

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 819 277**

51 Int. Cl.:

C12Q 1/68 (2008.01)

C12N 15/10 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **10.02.2015 PCT/EP2015/052723**

87 Fecha y número de publicación internacional: **20.08.2015 WO15121236**

96 Fecha de presentación y número de la solicitud europea: **10.02.2015 E 15708126 (6)**

97 Fecha y número de publicación de la concesión europea: **15.07.2020 EP 3105349**

54 Título: **Secuenciación dirigida y filtrado de UID**

30 Prioridad:

11.02.2014 US 201461938227 P
31.07.2014 US 201462031405 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
15.04.2021

73 Titular/es:

F. HOFFMANN-LA ROCHE AG (100.0%)
Grenzacherstrasse 124
4070 Basel, CH

72 Inventor/es:

VIGNEAULT, FRANCOIS y
DONAHUE, WILLIAM

74 Agente/Representante:

LINAGE GONZÁLEZ, Rafael

ES 2 819 277 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

Secuenciación dirigida y filtrado de UID

5 ANTECEDENTES

10 Muchas tecnologías actuales de secuenciación de próxima generación (NGS) usan una forma de secuenciación por síntesis (SBS). Las tecnologías de NGS tienen la capacidad de secuenciar masivamente en paralelo millones de moldes de ADN. Para lograr un alto rendimiento, muchos millones de moldes monocatenarios se agrupan en un chip y la secuencia de cada molde se lee de forma independiente. Las plataformas de NGS de segunda generación amplifican clonalmente los moldes de ADN en un soporte sólido, seguido de secuenciación cíclica. Las plataformas de NGS de tercera generación emplean protocolos libres de PCR de molécula única y química libre de ciclos (Schadt y col., Hum Mol Genet., 19 (R2):R227-40, (2010)).

15 Las principales limitaciones de los procedimientos de NGS y otros procedimientos de secuenciación de alto rendimiento incluyen errores y sesgos de secuenciación y amplificación. Debido a errores y sesgos asociados con la amplificación y secuenciación, estas tecnologías de secuenciación se desvían de la distribución uniforme ideal de las lecturas y pueden afectar muchas aplicaciones médicas y científicas. Para aplicaciones clínicas, los laboratorios deben verificar la exactitud de una mutación o una lectura de polimorfismo de un solo nucleótido (SNP) antes de informar a un paciente. Típicamente, la verificación de secuencia se realiza haciendo una biblioteca de Sanger de la diana después de obtener las secuencias y "calificar según Sanger" los resultados de secuenciación de próxima generación (NGS). Para superar la mayor tasa de error de las plataformas de NGS en comparación con la secuenciación tradicional de Sanger, se requiere un alto nivel de redundancia o cobertura de secuencia para leer con exactitud las secuencias nucleotídicas. Típicamente, se requiere una cobertura de 30-50x para lecturas de nucleótidos exactas, aunque esto puede variar en base a la exactitud de la plataforma de secuenciación, los procedimientos de detección de variantes y el material que se está secuenciando (Koboldt DC y col., Brief Bioinform., 11:484-98 (2010)). En general, todas las plataformas de segunda generación producen datos de una exactitud similar (98 - 99,5 %), confiando en una profundidad de secuencia adecuada, por ejemplo, cobertura) para realizar secuencias nucleotídicas de mayor exactitud.

20 El sesgo de secuenciación se puede manifestar como un sesgo de cobertura (desviación de una distribución uniforme de lecturas) y un sesgo de error (desviaciones de las tasas uniformes de emparejamiento erróneo, inserción y delección). Las tecnologías de secuenciación actuales son limitadas porque las químicas usadas en procedimientos de secuenciación de alto rendimiento están sesgadas inherentemente. Algunas secuencias de nucleótidos se leen con más frecuencia que otras secuencias y tienen una tasa de error inherente. Dependiendo de muchos factores, incluyendo la plataforma de secuenciación usada, los errores de lectura (la mayoría de los cuales son bases mal identificadas debido a secuencias nucleotídicas de baja calidad) se pueden producir en cualquier lugar en el intervalo de un error por cada 100-2000 bases. Si bien el sesgo de cobertura es una medida de secuenciación importante, las variaciones en la exactitud de la secuencia también son importantes.

