en la misma ubicación sobre el sustrato utilizando la misma etiqueta de afinidad. Es decir, las sondas producto del Locus 1 y Locus 2 se capturarán con la misma eficacia, por lo que no se introducirá ningún sesgo específico del locus.

En algunas realizaciones, algunas o todas las sondas y/o las moléculas diana no unidas se eliminan antes de la

captura sobre la superficie utilizando métodos convencionales. Esto disminuye la interferencia entre las sondas y/o las moléculas diana no unidas y las sondas producto durante la captura sobre la superficie.

Una característica de esta invención según algunas realizaciones es que se pueden colocar múltiples tipos de etiquetas de afinidad en la misma región del sustrato (por ejemplo, el mismo punto de matriz o miembro de la matriz). Esto tiene muchas ventajas, incluyendo la colocación de marcadores de control o calibración. Las Figuras 22-46 describen realizaciones ilustrativas adicionales de esta invención. Estas figuras no representan todas las posibles realizaciones, y todas las demás variaciones de este ensayo se incluyen como parte de esta invención. Además, todas las características de la realización descrita en la Figura 21 son aplicables a todas las demás realizaciones adicionales del ensayo descrito en esta solicitud.

La Figura 22 representa una modificación del procedimiento general descrito en la Figura 21. La Figura 22 representa dos conjuntos de sondas, un conjunto de sondas para el Locus 1 y un conjunto de sondas para el Locus 2, aunque como se mencionó anteriormente, se pueden diseñar múltiples conjuntos de sondas para cada locus genómico. 207 y 214 son moléculas diana correspondientes al Locus 1 y Locus 2, respectivamente. Un primer conjunto de sondas contiene las sondas miembro 202, 204, 206. 202 contiene una marca (201) de tipo "A". 206 contiene una etiqueta de afinidad (205) que se puede utilizar para el aislamiento y la identificación de la sonda producto. Un segundo conjunto de sondas con las sondas miembro 209, 211, 231 tiene características respectivas como en el primer conjunto de sondas. Sin embargo, 209 contiene una marca (208) de tipo "B", distinguible del tipo "A". 213 contiene una etiqueta de afinidad (212) que puede ser idéntica o única con respecto a 205. Se pueden diseñar muchos conjuntos de sondas que tienen como diana el "Locus 1", que contienen secuencias de sondas únicas pero el mismo tipo de marca "A". De manera similar, se pueden diseñar muchos conjuntos de sondas que tienen como diana el "Locus 2", que contienen secuencias de sondas únicas pero el mismo tipo de marca "B". En esta realización, las etiquetas de afinidad para los muchos conjuntos de sondas para el Locus 1 pueden ser idénticas o únicas o una mezcla de idénticas y únicas, y las etiquetas de afinidad para los muchos conjuntos de sondas para el Locus 2 pueden ser idénticas o únicas o una mezcla de idénticas y únicas. En esta realización, las sondas 204 y 211 pueden contener una o más marcas (203, 210) de tipo "C". Por lo tanto, las sondas producto contendrán una combinación de marcas. Para el Locus 1, las sondas producto contendrán etiquetas de tipo "A" y de tipo "C", mientras que las sondas producto del Locus 2 contendrán marcas de tipo "B" y de tipo "C".

La Figura 23 representa una modificación del procedimiento general descrito en la Figura 21. La Figura 23 representa dos conjuntos de sondas, un conjunto de sondas para el Locus 1 y un conjunto de sondas para el Locus 2, aunque como se mencionó anteriormente, se pueden diseñar múltiples conjuntos de sondas para cada locus genómico. 307 y 314 son moléculas diana correspondientes al Locus 1 y Locus 2, respectivamente. Un primer conjunto de sondas contiene las sondas miembro 302, 303, 305. 302 contiene una marca (301) de tipo "A". 305 contiene una etiqueta de afinidad (306) que se puede utilizar para el aislamiento y la identificación de la sonda producto. Un segundo conjunto de sondas con sondas miembro 309, 310, 312 tiene características respectivas como en el primer conjunto de sondas. Sin embargo, 309 contiene una marca (308) de tipo "B", distinguible del tipo "A". 312 contiene una etiqueta de afinidad (313) que puede ser idéntica o única con respecto a 306. Se pueden diseñar muchos conjuntos de sondas que tienen como diana el "Locus 1", que contienen secuencias de sondas únicas pero el mismo tipo de marca "A". De manera similar, se pueden diseñar muchos conjuntos de sondas que tienen como diana el "Locus 2", que contienen secuencias de sondas únicas pero el mismo tipo de marca "B". En esta realización, las etiquetas de afinidad para los muchos conjuntos de sondas para el Locus 1 pueden ser idénticas o únicas, y las etiquetas de afinidad para los muchos conjuntos de sondas para el Locus 2 pueden ser idénticas o únicas. En esta realización, las sondas 305 y 312 contienen una o más etiquetas (304, 311) de tipo "C". Por lo tanto, las sondas producto contendrán una combinación de marcas. Para el Locus 1, las sondas producto contendrán marcas de tipo "A" y de tipo "C", mientras que las sondas producto del Locus 2 contendrán marcas de tipo "B" y de tipo "C".

La Figura 24 representa una modificación del procedimiento general descrito en la Figura 21. La Figura 24 representa dos conjuntos de sondas, un conjunto de sondas para el Locus 1 y un conjunto de sondas para el Locus 2, aunque como se mencionó anteriormente, se pueden diseñar múltiples conjuntos de sondas para cada locus genómico. 407 y 414 son moléculas diana correspondientes al Locus 1 y Locus 2, respectivamente.

Un primer conjunto de sondas contiene las sondas miembro 402, 405. 402 contiene una marca (401) de tipo "A". 405 contiene una etiqueta de afinidad (406) que se puede utilizar para el aislamiento y la identificación de la sonda producto.

Un segundo conjunto de sondas con sondas miembro 409, 412 tiene características respectivas como en el primer conjunto de sondas. Sin embargo, 409 contiene una marca (408) de tipo "B", distinguible del tipo "A". 412 contiene una etiqueta de afinidad (413) que puede ser idéntica o única con respecto a 406. Se pueden diseñar muchos conjuntos de sondas que tienen como diana el "Locus 1", que contienen secuencias de sondas únicas pero el mismo tipo de marca "A". De manera similar, se pueden diseñar muchos conjuntos de sondas que tienen como diana el "Locus 2", que contienen secuencias de sondas únicas pero el mismo tipo de marca "B". En esta realización, las etiquetas de afinidad para los muchos conjuntos de sondas para el Locus 1 pueden ser idénticas o únicas, y las etiquetas de afinidad para los muchos conjuntos de sondas para el Locus 2 pueden ser idénticas o únicas.

En esta realización, las sondas 402 y 405 hibridan con las secuencias correspondientes al Locus 1, pero hay una "hueco" en la molécula diana que consiste en uno o más nucleótidos entre las sondas hibridadas 402 y 405. En esta realización, se puede utilizar una ADN polimerasa u otra enzima para sintetizar una nueva especie de polinucleótido (404) que se une covalentemente a 402 y 405. Es decir, la sonda producto formada en este ejemplo es una única molécula de ácido nucleico contiguo con una secuencia correspondiente al Locus 1, y que porta las marcas y/o etiquetas de afinidad anteriores. Además, 404 puede contener una o más marcas de tipo "C", posiblemente como resultado de la incorporación de uno o más nucleótidos que portan una marca de tipo "C". Este ejemplo también se refiere a la sonda producto formada para el Locus 2, que contiene las sondas 409 y 412. Por lo tanto, las sondas producto contendrán una combinación de marcas. Para el Locus 1, las sondas producto contendrán marcas de tipo "A" y de tipo "C", mientras que las sondas producto del Locus 2 contendrán marcas de tipo "B" y de tipo "C".

La Figura 25 representa una modificación del procedimiento general descrito en la Figura 21. La Figura 25 representa dos conjuntos de sondas, un conjunto de sondas para el Locus 1 y un conjunto de sondas para el Locus 2, aunque como se mencionó anteriormente, se pueden diseñar múltiples conjuntos de sondas para cada locus genómico. 505 y 510 son moléculas diana correspondientes al Locus 1 y Locus 2, respectivamente. Un primer conjunto de sondas contiene sondas miembro 502, 503. 502 contiene una marca (501) de tipo "A". 503 contiene una etiqueta de afinidad (504) que se puede utilizar para el aislamiento y la identificación de la sonda producto. Un segundo conjunto de sondas con sondas miembro 507, 508 tiene características respectivas como en el primer conjunto de sondas. Sin embargo, 507 contiene una marca (506) de tipo "B", distinguible del tipo "A". 508 contiene una etiqueta de afinidad (509) que puede ser idéntica o única con respecto a 504. Se pueden diseñar muchos conjuntos de sondas que tienen como diana el "Locus 1", que contienen secuencias de sondas únicas pero el mismo tipo de marca "A". De manera similar, se pueden diseñar muchos conjuntos de sondas que tienen como diana el "Locus 2", que contienen secuencias de sondas únicas pero el mismo tipo de marca "B". En esta realización, las etiquetas de afinidad para los muchos conjuntos de sondas para el Locus 1 pueden ser idénticas o únicas, y las etiquetas de afinidad para los muchos conjuntos de sondas para el Locus 2 pueden ser idénticas o únicas.

La Figura 26 representa una modificación del procedimiento general descrito en la Figura 21. La Figura 26 representa dos conjuntos de sondas, un conjunto de sondas para el Locus 1 y un conjunto de sondas para el Locus 2, aunque como se mencionó anteriormente, se pueden diseñar múltiples conjuntos de sondas para cada locus genómico. 606 y 612 son moléculas diana correspondientes al Locus 1 y Locus 2, respectivamente. Un primer conjunto de sondas contiene sondas miembro 602, 603. 602 contiene una marca (601) de tipo "A". 603 contiene una etiqueta de afinidad (605) que se puede utilizar para el aislamiento y la identificación de la sonda producto. Un segundo conjunto de sondas con sondas miembro 608, 609 tiene características respectivas como en el primer conjunto de sondas. Sin embargo, 608 contiene una marca (607) de tipo "B", distinguible del tipo "A". 609 contiene una etiqueta de afinidad (611) que puede ser idéntica o única con respecto a 605. Se pueden diseñar muchos conjuntos de sondas que tienen como diana el "Locus 1", que contienen secuencias de sondas únicas pero el mismo tipo de marca "A". De manera similar, se pueden diseñar muchos conjuntos de sondas que tienen como diana el "Locus 2", que contienen secuencias de sondas únicas pero el mismo tipo de marca "B". En esta realización, las etiquetas de afinidad para los muchos conjuntos de sondas para el Locus 1 pueden ser idénticas o únicas, y las etiquetas de afinidad para los muchos conjuntos de sondas para el Locus 2 pueden ser idénticas o únicas.

En esta realización, las sondas 603 y 609 contienen una o más marcas (604, 610) de tipo "C". Por lo tanto, las sondas producto contendrán una combinación de marcas. Para el Locus 1, las sondas producto contendrán marcas de tipo "A" y de tipo "C", mientras que las sondas producto del Locus 2 contendrán marcas de tipo "B" y de tipo "C".

La Figura 27 representa una modificación del procedimiento general descrito en la Figura 21. La Figura 27 representa dos conjuntos de sondas para identificar varios alelos del mismo locus genómico. Por ejemplo, para distinguir alelos maternos y fetales, en el caso de ADN sin células aislado de una mujer embarazada, o para distinguir alelos de anfitrión y donante, en el caso de ADN sin células de un receptor de un trasplante de órganos. La Figura 27 representa dos conjuntos de sondas: un conjunto de sondas para el Alelo 1 y un conjunto de sondas para el Alelo 2. 706 y 707 son moléculas diana correspondientes al Alelo 1 y al Alelo 2, respectivamente. Un primer conjunto de sondas contiene las sondas miembro 702, 703, 704. 702 contiene una marca (701) de tipo "A". 704 contiene una etiqueta de afinidad (705) que se puede utilizar para el aislamiento y la identificación de la sonda producto. Un segundo conjunto de sondas con sondas miembro 709, 703, 704 tiene características respectivas como en el primer conjunto de sondas. En esta realización, 703 y 704 son idénticas para ambos conjuntos de sondas. Sin embargo, 709 contiene una marca (708) de tipo "B", distinguible del tipo "A". En esta realización, 702 y 709 contienen secuencias que son casi idénticas y difieren en un solo nucleótido en la secuencia. Por lo tanto, las secuencias de hibridación de estas dos sondas, que están configuradas para hibridar con las regiones para el Alelo 1 y el Alelo 2, contienen regiones complementarias para el Alelo 1 (702) y el Alelo 2 (709). Adicionalmente, la longitud de cada dominio de hibridación en 702 y 709, así como las condiciones de hibridación experimentales, están diseñadas de tal manera que la sonda 702 solo hibridará con el Alelo 1 y la sonda 709 solo hibridará con el Alelo 2. El propósito de este tipo de ensayo es cuantificar con precisión la frecuencia del Alelo 1 y el Alelo 2 en una muestra.

La Figura 28 representa una modificación del procedimiento general descrito en la Figura 21. La Figura 28 representa dos conjuntos de sondas para identificar varios alelos del mismo locus genómico. Por ejemplo, para distinguir alelos maternos y fetales, en el caso de ADN sin células aislado de una mujer embarazada, o para distinguir alelos de anfitrión y donante, en el caso de ADN sin células de un receptor de un trasplante de órganos. La

Figura 28 representa dos conjuntos de sondas: un conjunto de sondas para el Alelo 1 y un conjunto de sondas para el Alelo 2. 807 y 810 son moléculas diana correspondientes al Alelo 1 y Alelo 2, respectivamente. Un primer conjunto de sondas contiene las sondas miembro 802, 804, 805. 802 contiene una marca (801) de tipo "A". 805 contiene una etiqueta de afinidad (806) que se puede utilizar para el aislamiento y la identificación de la sonda producto. Un segundo conjunto de sondas con sondas miembro 809, 804, 805 tiene características respectivas como en el primer conjunto de sondas. En esta realización, 804 y 805 son idénticas para ambos conjuntos de sondas. Sin embargo, 809 contiene una marca (808) de tipo "B", distinguible del tipo "A". En esta realización, 802 y 809 contienen secuencias que son casi idénticas, y difieren en un solo nucleótido en la secuencia. Por lo tanto, las secuencias de hibridación de estas dos sondas contienen regiones complementarias para el Alelo 1 (802) y el Alelo 2 (809). Adicionalmente, la longitud de cada dominio de hibridación en 802 y 809, así como las condiciones de hibridación experimentales, están diseñadas de tal manera que la sonda 802 solo hibridará con el Alelo 1 y la sonda 809 solo hibridará con el Alelo 2. El propósito de este tipo de ensayo es capaz de cuantificar con precisión la frecuencia del Alelo 1 y el Alelo 2 en una muestra. En esta realización, la sonda 804 contiene una o más marcas (803) de tipo "C". Por lo tanto, las sondas producto contendrán una combinación de marcas. Para el Alelo 1, las sondas producto contendrán marcas de tipo "A" y de tipo "C", mientras que las sondas producto del Alelo 2 contendrán marcas de tipo "B" y de tipo "C".

La Figura 29 representa una modificación del procedimiento general descrito en la Figura 21. La Figura 29 representa dos conjuntos de sondas para identificar varios alelos del mismo locus genómico. Por ejemplo, para distinguir alelos maternos y fetales, en el caso de ADN sin células aislado de una mujer embarazada, o para distinguir alelos de anfitrión y donante, en el caso de ADN sin células de un receptor de un trasplante de órganos. La Figura 29 representa dos conjuntos de sondas, un conjunto de sondas para el Alelo 1 y un conjunto de sondas para el Alelo 2.

907 y 910 son moléculas diana correspondientes al Alelo 1 y Alelo 2, respectivamente. Un primer conjunto de sondas contiene las sondas miembro 902, 905. 902 contiene una marca (901) de tipo "A". El elemento 905 contiene una etiqueta de afinidad (906) que se puede utilizar para el aislamiento y la identificación de la sonda producto. Un segundo conjunto de sondas con sondas miembro 909, 905 tiene características respectivas como en el primer conjunto de sondas. En esta realización, 905 es idéntica para ambos conjuntos de sondas. Sin embargo, 909 contiene una marca (908) de tipo "B", distinguible del tipo "A". En esta realización, 902 y 909 contienen secuencias que son casi idénticas y difieren en un solo nucleótido en la secuencia. Por lo tanto, las secuencias de hibridación de estas dos sondas contienen regiones complementarias para el Alelo 1 (902) y el Alelo 2 (909). Adicionalmente, la longitud de cada dominio de hibridación en 902 y 909, así como las condiciones de hibridación experimentales, están diseñadas de tal manera que la sonda 902 solo hibridará con el Alelo 1 y la sonda 909 solo hibridará con el Alelo 2. El propósito de este tipo de ensayo es poder cuantificar con precisión la frecuencia del Alelo 1 y el Alelo 2 en una muestra.

En esta realización, las sondas 902 y 905 hibridan con secuencias correspondientes al Alelo 1, de manera que hay un "hueco" en la molécula diana que consiste en uno o más nucleótidos entre las sondas hibridadas 902 y 905. En esta realización, una ADN polimerasa u otra enzima se pueden utilizar para sintetizar una nueva especie de polinucleótido (904) que se une covalentemente a 902 y 905. Es decir, la sonda producto formada en este ejemplo es una única molécula de ácido nucleico contiguo con una secuencia correspondiente al Alelo 1, que porta las marcas y/o etiquetas de afinidad anteriores. Adicionalmente, 904 puede contener una o más marcas de tipo "C", posiblemente como resultado de la incorporación de un nucleótido que porta una marca de tipo "C". Este ejemplo también se refiere a la sonda producto formada para el Alelo 2, que contiene las sondas 909 y 905.

La Figura 30 representa una modificación del procedimiento general descrito en la Figura 21. La Figura 30 representa dos conjuntos de sondas para identificar varios alelos del mismo locus genómico. Por ejemplo, para distinguir alelos maternos y fetales, en el caso de ADN sin células aislado de una mujer embarazada, o para distinguir alelos de anfitrión y donante, en el caso de ADN sin células de un receptor de un trasplante de órganos. La Figura 30 representa dos conjuntos de sondas, un conjunto de sondas para el Alelo 1 y un conjunto de sondas para el Alelo 2.

1006 y 1007 son moléculas diana correspondientes al Alelo 1 y Alelo 2, respectivamente. Un primer conjunto de sondas contiene las sondas miembro 1001, 1003, 1004. 1003 contiene una marca (1002) de tipo "A". 1004 contiene una etiqueta de afinidad (1005) que se puede utilizar para el aislamiento y la identificación de la sonda producto.

Un segundo conjunto de sondas con sondas miembro 1001, 1009, 1004 tiene características respectivas como en el primer conjunto de sondas. En esta realización, 1001 es idéntica para ambos conjuntos de sondas y 1004 es idéntica para ambos conjuntos de sondas. Sin embargo, 1009 contiene una marca (1008) de tipo "B", distinguible del tipo "A".

En esta realización, 1003 y 1009 contienen secuencias que son casi idénticas y difieren en un solo nucleótido en la secuencia. Por lo tanto, las secuencias de hibridación de estas dos sondas contienen regiones complementarias para el Alelo 1 (1003) y el Alelo 2 (1009), respectivamente. Adicionalmente, la longitud de cada dominio de hibridación en 1003 y 1009, así como las condiciones de hibridación experimentales, están diseñadas de tal manera que la sonda 1003 solo hibridará con el Alelo 1 y la sonda 1009 solo hibridará con el Alelo 2. El propósito de este tipo de ensayo es poder cuantificar con precisión la frecuencia del Alelo 1 y el Alelo 2 en una muestra. En esta

realización, la sonda 1001 contiene una o más marcas (1000) de tipo "C". Por lo tanto, las sondas producto contendrán una combinación de marcas. Para el Alelo 1, las sondas producto contendrán marcas de tipo "A" y de tipo "C", mientras que las sondas producto del Alelo 2 contendrán marcas de tipo "B" y de tipo "C".

La Figura 31 representa una modificación del procedimiento general descrito en la Figura 21. La Figura 31 representa dos conjuntos de sondas para identificar varios alelos del mismo locus genómico. Por ejemplo, para distinguir alelos maternos y fetales, en el caso de ADN sin células aislado de una mujer embarazada, o para distinguir alelos de anfitrión y donante, en el caso de ADN sin células de un receptor de un trasplante de órganos. La Figura 31 representa dos conjuntos de sondas: un conjunto de sondas para el Alelo 1 y un conjunto de sondas para el Alelo 2. 1104 y 1105 son moléculas dianas correspondientes al Alelo 1 y Alelo 2, respectivamente. Un primer conjunto de sondas contiene las sondas miembro 1101, 1102. 1101 contiene una marca (1100) de tipo "A". 1102 contiene una etiqueta de afinidad (1103) que se puede utilizar para el aislamiento y la identificación de la sonda producto. Un segundo conjunto de sondas con sondas miembro 1107, 1102 tiene características respectivas como en el primer conjunto de sondas. En esta realización, 1102 es idéntica para ambos conjuntos de sondas. Sin embargo, 1107 contiene una marca (1106) de tipo "B", distinguible del tipo "A". En esta realización, 1101 y 1107 contienen secuencias que son casi idénticas y difieren en un solo nucleótido en la secuencia. Por lo tanto, las secuencias de hibridación de estas dos sondas contienen regiones complementarias para el Alelo 1 (1101) y el Alelo 2 (1107). Adicionalmente, la longitud de cada dominio de hibridación en 1101 y 1107, así como las condiciones de hibridación experimentales, están diseñadas de tal manera que la sonda 1101 solo hibridará con el Alelo 1 y la sonda 1107 solo hibridará con el Alelo 2. El propósito de este tipo de ensayo es poder cuantificar con precisión la frecuencia del Alelo 1 y el Alelo 2 en una muestra.

La Figura 32 representa una modificación del procedimiento general descrito en la Figura 21. La Figura 32 representa dos conjuntos de sondas para identificar varios alelos del mismo locus genómico. Por ejemplo, para distinguir alelos maternos y fetales, en el caso de ADN sin células aislado de una mujer embarazada, o para distinguir alelos de anfitrión y donante, en el caso de ADN sin células de un receptor de un trasplante de órganos. La Figura 32 representa dos conjuntos de sondas: un conjunto de sondas para el Alelo 1 y un conjunto de sondas para el Alelo 2. 1206 y 1207 son moléculas diana correspondientes al Alelo 1 y Alelo 2, respectivamente. Un primer conjunto de sondas contiene las sondas miembro 1202, 1203. 1202 contiene una marca (1201) de tipo "A". 1203 contiene una etiqueta de afinidad (1205) que se puede utilizar para el aislamiento y la identificación de la sonda producto. Un segundo conjunto de sondas con sondas miembro 1209, 1203 tiene características respectivas como en el primer conjunto de sondas. En esta realización, 1203 es idéntica para ambos conjuntos de sondas. Sin embargo, 1209 contiene una marca (1208) de tipo "B", distinguible del tipo "A". En esta realización, 1202 y 1209 contienen secuencias que son casi idénticas y difieren en un solo nucleótido en la secuencia. Por lo tanto, las secuencias de hibridación de estas dos sondas contienen regiones complementarias para el Alelo 1 (1202) y el Alelo 2 (1209). Adicionalmente, la longitud de cada dominio de hibridación en 1202 y 1209, así como las condiciones de hibridación experimentales, están diseñadas de tal manera que la sonda 1202 solo hibridará con el Alelo 1 y la sonda 1209 solo hibridará con el Alelo 2. El propósito de este tipo de ensayo es poder cuantificar con precisión la frecuencia del Alelo 1 y el Alelo 2 en una muestra. En esta realización, la sonda 1203 contiene una o más marcas (1204) de tipo "C". Por lo tanto, la sonda producto contendrá una combinación de etiquetas. Para el Alelo 1, las sondas producto contendrán marcas de tipo "A" y de tipo "C", mientras que las sondas producto del Alelo 2 contendrán marcas de tipo "B" y de tipo "C".

La Figura 33 representa una modificación del procedimiento general descrito en la Figura 21. La Figura 33 representa dos conjuntos de sondas, un conjunto de sondas para el Locus 1 y un conjunto de sondas para el Locus 2, aunque como se mencionó anteriormente, se pueden diseñar múltiples conjuntos de sondas para cada locus genómico. 1304 y 1305 son moléculas diana correspondientes al Locus 1 y Locus 2, respectivamente. Un primer conjunto de sondas contiene las sondas miembro 1301, 1302. 1301 contiene una marca (1300) de tipo "A". 1301 contiene una etiqueta de afinidad (1303) que se puede utilizar para el aislamiento y la identificación de la sonda producto. Un segundo conjunto de sondas con sondas miembro 1307, 1308 tiene características respectivas como en el primer conjunto de sondas. Sin embargo, 1307 contiene una marca (1306) de tipo "B", distinguible del tipo "A". 1307 contiene una etiqueta de afinidad (1309) que puede ser idéntica o única con respecto a 1303. Se pueden diseñar muchos conjuntos de sondas que tienen como diana el "Locus 1", que contienen secuencias de sondas únicas pero el mismo tipo de marca "A". De manera similar, se pueden diseñar muchos conjuntos de sondas que tienen como diana el "Locus 2", que contienen secuencias de sondas únicas pero el mismo tipo de marca "B". En esta realización, las etiquetas de afinidad para los muchos conjuntos de sondas para el Locus 1 pueden ser idénticas o únicas, y las etiquetas de afinidad para los muchos conjuntos de sondas para el Locus 2 pueden ser idénticas o únicas. En esta realización, las sondas 1301 y 1307 tienen estructuras similares. Por ejemplo, en la sonda 1301 hay dos dominios de hibridación distintos, de manera que la sonda 1302 puede estar ligada a cada extremo de 1301, formando una sonda producto que consiste en una molécula de ADN contigua, topológicamente cerrada (p. ej., una molécula circular). La secuencia no hibridante en la sonda 1301 puede contener características adicionales, posiblemente sitios de enzimas de restricción o sitios de unión de cebadores para la amplificación universal.

Una característica de esta realización es que todas las sondas producto son moléculas circulares contiguas. De esta manera, las sondas producto se pueden aislar de todos los demás ácidos nucleicos mediante la degradación enzimática de todas las moléculas de ácido nucleico lineal, p. ej., utilizando una exonucleasa.

La Figura 34 representa una modificación del procedimiento general descrito en la Figura 21. La Figura 34 representa dos conjuntos de sondas, un conjunto de sondas para el Locus 1 y un conjunto de sondas para el Locus 2, aunque como se mencionó anteriormente, se pueden diseñar múltiples conjuntos de sondas para cada locus genómico. 1405 y 1406 son moléculas diana correspondientes al Locus 1 y Locus 2, respectivamente. Un primer conjunto de sondas contiene las sondas miembro 1401, 1403. 1401 contiene una marca (1400) de tipo "A". 1401 contiene una etiqueta de afinidad (1404) que se puede utilizar para el aislamiento y la identificación de la sonda producto. Un segundo conjunto de sondas con sondas miembro 1408, 1410 tiene características respectivas como en el primer conjunto de sondas. Sin embargo, 1408 contiene una marca (1407) de tipo "B", distinguible del tipo "A". 1408 contiene una etiqueta de afinidad (1411) que puede ser idéntica o única con respecto a 1404. Se pueden diseñar muchos conjuntos de sondas que tienen como diana el "Locus 1", que contienen secuencias de sondas únicas pero el mismo tipo de marca "A". De manera similar, se pueden diseñar muchos conjuntos de sondas que tienen como diana el "Locus 2", que contienen secuencias de sondas únicas pero el mismo tipo de marca "B". En esta realización, las etiquetas de afinidad para los muchos conjuntos de sondas para el Locus 1 pueden ser idénticas o únicas, y las etiquetas de afinidad para los muchos conjuntos de sondas para el Locus 2 pueden ser idénticas o únicas. En esta realización, las sondas 1401 y 1408 tienen estructuras similares. Por ejemplo, en la sonda 1401 hay dos dominios de hibridación distintos, de manera que la sonda 1403 se puede ligar a cada extremo de 1401, formando una sonda producto que consiste en una molécula de ADN contigua, topológicamente cerrada (p. ej., una molécula circular). La secuencia no hibridante en la sonda 1401 puede contener características adicionales, posiblemente sitios de enzimas de restricción, o sitios de unión a cebadores para amplificación universal.

Una característica de esta realización es que todas las sondas producto son moléculas circulares contiguas. De esta manera, las sondas producto se pueden aislar de todos los demás ácidos nucleicos mediante la degradación enzimática de todas las moléculas de ácido nucleico lineal, por ejemplo, utilizando una exonucleasa. En esta realización, las sondas 1403 y 1410 contienen una o más marcas (1402, 1409) de tipo "C". Por lo tanto, las sondas producto contendrán una combinación de marcas. Para el Locus 1, las sondas producto contendrán marcas de tipo "A" y de tipo "C", mientras que las sondas producto del Locus 2 contendrán marcas de tipo "B" y de tipo "C".

La Figura 35 representa una modificación del procedimiento general descrito en la Figura 21. La Figura 35 representa dos conjuntos de sondas, un conjunto de sondas para el Locus 1 y un conjunto de sondas para el Locus 2, aunque como se mencionó anteriormente, se pueden diseñar múltiples conjuntos de sondas para cada locus genómico. 1505 y 1506 son moléculas diana correspondientes al Locus 1 y Locus 2, respectivamente. Un primer conjunto de sondas contiene la sonda miembro 1501. 1501 contiene una marca (1500) de tipo "A". 1501 contiene una etiqueta de afinidad (1504) que se puede utilizar para el aislamiento y la identificación de la sonda producto. Un segundo conjunto de sondas con la sonda miembro 1508 tiene características respectivas como en el primer conjunto de sondas. Sin embargo, 1508 contiene una marca (1507) de tipo "B", distinguible del tipo "A". 1508 contiene una etiqueta de afinidad (1511) que puede ser idéntica o única con respecto a 1504. Se pueden diseñar muchos conjuntos de sondas que tienen como diana el "Locus 1", que contienen secuencias de sondas únicas pero el mismo tipo de marca "A". De manera similar, se pueden diseñar muchos conjuntos de sondas que tienen como diana el "Locus 2", que contienen secuencias de sondas únicas pero el mismo tipo de marca "B". En esta realización, las etiquetas de afinidad para los muchos conjuntos de sondas para el Locus 1 pueden ser idénticas o únicas, y las etiquetas de afinidad para los muchos conjuntos de sondas para el Locus 2 pueden ser idénticas o únicas. En esta realización, las sondas 1501 y 1508 tienen estructuras similares.

Por ejemplo, en la sonda 1501 hay dos dominios de hibridación distintos, de modo que cuando hibrida contra una molécula diana, existe un hueco entre los dos dominios de hibridación. En esta realización, se puede utilizar una ADN polimerasa u otra enzima para sintetizar una nueva especie de polinucleótido (1503) que llena de manera covalente el hueco entre los dominios de hibridación de 1501. Es decir, la sonda producto formada en este ejemplo es una única molécula de ADN topológicamente cerrada, contigua (p. ej., una molécula circular) con una secuencia correspondiente al Locus 1, y que porta las marcas y/o etiquetas de afinidad anteriores. Adicionalmente, 1503 puede contener una o más marcas de tipo "C", posiblemente como resultado de la incorporación de un nucleótido que porta una marca de tipo "C". Este ejemplo también se refiere a la sonda producto formada para el Locus 2, que contiene la sonda 1508. La secuencia no hibridante en la sonda 1501 y la sonda 1508 puede contener características adicionales, posiblemente sitios de enzimas de restricción. Una característica de esta realización es que todas las sondas producto son moléculas circulares contiguas. De esta manera, las sondas producto se pueden aislar de todos los demás ácidos nucleicos mediante la degradación enzimática de todas las moléculas de ácido nucleico lineales, por ejemplo, utilizando una exonucleasa. Las sondas producto contendrán una combinación de marcas. Para el Locus 1, las sondas producto contendrán marcas de tipo "A" y de tipo "C", mientras que las sondas producto del Locus 2 contendrán marcas de tipo "B" y de tipo "C".

La Figura 36 representa una modificación del procedimiento general descrito en la Figura 21. La Figura 36 representa dos conjuntos de sondas, un conjunto de sondas para el Locus 1 y un conjunto de sondas para el Locus 2, aunque como se mencionó anteriormente, se pueden diseñar múltiples conjuntos de sondas para cada locus genómico. 1605 y 1606 son moléculas diana correspondientes al Locus 1 y Locus 2, respectivamente.

Un primer conjunto de sondas contiene la sonda miembro 1602. 1602 contiene una marca (1600) de tipo "A". 1602 contiene una etiqueta de afinidad (1601) que se puede utilizar para el aislamiento y la identificación de la sonda producto.

Un segundo conjunto de sondas con la sonda miembro 1609 tiene características respectivas como en el primer conjunto de sondas. Sin embargo, 1609 contiene una marca (1608) de tipo "B", distinguible del tipo "A". 1609 contiene una etiqueta de afinidad (1607) que puede ser idéntica o única con respecto a 1601. Se pueden diseñar muchos conjuntos de sondas que tienen como diana el "Locus 1", que contienen secuencias de sondas únicas pero el mismo tipo de marca "A". De manera similar, se pueden diseñar muchos conjuntos de sondas que tienen como diana el "Locus 2", que contienen secuencias de sondas únicas pero el mismo tipo de marca "B". En esta realización, las etiquetas de afinidad para los muchos conjuntos de sondas para el Locus 1 pueden ser idénticas o únicas, y las etiquetas de afinidad para los muchos conjuntos de sondas para el Locus 2 pueden ser idénticas o únicas.

En esta realización, las sondas 1602 y 1609 hibridan con secuencias correspondientes al Locus 1 o Locus 2 respectivamente, y se puede utilizar una ADN polimerasa u otra enzima para sintetizar una nueva secuencia de polinucleótidos, por ejemplo 1603 en el caso de Locus 1 o 1611 en el caso de Locus 2. En esta realización, 1603 y 1611 pueden contener una o más marcas (1604) de tipo "C", posiblemente como resultado de la incorporación de uno o más nucleótidos que portan una marca de tipo "C". Este ejemplo también se refiere a la sonda producto formada para el Locus 2. Por lo tanto, las sondas producto contendrán una combinación de marcas. Para el Locus 1, las sondas producto contendrán marcas de tipo "A" y de tipo "C", mientras que las sondas producto del Locus 2 contendrán marcas de tipo "B" y de tipo "C". Esta realización da como resultado sondas producto con alta especificidad para secuencias en el Locus 1 o Locus 2, respectivamente.

La Figura 37 representa una modificación del procedimiento general descrito en la Figura 21. La Figura 37 representa dos conjuntos de sondas, un conjunto de sondas para el Locus 1 y un conjunto de sondas para el Locus 2, aunque como se mencionó anteriormente, se pueden diseñar múltiples conjuntos de sondas para cada locus genómico. 1704 y 1705 son moléculas diana correspondientes al Locus 1 y Locus 2, respectivamente.

Un primer conjunto de sondas contiene la sonda miembro 1702. 1702 contiene una etiqueta de afinidad (1700) que se puede utilizar para el aislamiento y la identificación de la sonda producto.

Un segundo conjunto de sondas con la sonda miembro 1708 tiene características respectivas como en el primer conjunto de sondas. 1708 contiene una etiqueta de afinidad (1706) que puede ser idéntica o única con respecto a 1700. Se pueden diseñar muchos conjuntos de sondas que tienen como diana el "Locus 1", que contienen secuencias de sondas únicas. De manera similar, se pueden diseñar muchos conjuntos de sondas que tienen como diana el "Locus 2", que contienen secuencias de sondas únicas. En esta realización, las etiquetas de afinidad para los muchos conjuntos de sondas para el Locus 1 pueden ser idénticas o únicas, y las etiquetas de afinidad para los muchos conjuntos de sondas para el Locus 2 pueden ser idénticas o únicas.

En esta realización, las sondas 1702 y 1708 hibridan con secuencias correspondientes al Locus 1 y Locus 2, respectivamente. Los diseños de cada sonda para el Locus 1 y el Locus 2 son tales que el primer nucleótido adyacente junto a los dominios de hibridación contiene un nucleótido diferente para el Locus 1 que para el Locus 2. En este ejemplo, el primer nucleótido adyacente junto al dominio de hibridación de 1702 es una "A", mientras que el primer nucleótido adyacente al lado del dominio de hibridación de 1708 es una "T". En esta realización, todas las sondas para el Locus 1 se diseñarán de tal manera que el primer nucleótido inmediatamente adyacente al dominio de hibridación consistirá en uno o varios nucleótidos diferentes del primer nucleótido inmediatamente adyacente al dominio de hibridación de las sondas para el Locus 2. Eso es, por diseño, los conjuntos de sondas de Locus 1 y Locus 2 se pueden distinguir entre sí en función de la identidad del primer nucleótido inmediatamente adyacente al dominio de hibridación.

En esta realización, se utilizará una ADN polimerasa u otra enzima para añadir al menos un nucleótido adicional a cada una de las secuencias de la sonda. En este ejemplo, los sustratos de nucleótidos para la ADN polimerasa son competentes para una sola adición, por ejemplo, los nucleótidos pueden ser terminadores de la cadena dideoxi. Es decir, solo se añadirá un nuevo nucleótido a cada secuencia de la sonda. En este ejemplo, el nucleótido añadido a la sonda 1702 contendrá una o más marcas (1703) de tipo "A". El nucleótido añadido a la sonda 1708 contendrá una o más marcas (1709) de tipo "B", de manera que las sondas producto para el Locus 1 se pueden distinguir de las sondas producto del Locus 2.

La Figura 38 representa una modificación del procedimiento general descrito en la Figura 21. La Figura 38 representa dos conjuntos de sondas, un conjunto de sondas para el Locus 1 y un conjunto de sondas para el Locus 2, aunque como se mencionó anteriormente, se pueden diseñar múltiples conjuntos de sondas para cada locus genómico. 1804 y 1805 son moléculas diana correspondientes al Locus 1 y Locus 2, respectivamente.

Un primer conjunto de sondas contiene la sonda miembro 1802. 1802 contiene una etiqueta de afinidad (1800) que se puede utilizar para el aislamiento y la identificación de la sonda producto.

Un segundo conjunto de sondas con la sonda miembro 1808 tiene características respectivas como en el primer conjunto de sondas. 1808 contiene una etiqueta de afinidad (1806) que puede ser idéntica o única con respecto a 1800. Se pueden diseñar muchos conjuntos de sondas que tienen como diana el "Locus 1", que contienen secuencias de sondas únicas. De manera similar, se pueden diseñar muchos conjuntos de sondas que tienen como

diana el "Locus 2", que contienen secuencias de sondas únicas. En esta realización, las etiquetas de afinidad para los muchos conjuntos de sondas para el Locus 1 pueden ser idénticas o únicas, y las etiquetas de afinidad para los muchos conjuntos de sondas para el Locus 2 pueden ser idénticas o únicas.

5 En esta realización, las sondas 1802 y 1808 hibridan con secuencias correspondientes al Locus 1 y Locus 2, respectivamente. Los diseños de cada sonda para el Locus 1 y el Locus 2 son tales que el primer nucleótido adyacente junto a los dominios de hibridación contiene un nucleótido diferente para el Locus 1 que para el Locus 2. En este ejemplo, el primer nucleótido adyacente junto al dominio de hibridación de 1802 es una "A", mientras que el primer nucleótido adyacente junto al dominio de hibridación de 1808 es una "T". En esta realización, todas las sondas para el Locus 1 se diseñarán de tal manera que el primer nucleótido inmediatamente adyacente al dominio de hibridación consista en uno o varios nucleótidos diferentes del primer nucleótido inmediatamente adyacente al dominio de hibridación de las sondas para el Locus 2. Eso es, por diseño, los conjuntos de sondas del Locus 1 y el Locus 2 se pueden distinguir entre sí en función de la identidad del primer nucleótido inmediatamente adyacente al dominio de hibridación.

15 En esta realización, se utilizará una ADN polimerasa u otra enzima para añadir al menos un nucleótido adicional a cada una de las secuencias de la sonda. En este ejemplo, los sustratos de nucleótidos para la ADN polimerasa son competentes para una sola adición, quizás porque los nucleótidos añadidos a la mezcla de reacción son didesoxi nucleótidos. Es decir, solamente se añadirá un nuevo nucleótido a cada secuencia de la sonda. En este ejemplo, el nucleótido añadido a la sonda 1802 contendrá una o más marcas (1803) de tipo "A". El nucleótido añadido a la sonda 1808 contendrá una o más marcas (1809) de tipo "B", de manera que las sondas producto para el Locus 1 se pueden distinguir de las sondas producto del Locus 2.

20 En esta realización, las sondas 1802 y 1808 contienen una o más marcas (1801, 1806) de tipo "C". Por lo tanto, las sondas producto contendrán una combinación de marcas. Para el Locus 1, las sondas producto contendrán marcas de tipo "A" y de tipo "C", mientras que las sondas producto del Locus 2 contendrán marcas de tipo "B" y de tipo "C".

25 La Figura 39 representa una modificación del procedimiento general descrito en la Figura 21. La Figura 39 representa dos conjuntos de sondas, un conjunto de sondas para el Locus 1 y un conjunto de sondas para el Locus 2, aunque como se mencionó anteriormente, se pueden diseñar múltiples conjuntos de sondas para cada locus genómico. 1906 y 1907 son moléculas diana correspondientes al Locus 1 y Locus 2, respectivamente.

Un primer conjunto de sondas contiene la sonda miembro 1902. 1902 contiene una etiqueta de afinidad (1901) que se puede utilizar para el aislamiento y la identificación de la sonda producto.

30 Un segundo conjunto de sondas con la sonda miembro 1910 tiene características respectivas como en el primer conjunto de sondas. 1910 contiene una etiqueta de afinidad (1908) que puede ser idéntica o única a partir de 1901. Se pueden diseñar muchos conjuntos de sondas que tienen como diana el "Locus 1", que contienen secuencias de sondas únicas. De manera similar, se pueden diseñar muchos conjuntos de sondas que tienen como diana el "Locus 2", que contienen secuencias de sondas únicas. En esta realización, las etiquetas de afinidad para los muchos conjuntos de sondas para el Locus 1 pueden ser idénticas o únicas, y las etiquetas de afinidad para los muchos conjuntos de sondas para el Locus 2 pueden ser idénticas o únicas.

40 En esta realización, las sondas 1902 y 1910 hibridan con secuencias correspondientes al Locus 1 y Locus 2, respectivamente. Los diseños de cada sonda para el Locus 1 y el Locus 2 son tales que el primer nucleótido adyacente junto a los dominios de hibridación contiene un nucleótido diferente para el Locus 1 que el Locus 2. En este ejemplo, el primer nucleótido adyacente junto al dominio de hibridación de 1902 es una "A", mientras que el primer nucleótido adyacente junto al dominio de hibridación de 1910 es una "T". En esta realización, todas las sondas para el Locus 1 se diseñarán de tal manera que el primer nucleótido inmediatamente adyacente al dominio de hibridación consista en uno o varios nucleótidos diferentes del primer nucleótido inmediatamente adyacente al dominio de hibridación de las sondas para el Locus 2. Esto es, por diseño, los conjuntos de sondas de Locus 1 y Locus 2 se pueden distinguir de uno a otro nucleótido en la identidad del primer nucleótido inmediatamente adyacente al dominio de hibridación. Un nucleótido diferente, que no se utiliza para distinguir las sondas del Locus 1 o Locus 2, servirá como un terminador de cadena. En este ejemplo particular, se utiliza un nucleótido "A" en una molécula diana para distinguir las sondas del Locus 1 y se utiliza un nucleótido "T" para distinguir las sondas del Locus 2. En este ejemplo, un nucleótido "C" puede servir como terminador de cadena. En este caso, se añadirá un nucleótido "C" al ensayo que no es capaz de alargar la cadena (p. ej., un didesoxi C). Una restricción adicional es que las secuencias de la sonda están diseñadas de tal manera que no hay casos de un nucleótido de identificación para el Locus 2 presente en 1906 entre el nucleótido distintivo para el Locus 1 y el nucleótido de terminación de la cadena. En este ejemplo, no habrá nucleótidos "T" presentes en 1906 después del dominio de hibridación de 1902 y antes de la G, que se emparejará con el terminador de cadena C.

55 En esta realización, se utilizará ADN polimerasa o una enzima similar para sintetizar nuevas secuencias de nucleótidos, y el nucleótido añadido en la ubicación del nucleótido distintivo para el Locus 1 contendrá una o más marcas (1903) de tipo "A". El nucleótido añadido en la ubicación del nucleótido distintivo para el Locus 2 contendrá 1 o más marcas (1911) de tipo "B", de tal manera que las sondas producto para el Locus 1 se pueden distinguir de las sondas producto del Locus 2. En esta realización, el nucleótido añadido en la posición de terminación de la cadena

contendrá una o más marcas (1912) de tipo "C". Por lo tanto, las sondas producto contendrán una combinación de marcas. Para el Locus 1, las sondas producto contendrán marcas de tipo "A" y de tipo "C", mientras que las sondas producto del Locus 2 contendrán marcas de tipo "B" y de tipo "C".

5 En otra realización, el terminador de cadena no puede contener ninguna marca. En esta realización, se puede añadir un cuarto nucleótido al ensayo que contiene una o más marcas de tipo "C". Este cuarto nucleótido no se empareja con el nucleótido de identificación para el Alelo 1 (en este ejemplo, A), no se empareja con el nucleótido de identificación para el Alelo 2 (en este ejemplo, T), no se empareja con el nucleótido de terminación de la cadena (en este ejemplo G). En este ejemplo, el cuarto nucleótido que llevaría una o más marcas de tipo "C" es G, y se emparejará con las ubicaciones C en 1906 y 1907. Por lo tanto, las sondas producto contendrán una combinación de
10 marcas. Para el Locus 1, las sondas producto contendrán marcas de tipo "A" y de tipo "C", mientras que las sondas producto del Locus 2 contendrán marcas de tipo "B" y de tipo "C".

La Figura 40 representa una modificación del procedimiento general descrito en la Figura 21. La Figura 40 representa dos conjuntos de sondas, un conjunto de sondas para el Locus 1 y un conjunto de sondas para el Locus 2, aunque como se mencionó anteriormente, se pueden diseñar múltiples conjuntos de sondas para cada locus genómico. 2005 y 2006 son moléculas diana correspondientes al Locus 1 y Locus 2, respectivamente.
15

Un primer conjunto de sondas contiene la sonda miembro 2001. 2001 contiene una etiqueta de afinidad (2000) que se puede utilizar para el aislamiento y la identificación de la sonda producto.

Un segundo conjunto de sondas con la sonda miembro 2008 incluye características respectivas como en el primer conjunto de sondas. 2008 contiene una etiqueta de afinidad (2007) que puede ser idéntica o única con respecto a 2000. Se pueden diseñar muchos conjuntos de sondas que tienen como diana el "Locus 1", que contienen secuencias de sondas únicas. De manera similar, se pueden diseñar muchos conjuntos de sondas que tienen como
20 diana el "Locus 2", que contienen secuencias de sondas únicas. En esta realización, las etiquetas de afinidad para los muchos conjuntos de sondas para el Locus 1 pueden ser idénticas o únicas, y las etiquetas de afinidad para los muchos conjuntos de sondas para el Locus 2 pueden ser idénticas o únicas.

25 En esta realización, las sondas 2001 y 2008 hibridan con secuencias correspondientes al Locus 1 y Locus 2 respectivamente. Los diseños de cada sonda para el Locus 1 y el Locus 2 son tales que hay uno o más casos de un nucleótido distintivo (en este ejemplo, "A" es un nucleótido distintivo para el Locus 1 y "T" es un nucleótido distintivo para el Locus 2) seguido de un nucleótido de terminación de cadena (en este ejemplo "G") adyacente al dominio de hibridación de las sondas. Es importante destacar que no habrá casos del nucleótido distintivo para el Locus 2 (en este ejemplo, "T") presente entre el dominio de hibridación de 2001 en 2005 y el nucleótido de terminación de
30 cadena en 2005. Del mismo modo, no habrá ningún caso de nucleótido distintivo para el Locus 1 (en este ejemplo, "A") presente entre el dominio de hibridación de 2008 en 2006 y el nucleótido de terminación de cadena en 2006.

En esta realización, se utilizará ADN polimerasa o una enzima similar para sintetizar nuevas secuencias de nucleótidos (2004, 2011) hasta la adición de un nucleótido de terminación de cadena, un ejemplo posible sería una
35 didesoxi C. En esta realización, los nucleótidos añadidos en las ubicaciones de nucleótidos distintivos para el Locus 1 contendrán una o más marcas (2003) de tipo "A". Los nucleótidos añadidos en las ubicaciones de nucleótidos distintivos para el Locus 2 contendrán 1 o más marcas (2010) de tipo "B", de manera que las sondas producto para el Locus 1 se pueden distinguir claramente de las sondas producto del Locus 2.

La Figura 41 representa una modificación del procedimiento general descrito en la Figura 21. La Figura 41 representa dos conjuntos de sondas, un conjunto de sondas para el Locus 1 y un conjunto de sondas para el Locus 2, aunque como se mencionó anteriormente, se pueden diseñar múltiples conjuntos de sondas para cada locus genómico. 2105 y 2106 son moléculas diana correspondientes al Locus 1 y al Locus 2, respectivamente.
40

Un primer conjunto de sondas contiene la sonda miembro 2102. 2102 contiene una etiqueta de afinidad (2100) que se puede utilizar para el aislamiento y la identificación de la sonda producto.

45 Un segundo conjunto de sondas con la sonda miembro 2109 tiene características respectivas como en el primer conjunto de sondas. 2109 contiene una etiqueta de afinidad (2107) que puede ser idéntica o única con respecto a 2100. Se pueden diseñar muchos conjuntos de sondas que tienen como diana el "Locus 1", que contienen secuencias de sondas únicas. De manera similar, se pueden diseñar muchos conjuntos de sondas que tienen como diana el "Locus 2", que contienen secuencias de sondas únicas. En esta realización, las etiquetas de afinidad para
50 los muchos conjuntos de sondas para el Locus 1 pueden ser idénticas o únicas, y las etiquetas de afinidad para los muchos conjuntos de sondas para el Locus 2 pueden ser idénticas o únicas.

En esta realización, las sondas 2102 y 2109 hibridan con secuencias correspondientes al Locus 1 y Locus 2, respectivamente. Los diseños de cada sonda para el Locus 1 y el Locus 2 son tales que hay uno o más casos de un nucleótido distintivo (en este ejemplo, "A" es un nucleótido distintivo para el Locus 1 y "T" es un nucleótido distintivo para el Locus 2) seguido de un nucleótido de terminación de cadena (en este ejemplo "G") adyacente al dominio de
55 hibridación de las sondas. Es importante destacar que no habrá casos del nucleótido distintivo para el Locus 2 (en este ejemplo, "T") presente entre el dominio de hibridación 2102 en 2105 y el nucleótido de terminación de la cadena en 2105. Del mismo modo, no habrá ningún caso de nucleótido distintivo para el Locus 1 (en este ejemplo, "A")

presente entre el dominio de hibridación 2109 en 2106 y el nucleótido de terminación de la cadena en 2106.

En esta realización, se utilizará ADN polimerasa o una enzima similar para sintetizar nuevas secuencias de nucleótidos (2104, 2110) hasta la adición de un nucleótido de terminación de cadena, un posible ejemplo sería una didesoxi C. En esta realización, los nucleótidos añadidos en las ubicaciones del nucleótido distintivo para el Locus 1 contendrán una o más marcas (2103) de tipo "A". Los nucleótidos añadidos en las ubicaciones del nucleótido distintivo para el Locus 2 contendrán 1 o más marcas (2110) de tipo "B", de modo que las sondas producto para el Locus 1 se puedan distinguir claramente de las sondas producto del Locus 2.

En esta realización, las sondas 2102 y 2109 contienen una o más marcas (2101, 2108) de tipo "C". Por lo tanto, las sondas producto contendrán una combinación de marcas. Para el Locus 1, las sondas producto contendrán marcas de tipo "A" y de tipo "C", mientras que las sondas producto del Locus 2 contendrán marcas de tipo "B" y de tipo "C".

La Figura 42 representa una modificación del procedimiento general descrito en la Figura 21. La Figura 42 representa dos conjuntos de sondas para identificar varios alelos del mismo locus genómico. Por ejemplo, para distinguir alelos maternos y fetales, en el caso de ADN sin células aislado de una mujer embarazada, o para distinguir alelos de anfitrión y donante, en el caso de ADN sin células de un receptor de un trasplante de órganos. La Figura 42 representa dos conjuntos de sondas - un conjunto de sondas para el Alelo 1 y un conjunto de sondas para el Alelo 2. 2203 y 2204 son moléculas diana correspondientes al Alelo 1 y Alelo 2, respectivamente.

Un primer conjunto de sondas contiene la sonda miembro 2201. 2201 contiene una etiqueta de afinidad (2200) que se puede utilizar para el aislamiento y la identificación de la sonda producto. En esta realización, los conjuntos de sondas utilizados para la identificación de los dos alelos diferentes son los mismos. Es decir, el conjunto de sondas para el Alelo 2 consiste en la sonda miembro 2201. En esta realización, la sonda 2201 hibrida con una secuencia correspondiente al Alelo 1 y al Alelo 2, respectivamente, en la Figura 42. El diseño de la sonda 2201 es tal que el primer nucleótido adyacente junto al dominio de hibridación contiene un nucleótido diferente para el Alelo 1 que para el Alelo 2. En otras palabras, el primer nucleótido adyacente al dominio de hibridación puede ser un polimorfismo de un solo nucleótido o SNP. En este ejemplo, el primer nucleótido adyacente en 2203 junto al dominio de hibridación de 2201 es una "A", mientras que el primer nucleótido adyacente en 2204 junto al dominio de hibridación de 2201 es una "T". Es decir, las sondas producto del Alelo 1 y el Alelo 2 se pueden distinguir entre sí en función de la identidad del primer nucleótido inmediatamente adyacente al dominio de hibridación.

En esta realización, se utilizará una ADN polimerasa u otra enzima para añadir al menos un nucleótido adicional a cada una de las secuencias de la sonda. En este ejemplo, los sustratos de nucleótidos para la ADN polimerasa son competentes para una sola adición, quizás porque los nucleótidos añadidos a la mezcla de reacción son didesoxi nucleótidos. Es decir, solo se añadirá un nuevo nucleótido a cada secuencia de la sonda. En este ejemplo, el nucleótido añadido a la sonda 2201 para el Alelo 1 contendrá una o más marcas (2202) de tipo "A". El nucleótido añadido a la sonda 2201 para el Alelo 2 contendrá una o más marcas (2205) de tipo "B", de manera que las sondas producto para el Alelo 1 se pueden distinguir claramente de las sondas producto del Alelo 2. Es decir, la sonda producto para el Alelo 1 consiste en la sonda 2201 más un nucleótido adicional que porta una o más marcas de tipo "A", y las sondas producto para el Alelo 2 consisten en la sonda 2201 más un nucleótido adicional que porta una o más marcas de tipo "B".

La Figura 43 representa una modificación del procedimiento general descrito en la Figura 21. La Figura 43 representa dos conjuntos de sondas para identificar varios alelos del mismo locus genómico. Por ejemplo, para distinguir alelos maternos y fetales, en el caso de ADN sin células aislado de una mujer embarazada, o para distinguir alelos de anfitrión y donante, en el caso de ADN sin células de un receptor de un trasplante de órganos. La Figura 43 representa dos conjuntos de sondas - un conjunto de sondas para el Alelo 1 y un conjunto de sondas para el Alelo 2. 2304 y 2305 son moléculas dianas correspondientes al Alelo 1 y Alelo 2, respectivamente.

Un primer conjunto de sondas contiene la sonda miembro 2302. 2302 contiene una etiqueta de afinidad (2300) que se puede utilizar para el aislamiento y la identificación de la sonda producto. En esta realización, los conjuntos de sondas utilizados para la identificación de los dos alelos diferentes son los mismos. Es decir, el conjunto de sondas para el Alelo 2 consiste en la sonda miembro 2302. En esta realización, la sonda 2302 hibrida con una secuencia correspondiente al Alelo 1 y al Alelo 2, respectivamente, en la Figura 43. El diseño de la sonda 2302 es tal que el primer nucleótido adyacente junto a los dominios de hibridación contienen un nucleótido diferente para el Alelo 1 que el Alelo 2. En otras palabras, el primer nucleótido adyacente al dominio de hibridación puede ser un polimorfismo de un solo nucleótido, o SNP. En este ejemplo, el primer nucleótido adyacente en 2304 junto al dominio de hibridación de 2302 es una "A", mientras que el primer nucleótido adyacente en 2305 junto al dominio de hibridación de 2302 es una "T". Es decir, las sondas producto del Alelo 1 y el Alelo 2 se pueden distinguir entre sí en función de la identidad del primer nucleótido inmediatamente adyacente al dominio de hibridación.

En esta realización, se utilizará una ADN polimerasa u otra enzima para añadir al menos un nucleótido adicional a cada una de las secuencias de la sonda. En este ejemplo, los sustratos de nucleótidos para la ADN polimerasa son competentes para una sola adición, quizás porque los nucleótidos añadidos a la mezcla de reacción son didesoxi nucleótidos. Es decir, solo se añadirá un nuevo nucleótido a cada secuencia de la sonda. En este ejemplo, el nucleótido añadido a la sonda 2302 para el Alelo 1 contendrá una o más marcas (2303) de tipo "A". El nucleótido

añadido a la sonda 2302 para el Alelo 2 contendrá una o más marcas (2306) de tipo "B", de manera que las sondas producto para el Alelo 1 se puedan distinguir claramente de las sondas producto del Alelo 2. Es decir, la sonda producto para el Alelo 1 consiste en la sonda 2302 más un nucleótido adicional que porta una o más marcas de tipo "A", y las sondas producto para el Alelo 2 consiste en la sonda 2302 más un nucleótido adicional que porta una o más marcas de tipo "B".

En esta realización, las sondas 2302 contienen una o más marcas (2301) de tipo "C". Por lo tanto, las sondas producto contendrán una combinación de marcas. Para el Alelo 1, las sondas producto contendrán marcas de tipo "A" y de tipo "C", mientras que las sondas producto del Alelo 2 contendrán marcas de tipo "B" y de tipo "C".

La Figura 44 representa una modificación del procedimiento general descrito en la Figura 21. La Figura 44 representa dos conjuntos de sondas para identificar varios alelos del mismo locus genómico. Por ejemplo, para distinguir alelos maternos y fetales, en el caso de ADN sin células aislado de una mujer embarazada, o para distinguir alelos de anfitrión y donante, en el caso de ADN sin células de un receptor de un trasplante de órganos. La Figura 44 representa dos conjuntos de sondas - un conjunto de sondas para el Alelo 1 y un conjunto de sondas para el Alelo 2. 2405 y 2406 son moléculas diana correspondientes al Alelo 1 y Alelo 2, respectivamente.

Un primer conjunto de sondas contiene la sonda miembro 2401. 2401 contiene una etiqueta de afinidad (2400) que se puede utilizar para el aislamiento y la identificación de la sonda producto. En esta realización, los conjuntos de sondas utilizados para la identificación de dos alelos diferentes son los mismos. Es decir, el conjunto de sondas para el Alelo 2 consiste en la sonda miembro 2401. En esta realización, la sonda 2401 hibrida con una secuencia correspondiente al Alelo 1 y Alelo 2, respectivamente, en la Figura 44. El diseño de la sonda para 2401 es tal que el primer nucleótido adyacente junto a los dominios de hibridación contiene un nucleótido diferente para el Alelo 1 que para el Alelo 2. En otras palabras, el primer nucleótido adyacente al dominio de hibridación puede ser un polimorfismo de un solo nucleótido, o SNP. En este ejemplo, el primer nucleótido adyacente en 2405 junto al dominio de hibridación de 2401 es una "A", mientras que el primer nucleótido adyacente en 2406 junto al dominio de hibridación de 2401 es una "T". Es decir, las sondas producto del Alelo 1 y el Alelo 2 se pueden distinguir entre sí en función de la identidad del primer nucleótido inmediatamente adyacente al dominio de hibridación.

En esta realización, se utilizará una ADN polimerasa u otra enzima para añadir al menos un nucleótido adicional a cada una de las secuencias de la sonda. En este ejemplo, el nucleótido añadido a la sonda 2401 para el Alelo 1 contendrá una o más marcas (2402) de tipo "A". El nucleótido añadido a la sonda 2401 para el Alelo 2 contendrá una o más marcas (2407) de tipo "B", de manera que las sondas producto para el Locus 1 se pueden distinguir claramente de las sondas producto del Locus 2. Es decir, la sonda producto para el Alelo 1 contiene la sonda 2401 más un nucleótido adicional que porta una o más marcas de tipo "A", y la sonda producto para el Alelo 2 contiene la sonda 2401 más un nucleótido adicional que tiene una o más marcas de tipo "B". Un nucleótido diferente, que no se use para distinguir el Alelo 1 del Alelo 2, servirá como terminador de la cadena. En este ejemplo particular, se utiliza un nucleótido "A" en una molécula diana para identificar el Alelo 1 y se utiliza un nucleótido "T" para identificar el Alelo 2. En este ejemplo, un nucleótido "C" puede servir como un terminador de cadena. En este caso, se añadirá un nucleótido "C" al ensayo que no es capaz de alargar la cadena (p. ej., una dideoxi C). Una restricción adicional es que las secuencias de la sonda están diseñadas de tal manera que no hay casos de un nucleótido identificativo para el Alelo 2 que está presente en 2405 entre el nucleótido distintivo para el Alelo 1 y el nucleótido d terminación de la cadena. En este ejemplo, no habrá nucleótidos "T" presentes en 2405 después del dominio de hibridación de 2401 y antes de una G, que se empareje con el terminador de cadena C.

En esta realización, se utilizará ADN polimerasa o una enzima similar para sintetizar nuevas secuencias de nucleótidos, y el nucleótido añadido en la ubicación del nucleótido distintivo para el Alelo 1 contendrá una o más marcas (2402) de tipo "A". El nucleótido añadido en la ubicación del nucleótido distintivo para el Alelo 2 contendrá 1 o más marcas (2407) de tipo "B", de tal manera que las sondas producto para el Alelo 1 se pueden distinguir claramente de las sondas producto del Alelo 2. En esta realización, el nucleótido añadido en la posición de terminación de la cadena contendrá una o más marcas (2403) de tipo "C". Por lo tanto, las sondas producto contendrán una combinación de marcas. Para el Alelo 1, las sondas producto contendrán marcas de tipo "A" y de tipo "C", mientras que las sondas producto del Alelo 2 contendrán marcas de tipo "B" y de tipo "C".

La Figura 45 representa una modificación del procedimiento general descrito en la Figura 21. La Figura 45 representa dos conjuntos de sondas para identificar varios alelos del mismo locus genómico. Por ejemplo, para distinguir alelos maternos y fetales, en el caso de ADN sin células aislado de una mujer embarazada, o para distinguir alelos de anfitrión y donante, en el caso de ADN sin células de un receptor de un trasplante de órganos. La Figura 45 representa dos conjuntos de sondas - un conjunto de sondas para el Alelo 1 y un conjunto de sondas para el Alelo 2. 2505 y 2506 son moléculas dianas correspondientes al Alelo 1 y Alelo 2, respectivamente.

Un primer conjunto de sondas contiene la sonda miembro 2501. 2501 contiene una etiqueta de afinidad (2500) que se puede utilizar para el aislamiento y la identificación de la sonda producto. En esta realización, los conjuntos de sondas utilizados para la identificación de dos alelos diferentes son los mismos. Es decir, el conjunto de sondas para el Alelo 2 consiste en la sonda miembro 2501. En esta realización, la sonda 2501 hibrida con una secuencia correspondiente al Alelo 1 y al Alelo 2, respectivamente, en la Figura 45. El diseño de la sonda para 2501 es tal que el primer nucleótido adyacente junto a los dominios de hibridación contiene un nucleótido diferente para el Alelo 1

que para el Alelo 2. En otras palabras, el primer nucleótido adyacente al dominio de hibridación puede ser un polimorfismo de un solo nucleótido, o SNP. En este ejemplo, el primer nucleótido adyacente en 2505 junto al dominio de hibridación de 2501 es una "A", mientras que el primer nucleótido adyacente en 2506 junto al dominio de hibridación de 2501 es una "T". Es decir, las sondas producto del Alelo 1 y el Alelo 2 se pueden distinguir entre sí en función de la identidad de la primera base inmediatamente adyacente al dominio de hibridación.

En esta realización, se utilizará una ADN polimerasa u otra enzima para añadir al menos un nucleótido adicional a cada una de las secuencias de la sonda. En este ejemplo, el nucleótido añadido a la sonda 2501 para el Alelo 1 contendrá una o más marcas (2502) de tipo "A". El nucleótido añadido a la sonda 2501 para el Alelo 2 contendrá una o más marcas (2507) de tipo "B", de manera que las sondas producto para el Locus 1 se pueden distinguir claramente de las sondas producto del Locus 2. Es decir, la sonda producto para el Alelo 1 contiene la sonda 2501 más un nucleótido adicional que porta una o más marcas de tipo "A", y la sonda producto para el Alelo 2 contiene la sonda 2501 más un nucleótido adicional que tiene una o más marcas de tipo "B". Un nucleótido diferente, que no se use para distinguir el Alelo 1 del Alelo 2, servirá como terminador de la cadena. En este ejemplo particular, se utiliza un nucleótido "A" en una molécula diana para identificar el Alelo 1 y se utiliza un nucleótido "T" para identificar el Alelo 2. En este ejemplo, un nucleótido "C" puede servir como un terminador de cadena. En este caso, se añadirá un nucleótido "C" al ensayo que no sea capaz de alargar la cadena (p. ej., una didesoxi C). Una restricción adicional es que las secuencias de la sonda se diseñan de tal manera que no haya casos de un nucleótido identificativo para el Alelo 2 en 2505 entre el nucleótido distintivo para el Alelo 1 y el nucleótido de terminación de la cadena. En este ejemplo, no habrá nucleótidos "T" presentes en 2505 después del dominio de hibridación de 2501 y antes de una G, que se emparejará con el terminador de cadena C.

En esta realización, se utilizará ADN polimerasa o una enzima similar para sintetizar nuevas secuencias de nucleótidos, y el nucleótido añadido en la ubicación del nucleótido distintivo para el Alelo 1 contendrá una o más marcas (2502) de tipo "A". El nucleótido añadido en la ubicación del nucleótido distintivo para el Alelo 2 contendrá 1 o más marcas (2507) de tipo "B", de manera que las sondas producto para el Alelo 1 se pueden distinguir claramente de las sondas producto del Alelo 2. En esta realización, se puede añadir un cuarto nucleótido al ensayo que contiene una o más marcas (2508, 2503) de tipo "C". Este cuarto nucleótido no se empareja con el nucleótido identificativo para el Alelo 1 (en este ejemplo, A), no se empareja con el nucleótido identificativo para el Alelo 2 (en este ejemplo, T), no se empareja con el nucleótido de terminación de la cadena (en este ejemplo G). En este ejemplo, el cuarto nucleótido que llevaría una o más marcas de tipo "C" es G, y se emparejaría con las ubicaciones C en 2505 y 2506. Por lo tanto, las sondas producto contendrán una combinación de marcas. Para el Alelo 1, las sondas producto contendrán marcas de tipo "A" y de tipo "C", mientras que las sondas producto del Alelo 2 contendrán marcas de tipo "B" y de tipo "C".

La Figura 46 representa una modificación del procedimiento general descrito en la Figura 21. La Figura 46 representa dos conjuntos de sondas para identificar varios alelos del mismo locus genómico. Por ejemplo, para distinguir alelos maternos y fetales, en el caso de ADN sin células aislado de una mujer embarazada, o para distinguir alelos de anfitrión y donante, en el caso de ADN sin células de un receptor de un trasplante de órganos. La Figura 46 representa dos conjuntos de sondas - un conjunto de sondas para el Alelo 1 y un conjunto de sondas para el Alelo 2. 2605 y 2606 son moléculas dianas correspondientes al Alelo 1 y Alelo 2, respectivamente. Un primer conjunto de sondas contiene la sonda miembro 2602. 2602 contiene una marca (2601) de tipo "A". 2602 contiene una etiqueta de afinidad (2600) que se puede utilizar para el aislamiento y la identificación de la sonda producto.

Un segundo conjunto de sondas con la sonda miembro 2609 tiene características respectivas como en el primer conjunto de sondas. Sin embargo, 2609 contiene una marca (2608) de tipo "B", distinguible del tipo "A". 2609 contiene una etiqueta de afinidad (2607) que puede ser idéntica o única con respecto a 2600.

En esta realización, 2602 y 2609 contienen secuencias que son casi idénticas, y difieren en un solo nucleótido en la secuencia. Por lo tanto, las secuencias de hibridación de estas dos sondas son complementarias al Alelo 1 (2605), o al Alelo 2 (2606). Adicionalmente, la longitud de cada dominio de hibridación en 2602 y 2609, así como las condiciones de hibridación experimentales, están diseñadas de tal manera que la sonda 2602 solo hibridará con el Alelo 1 y la sonda 2609 solo hibridará con el Alelo 2. El propósito de este tipo de ensayo es poder cuantificar con precisión la frecuencia del Alelo 1 y el Alelo 2 en una muestra.

En esta realización, se puede utilizar ADN polimerasa u otra enzima para sintetizar una nueva secuencia de polinucleótidos, p. ej., 2604 en el caso del Alelo 1 o 2611 en el caso del Alelo 2. En esta realización, 2604 y 2611 pueden contener una o más marcas (2603, 2610) de tipo "C", posiblemente como resultado de la incorporación de uno o más nucleótidos que portan una marca de tipo "C". Por lo tanto, las sondas producto contendrán una combinación de marcas. Para el Alelo 1, las sondas producto contendrán marcas de tipo "A" y de tipo "C", mientras que las sondas producto del Alelo 2 contendrán marcas de tipo "B" y de tipo "C". Esta realización da como resultado sondas producto con alta especificidad para secuencias en el Alelo 1 o el Alelo 2, respectivamente.

Las Figuras 55-58 ilustran una modificación del procedimiento general descrito con respecto a las Figuras 21-46. La Figura 55 representa dos conjuntos de sondas; un conjunto de sondas para el Locus 1 y un conjunto de sondas para el Locus 2 - aunque como se mencionó anteriormente, se pueden diseñar múltiples conjuntos de sondas para cada locus genómico. El brazo izquierdo del conjunto de sondas del Locus 1 consiste en una secuencia de cebado

directo, una secuencia de etiqueta de afinidad y un homólogo para la secuencia del Locus 1. El brazo derecho del conjunto de sondas del Locus 1 consiste en un homólogo para la secuencia del Locus 1 y una secuencia de cebado inverso para el marcaje del conjunto de sondas del Locus 1 con la marca A. El brazo izquierdo del conjunto de sondas del Locus 2 consiste en una secuencia de cebado directo, una secuencia de la etiqueta de afinidad y un homólogo para la secuencia del Locus 2. El brazo derecho del conjunto de sondas del Locus 2 consiste en un homólogo para la secuencia del Locus 2 y una secuencia de cebado inverso para el marcaje del conjunto de sondas del Locus 2 con la marca B. La secuencia de cebado directo y la secuencia de la etiqueta de afinidad son idénticas para los conjuntos de sondas para el Locus 1 y el Locus 2. Las secuencias homólogas son específicas de un solo locus genómico. Las secuencias homólogas del locus para cada conjunto de sondas están inmediatamente adyacentes entre sí, de modo que cuando hibridan con sus loci diana, se apoyan entre sí inmediatamente y, por lo tanto, se pueden ligar para formar una molécula continua. La secuencia de cebado inverso es específica para la marca (p. ej., la marca A o la marca B) que se utilizará en el marcaje de sondas producto para un locus en particular para una secuencia de etiqueta de afinidad particular.

La Figura 56 representa el flujo de trabajo de procedimiento que se aplicaría a la colección de conjuntos de sondas, tales como los conjuntos de sondas ilustrados en la Figura 55. Esta representación se basa en un conjunto de sondas para un locus genómico (p. ej., el conjunto de sondas para el Locus 1 que se muestra en Figura 55). En la Etapa 1, la colección de conjuntos de sondas se mezcla con ADN purificado sin células. En la Etapa 2, las secuencias específicas del locus en cada conjunto de sondas hibridan con sus correspondientes secuencias homólogas en la muestra de ADN sin células. En la Etapa 3, se añade una enzima ligasa para catalizar la formación de un enlace fosfodiéster entre la base 3' en el homólogo del brazo izquierdo y el brazo 5' del homólogo derecho, cerrando la muesca entre los dos brazos y formando así una molécula continua que es la sonda producto. En la Etapa 4, se añaden los cebadores modificados y los componentes de la reacción de PCR (*Taq* polimerasa, dNTP y tampón de reacción) para amplificar la sonda producto ligada. El Cebador Directo se modifica porque tiene un grupo 5' fosfato que lo convierte en un molde preferido para la exonucleasa Lambda utilizada en la Etapa 6 y el Cebador Inverso se modifica porque contiene la marca (círculo de color azul) que es específica para las sondas producto para un locus particular para una etiqueta de afinidad particular. En la Etapa 5, la sonda producto se amplifica por PCR para producir un producto de PCR de doble hebra en el que la hebra directa contiene un grupo fosfato 5' y la hebra inversa contiene una marca 5'. En la Etapa 6, se añade la exonucleasa Lambda para digerir la hebra directa en una dirección 5' a 3'; el grupo fosfato 5' en la hebra directa lo convierte en un molde preferido para la digestión con la exonucleasa Lambda. El material resultante es de hebra sencilla (solo hebra inversa) con una marca 5'. Esto representa el material diana marcado para la hibridación con una micromatriz o monocapa.

La Figura 57 representa una versión modificada del flujo de trabajo de procedimiento ilustrado en la Figura 56. En esta realización, el brazo izquierdo de cada conjunto de sondas contiene una molécula de biotina terminal que se indica con una "B" en las Etapas 1 a 6 de la Figura. Esta biotilación permite la purificación de la colección de sondas producto después de completar la reacción de hibridación-ligación y antes de la amplificación por PCR. El flujo de trabajo para esta realización es idéntico al descrito en la Figura 57 para las Etapas 1 a 3. En la Etapa 4, se añaden cuentas magnéticas recubiertas con estreptavidina a la reacción de hibridación-ligación. La molécula de biotina contenida en las sondas producto unirá los productos a la estreptavidina. En la Etapa 5, las cuentas magnéticas se lavan para eliminar el ADN no biotilado (ADN genómico sin células y oligonucleótidos del brazo derecho), dando como resultado una sonda producto purificada. Los Etapas 6 a 9 se realizan de la misma manera que se ha descrito para las Etapas 4 a 7 en la Figura 56.

La Figura 58 proporciona un ejemplo de cómo las sondas producto para el Locus 1 y el Locus 2 se pueden marcar con diferentes moléculas de marcaje. En la Figura 58A, las sondas producto del Locus 1 están marcadas con la marca A (color verde) y las sondas producto del Locus 2 están marcadas con la marca B (color rojo) en una reacción de amplificación por PCR. Las sondas producto para ambos loci contienen la secuencia de la etiqueta de afinidad A. En la Figura 58B, la mezcla de sondas producto marcadas diferencialmente hibrida con una ubicación de micromatriz en la que la secuencia de la sonda de captura es complementaria a la secuencia de la etiqueta de afinidad A. En la Figura 58C, se genera una imagen de la ubicación de la micromatriz y se cuenta el número de moléculas de la marca A y la marca B para proporcionar una medida relativa de los niveles de Locus 1 y Locus 2 presentes en la muestra.

La Figura 59 proporciona evidencia de que las sondas producto que representan una multitud de ubicaciones genómicas para un locus se pueden generar de una manera específica de la enzima ligasa utilizando el procedimiento de hibridación-ligación. Ocho conjuntos de sondas, que consistían cada uno en un componente del brazo izquierdo y uno del brazo derecho como se describe en la Figura 55 y, que contenían homólogos para ocho ubicaciones del cromosoma 18, se hibridaron con moldes de oligonucleótidos sintéticos (aproximadamente 48 nucleótidos) y se ligaron utilizando una enzima ligasa para unirse a los brazos izquierdo y derecho. Los productos de reacción se analizaron utilizando electroforesis en gel de poliacrilamida desnaturalizante. La Calle 1 del gel contiene una escala de peso molecular para indicar el tamaño de las bandas de ADN. Las Calles 2 a 9 contienen productos de la reacción de hibridación-ligación para los ocho conjuntos de sondas del cromosoma 18. Una banda de ADN de aproximadamente 100 nucleótidos, que representa la sonda producto del brazo izquierdo de aproximadamente 60 nucleótidos y del brazo derecho de aproximadamente 40 nucleótidos, se encuentra presente en cada una de las Calles 2 a 9. Las Calles 10 y 11 contienen reacciones de control negativo a las que no se les añadió enzima ligasa. No se encuentra presente ninguna banda de ADN de aproximadamente 100 nucleótidos en las Calles 10 y 11.

La Figura 60 proporciona datos que indican que los conjuntos de sondas se pueden utilizar para detectar cambios relativos en el estado del número de copias. Se utilizó una mezcla de ocho conjuntos de sondas que contenían homólogos para ocho ubicaciones distintas del cromosoma X para someter a ensayo las líneas celulares que contenían diferentes números de cromosoma X indicados en la Tabla 1.

5 **Tabla 1:** Líneas celulares que contienen diferentes números de copias del cromosoma X

ID Línea Celular Coriell	Número de copias del cromosoma X
NA12138	1
NA13783	2
NA00254	3
NA01416	4
NA06061	5

10 La PCR cuantitativa se utilizó para determinar la cantidad sonda producto presente para cada línea celular después de los procedimientos de hibridación-ligación y purificación descritos en la Figura 57 (Etapas 1 a 5). Como se ilustra en la Figura 60A, el estado del número de copias medido para las diversas líneas celulares siguió la tendencia esperada indicada en la Tabla 1. Por ejemplo, qPCR indicó un estado de número de copias de menos de dos para NA12138, que tiene una copia del cromosoma X. El estado del número de copias medido para NA00254 (tres copias de X) fue mayor que dos, para NA01416 (cuatro copias de X) fue mayor que tres, y para NA06061 (cinco copias de X) fue mayor que cuatro. La capacidad de respuesta del procedimiento para detectar diferencias en el estado del número de copias se ilustra con más detalle en la Figura 60B, en la que se representa el estado del número de copias medido frente al estado del número de copias teórico.

15 La Figura 61 proporciona evidencia de que se pueden utilizar mezclas de sondas producto para generar datos cuantitativos de micromatrices como se describe en las Figuras 56 y 57.

20 La Figura 61A muestra imágenes representativas de fluorescencia de dos puntos de la matriz en dos canales de imágenes ortogonales (Alexa 488: color verde, Alexa 594; color rojo). Una región de interés (ROI) se selecciona automáticamente (círculo grande), siendo enmascarado cualquier contaminante brillante no deseado de la imagen (regiones esbozadas más pequeñas dentro de la ROI). Los fluoróforos individuales en los productos de ensayo hibridados únicos se visualizan como pequeñas características punteadas dentro del punto de la matriz. (i) Un punto "equilibrado" generó imágenes (que representan la entrada de dianas genómicas a una razón de concentración 1:1 con respecto al ensayo) en el canal verde y (ii) el mismo punto generó imágenes en el canal rojo. (iii) Un punto "aumentado" (que representa la entrada de dianas genómicas a una razón de concentración > 1:1 con respecto al ensayo) generó imágenes en el canal verde y (iv) el mismo punto generó imágenes en el canal rojo.

25 La Figura 61B presenta los recuentos sin procesar de los fluoróforos detectados en dos canales para cinco puntos, cada uno de los estados "equilibrado" y "aumentado". A pesar de algunas variaciones en el número absoluto de fluoros, los números en los dos canales siguen de cerca el caso "Equilibrado", pero demuestran una clara separación en el caso "Aumentado".

30 La Figura 61C presenta valores de la razón calculados para el número de fluoros en el canal verde dividido por el número de fluoros en el canal rojo, para los cinco puntos de cada una de las condiciones "Equilibrado" y "Aumentado". El caso "Equilibrado" se centra en una razón de 1,0 y el caso "Aumentado" se encuentra en una razón elevada. Considerando el caso "Equilibrado" como la comparación de dos loci genómicos equilibrados y el caso "Aumentado" como uno en el que un locus aumenta con respecto al otro, se puede calcular la confianza de separación de las dos condiciones mediante un test T, de 2 grupos independiente, que proporciona un valor de p de 8×10^{-14} .

35 La Figura 62 ilustra una modificación del procedimiento general descrito en las Figuras 55 a 58. En esta realización, un segundo conjunto de sondas, el Conjunto de Sondas B está diseñado para cada ubicación genómica de tal manera que las secuencias homólogas del genoma en el Conjunto de Sondas B son un complemento inverso del secuencias homólogas del genoma en el Conjunto de Sondas A. El Conjunto de Sondas A hibridará con la hebra inversa del ADN genómico y el Conjunto de Sondas B hibridará con la hebra directa del ADN genómico. Esta realización proporcionará una mayor sensibilidad con respecto a la realización descrita en las Figuras 55 a 58, ya que producirá aproximadamente el doble del número de sondas producto por locus.

La Figura 63 ilustra una modificación del procedimiento general descrito en la Figura 57. En esta realización, el Cebador Inverso utilizado en la Etapa 6 se modifica adicionalmente porque los cuatro enlaces que conectan los cinco primeros nucleótidos en la secuencia de oligonucleótidos son enlaces fosforotioato. Esta modificación dará como resultado que todos los productos de PCR generados durante la amplificación por PCR (Etapa 7) tengan una modificación de fosforotioato en el extremo 5'. Esta modificación protegerá la hebra inversa de cualquier digestión que pueda ocurrir durante el tratamiento con la exonucleasa Lambda en la Etapa 8.

Aunque el grupo fosfato 5' en la hebra directa la convierte en un molde preferido para la digestión con exonucleasa Lambda, la hebra inversa aún puede tener alguna vulnerabilidad a la digestión. La modificación con fosforotioato del extremo 5' de la hebra inversa reducirá su vulnerabilidad a la digestión con exonucleasa Lambda.

La Figura 64 ilustra una modificación del procedimiento general descrito en las Figuras 55 a 58. En esta realización, la amplificación por PCR de la sonda producto se reemplaza por una amplificación lineal al añadir el Cebador inverso pero no el Cebador Directo a la reacción de amplificación en la Etapa 6. Si solo está presente el Cebador Inverso, el producto de amplificación será de una sola hebra - la hebra inversa con una marca del extremo 5'. Como el producto de amplificación ya es de una sola hebra, no requiere un procesamiento adicional antes de la hibridación con una micromatriz, es decir, se puede omitir la digestión con la exonucleasa Lambda. Como no se utiliza un cebador directo en esta realización, no es necesario que el brazo izquierdo del conjunto de sondas contenga una secuencia de cebado directo. El brazo izquierdo consistiría en una secuencia de etiqueta de afinidad y una secuencia homóloga para el locus solamente como se ilustra en la Figura 64.

Una realización adicional del procedimiento general descrito en las Figuras 55 a 58 es una en la que el procedimiento de la reacción de ligación única en la Etapa 3 se reemplaza por un procedimiento de reacción de ligación ciclada. Esto se logra reemplazando la enzima ligasa termolábil (p. ej., ligasa T4) utilizada para catalizar la reacción de ligación por una ligasa termoestable (p. ej., ligasa *Taq*). Cuando se utiliza una ligasa termoestable, la reacción de hibridación-ligación se puede calentar a una temperatura que funda todos los dúplex de ADN (p. ej., 95°C) después de que haya ocurrido el ciclo inicial de hibridación y ligación. Esto hará que el ADN molde genómico esté completamente disponible para otra hibridación y ligación de conjuntos de sondas. La reducción posterior de la temperatura (p. ej., a 45°C) permitirá que ocurra este próximo evento de hibridación y ligación. Cada termociclación de la reacción de hibridación y ligación entre una temperatura que fundirá los dúplex de ADN y una que permitirá que se produzca la hibridación y la ligación aumentará linealmente la cantidad de sonda producto producida por la reacción. Si la reacción se expone a 30 ciclos de este tipo, se producirá hasta 30 veces la cantidad de sonda producto que a partir de un procedimiento en el que se utilice una única reacción de ligación.

La Figura 65 representa una realización adicional del procedimiento modificado descrito en la Figura 62. Esta realización aprovecha la reacción en cadena de la ligasa (LCR) al combinar la presencia del complemento inverso para cada conjunto de sondas con el uso de una ligasa termoestable para permitir una reacción de ligación ciclada en la que el producto se amplifica exponencialmente. La Figura 65 representa dos conjuntos de sondas, el Conjunto de Sondas A y el Conjunto de Sondas B para un locus; donde las secuencias homólogas del genoma en el Conjunto de Sondas B son el complemento inverso de las secuencias homólogas del genoma en el Conjunto de Sondas A. El brazo 5' de cada Conjunto de Sondas consiste en una secuencia de etiqueta de afinidad y un homólogo, mientras que el brazo 3' de cada Conjunto de Sondas consiste en una secuencia homóloga con una marca adjunta. En el primer ciclo de una reacción de termociclación, el ADN genómico será la única molde disponible para permitir que se produzcan la hibridación y la ligación para generar una sonda producto como se ilustra en la Figura 65A. Sin embargo, en el segundo ciclo, la Sonda Producto B generada en el primer ciclo actuará como un molde adicional para el Conjunto de Sondas A y, de la misma manera, la Sonda Producto A generada en el primer ciclo actuará como un molde adicional para el Conjunto de Sondas B, como se ilustra en la Figura 65B. De esta misma manera, las sondas producto de cada ciclo sucesivo actuarán como molde para la hibridación y la ligación del conjunto de sondas en el ciclo siguiente. Este procedimiento eliminaría la necesidad de amplificación por PCR de la sonda producto, que se puede utilizar directamente como diana de la micromatriz.

Otra realización del procedimiento representado en la Figura 65 es aquella que emplea la LCR pero utiliza conjuntos de sondas que tienen la estructura descrita en la Figura 55, es decir, los brazos izquierdo y derecho están flanqueados por secuencias de cebado, el brazo izquierdo contiene una molécula de biotina y el brazo derecho no contiene una marca. Una vez completada la LCR, las sondas producto se purifican utilizando cuentas magnéticas (opcional) y a continuación se amplifican por PCR y se prepara la diana de la micromatriz como se ilustra en las Figuras 56 y 57.

La Figura 66 representa otra realización del procedimiento representado en la Figura 65. El brazo 5' de cada Conjunto de Sondas consiste en una secuencia de etiqueta de afinidad y un homólogo, mientras que el brazo 3' de cada Conjunto de Sondas consiste en una secuencia homóloga y una secuencia de cebado sin una marca anclada como se ilustra en la Figura 66A. Una vez completada la LCR, la sonda producto se puede purificar. El producto de LCR se amplificaría a continuación de manera lineal mediante la adición de un cebador único que tiene una marca anclada, junto con los componentes de la reacción (polimerasa *Taq*, dNTP y tampón de reacción) como se ilustra en la Figura 66B. El producto de esta amplificación sería de una sola hebra (solo hebra inversa) con una marca 5' como se ilustra en la Figura 66C. En consecuencia, no sería necesario tratarlo con la exonucleasa Lambda, sino que podría utilizarse directamente como diana de la micromatriz.

En otro aspecto, la variación genética determinada por los métodos descritos en la presente memoria indica la presencia o ausencia de cáncer, variabilidad farmacocinética, toxicidad de fármaco, rechazo del trasplante o aneuploidía en el sujeto. En otro aspecto, la variación genética determinada indica presencia o ausencia de cáncer. Por consiguiente, los métodos descritos en la presente memoria se pueden realizar para diagnosticar cáncer.

- 5 Un reto importante en oncología es la detección temprana del cáncer. Esto es particularmente cierto en los cánceres en los que es difícil obtener imágenes o biopsias (p. ej., cáncer de páncreas, cáncer de pulmón). El ADN tumoral sin células (ADNsc tumoral) en la sangre de un paciente ofrece un método para detectar un tumor de forma no invasiva. Estos pueden ser tumores sólidos, tumores benignos, micro tumores, tumores líquidos, metástasis u otros crecimientos somáticos. La detección se puede realizar en cualquier estadio del desarrollo del tumor, aunque
- 10 idealmente en un estadio temprano (Estadio I o Estadio II). La detección temprana permite una intervención (p. ej., cirugía, quimioterapia, tratamiento farmacéutico) que puede prolongar la vida o llevar a la remisión. Otros problemas en oncología incluyen la verificación de la eficacia del tratamiento, la titulación de la dosis de un agente terapéutico, la recurrencia de un tumor en el mismo órgano que el tumor primario o en ubicaciones distales y la detección de metástasis. La presente invención puede ser utilizada para todas estas aplicaciones.
- 15 En algunas realizaciones, los conjuntos de sondas de la presente descripción se pueden configurar para dirigirlos a variaciones genéticas conocidas asociadas con tumores. Estas pueden incluir mutaciones, SNP, variantes de número de copias (p. ej., amplificaciones, deleciones), variantes de copia neutra (p. ej., inversiones, translocaciones) y/o combinaciones complejas de estas variantes. Por ejemplo, las variaciones genéticas conocidas asociadas con los tumores incluyen las enumeradas en cancer.sanger.ac.uk/cancergenome/projects/cosmic;
- 20 nature.com/ng/journal/v45/n10/full/ng.2760.html#supplementary-information; y las Tablas 2 y 3 a continuación: ^BGEN = valor de p de corregido para FDR dentro del pico; ^KOncogén amplificado frecuentemente conocido o TSG eliminado; ^PGen supuesto de cáncer; ^ERegulador epigenético; ^MGen asociado a las mitocondrias; ** Inmediatamente adyacente a la región del pico; ^TAdyacente al telómero o centrómero del cromosoma acrocéntrico.

Tabla 2: Variaciones genéticas ilustrativas asociadas a tumores (Amplificación del gen)

Nombre del pico	Rango	Ubicación genómica	Región pico	Valor q GISTIC	Recuento de genes	Dianas	Genes frecuentemente mutados ^B
CCND1	1	11q13.3	chr11:69464719-69502928	2,05E-278	2	CCND1 ^K	CCND1 = 6,6e-08
EGFR	2	7p11.2	chr7:55075808-55093954	2,30E-240	1	EGFR ^K	EGFR = 2,2e-15
MI C	3	8q24.21	chr8:128739772-128762863	6,50E-180	1	MI C ^K	
TERC	4	3q26.2	chr3:169389459-169490555	5,40E-117	2	TERC ^P	
ERBB2	5	17q12	chr17:37848534-37877201	1,59E-107	1	ERBB2 ^K	ERBB2 = 1,3e-06
CCNE1	6	19q12	chr19:30306758-30316875	4,77E-90	1	CCNE1 ^K	
MCL1	7	1q21.3	chr1:150496857-150678056	1,25E-80	6	MCL1 ^K	
MDM2	8	12q15	chr12:69183279-69260755	2,59E-62	2	MDM2 ^K	
INTS4	9	11q14.1	chr11:77610143-77641464	1,01E-54	1	INTS4	
WHSC1L 1	10	8p11.23	chr8:38191804-38260814	3,43E-46	2	WHSC1L1 ^E , LETM2 ^M	
CDK4	11	12q14.1	chr12:58135797-58156509	5,14E-41	5	CDK4 ^K	CDK4 = 0,0048
KAT6A	12	8p11.21	chr8:41751300-41897859	2,97E-39	2	KAT6A ^{P,E} /KKB ^B **	
SOX2	13	3q26.33	chr3:181151312-181928394	1,21E-38	2	SOX2 ^K	
PDGFRA	14	4q12	chr4:54924794-55218386	1,08E-37	3	PDGFRA ^K	
BDH1	15	3q29	chr3:197212101-197335320	1,21E-31	1	BDH1 ^M	
1q44	16	1q44 ^T	chr1:242979907-249250621	4,48E-31	83	SMYD3 ^E	

Nombre del pico	Rango	Ubicación genómica	Región pico	Valor q GISTIC	Recuento de genes	Dianas	Genes frecuentemente mutados ^B
MDM4	17	1q32.1	chr1:204367383-204548517	1,98E-29	3	MDM4 ^K	
TERT	18	5p15.33	chr5:1287704-1300024	9,34E-27	1	TERT ^K	
KDM5A	19	12p13.33 ^T	chr12:1-980639	1,59E-25	11	KDM5A ^F	
MYCL1	20	1p34.2	chr1:40317971-40417342	3,99E-25	2	MYCL1 ^K	
IGF1R	21	15q26.3	chr15:98667475-100292401	8,62E-25	9	IGF1R ^K	
PARP10	22	8q24.3	chr8:144925436-145219779	5,44E-20	15	PARP10 ^{F,E} , CYC1 ^M	
G6PD	23	Xq28	chrX:153760870-153767853	3,66E-19	1	G6PD	
PHF12	24	17q11.2	chr17:27032828-27327946	1,75E-16	21	PHF12 ^E , ERAL1 ^M	
20q13.33	25	20q13.33	chr20:62187847-62214354	2,96E-16	2		
PAF1	26	19q13.2	chr19:39699366-39945515	1,66E-15	13	PAF1 ^{F,E}	IL28A=0,021, SUPT5H=0,084
BCL2L1	27	20q11.21	chr20:30179028-30320705	2,85E-15	4	BCL2L1 ^K	
TUBD1	28	17q23.1	chr17:57922443-57946458	7,19E-15	1	TUBD1	TUBD1 = 0,009
[ZNF703]	29	8p11.23	chr8:37492669-37527108	2,44E-14	0		
1q23.3	30	1q23.3	chr1:160949115-161115281	7,73E-13	9		
8q22.2	31	8q22.2	chr8:101324079-101652657	4,22E-11	3		SNX31 = 0,015
BRD4	32	19p13.12	chr19:15310246-15428182	5,04E-10	3	NOTCH3 ^P , BRD4 ^{F,E}	
KRAS	33	12p12.1	chr12:24880663-25722878	9,47E-10	7	KRAS ^K	KRAS = 1,5e-14

Nombre del pico	Rango	Ubicación genómica	Región pico	Valor q GISTIC	Recuento de genes	Dianas	Genes frecuentemente mutados ^B
NKX2-1	34	14q13.2	chr14:35587755-37523513	1,33E-09	14	NKX2-1 ^K	NFKBIA=0,0098, RALGAPA1=0,027
NFE2L2	35	2q31.2	chr2:178072322-178171101	5,48E-09	5	NFE2L2	NFE2L2 = 3,9e-14
ZNF217	36	20q13.2	chr20:52148496-52442225	5,83E-08	1	ZNF217 ^K	ZNF217 = 0,0082
13q34	37	13q34 ^T	chr13:108818892-115169878	6,28E-08	45	ING1 ^E	ING1 = 0,00026
KAT6B	38	10q22.2	chr10:76497097-77194071	1,41E-07	9	KAT6B ^E , VDAC2 ^M	
NSD1	39	5q35.3	chr5:176337344-177040112	1,75E-06	22	NSD1 ^E , PRELID1 ^M	NSD1 = 4,9e-10
FGFR3	40	4p16.3	chr4:1778797-1817427	2,14E-06	2	FGFR3 ^P , LETM1 ^M	FGFR3 = 0,00018
9p13.3	41	9p13.3	chr9:35652385-35739486	2,55E-06	8		
COX18	42	4q13.3	chr4:73530210-74658151	2,68E-06	7	COX18 ^M	
7q36.3	43	7q36.3 ^T	chr7:153768037-159138663	3,19E-06	30	PTPRN2 ^L , DPP6 ^L	
18q11.2	44	18q11.2	chr18:23857484-24119078	3,83E-06	2		
SOX17	45	8q11.23	chr8:55069781-55384342	2,02E-05	1	SOX17	SOX17 = 0,00092
11q22.2	46	11q22.2	chr11:102295593-102512085	0,00015337	3		
CBX8	47	17q25.3	chr17:77770110-77795534	0,00023029	1	CBX8 ^E	
AKT1	48	14q32.33	chr14:105182581-105333748	0,00028451	7	AKT1 ^K	AKT1 = 1,1e-14
CDK6	49	7q21.2	chr7:92196092-92530348	0,00069831	3	CDK6 ^K	
6p21.1	50	6p21.1	chr6:41519930-44297771	0,0010459	70		

Nombre del pico	Rango	Ubicación genómica	Región pico	Valor q GISTIC	Recuento de genes	Dianas	Genes frecuentemente mutados ^B
Ehf	51	11p13	chr11:34574296-34857324	0,0011002	1	Ehf	
6q21	52	6q21	chr6:107098934-107359899	0,0011806	4		
19q13.42	53	19q13.42 ^T	chr19:55524376-59128983	0,0013319	138	TRIM28 ^E , SUV420H2 ^E	ZNF471 = 5,4e-05
17q21.33	54	17q21.33	chr17:47346425-47509605	0,0025775	2		
BPTF	55	17q24.2	chr17:65678858-66288612	0,0028375	11	BPTF ^E	
E2F3	56	6p22.3	chr6:19610794-22191922	0,0033658	7	E2F3 ^K	
19p13.2	57	19p13.2	chr19:10260457-10467501	0,0038041	12	MRPL4 ^M	DNMT1 = 0,099
17q25.1	58	17q25.1	chr17:73568926-73594884	0,012337	2		
KDM2A	59	11q13.2	chr11:67025375-67059633	0,012445	3	KDM2A ^E	
8q21.13	60	8q21.13	chr8:80432552-81861219	0,020548	6	MRPS2s ^M	
2p15	61	2p15	chr2:59143237-63355557	0,021056	25		XPO1 = 1,1e-05
14q11.2	62	14q11.2 ^T	chr14:1-21645085	0,027803	57		
NEDD9	63	6p24.2	chr6:11180426-11620845	0,082606	2	NEDD9 ^K	
5p13.1	64	5p13.1	chr5:35459650-50133375	0,094657	61		SLC1A3=0,0021, IL7R=0,0021
LINC00536	65	8q23.3	chr8:116891361-117360815	0,095294	1	LINC00536	
10p15.1	66	10p15.1	chr10:4190059-6130004	0,10391	21		
22q11.21	67	22q11.21	chr22:18613558-23816427	0,13213	105		

Nombre del pico	Rango	Ubicación genómica	Región pico	Valor q GISTIC	Recuento de genes	Dianas	Genes frecuentemente mutados ^B
PHF3	68	6q12	chr6:63883156-64483307	0,17851	4	PHF3 ^E , EYS ^L	PHF3 = 0,051
PAX8	69	2q13	chr2:113990138-114122826	0,19717	2	PAX8 ^K	
						SMARCA2 ^E ,	
						KDM4 ^E ,	
						UHRF2 ^E ,	
9p24.2	70	9p24.2 ^I	chr9:1-7379570	0,20405	45	KIAA2026 ^F	

Tabla 3: Variaciones genéticas ilustrativas asociadas a tumores (Supresión del gen)

Nombre del pico	Rango	Ubicación genómica	Región pico	Valor q GISTIC	Recuento de genes	Dianas	Genes frecuentemente mutados ^B
CDKN2A	1	9p21.3	chr9:21865498-22448737	0	4	CDKN2A ^K	CDKN2A = 4,4e-15
STK11	2	19p13.3	chr19:1103715-1272039	1,46E-238	7	STK11 ^K	STK11 = 2,5e-13
PDE4D	3	5q11.2	chr5:58260298-59787985	2,02E-143	3	PDE4D ^L	
PARK2	4	6q26	chr6:161693099-163153207	5,85E-137	1	PARK2 ^{L,K}	
LRP1B	5	2q22.1	chr2:139655617-143637838	4,25E-107	1	LRP1B ^L	
CSMD1	6	8p23.2	chr8:2079140-6262191	2,39E-96	1	CSMD1 ^L	
1p36.23	7	1p36.23	chr1:7829287-8925111	1,23E-93	8		
ARID1A	8	1p36.11	chr1:26900639-27155421	5,74E-87	2	ARID1A ^K	ARID1A = 1,5e-14

Nombre del pico	Rango	Ubicación genómica	Región pico	Valor q GISTIC	Recuento de genes	Dianas	Genes frecuentemente mutados ^B
PTEN	9	10q23.31	chr10:89615138-90034038	1,12E-79	2	PTEN ^K	PTEN = 2,2e-15
WWOX	10	16q23.1	chr16:78129058-79627770	8,14E-76	1	WWOX ^L	WWOX = 0,092
RB1	11	13q14.2	chr13:48833767-49064807	3,88E-75	2	RB1 ^K	RB1 = 1,7e-13
FAM190A	12	4q22.1	chr4:90844993-93240505	9,26E-75	1	FAM190A ^L	
2q37.3	13	2q37.3 ^T	chr2:241544527-243199373	1,77E-70	29	ING5 ^E	
22q13.32	14	22q13.32 ^T	chr22:48026910-51304566	8,20E-65	45	BRD1 ^E , HDAC10 ^E	
11p15.5	15	11p15.5 ^T	chr11:1-709860	1,02E-62	34	SIRT3 ^E , PHRF1 ^E	HRAS = 7,8e-13
LINC00290	16	4q34.3	chr4:178911874-183060693	1,21E-55	1	LINC00290	
FHIT	17	3p14.2	chr3:59034763-61547330	3,01E-55	1	FHIT ^L	
RBFOX1	18	16p13.3	chr16:5144019-7771745	1,00E-45	1	RBFOX1 ^L	
PTPRD	19	9p24.1	chr9:8310705-12693402	3,24E-38	1	PTPRD ^L	
18q23	20	18q23 ^T	chr18:74979706-78077248	1,69E-37	12		
FAT1	21	4q35.2	chr4:187475875-188227950	6,81E-36	1	FAT1 ^K	FAT1 = 2,4e-15
MPHOSP H8	22	13q12.11 ^T	chr13:1-20535070	2,57E-31	10	MPHOSP H8 ^F	
15q15.1	23	15q15.1	chr15:41795901-42068054	2,71E-29	4		MGA = 0,0083, RPAP1=0,035
11q25	24	11q25 ^T	chr11:133400280-135006516	4,93E-26	14		

Nombre del pico	Rango	Ubicación genómica	Región pico	Valor q GISTIC	Recuento de genes	Dianas	Genes frecuentemente mutados ^B
1p13.2	25	1p13.2	chr1:110048528-117687124	1,69E-25	100	TRIM33 ^E	NRAS= 1,8e-13, CD58=0,079
NF1	26	17q11.2	chr17:29326736-29722618	6,59E-23	5	NF1 ^K	NF1 = 3,3e-13
MACROD2	27	20p12.1	chr20:14302876-16036135	9,00E-19	3	MACROD2 ^L	
7p22.3	28	7p22.3 ^T	chr7:1-1496620	1,04E-17	18		
6p25.3	29	6p25.3	chr6:1608837-2252425	3,01E-17	2		
21q11.2	30	21q11.2 ^T	chr21:1-15482604	2,34E-14	14		
9p13.1	31	9p13.1	chr9:38619152-71152237	9,75E-14	48		
ZNF132	32	19q13.43 ^T	chr19:58661582-59128983	3,77E-13	24	TRIM28 ^E , ZNF132	
5q15	33	5q15	chr5:73236070-114508587	8,15E-13	156	APC ^K , CHD1 ^F	APC= 2,6e-13, RASA1=0,0029
MLL3	34	7q36.1	chr7:151817415-152136074	9,26E-13	1	MLL3 ^{K, E}	MLL3 = 1,1e-05
19q13.32	35	19q13.32	chr19:47332686-47763284	2,38E-12	10		
15q12	36	15q12 ^T	chr15:1-32929863	3,40E-11	155		OTUD7A = 0,027
12q24.33	37	12q24.33 ^T	chr12:131692956-133851895	1,24E-10	27		POLE= 3,9E-05, PGAM5=0,038
10q26.3	38	10q26.3 ^T	chr10:135190263-135534747	2,09E-10	14		

Nombre del pico	Rango	Ubicación genómica	Región pico	Valor q GISTIC	Recuento de genes	Dianas	Genes frecuentemente mutados ^B
6q21	39	6q21	chr6:86319089-117076132	4,56E-10	141	PRDM1 ^E , HDAC2 ^E , PRDM13 ^E	PRDM1 = 0,00054
PPP2R2A	40	8p21.2	chr8:25896447-26250295	1,78E-09	1	PPP2R2A	
IKZF2	41	2q34	chr2:211542637-214143899	3,24E-09	4	IKZF2 ^K , ERBB4 ^L	ERBB4 = 0,00058
CNTN4	42	3p26.3 ^T	chr3:1-3100786	6,44E-09	3	CNTN4 ^L	
3p12.2	43	3p12.2	chr3:75363575-86988125	1,22E-07	12	ROBO1 ^L , CADM2 ^L	
RAD51B	44	14q24.1	chr14:68275375-69288431	1,38E-07	2	RAD51B ^L	ZFP36L1 = 0,0016
11q23.1	45	11q23.1	chr11:105849158-117024891	5,31E-07	84	ATM ^K	ATM= 1,4e-06, POU2AF1=0,082
IMMP2L	46	7q31.1	chr7:109599468-111366370	5,74E-07	2	IMMP2L ^L	
NEGRI	47	1p31.1	chr1:71699756-74522473	7,25E-07	2	NEGR1 ^L	
BRCA1	48	17q21.31	chr17:41178765-41336147	7,25E-07	2	BRCA1 ^K	BRCA1 = 3,5e-08
9q34.3	49	9q34.3	chr9:135441810-139646221	8,73E-06	94	NOTCH1 ^K , BRD3 ^E , GTF3C4 ^E	NOTCH1= 1e-08, RXRA= 2,1e-05, COL5A1=0,0022, TSC1=0,012
ANKS1B	50	12q23.1	chr12:99124001-100431272	8,73E-06	2	ANKS1B ^L	
DMD	51	Xp21.2	chrX:30865118-34644819	5,15E-05	4	DMD ^L	

Nombre del pico	Rango	Ubicación genómica	Región pico	Valor q GISTIC	Recuento de genes	Dianas	Genes frecuentemente mutados ^B
ZMYND11	52	10p15.3 ^T	chr10:1-857150	7,12E-05	4	ZMYND11 ^E	
PRKG1	53	10q11.23	chr10:52644085-54061437	9,79E-05	3	PRKG1 ^L	
FOXK2	54	17q25.3	chr17:80443432-80574531	0,00019271	1	FOXK2	
AGBL4	55	1p33	chr1:48935280-50514967	0,000219	2	AGBL4 ^L	
CDKN1B	56	12p13.1	chr12:12710990-12966966	0,00035777	5	CDKN1B ^K	CDKN1B = 2,2e-06
						SETD3 ^E ,	AKT1= 2,1e-13,
14q32.33	57	14q32.33 ^T	chr14:94381429-107349540	0,00074358	227	TDRD9 ^E	TRAF3= 9,1e-05
14q11.2	58	14q11.2 ^T	chr14:1-30047530	0,0010181	162	PRMT5 ^E , CHD8 ^E	CHD8 = 0,034
2p25.3	59	2p25.3 ^T	chr2:1-20072169	0,0011137	86	MYCN ^K	MYCN=0,068
5q35.3	60	5q35.3 ^T	chr5:153840473-180915260	0,0028515	212	NSD1 ^E , ODZ2 ^L	NPM1= 3,5e-13, NSD1= 1,9e-09, ZNF454=0,0019, UBLCP1=0,03, GABRB2=0,07
PTTG1IP	61	21q22.3	chr21:46230687-46306160	0,012227	1	PTTG1IP	
22q11.1	62	22q11.1 ^T	chr22:1-17960585	0,020332	15		
SMAD4	63	18q21.2	chr18:48472083-48920689	0,036866	3	SMAD4 ^K	SMAD4 = 6,6e-15
17p13.3	64	17p13.3 ^T	chr17:1-1180022	0,040814	16		

Nombre del pico	Rango	Ubicación genómica	Región pico	Valor q GISTIC	Recuento de genes	Dianas	Genes frecuentemente mutados ^B
4p16.3	65	4p16.3 ^T	chr4:1-1243876	0,056345	27		
9p21.2	66	9p21.2	chr9:27572512-28982153	0,091742	3		
10q25.1	67	10q25.1	chr10:99340084-113910615	0,11879	137	HPSE2 ^L , SMNDC1 ^E	SMC3=0,00031, GSTO2=0,086
SMYD3	68	1q44	chr1:245282267-247110824	0,15417	8	SMYD3 ^E	
8p11.21	69	8p11.21	chr8:42883855-47753079	0,17382	4		
Xp22.33	70	Xp22.33 ^T	chrX:1-11137490	0,21462	52		MXRA5 = 0,031

En el método de diagnóstico del cáncer de acuerdo con algunas realizaciones, las inversiones que se producen en lugares conocidos (Figura 67A) se pueden dirigir fácilmente diseñando sondas que se solapan al menos parcialmente con el punto de ruptura en un brazo de la sonda. Una primera sonda que se une a la secuencia "normal" se dirige al material genómico no invertido (Figura 67B) y porta un primer tipo de marca. Una segunda sonda que se une a la diana "invertida" porta un segundo tipo de marca (Figura 67C). Un brazo de sonda derecho común se une a una secuencia nativa que no es susceptible de inversión, inmediatamente adyacente a las dos primeras sondas. Este brazo derecho de la sonda porta adicionalmente una etiqueta desplegable común que localiza las sondas producto en la misma región de un sustrato de generación de imágenes. De esta manera, los pares de sondas pueden hibridar con las dianas genómicas, ligarse y generar imágenes para obtener recuentos relativos de las dos especies subyacentes.