25 Otra limitación importante es el sesgo de amplificación por PCR, porque las condiciones durante la construcción de la biblioteca de moldes de nucleótidos para la secuenciación pueden influir significativamente en el sesgo de secuenciación. Se ha demostrado que la amplificación por PCR para la construcción de bibliotecas es una fuente de error de datos de secuenciación (Keohavong P y col., PNAS86:9253-9257 (1989); Cariello y col., Nucleic Acids Res., 19:4193-4198 (1991); Cline y col., Nucleic Acids Res., 24:3546-3551 (1996)). Los procedimientos de construcción de la biblioteca pueden afectar la uniformidad de la cobertura. Por ejemplo, la amplificación por PCR también es una fuente conocida de subcobertura de regiones extremas de GC durante la construcción de la biblioteca (Aird y col., Genome Biol., 12: R18 (2011); Oyola y col., BMC Genomics, 13:1; 22 (2012); Benjamini y col., Nucleic Acids Res., 40:e72 (2012)). También se pueden introducir sesgos similares durante la PCR de puente para la amplificación de racimos y, en algunas plataformas de NGS, los errores específicos de cadena pueden dar lugar a sesgos de cobertura al afectar el rendimiento del alineador (Nakamura y col., Nucleic Acids Res., 39:e90 (2011)). Otras plataformas que utilizan una química libre de finalizadores pueden estar limitadas en su capacidad para secuenciar con exactitud homopolímeros largos, y también pueden ser sensibles a los sesgos de cobertura introducidos por la PCR de emulsión en la construcción de la biblioteca (Rothberg y col., Nature, 475:348-352 (2011); Margulies y col., Nature 2005, 437:376-380 (2005); Huse y col., Genome Biol., 8:R143 (2007); Merriman y col., Electrophoresis, 33:3397-3417 (2012)). Hashimony y col., Cell reports vol. 2, N.º 3 págs. 666-673 (2012), Head y col. BIOTECHNIQUES vol. 56, N.º 2, pág. 61 (2014), y el documento WO2013/13051 2 A2 han sugerido mejoras recientemente.

SUMARIO

60 La presente invención proporciona un procedimiento para generar una biblioteca de polinucleótidos que comprende:

(a) generar una primera secuencia del complemento (CS) de un polinucleótido diana a partir de una muestra usando un primer cebador, comprendiendo el primer cebador una secuencia específica de diana;

65 (b) unir por medio de ligadura a la primera CS un adaptador que comprende una primera secuencia de unión de cebador (PBS) o una parte de la misma, formando de este modo una secuencia del complemento modificada (MCS);

(c) extender un segundo cebador hibridado a la MCS, formando de este modo una segunda CS, en el que el segundo cebador comprende:

5 (i) una región específica de diana, y

(ii) una segunda PBS; y

10 (d) amplificar la segunda CS usando cebadores que se hibridan con la primera PBS y la segunda PBS respectivamente, en el que el primer o el segundo cebador comprende una secuencia de identificación única (UID).

15 El primer cebador puede comprender una secuencia de ligadura universal (ULS). El segundo cebador comprende además una secuencia de cebado universal (UPS). El adaptador puede comprender además una secuencia de código de barras de muestra (SBC). La MCS puede comprender además una molécula de afinidad o una secuencia de captura.

20 En un modo de realización, la UID comprende la secuencia NNNNNNNNNNNNNNNN (SEQ ID NO: 1), en la que N es cualquier residuo de ácido nucleico. En otro modo de realización, la UID comprende la secuencia NNNNNWNNNNNNWNNNNN (SEQ ID NO: 2), en la que N es cualquier residuo de ácido nucleico y W es adenina o timina.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

25 Los rasgos característicos novedosos descritos en el presente documento se exponen con particularidad en las reivindicaciones adjuntas. Se obtendrá una mejor comprensión de los rasgos característicos y ventajas de los rasgos característicos descritos en el presente documento haciendo referencia a la siguiente descripción detallada que expone ejemplos ilustrativos, en los que se utilizan los principios de los rasgos característicos descritos en el presente documento, y los dibujos adjuntos, en los que:

30 La **FIG. 1** representa un esquema de un procedimiento ejemplar para secuenciación dirigida descrito en el presente documento.

La **FIG. 2** representa un esquema de un procedimiento ejemplar para secuenciación dirigida descrito en el presente documento.

35 La **FIG. 3** representa un esquema de un proceso ejemplar para generar procedimientos de secuenciación dirigida mejorados. Se representan los tiempos de procesamiento.

40 La **FIG. 4** representa una tabla que muestra el porcentaje de especificidad en la diana usando los paneles de cebadores indicados de los procedimientos de secuenciación dirigida no mejorados y mejorados descritos en el presente documento y comparados con otros paneles de cebadores conocidos en la técnica (Otro N.º 1 y otro N.º 2). El panel Tex-1 es un panel portador de 23 genes (todos exón). CS-23 es un panel de cebadores centrado en rsSNP para 18 genes.