De manera similar, también se pueden someter a ensayo translocaciones que tienen puntos de ruptura conocidos. La Figura 68A muestra dos elementos genéticos que están en su orden nativo o están translocados. Los brazos de sondas que se solapan al menos parcialmente con estos puntos de ruptura de translocación permiten la diferenciación entre los órdenes normales y transpuestos de material genético. Como se muestra en las Figuras 68B y 68C, al elegir marcas únicas en los dos brazos izquierdos, las sondas producto ligadas resultantes se pueden distinguir y contar durante la generación de imágenes.

Estos métodos para detectar cambios de copia neutral (p. ej., inversiones, translocación) también se pueden utilizar para detectar variantes de la línea germinal en el cáncer o en otras enfermedades o afecciones.

Las mutaciones o los SNP también están implicados en numerosos cánceres, y se eligen como diana de manera similar a los que son interrogados para determinar la fracción fetal en la aplicación de diagnóstico prenatal. En algunas realizaciones mostradas en las Figuras 69A y 69B, los brazos de la sonda izquierda están diseñados para aprovechar un desequilibrio energético causado por uno o más SNP no coincidentes. Esto hace que un brazo de sonda (1101, que porta una marca) se una más favorablemente que un segundo brazo de sonda (1107, que porta un segundo tipo de marca). Ambos diseños están ligados al mismo brazo derecho de la sonda (1102) que porta la etiqueta desplegable universal.

La sangre de un paciente dado puede ser sondeada por un método, o un híbrido de más de un método. Adicionalmente, en algunos casos, la personalización de sondas específicas para un paciente puede ser valiosa. Esto implicaría caracterizar las características del tumor (SNP, translocaciones, inversiones, etc.) en una muestra del tumor primario (p. ej., una biopsia) y crear uno o más conjuntos de sondas personalizadas que estén optimizados para detectar aquellas variaciones genéticas específicas del paciente en la sangre del paciente, lo que proporciona un método de verificación no invasivo, de bajo coste. Esto podría tener un valor significativo en el caso de recaída, donde es ideal detectar una recurrencia a bajo nivel de un tipo de tumor (idéntico o relacionado con el tumor original) lo antes posible.

Para vías comunes de progresión de la enfermedad, se pueden diseñar paneles adicionales para anticipar y verificar el avance de la enfermedad. Por ejemplo, si las mutaciones tienden a acumularse en un orden dado, las sondas se pueden diseñar para verificar el estado actual y los "puntos de control" de progresión, y guiar las opciones de terapia.

Detección precoz del cáncer: Por ejemplo, la translocación ALK se ha asociado con el cáncer de pulmón. Se puede utilizar una sonda diseñada para interrogar la translocación ALK para detectar tumores de este tipo a través de una muestra de sangre. Esto sería altamente ventajoso, ya que el método convencional para detectar tumores de pulmón es a través de una radiografía de tórax, un procedimiento costoso que puede ser perjudicial para la salud del paciente y, por lo tanto, no se realiza de manera convencional.

Detección de recurrencia del tipo de tumor primario: Por ejemplo, un tumor mamario HER2+ se extrae mediante cirugía y el paciente está en remisión. Se puede utilizar una sonda dirigida al gen HER2 para verificar las amplificaciones del gen HER2 en uno o más puntos temporales. Si se detectan, el paciente puede tener un segundo tumor HER2+ en el sitio del primario o en otro lugar.

Detección de tipos de tumores no primarios: Por ejemplo, un tumor mamario HER2+ se extrae mediante cirugía y el paciente está en remisión. Se puede utilizar una sonda dirigida al gen EGFR para verificar los tumores EGFR+. Si se detectan, el paciente puede tener un segundo tumor EGFR+ en el sitio del primario o en otro lugar.

Detección de metástasis: Por ejemplo, la paciente tiene un tumor mamario HER2+. Se puede utilizar una sonda diseñada para interrogar la translocación ALK para detectar tumores de este tipo a través de una muestra de sangre. Este tumor puede no estar en la mama y es más probable que esté en el pulmón. Si se detectan, el paciente puede tener un tumor metastásico distal con respecto al órgano primario.

Determinación de la heterogeneidad tumoral: Muchos tumores tienen múltiples poblaciones clonales caracterizadas por diferentes variantes genéticas. Por ejemplo, un tumor mamario puede tener una población de células que son HER2+ y otra población de células que son EGFR+. El uso de sondas diseñadas para dirigirse a ambas variantes permitiría identificar esta heterogeneidad genética subyacente.

Medición de la carga tumoral: En todos los ejemplos anteriores, la cantidad de ADNsc del tumor se puede medir y se puede utilizar para determinar el tamaño, la tasa de crecimiento, la agresividad, el estadio, el pronóstico, el diagnóstico y otros atributos del tumor y del paciente. Lo ideal es que las mediciones se realicen en más de un punto temporal para mostrar los cambios en la cantidad de ADNsc del tumor.

- 5 Verificación del tratamiento: Por ejemplo, un tumor mamario HER2+ se trata con Herceptin. Se puede utilizar una sonda dirigida al gen HER2 para verificar la cantidad de ADNsc del tumor, que puede ser un indicador del tamaño del tumor. Esto se puede utilizar para determinar si el tumor está cambiando de tamaño y el tratamiento se puede modificar para optimizar el resultado en el paciente. Esto puede incluir cambiar la dosis, interrumpir el tratamiento, cambiar a otra terapia, combinar múltiples terapias.
- 10 Escrutinio del ADN tumoral: Actualmente no existe un escrutinio universal para detectar el cáncer. La presente invención ofrece una forma de detectar tumores en algunas o todas las ubicaciones del cuerpo. Por ejemplo, se desarrolla un panel de sondas con un espaciamiento de 100 kb a través del genoma. Este panel se puede utilizar como una forma de detectar la variación genética en todo el genoma. En un ejemplo, el panel detecta cambios en el número de copias de un cierto tamaño en todo el genoma. Tales cambios en el número de copias están asociados con células tumorales y, por lo tanto, la prueba detecta la presencia de células tumorales. Los diferentes tipos de tumores pueden producir diferentes cantidades de ADNsc tumoral o pueden tener variaciones en diferentes partes del genoma. Como tal, la prueba puede identificar qué órgano está afectado. Adicionalmente, la cantidad de ADNsc tumoral medida puede indicar el estadio o el tamaño del tumor o la ubicación del tumor. De esta manera, la prueba es un escrutinio del genoma completo para muchos o todos los tipos de tumores.
- 15
- 20 Para todas las pruebas anteriores, con el fin de mitigar los falsos positivos, se puede utilizar un umbral para determinar la presencia o la certeza de un tumor. Adicionalmente, la prueba puede repetirse en múltiples muestras o en múltiples puntos temporales para aumentar la certeza de los resultados. Los resultados también se pueden combinar con otra información o síntomas para proporcionar más información o información más segura sobre el tumor.
- 25 Los conjuntos de sondas y cebadores ilustrativos que se pueden utilizar en el método descrito en la presente memoria para medir el número de copias de las regiones de ácido nucleico de interés se enumeran en la Tabla 4 a continuación. Cada uno de los conjuntos de sondas ilustrativos en la Tabla 4 comprende dos sondas. La primera sonda (de etiquetado) tiene una estructura que incluye un sitio de cebado directo, etiqueta y homología 1. La segunda sonda (de marcaje) tiene una estructura, que incluye la homología 2 y el sitio del cebador inverso, que se utiliza en el marcaje. También se muestran las secuencias componentes de las sondas (etiqueta, secuencia de homología, etc.).
- 30

Tabla 4: Sondas y cebadores ilustrativos

Cromosoma	ID de locus	Sonda de Etiquetado (Cebador Directo + Etiqueta + 5pHom)	Sonda de Marcaje (3'-Hop + Cebador Inverso)	Cebador Directo	Etiqueta	Hom 5p	Hom 3p	Cebador Inverso
18	18-1	GCCCTCATCTTC TTCCCTGGGTTT TCACCAACCCTCA CCAAAGGAAGAA GTGAGGGCTTCT C (SEQ ID NO: 1)	CGTGCTAATAGTC TCAGGGCTTCCTC CACCGAACGTGT CT (SEQ ID NO: 17)	GCCCTCAT CTTCTTCC CTGC (SEQ ID NO: 33)	GTTCTCA CCACCCCT CACCAA (SEQ ID NO: 34)	GGAAGAA GTGAGGG CTTCTC (SEQ ID NO: 35)	CGTGCTA ATAGTCT CAGGGC (SEQ ID NO: 51)	TTCCTCCA CCGAACGT GTCT (SEQ ID NO: 67)
18	18-2	GCCCTCATCTTC TTCCCTGGGTTT TCACCAACCCTCA CCAAAAATCAAG GTGACCAAGCTCC (SEQ ID NO: 2)	CGACGGTTCATTG CTTCATTTTCCTC CACCGAACGTGT CT (SEQ ID NO: 18)	GCCCTCAT CTTCTTCC CTGC (SEQ ID NO: 33)	GTTCTCA CCACCCCT CACCAA (SEQ ID NO: 34)	AAATCAA GGTGACC AGCTCC (SEQ ID NO: 36)	CGACGC TTCATTG CTTCATT (SEQ ID NO: 52)	TTCCTCCA CCGAACGT GTCT (SEQ ID NO: 67)
18	18-3	GCCCTCATCTTC TTCCCTGGGTTT TCACCAACCCTCA CCAATCATCTGC CAAGACAGAAG TTC (SEQ ID NO: 3)	CTTGGCCAAAC AATTGTCTTCTCCT CCACCGAACGTG TCT (SEQ ID NO: 19)	GCCCTCAT CTTCTTCC CTGC (SEQ ID NO: 33)	GTTCTCA CCACCCCT CACCAA (SEQ ID NO: 34)	TCATCTG CCAAGAC AGAAATT C (SEQ ID NO: 37)	CTTGCGC CAAACA ATTGTCC (SEQ ID NO: 53)	TTCCTCCA CCGAACGT GTCT (SEQ ID NO: 67)
18	18-4	GCCCTCATCTTC TTCCCTGGGTTT TCACCAACCCTCA CCAAAGCAGGAG AGTCAAAGGTCT G (SEQ ID NO: 4)	GCTGCAGAGTTTG CATTCATTTCCCTC CACCGAACGTGT CT (SEQ ID NO: 20)	GCCCTCAT CTTCTTCC CTGC (SEQ ID NO: 33)	GTTCTCA CCACCCCT CACCAA (SEQ ID NO: 34)	GCAGGAG AGTCAAA GGTCTG (SEQ ID NO: 38)	GCTGCA GAGTTTG CATTCAT (SEQ ID NO: 54)	TTCCTCCA CCGAACGT GTCT (SEQ ID NO: 67)
18	18-5	GCCCTCATCTTC TTCCCTGGGTTT TCACCAACCCTCA CCAAAGTTGCCAT	CATACACACAGA CCGAGAGTCTTCC TCCACCGAACGT GTCT (SEQ ID NO:	GCCCTCAT CTTCTTCC CTGC (SEQ ID NO: 33)	GTTCTCA CCACCCCT CACCAA (SEQ ID NO: 34)	GTTGCCA TGGAGAT TGTTGC (SEQ ID NO: 39)	CATACA CACAGA CCGAGA GTC (SEQ	TTCCTCCA CCGAACGT GTCT (SEQ ID NO: 67)

Cromosoma	ID de locus	Sonda de Etiquetado (Cebador Directo + Etiqueta + 5pHom)	Sonda de Marcaje (3'-Hop + Cebador Inverso)	Cebador Directo	Etiqueta	Hom 5p	Hom 3p	Cebador Inverso
		GGAGATTGTTGC (SEQ ID NO: 5)	21)		NO: 34)	NO: 39)	ID NO: 55)	
18	18-6	GCCCTCATCTTC TTCCCTGCGTTC TCACCACCCCTCA CCAACAGCTCAG TGATGTCATTGC (SEQ ID NO: 6)	GGATGTCAGCCA GCATAAGTTTCCT CCACCGAACGTG TCT (SEQ ID NO: 22)	GCCCTCAT CTTCTTCC CTGC (SEQ ID NO: 33)	GTTCTCA CCACCCCT CACCAA (SEQ ID NO: 34)	CAGCTCA GTGATGT CATTGC (SEQ ID NO: 40)	GGATGT CAGCCA GCATAA GT (SEQ ID NO: 56)	TTCCTCCA CCGAACGT GTCT (SEQ ID NO: 67)
18	18-7	GCCCTCATCTTC TTCCCTGCGTTC TCACCACCCCTCA CCAACCTTGACC TCTGCTAAATGIG G (SEQ ID NO: 7)	GCAAGTGCCAAA CAGTCTCTTCT CCACCGAACGTG TCT (SEQ ID NO: 23)	GCCCTCAT CTTCTTCC CTGC (SEQ ID NO: 33)	GTTCTCA CCACCCCT CACCAA (SEQ ID NO: 34)	CCTTGAC CTCTGCT AATGTGG (SEQ ID NO: 41)	GCAAAGT GCCAAA CAGTTCT C (SEQ ID NO: 57)	TTCCTCCA CCGAACGT GTCT (SEQ ID NO: 67)
18	18-8	GCCCTCATCTTC TTCCCTGCGTTC TCACCACCCCTCA CCAACACCTGTC CAACAGCTACAG (SEQ ID NO: 8)	GATTCAGCACA CTTGAGTCTTCC TCCACCGAACGT GTCT (SEQ ID NO: 24)	GCCCTCAT CTTCTTCC CTGC (SEQ ID NO: 33)	GTTCTCA CCACCCCT CACCAA (SEQ ID NO: 34)	CACCTGT CCAACAG CTACAG (SEQ ID NO: 42)	GATTCCA GCACAC TTGAGTC T (SEQ ID NO: 58)	TTCCTCCA CCGAACGT GTCT (SEQ ID NO: 67)
X	X-1	GCCCTCATCTTC TTCCCTGCGTTC TCACCACCCCTCA CCAAGAATGTA TCTCAGGCCTG C (SEQ ID NO: 9)	CCGTTGCAGGTTT AAATGGCGCCT ATTGCAAGCCCTC TT (SEQ ID NO: 25)	GCCCTCAT CTTCTTCC CTGC (SEQ ID NO: 33)	GTTCTCA CCACCCCT CACCAA (SEQ ID NO: 34)	AGAAATGT ATCTTCA GGCCTGC (SEQ ID NO: 43)	CCGTTGC AGGTTTA AATGGC (SEQ ID NO: 59)	GCCCTATT GCAAGCC TCCT (SEQ ID NO: 68)

Cromosoma	ID de locus	Sonda de Etiquetado (Cebador Directo + Etiqueta + 5pHom)	Sonda de Marcaje (3'-Hop + Cebador Inverso)	Cebador Directo	Etiqueta	Hom 5p	Hom 3p	Cebador Inverso
X	X-2	GCCCTCATCTTC TTCCCTGCGTTC TCACCACCCCTCA CCAAAAGTAAATC ACTCTGGGTGGC (SEQ ID NO: 10)	CAAGAGTGCTTTA TGGGCTGCCCTA TTGCAAGCCCTCT T (SEQ ID NO: 26)	GCCCTCAT CTTCTTCC CTGC (SEQ ID NO: 33)	GTTCTCA CCACCCCT CACCAA (SEQ ID NO: 34)	AAGTAAT CACTCTG GGTGGC (SEQ ID NO: 44)	CAAGAG TGCTTTA TGGGCTT (SEQ ID NO: 60)	GCCCTATT GCAAAGCCC TCCTT (SEQ ID NO: 68)
X	X-3	GCCCTCATCTTC TTCCCTGCGTTC TCACCACCCCTCA CCAAAAGTCAACA GACAACTTGTG (SEQ ID NO: 11)	GCACTCAAGGAG ATCAGACTGGCC CTATTGCAAGCCC TCCTT (SEQ ID NO: 27)	GCCCTCAT CTTCTTCC CTGC (SEQ ID NO: 33)	GTTCTCA CCACCCCT CACCAA (SEQ ID NO: 34)	AGCTCAC AGACAAC CTTGTG (SFO ID NO: 45)	GCACTC AAGGAG ATCAGA CTG (SEQ ID NO: 61)	GCCCTATT GCAAAGCCC TCCTT (SEQ ID NO: 68)
X	X-4	GCCCTCATCTTC TTCCCTGCGTTC TCACCACCCCTCA CCAAAGCAATAGA CACCTACAGGCG (SEQ ID NO: 12)	GGTATCGAACT ACAACCACAGCC CTATTGCAAGCCC TCCTT (SEQ ID NO: 28)	GCCCTCAT CTTCTTCC CTGC (SEQ ID NO: 33)	GTTCTCA CCACCCCT CACCAA (SEQ ID NO: 34)	GCAATAG ACACCTA CAGGCG (SEQ ID NO: 46)	GGCTATC GAACTA CAACCA CA (SEQ ID NO: 62)	GCCCTATT GCAAAGCCC TCCTT (SEQ ID NO: 68)
X	X-5	GCCCTCATCTTC TTCCCTGCGTTC TCACCACCCCTCA CCAAAGCACATTA TCAAAGGCCACG (SEQ ID NO: 13)	GTAGTGTCTGTG GTGTGATCGCCCT ATTGCAAGCCCTC TT (SEQ ID NO: 29)	GCCCTCAT CTTCTTCC CTGC (SEQ ID NO: 33)	GTTCTCA CCACCCCT CACCAA (SEQ ID NO: 34)	GCACATT ATCAAAG GCCACG (SEQ ID NO: 47)	GTAGCT GTCTGTG GTGTGAT C (SEQ ID NO: 63)	GCCCTATT GCAAAGCCC TCCTT (SEQ ID NO: 68)
X	X-6	GCCCTCATCTTC TTCCCTGCGTTC TCACCACCCCTCA CCAAACAACGACC TAAAGCATGTGC (SEQ ID NO: 14)	CAAGAAACTTGG AGCCTTAGCAGC CCTATTGCAAGCC CTCTT (SEQ ID NO: 30)	GCCCTCAT CTTCTTCC CTGC (SEQ ID NO: 33)	GTTCTCA CCACCCCT CACCAA (SEQ ID NO: 34)	CAACGAC CTAAAGC ATGTGC (SEQ ID NO: 48)	CAAGAA ACTTCGA GCCTTAG CA (SEQ ID NO: 64)	GCCCTATT GCAAAGCCC TCCTT (SEQ ID NO: 68)

Cromosoma	ID de locus	Sonda de Etiquetado (Cebador Directo + Etiqueta + 5pHom)	Sonda de Marcaje (3'-Hop + Cebador Inverso)	Cebador Directo	Etiqueta	Hom 5p	Hom 3p	Cebador Inverso
X	X-7	GCCCTCATCTTC TTCCCTGCGTTC TCACCAACCCTCA CCAAGACATACA TGGCTTTGGCAG (SEQ ID NO: 15)	GTGAACCAGTCC GAGTAAAAGCCC TATTGCAAGCCCT CTT (SEQ ID NO: 31)	GCCCTCAT CTTCTTCC CTGC (SEQ ID NO: 33)	GTTCTCA CCACCCCT CACCAA (SEQ ID NO: 34)	GACATAC ATGGCTT TGGCAG (SEQ ID NO: 49)	GTGAAC CAGTCC GAGTGA AA (SEQ ID NO: 65)	GCCCTATT GCAAAGCCC TCCTT (SEQ ID NO: 68)
X	X-8	GCCCTCATCTTC TTCCCTGCGTTC TCACCAACCCTCA CCAAGAGATACT GCCACTTATGCA CG (SEQ ID NO: 16)	GCAAATGATGTTT AGCACCAAGCCC TATTGCAAGCCCT CTT (SEQ ID NO: 32)	GCCCTCAT CTTCTTCC CTGC (SEQ ID NO: 33)	GTTCTCA CCACCCCT CACCAA (SEQ ID NO: 34)	GAGATAC TGCCACT TATGCAC G (SEQ ID NO: 50)	GCAAAT GATGTTT AGCACCC AC (SEQ ID NO: 66)	GCCCTATT GCAAAGCCC TCCTT (SEQ ID NO: 68)

5 Los conjuntos de sondas y cebadores ilustrativos que se pueden utilizar en el método descrito en la presente memoria para detectar un polimorfismo en un sitio SNP se enumeran en la Tabla 5 a continuación. Cada uno de los conjuntos de sondas ilustrativos en la Tabla 5 comprende tres sondas, dos sondas específicas del alelo (que se utilizan para el marcaje) y una sonda de etiquetado. En estos ejemplos, las dos sondas específicas de alelos tienen secuencias de homología que son diferentes en uno o más nucleótidos. La estructura de la primera sonda alélica incluye un Alelo 1 del Sitio de Cebador Directo y un Alelo 1 de Homología; y la estructura de la segunda sonda alélica incluye un Alelo 2 del Sitio de Cebador Directo y un Alelo 2 de Homología. En la práctica, los cebadores marcados se pueden utilizar con diferentes marcas en los dos cebadores (por lo tanto, las marcas son específicas del alelo). En estos ejemplos, también hay una sonda 3'universal que incluye una región de homología (sin ningún SNP), la secuencia de etiquetado y un sitio de cebador inverso. También se muestran las secuencias componentes de las sondas (etiqueta, secuencia de homología, etc.).

10

Tabla 5 - Sondas y cebadores ilustrativos

Cromosoma	Sonda Marcaje - Alelo 1 (Cebador Directo Alelo 1 + Hom 5p Alelo 1)	Sonda Marcaje - Alelo 2 (Cebador Directo Alelo 2 + Hom 5p Alelo 2)	Sonda Etiquetado (Hom 3p + Etiqueta + Cebador Inverso)	Cebador Directo - Alelo 1	Cebador Directo - Alelo 2	Hom 5p - Alelo 1	Hom 5p - Alelo 2	Hom 3p	Etiqueta	Cebador Inverso
chr21	TTCCTCCACC GAACGTGTCT AGACCAGCAC AACTTACTcg (SEQ ID NO: 69)	GCCCTATTGCA AGCCCTCTTAG ACCAGCACAAAC TTACTta (SEQ ID NO: 112)	CAC TTGACAA AGTICTCACG CGCCGAAGTT CTCCGAAGGA TGCCCTCATC TTCTTCCCCTGC (SEQ ID NO: 155)	TTCCTC CACCGA ACGTGT CT (SEQ ID NO: 67)	GCCCTAT TGCAAG CCCTCTT (SEQ ID NO: 68)	AGACCA GCACAA CTTACTc (SEQ ID NO: 198)	AGACC AGCAC AACTT ACTta (SEQ ID NO: 241)	CACTT GACA AAGTT CTCAC GC (SEQ ID NO: 284)	GCCGA AGTIC TCCGA AGGAT (SEQ ID NO: 327)	GCCCT CATCT TCTTC CCTGC (SEQ ID NO: 33)
chr3	TTCCTCCACC GAACGTGTCT CCAAATgCAC CTGCCcg (SEQ ID NO: 70)	GCCCTATTGCA AGCCCTCTTCCA AATcACCTGCCca (SEQ ID NO: 113)	CATTAGGGAT TAACGGCTTG GGACAGACTG ACGGAGCTTC AGCCCTCATC TTCTTCCCCTGC (SEQ ID NO: 156)	TTCCTC CACCGA ACGTGT CT (SEQ ID NO: 67)	GCCCTAT TGCAAG CCCTCTT (SEQ ID NO: 68)	CCAAAT gCACCT GCCcg (SEQ ID NO: 199)	CCAAA TtCAC TGCCca (SEQ ID NO: 242)	CATTA GGGAT TAACG GCTTG G (SEQ ID NO: 285)	GACA GACTG ACGG AGCTT CA (SEQ ID NO: 328)	GCCCT CATCT TCTTC CCTGC (SEQ ID NO: 33)
chr13	TTCCTCCACC GAACGTGTCT AGTTTGGACA AAGGCaATTcg (SEQ ID NO: 71)	GCCCTATTGCA AGCCCTCTTAGT TTGGACAAAAGG CgATTta (SEQ ID NO: 114)	CACACGTTAA GAAGACTTTC TGCTGACTCT GCCGCACATG ATCGCCCTCA TCTTCTTCCCCT GC (SEQ ID NO: 157)	TTCCTC CACCGA ACGTGT CT (SEQ ID NO: 67)	GCCCTAT TGCAAG CCCTCTT (SEQ ID NO: 68)	AGTTTG GACAAA GGCaAT Tcg (SEQ ID NO: 200)	AGTTT GGACA AAGGC gATTta (SEQ ID NO: 243)	CACAC GTTAA GAAG ACTTT CTGC (SEQ ID NO: 286)	TGACT CTGCC GCACA TGATC (SEQ ID NO: 329)	GCCCT CATCT TCTTC CCTGC (SEQ ID NO: 33)
chr3	TTCCTCCACC GAACGTGTCT TGAGCTTAGC CAATATCAAG AAGg (SEQ ID NO: 72)	GCCCTATTGCA AGCCCTCTTGA GCTTAGCCAAT ATCAAcAAGa (SEQ ID NO: 115)	CTAAGTGCC TCCATGAGAA AGGATCCGAT AGCCCTCTGC AGGCCCTCAT CTTCTTCCCCTG (SEQ ID NO: 158)	TTCCTC CACCGA ACGTGT CT (SEQ ID NO: 67)	GCCCTAT TGCAAG CCCTCTT (SEQ ID NO: 68)	TGAGCT TAGCCA ATATCA AgAAGg (SEQ ID NO: 201)	TGAGC TTAGC CAATA TCAAcA AGa (SEQ ID NO: 244)	CTAAG TGCCC TCCAT GAGA AAG (SEQ ID NO: 287)	GATCC GATAG CCCTC TGcAG (SEQ ID NO: 330)	GCCCT CATCT TCTTC CCTGC (SEQ ID NO: 33)

chr9	TTCCTCCACC GAACGTGTCT ACGTGAACCTT TCCTTGGTAcA c (SEQ ID NO: 73)	GCCCTATTGCA AGCCCTCTTAC GTGAACTTTCCCT TGGTAaAt (SEQ ID NO: 116)	C (SEQ ID NO: 158)	TTCCTC CACCGA ACGTGT CT (SEQ ID NO: 67)	GCCCTAT TGCAAG CCCTCTT (SEQ ID NO: 68)	ACGTGA ACTTTC CTTGGT AcAc (SEQ ID NO: 202)	NO: 244)	ID NO: 287)	330)	33)
chr3	TTCCTCCACC GAACGTGTCT TGAAGATGTT CTAATACCTT GCcg (SEQ ID NO: 74)	GCCCTATTGCA AGCCCTCTTGA AGATGTTCTAA TACCTTGCta (SEQ ID NO: 117)	C (SEQ ID NO: 159)	TTCCTC CACCGA ACGTGT CT (SEQ ID NO: 67)	GCCCTAT TGCAAG CCCTCTT (SEQ ID NO: 68)	TGAAGA TGTTCT AATACC TTGCcg (SEQ ID NO: 203)	TGAAG ATGTT CTAAT ACCTT GCta (SEQ ID NO: 246)	CITTAC AGGA GGTCT GGCAT CA (SEQ ID NO: 332)	GGTCA ACAAC CGAG GGACT C (SEQ ID NO: 333)	GCCCT CATCT TCCTC CCTGC (SEQ ID NO: 33)
chr17	TTCCTCCACC GAACGTGTCT CAGTGTGGAG ACtGAACg (SEQ ID NO: 75)	GCCCTATTGCA AGCCCTCTTCA GTGTGGAGAc GAACa (SEQ ID NO: 118)	C (SEQ ID NO: 161)	TTCCTC CACCGA ACGTGT CT (SEQ ID NO: 67)	GCCCTAT TGCAAG CCCTCTT (SEQ ID NO: 68)	CAGTGT GGAGAC tGAACg (SEQ ID NO: 204)	CAGTG TGGAG ACcGA ACa (SEQ ID NO: 247)	CCACA ATGAG AAGG CAGA G (SEQ ID NO: 290)	TTGTC ATTAA TGCTG GCcGC (SEQ ID NO: 333)	GCCCT CATCT TCCTC CCTGC (SEQ ID NO: 33)
chr16	TTCCTCCACC GAACGTGTCT AGGCAGGGTA ATGTCATGAA aTg (SEQ ID NO: 76)	GCCCTATTGCA AGCCCTCTTAG GCAGGGTAATG TCATGAAGt (SEQ ID NO: 119)	C (SEQ ID NO: 162)	TTCCTC CACCGA ACGTGT CT (SEQ ID NO: 67)	GCCCTAT TGCAAG CCCTCTT (SEQ ID NO: 68)	AGGCAG GGTAAT GTCATG AAaTg (SEQ ID NO: 205)	AGGCA GGGTA ATGTC ATGAA gTt (SEQ ID NO: 248)	GCTGT GGCAT AGCTA CACTC (SEQ ID NO: 291)	CGGTG ACGGT TTGCA ACTTT (SEQ ID NO: 334)	GCCCT CATCT TCCTC CCTGC (SEQ ID NO: 33)
chr21	TTCCTCCACC GAACGTGTCT	GCCCTATTGCA AGCCCTCTTGAT	C (SEQ ID NO: 162)	TTCCTC CACCGA	GCCCTAT TGCAAG	GATTGT CTGGAG	GATTG TCTGG	CAGG GTAAT	GTCCG GCAGT	GCCCT CATCT

	GATTGTCTGG AGcGCTg (SEQ ID NO: 77)	TGTCTGGAGgGC Tc (SEQ ID NO: 120)	GTCCGGCAGT TAAGGGTCTC GCCCTCATCT TCTTCCCTGC (SEQ ID NO: 163)	ACGTGT CT (SEQ ID NO: 67)	CCCTCTT (SEQ ID NO: 68)	cGCTg (SEQ ID NO: 206)	AGgGC Tc (SEQ ID NO: 249)	TTGTG GGTCT G (SEQ ID NO: 292)	TAAGG GTCTC (SEQ ID NO: 335)	TCTTC CCTGC (SEQ ID NO: 33)
chr2	TTCCTCCACC GAACGTGTCT AGGAGCAAT AGGcG (SEQ ID NO: 78)	GCCCTATTGCA AGCCCTCTTAG GGAGCAATAGG Cta (SEQ ID NO: 121)	GGGCTATCCA GAAAGATAAG AATACTCAC AACGACTGCG CAGCCCTCAT CTTCTCCCTG C (SEQ ID NO: 164)	TTCCTC CACCGA ACGTGT CT (SEQ ID NO: 67)	GCCCTAT TGCAAG CCCTCTT (SEQ ID NO: 68)	AGGGAG CAATAG GCcg (SEQ ID NO: 207)	AGGGA GCAAT AGGcA (SEQ ID NO: 250)	GGGCT ATCCA GAA GATAA GAA (SEQ ID NO: 293)	TACTC ACAA ACGAC TGCGC A (SEQ ID NO: 336)	GCCCT CATCT TCTTC CCTGC (SEQ ID NO: 33)
chr2	TTCCTCCACC GAACGTGTCT CTGCAGGGTA CAAcAcg (SEQ ID NO: 79)	GCCCTATTGCA AGCCCTCTTCTG CAGGGTACAAg ACa (SEQ ID NO: 122)	CATAAAGTGT GGAGTATTC ACTCGTATAT GGCCGACTGG AGGGCCCTCA TCTTCTCCCT GC (SEQ ID NO: 165)	TTCCTC CACCGA ACGTGT CT (SEQ ID NO: 67)	GCCCTAT TGCAAG CCCTCTT (SEQ ID NO: 68)	CTGCAG GGTACA AcAcg (SEQ ID NO: 208)	CTGCA GGTA CAAgA Ca (SEQ ID NO: 251)	CATAA CTGGT GGAGT ATTTC ACT (SEQ ID NO: 294)	CGTAT ATGGC CGACT GGAG G (SEQ ID NO: 337)	GCCCT CATCT TCTTC CCTGC (SEQ ID NO: 33)
chr19	TTCCTCCACC GAACGTGTCT CGTATCTGGG AAGAcGc (SEQ ID NO: 80)	GCCCTATTGCA AGCCCTCTTCTG ATCTGGGAAGAT GGg (SEQ ID NO: 123)	CTTCAAGGAA GAAATTCAAC AGGGTAGGGT TTGCCGGGAT AAGGGCCCTC ATCTTCTCCC TGC (SEQ ID NO: 166)	TTCCTC CACCGA ACGTGT CT (SEQ ID NO: 67)	GCCCTAT TGCAAG CCCTCTT (SEQ ID NO: 68)	CGTATC TGGGA GAcGc (SEQ ID NO: 209)	CGTAT CTGGG AAGAIG Gg (SEQ ID NO: 252)	CTTCA AGGA AGAA ATTCA ACAG GG (SEQ ID NO: 295)	TAGGG TTTTGC GGCG ATAAG G (SEQ ID NO: 338)	GCCCT CATCT TCTTC CCTGC (SEQ ID NO: 33)
chr9	TTCCTCCACC GAACGTGTCT CCTGTAATCC CTTGCAAJgc (SEQ ID NO: 81)	GCCCTATTGCA AGCCCTCTTCT GTAATCCCTTGC AAJaa (SEQ ID NO: 124)	CATGGATTCA ACACAGCAA CACCAAGTCA ACCACCCGAG ACGCCCTCAT	TTCCTC CACCGA ACGTGT CT (SEQ ID NO: 67)	GCCCTAT TGCAAG CCCTCTT (SEQ ID NO: 68)	CCTGTA ATCCCT TGCAAT gc (SEQ ID NO: 209)	CCTGT AATCC CTTGC AAJaa (SEQ ID NO: 252)	CATGG ATTCA ACACA GCAA ACA	CCAAG TCAAC CACCC GAGA C (SEQ ID NO: 338)	GCCCT CATCT TCTTC CCTGC (SEQ ID NO: 33)

	TGCTTTAGAA TTTGGCAACTg Gc (SEQ ID NO: 86)	CTTAGAATTTG GCAACTaGt (SEQ ID NO: 129)	CCCCCACCCT AGCACTCCTT CTTGGCCCTCA TCTTCTCCCT Gc (SEQ ID NO: 172)	ACGTGT CT (SEQ ID NO: 67)	CCCTCTT (SEQ ID NO: 68)	TGGCAA CTgGc (SEQ ID NO: 215)	TTGGC AACTaG t (SEQ ID NO: 258)	ACTCC ACAG AACCC (SEQ ID NO: 301)	ACTCC TTCTT (SEQ ID NO: 344)	TCTTC CCTGC (SEQ ID NO: 33)
chr7	TTCCTCCACC GAACGTGTCT GCAGGAAAGC CTAcTGAac (SEQ ID NO: 87)	GCCCTATTGCA AGCCCTTTGC AGGAAAGCCTAT TGAAt (SEQ ID NO: 130)	GTGCAGAGGA CAGGAAGAAC GGAGCGTCGG TAGTGIAAAG CCCTCATCTTC TTCCTGC (SEQ ID NO: 173)	TTCCTC CACCGA ACGTGT CT (SEQ ID NO: 67)	GCCCTAT TGCAAG CCCTCTT (SEQ ID NO: 68)	GCAGGA AAGCCT AcTGAA c (SEQ ID NO: 216)	GCAGG AAAGC CTATG AAt (SEQ ID NO: 259)	GTGCA GAGG ACAG GAAG AA (SEQ ID NO: 302)	CGGA GCGTC GGTAG TGTA A (SEQ ID NO: 345)	GCCCT CATCT TCTTC CCTGC (SEQ ID NO: 33)
chr3	TTCCTCCACC GAACGTGTCT GGGAGCCAGA GAAATgTcc (SEQ ID NO: 88)	GCCCTATTGCA AGCCCTTTGG GAGCCAGAGAA ATtTc (SEQ ID NO: 131)	GGTGCTCAA GACATACACC TTAAACAACCTC GACGAACCTA CCGGCCCTCA TCTTCTCCCT GC (SEQ ID NO: 174)	TTCCTC CACCGA ACGTGT CT (SEQ ID NO: 67)	GCCCTAT TGCAAG CCCTCTT (SEQ ID NO: 68)	GGGAGC CAGAGA AATgTcc (SEQ ID NO: 217)	GGGAG CCAGA GAAAt Tct (SEQ ID NO: 260)	GGTGC TTCAA GACAT ACACC TTA (SEQ ID NO: 303)	ACAAC TCGAC GAACC TACCG (SEQ ID NO: 346)	GCCCT CATCT TCTTC CCTGC (SEQ ID NO: 33)
chr2	TTCCTCCACC GAACGTGTCT TGCTCCAGT TCCACTTCAt Tag (SEQ ID NO: 89)	GCCCTATTGCA AGCCCTTTGT CTCCAGTTCCAC TTCATgTaa (SEQ ID NO: 132)	GGAACCTCTG TGACCTTGGGA TGGCCCATCC TTATGTGCTG GCCCTCATCT TCTTCCCTGC (SEQ ID NO: 175)	TTCCTC CACCGA ACGTGT CT (SEQ ID NO: 67)	GCCCTAT TGCAAG CCCTCTT (SEQ ID NO: 68)	TGTCTC CAGTT CCACT TCATgT Aa (SEQ ID NO: 261)	GGAA CCTCT GTGAC CTTGG A (SEQ ID NO: 304)	TGGCC CATCC TTATG TGCTG (SEQ ID NO: 347)	GCCCT CATCT TCTTC CCTGC (SEQ ID NO: 33)	
chr15	TTCCTCCACC GAACGTGTCT CCCGTAAAT GCCTAcTeg (SEQ ID NO: 90)	GCCCTATTGCA AGCCCTTTCCC GTTAATTGCCTA tTta (SEQ ID NO: 133)	CCAGTGGTA CCTTCTGAAG GTCGTTATTG CTCAAGCCCG CCCTCATCTTC TTCCTGC	TTCCTC CACCGA ACGTGT CT (SEQ ID NO: 67)	GCCCTAT TGCAAG CCCTCTT (SEQ ID NO: 68)	CCCGTT AATTGC CTAcTeg (SEQ ID NO: 219)	CCCCT TAATT GCCTAt Tta (SEQ ID NO: 262)	CCCAG TGGTA CCTTC TGAA (SEQ ID NO: 305)	GGTGC TTATT GCCTA AGCCC (SEQ ID NO: 348)	GCCCT CATCT TCTTC CCTGC (SEQ ID NO: 33)

chr15	TTCCTCCACC GAACGTGTCT CTCGGTCCTCA CTGGaAaG (SEQ ID NO: 91)	GCCCTATTGCA AGCCCTCTTCTC GGTCCACACTGGg AAa (SEQ ID NO: 134)	(SEQ ID NO: 176) CTTCTGTGCT TATTTGGGTA ACTTGATICT GGCCCTCCCA TCGCCCTCAT CTTCTTCCCTG C (SEQ ID NO: 177)	TTCCTC CACCGA ACGTGT CT (SEQ ID NO: 67)	GCCCTAT TGCAAG CCCTCTT (SEQ ID NO: 68)	CTCGGT CCCACT GGaAaG (SEQ ID NO: 220)	CTCGG TCCCA CTGGg AAa (SEQ ID NO: 263)	305) CTTCT GTTGC TTATT TGGGT AAC (SEQ ID NO: 349)	348) TTGAT TCTGG CCCTC CCATC (SEQ ID NO: 349)	33) GCCCT CATCT TCCTC CCTGC (SEQ ID NO: 33)
chr2	TTCCTCCACC GAACGTGTCT ACACCCATGA TTCAGTTACfg (SEQ ID NO: 92)	GCCCTATTGCA AGCCCTCTTAC ACCCATGATTC AGTTACca (SEQ ID NO: 135)	CCCACCTGGAT GCCCTCCCTCA CGCCGGCTAT TTAGGTGCC TCATCTTCTC CCTGC (SEQ ID NO: 178)	TTCCTC CACCGA ACGTGT CT (SEQ ID NO: 67)	GCCCTAT TGCAAG CCCTCTT (SEQ ID NO: 68)	ACACCC ATGATT CAGTIA Ctg (SEQ ID NO: 221)	ACACC CATGA TTCAG TTACca (SEQ ID NO: 264)	306) CCCAC TGGAT GCCTC C (SEQ ID NO: 307)	CTCAC GCCGG CTATT TAGGT (SEQ ID NO: 350)	GCCCT CATCT TCCTC CCTGC (SEQ ID NO: 33)
chr9	TTCCTCCACC GAACGTGTCT GCTAGTATGA ACATCACAgGc (SEQ ID NO: 93)	GCCCTATTGCA AGCCCTCTTGT AGTATGAACAT CACaGt (SEQ ID NO: 136)	CGGAGAGAG CATTGAAAG TCTGGTAGG TGGAGGACGC CCTCATCTTCT TCCCTGC (SEQ ID NO: 179)	TTCCTC CACCGA ACGTGT CT (SEQ ID NO: 67)	GCCCTAT TGCAAG CCCTCTT (SEQ ID NO: 68)	GCTAGT ATGAAC ATCACA gGc (SEQ ID NO: 222)	GCTAG TATGA ACATC ACAaGt (SEQ ID NO: 265)	308) CGGA GAGA CGCAT CTGAA (SEQ ID NO: 308)	AGTCT GGGTA GGTGG AGGA C (SEQ ID NO: 351)	GCCCT CATCT TCCTC CCTGC (SEQ ID NO: 33)
chr7	TTCCTCCACC GAACGTGTCT ACAAATGAGT AAGAAGCGAG Tcg (SEQ ID NO: 94)	GCCCTATTGCA AGCCCTCTTAC AAATGAGTAAG AAGCGAGTta (SEQ ID NO: 137)	CAGGATTTCC AGCTTACAGG GCGACTGAGC CACATCCAAC TGCCCTCATC TTCTTCCCTGC (SEQ ID NO: 180)	TTCCTC CACCGA ACGTGT CT (SEQ ID NO: 67)	GCCCTAT TGCAAG CCCTCTT (SEQ ID NO: 68)	ACAAAT GAGTAA GAAGCG AGTcg (SEQ ID NO: 223)	ACAAA TGAGT AAGAA GCGAG Tta (SEQ ID NO: 266)	309) CAGG ATTTC CAGCT TACAG GG (SEQ ID NO: 309)	CGACT GAGCC ACATC CAACT (SEQ ID NO: 352)	GCCCT CATCT TCCTC CCTGC (SEQ ID NO: 33)
chr20	TTCCTCCACC GAACGTGTCT GATAAGGGTT GCTCTgGg	GCCCTATTGCA AGCCCTCTTGT AAGGGTTGCTC TaCa (SEQ ID NO: 138)	CTTGCAAGAT GTGCCCTTA GAGCCCTCAGC CGGAATTGAA	TTCCTC CACCGA ACGTGT CT (SEQ ID NO: 67)	GCCCTAT TGCAAG CCCTCTT (SEQ ID NO: 68)	GATAAG GGTTGC TCTgGg (SEQ ID NO: 223)	GATAA GGGTT GCTCTa Ca (SEQ ID NO: 266)	309) CTTGC AAGAT GTGCC TCTTA	GAGCC TCAGC CGGA ATTGA	GCCCT CATCT TCCTC CCTGC

	(SEQ ID NO: 95)	138)	GCCCTCATCT TCTTCCCTGC (SEQ ID NO: 181)	ID NO: 67)	NO: 68)	NO: 224)	ID NO: 267)	(SEQ ID NO: 310)	A (SEQ ID NO: 353)	(SEQ ID NO: 33)
chr20	TTCCTCCACC GAACGTGTCT CCATGCACCA GCTACcc (SEQ ID NO: 96)	GCCCTATTGCA AGCCCTCTTCCA TGCACCCAGCTA Cta (SEQ ID NO: 139)	GGGTGGTTTC TCTAAACACA AATTGCCAAT CTGCACCAAT GGCCCTCAT CTTCTTCCCTG C (SEQ ID NO: 182)	TTCCTC CACCGA ACGTGT CT (SEQ ID NO: 67)	GCCCTAT TGCAAG CCCTCTT (SEQ ID NO: 68)	CCAATG ACCAGC TACcc (SEQ ID NO: 225)	CCAATG CACCA GCTACT a (SEQ ID NO: 268)	GGGTG GTTTC TCTAA ACACA AA (SEQ ID NO: 311)	TIGCC ATTCT GCACC AATGC (SEQ ID NO: 354)	GCCCT CATCT TCTTC CCTGC (SEQ ID NO: 33)
chr1	TTCCTCCACC GAACGTGTCT AACTGTACCC TACTCCCAgc (SEQ ID NO: 97)	GCCCTATTGCA AGCCCTCTTAA CTGTACCCTACT CCCaat (SEQ ID NO: 140)	GCAAGGTATT GAGAGAAGG ATCTAATGGT GTTCCGGGCT GATGCCCTCA TCTTCTTCCCT GC (SEQ ID NO: 183)	TTCCTC CACCGA ACGTGT CT (SEQ ID NO: 67)	GCCCTAT TGCAAG CCCTCTT (SEQ ID NO: 68)	AACTGT ACCCTA CTCCCA gc (SEQ ID NO: 226)	AACTG TACCC TACTC CCAat (SEQ ID NO: 269)	GCAG GGTAT TGAGA GAAG GATC (SEQ ID NO: 312)	TATTG GTGTT CGCGG CTGAT (SEQ ID NO: 355)	GCCCT CATCT TCTTC CCTGC (SEQ ID NO: 33)
chr2	TTCCTCCACC GAACGTGTCT AGGACCAAGG GACCAGTtag (SEQ ID NO: 98)	GCCCTATTGCA AGCCCTCTTAG GACCAAGGGAC CAGTTcAc (SEQ ID NO: 141)	GTGCACATTT CTTGATGAAG GGATGGCGT AACAGGAGGA CTGCCCTCAT CTTCTTCCCTG C (SEQ ID NO: 184)	TTCCTC CACCGA ACGTGT CT (SEQ ID NO: 67)	GCCCTAT TGCAAG CCCTCTT (SEQ ID NO: 68)	AGGACC AAGGGA CCAGTtt Ag (SEQ ID NO: 227)	AGGAC CAAGG GACCA GTTcAc (SEQ ID NO: 270)	GTGCA CATTT CTTGA TGAAG GG (SEQ ID NO: 313)	ATGGG CGTAA CAGG AGGA CT (SEQ ID NO: 356)	GCCCT CATCT TCTTC CCTGC (SEQ ID NO: 33)
chr7	TTCCTCCACC GAACGTGTCT AGAGTTCCTC CAAGAAATTG cg (SEQ ID NO: 99)	GCCCTATTGCA AGCCCTCTTAG AGTTCCTCCAA GAAATTGta (SEQ ID NO: 142)	GAGCAATGCC TGTTTCATGA GAGGAATGGC CTACCTGCAT CAGCCCTCAT CTTCTTCCCTG C (SEQ ID NO: 185)	TTCCTC CACCGA ACGTGT CT (SEQ ID NO: 67)	GCCCTAT TGCAAG CCCTCTT (SEQ ID NO: 68)	AGAGTT CCTCCA AGAAAT TGcg (SEQ ID NO: 228)	AGAGT TCCTCC AAGAA ATTGta (SEQ ID NO: 271)	GAGC AATGC CTGTT TCATG AGA (SEQ ID NO: 314)	GGAAT GGCCT ACCTG CATCA (SEQ ID NO: 357)	GCCCT CATCT TCTTC CCTGC (SEQ ID NO: 33)

chr5	TTCCTCCACC GAACGTGTCT ACATTATACA GCATGCTGc TAic (SEQ ID NO: 100)	GCCCTATTGCA AGCCCTCTTAC ATTATACAGCA TGCTGGtAga (SEQ ID NO: 143)	GTAAACATTA TACAGCATGG TGGCCCCGTT GTTGTCAICG CAICGGCCCTC AICTTCTTCCC TGC (SEQ ID NO: 186)	TTCCTC CACCGA ACGTGT CT (SEQ ID NO: 67)	GCCCTAT TGCAAG CCCTCTT (SEQ ID NO: 68)	ACATTA TACAGC ATGCTG GeTAic (SEQ ID NO: 229)	ACATT ATACA GCATG CTGGt Aga (SEQ ID NO: 272)	GTAA CATT TACAG CATGG TGGC (SEQ ID NO: 315)	CCCCG TGTTG TCATC GCATC (SEQ ID NO: 358)	GCCCT CATCT TCCTC CCTGC (SEQ ID NO: 33)
	chr2	TTCCTCCACC GAACGTGTCT GAGGAAGAA AGTGAGgTTT Gc (SEQ ID NO: 101)	GCCCTATTGCA AGCCCTCTTGA GGAAGAAAGTG AGaTTTgT (SEQ ID NO: 144)	GCAGAACAIG TCCTGAAGCG TTCGATGCCG CCCATGAGTG CCCTCAICTTC TTCCCTGC (SEQ ID NO: 187)	TTCCTC CACCGA ACGTGT CT (SEQ ID NO: 67)	GCCCTAT TGCAAG CCCTCTT (SEQ ID NO: 68)	GAGGAA GAAAGT GAGgTT TGc (SEQ ID NO: 230)	GAGGA AGAAA GTGAG aTTTgT (SEQ ID NO: 273)	GCAG AACAT GTCCT GAAG C (SEQ ID NO: 316)	GTTCG ATGCG TCCCA TGAGT (SEQ ID NO: 359)
chr15	TTCCTCCACC GAACGTGTCT CTGAATTATG TGCTTACCAAG AGc (SEQ ID NO: 102)	GCCCTATTGCA AGCCCTCTTCTG AATTATGTGCTT ACCAgGAGt (SEQ ID NO: 145)	CAGCTTGTTT CCAAACCCAT CAACCCGCGT AGATGTTCTT GCCCTCATCT TCTTCCCTGC (SEQ ID NO: 188)	TTCCTC CACCGA ACGTGT CT (SEQ ID NO: 67)	GCCCTAT TGCAAG CCCTCTT (SEQ ID NO: 68)	CTGAAT TATGTG CTTACC AaGAGc (SEQ ID NO: 231)	CTGAA TTATGT GCCTA CCAgG AGt (SEQ ID NO: 274)	CAGCT TGTTT CCAAA CCCAT (SEQ ID NO: 317)	CAACC CGCGT AGATG TTCCT (SEQ ID NO: 360)	GCCCT CATCT TCCTC CCTGC (SEQ ID NO: 33)
	chr9	TTCCTCCACC GAACGTGTCT TGGGTTCTGA TAACCTTATC AAGc (SEQ ID NO: 103)	GCCCTATTGCA AGCCCTCTTGG GTTCTGATAAC CTTATCAAct (SEQ ID NO: 146)	CAAAGTGTG AAGTTGCTT CGCCAGCTCA AGAGTGTAGC CGCCCTCATC TTCCTCCCTGC (SEQ ID NO: 189)	TTCCTC CACCGA ACGTGT CT (SEQ ID NO: 67)	GCCCTAT TGCAAG CCCTCTT (SEQ ID NO: 68)	TGGGTT CTGATA ACCTTA TCAAGc (SEQ ID NO: 232)	TGGGT TCTGA TAACC TTATC AAct (SEQ ID NO: 275)	CAAA GTGTG GAAAGT TGCTT CC (SEQ ID NO: 318)	GCCAG CTCAA GAGTG TAGCC (SEQ ID NO: 361)
chr2	TTCCTCCACC GAACGTGTCT GGTTAGTCAA ACATGcTGc	GCCCTATTGCA AGCCCTCTTGGT TAGTCAAACAT GtGt (SEQ ID	GGTCGACTT GTCCATCCTT CTTGATCCIG CGCGAIGTGC	TTCCTC CACCGA ACGTGT CT (SEQ	GCCCTAT TGCAAG CCCTCTT (SEQ ID	GGTTAG TCAAAC ATGcTGc (SEQ ID	GGTTA GTCAA ACATGt TgT	GGTCG ACTTT GTCCA TCC	TTCCT GATCC TGCGC GATGT	GCCCT CATCT TCCTC CCTGC