45 La **FIG. 5** representa un gráfico de cobertura de lectura diana con las condiciones de reacción indicadas. La fracción de genes por encima de la cobertura frente a la profundidad de lectura se muestra con las condiciones de reacción indicadas. Las condiciones que tienen un efecto positivo en la cobertura de secuencia se representan en negrita.

50 La **FIG. 6A** representa esquemas de las condiciones de rampa e hibridación para las etapas indicadas de un procedimiento ejemplar para secuenciación dirigida en condiciones menos rigurosas usado para un panel de 30 cebadores.

La **FIG. 6B** representa las concentraciones de cebadores en un panel de aproximadamente 350 cebadores usados en las condiciones de rampa e hibridación en 6A que eran insuficientes para generar suficiente producción diana.

55 La **FIG. 7** representa esquemas de procedimientos ejemplares para la secuenciación dirigida en condiciones de rampa e hibridación menos rigurosas (arriba) y más rigurosas (abajo). La rigurosidad se incrementó disminuyendo las velocidades de rampa para la segunda etapa de extensión del cebador. La rigurosidad se incrementó añadiendo una etapa de retención de 68 °C para la primera etapa de extensión del cebador. La rigurosidad se incrementó reduciendo la temperatura mínima de hibridación a 55 °C.

60 La **FIG. 8** representa las concentraciones de cebadores y los resultados donde se fijó la concentración por cebador (Grupo 1) y donde se fijó la concentración total de cebadores (Grupo 2) usando la fracción completa, media, cuarta o pequeña de cebadores de un panel de aproximadamente 350 cebadores en las condiciones menos rigurosas representadas en la FIG. 8.

65 La **FIG. 9A** representa los productos después de los ciclos de PCR indicados de un procedimiento de secuenciación dirigida ejemplar en un gel de agarosa junto con una escalera de 100 pares de bases (pb) sin aditivos añadidos. Se

muestran el producto diana y el producto de dímero.

La **FIG. 9B** representa los productos después de los ciclos de PCR indicados de un procedimiento de secuenciación dirigida ejemplar en un gel de agarosa junto con una escalera de 100 pares de bases (pb) con el aditivo betaína. Se muestran el producto diana y el producto de dímero. La **FIG. 9C** representa los productos después de los ciclos de PCR indicados de un procedimiento de secuenciación dirigida ejemplar en un gel de agarosa junto con una escalera de 100 pares de bases (pb) con el aditivo trehalosa. Se muestran el producto diana y el producto de dímero. La **FIG. 9D** representa los productos después de los ciclos de PCR indicados de un procedimiento de secuenciación dirigida ejemplar en un gel de agarosa junto con una escalera de 100 pares de bases (pb) con el aditivo cloruro de magnesio. Se muestran el producto diana y el producto de dímero. La **FIG. 9E** representa los productos después de los ciclos de PCR indicados de un procedimiento de secuenciación dirigida ejemplar en un gel de agarosa junto con una escalera de 100 pares de bases (pb) con el aditivo sulfato de amonio. Se muestran el producto diana y el producto de dímero.

La **FIG. 10** representa los productos después de 33 ciclos de PCR indicados de un procedimiento de secuenciación dirigida ejemplar en un gel de agarosa junto con una escalera de 100 pares de bases (pb) en las condiciones indicadas.

La **FIG. 11** representa un gráfico de análisis de secuencia de dímero de la longitud de la secuencia frente a la longitud de secuencia. Los productos de dímero correspondientes secuenciados se muestran en el gel de agarosa a la derecha.

La **FIG. 12** representa un diagrama que representa el mecanismo propuesto de formación de producto no deseado durante la segunda etapa de extensión del cebador como se determina mediante análisis de secuenciación de dímero. La formación de dímeros es facilitada por cebadores con altas temperaturas de fusión a una temperatura de hibridación baja. La formación de dímeros es facilitada por cebadores con alto contenido de GC que interactúan con la UID. La figura divulga las SEQ ID NO: 90-91, 92, 91 y 93, respectivamente, en orden de aparición.

La **FIG. 13** representa una tabla que muestra los genes y las enfermedades asociadas, el número de exones y el número de conjuntos de sondas del panel de cebadores ejemplar CS-350. La lista de exones sin cebador a la derecha indica exones para los que no se produjeron secuencias de cebador usando otros procedimientos de diseño de cebadores distintos de los descritos en el presente documento.