	(SEQ ID NO: 104)	NO: 147)	CCTCATCTTCT TCCCTGC (SEQ ID NO: 190)	ID NO: 67)	NO: 68)	NO: 233)	(SEQ ID NO: 276)	(SEQ ID NO: 319)	(SEQ ID NO: 362)	(SEQ ID NO: 33)
chr17	TTCCTCCACC GAACGTGTCT GACACTGGCA GAATCAAATC Ac (SEQ ID NO: 105)	GCCCTATTGCA AGCCCTCTTGA CACTGGCAGAA TCAAACcAa (SEQ ID NO: 148)	CTCTGTTGCCT GTGACTCAT CGCAGGCGTT CCCTATACGC CCTCATCTTCT TCCCTGC (SEQ ID NO: 191)	TTCCTC CACCGA ACGTGT CT (SEQ ID NO: 67)	GCCCTAT TGCAAG CCCTCTT (SEQ ID NO: 68)	GACACT GGCAGA ATCAAA tAc (SEQ ID NO: 234)	GACAC TGGCA GAATC AAAcC Aa (SEQ ID NO: 277)	CTCTG TTGCC TGTGG ACTC (SEQ ID NO: 320)	ATCGC AGGC GTTCC CTATA C (SEQ ID NO: 363)	GCCCT CATCT TCCTC CCTGC (SEQ ID NO: 33)
chr6	TTCCTCCACC GAACGTGTCT AGAGTTACAC CTTAGCTAA CcAc (SEQ ID NO: 106)	GCCCTATTGCA AGCCCTCTTAG AGTTACACCTTT AGCTAACtAg (SEQ ID NO: 149)	CTAACTAGAA TTAGTCTGCC TGCCATATGG ACCTCCGACC ACGAGCCCTC ATCTTCTTCCC TGC (SEQ ID NO: 192)	TTCCTC CACCGA ACGTGT CT (SEQ ID NO: 67)	GCCCTAT TGCAAG CCCTCTT (SEQ ID NO: 68)	AGAGTT ACACCT TTAGCT AACcAc (SEQ ID NO: 235)	AGAGT TACAC CTTTA GCTAA CtAg (SEQ ID NO: 278)	CTAAC TAGAA TTAGT CTGCC TGCC (SEQ ID NO: 321)	TATIG GACCT CCGAC CACGA (SEQ ID NO: 364)	GCCCT CATCT TCCTC CCTGC (SEQ ID NO: 33)
chr7	TTCCTCCACC GAACGTGTCT CCAGGAGTTC AAGaAGCg (SEQ ID NO: 107)	GCCCTATTGCA AGCCCTCTTCCA GGAGTTCAAAGg AGCa (SEQ ID NO: 150)	GTGAGCCATA ATCGTGICAA GCCACCATTT AGATCCGCGG CCCTCATCTTC TTCCCTGC (SEQ ID NO: 193)	TTCCTC CACCGA ACGTGT CT (SEQ ID NO: 67)	GCCCTAT TGCAAG CCCTCTT (SEQ ID NO: 68)	CCAGGA GTTCAA GaAGCg (SEQ ID NO: 236)	CCAGG AGTTC AAGgA GCa (SEQ ID NO: 279)	GTGAG CCATA ATCGT GTCA (SEQ ID NO: 322)	AGCCA CCATT TAGAT CCGCG (SEQ ID NO: 365)	GCCCT CATCT TCCTC CCTGC (SEQ ID NO: 33)
chr4	TTCCTCCACC GAACGTGTCT ACCACTCCTT TCTCCCaTCTc (SEQ ID NO: 108)	GCCCTATTGCA AGCCCTCTTACC ACTCCTTCTCC CgTCTt (SEQ ID NO: 151)	GAGAATTAAT GTCCCTCTC CTGGACCAGT AGAAGTCTGC CCGGCCCTCA TCTTCTTCCCT GC (SEQ ID NO: 194)	TTCCTC CACCGA ACGTGT CT (SEQ ID NO: 67)	GCCCTAT TGCAAG CCCTCTT (SEQ ID NO: 68)	ACCACT CCTTTC TCCCAtC Tc (SEQ ID NO: 237)	ACCAC TCCTTT CTCCCCg TCTt (SEQ ID NO: 280)	GAGA ATTAA TGCTC CCTCT CCTG (SEQ ID NO: 323)	GACCA GTAGA AGTCT GCCCG (SEQ ID NO: 366)	GCCCT CATCT TCCTC CCTGC (SEQ ID NO: 33)
chr2	TTCCTCCACC GAACGTGTCT	GCCCTATTGCA AGCCCTCTTGTc	GTGGTCTGCT GTTGACCAAT	TTCCTC CACCGA	GCCCTAT TGCAAG	GTCTT TGGGAC	GTCTT ATGGG	GTGGT CTGCT	TTTCA GAATG	GCCCT CATCT

	GTCTTATGGG ACAAATGGTg ATAg (SEQ ID NO: 109)	TTATGGGACAA TGGTcGATAt (SEQ ID NO: 152)	TTCAGAATGG CCGAGCTGTG CCCTCATCTTC TTCCCTGC (SEQ ID NO: 195)	ACGTGT CT (SEQ ID NO: 67)	CCCTCTT (SEQ ID NO: 68)	AATGGT tGATAg (SEQ ID NO: 238)	ACAAT GGTcG ATAt (SEQ ID NO: 281)	GTTGA CCAA (SEQ ID NO: 324)	GCCGA GCTGT (SEQ ID NO: 367)	TCTTC CCTGC (SEQ ID NO: 33)
chr17	TTCCTCCACC GAACGTGTCT CTACCCCTCAA CCCTCgTc (SEQ ID NO: 110)	GCCCTATTGCA AGCCCTCTTCTA CCCTCAACCCTC aTt (SEQ ID NO: 153)	GGTTGCAACT GCTGATCTAT AGGTGACCTT CTTGTAAGCC GCCCTCATCT TCTTCCCTGC (SEQ ID NO: 196)	TTCCCTC CACCGA ACGTGT CT (SEQ ID NO: 67)	GCCCTAT TGCAAG CCCTCTT (SEQ ID NO: 68)	C'TACCC TCAAACC CTCgTc (SEQ ID NO: 239)	CTACC CTCAA CCCTCa Tt (SEQ ID NO: 282)	GGTTG CAACT GCTGA TCTAT (SEQ ID NO: 325)	AGGTG ACCTT CTTGT ACGCC (SEQ ID NO: 368)	GCCCT CATCT TCTTC CCTGC (SEQ ID NO: 33)
chr7	TTCCTCCACC GAACGTGTCT CCAAGACTGA TCATGCcg (SEQ ID NO: 111)	GCCCTATTGCA AGCCCTCTTCCA AGACTGATCAT GCta (SEQ ID NO: 154)	CTTCCCAGT CAAGGCAGGG CGCGTCCCTTA TTTCCATCGC CCTCATCTTCT TCCCTGC (SEQ ID NO: 197)	TTCCCTC CACCGA ACGTGT CT (SEQ ID NO: 67)	GCCCTAT TGCAAG CCCTCTT (SEQ ID NO: 68)	CCAAGA CTGATC ATGCcg (SEQ ID NO: 240)	CCAAG ACTGA TCATG Cta (SEQ ID NO: 283)	CTTTC CCAGT CAAG GCAG (SEQ ID NO: 326)	GGCGC GTCCT TATTT CCATC (SEQ ID NO: 369)	GCCCT CATCT TCTTC CCTGC (SEQ ID NO: 33)

EJEMPLOS

El siguiente protocolo describe el procesamiento de hasta 24 muestras de ADN sin células a través de la hibridación-ligación, purificación, amplificación, preparación de dianas de micromatrices, hibridación de micromatrices y lavado de micromatrices.

- 5 Se prepararon u obtuvieron los siguientes materiales: ADN sin células (ADNsc) en un volumen de 20 µL de agua; Mezcla de Sonda: mezcla de todos los oligonucleótidos de la sonda de Etiquetado y Marcaje a una concentración 2 nM cada uno; Ligasa *Taq* (40 U/µL); Cuentas Magnéticas: MyOne Streptavidin C1 Dynabeads; Tampón de Unión y Lavado de las Cuentas, concentraciones 1X y 2X; Cebador de amplificación directo, modificado en fosfato 5' Cebador de amplificación inverso, marcado; Enzima AmpliTaq Gold (5 U/µL); Mezcla de dNTP; Exonucleasa Lambda (5 U/µL); Tampón de Hibridación, 1,25X; Oligonucleótidos de control de hibridación; Tampón de Lavado de la Micromatriz A; Tampón de Lavado de la Micromatriz B; Tampón de Lavado de la Micromatriz C.

Reacción de hibridación-ligación: Las muestras de ADNsc (20 µL) se añadieron a los pocillos A3-H3 de una placa de reacción de 96 pocillos. Se añadieron los siguientes reactivos a cada muestra de ADNsc para un volumen de reacción total de 50 µL y se mezclaron pipeteando arriba y abajo 5-8 veces.

Componente	Volumen
H ₂ O	19,33 µL
Mezcla de Sonda	5 µL
Tampón Ligasa <i>Taq</i> 10X	5 µL
Ligasa <i>Taq</i>	0,67 µL

15

La placa se colocó en un termociclador y se ligó utilizando el siguiente perfil de ciclos: (i) 95°C durante 5 minutos; (ii) 95°C durante 30 segundos; (iii) 45°C durante 25 minutos; (iv) Repetir las etapas b a c 4 veces; y (v) mantener a 4°C.

Purificación del producto de hibridación-ligación:

- 20 Lavado de Dynabeads: se sometió a agitación vorticial un frasco de Dynabeads en el ajuste más alto durante 30 segundos. Se transfirieron 260 µL de cuentas a un tubo de 1,5 mL. Se mezclaron 900 µL de Tampón de Unión y Lavado de Cuentas 2X y las cuentas de la mezcla pipeteando arriba y abajo 5-8 veces. El tubo se colocó en un soporte magnético durante 1 minuto, y el sobrenadante se descartó. Se retiró el tubo del soporte magnético y se resuspendieron las cuentas magnéticas lavadas en 900 µL de 2X Tampón de Unión y Lavado de las Cuentas pipeteando arriba y abajo 5-8 veces. El tubo se colocó en el soporte magnético durante 1 minuto y se descartó el sobrenadante. El tubo se retiró del soporte magnético y se añadieron 1.230 µL de 2X Tampón de Unión y Lavado de las cuentas. Las cuentas se resuspendieron pipeteando arriba y abajo 5-8 veces.

- 30 Inmovilización de productos HL: se transfirieron 50 µL de cuentas lavadas a cada producto de reacción de hibridación-ligación en la placa de reacción de 96 pocillos y la mezcla se pipeteó arriba y abajo 8 veces, se incubó durante 15 minutos a temperatura ambiente, se mezcló dos veces en un imán de placa durante el tiempo de incubación. Las cuentas se separaron con un imán de placa durante 3 minutos y después se retiraron y se descartó el sobrenadante. La placa se retiró del imán de la placa, se añadieron 200 µL de Tampón de Unión y Lavado de Cuentas 1X, y las cuentas se resuspendieron pipeteando arriba y abajo 5-8 veces. La placa se colocó en el imán de placa durante 1 minuto, y el sobrenadante se descartó. La placa se retiró del imán de placa, se añadieron 180 µL de 1X SSC y las cuentas se resuspendieron pipeteando arriba y abajo 5-8 veces. La placa se colocó sobre el imán de placa durante 1 minuto, y el sobrenadante se descartó.

- 35 Purificación de productos de Hib-Ligación: se añadieron 50 µL de NaOH 0,15 M recién preparado a cada pocillo y, las cuentas se resuspendieron pipeteando arriba y abajo 5-8 veces, y se incubaron a temperatura ambiente durante 10 minutos. La placa se colocó sobre el imán de placa durante 2 minutos y a continuación se retiró, y el sobrenadante se descartó. La placa se retiró del imán de la placa, se añadieron 200 µL de NaOH 0.1 M recién preparado y las cuentas se resuspendieron pipeteando arriba y abajo 5-8 veces. La placa se colocó sobre el imán de placa durante 1 minuto, y el sobrenadante se descartó. La placa se retiró del imán de la placa y se añadieron 180 µL de NaOH 0,1 M, y las cuentas se resuspendieron pipeteando arriba y abajo 5-8 veces. La placa se colocó sobre el imán de la placa durante 1 minuto, y el sobrenadante se descartó. La placa se retiró del imán de placa, se añadieron 200 µL de 1X Tampón de Unión y Lavado, y las cuentas se resuspendieron pipeteando arriba y abajo 5-8 veces. Se colocó la placa sobre el imán de placa durante 1 minuto y el sobrenadante se descartó. Se retiró la placa del imán de placa, se añadieron 180 µL de TE, y las cuentas se resuspendieron pipeteando arriba y abajo 5-8 veces. La placa se

45

colocó sobre el imán de placa durante 1 minuto, y el sobrenadante se descartó. Se añadieron 20 µL de agua a cada pocillo y las cuentas se resuspendieron pipeteando arriba y abajo 5-8 veces. La placa se selló y se almacenó a 4°C hasta que se utilizó en las etapas posteriores.

- 5 **Amplificación:** Se añadieron los siguientes reactivos a cada producto de reacción de hibridación-ligación en la placa de reacción de 96 pocillos para un volumen de reacción total de 50 µL.

Componente	Volumen
H2O	17,25 µL
Cebador Directo, 10 µM	2,5 µL
Cebador Inverso, 10 µM	2,5 µL
Mezcla de dNTP 4 mM (L/N 052114)	2,5 µL
10X Tampón AmpliTaq Gold	5 µL
Enzima AmpliTaq Gold	0,25 µL

La placa se colocó en un termociclador y las sondas se ligaron utilizando el siguiente perfil de ciclos: (i) 95°C durante 5 minutos; (ii) 95°C durante 30 segundos; (iii) 45°C durante 25 minutos; (iv) Repetir las etapas b a c 4 veces; y (v) mantener a 4°C.

- 10 Purificación de producto de hibridación-ligación: los reactivos se mezclaron pipeteando arriba y abajo 5-8 veces. La placa se colocó en un termociclador y las sondas se amplificaron utilizando el siguiente perfil de ciclos: (i) 95°C durante 5 minutos; (ii) 95°C durante 30 segundos; (iii) 54°C durante 30 segundos; (iv) 72°C durante 60 segundos, (v) Repetir las etapas b a d 29 veces; (vi) 72°C durante 5 minutos; (vii) Repetir las etapas b a c 4 veces; y (v) mantener a 4°C.

- 15 **Preparación de la diana de la micromatriz (digestión de una sola hebra):** se añadieron los siguientes reactivos a cada producto de reacción amplificado en la placa de reacción de 96 pocillos para un volumen de reacción total de 60 µL.

Componente	Volumen
H2O	3 µL
10X Tampón Exonucleasa Lambda	6 µL
Enzima Exonucleasa Lambda	1 µL

- 20 Los reactivos se mezclaron pipeteando arriba y abajo 5-8 veces. La placa se colocó en un termociclador y las sondas se digirieron utilizando el siguiente perfil de ciclos: (i) 37°C durante 60 minutos; (ii) 80°C durante 30 minutos; (iii) mantener a 4°C. La placa se colocó en Speed-vac y las muestras se secaron utilizando un ajuste de calor medio durante aproximadamente 60 minutos o hasta que todo el líquido se evaporó. Las muestras se almacenaron a 4°C en la oscuridad hasta que se utilizaron en las etapas posteriores.

- 25 **Hibridación de micromatrices:** se añadieron los siguientes reactivos a cada Diana de Micromatriz seca en la placa de reacción de 96 pocillos para un volumen de reacción total de 20 µL.

Componente	Volumen
H ₂ O	3 µL
1,25X Tampón de Hibridación	16 µL

ES 2 739 430 T3

Componente	Volumen
Oligonucleótidos de control de hibridación	1 μ L

Los reactivos se mezclaron pipeteando arriba y abajo 10-20 veces para que se resuspendieran y se centrifugaron brevemente para llevar el contenido al fondo de los pocillos de la placa. La placa se colocó en un termociclador y las sondas se desnaturalizaron utilizando el siguiente perfil de ciclos: (i) 70°C durante 3 minutos; (ii) mantener a 42°C. El código de barras de la micromatriz que se iba a utilizar se registró para cada muestra en la Hoja de Seguimiento. Se preparó una cámara de hibridación que contenía un Lifter Slip (cubreobjetos con el borde elevado recubierto de Teflón) para cada micromatriz que se iba a procesar. Para cada muestra, se añadieron 15 μ L de Diana de Micromatriz al centro de un Lifter Slip en una cámara de hibridación, y la micromatriz apropiada se colocó de inmediato en el fluido diana colocando el borde superior hacia abajo sobre el Lifter Slip y dejando que caiga lentamente. Las cámaras de hibridación se cerraron y se incubaron a 42°C durante 60 minutos. Las cámaras de hibridación se abrieron y cada micromatriz se retiró de los Lifter Slips y se colocó en una gradilla sumergida en el Tampón de Lavado de Micromatriz A. Una vez que todas las micromatrices estaban en la gradilla, la gradilla se agitó a 650 rpm durante 5 minutos. La gradilla de micromatrices se retiró del Tampón de Lavado de Micromatriz A, se eliminó el exceso de líquido en una toallita limpia y la gradilla se colocó rápidamente en el Tampón de Lavado de Micromatriz B. La gradilla se agitó a 650 rpm durante 5 minutos. La gradilla de micromatrices se retiró del Tampón de Lavado de Micromatriz B, el exceso de líquido se extrajo con una toallita limpia y se colocó rápidamente en el Tampón de Lavado de Micromatriz C. La gradilla se agitó a 650 rpm durante 5 minutos. Inmediatamente después de completar el lavado de 5 minutos en el Tampón de Lavado de Micromatriz C, la gradilla de micromatrices se retiró lentamente del tampón. Esto llevó de 5 a 10 segundos para maximizar la formación de láminas del tampón de lavado desde la superficie del cubreobjetos. El exceso de líquido se eliminó en una toallita limpia. Se utilizó un aspirador de vacío para eliminar las gotas de tampón restantes presentes en cualquiera de las superficies de cada micromatriz. Las micromatrices se almacenaron en una gradilla para portaobjetos en nitrógeno y en la oscuridad hasta que se analizaron las micromatrices.

LISTADO DE SECUENCIAS

<110> SINGULAR BIO, INC. FEHR, Adrian Nielsen COLLINS, Patrick James HERSCHLEB, Jill Lyndon JONES, Hywel Bowden

5 <120> ENSAYOS PARA LA DETECCIÓN DE UNA SOLA MOLÉCULA Y USO DE LOS MISMOS

<130> 31753-5003WO

10 <150> US 61/867,554
<151> 19-08-2013

<150> US 61/867,559
<151> 19-08-2013

15 <160> 369

<170> PatentIn versión 3.5

20 <210> 1
<211> 60
<212> ADN
<213> Secuencia Artificial

25 <220>
<223> Secuencia sintética

<400> 1
gcctcatct tcttcctgc gttctacca cctcaccaa ggaagaagt agggcttctc 60

30 <210> 2
<211> 60
<212> ADN
<213> Secuencia Artificial

35 <220>
<223> Secuencia sintética

<400> 2
gcctcatct tcttcctgc gttctacca cctcaccaa aaatcaagt gaccagctcc 60

40 <210> 3
<211> 62
<212> ADN
<213> Secuencia Artificial

45 <220>
<223> Secuencia sintética

<400> 3
gccctcatct tcttcctgc gttctacca cctcaccaa tcatctgcca agacagaagt 60

50 tc 62

<210> 4
<211> 60
<212> ADN
<213> Secuencia Artificial

55 <220>
<223> Secuencia sintética

60 <400> 4
gcctcatct tcttcctgc gttctacca cctcaccaa gcaggagagt caaaggtctg 60

ES 2 739 430 T3

<210> 5
 <211> 60
 <212> ADN
 <213> Secuencia Artificial
 5
 <220>
 <223> Secuencia sintética
 10
 <400> 5
 gcctcatct tctccctgc gttctacca cctcacaa gttgcatgg agattgtgc 60
 <210> 6
 <211> 60
 <212> ADN
 <213> Secuencia Artificial
 15
 <220>
 <223> Secuencia sintética
 20
 <400> 6
 gcctcatct tctccctgc gttctacca cctcacaa cagctcagtg atgcattgc 60
 <210> 7
 <211> 61
 <212> ADN
 <213> Secuencia Artificial
 25
 <220>
 <223> Secuencia sintética
 30
 <400> 7
 gcctcatct tctccctgc gttctacca cctcacaa ccttgacctc tgctaattg 60
 g 61
 35
 <210> 8
 <211> 60
 <212> ADN
 <213> Secuencia Artificial
 40
 <220>
 <223> Secuencia sintética
 <400> 8
 gcctcatct tctccctgc gttctacca cctcacaa cacctgtcca acagctacag 60
 45
 <210> 9
 <211> 61
 <212> ADN
 <213> Secuencia Artificial
 50
 <220>
 <223> Secuencia sintética
 <400> 9
 gcctcatct tctccctgc gttctacca cctcacaa agaattgatc ttcaggcctg 60
 55
 c 61
 <210> 10
 <211> 60
 <212> ADN
 <213> Secuencia Artificial
 60
 <220>

ES 2 739 430 T3

	<223> Secuencia sintética	
5	<400> 10 gcctcatct tcttcctgc gttctacca cctcacca aagtaatcac tctgggtg	60
	<210> 11 <211> 60 <212> ADN <213> Secuencia Artificial	
10	<220> <223> Secuencia sintética	
15	<400> 11 gcctcatct tcttcctgc gttctacca cctcacca agctcacaga caacctgtg	60
	<210> 12 <211> 60 <212> ADN <213> Secuencia Artificial	
20	<220> <223> Secuencia sintética	
25	<400> 12 gcctcatct tcttcctgc gttctacca cctcacca gcaatagaca cctacaggcg	60
	<210> 13 <211> 60 <212> ADN <213> Secuencia Artificial	
30	<220> <223> Secuencia sintética	
35	<400> 13 gcctcatct tcttcctgc gttctacca cctcacca gcacattatc aaaggccacg	60
	<210> 14 <211> 60 <212> ADN <213> Secuencia Artificial	
40	<220> <223> Secuencia sintética	
45	<400> 14 gcctcatct tcttcctgc gttctacca cctcacca caacgacctc aagcatgtg	60
	<210> 15 <211> 60 <212> ADN <213> Secuencia Artificial	
50	<220> <223> Secuencia sintética	
55	<400> 15 gcctcatct tcttcctgc gttctacca cctcacca gacatacatg gcttggcag	60
	<210> 16 <211> 62 <212> ADN <213> Secuencia Artificial	
60	<220>	
65		

ES 2 739 430 T3

	<223> Secuencia sintética	
	<400> 16	
	gccctcatct tcttccctgc gttctcacca ccctcaccaa gagatactgc cacttatgca	60
	cg	62
5	<210> 17	
	<211> 40	
	<212> ADN	
10	<213> Secuencia Artificial	
	<220>	
	<223> Secuencia sintética	
	<400> 17	
15	cgctgtaata gtctcagggc ttctccacc gaacgtgtct	40
	<210> 18	
	<211> 40	
	<212> ADN	
20	<213> Secuencia Artificial	
	<220>	
	<223> Secuencia sintética	
	<400> 18	
25	cgagcctca ttgcttcatt ttctccacc gaacgtgtct	40
	<210> 19	
	<211> 40	
30	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
	<223> Secuencia sintética	
35	<400> 19	
	cttgcgcaa acaattgtcc ttctccacc gaacgtgtct	40
	<210> 20	
40	<211> 40	
	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
45	<223> Secuencia sintética	
	<400> 20	
	gctgcagagt ttgattcat ttctccacc gaacgtgtct	40
50	<210> 21	
	<211> 41	
	<212> ADN	
	<213> Secuencia Artificial	
55	<220>	
	<223> Secuencia sintética	
	<400> 21	
60	catacacaca gaccgagagt cttctccac cgaacgtgtc t	41
	<210> 22	
	<211> 40	
	<212> ADN	
	<213> Secuencia Artificial	

ES 2 739 430 T3

<220>
<223> Secuencia sintética

5 <400> 22
ggatgfcagc cagcataagt ttctccacc gaacgtgtct 40

<210> 23
<211> 40
10 <212> ADN
<213> Secuencia Artificial

<220>
<223> Secuencia sintética

15 <400> 23
gcaagtgcc aacagttctc ttctccacc gaacgtgtct 40

<210> 24
<211> 41
20 <212> ADN
<213> Secuencia Artificial

<220>
<223> Secuencia sintética

25 <400> 24
gattccagca cactgagtc ttctccacc cgaacgtgtc t 41

<210> 25
<211> 40
30 <212> ADN
<213> Secuencia Artificial

<220>
<223> Secuencia sintética

35 <400> 25
ccgtgfcagg tttaaatgc gccctattgc aagccctctt 40

<210> 26
<211> 40
40 <212> ADN
<213> Secuencia Artificial

<220>
<223> Secuencia sintética

45 <400> 26
caagagtgtc ttatggcct gccctattgc aagccctctt 40

<210> 27
<211> 41
50 <212> ADN
<213> Secuencia Artificial

<220>
<223> Secuencia sintética

60 <400> 27
gcactcaagg agatcagact gccctattgc caagccctct t 41

<210> 28
<211> 41
65 <212> ADN
<213> Secuencia Artificial

<220>
 <223> Secuencia sintética

5 <400> 28
 ggctatcgaa ctacaaccac agccctattg caagccctct t 41

<210> 29
 <211> 41
 <212> ADN
 <213> Secuencia Artificial

10

<220>
 <223> Secuencia sintética

15 <400> 29
 gtagctgtct ggggtgat cgcctattg caagccctct t 41

<210> 30
 <211> 42
 <212> ADN
 <213> Secuencia Artificial

20

<220>
 <223> Secuencia sintética

25 <400> 30
 caagaaactt cgagccttag cagccctatt gcaagccctc tt 42

<210> 31
 <211> 40
 <212> ADN
 <213> Secuencia Artificial

30

<220>
 <223> Secuencia sintética

<400> 31
 gtgaaccagt ccgagtgaaa gccctattgc aagccctctt 40

40

<210> 32
 <211> 41
 <212> ADN
 <213> Secuencia Artificial

45

<220>
 <223> Secuencia sintética

<400> 32
 gcaaattgatg ttcagcacca cgcctattg caagccctct t 41

50

<210> 33
 <211> 20
 <212> ADN
 <213> Secuencia Artificial

55

<220>
 <223> Secuencia sintética

<400> 33
 gccctcatct tctccctgc 20

60

<210> 34
 <211> 20
 <212> ADN
 <213> Secuencia Artificial

65

<220>
 <223> Secuencia sintética

5 <400> 34
 gttctcacca ccctcacaa 20

<210> 35
 <211> 20
 <212> ADN
 <213> Secuencia Artificial

10

<220>
 <223> Secuencia sintética

15

<400> 35
 ggaagaagtg agggcttctc 20

<210> 36
 <211> 20
 <212> ADN
 <213> Secuencia Artificial

20

<220>
 <223> Secuencia sintética

25

<400> 36
 aaatcaaggt gaccagctcc 20

30

<210> 37
 <211> 22
 <212> ADN
 <213> Secuencia Artificial

35

<220>
 <223> Secuencia sintética

<400> 37
 tcatctgccca agacagaagt tc 22

40

<210> 38
 <211> 20
 <212> ADN
 <213> Secuencia Artificial

45

<220>
 <223> Secuencia sintética

<400> 38
 gcaggagagt caaaggtctg 20

50

<210> 39
 <211> 20
 <212> ADN
 <213> Secuencia Artificial

55

<220>
 <223> Secuencia sintética

<400> 39
 gttgccatgg agattgttc 20

60

<210> 40
 <211> 20
 <212> ADN
 <213> Secuencia Artificial

65

<220>
 <223> Secuencia sintética
 5 <400> 40
 cagctcagtg atgcattgc 20
 <210> 41
 <211> 21
 10 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 15 <400> 41
 ccttgacctc tgctaattg g 21
 <210> 42
 <211> 20
 20 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 25 <400> 42
 cacctgtcca acagctacag 20
 <210> 43
 <211> 21
 30 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 35 <400> 43
 agaattatc ttcaggcctg c 21
 <210> 44
 <211> 20
 40 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 45 <400> 44
 aagtaatcac tctgggtggc 20
 <210> 45
 <211> 20
 50 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 55 <400> 45
 agctcacaga caacctgtg 20
 <210> 46
 <211> 20
 60 <212> ADN
 <213> Secuencia Artificial
 65

<220>
 <223> Secuencia sintética
 5 <400> 46
 gcaatagaca cctacaggcg 20
 <210> 47
 <211> 20
 10 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 15 <400> 47
 gcacattatc aaagccacg 20
 <210> 48
 <211> 20
 20 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 25 <400> 48
 caacgaccta aagcatgtgc 20
 <210> 49
 <211> 20
 30 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 35 <400> 49
 gacatacatg gcttggcag 20
 <210> 50
 <211> 22
 40 <212> ADN
 <213> Secuencia Artificial
 45 <220>
 <223> Secuencia sintética
 <400> 50
 50 gagatactgc cacttatgca cg 22
 <210> 51
 <211> 20
 55 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 60 <400> 51
 cgtgctaata gtctcagggc 20
 <210> 52
 <211> 20
 65 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética
 5 <400> 52
 cgacgctca ttgcttcatt 20
 <210> 53
 <211> 20
 10 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 15 <400> 53
 cttgcgcca acaattgtcc 20
 <210> 54
 <211> 20
 20 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 25 <400> 54
 gctgcagagt ttgcattcat 20
 <210> 55
 <211> 21
 30 <212> ADN
 <213> Secuencia Artificial
 35 <220>
 <223> Secuencia sintética
 <400> 55
 40 catacacaca gaccgagagt c 21
 <210> 56
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 45 <220>
 <223> Secuencia sintética
 <400> 56
 50 ggatgtcagc cagcataagt 20
 <210> 57
 <211> 20
 <212> ADN
 55 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 <400> 57
 60 gcaagtgcc aacagtctc 20
 <210> 58
 <211> 21
 65 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética
 5 <400> 58
 gattccagca cacttgagtc t 21
 <210> 59
 <211> 20
 10 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 15 <400> 59
 ccggtgcagg tttaaagtc 20
 <210> 60
 <211> 20
 20 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 25 <400> 60
 caagagtgct ttatggcct 20
 <210> 61
 <211> 21
 30 <212> ADN
 <213> Secuencia Artificial
 35 <220>
 <223> Secuencia sintética
 <400> 61
 40 gcactcaagg agatcagact g 21
 <210> 62
 <211> 21
 <212> ADN
 <213> Secuencia Artificial
 45 <220>
 <223> Secuencia sintética
 <400> 62
 50 ggctatcgaa ctacaaccac a 21
 <210> 63
 <211> 21
 <212> ADN
 55 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 60 <400> 63
 gtatgtgtct ggggtgat c 21
 <210> 64
 <211> 22
 65 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

5 <400> 64
 caagaaactt cgagccttag ca 22

<210> 65
 <211> 20
 <212> ADN
 <213> Secuencia Artificial

10

<220>
 <223> Secuencia sintética

15

<400> 65
 gtgaaccagt ccgagtgaaa 20

<210> 66
 <211> 21
 <212> ADN
 <213> Secuencia Artificial

20

<220>
 <223> Secuencia sintética

25

<400> 66
 gcaaattgatg ttcagcacca c 21

30

<210> 67
 <211> 20
 <212> ADN
 <213> Secuencia Artificial

35

<220>
 <223> Secuencia sintética

<400> 67
 ttctccacc gaacgtgtct 20

40

<210> 68
 <211> 20
 <212> ADN
 <213> Secuencia Artificial

45

<220>
 <223> Secuencia sintética

<400> 68
 gcctattgc aagccctctt 20

50

<210> 69
 <211> 40
 <212> ADN
 <213> Secuencia Artificial

55

<220>
 <223> Secuencia sintética

<400> 69
 ttctccacc gaacgtgtct agaccagcac aacttactcg 40

60

<210> 70
 <211> 37
 <212> ADN
 <213> Secuencia Artificial

65

<220>
 <223> Secuencia sintética

5 <400> 70
 ttctccacc gaacgtgtct ccaaatgcac ctgctg 37

<210> 71
 <211> 41
 <212> ADN
 <213> Secuencia Artificial

10

<220>
 <223> Secuencia sintética

15

<400> 71
 ttctccacc gaacgtgtct agtttgaca aaggcaattc g 41

<210> 72
 <211> 44
 <212> ADN
 <213> Secuencia Artificial

20

<220>
 <223> Secuencia sintética

25

<400> 72
 ttctccacc gaacgtgtct tgagcttagc caatatcaag aagg 44

30

<210> 73
 <211> 42
 <212> ADN
 <213> Secuencia Artificial

35

<220>
 <223> Secuencia sintética

<400> 73
 ttctccacc gaacgtgtct acgtgaactt tccttgtagc ac 42

40

<210> 74
 <211> 44
 <212> ADN
 <213> Secuencia Artificial

45

<220>
 <223> Secuencia sintética

<400> 74
 ttctccacc gaacgtgtct tgaagatggt ctaatacctt gccg 44

50

<210> 75
 <211> 38
 <212> ADN
 <213> Secuencia Artificial

55

<220>
 <223> Secuencia sintética

<400> 75
 ttctccacc gaacgtgtct cagtgtggag actgaacg 38

60

<210> 76
 <211> 43
 <212> ADN
 <213> Secuencia Artificial

65

<220>
 <223> Secuencia sintética

5 <400> 76
 ttctccacc gaacgtgtct aggcaggga atgcatgaa atg 43

<210> 77
 <211> 37
 10 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

15 <400> 77
 ttctccacc gaacgtgtct gattgtctgg agcgctg 37

<210> 78
 <211> 36
 20 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

25 <400> 78
 ttctccacc gaacgtgtct agggagcaat aggccg 36

<210> 79
 <211> 37
 30 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

<400> 79
 40 ttctccacc gaacgtgtct ctgcaggga caacacg 37

<210> 80
 <211> 38
 <212> ADN
 45 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

<400> 80
 50 ttctccacc gaacgtgtct cgtatctggg aagacggc 38

<210> 81
 <211> 40
 <212> ADN
 55 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

<400> 81
 60 ttctccacc gaacgtgtct cctgtaatcc cttgcaatgc 40

<210> 82
 <211> 38
 <212> ADN
 65 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

5 <400> 82
 ttctccacc gaacgtgtct ggtctcagca cggttctg 38

<210> 83
 <211> 37
 <212> ADN
 <213> Secuencia Artificial

10

<220>
 <223> Secuencia sintética

15 <400> 83
 ttctccacc gaacgtgtct gcacctcct accacac 37

<210> 84
 <211> 41
 <212> ADN
 <213> Secuencia Artificial

20

<220>
 <223> Secuencia sintética

25 <400> 84
 ttctccacc gaacgtgtct gcctctagct agagagaagt c 41

30 <210> 85
 <211> 39
 <212> ADN
 <213> Secuencia Artificial

35

<220>
 <223> Secuencia sintética

<400> 85
 ttctccacc gaacgtgtct ctggcagtct agccgttac 39

40 <210> 86
 <211> 43
 <212> ADN
 <213> Secuencia Artificial

45

<220>
 <223> Secuencia sintética

<400> 86
 ttctccacc gaacgtgtct tgtcttagaa ttggcaact ggc 43

50 <210> 87
 <211> 39
 <212> ADN
 <213> Secuencia Artificial

55

<220>
 <223> Secuencia sintética

<400> 87
 ttctccacc gaacgtgtct gcaggaaagc ctactgaac 39

60 <210> 88
 <211> 39
 <212> ADN
 <213> Secuencia Artificial

65

<220>
 <223> Secuencia sintética

5 <400> 88
 ttctccacc gaacgtgtct gggagccaga gaaatgtcc 39

<210> 89
 <211> 44
 10 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

15 <400> 89
 ttctccacc gaacgtgtct tgtctccagt tccactcat ttag 44

<210> 90
 <211> 39
 20 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

25 <400> 90
 ttctccacc gaacgtgtct cccgtaatt gcctactcg 39

<210> 91
 <211> 38
 30 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

<400> 91
 40 ttctccacc gaacgtgtct ctcggtccca ctggaag 38

<210> 92
 <211> 41
 <212> ADN
 <213> Secuencia Artificial

45 <220>
 <223> Secuencia sintética

<400> 92
 50 ttctccacc gaacgtgtct acacccatga ttcagtact g 41

<210> 93
 <211> 41
 <212> ADN
 55 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

<400> 93
 60 ttctccacc gaacgtgtct gctagtatga acatcacagg c 41

<210> 94
 <211> 43
 65 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

5 <400> 94
 ttctccacc gaacgtgtct acaaatgagt aagaagcgag tog 43

<210> 95
 <211> 38
 <212> ADN
 <213> Secuencia Artificial

10

<220>
 <223> Secuencia sintética

15

<400> 95
 ttctccacc gaacgtgtct gataagggtt gctctgcg 38

<210> 96
 <211> 37
 <212> ADN
 <213> Secuencia Artificial

20

<220>
 <223> Secuencia sintética

25

<400> 96
 ttctccacc gaacgtgtct ccatgcacca gctaccc 37

30

<210> 97
 <211> 40
 <212> ADN
 <213> Secuencia Artificial

35

<220>
 <223> Secuencia sintética

<400> 97
 ttctccacc gaacgtgtct aactgtaccc tactcccagc 40

40

<210> 98
 <211> 41
 <212> ADN
 <213> Secuencia Artificial

45

<220>
 <223> Secuencia sintética

<400> 98
 ttctccacc gaacgtgtct aggaccaagg gaccagtta g 41

50

<210> 99
 <211> 42
 <212> ADN
 <213> Secuencia Artificial

55

<220>
 <223> Secuencia sintética

<400> 99
 ttctccacc gaacgtgtct agagtcttc caagaaattg cg 42

60

<210> 100
 <211> 44
 <212> ADN
 <213> Secuencia Artificial

65

<220>
 <223> Secuencia sintética

5 <400> 100
 ttctccacc gaacgtgtct acattataca gcatgctggc tatc 44

<210> 101
 <211> 41
 <212> ADN
 <213> Secuencia Artificial

10

<220>
 <223> Secuencia sintética

15 <400> 101
 ttctccacc gaacgtgtct gaggaagaaa gtgaggtttg c 41

<210> 102
 <211> 44
 <212> ADN
 <213> Secuencia Artificial

20

<220>
 <223> Secuencia sintética

25 <400> 102
 ttctccacc gaacgtgtct ctgaattatg tgcttaccaa gagc 44

<210> 103
 <211> 44
 <212> ADN
 <213> Secuencia Artificial

30

<220>
 <223> Secuencia sintética

35 <400> 103
 ttctccacc gaacgtgtct tgggttctga taacctatc aagc 44

<210> 104
 <211> 39
 <212> ADN
 <213> Secuencia Artificial

40

<220>
 <223> Secuencia sintética

45 <400> 104
 ttctccacc gaacgtgtct ggtagtcaa acatgctgc 39

<210> 105
 <211> 42
 <212> ADN
 <213> Secuencia Artificial

50

<220>
 <223> Secuencia sintética

55 <400> 105
 ttctccacc gaacgtgtct gacactggca gaatcaaac ac 42

<210> 106
 <211> 44
 <212> ADN
 <213> Secuencia Artificial

60

65

<220>
 <223> Secuencia sintética

5 <400> 106
 ttctccacc gaacgtgtct agagttacac ctttagctaa ccac 44

<210> 107
 <211> 38
 10 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

15 <400> 107
 ttctccacc gaacgtgtct ccaggagttc aagaagcg 38

<210> 108
 <211> 41
 20 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

25 <400> 108
 ttctccacc gaacgtgtct accactcctt tctccatct c 41

<210> 109
 <211> 44
 30 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

<400> 109
 40 ttctccacc gaacgtgtct gtcttatggg acaatgggtg atag 44

<210> 110
 <211> 38
 <212> ADN
 <213> Secuencia Artificial

45 <220>
 <223> Secuencia sintética

<400> 110
 50 ttctccacc gaacgtgtct ctaccctcaa cctcgtc 38

<210> 111
 <211> 38
 <212> ADN
 55 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

<400> 111
 60 ttctccacc gaacgtgtct ccaagactga tcatgccg 38

<210> 112
 <211> 40
 65 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

5 <400> 112
 gccctattgc aagccctctt agaccagcac aacttactta 40

<210> 113
 <211> 37
 <212> ADN
 <213> Secuencia Artificial

10

<220>
 <223> Secuencia sintética

15

<400> 113
 gccctattgc aagccctctt ccaaattcac ctgccca 37

<210> 114
 <211> 41
 <212> ADN
 <213> Secuencia Artificial

20

<220>
 <223> Secuencia sintética

25

<400> 114
 gccctattgc aagccctctt agtttgaca aaggcgatt a 41

30

<210> 115
 <211> 44
 <212> ADN
 <213> Secuencia Artificial

35

<220>
 <223> Secuencia sintética

<400> 115
 gccctattgc aagccctctt tgagcttagc caatatcaac aaga 44

40

<210> 116
 <211> 42
 <212> ADN
 <213> Secuencia Artificial

45

<220>
 <223> Secuencia sintética

<400> 116
 gccctattgc aagccctctt acgtgaactt tccttgtaa at 42

50

<210> 117
 <211> 44
 <212> ADN
 <213> Secuencia Artificial

55

<220>
 <223> Secuencia sintética

<400> 117
 gccctattgc aagccctctt tgaagatggt ctaatacctt gcta 44

60

<210> 118
 <211> 38
 <212> ADN
 <213> Secuencia Artificial

65

<220>
 <223> Secuencia sintética

5 <400> 118
 gccctattgc aagccctctt cagtgtggag accgaaca 38

<210> 119
 <211> 43
 10 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

15 <400> 119
 gccctattgc aagccctctt aggcaggga atgcatgaa gtt 43

<210> 120
 <211> 37
 20 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

25 <400> 120
 gccctattgc aagccctctt gattgtctgg agggctc 37

<210> 121
 <211> 36
 30 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

<400> 121
 40 gccctattgc aagccctctt agggagcaat aggcta 36

<210> 122
 <211> 37
 <212> ADN
 <213> Secuencia Artificial

45 <220>
 <223> Secuencia sintética

<400> 122
 50 gccctattgc aagccctctt ctgcaggga caagaca 37

<210> 123
 <211> 38
 <212> ADN
 55 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

<400> 123
 60 gccctattgc aagccctctt cgtatctggg aagatggg 38

<210> 124
 <211> 40
 65 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

5 <400> 124
 gccctattgc aagccctctt cctgtaatcc cttgcaataa 40

<210> 125
 <211> 38
 10 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

15 <400> 125
 gccctattgc aagccctctt ggtctcagca cggctcct 38

<210> 126
 <211> 37
 20 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

25 <400> 126
 gccctattgc aagccctctt gcacctcct atcacat 37

<210> 127
 <211> 41
 30 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

<400> 127
 40 gccctattgc aagccctctt gcctctagct agagagaagc g 41

<210> 128
 <211> 39
 <212> ADN
 <213> Secuencia Artificial

45 <220>
 <223> Secuencia sintética

<400> 128
 50 gccctattgc aagccctctt ctggcagtct agccattat 39

<210> 129
 <211> 43
 <212> ADN
 55 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

<400> 129
 60 gccctattgc aagccctctt tgtcttagaa ttggcaact agt 43

<210> 130
 <211> 39
 65 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

5 <400> 130
 gccctattgc aagccctctt gcaggaaagc ctattgaat 39

<210> 131
 <211> 39
 <212> ADN
 <213> Secuencia Artificial

10

<220>
 <223> Secuencia sintética

15

<400> 131
 gccctattgc aagccctctt gggagccaga gaaattct 39

<210> 132
 <211> 44
 <212> ADN
 <213> Secuencia Artificial

20

<220>
 <223> Secuencia sintética

25

<400> 132
 gccctattgc aagccctctt tgtctccagt tccactcat gtaa 44

30

<210> 133
 <211> 39
 <212> ADN
 <213> Secuencia Artificial

35

<220>
 <223> Secuencia sintética

<400> 133
 gccctattgc aagccctctt cccgtaatt gcctattta 39

40

<210> 134
 <211> 38
 <212> ADN
 <213> Secuencia Artificial

45

<220>
 <223> Secuencia sintética

<400> 134
 gccctattgc aagccctctt ctcggtocca ctgggaaa 38

50

<210> 135
 <211> 41
 <212> ADN
 <213> Secuencia Artificial

55

<220>
 <223> Secuencia sintética

<400> 135
 gccctattgc aagccctctt acacccatga ttcagtacc a 41

60

<210> 136
 <211> 41
 <212> ADN
 <213> Secuencia Artificial

65

<220>
 <223> Secuencia sintética

5 <400> 136
 gccctattgc aagccctctt gctagtatga acatcacaag t 41

<210> 137
 <211> 43
 10 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

15 <400> 137
 gccctattgc aagccctctt acaaatgagt aagaagcgag tta 43

<210> 138
 <211> 38
 20 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

25 <400> 138
 gccctattgc aagccctctt gataagggtt gctctaca 38

<210> 139
 <211> 37
 30 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

<400> 139
 40 gccctattgc aagccctctt ccatgcacca gctacta 37

<210> 140
 <211> 40
 <212> ADN
 <213> Secuencia Artificial

45 <220>
 <223> Secuencia sintética

<400> 140
 50 gccctattgc aagccctctt aactgtacc tactccaat 40

<210> 141
 <211> 41
 <212> ADN
 55 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

<400> 141
 60 gccctattgc aagccctctt aggaccaagg gaccagtta c 41

<210> 142
 <211> 42
 65 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

5 <400> 142
 gccctattgc aagccctctt agagttcttc caagaaattg ta 42

<210> 143
 <211> 44
 <212> ADN
 <213> Secuencia Artificial

10

<220>
 <223> Secuencia sintética

15

<400> 143
 gccctattgc aagccctctt acattataca gcatgctggt taga 44

<210> 144
 <211> 41
 <212> ADN
 <213> Secuencia Artificial

20

<220>
 <223> Secuencia sintética

25

<400> 144
 gccctattgc aagccctctt gaggaagaaa gtgagattg t 41

30

<210> 145
 <211> 44
 <212> ADN
 <213> Secuencia Artificial

35

<220>
 <223> Secuencia sintética

<400> 145
 gccctattgc aagccctctt ctgaattatg tgctaccag gagt 44

40

<210> 146
 <211> 44
 <212> ADN
 <213> Secuencia Artificial

45

<220>
 <223> Secuencia sintética

<400> 146
 gccctattgc aagccctctt tgggttctga taacctatc aact 44

50

<210> 147
 <211> 39
 <212> ADN
 <213> Secuencia Artificial

55

<220>
 <223> Secuencia sintética

<400> 147
 gccctattgc aagccctctt ggtagtcaa acatgttg t 39

60

<210> 148
 <211> 42
 <212> ADN
 <213> Secuencia Artificial

65

<220>
 <223> Secuencia sintética

5 <400> 148
 gccctattgc aagccctctt gacactggca gaatcaaacc aa 42

<210> 149
 <211> 44
 10 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

15 <400> 149
 gccctattgc aagccctctt agagtacac ctttagctaa ctag 44

<210> 150
 <211> 38
 20 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

25 <400> 150
 gccctattgc aagccctctt ccaggagtc aaggagca 38

30 <210> 151
 <211> 41
 <212> ADN
 <213> Secuencia Artificial

35 <220>
 <223> Secuencia sintética

<400> 151
 40 gccctattgc aagccctctt accactcctt tctcccgct t 41

<210> 152
 <211> 44
 <212> ADN
 <213> Secuencia Artificial

45 <220>
 <223> Secuencia sintética

<400> 152
 50 gccctattgc aagccctctt gtcttatggg acaatggtcg atat 44

<210> 153
 <211> 38
 <212> ADN
 55 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

<400> 153
 60 gccctattgc aagccctctt ctaccctcaa cctcatt 38

<210> 154
 <211> 38
 65 <212> ADN
 <213> Secuencia Artificial

ES 2 739 430 T3

<220>
<223> Secuencia sintética

5 <400> 154
gccctattgc aagcctctt ccaagactga tcatgcta 38

<210> 155
<211> 61
10 <212> ADN
<213> Secuencia Artificial

<220>
<223> Secuencia sintética

15 <400> 155
cacttgacaa agttctcacg cgccgaagtt ctccgaagga tgcctcatc ttcttcctg 60
c 61

<210> 156
20 <211> 61
<212> ADN
<213> Secuencia Artificial

<220>
25 <223> Secuencia sintética

<400> 156
cattagggat taacggcttg ggacagactg acggagcttc agcctcatc ttcttcctg 60
c 61

<210> 157
30 <211> 63
<212> ADN
<213> Secuencia Artificial

<220>
35 <223> Secuencia sintética

<400> 157
cacacgtaa gaagactttc tgctgactct gccgcacatg atcgccctca tcttcttccc 60
tgc 63

<210> 158
40 <211> 62
<212> ADN
<213> Secuencia Artificial

<220>
45 <223> Secuencia sintética

<400> 158
ctaagtgcc tccatgagaa aggatccgat agcctctgc aggccctcat cttcttcct 60

50 gc 62

<210> 159
55 <211> 60
<212> ADN
<213> Secuencia Artificial

<220>
<223> Secuencia sintética

ES 2 739 430 T3

<400> 159
 gcacagattt cccacactct caacaggcct gctaaacacc gccctcatct tcttccctgc 60

5 <210> 160
 <211> 61
 <212> ADN
 <213> Secuencia Artificial

10 <220>
 <223> Secuencia sintética

<400> 160
 cttacaggag gtctggcatc aggtcaacaa cggagggact cggcctcatc ttcttccctg 60

c 61

15 <210> 161
 <211> 59
 <212> ADN
 <213> Secuencia Artificial

20 <220>
 <223> Secuencia sintética

<400> 161
 ccacaatgag aaggcagagt tgcattaat gctggcggcg ccctcatctt cttccctgc 59

25 <210> 162
 <211> 60
 <212> ADN
 <213> Secuencia Artificial

30 <220>
 <223> Secuencia sintética

<400> 162
 gctgtggcat agctacactc cggtgacggt ttgcaacttt gccctcatct tcttccctgc 60

40 <210> 163
 <211> 60
 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

45 <400> 163
 cagggaatt tggggtctg gtcggcaggt taagggtctc gccctcatct tcttccctgc 60

50 <210> 164
 <211> 62
 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Secuencia sintética

55 <400> 164
 gggctatcca gaaagataag aataactcaca aacgactgcg cagccctcat cttcttccct 60

gc 62

60 <210> 165
 <211> 63
 <212> ADN
 <213> Secuencia Artificial

ES 2 739 430 T3

	<220>		
	<223>	Secuencia sintética	
5	<400>	165	
		cataactggt ggagtatctc actcgtatat gcccgactgg agggccctca tcttctccc	60
		tgc	63
	<210>	166	
10	<211>	64	
	<212>	ADN	
	<213>	Secuencia Artificial	
	<220>		
	<223>	Secuencia sintética	
15	<400>	166	
		cttcaaggaa gaaattcaac agggtagggt ttgcggcgat aagggccctc atcttctcc	60
		ctgc	64
	<210>	167	
20	<211>	62	
	<212>	ADN	
	<213>	Secuencia Artificial	
	<220>		
25	<223>	Secuencia sintética	
	<400>	167	
		catggattca acacagcaaa caccaagtca accaccogag acgcccctcat cttcttccct	60
		gc	62
30	<210>	168	
	<211>	62	
	<212>	ADN	
	<213>	Secuencia Artificial	
35	<220>		
	<223>	Secuencia sintética	
	<400>	168	
		ctctgacctc cttcactctt acacttccct ggccttccct ctgcccctcat cttcttccct	60
		gc	62
40	<210>	169	
	<211>	62	
	<212>	ADN	
	<213>	Secuencia Artificial	
45	<220>		
	<223>	Secuencia sintética	
	<400>	169	
		gctttcattt gtgctaaacc togcttgggt cctctcctga acgcccctcat cttcttccct	60
50		gc	62
	<210>	170	
55	<211>	58	
	<212>	ADN	
	<213>	Secuencia Artificial	
	<220>		

ES 2 739 430 T3

<223> Secuencia sintética

<400> 170
 catcccagat gccctcataa cgccgaacc acaatgctgc cctcatctt ttcctgc 58

5

<210> 171
 <211> 62
 <212> ADN
 <213> Secuencia Artificial

10

<220>
 <223> Secuencia sintética

<400> 171
 gtagaaatcc caaggcaatc agctcctcgc atccaacagt cggccctcat cttcttccct 60

15

gc 62

<210> 172
 <211> 63
 <212> ADN
 <213> Secuencia Artificial

20

<220>
 <223> Secuencia sintética

25

<400> 172
 gaacaactaa ctccacagaa cccccaccgt agcactcctt cttgccctca tcttcttccc 60

tgc 63

<210> 173
 <211> 59
 <212> ADN
 <213> Secuencia Artificial

30

<220>
 <223> Secuencia sintética

35

<400> 173
 gtgcagagga caggaagaac ggagcgtcgg tagtgtaaag ccctcatctt cttccctgc 59

40

<210> 174
 <211> 63
 <212> ADN
 <213> Secuencia Artificial

45

<220>
 <223> Secuencia sintética

<400> 174
 ggtgcttcaa gacatacacc ttaacaactc gacgaacctc cggccctca tcttcttccc 60

50

tgc 63

<210> 175
 <211> 60
 <212> ADN
 <213> Secuencia Artificial

55

<220>
 <223> Secuencia sintética

60

<400> 175
 ggaacctctg tgacctgga tggccatcc ttatgtgctg gccctcatct tcttccctgc 60

ES 2 739 430 T3

<210> 176
 <211> 59
 <212> ADN
 <213> Secuencia Artificial
 5
 <220>
 <223> Secuencia sintética
 <400> 176
 10 cccagtggtta cctctgaag gtcggtattg ctcaagcccg ccctcatctt ctccctgc 59
 <210> 177
 <211> 63
 <212> ADN
 15 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 20 <400> 177
 cttctgttgc ttatttgggt aacttgattc tggccctccc atcgccctca tcttcttccc 60
 tgc 63
 <210> 178
 <211> 56
 <212> ADN
 25 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 30 <400> 178
 cccactggat gctccctca cgccggctat ttaggtgcc tcattctt cctgc 56
 <210> 179
 <211> 58
 <212> ADN
 35 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 40 <400> 179
 cggagagacg catctgaaag tctggtagg tggaggacgc cctcatctt tccctgc 58
 <210> 180
 <211> 61
 <212> ADN
 45 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 50 <400> 180
 caggatttcc agcttacagg gcgactgagc cacatccaac tggcctcatc ttcttccctg 60
 c 61
 55 <210> 181
 <211> 60
 <212> ADN
 <213> Secuencia Artificial
 60 <220>
 <223> Secuencia sintética

ES 2 739 430 T3

<400> 181
cttgcaagat gtgcctotta gagcctcagc cggaattgaa gccctcatct tcttcctgc 60

5 <210> 182
<211> 62
<212> ADN
<213> Secuencia Artificial

10 <220>
<223> Secuencia sintética

<400> 182
gggtggtttc tctaaacaca aattgccatt ctgcaccaat ggcacctcat cttcttcctt 60

15 gc 62

<210> 183
<211> 62
<212> ADN
<213> Secuencia Artificial

20 <220>
<223> Secuencia sintética

<400> 183
gcagggtatt gagagaagga tctattggtg ttgcggcgtg atgccctcat cttcttcctt 60

25 gc 62

<210> 184
<211> 62
<212> ADN
<213> Secuencia Artificial

30 <220>
<223> Secuencia sintética

35 <400> 184
gtgcacattt ctgatgaag ggatggcgt aacaggagga ctgccctcat cttcttcctt 60

gc 62

<210> 185
<211> 62
<212> ADN
<213> Secuencia Artificial

40 <220>
<223> Secuencia sintética

45 <400> 185
gagcaatgcc tgtttcatga gaggaatggc ctacctgcat cagacctcat cttcttcctt 60

gc 62

<210> 186
<211> 64
<212> ADN
<213> Secuencia Artificial

50 <220>
<223> Secuencia sintética

55 <400> 186

ES 2 739 430 T3

	gttaacatta tacagcatgg tggccccggt gttgtcatcg catcgccctc atcttcttcc	60
	ctgc	64
5	<210> 187 <211> 59 <212> ADN <213> Secuencia Artificial	
10	<220> <223> Secuencia sintética	
	<400> 187 gagaacatg tctgaagcg ttcgatgctg cccatgagtg ccctcatct cttccctgc	59
15	<210> 188 <211> 60 <212> ADN <213> Secuencia Artificial	
20	<220> <223> Secuencia sintética	
	<400> 188 cagctgttc ccaaaccat caaccgcgt agatgttct gcctcatct tctccctgc	60
25	<210> 189 <211> 61 <212> ADN <213> Secuencia Artificial	
30	<220> <223> Secuencia sintética	
	<400> 189 caaagtgtgg aagttgcttc cgccagctca agagtgtagc cgccctcatc ttcttccctg	60
35	c	61
40	<210> 190 <211> 58 <212> ADN <213> Secuencia Artificial	
	<220> <223> Secuencia sintética	
45	<400> 190 ggtcgacttt gtccatcct ctgatcctg cgcgatgtgc cctcatctc ttcctgc	58
50	<210> 191 <211> 59 <212> ADN <213> Secuencia Artificial	
	<220> <223> Secuencia sintética	
55	<400> 191 ctctgtgcc tgtgactca tcgaggcgt tcctatacg ccctcatct cttccctgc	59
60	<210> 192 <211> 64 <212> ADN <213> Secuencia Artificial	

ES 2 739 430 T3

<220>
 <223> Secuencia sintética

<400> 192
 ctaactagaa ttagtctgcc tgcctattgg acctccgacc acgagccctc atcttcttcc 60

5 ctgc 64

<210> 193
 <211> 59
 <212> ADN
 <213> Secuencia Artificial

10

<220>
 <223> Secuencia sintética

15 <400> 193
 gtgagcata atcgtgtcaa gccaccattt agatccgagg ccctcatctt cttccctgc 59

20 <210> 194
 <211> 63
 <212> ADN
 <213> Secuencia Artificial

25 <220>
 <223> Secuencia sintética

<400> 194
 gagaattaat gctccctctc ctggaccagt agaagtctgc ccggccctca tcttcttccc 60

tgc 63

30 <210> 195
 <211> 59
 <212> ADN
 <213> Secuencia Artificial

35 <220>
 <223> Secuencia sintética

<400> 195
 gtggctgct gttgaccaat ttcagaatgg ccgagctgtg ccctcatctt cttccctgc 59

40 <210> 196
 <211> 60
 <212> ADN
 <213> Secuencia Artificial

45 <220>
 <223> Secuencia sintética

<400> 196
 ggttgaact gctgatctat aggtgacctt ctgtacgcc gcctcatct tcttccctgc 60

50 <210> 197
 <211> 58
 <212> ADN
 <213> Secuencia Artificial

55 <220>
 <223> Secuencia sintética

<400> 197
 ctttccagct caaggcaggg cgcgctccta ttccatcgc cctcatctc ttcctgc 58

60 <210> 198

<211> 20
 <212> ADN
 <213> Secuencia Artificial
 5 <220>
 <223> Secuencia sintética
 <400> 198
 agaccagcac aacttactcg 20
 10 <210> 199
 <211> 17
 <212> ADN
 <213> Secuencia Artificial
 15 <220>
 <223> Secuencia sintética
 <400> 199
 ccaaatgcac ctgcctg 17
 <210> 200
 <211> 21
 <212> ADN
 25 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 30 <400> 200
 agtttgaca aaggcaattc g 21
 <210> 201
 <211> 24
 <212> ADN
 35 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 40 <400> 201
 tgagcttagc caatatcaag aagg 24
 <210> 202
 <211> 22
 <212> ADN
 <213> Secuencia Artificial
 45 <220>
 <223> Secuencia sintética
 50 <400> 202
 acgtgaactt tccttggtac ac 22
 <210> 203
 <211> 24
 <212> ADN
 <213> Secuencia Artificial
 55 <220>
 <223> Secuencia sintética
 <400> 203
 tgaagatggt ctaatacctt gccg 24
 60 <210> 204
 65

<211> 18
 <212> ADN
 <213> Secuencia Artificial
 5 <220>
 <223> Secuencia sintética
 <400> 204
 cagtgtggag actgaacg 18
 10 <210> 205
 <211> 23
 <212> ADN
 <213> Secuencia Artificial
 15 <220>
 <223> Secuencia sintética
 <400> 205
 20 aggcaggta atgtcatgaa atg 23
 <210> 206
 <211> 17
 <212> ADN
 25 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 30 <400> 206
 gattgtctgg agcgctg 17
 <210> 207
 <211> 16
 35 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 40 <400> 207
 agggagcaat aggccg 16
 <210> 208
 <211> 17
 45 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 50 <400> 208
 ctgcaggta caacacg 17
 <210> 209
 <211> 18
 55 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 60 <400> 209
 cgatatctggg aagacggc 18
 65 <210> 210

<211> 20
 <212> ADN
 <213> Secuencia Artificial

5

<220>
 <223> Secuencia sintética

<400> 210
 cctgtaatcc ctgcaatgc 20

10

<210> 211
 <211> 18
 <212> ADN
 <213> Secuencia Artificial

15

<220>
 <223> Secuencia sintética

<400> 211
 ggtctcagca cggttctg 18

20

<210> 212
 <211> 17
 <212> ADN
 <213> Secuencia Artificial

25

<220>
 <223> Secuencia sintética

<400> 212
 gcacctccct accacac 17

30

<210> 213
 <211> 21
 <212> ADN
 <213> Secuencia Artificial

35

<220>
 <223> Secuencia sintética

<400> 213
 gcctctagct agagagaagt c 21

40

<210> 214
 <211> 19
 <212> ADN
 <213> Secuencia Artificial

45

<220>
 <223> Secuencia sintética

<400> 214
 ctggcagtct agccggtac 19

50

<210> 215
 <211> 23
 <212> ADN
 <213> Secuencia Artificial

55

<220>
 <223> Secuencia sintética

<400> 215
 tgtcttagaa ttggcaact ggc 23

60

<210> 216

65

<211> 19
 <212> ADN
 <213> Secuencia Artificial
 5 <220>
 <223> Secuencia sintética
 <400> 216
 gcaggaaagc ctactgaac 19
 10 <210> 217
 <211> 19
 <212> ADN
 <213> Secuencia Artificial
 15 <220>
 <223> Secuencia sintética
 <400> 217
 20 gggagccaga gaaatgtcc 19
 <210> 218
 <211> 24
 <212> ADN
 25 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 30 <400> 218
 tgtctccagt tccactcat ttag 24
 <210> 219
 <211> 19
 35 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 40 <400> 219
 cccgtaatt gcctactcg 19
 <210> 220
 <211> 18
 45 <212> ADN
 <213> Secuencia Artificial
 <220>
 50 <223> Secuencia sintética
 <400> 220
 ctcggtccca ctggaaag 18
 55 <210> 221
 <211> 21
 <212> ADN
 <213> Secuencia Artificial
 60 <220>
 <223> Secuencia sintética
 <400> 221
 65 acacctatga tcagttact g 21
 <210> 222

<211> 21
 <212> ADN
 <213> Secuencia Artificial

5 <220>
 <223> Secuencia sintética

<400> 222
 gctagtatga acatcacagg c 21

10 <210> 223
 <211> 23
 <212> ADN
 <213> Secuencia Artificial

15 <220>
 <223> Secuencia sintética

<400> 223
 acaaatgagt aagaagcgag tcg 23

20 <210> 224
 <211> 18
 <212> ADN
 <213> Secuencia Artificial

25 <220>
 <223> Secuencia sintética

<400> 224
 gataagggtt gctctgcg 18

30 <210> 225
 <211> 17
 <212> ADN
 <213> Secuencia Artificial

35 <220>
 <223> Secuencia sintética

<400> 225
 ccatgcacca gctaccc 17

40 <210> 226
 <211> 20
 <212> ADN
 <213> Secuencia Artificial

45 <220>
 <223> Secuencia sintética

<400> 226
 aactgtaccc tactcccagc 20

50 <210> 227
 <211> 21
 <212> ADN
 <213> Secuencia Artificial

55 <220>
 <223> Secuencia sintética

<400> 227
 aggaccaagg gaccagtta g 21

60 <210> 228

65

<211> 22
 <212> ADN
 <213> Secuencia Artificial

5 <220>
 <223> Secuencia sintética

<400> 228
 agagttcctc caagaaattg cg 22

10 <210> 229
 <211> 24
 <212> ADN
 <213> Secuencia Artificial

15 <220>
 <223> Secuencia sintética

<400> 229
 acattataca gcatgctggc tatc 24

20 <210> 230
 <211> 21
 <212> ADN
 <213> Secuencia Artificial

25 <220>
 <223> Secuencia sintética

<400> 230
 gaggaagaaa gtgaggttg c 21

30 <210> 231
 <211> 24
 <212> ADN
 <213> Secuencia Artificial

35 <220>
 <223> Secuencia sintética

<400> 231
 ctgaattatg tgcttaccaa gage 24

40 <210> 232
 <211> 24
 <212> ADN
 <213> Secuencia Artificial

45 <220>
 <223> Secuencia sintética

<400> 232
 tgggttctga taaccttatc aagc 24

50 <210> 233
 <211> 19
 <212> ADN
 <213> Secuencia Artificial

55 <220>
 <223> Secuencia sintética

<400> 233
 ggtagtcaa acatgctgc 19

60 <210> 234

65

<211> 22
 <212> ADN
 <213> Secuencia Artificial
 5 <220>
 <223> Secuencia sintética
 <400> 234
 gacactggca gaatcaaac ac 22
 10 <210> 235
 <211> 24
 <212> ADN
 <213> Secuencia Artificial
 15 <220>
 <223> Secuencia sintética
 <400> 235
 20 agagttacac ctttagctaa ccac 24
 <210> 236
 <211> 18
 <212> ADN
 25 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 30 <400> 236
 ccaggagttc aagaagcg 18
 <210> 237
 <211> 21
 35 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 40 <400> 237
 accactcctt tctccatct c 21
 <210> 238
 <211> 24
 45 <212> ADN
 <213> Secuencia Artificial
 <220>
 50 <223> Secuencia sintética
 <400> 238
 gtcttatggg acaatggtg atag 24
 55 <210> 239
 <211> 18
 <212> ADN
 <213> Secuencia Artificial
 60 <220>
 <223> Secuencia sintética
 <400> 239
 65 ctaccctcaa ccctcgtc 18
 <210> 240

<211> 18
 <212> ADN
 <213> Secuencia Artificial
 5 <220>
 <223> Secuencia sintética
 <400> 240
 ccaagactga tcatgccg 18
 10 <210> 241
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 15 <220>
 <223> Secuencia sintética
 <400> 241
 20 agaccagcac aactactta 20
 <210> 242
 <211> 17
 <212> ADN
 25 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 30 <400> 242
 ccaaattcac ctgccca 17
 <210> 243
 <211> 21
 35 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 40 <400> 243
 agtttgaca aaggcgattt a 21
 <210> 244
 <211> 24
 45 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 50 <400> 244
 tgagcttagc caatatcaac aaga 24
 <210> 245
 <211> 22
 55 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 <400> 245
 60 acgtgaactt tccttgtaa at 22
 65 <210> 246

<211> 24
 <212> ADN
 <213> Secuencia Artificial

5

<220>
 <223> Secuencia sintética

<400> 246
 tgaagatggt ctaatacctt gcta 24

10

<210> 247
 <211> 18
 <212> ADN
 <213> Secuencia Artificial

15

<220>
 <223> Secuencia sintética

<400> 247
 cagtgtggag accgaaca 18

20

<210> 248
 <211> 23
 <212> ADN
 <213> Secuencia Artificial

25

<220>
 <223> Secuencia sintética

<400> 248
 aggcaggta atgtcatgaa gtt 23

30

<210> 249
 <211> 17
 <212> ADN
 <213> Secuencia Artificial

35

<220>
 <223> Secuencia sintética

<400> 249
 gattgtctgg agggctc 17

40

<210> 250
 <211> 16
 <212> ADN
 <213> Secuencia Artificial

45

<220>
 <223> Secuencia sintética

<400> 250
 agggagcaat aggcta 16

50

<210> 251
 <211> 17
 <212> ADN
 <213> Secuencia Artificial

55

<220>
 <223> Secuencia sintética

<400> 251
 ctgcaggta caagaca 17

60

<210> 252

65

<211> 18
 <212> ADN
 <213> Secuencia Artificial
 5 <220>
 <223> Secuencia sintética
 <400> 252
 cgatatctggg aagatggg 18
 10 <210> 253
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 15 <220>
 <223> Secuencia sintética
 <400> 253
 20 cctgtaatcc ctgcaataa 20
 <210> 254
 <211> 18
 <212> ADN
 25 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 30 <400> 254
 ggtctcagca cggctctt 18
 <210> 255
 <211> 17
 35 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 40 <400> 255
 gcacctccct atcacat 17
 <210> 256
 <211> 21
 <212> ADN
 <213> Secuencia Artificial
 45 <220>
 <223> Secuencia sintética
 50 <400> 256
 gcctctagct agagagaagc g 21
 <210> 257
 <211> 19
 <212> ADN
 <213> Secuencia Artificial
 55 <220>
 <223> Secuencia sintética
 <400> 257
 60 ctggcagtct agccattat 19
 <210> 258

<211> 23
 <212> ADN
 <213> Secuencia Artificial

5 <220>
 <223> Secuencia sintética

<400> 258
 tgtcttagaa tttggcaact agt 23

10 <210> 259
 <211> 19
 <212> ADN
 <213> Secuencia Artificial

15 <220>
 <223> Secuencia sintética

<400> 259
 gcaggaaagc ctattgaat 19

20 <210> 260
 <211> 19
 <212> ADN
 <213> Secuencia Artificial

25 <220>
 <223> Secuencia sintética

<400> 260
 gggagccaga gaaatttct 19

30 <210> 261
 <211> 24
 <212> ADN
 <213> Secuencia Artificial

35 <220>
 <223> Secuencia sintética

<400> 261
 tgtctccagt tccactcat gtaa 24

40 <210> 262
 <211> 19
 <212> ADN
 <213> Secuencia Artificial

45 <220>
 <223> Secuencia sintética

<400> 262
 cccgtaatt gcctattta 19

50 <210> 263
 <211> 18
 <212> ADN
 <213> Secuencia Artificial

55 <220>
 <223> Secuencia sintética

<400> 263
 ctcggtccca ctgggaaa 18

60 <210> 264

65

<211> 21
 <212> ADN
 <213> Secuencia Artificial
 5 <220>
 <223> Secuencia sintética
 <400> 264
 acacccatga ttcagttacc a 21
 10 <210> 265
 <211> 21
 <212> ADN
 <213> Secuencia Artificial
 15 <220>
 <223> Secuencia sintética
 <400> 265
 20 gctagtatga acatcacaag t 21
 <210> 266
 <211> 23
 <212> ADN
 25 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 30 <400> 266
 acaaatgagt aagaagcgag tta 23
 <210> 267
 <211> 18
 35 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 40 <400> 267
 gataagggtt gctctaca 18
 <210> 268
 <211> 17
 45 <212> ADN
 <213> Secuencia Artificial
 <220>
 50 <223> Secuencia sintética
 <400> 268
 ccatgcacca gctacta 17
 55 <210> 269
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 60 <220>
 <223> Secuencia sintética
 <400> 269
 65 aactgtaccc tactccaat 20
 <210> 270

	<211> 21	
	<212> ADN	
	<213> Secuencia Artificial	
5	<220>	
	<223> Secuencia sintética	
	<400> 270	
10	aggaccaagg gaccagttca c	21
	<210> 271	
	<211> 22	
	<212> ADN	
15	<213> Secuencia Artificial	
	<220>	
	<223> Secuencia sintética	
	<400> 271	
20	agagttcctc caagaaattg ta	22
	<210> 272	
	<211> 24	
	<212> ADN	
25	<213> Secuencia Artificial	
	<220>	
	<223> Secuencia sintética	
	<400> 272	
30	acattataca gcatgctggt taga	24
	<210> 273	
	<211> 21	
35	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
	<223> Secuencia sintética	
40	<400> 273	
	gaggaagaaa gtgagattg t	21
	<210> 274	
45	<211> 24	
	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
50	<223> Secuencia sintética	
	<400> 274	
	ctgaattatg tgctaccag gagt	24
55	<210> 275	
	<211> 24	
	<212> ADN	
	<213> Secuencia Artificial	
60	<220>	
	<223> Secuencia sintética	
	<400> 275	
65	tgggttctga taacctatc aact	24
	<210> 276	

<211> 19
 <212> ADN
 <213> Secuencia Artificial
 5 <220>
 <223> Secuencia sintética
 <400> 276
 ggtagtcaa acatgtgt 19
 10 <210> 277
 <211> 22
 <212> ADN
 <213> Secuencia Artificial
 15 <220>
 <223> Secuencia sintética
 <400> 277
 20 gacactggca gaatcaaacc aa 22
 <210> 278
 <211> 24
 <212> ADN
 25 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 30 <400> 278
 agagttacac ctttagctaa ctg 24
 <210> 279
 <211> 18
 35 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 40 <400> 279
 ccaggagttc aaggagca 18
 <210> 280
 <211> 21
 <212> ADN
 <213> Secuencia Artificial
 50 <220>
 <223> Secuencia sintética
 <400> 280
 accactcctt tctccgtct t 21
 55 <210> 281
 <211> 24
 <212> ADN
 <213> Secuencia Artificial
 60 <220>
 <223> Secuencia sintética
 <400> 281
 65 gtcttatggg acaatggtcg atat 24
 <210> 282

<211> 18
 <212> ADN
 <213> Secuencia Artificial
 5 <220>
 <223> Secuencia sintética
 <400> 282
 ctaccctcaa ccctcatt 18
 10 <210> 283
 <211> 18
 <212> ADN
 <213> Secuencia Artificial
 15 <220>
 <223> Secuencia sintética
 <400> 283
 ccaagactga tcatgcta 18
 <210> 284
 <211> 21
 <212> ADN
 25 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 30 <400> 284
 cacttgacaa agttctcag c 21
 <210> 285
 <211> 21
 <212> ADN
 35 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 40 <400> 285
 cattagggat taacggcttg g 21
 <210> 286
 <211> 23
 <212> ADN
 <213> Secuencia Artificial
 45 <220>
 <223> Secuencia sintética
 50 <400> 286
 cacacgtaa gaagacttc tgc 23
 <210> 287
 <211> 22
 <212> ADN
 <213> Secuencia Artificial
 55 <220>
 <223> Secuencia sintética
 <400> 287
 ctaagtgcc tccatgagaa ag 22
 60 <210> 288
 65

<211> 20
 <212> ADN
 <213> Secuencia Artificial
 5 <220>
 <223> Secuencia sintética
 <400> 288
 gcacagattt cccacactct 20
 10 <210> 289
 <211> 21
 <212> ADN
 <213> Secuencia Artificial
 15 <220>
 <223> Secuencia sintética
 <400> 289
 cttacaggag gtctggcatc a 21
 <210> 290
 <211> 19
 <212> ADN
 25 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 30 <400> 290
 ccacaatgag aaggcagag 19
 <210> 291
 <211> 20
 <212> ADN
 35 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 40 <400> 291
 gctgtggcat agctacactc 20
 <210> 292
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 45 <220>
 <223> Secuencia sintética
 50 <400> 292
 cagggttaatt tgtgggtctg 20
 <210> 293
 <211> 22
 <212> ADN
 <213> Secuencia Artificial
 55 <220>
 <223> Secuencia sintética
 <400> 293
 gggctatcca gaaagataag aa 22
 60 <210> 294
 65

<211> 23
 <212> ADN
 <213> Secuencia Artificial
 5 <220>
 <223> Secuencia sintética
 <400> 294
 cataactggt ggagtatttc act 23
 10 <210> 295
 <211> 24
 <212> ADN
 <213> Secuencia Artificial
 15 <220>
 <223> Secuencia sintética
 <400> 295
 20 ctcaaggaa gaaattcaac aggg 24
 <210> 296
 <211> 22
 <212> ADN
 25 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 <400> 296
 30 catggattca acacagcaaa ca 22
 <210> 297
 <211> 22
 <212> ADN
 35 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 <400> 297
 40 ctctgacctc cttcacttt ac 22
 <210> 298
 <211> 22
 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 <400> 298
 50 gctttcattt gtgctaaacc tc 22
 <210> 299
 <211> 18
 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 <400> 299
 60 catcccagat gccctcat 18
 <210> 300

<211> 22
 <212> ADN
 <213> Secuencia Artificial
 5 <220>
 <223> Secuencia sintética
 <400> 300
 gtagaatcc caaggcaatc ag 22
 10 <210> 301
 <211> 23
 <212> ADN
 <213> Secuencia Artificial
 15 <220>
 <223> Secuencia sintética
 <400> 301
 20 gaacaactaa ctccacagaa ccc 23
 <210> 302
 <211> 19
 <212> ADN
 25 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 30 <400> 302
 gtcgagagga caggaagaa 19
 <210> 303
 <211> 23
 35 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 40 <400> 303
 ggtgctcaa gacatacacc tta 23
 <210> 304
 <211> 20
 <212> ADN
 45 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 50 <400> 304
 ggaaccttg tgacctgga 20
 <210> 305
 <211> 19
 <212> ADN
 55 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 <400> 305
 60 cccagtggta ccttctgaa 19
 <210> 306

<211> 23
 <212> ADN
 <213> Secuencia Artificial
 5 <220>
 <223> Secuencia sintética
 <400> 306
 ctctgttgc ttattgggt aac 23
 10 <210> 307
 <211> 16
 <212> ADN
 <213> Secuencia Artificial
 15 <220>
 <223> Secuencia sintética
 <400> 307
 20 cccactggat gcctcc 16
 <210> 308
 <211> 18
 <212> ADN
 25 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 30 <400> 308
 cggagagacg catctgaa 18
 <210> 309
 <211> 21
 35 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 40 <400> 309
 caggattcc agcttacagg g 21
 <210> 310
 <211> 20
 45 <212> ADN
 <213> Secuencia Artificial
 <220>
 50 <223> Secuencia sintética
 <400> 310
 cttgcaagat gtgccttta 20
 55 <210> 311
 <211> 22
 <212> ADN
 <213> Secuencia Artificial
 60 <220>
 <223> Secuencia sintética
 <400> 311
 65 gggtggttc tctaaacaca aa 22
 <210> 312

<211> 22
 <212> ADN
 <213> Secuencia Artificial

5 <220>
 <223> Secuencia sintética

10 <400> 312
 gcagggtatt gagagaagga tc 22

15 <210> 313
 <211> 22
 <212> ADN
 <213> Secuencia Artificial

20 <220>
 <223> Secuencia sintética

25 <400> 313
 gtgcacattt ctgatgaag gg 22

30 <210> 314
 <211> 22
 <212> ADN
 <213> Secuencia Artificial

35 <220>
 <223> Secuencia sintética

40 <400> 314
 gagcaatgcc tgttcatga ga 22

45 <210> 315
 <211> 24
 <212> ADN
 <213> Secuencia Artificial

50 <220>
 <223> Secuencia sintética

55 <400> 315
 gttaacatta tacagcatgg tggc 24

60 <210> 316
 <211> 19
 <212> ADN
 <213> Secuencia Artificial

65 <220>
 <223> Secuencia sintética

70 <400> 316
 gcagaacatg tctgaagc 19

75 <210> 317
 <211> 20
 <212> ADN
 <213> Secuencia Artificial

80 <220>
 <223> Secuencia sintética

85 <400> 317
 cagctgttc ccaaaccat 20

90 <210> 318

<211> 21
 <212> ADN
 <213> Secuencia Artificial
 5 <220>
 <223> Secuencia sintética
 <400> 318
 caaagtgtgg aagttgcttc c 21
 10 <210> 319
 <211> 18
 <212> ADN
 <213> Secuencia Artificial
 15 <220>
 <223> Secuencia sintética
 <400> 319
 20 ggtcgacttt gtccatcc 18
 <210> 320
 <211> 19
 <212> ADN
 25 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 30 <400> 320
 ctctgttgcc tgtggactc 19
 <210> 321
 <211> 24
 35 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 40 <400> 321
 ctaactagaa ttagtctgcc tgcc 24
 <210> 322
 <211> 19
 <212> ADN
 <213> Secuencia Artificial
 45 <220>
 <223> Secuencia sintética
 50 <400> 322
 gtgagccata atcgtgtca 19
 <210> 323
 <211> 23
 <212> ADN
 <213> Secuencia Artificial
 55 <220>
 <223> Secuencia sintética
 <400> 323
 60 gagaattaat gctccctctc ctg 23
 65 <210> 324

<211> 19
 <212> ADN
 <213> Secuencia Artificial
 5 <220>
 <223> Secuencia sintética
 <400> 324
 gtggctgct gttgaccaa 19
 10 <210> 325
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 15 <220>
 <223> Secuencia sintética
 <400> 325
 20 ggttgcaact gctgatctat 20
 <210> 326
 <211> 18
 <212> ADN
 25 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 30 <400> 326
 ctttcccagt caaggcag 18
 <210> 327
 <211> 20
 35 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 40 <400> 327
 gccgaagtc tccgaaggat 20
 <210> 328
 <211> 20
 45 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 50 <400> 328
 gacagactga cggagctca 20
 <210> 329
 <211> 20
 55 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 60 <400> 329
 tgactctgcc gcacatgatc 20
 65 <210> 330

<211> 20
 <212> ADN
 <213> Secuencia Artificial
 5 <220>
 <223> Secuencia sintética
 <400> 330
 gatccgatag ccctctgcag 20
 10 <210> 331
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 15 <220>
 <223> Secuencia sintética
 <400> 331
 caacaggcct gctaaacacc 20
 20 <210> 332
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 25 <220>
 <223> Secuencia sintética
 <400> 332
 ggtcaacaac cgagggactc 20
 30 <210> 333
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 35 <220>
 <223> Secuencia sintética
 <400> 333
 ttgtcattaa tgctggcggc 20
 40 <210> 334
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 45 <220>
 <223> Secuencia sintética
 <400> 334
 cggtgacggt ttgcaactt 20
 50 <210> 335
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 55 <220>
 <223> Secuencia sintética
 <400> 335
 gtccggcagt taagggtctc 20
 60 <210> 336

<211> 20
 <212> ADN
 <213> Secuencia Artificial
 5 <220>
 <223> Secuencia sintética
 <400> 336
 tactcacaaa cgactgca 20
 10 <210> 337
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 15 <220>
 <223> Secuencia sintética
 <400> 337
 20 cgtatatggc cgactggagg 20
 <210> 338
 <211> 20
 <212> ADN
 25 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 <400> 338
 30 tagggtttgc ggcgataagg 20
 <210> 339
 <211> 20
 <212> ADN
 35 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 <400> 339
 40 ccaagtcaac cacccgagac 20
 <210> 340
 <211> 20
 <212> ADN
 45 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 <400> 340
 50 actccctgg cctcctct 20
 <210> 341
 <211> 20
 <212> ADN
 55 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 <400> 341
 60 gcttgggtcc tctcctgaac 20
 65 <210> 342

<211> 20
 <212> ADN
 <213> Secuencia Artificial
 5 <220>
 <223> Secuencia sintética
 <400> 342
 aacgtccgaa ccacaatgct 20
 10 <210> 343
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 15 <220>
 <223> Secuencia sintética
 <400> 343
 ctctctgcat ccaacagtcg 20
 <210> 344
 <211> 20
 <212> ADN
 25 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 <400> 344
 ccaccgtagc actccttct 20
 <210> 345
 <211> 20
 <212> ADN
 35 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 <400> 345
 cggagcgtcg gtagtgtaaa 20
 <210> 346
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 <400> 346
 acaactgac gaacctaccg 20
 <210> 347
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 <400> 347
 tggccatcc ttatgtgctg 20
 <210> 348

<211> 20
 <212> ADN
 <213> Secuencia Artificial
 5 <220>
 <223> Secuencia sintética
 <400> 348
 ggtcgttatt gctcaagccc 20
 10 <210> 349
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 15 <220>
 <223> Secuencia sintética
 <400> 349
 ttgattctgg ccctcccatc 20
 20 <210> 350
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 25 <220>
 <223> Secuencia sintética
 <400> 350
 ctcacgccgg ctatttaggt 20
 30 <210> 351
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 35 <220>
 <223> Secuencia sintética
 <400> 351
 agtctgggta ggtggaggac 20
 40 <210> 352
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 45 <220>
 <223> Secuencia sintética
 <400> 352
 cgactgagcc acatccaact 20
 50 <210> 353
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 55 <220>
 <223> Secuencia sintética
 <400> 353
 gagcctcagc cgaattgaa 20
 60 <210> 354

<211> 20
 <212> ADN
 <213> Secuencia Artificial
 5 <220>
 <223> Secuencia sintética
 <400> 354
 ttgccattct gcaccaatgc 20
 10 <210> 355
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 15 <220>
 <223> Secuencia sintética
 <400> 355
 20 tattggtgtt cgcggtgat 20
 <210> 356
 <211> 20
 <212> ADN
 25 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 30 <400> 356
 atgggcgtaa caggaggact 20
 <210> 357
 <211> 20
 35 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Secuencia sintética
 40 <400> 357
 ggaatggcct acctgcatca 20
 <210> 358
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 45 <220>
 <223> Secuencia sintética
 50 <400> 358
 cccgtgttg tcatgcatc 20
 <210> 359
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 55 <220>
 <223> Secuencia sintética
 <400> 359
 60 gttcgatgcg tccatgagt 20
 <210> 360

<211> 20
 <212> ADN
 <213> Secuencia Artificial
 5 <220>
 <223> Secuencia sintética
 <400> 360
 caaccgcggt agatgttct 20
 10 <210> 361
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 15 <220>
 <223> Secuencia sintética
 <400> 361
 gccagctcaa gagttagcc 20
 20 <210> 362
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 25 <220>
 <223> Secuencia sintética
 <400> 362
 ttcttgatcc tgcgcatgt 20
 30 <210> 363
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 35 <220>
 <223> Secuencia sintética
 <400> 363
 atcgcaggcg ttcctatac 20
 40 <210> 364
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 45 <220>
 <223> Secuencia sintética
 <400> 364
 tattggacct cggaccacga 20
 50 <210> 365
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 55 <220>
 <223> Secuencia sintética
 <400> 365
 agccaccatt tagatccgcg 20
 60 <210> 366

<211> 20
 <212> ADN
 <213> Secuencia Artificial
 5 <220>
 <223> Secuencia sintética
 <400> 366
 gaccagtaga agtctgcccg 20
 10 <210> 367
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 15 <220>
 <223> Secuencia sintética
 <400> 367
 tttcagaatg gccgagctgt 20
 20 <210> 368
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 25 <220>
 <223> Secuencia sintética
 <400> 368
 aggtgacctt cttgtacgcc 20
 30 <210> 369
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 35 <220>
 <223> Secuencia sintética
 <400> 369
 ggcggtcct tattccatc 20
 40

REIVINDICACIONES

1. Un método para detectar una variación genética en una muestra genética de un sujeto, que comprende

5 poner en contacto el primer y el segundo conjuntos de sondas con la muestra genética, en donde el primer conjunto de sondas comprende una primera sonda de marcaje y una primera sonda de etiquetado, y el segundo conjunto de sondas comprende una segunda sonda de marcaje y una segunda sonda de etiquetado,

hibridar al menos partes del primer y el segundo conjuntos de sondas con la primera y la segunda regiones de ácido nucleico de interés en moléculas de nucleótidos de la muestra genética, respectivamente,

10 ligar el primer conjunto de sondas ligando al menos la primera sonda de marcaje y la primera sonda de etiquetado,

ligar el segundo conjunto de sondas ligando al menos la segunda sonda de marcaje y la segunda sonda de etiquetado,

opcionalmente amplificar los conjuntos de sondas ligadas,

inmovilizar las sondas de etiquetado y/o las sondas de etiquetado amplificadas de las mismas en una ubicación predeterminada sobre un sustrato, en donde

15 la primera y la segunda sondas de marcaje y/o las sondas de marcaje amplificadas de las mismas ligadas a las sondas de etiquetado inmovilizadas y/o las sondas de etiquetado amplificadas de las mismas comprenden una primera y una segunda marcas, respectivamente,

la primera y la segunda marcas son diferentes,

las marcas inmovilizadas se pueden resolver ópticamente,

20 la primera y la segunda sondas de etiquetado inmovilizadas y/o las sondas de etiquetado amplificadas de las mismas comprenden una primera y una segunda etiquetas, respectivamente, y

la etapa de inmovilización se realiza inmovilizando las etiquetas en la ubicación predeterminada,

contar (i) un primer número de la primera marca inmovilizada sobre el sustrato, y (ii) un segundo número de la segunda marca inmovilizada sobre el sustrato, y

25 comparar el primer y el segundo números para determinar la variación genética en la muestra genética.

2. El método según la reivindicación 1, que comprende adicionalmente marcar la primera y la segunda sondas de marcaje con la primera y la segunda marcas, respectivamente, antes de la etapa de contacto; y/o que comprende etiquetar la primera y la segunda sondas de etiquetado con la primera y la segunda etiquetas, respectivamente, antes de la etapa de contacto.

30 3. El método de acuerdo con cualquiera de las reivindicaciones 1-2, en donde el método comprende amplificar los conjuntos de sondas ligadas con o sin marcar las sondas durante la amplificación.

4. El método según la reivindicación 1, en donde

35 cada una de las primeras y las segundas sondas de marcaje comprende una secuencia de cebado directo o inverso, y cada una de las primeras y las segundas sondas de etiquetado comprende una secuencia de cebado inversa o directa correspondiente y una secuencia de nucleótidos de etiquetado como etiqueta,

el método comprende amplificar los conjuntos de sondas ligadas,

40 la etapa de amplificación comprende amplificar (i) las primeras sondas de marcaje y etiquetado ligadas con los primeros cebadores directo e inverso que hibridan con las secuencias de cebado directo e inverso, respectivamente, en donde el primer cebador directo o inverso que hibrida con la primera sonda de marcaje comprende la primera marca, y (ii) las segundas sondas de marcaje y etiquetado ligadas con segundos cebadores directos e inversos que hibridan con las secuencias de cebado directo e inverso, respectivamente, en donde el segundo cebador directo o inverso que hibrida con la segunda sonda de marcaje comprende la segunda marca,

45 las secuencias de nucleótidos de etiquetado amplificadas de las sondas de etiquetado se inmovilizan en una ubicación predeterminada sobre un sustrato, en donde las secuencias de nucleótidos de etiquetado amplificadas de la primera y la segunda sondas de etiquetado son la primera y la segunda etiquetas,

el primer número es el número de la primera marca en el primer conjunto de sondas amplificadas inmovilizadas sobre el sustrato, y el segundo número es el número de la segunda marca en el segundo conjunto de sondas amplificadas inmovilizadas sobre el sustrato.

5. El método según cualquiera de las reivindicaciones 1-4, en donde la etapa de amplificación comprende poner en contacto una exonucleasa con la sonda amplificada, digiriendo el extremo 5' del conjunto de sondas amplificadas que no tiene ninguna marca en el extremo 5', en donde el extremo 5' del conjunto de sondas amplificadas que comprenden la marca en el extremo 5' está protegido de la digestión con exonucleasa.
- 5 6. El método según cualquiera de las reivindicaciones 1-5, en donde la variación genética determinada indica la presencia o ausencia de cáncer, la variabilidad farmacocinética, la toxicidad del fármaco, el rechazo de trasplante o la aneuploidía en el sujeto.
7. El método según cualquiera de las reivindicaciones 1-5, en donde el sujeto es un sujeto gestante, y la variación genética es una variación genética en el feto del sujeto gestante.
- 10 8. El método según cualquiera de las reivindicaciones 1-5, en donde el sujeto es un sujeto gestante y la variación genética se selecciona del grupo que consiste en la trisomía 13, la trisomía 18, la trisomía 21, la aneuploidía de X y la aneuploidía de Y en el feto del sujeto gestante.
9. El método según cualquiera de las reivindicaciones 1-8, en donde la muestra genética se selecciona del grupo que consiste en sangre completa, suero, plasma, orina, saliva, sudor, materia fecal y lágrimas.
- 15 10. El método según cualquiera de las reivindicaciones 1-9, en donde
 la primera sonda de marcaje y la primera sonda de etiquetado hibridan con la primera región de ácido nucleico de interés, y
 la segunda sonda de marcaje y las segundas sondas de etiquetado hibridan con la segunda región de ácido nucleico de interés.
- 20 11. El método según cualquiera de las reivindicaciones 1-10, en donde la etapa de recuento comprende el filtrado espacial; o la etapa de recuento comprende el análisis de segmentación de imágenes.
12. El método según cualquiera de las reivindicaciones 1-11, en donde la etapa de comparación comprende obtener una estimación de un número relativo de las moléculas de nucleótido que tienen la primera y segunda regiones de ácido nucleico de interés.
- 25 13. El método según cualquiera de las reivindicaciones 1-12, en donde las marcas inmovilizadas del mismo tipo están separadas por una distancia de al menos 250 nm en ambas dimensiones.
14. El método según cualquiera de las reivindicaciones 1-13, en donde las marcas son colorantes fluorescentes.
15. El método según cualquiera de las reivindicaciones 1-14, en donde el sustrato comprende un compañero de unión que se pone en contacto, e inmoviliza las etiquetas.
- 30 16. El método según cualquiera de las reivindicaciones 1-15, en donde la etapa de inmovilización comprende hibridar al menos una parte de la etiqueta o secuencia de nucleótidos de etiquetado con una molécula de nucleótido correspondiente inmovilizada sobre el sustrato.
17. El método según cualquiera de las reivindicaciones 1-16, en donde las sondas de etiquetado se inmovilizan en una matriz de múltiples ubicaciones predeterminadas sobre un sustrato.
- 35 18. El método según cualquiera de las reivindicaciones 1-17, en donde el sujeto es un sujeto gestante, la variación genética es una variación genética en el feto del sujeto gestante y el método comprende adicionalmente
 poner en contacto los conjuntos de sondas maternas y paternas con la muestra genética, en donde el conjunto de sondas maternas comprende una sonda de marcaje materna y una sonda de etiquetado materna, y el conjunto de sondas paternas comprende una sonda de marcaje paterna y una sonda de etiquetado paterna,
- 40 hibridar al menos una parte de cada uno de los conjuntos de sondas maternas y paternas con una región de ácido nucleico de interés en las moléculas de nucleótidos de la muestra genética, comprendiendo la región de ácido nucleico de interés un sitio predeterminado de Polimorfismo de un Solo Nucleótido (SNP), en donde
 la al menos una parte del conjunto de sondas maternas hibrida con un primer alelo en el sitio SNP, la al menos una parte del conjunto de sondas paternas hibrida con un segundo alelo en el sitio SNP, y el primer y
- 45 segundo alelos son diferentes entre sí,
 ligar los conjuntos de sondas maternas y paternas ligando al menos (i) las sondas de marcaje y etiquetado maternas, y (ii) las sondas de marcaje y etiquetado paternas,
 opcionalmente amplificar los conjuntos de sondas ligadas,
 inmovilizar las sondas de etiquetado y/o las sondas de etiquetado amplificadas de las mismas en una

ubicación predeterminada sobre un sustrato, en donde

las sondas de marcaje maternas y paternas y/o las sondas de marcaje amplificadas de las mismas ligadas a las sondas de etiquetado inmovilizadas y/o las sondas de etiquetado amplificadas de las mismas comprenden marcas maternas y paternas, respectivamente,

5 las marcas maternas y paternas son diferentes, y

las marcas inmovilizadas se pueden resolver ópticamente,

contar los números de las marcas maternas y paternas, y

determinar si una razón de material fetal en la muestra genética es suficiente para detectar la variación genética en el feto en función de los números de las marcas maternas y paternas.

10 19. El método según cualquiera de las reivindicaciones 1-18, en donde el sujeto es un sujeto gestante, la variación genética es una variación genética en el feto del sujeto gestante y el método comprende adicionalmente

poner en contacto conjuntos de sondas del alelo A y B que son específicos del alelo para la muestra genética, en donde el conjunto de sondas del alelo A comprende una sonda de marcaje del alelo A y una sonda de etiquetado del alelo A, y el conjunto de sondas del alelo B comprende una sonda de marcaje del alelo B y una sonda de etiquetado del alelo B,

15 hibridar al menos una parte de cada uno de los conjuntos de sondas del alelo A y el alelo B con una región de ácido nucleico de interés en moléculas de nucleótidos de la muestra genética, comprendiendo la región de ácido nucleico de interés un sitio de polimorfismo de un solo nucleótido (SNP) predeterminado para el cual el perfil alélico materno difiere de un perfil alélico fetal en el sitio SNP, en donde

20 la al menos una parte del conjunto de sondas del alelo A hibrida con un primer alelo en el sitio SNP, la al menos una parte del conjunto de sondas del alelo B hibrida con un segundo alelo en el sitio SNP, y el primer y segundo alelos son diferentes entre sí,

ligar los conjuntos de sondas del alelo A y B ligando al menos (i) las sondas de marcaje y etiquetado del alelo A, y (ii) las sondas de marcaje y etiquetado del alelo B,

25 opcionalmente amplificar los conjuntos de sondas ligadas,

inmovilizar las sondas de etiquetado y/o las sondas de etiquetado amplificadas de las mismas en una ubicación predeterminada sobre un sustrato, en donde

las sondas de marcaje del alelo A y el alelo B y/o las sondas de marcaje amplificadas de las mismas ligadas a las sondas de etiquetado inmovilizadas y/o las sondas de etiquetado amplificadas de las mismas comprenden marcas del alelo A y el alelo B, respectivamente,

30 las marcas del alelo A y el alelo B son diferentes, y

las marcas inmovilizadas se pueden resolver ópticamente,

contar los números de las marcas del alelo A y del alelo B, y

35 determinar si una razón de un material fetal en la muestra genética es suficiente para detectar la variación genética en el feto en función de los números de las marcas del alelo A y del alelo B.

20. El método según cualquiera de las reivindicaciones 1-19, en donde el método excluye la secuenciación de las regiones de ácido nucleico de interés y/o las sondas.

21. El método según cualquiera de las reivindicaciones 1-20, en donde el método excluye la lectura de matriz global o la cuantificación analógica.

40 22. El método según cualquiera de las reivindicaciones 1-21, en donde la etapa de recuento comprende

medir señales ópticas de las marcas inmovilizadas, y

calibrar los números contados que distinguen una señal óptica de una marca única del resto de las señales ópticas del fondo y/o múltiples marcas.

45 23. El método según cualquiera de las reivindicaciones 1-22, en donde la etapa de recuento comprende adicionalmente confirmar el recuento, en donde la confirmación comprende

poner en contacto el primer y segundo conjunto de sondas de control con la muestra genética, en donde el

primer conjunto de sondas de control comprende una primera sonda de marcaje de control y la primera sonda de etiquetado, y el segundo conjunto de sondas de control comprende una segunda sonda de marcaje de control y la segunda sonda de etiquetado,

5 hibridar al menos una parte del primer y el segundo conjuntos de sondas de control con la primera y la segunda regiones de ácido nucleico de interés en las moléculas de nucleótidos de la muestra genética, respectivamente,

 ligar el primer conjunto de sondas de control ligando al menos la primera sonda de marcaje de control y la primera sonda de etiquetado,

10 ligar el segundo conjunto de sondas de control ligando al menos la segunda sonda de marcaje de control y la segunda sonda de etiquetado,

 opcionalmente amplificar los conjuntos de sondas ligadas,

 inmovilizar cada una de las sondas de etiquetado y/o las sondas de etiquetado amplificadas de las mismas en una ubicación predeterminada sobre un sustrato, en donde

15 la primera y segunda sondas de marcaje de control y/o las sondas de marcaje amplificadas de las mismas ligadas a las sondas de etiquetado inmovilizadas y/o las sondas de etiquetado amplificadas de las mismas comprenden la primera y segunda marcas de control, respectivamente,

 la primera y segunda marcas de control son diferentes, y

 las marcas inmovilizadas se pueden resolver ópticamente,

 medir las señales ópticas de las marcas de control inmovilizadas sobre el sustrato, y

20 comparar las señales ópticas de la primera y la segunda marcas de control inmovilizadas con las señales ópticas de la primera y la segunda marcas inmovilizadas para determinar si existe un error basado en las marcas.

24. El método según la reivindicación 23, en donde la primera marca y la segunda marca de control son iguales, y la segunda marca y la primera marca de control son iguales.