La **FIG. 14** representa un diagrama de criterios de exclusión usados para generar subpaneles de cebadores a partir de un panel de cebadores que contiene aproximadamente 350 cebadores.

La **FIG. 15A** representa un gráfico que muestra la especificidad en la diana y la uniformidad de cobertura en un límite de 100x del panel de cebadores indicado de aproximadamente 350 cebadores y subpaneles generados a partir del mismo usando los criterios de exclusión mostrados en la FIG. 14.

La **FIG. 15B** representa una tabla que muestra la especificidad en la diana, la uniformidad de cobertura y la profundidad de lectura media por amplicón en un límite de 100x del panel de cebadores indicado de aproximadamente 350 cebadores y subpaneles generados a partir del mismo usando los criterios de exclusión mostrados en la FIG. 14.

La **FIG. 16** representa un gráfico que muestra la uniformidad de cobertura sobre el intervalo límite *in silico* del panel indicado de aproximadamente 350 cebadores y subpaneles generados a partir del mismo usando los criterios de exclusión mostrados en la FIG. 14.

La **FIG. 17A** representa una gráfica que muestra la especificidad en la diana de un procedimiento ejemplar para secuenciación dirigida, descrito en el presente documento usando tres UID diferentes (BC_01, BC_02, BC_03).

La **FIG. 17B** representa los productos después de la PCR de un procedimiento de secuenciación dirigida ejemplar descrito en el presente documento usando tres UID diferentes.

La **FIG. 17C** representa una tabla con los valores correspondientes de la FIG. 17A.

La **FIG. 18** representa un gráfico del porcentaje de amplicones mayor que el 20 % de la media sobre el intervalo límite *in silico* indicado que muestra la uniformidad de cobertura de un procedimiento ejemplar para secuenciación dirigida, descrito en el presente documento usando tres UID diferentes.

La **FIG. 19** representa un gráfico que compara las lecturas sin procesar (sin UID) con la exactitud mejorada de UID.

La **FIG. 20** representa un esquema del flujo de trabajo de detección de SNP y análisis de secuencia usando un procedimiento ejemplar para secuenciación dirigida.

La **FIG. 21** representa un gráfico y una tabla correspondiente que muestran el porcentaje relativo de secuencias nucleotídicas de SNP que coinciden entre muestras usando un procedimiento ejemplar para secuenciación dirigida, descrito en el presente documento usando las UID indicadas. La reducción del ciclo con BC_6 dio como resultado un mayor número de moléculas únicas.