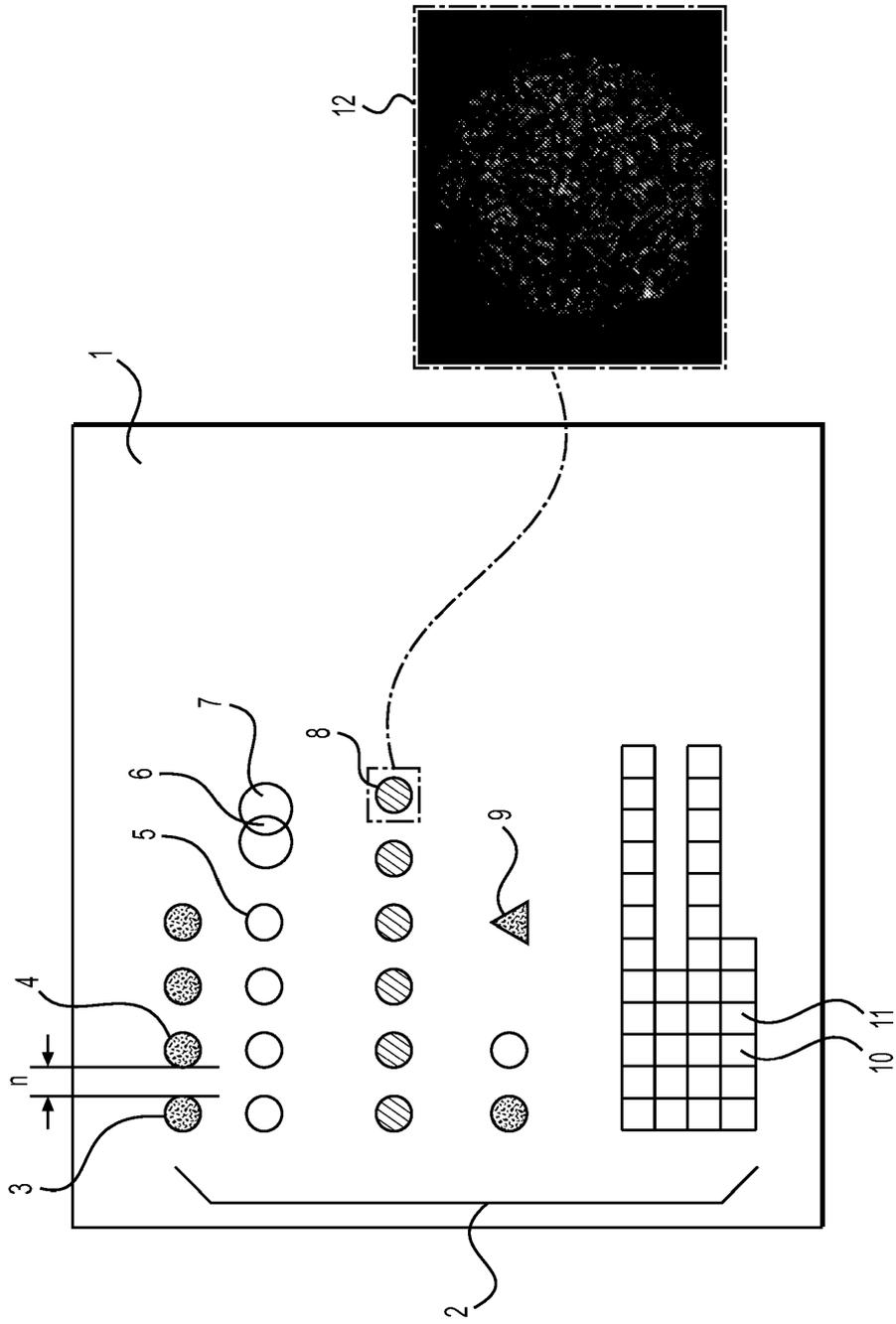


Figure 1

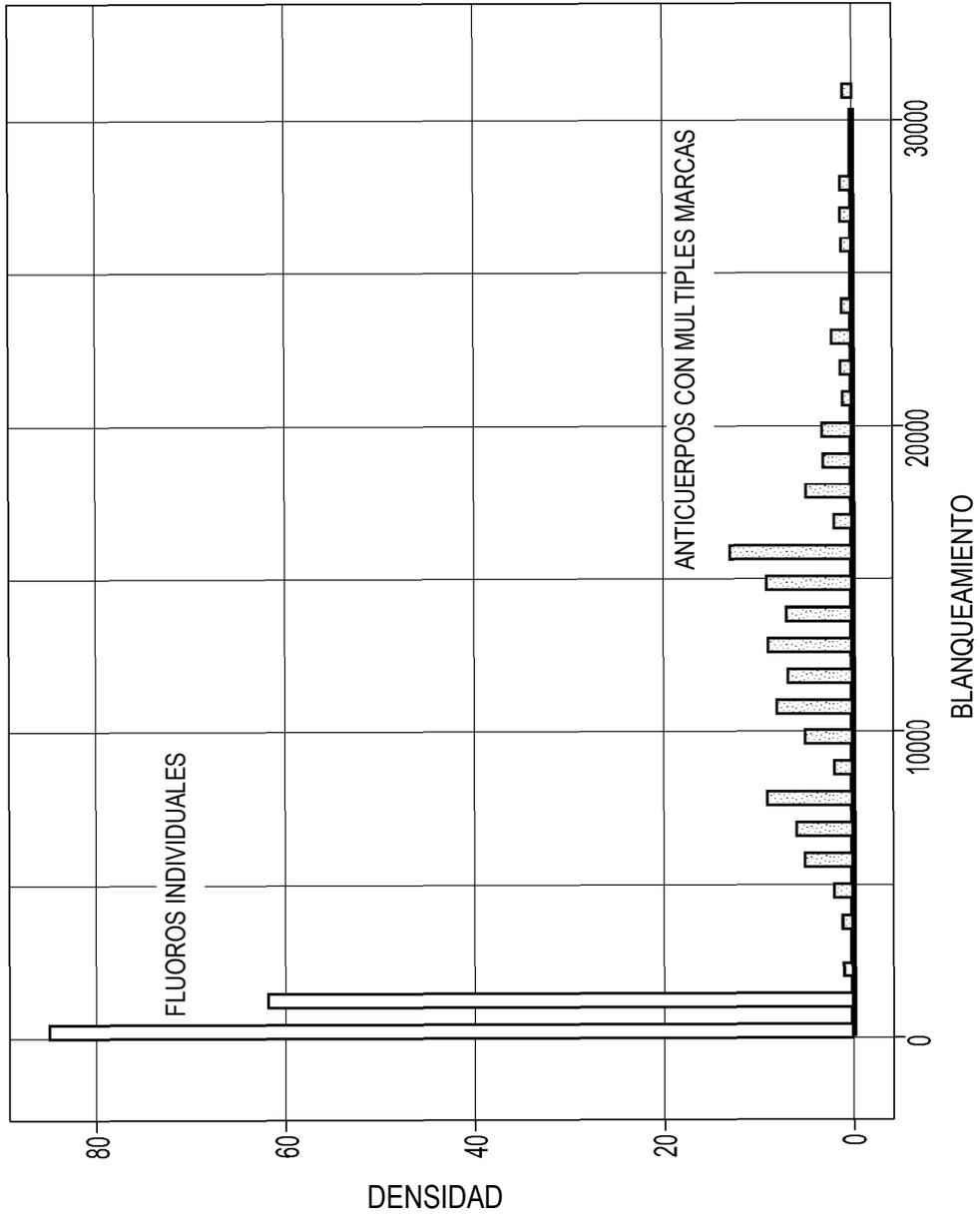


Figura 2

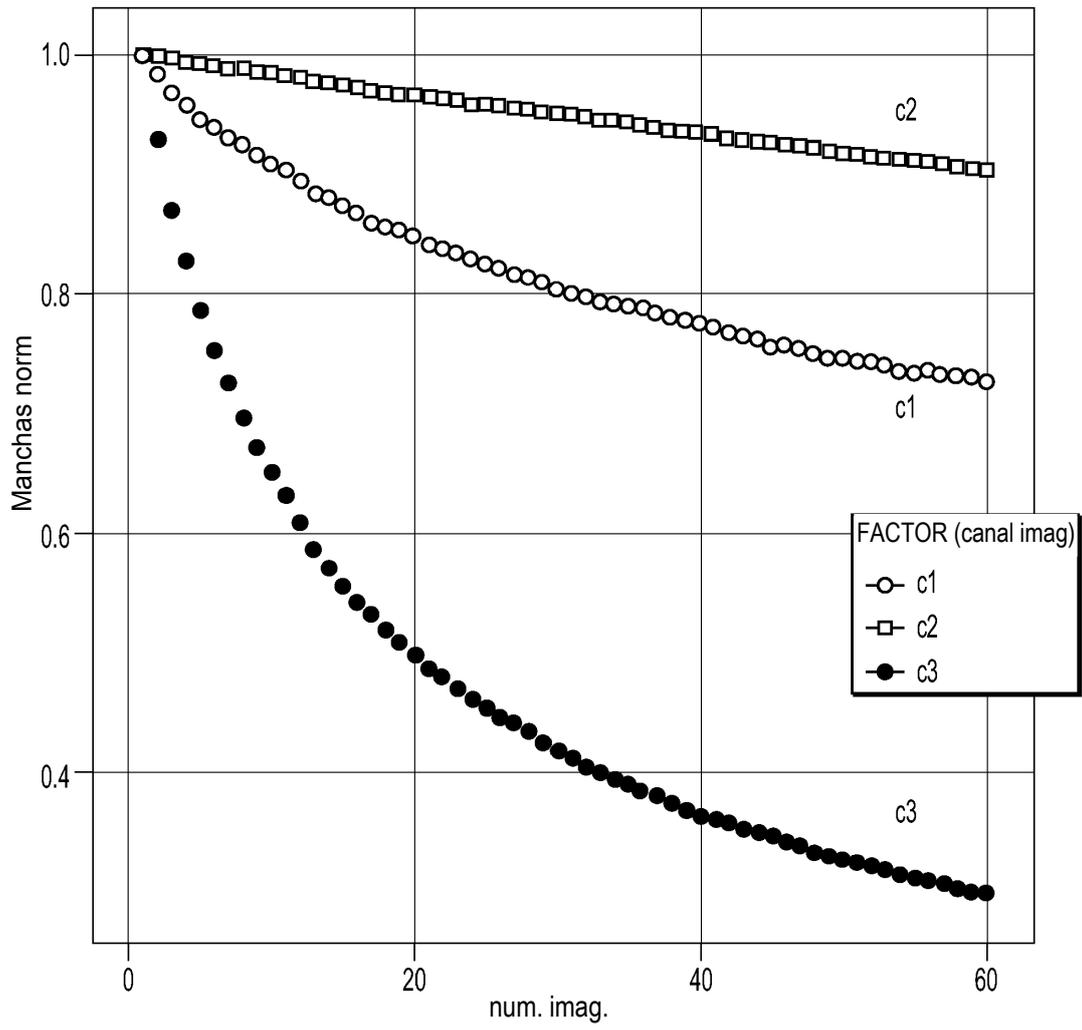


Figura 3

Figura 4

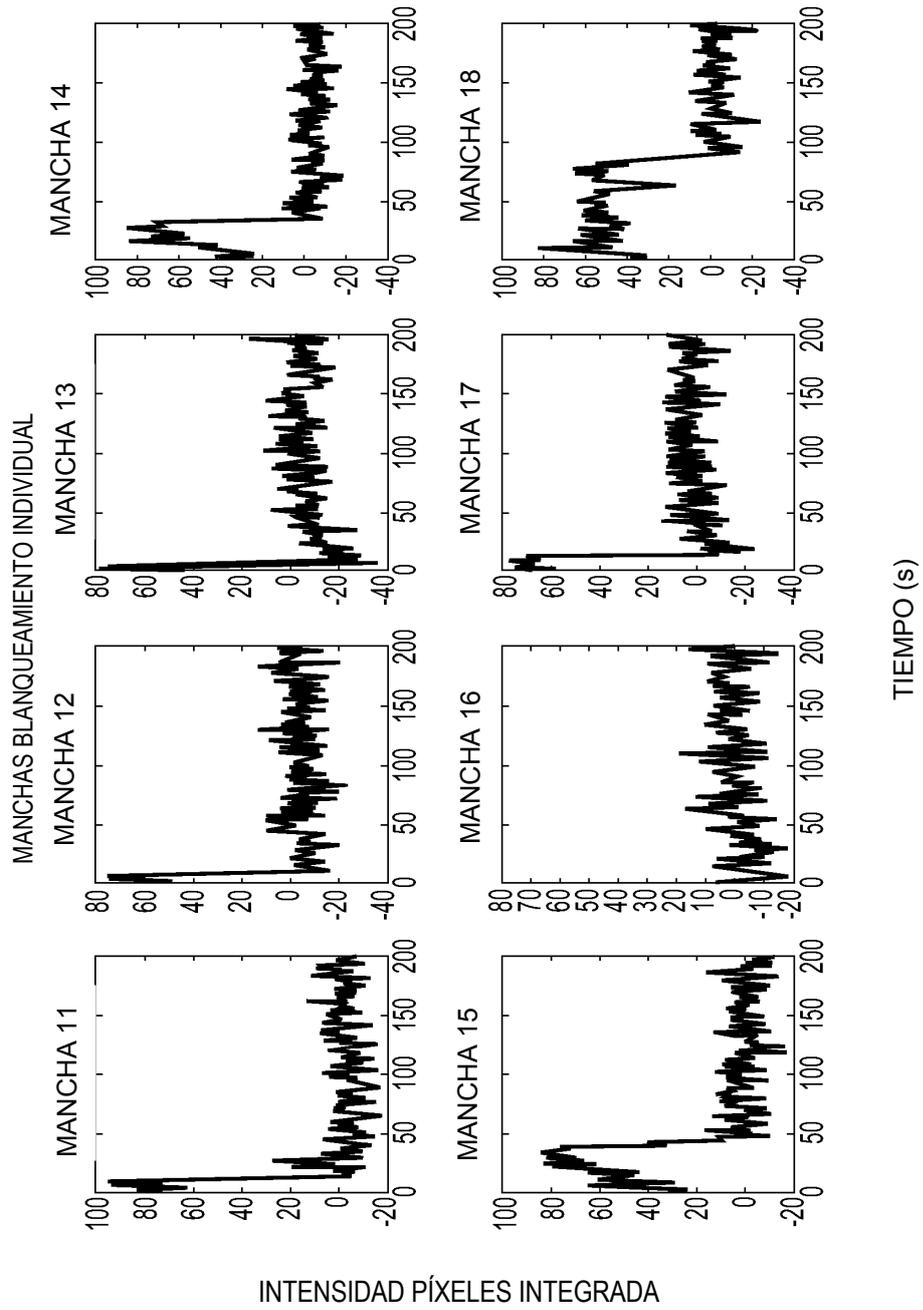


Figura 5

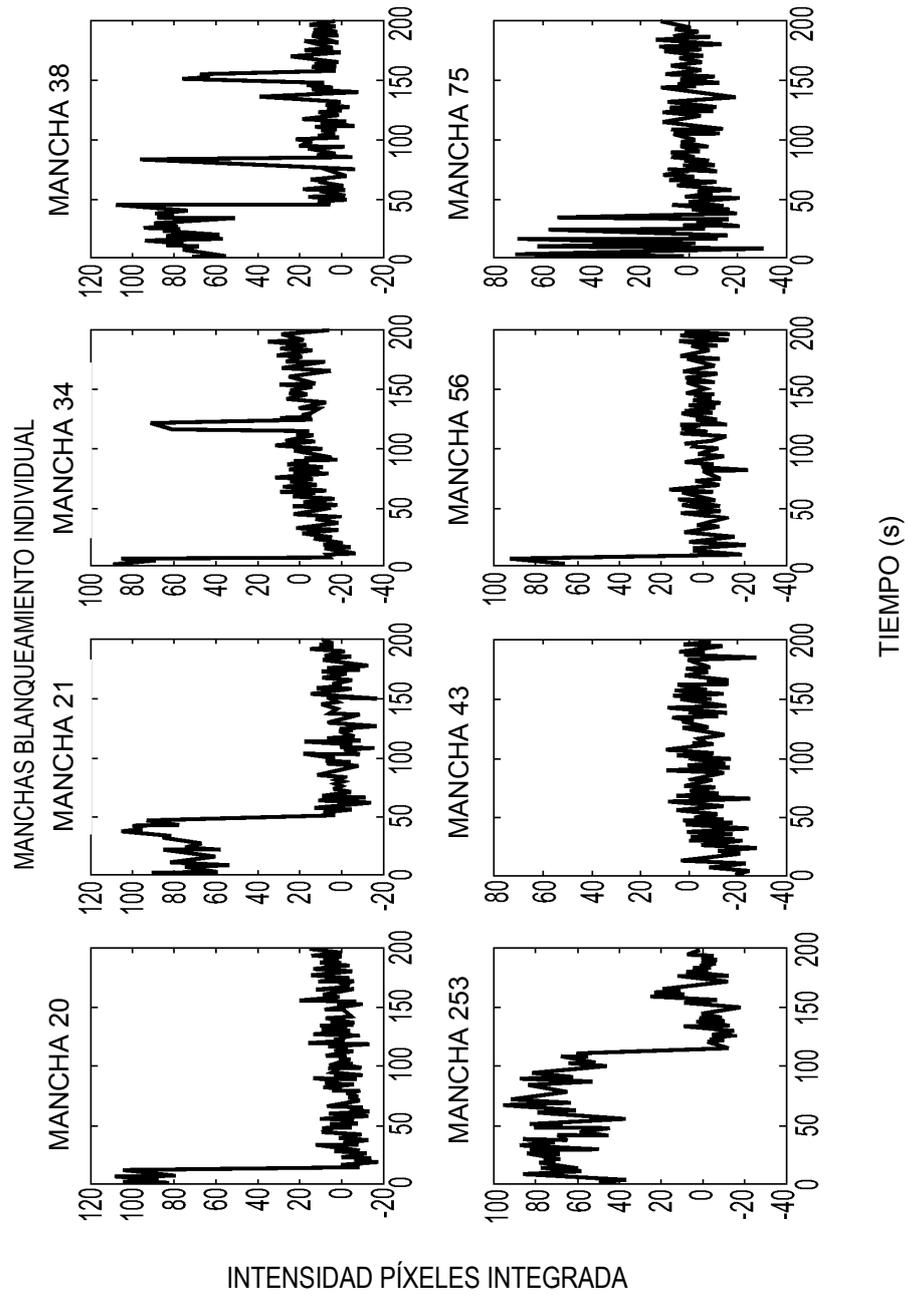


Figura 6

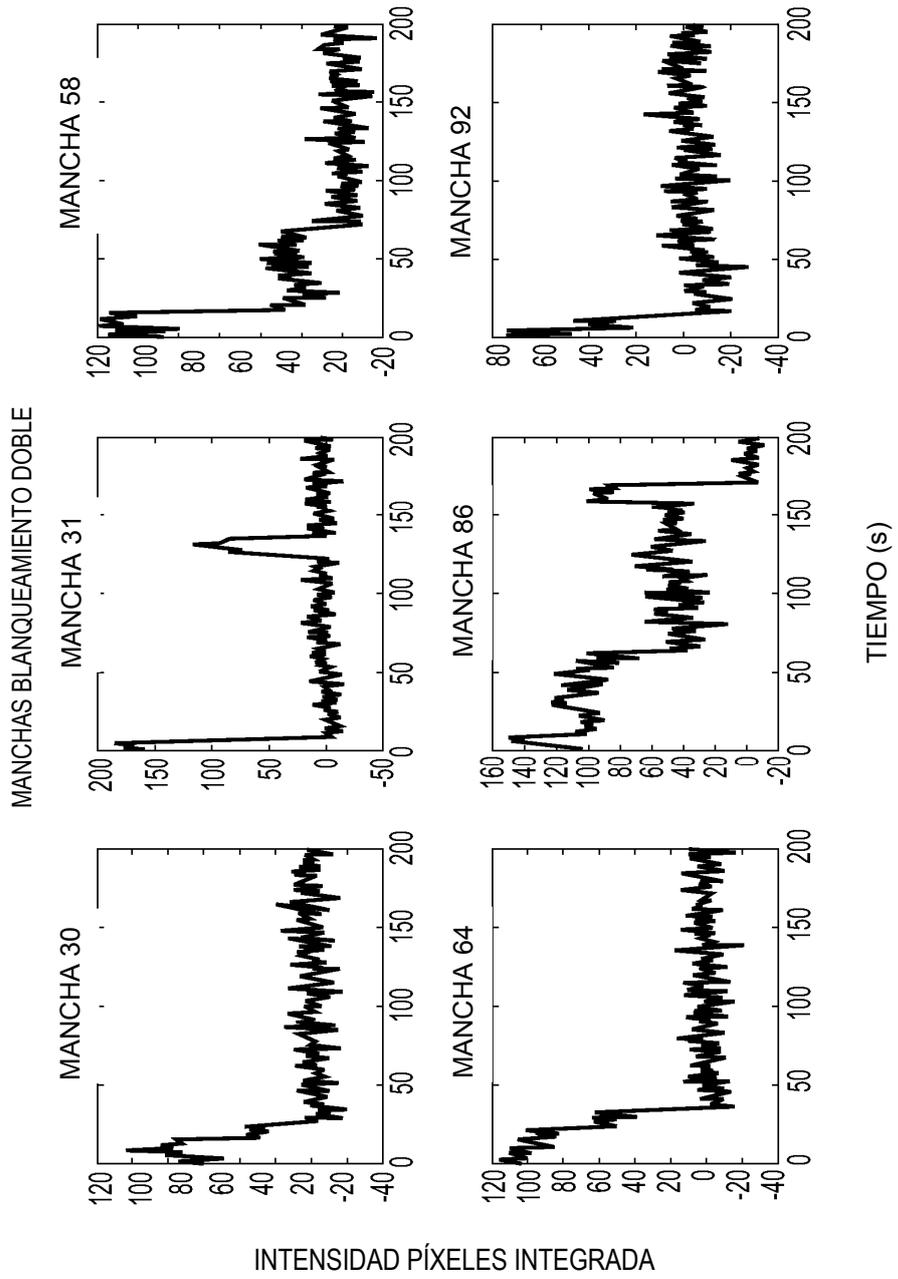


Figura 7

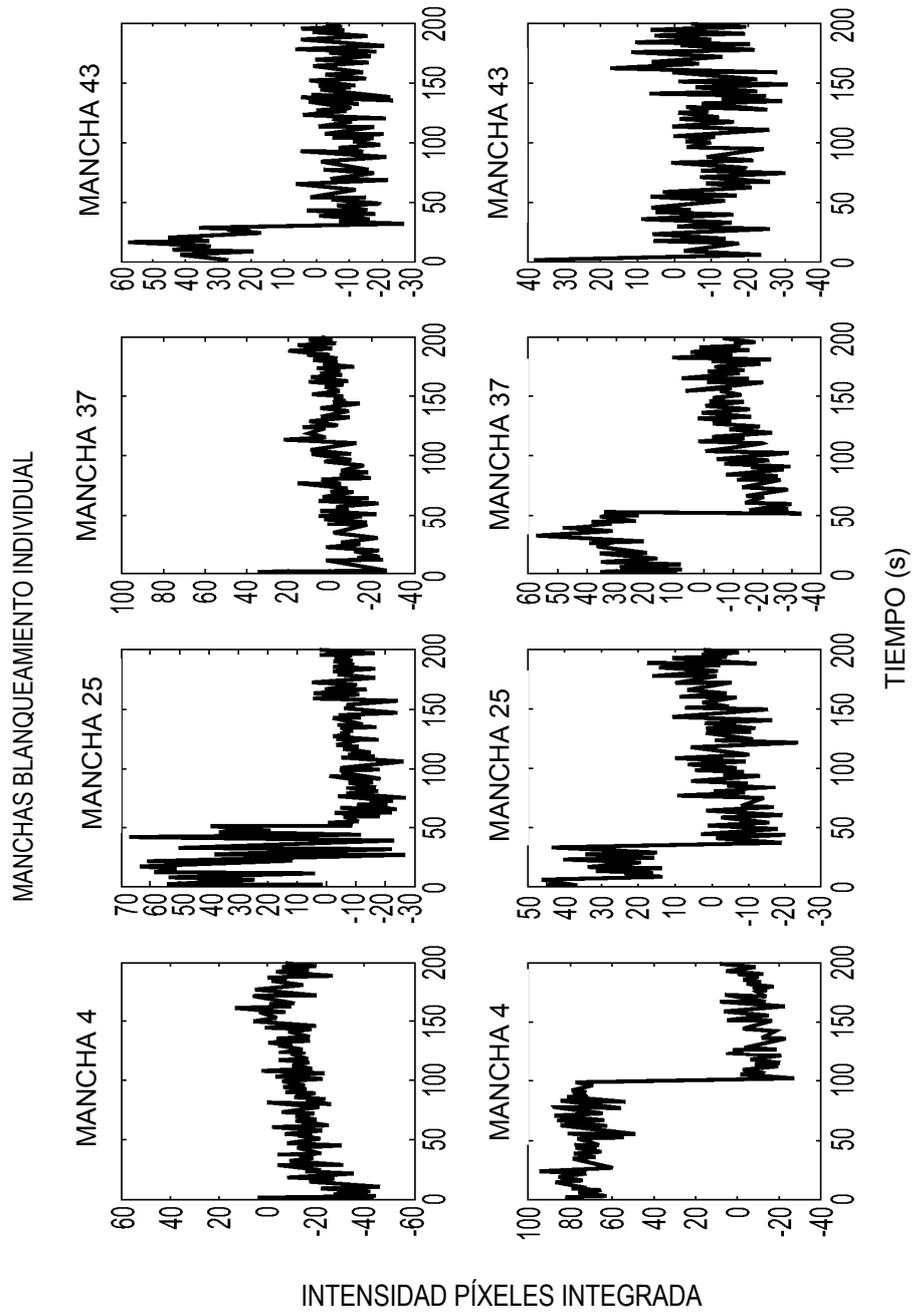


Figura 8

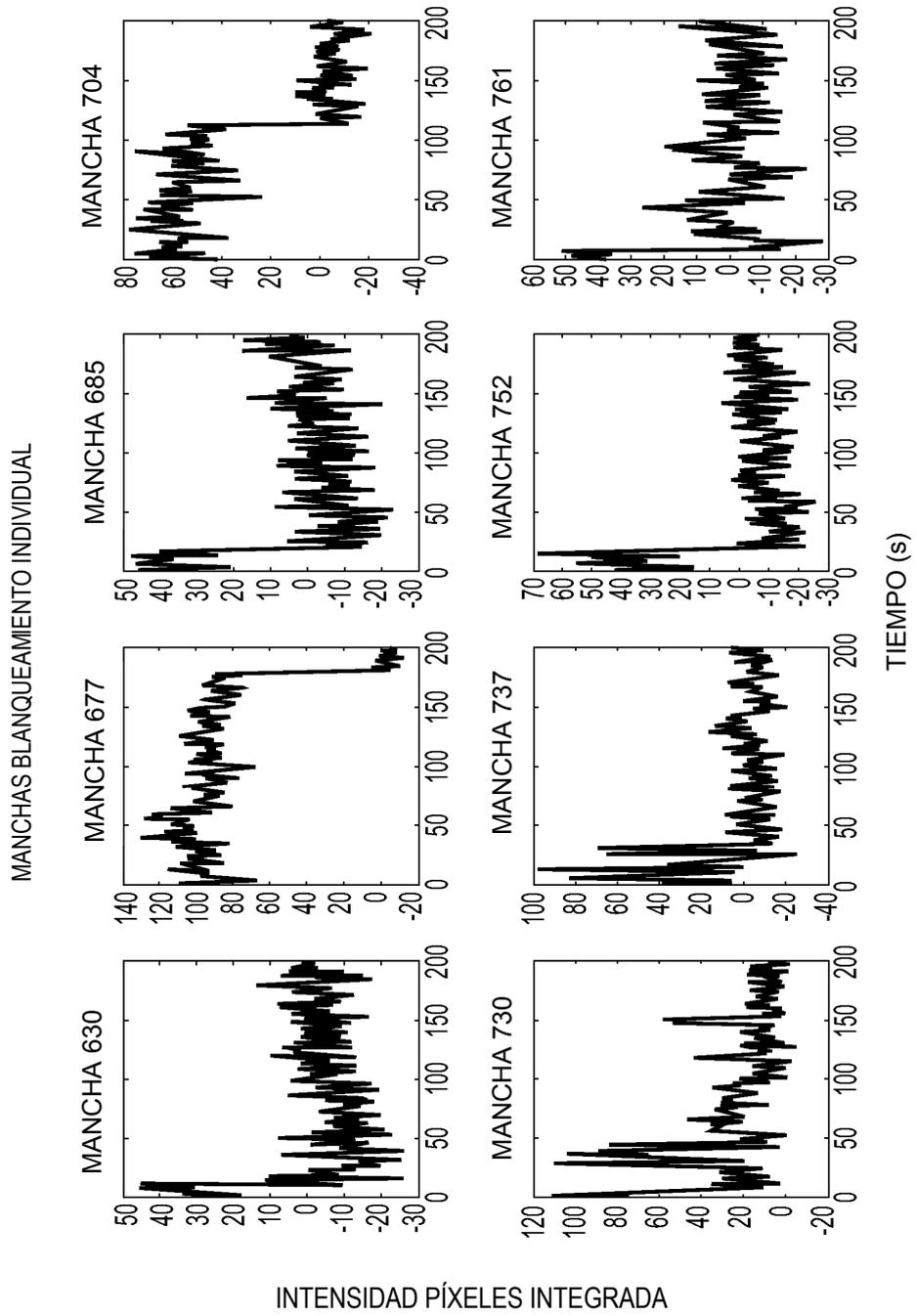


Figura 9

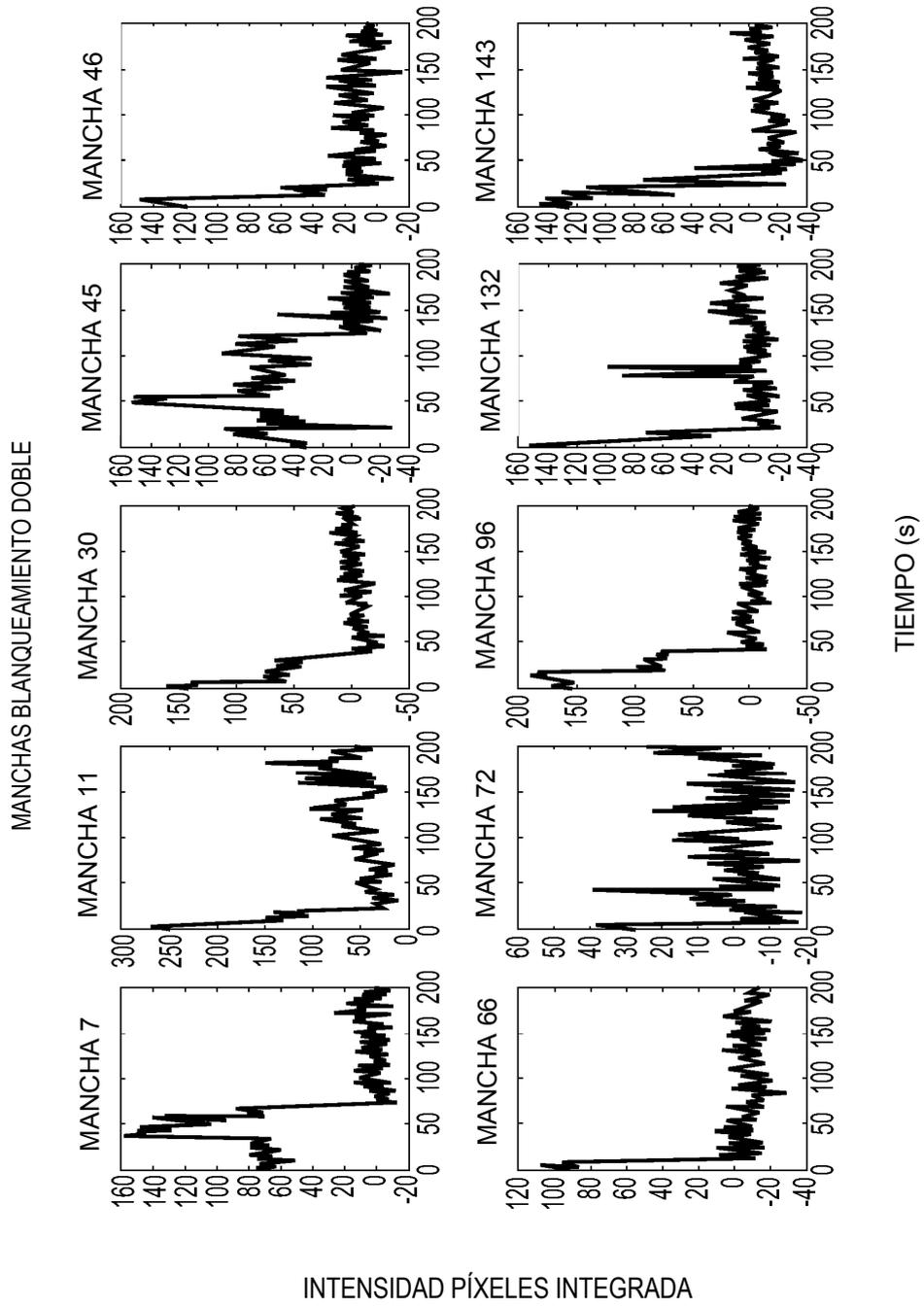


Figura 10

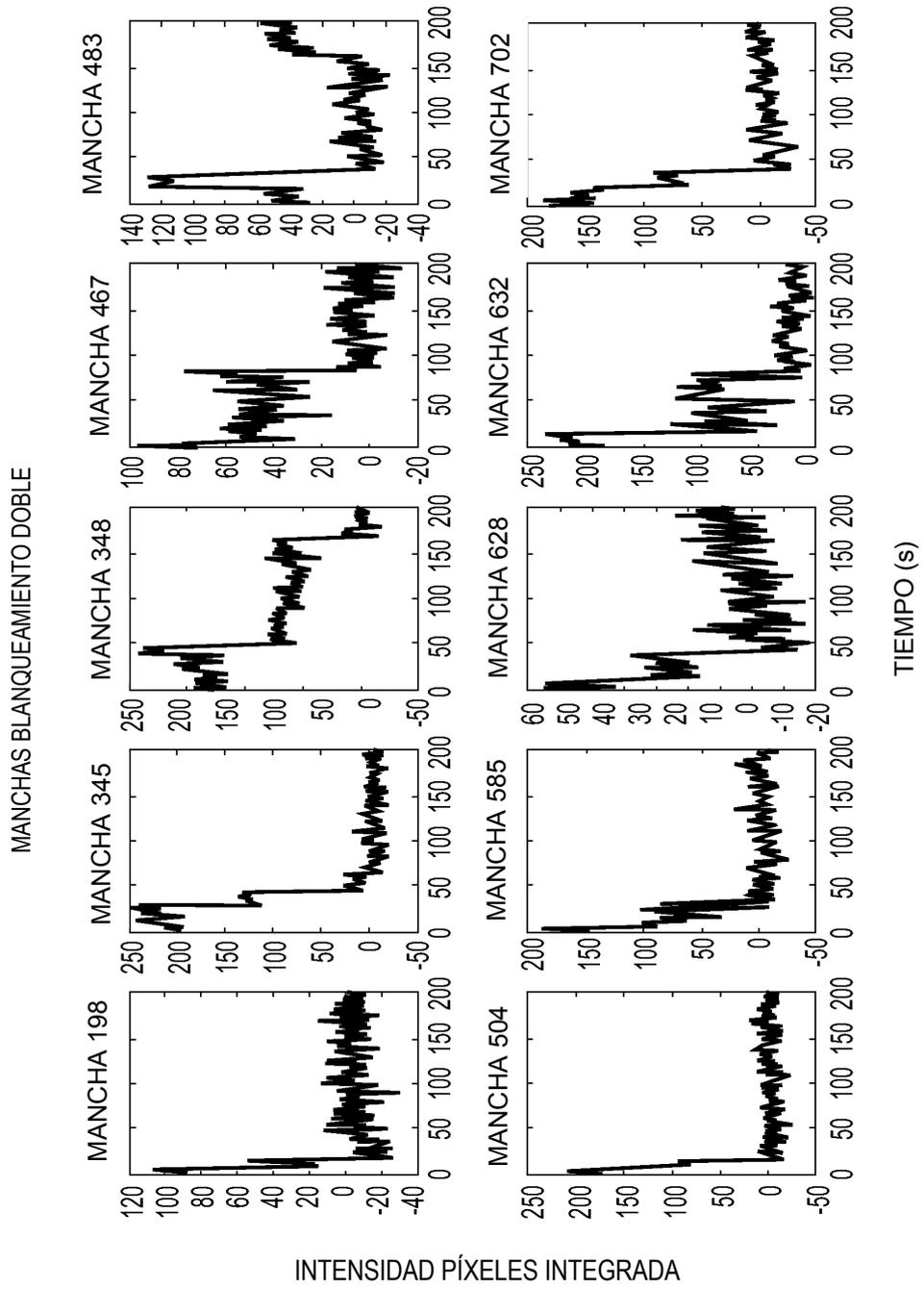


Figura 11

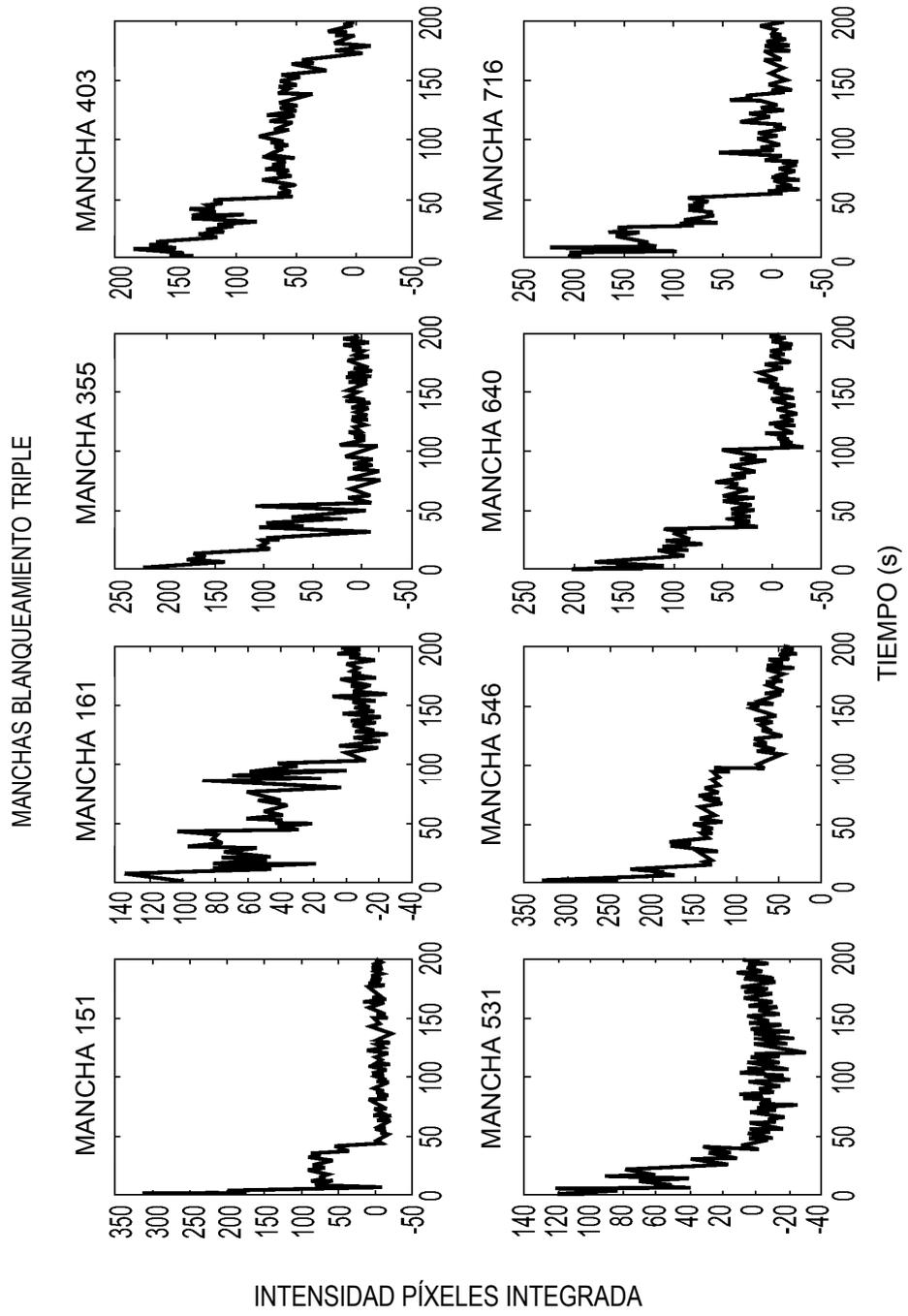


Figura 12

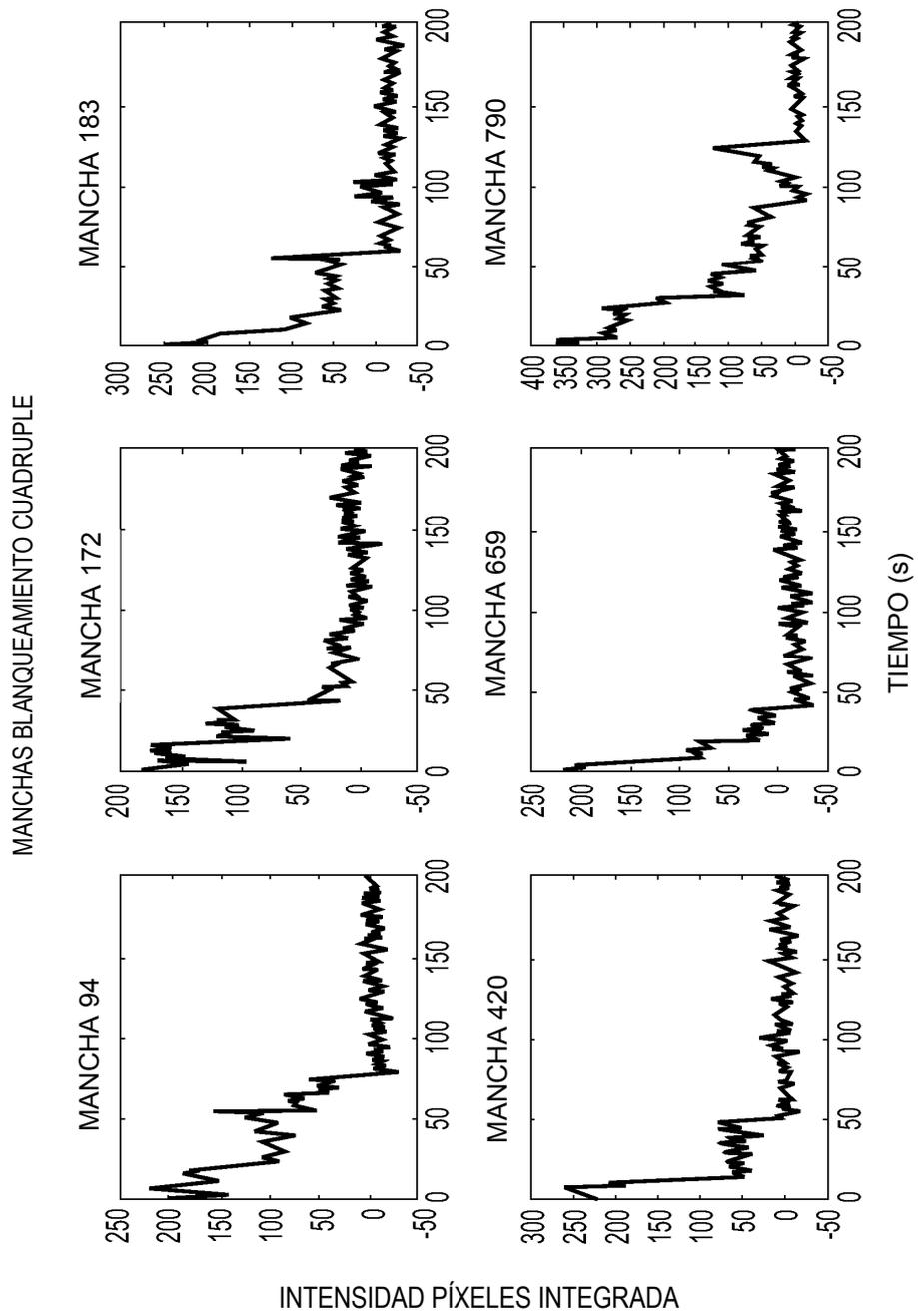
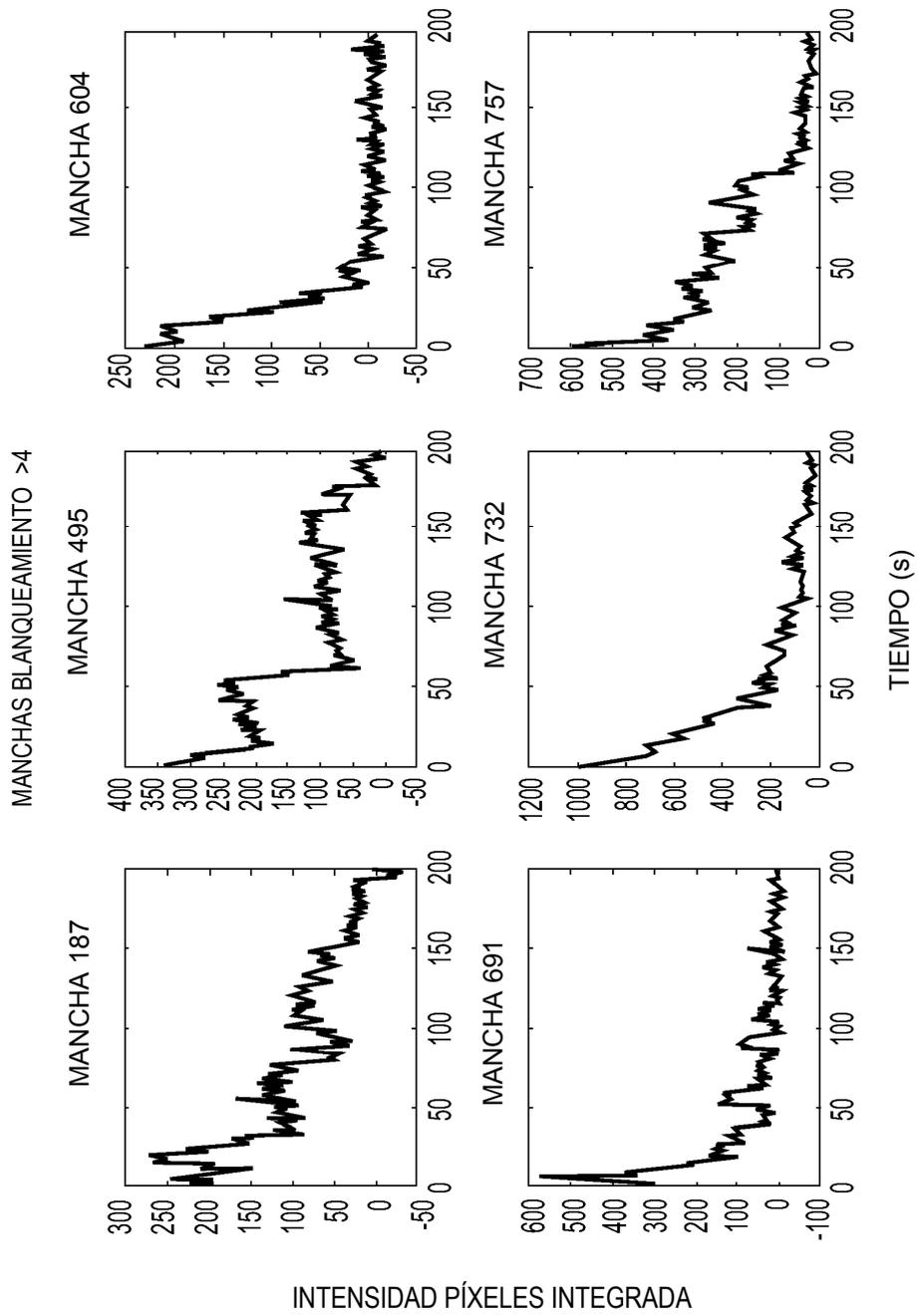


Figura 13



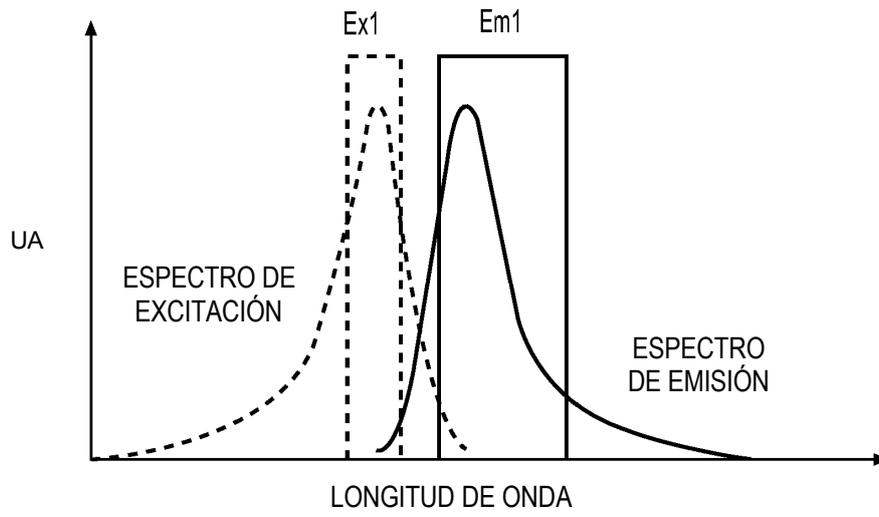


Figura 14

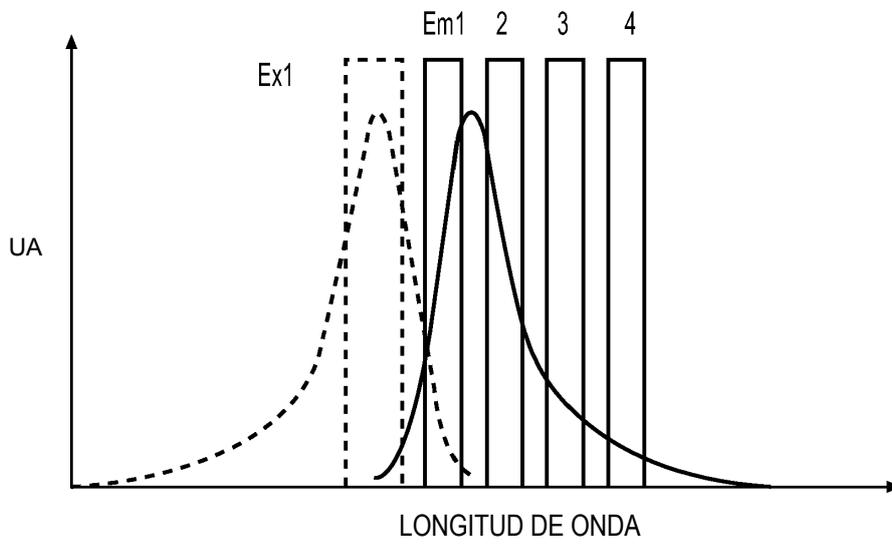


Figura 15

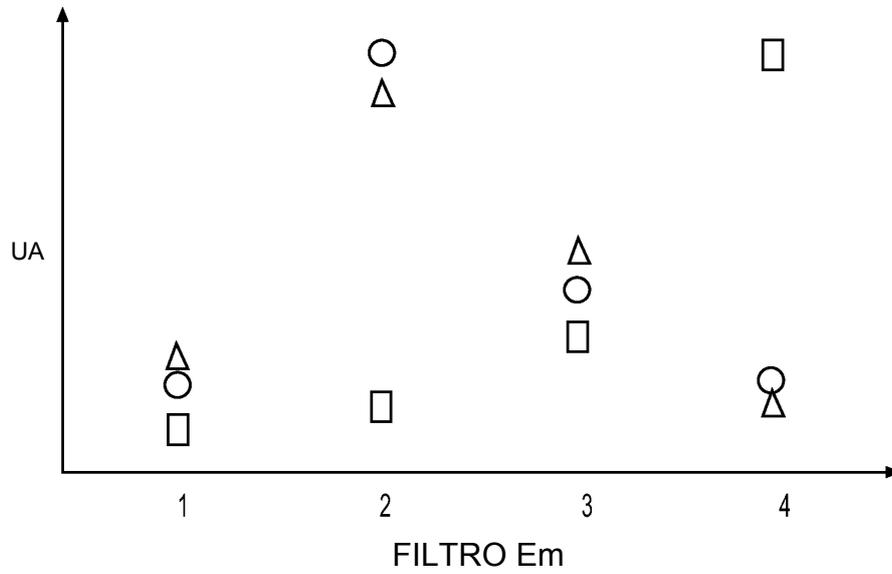


Figura 16

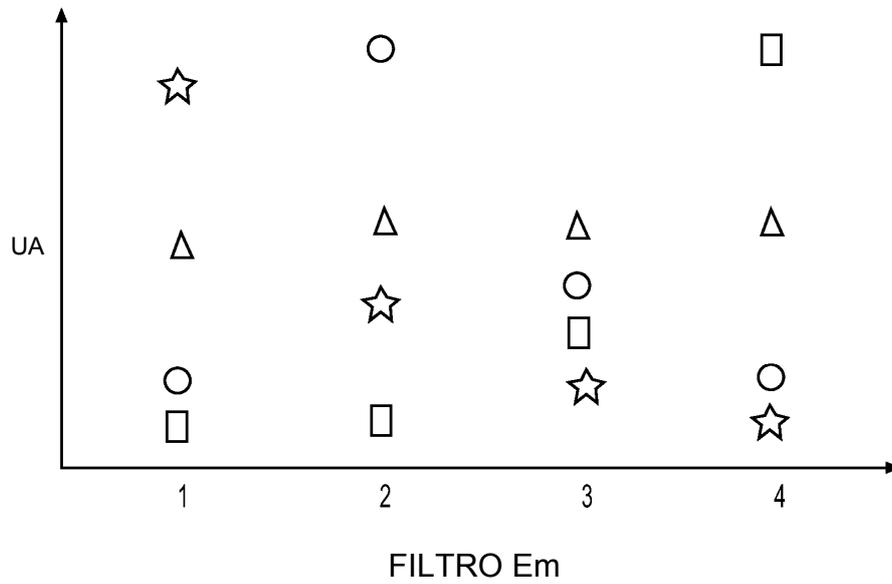


Figura 17

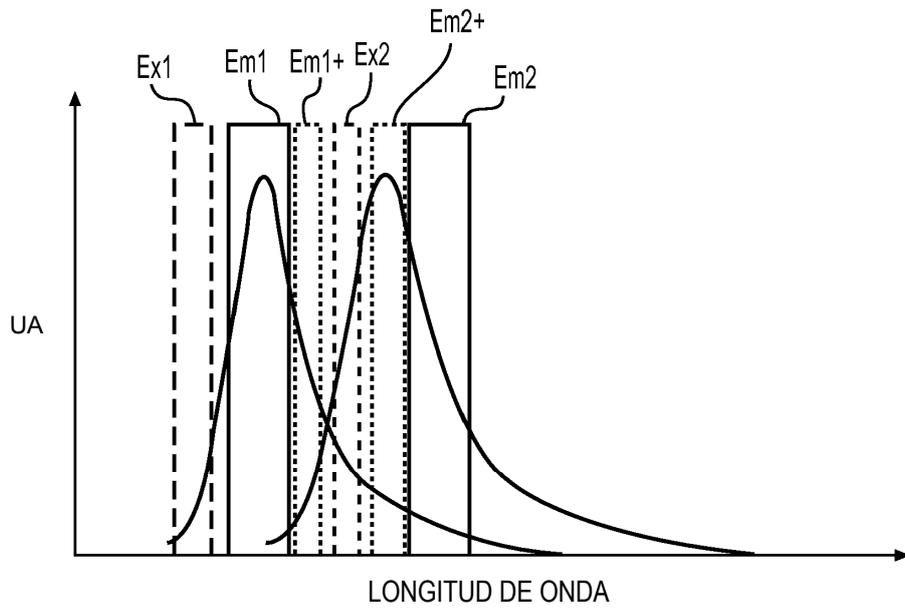


Figura 18

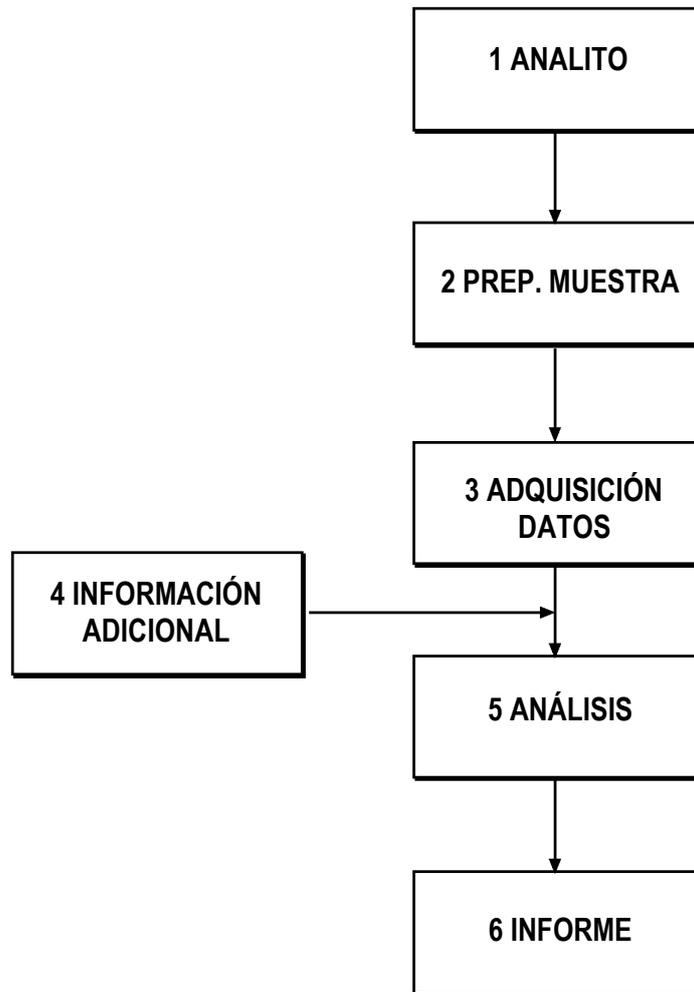


Figura 19

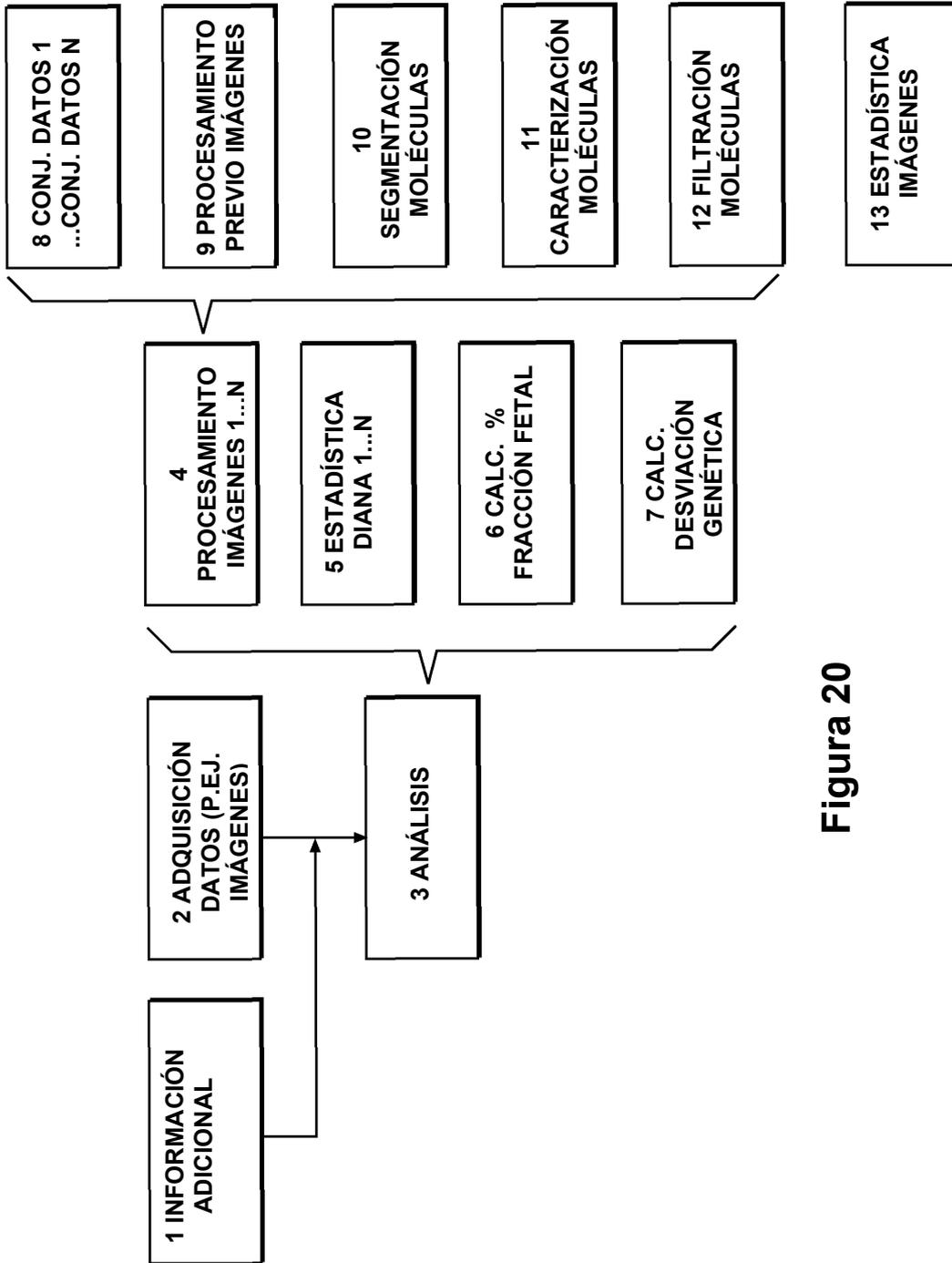


Figura 20

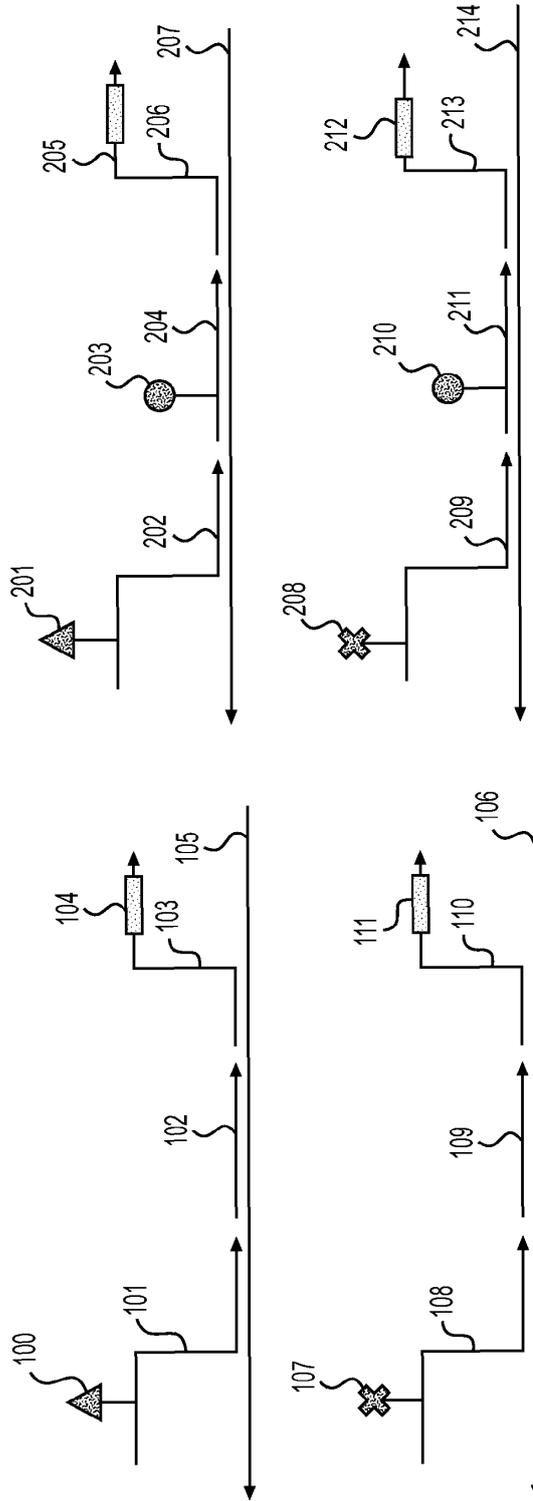


Figure 22

Figure 21

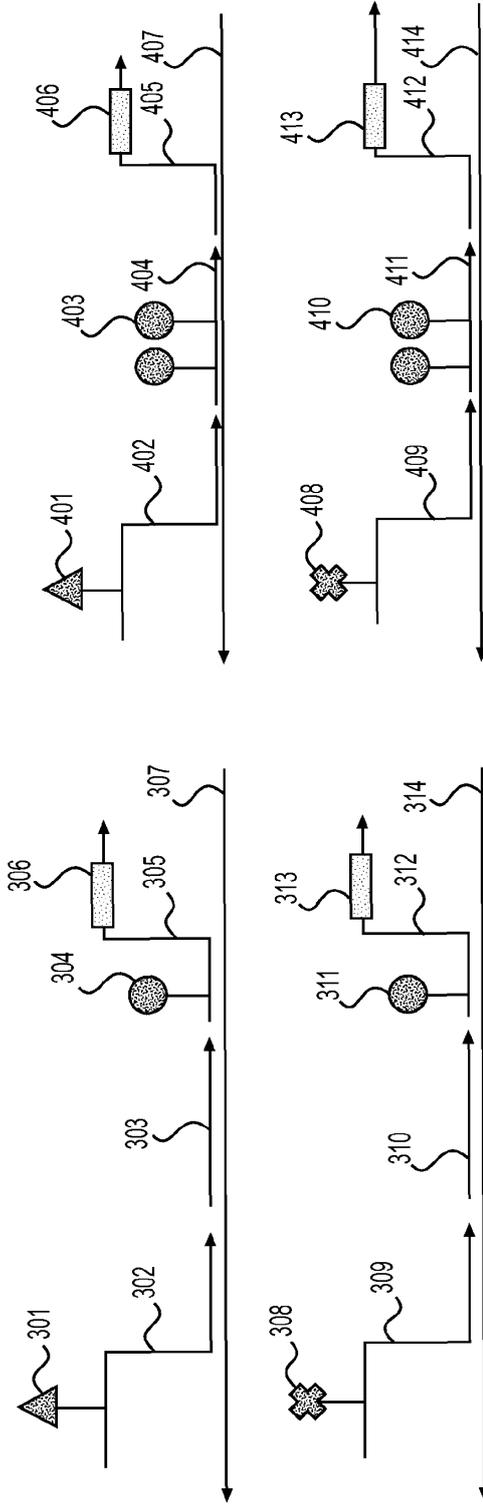


Figura 23

Figura 24

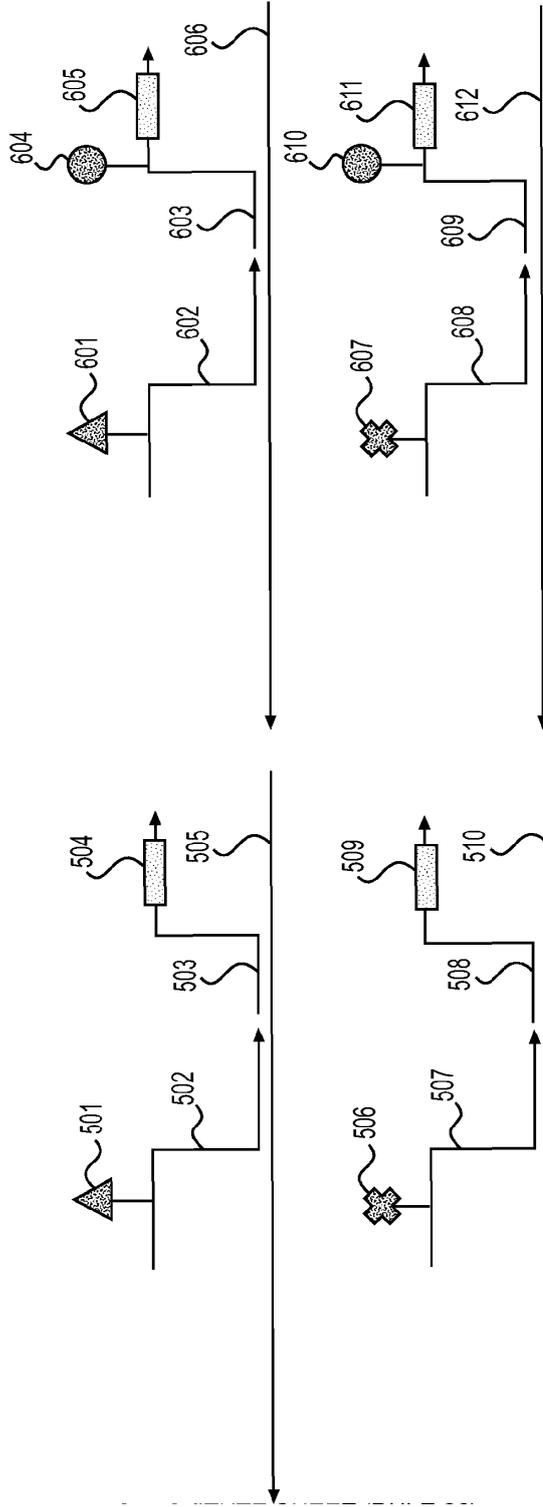


Figura 26

Figura 25

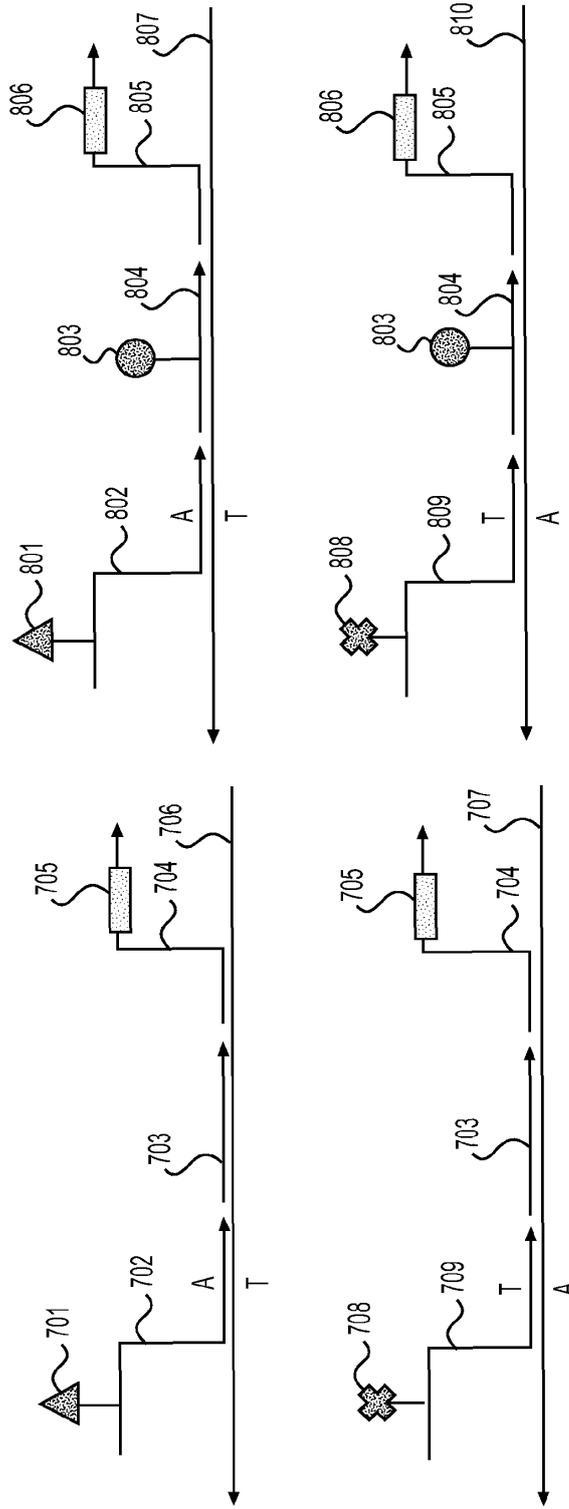


Figura 28

Figura 27

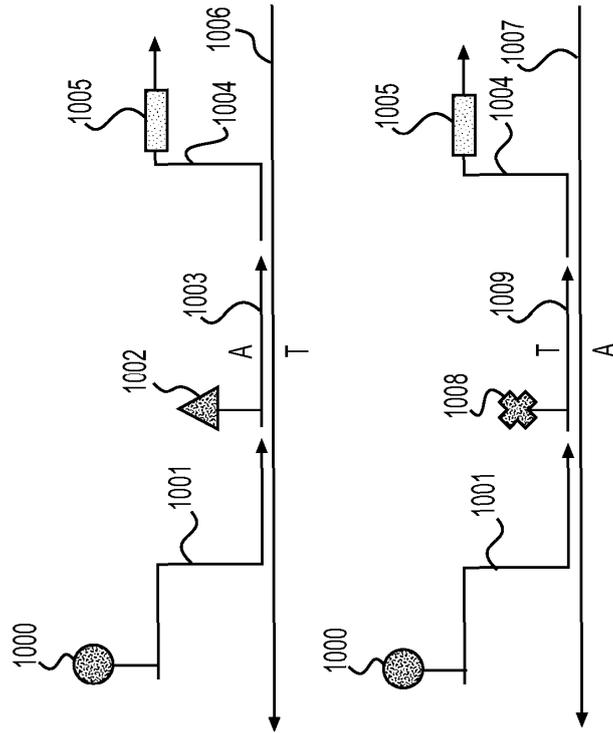


Figura 30

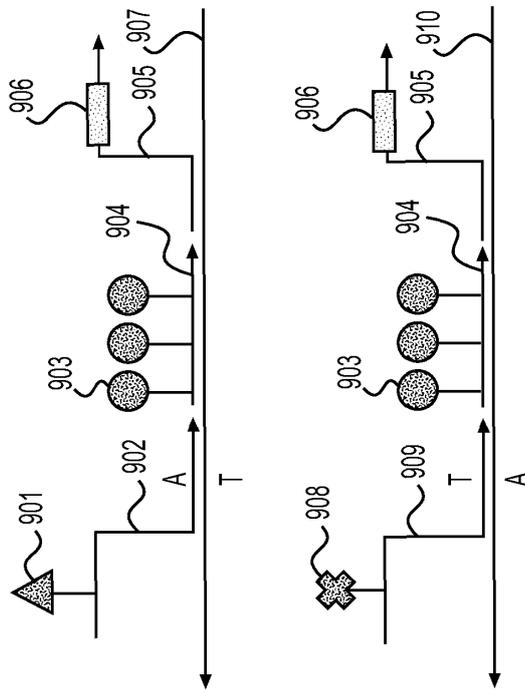


Figura 29

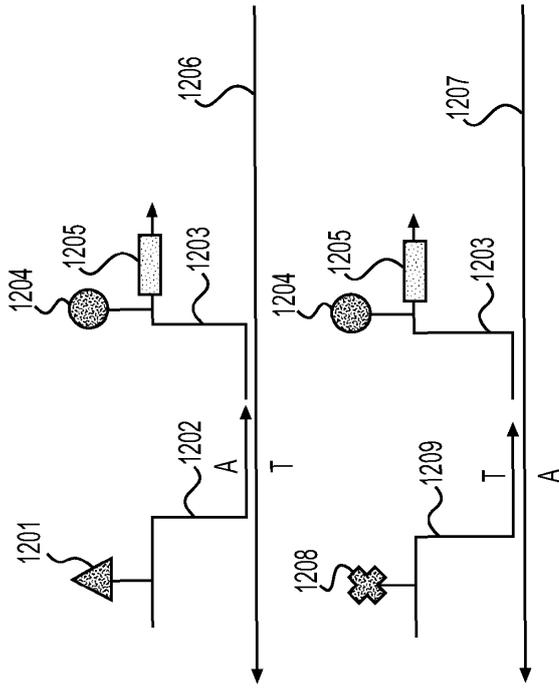


Figure 32

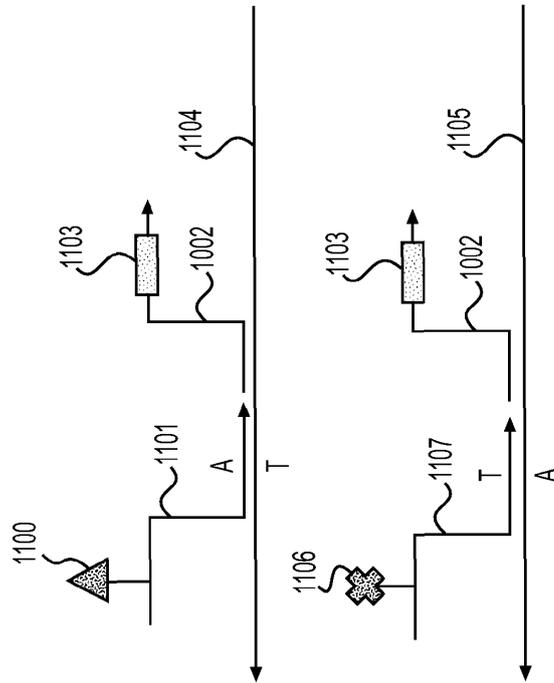


Figure 31

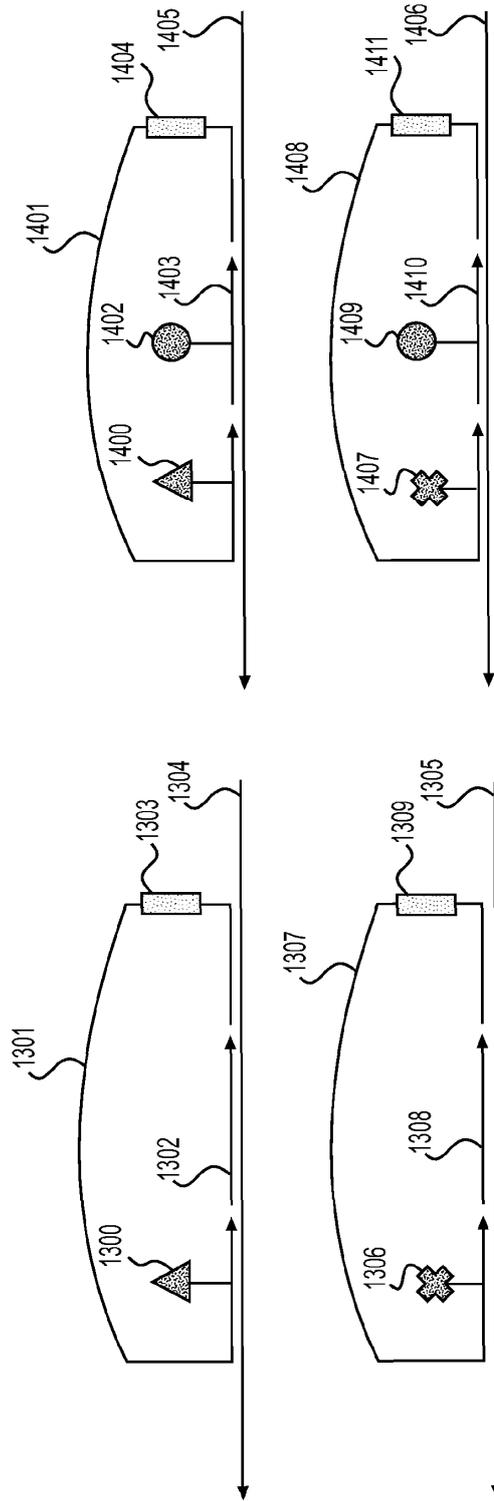


Figura 34

Figura 33

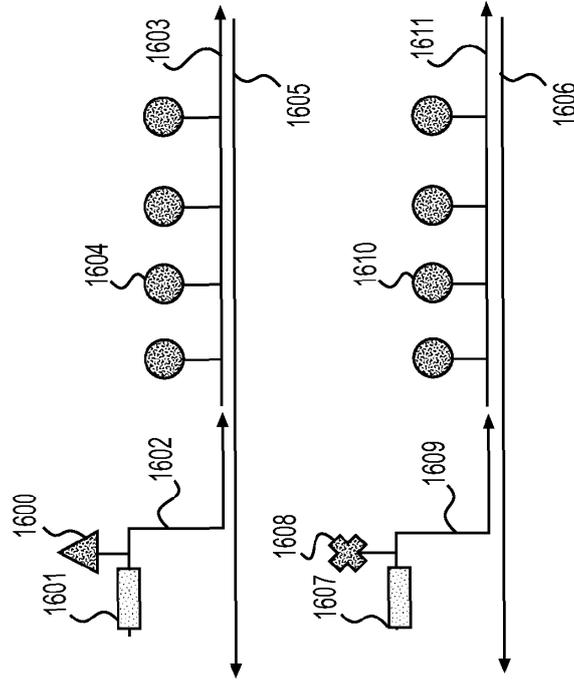


Figura 36

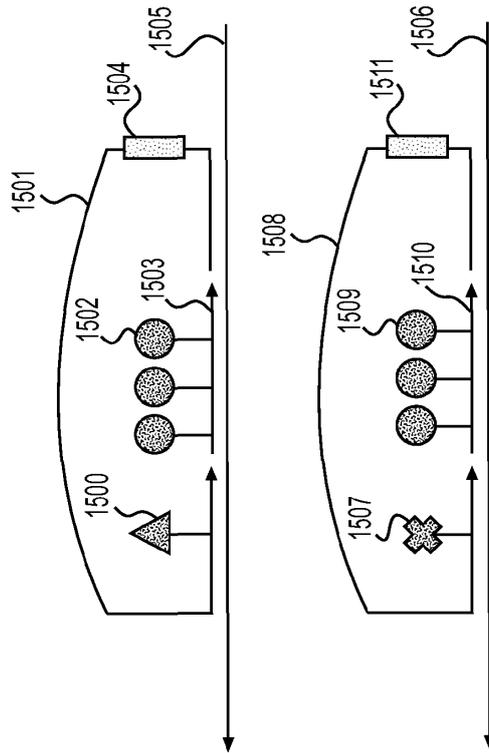


Figura 35

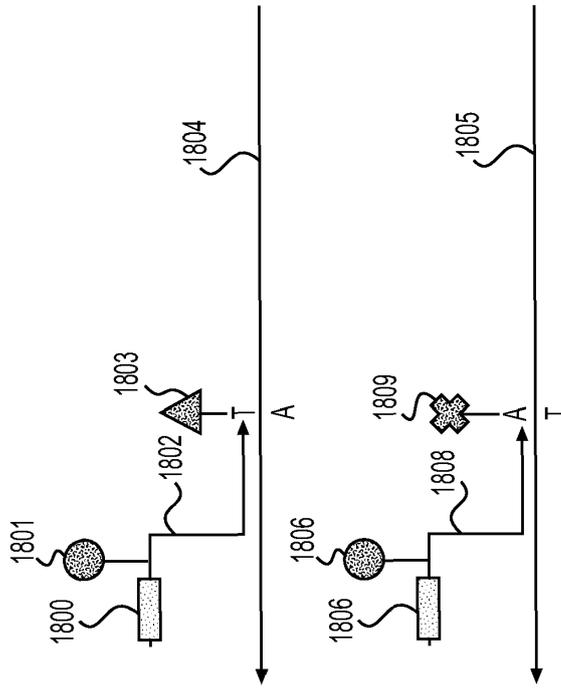


Figure 38

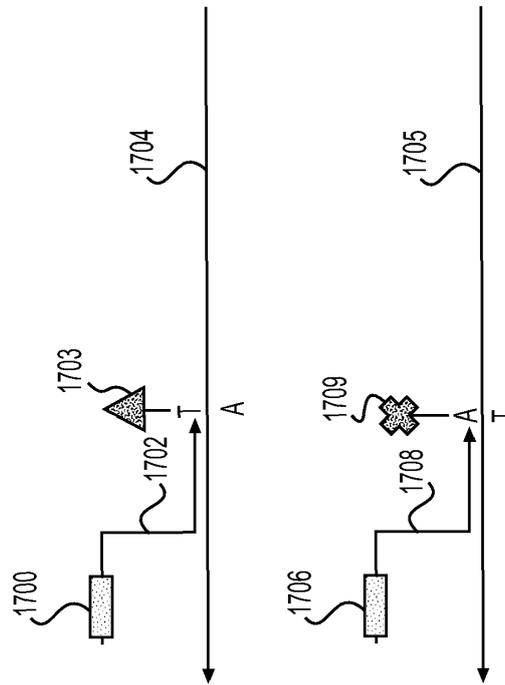


Figure 37

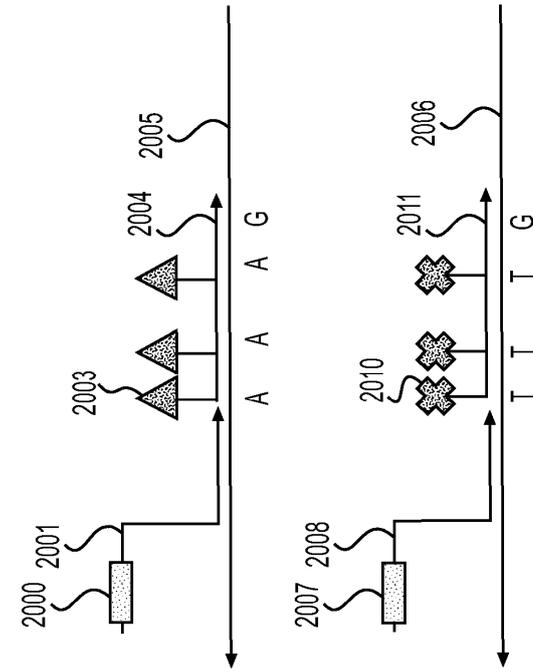


Figura 40

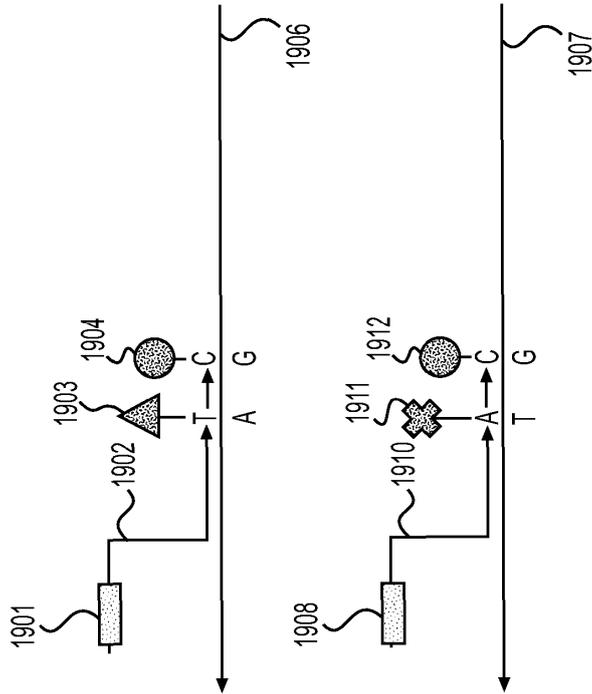


Figura 39

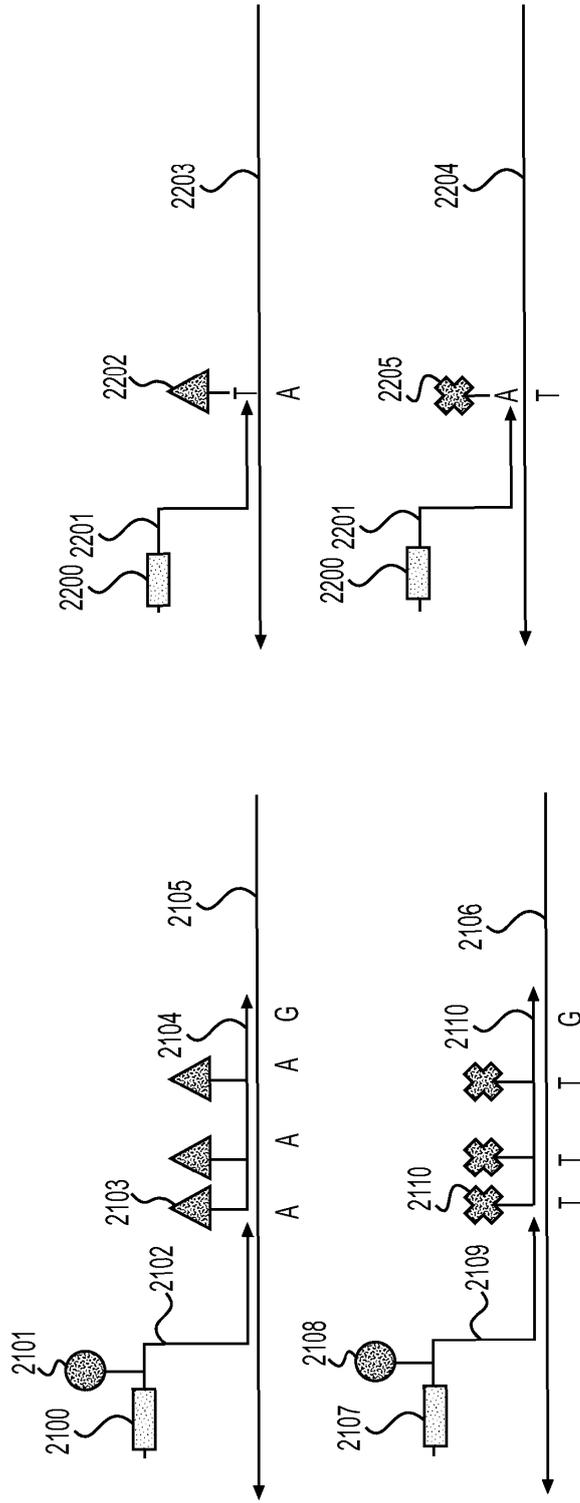


Figura 41

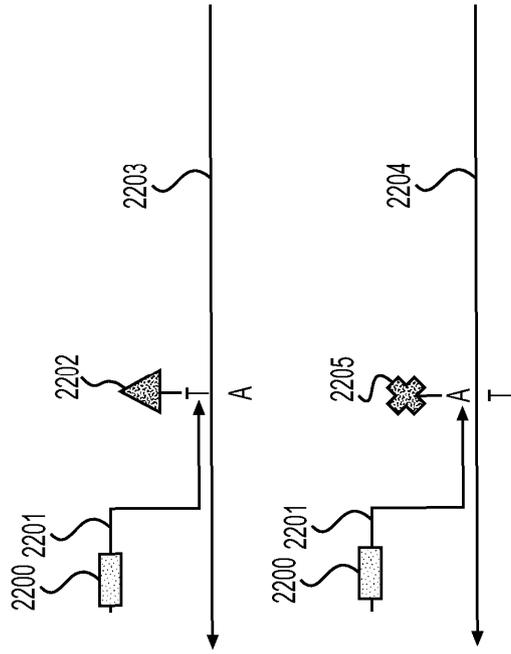


Figura 42

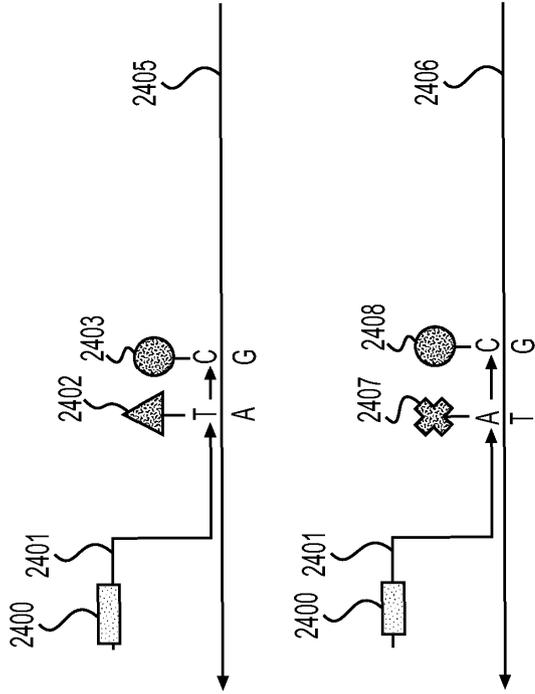


Figura 44

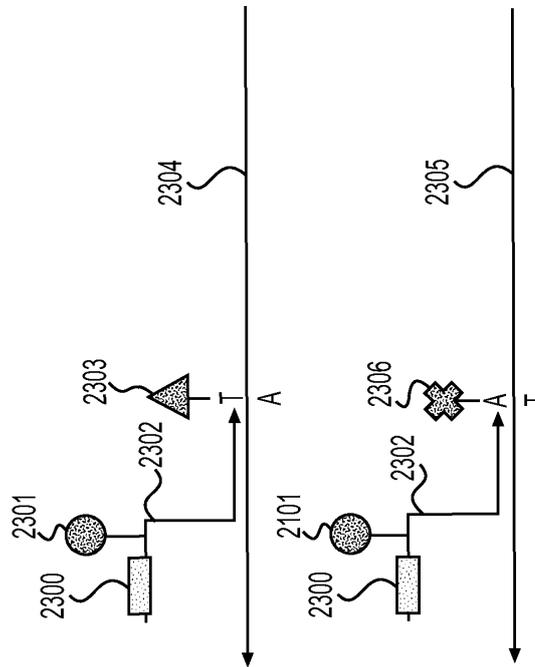


Figura 43

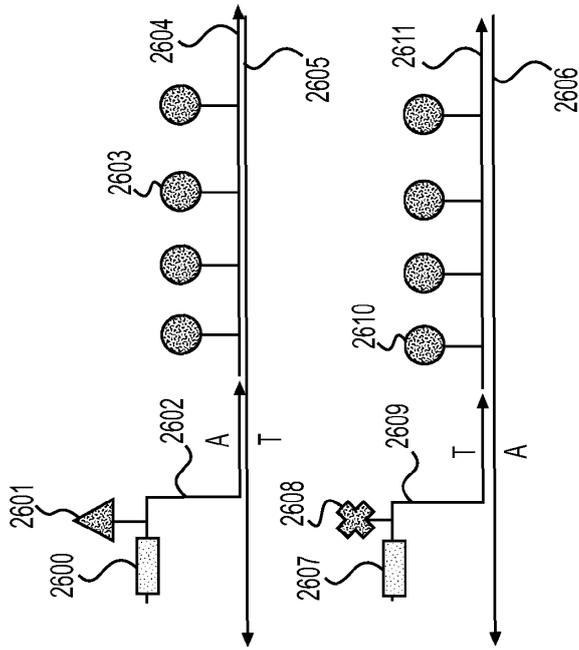


Figura 46

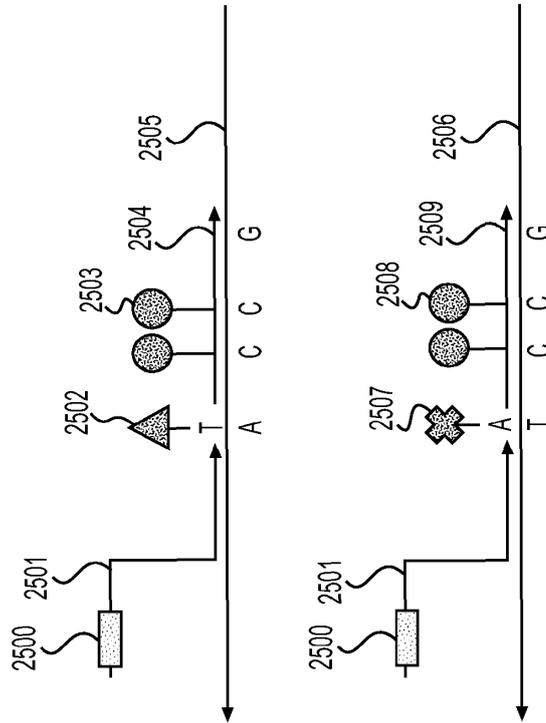


Figura 45

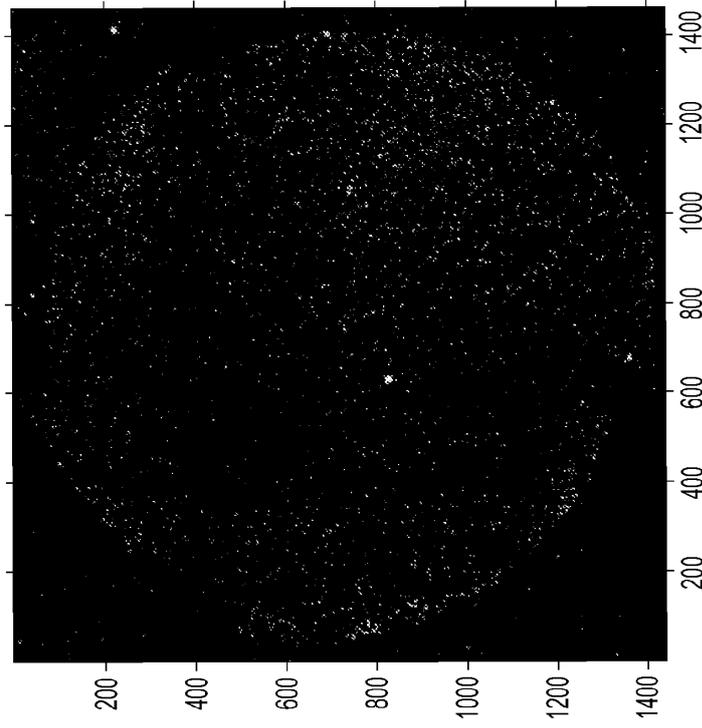


Figura 47

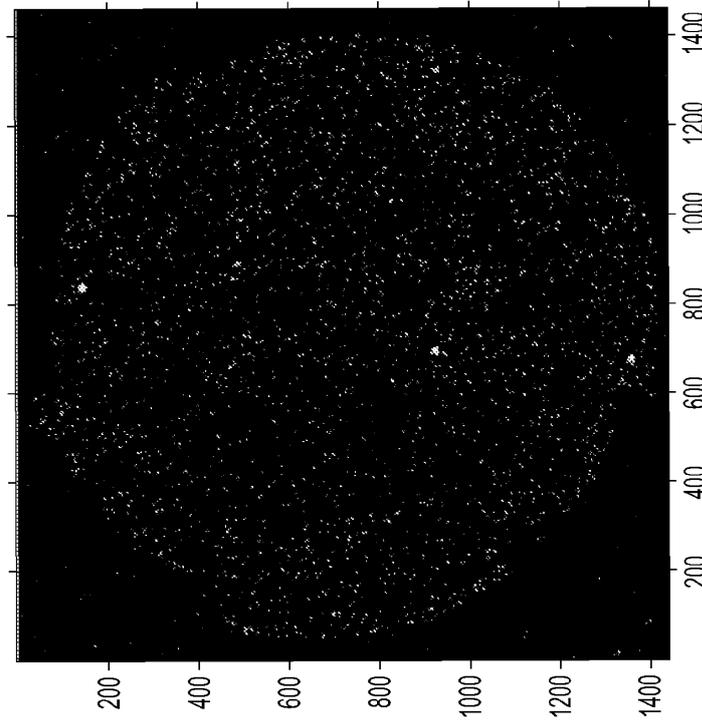


Figura 48

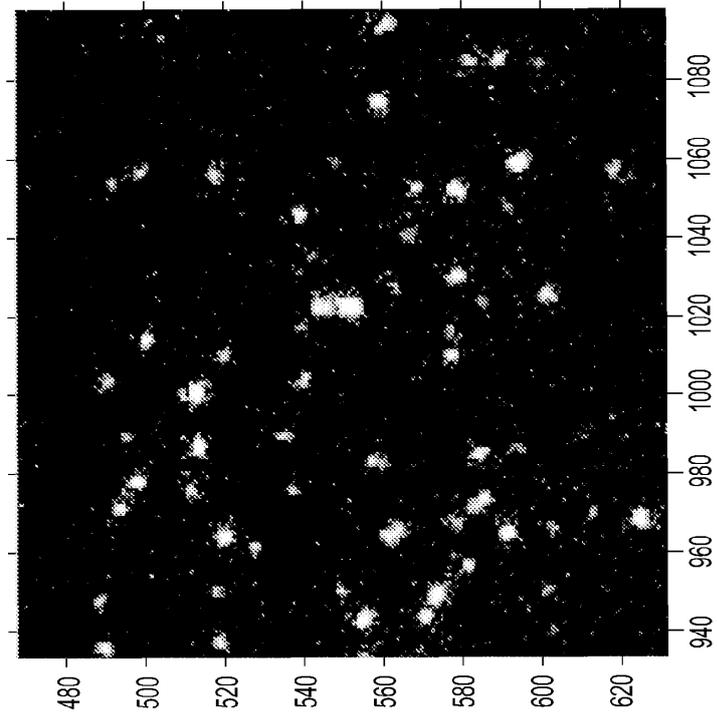


Figura 49

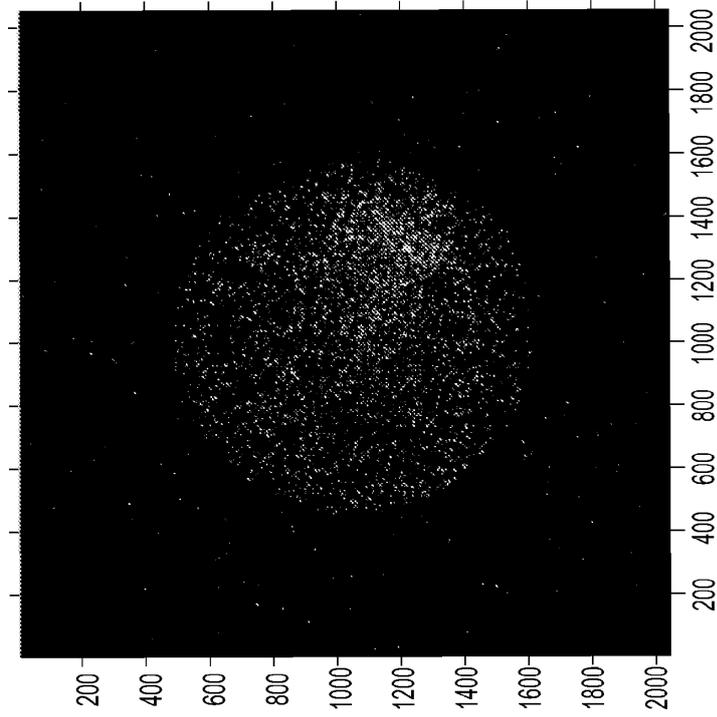


Figura 50

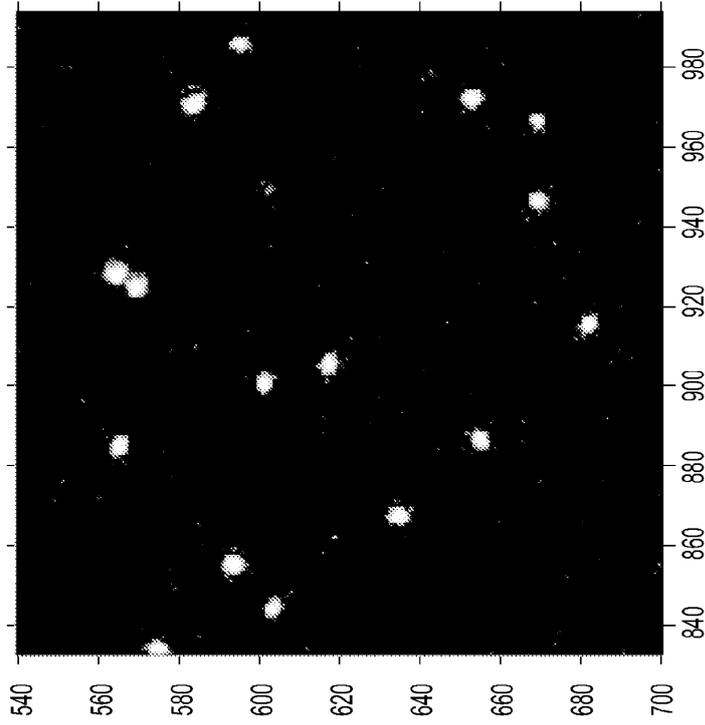


Figura 51

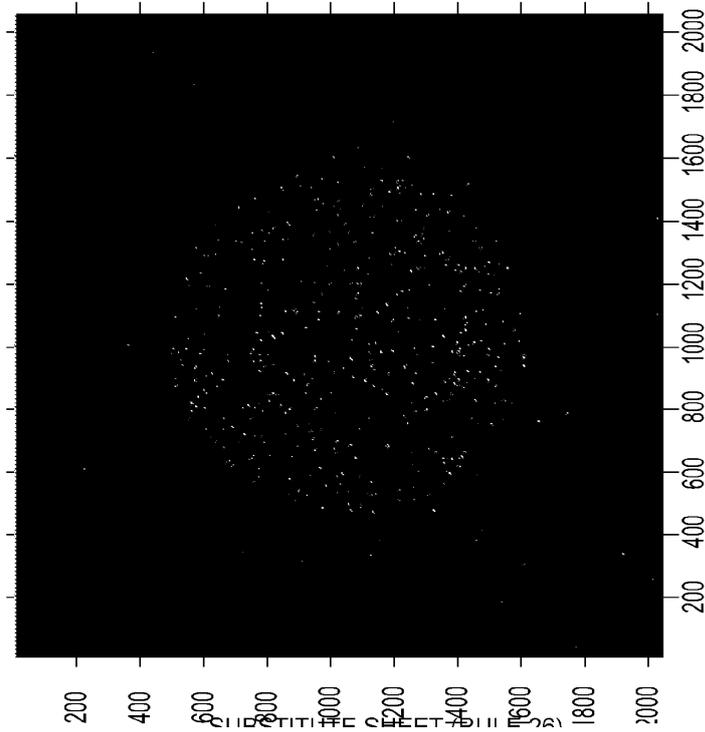


Figura 52

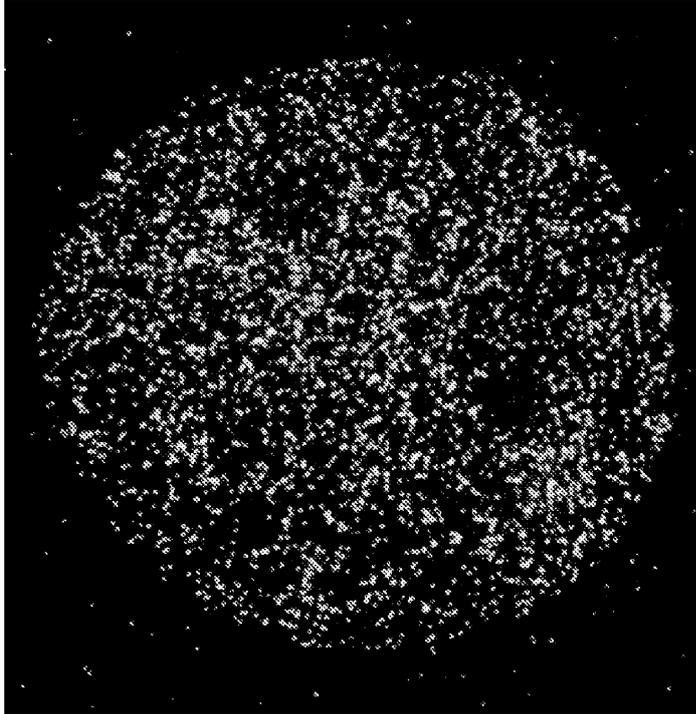


Figura 54

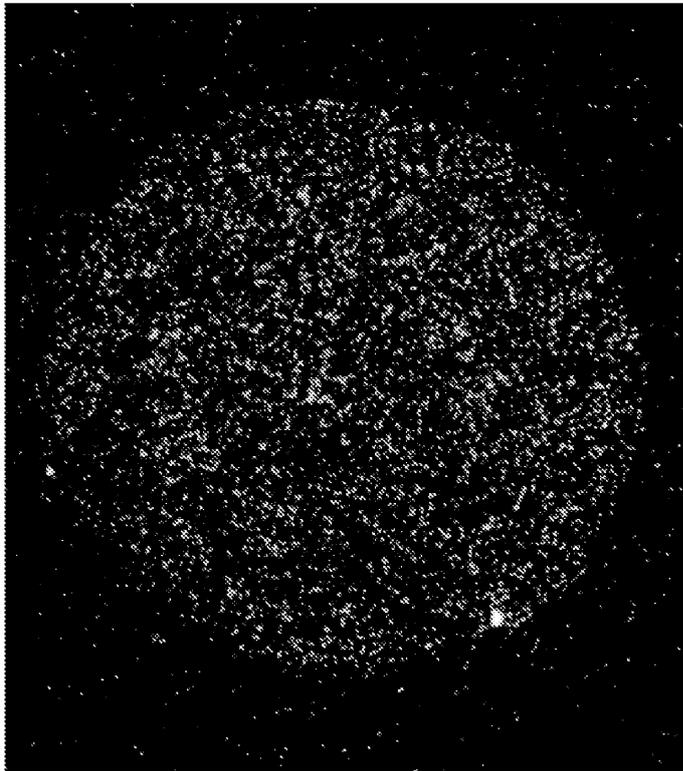


Figura 53

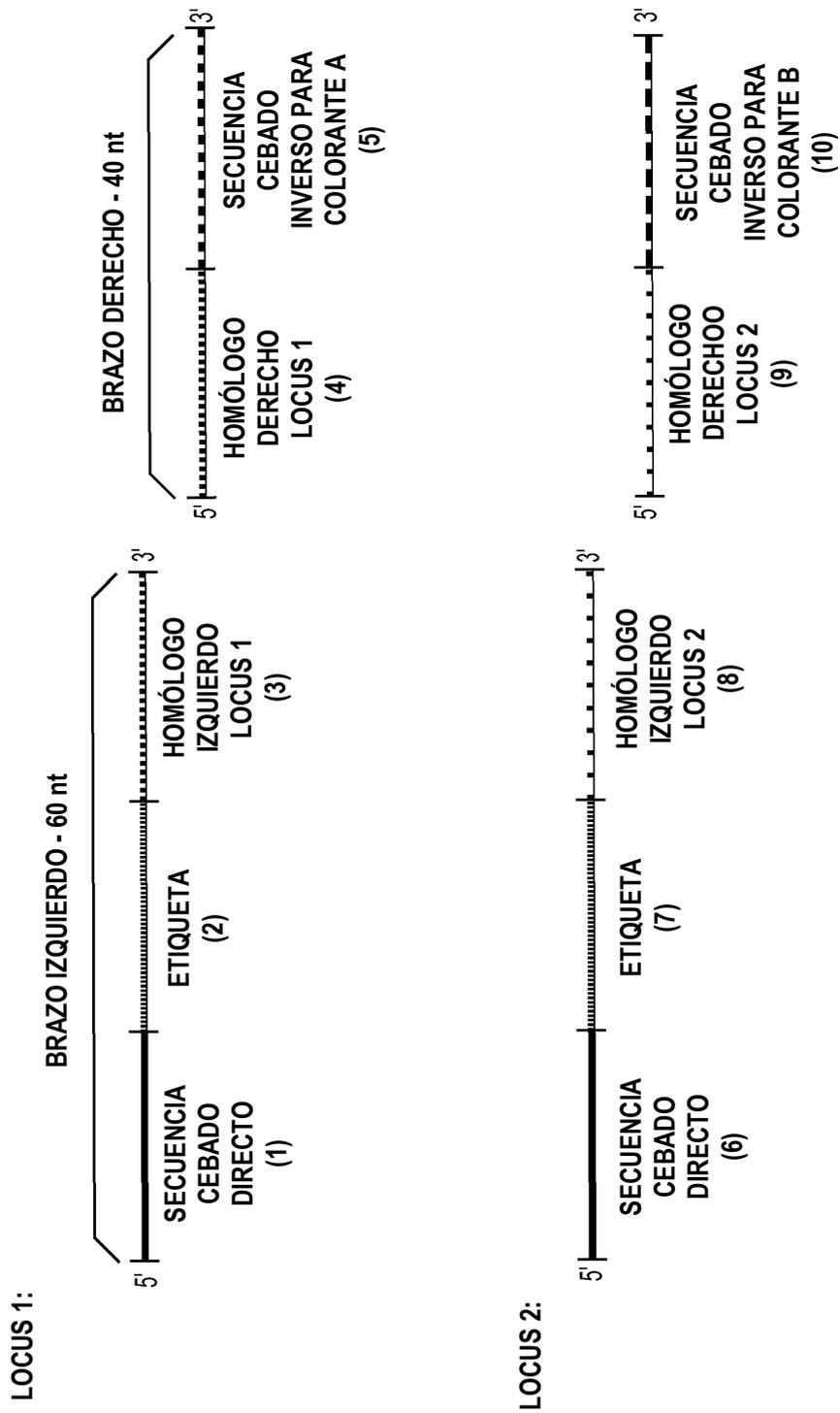


Figura 55

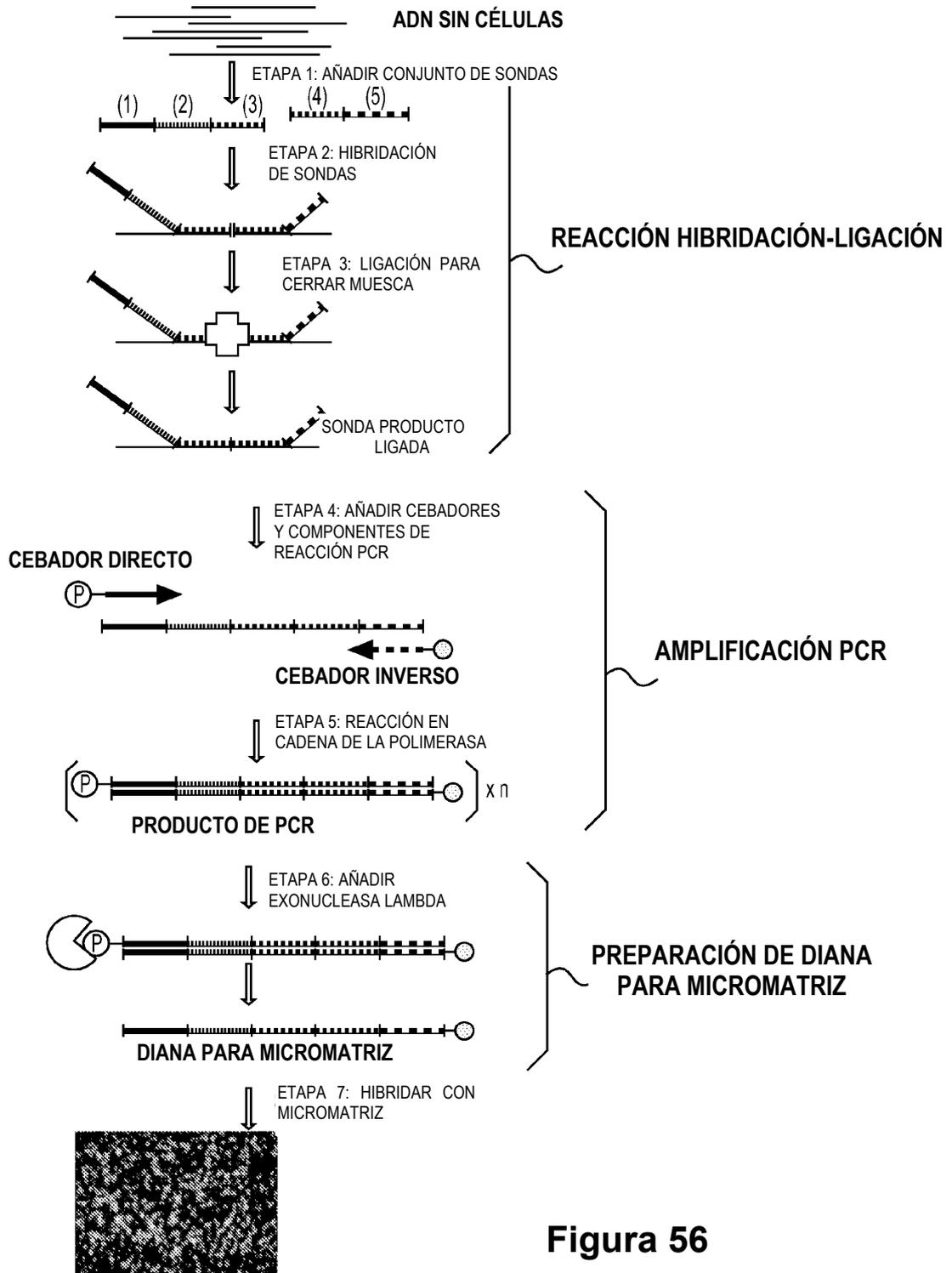


Figura 56

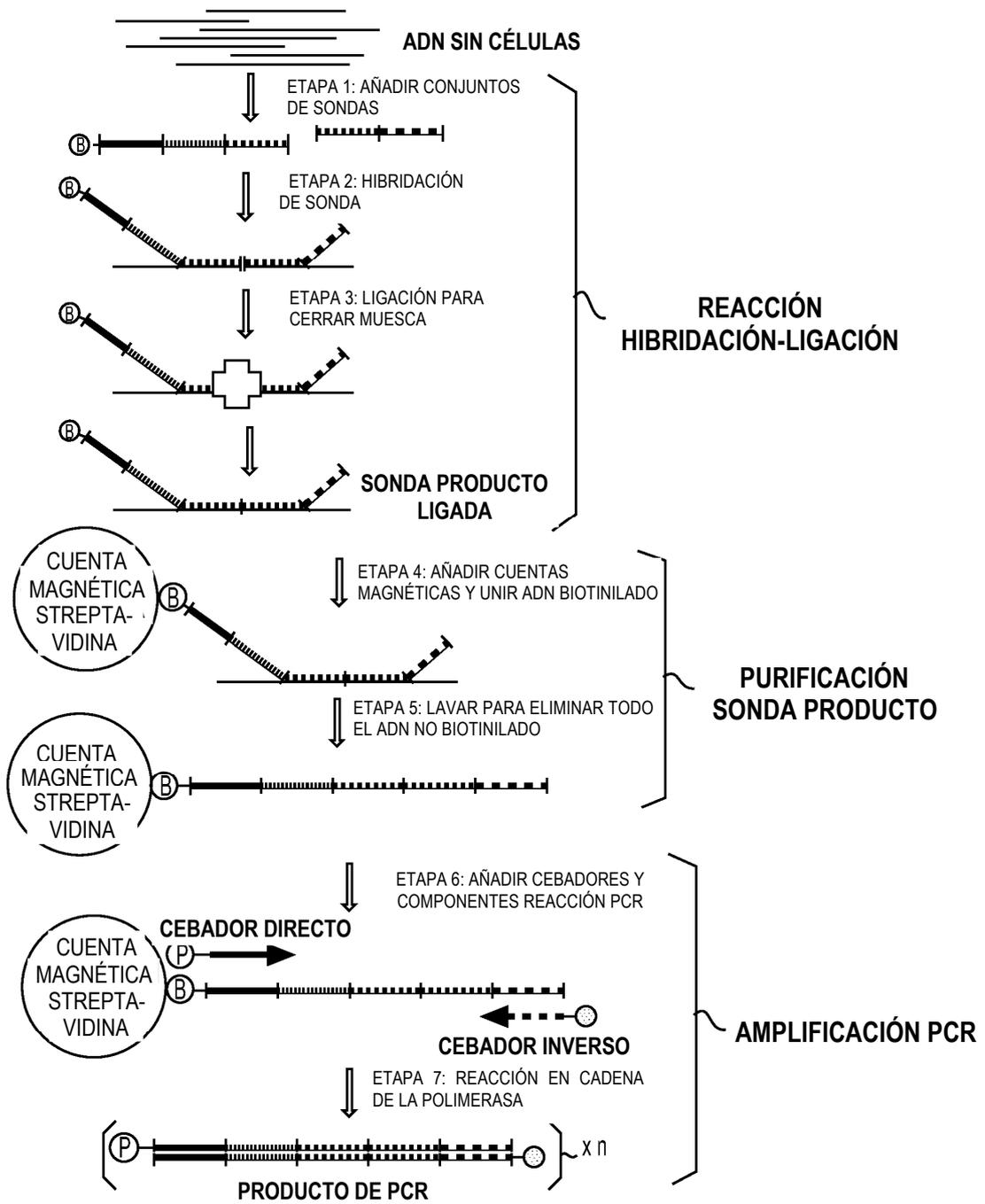


Figura 57

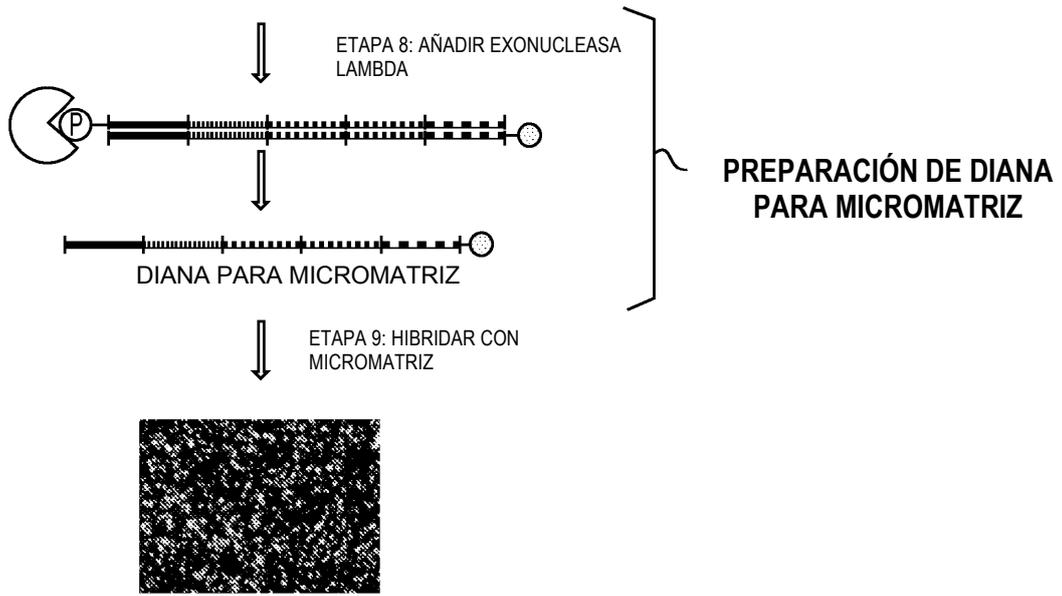


Figura 57 Continuación

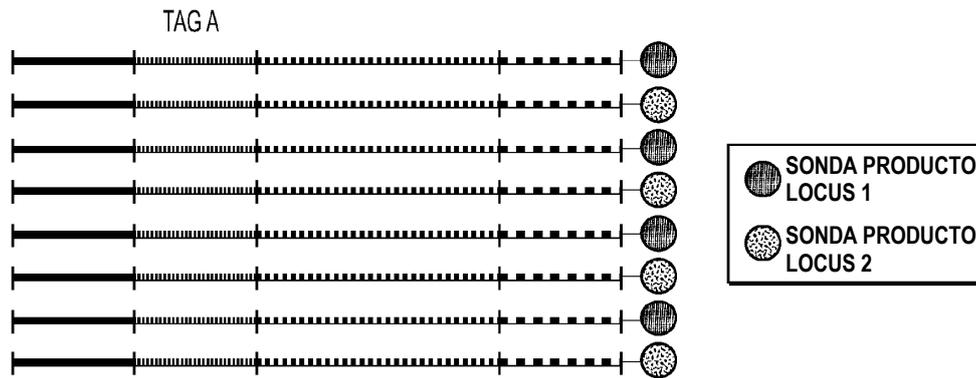


Figura 58A

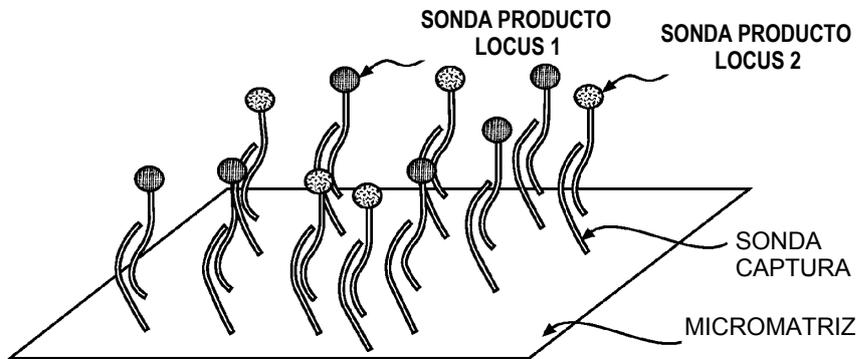


Figura 58B

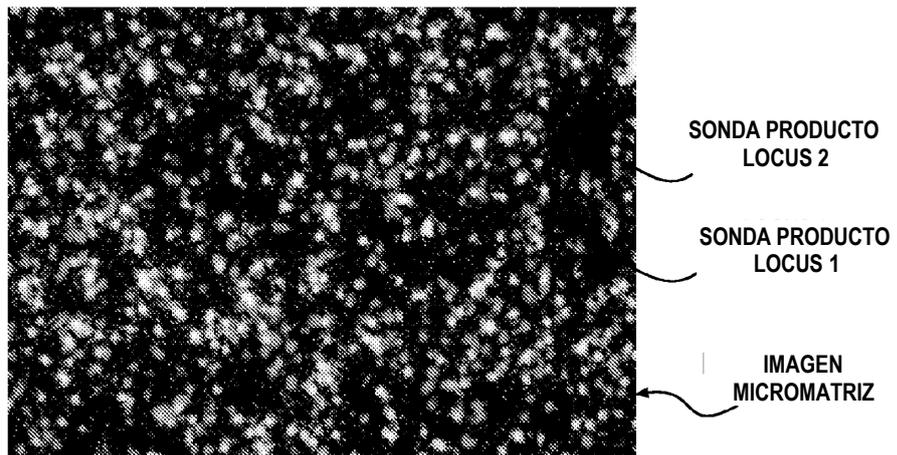


Figura 58C

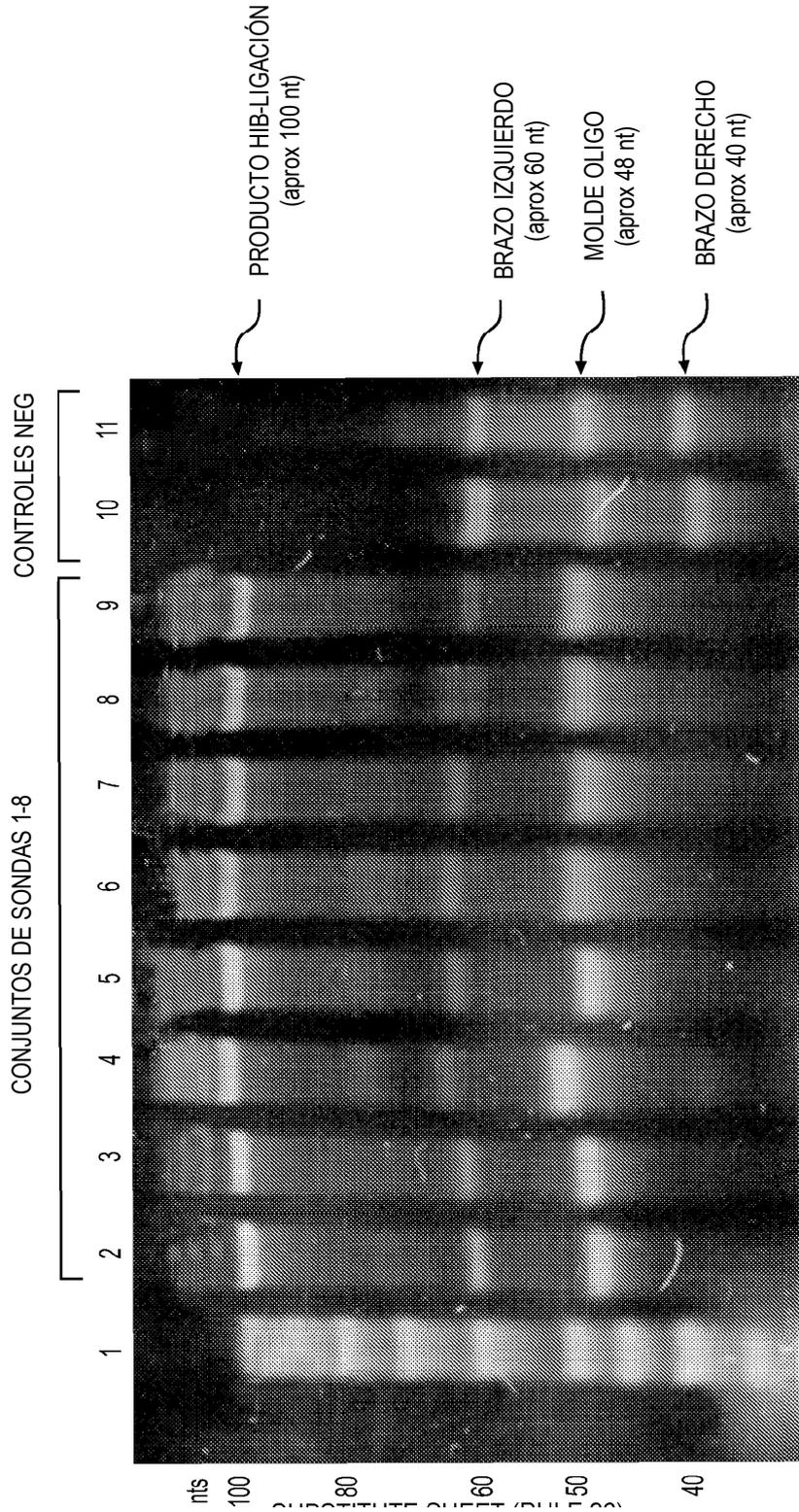


Figura 59

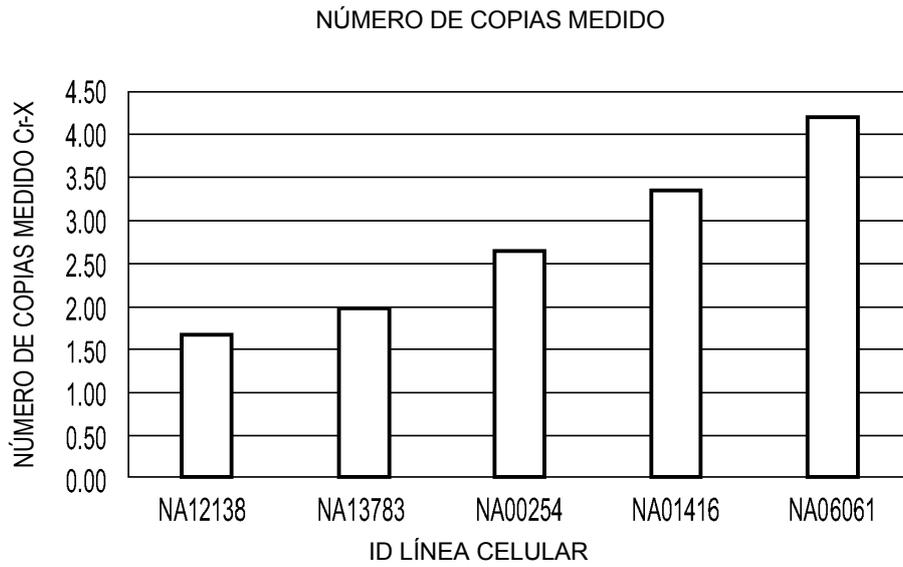


Figura 60A

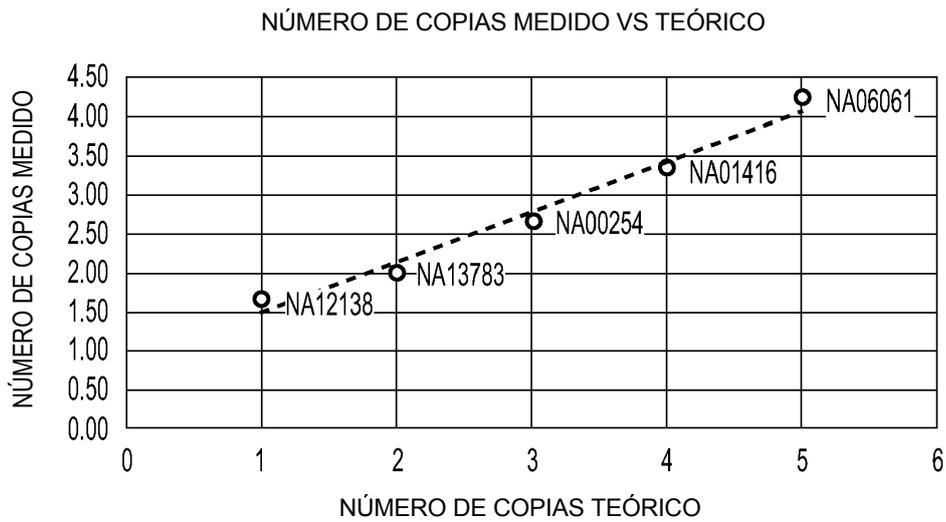


Figura 60B

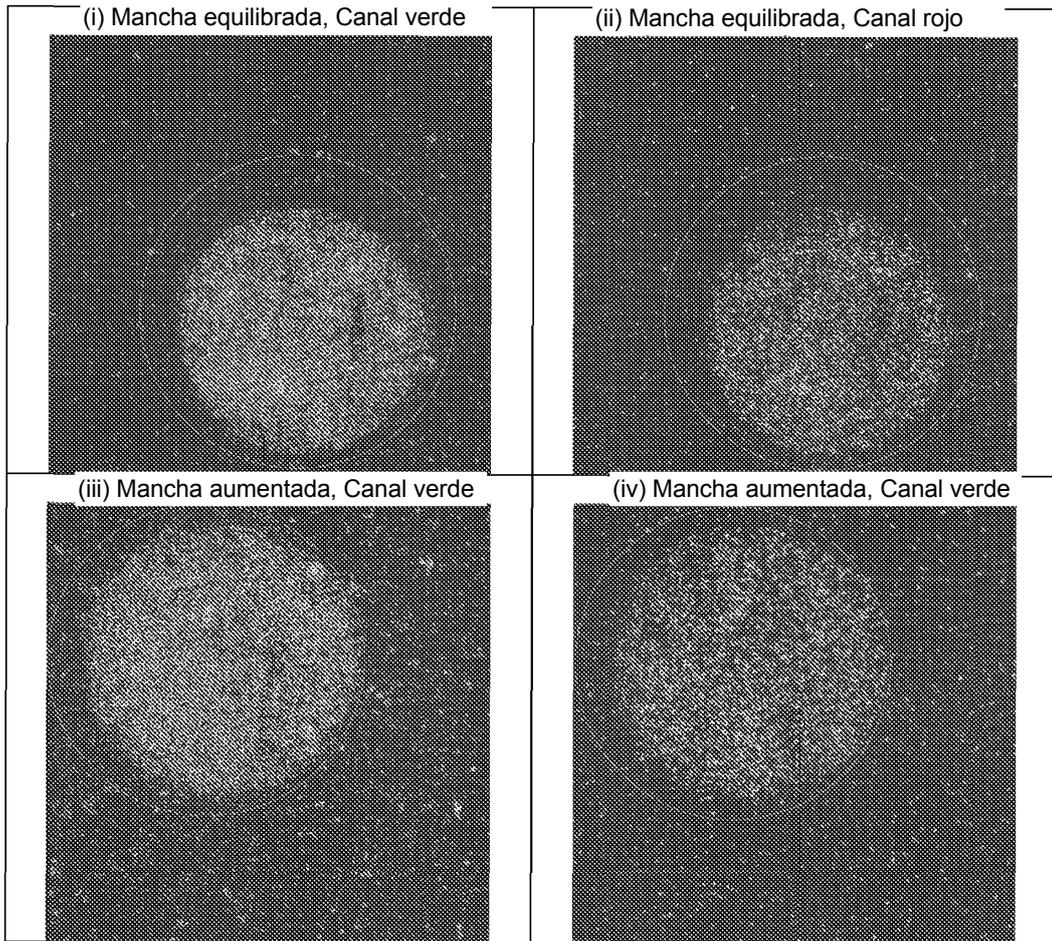


Figura 61A

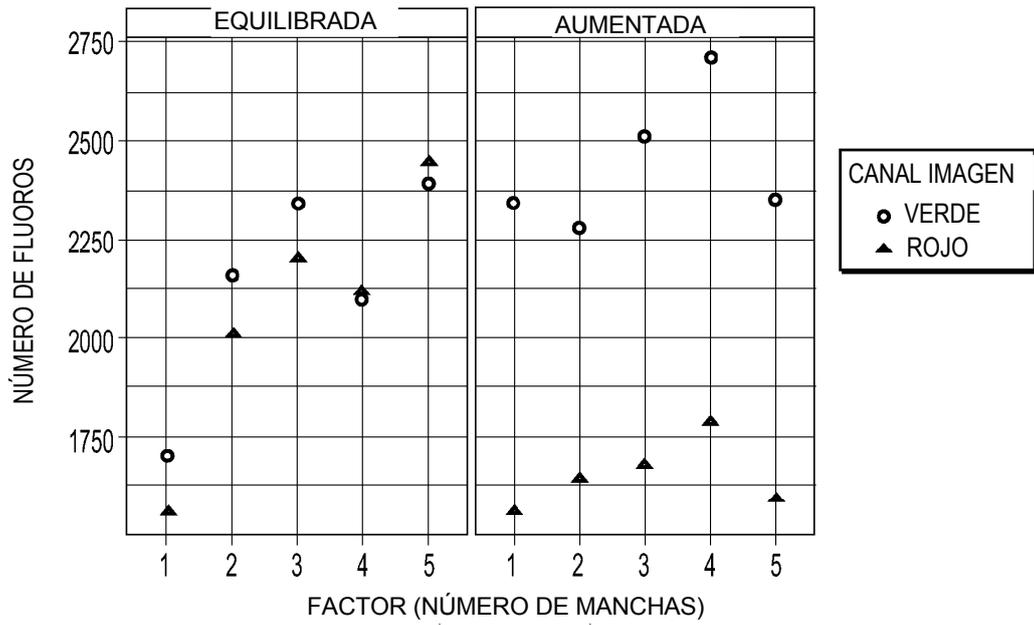


Figura 61B

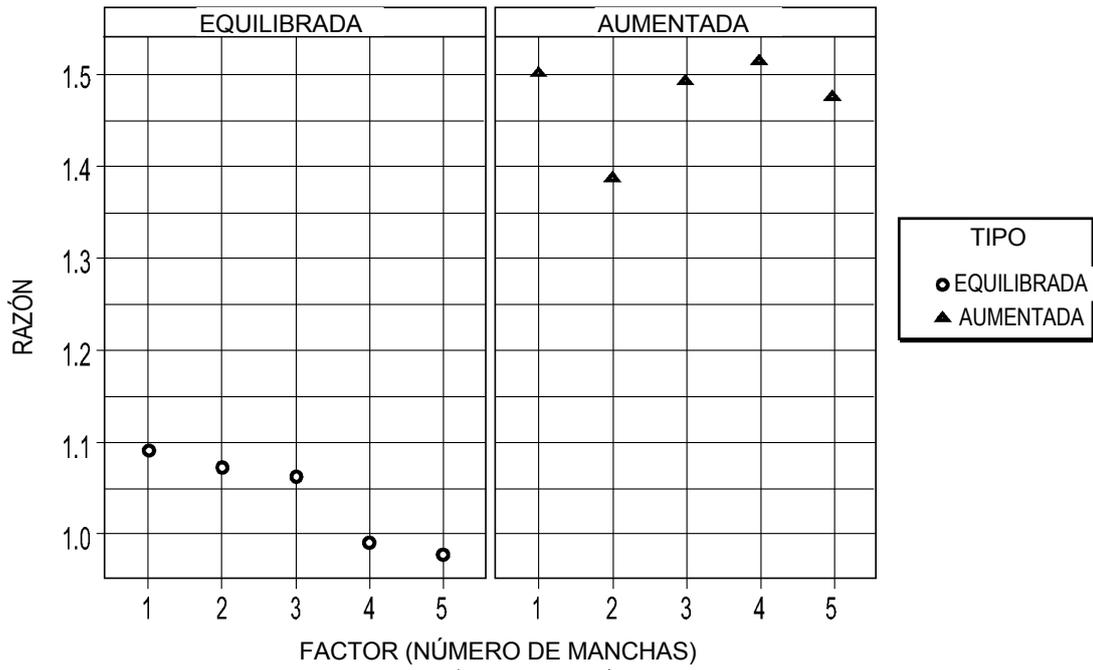


Figura 61C

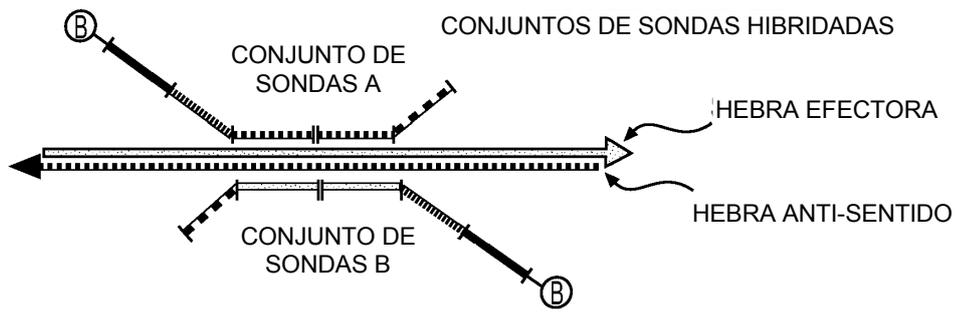


Figura 62

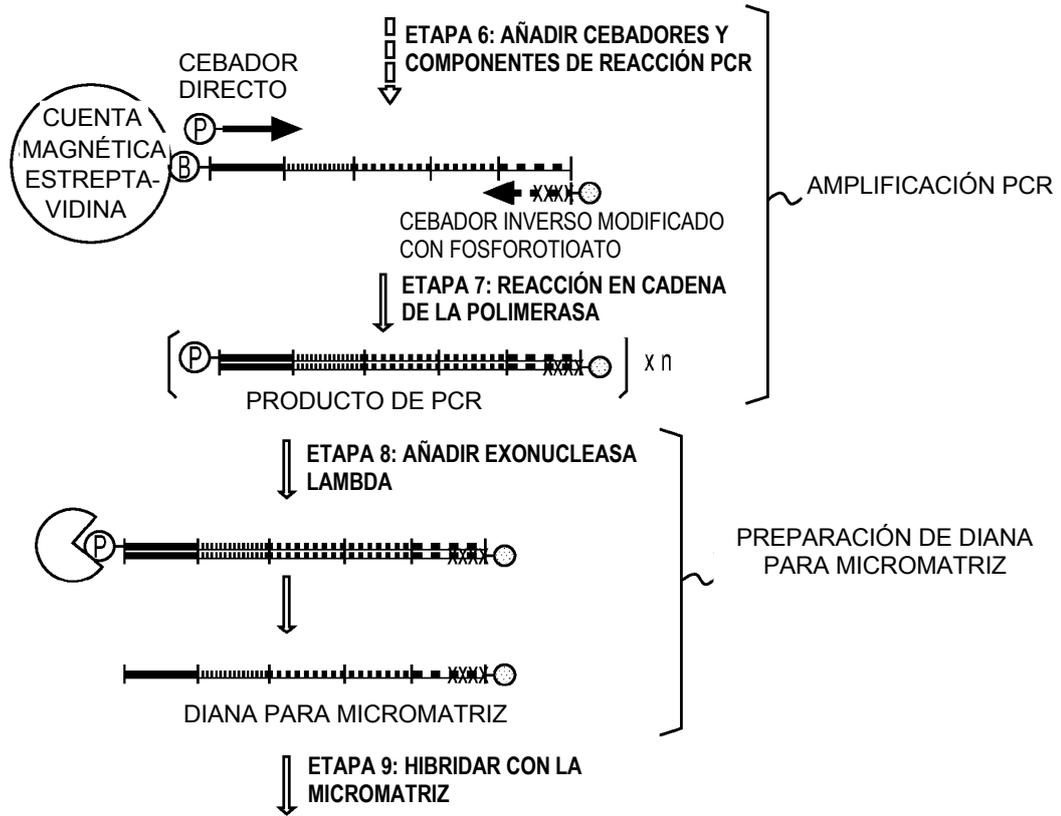


Figura 63

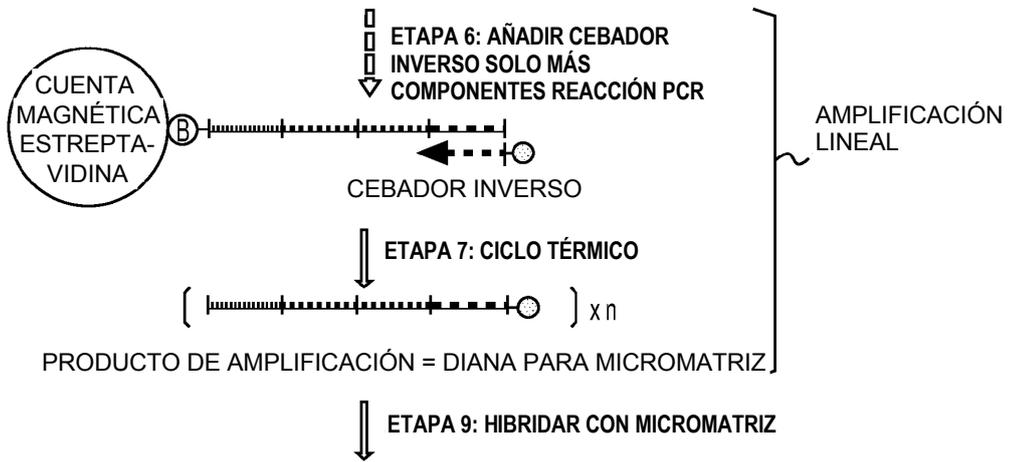


Figura 64

EVENTOS DE HIBRIDACIÓN-LIGACIÓN CICLO 1

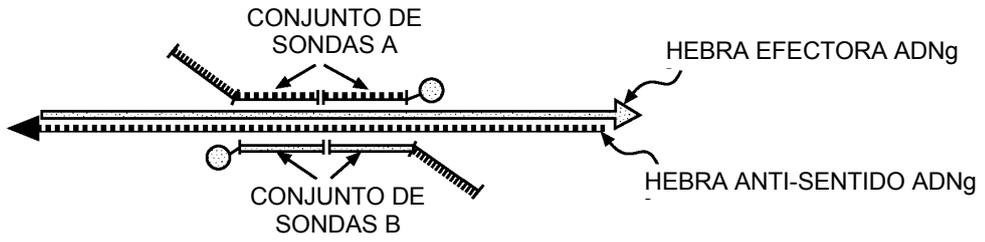


Figura 65A

EVENTOS DE HIBRIDACIÓN-LIGACIÓN CICLO 2

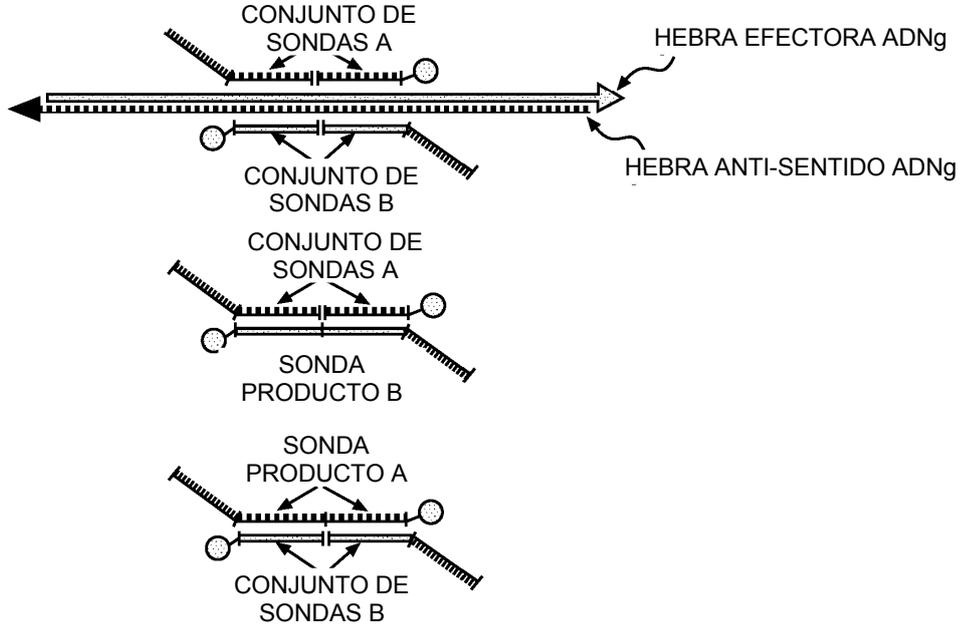


Figura 65B

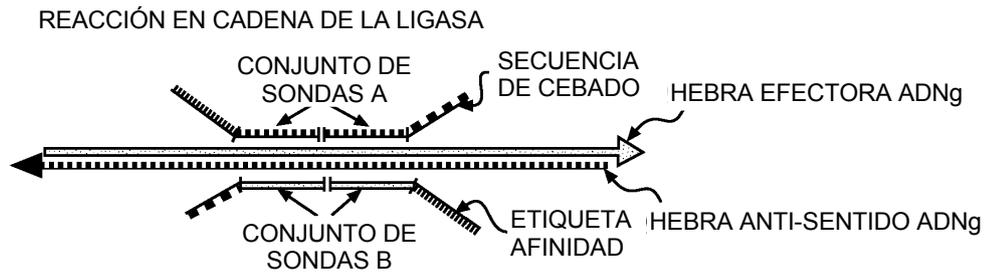


Figura 66A

AMPLIFICACIÓN LINEAL Y MARCAJE CON UN CEBADOR

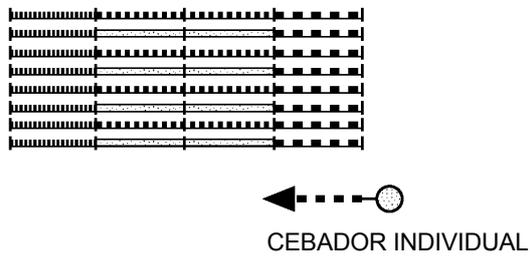


Figura 66B

DIANA PARA MICROMATRIZ

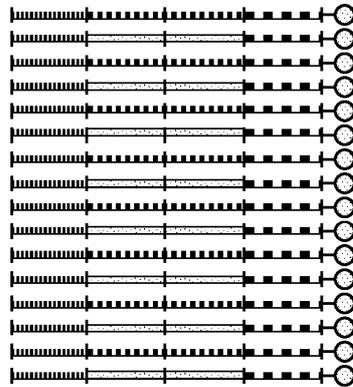


Figura 66C

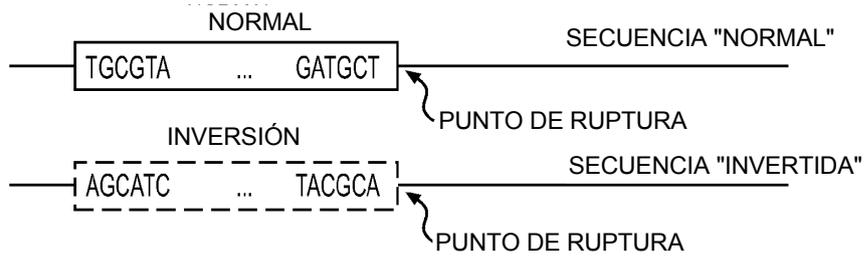


Figura 67A

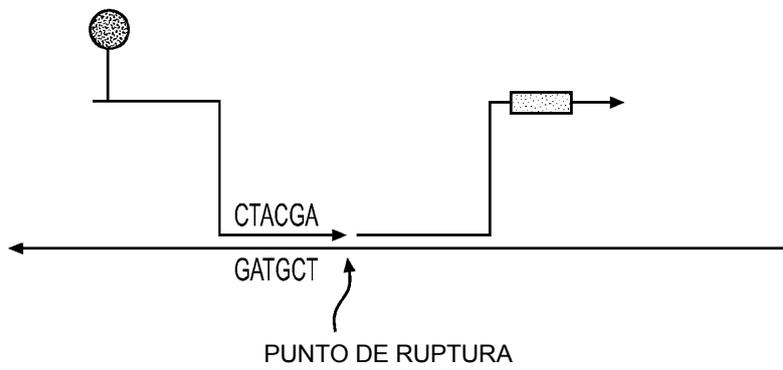


Figura 67B

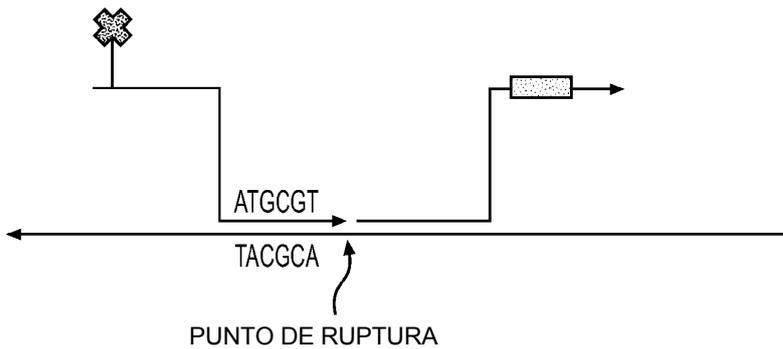


Figura 67C

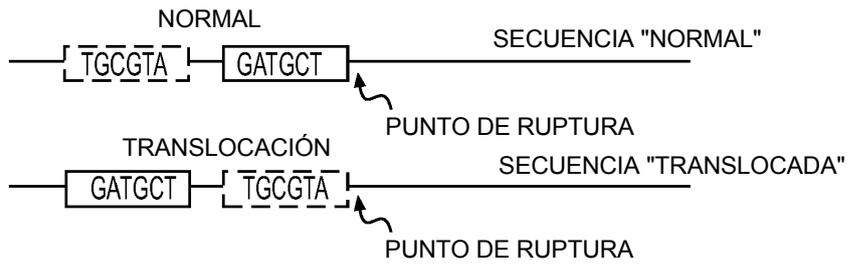


Figura 68A

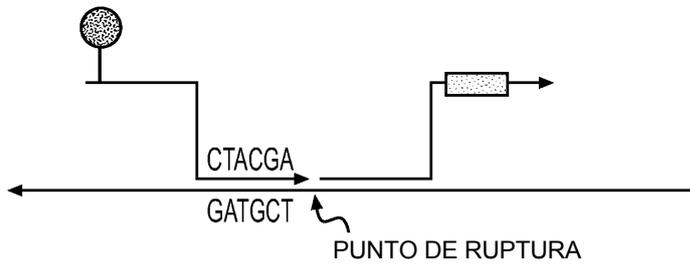


Figura 68B

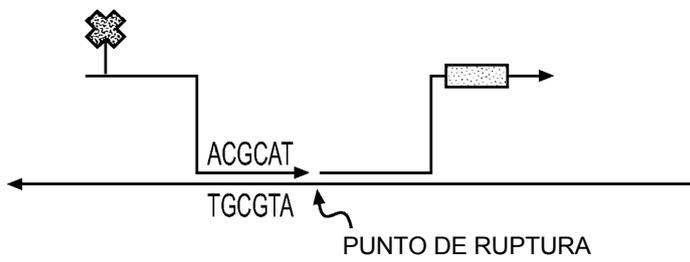


Figura 68C

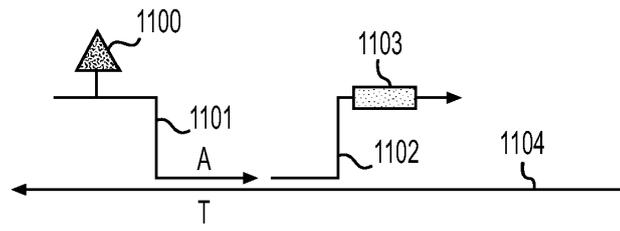


Figura 69A

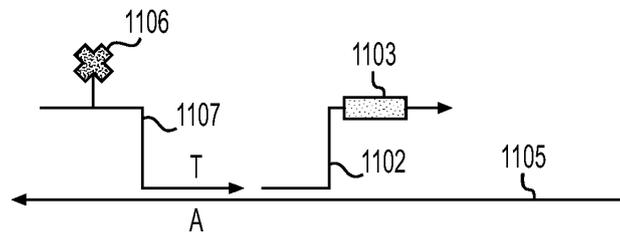


Figura 69B