- 5 La **FIG. 22A** representa un gráfico del porcentaje de lectura de cada amplicón frente al contenido en % de GC de amplicón usando un procedimiento ejemplar para secuenciación dirigida, descrito en el presente documento. Una gran cantidad de ejecutores de bajo rendimiento están presentes en el Grupo A.
- La **FIG. 22B** representa un gráfico del porcentaje de lectura de cada amplicón frente al contenido en % de GC de amplicón usando un procedimiento ejemplar para secuenciación dirigida, descrito en el presente documento.
- 10 La **FIG. 23A** representa un gráfico de amplicones de bajo rendimiento mapeados por su respectiva temperatura de fusión de cebador.
- La **FIG. 23B** representa un gráfico de amplicones de bajo rendimiento mapeados por su respectiva temperatura de fusión de cebador.
- 15 La **FIG. 24** representa un gráfico de amplicones de bajo, medio y alto rendimiento mapeados por su respectiva temperatura de fusión de cebador.
- La **FIG. 25** representa una tabla que resume la configuración para un diseño de cebadores mejorado para su uso en procedimientos para secuenciación dirigida.
- 20 La **FIG. 26** representa esquemas de criterios mejorados de lectura de aciertos fuera de diana para su uso en procedimientos para secuenciación dirigida.
- La **FIG. 27** representa un esquema de diseño de cebadores mejorado para su uso en procedimientos para secuenciación dirigida. Un diseño de cebadores mejorado es añadir una secuencia de tampón de intrón de aproximadamente 20 nt. Un diseño de cebadores mejorado consiste en exones divididos uniformemente para una mejor cobertura y una mayor flexibilidad.
- 25 La **FIG. 28** representa una gráfica que muestra un panel de cebadores mejorado (v.3.0) diseñado usando el procedimiento de diseño de cebadores descrito en el presente documento en comparación con el panel de cebadores diseñado usando un procedimiento de la técnica anterior (v.1.0). El panel mejorado da lugar a una mayor eficacia de amplificación y aumenta el número de moléculas únicas detectadas, dando lugar a capacidades mejoradas de lecturas de SNP, una reducción en los requisitos de cobertura de secuenciación y reduce los requisitos de entrada de muestra.
- 30 La **FIG. 29** representa un gráfico que compara la uniformidad de cobertura de cebadores en el subpanel indicado que se ajustan a los criterios de diseño de cebadores mejorados frente a los cebadores en el mismo subpanel que no se ajustan a los criterios de diseño de cebadores mejorados. Los cebadores en el subpanel que se ajustan a los criterios de diseño de cebadores mejorados demuestran una mayor uniformidad de cobertura, mayor especificidad en la diana y mayores recuentos de lectura.
- 35 La **FIG. 30A** representa una gráfica que muestra el rendimiento de la muestra de sangre completa con respecto a la uniformidad y la cobertura y la especificidad en la diana en comparación con una muestra de ADN extraída de sangre completa.
- 40 La **FIG. 30B** representa los productos después de los ciclos de PCR indicados de un procedimiento de secuenciación dirigida ejemplar usando los volúmenes indicados de una muestra de sangre completa en un gel de agarosa junto con una escalera de 100 pares de bases (pb). Se puede usar tan solo 1 μ L de sangre completa.
- 45 La **FIG. 31** representa una gráfica (arriba) y la tabla correspondiente (abajo) de un análisis que compara el número de amplicones con más de 10 moléculas únicas de una muestra de sangre completa y una muestra de ADN extraída de la sangre completa. La muestra de sangre completa 3x combina tres primeras reacciones de extensión del cebador antes de la ligadura del adaptador.
- 50 La **FIG. 32** representa gráficos que muestran diferencias de lecturas de SNP entre una muestra de sangre completa y una muestra de ADN extraída de la sangre completa. El gráfico superior muestra las lecturas de SNP perdidas usando la muestra de sangre completa. El gráfico superior muestra las lecturas de SNP perdidas usando la muestra de ADN extraída de la sangre completa. La figura divulga las SEQ ID NO: 94-95, 94-96 y 96- 97, respectivamente, en orden de aparición.
- 55 La **FIG. 33** representa una gráfica que muestra el rendimiento de la muestra de tejido prostático con FFPE con respecto a la uniformidad y cobertura y la especificidad en la diana en comparación con una muestra de sangre completa y una muestra de ADN extraída de sangre completa.
- 60 La **FIG. 34** representa los productos de un procedimiento de secuenciación dirigida ejemplar usando una variedad de muestras en un gel de agarosa junto con una escalera de 100 pares de bases (pb). Los procedimientos descritos en el presente documento pueden adaptarse a una variedad de muestras que incluyen la entrada directa de sangre completa o saliva en la primera reacción de extensión del cebador sin extracción previa de nucleótidos, muestras bucales y muestras
- 65

de FFPE.

5 La **FIG. 35** representa una gráfica (arriba) y la tabla correspondiente (abajo) de un análisis que compara el número de amplicones con más de 10 moléculas únicas de una muestra de FFPE, una muestra de sangre completa y una muestra de ADN extraída de la sangre completa. La muestra de sangre completa 3x combina tres primeras reacciones de extensión del cebador antes de la ligadura del adaptador.

10 La **FIG. 36** representa una gráfica del número de moléculas únicas detectadas usando el número indicado de ciclos de PCR y las UID indicadas. La gráfica demuestra que la reducción del número de ciclos de PCR evita la formación de productos más grandes sobreamplificados, reduce la duplicación de PCR, puede permitir reducciones en la profundidad de secuenciación requerida, puede mejorar los datos para muestras de baja entrada y puede aprovechar la amplificación lineal para compensar los ciclos de PCR reducidos.

15 La **FIG. 37** representa una gráfica que representa la calidad de los datos de secuenciación usando una biblioteca producida a partir de un procedimiento de secuenciación dirigida ejemplar que ha sido purificada en gel en comparación con una biblioteca que ha sido purificada en Ampure.

20 La **FIG. 38** representa una gráfica de una titulación de lectura *in silico* que muestra el porcentaje de amplicones con una cobertura de molécula única mayor o igual a 10. La secuenciación a una profundidad de lectura promedio de 500x por amplicón proporcionó una cobertura de molécula única adecuada para el 95 % del amplicón en el panel de cebadores Tex_01 (336 amplicones). Esto puede permitir una multiplexación de 90 muestras por ejecución (336 x 500 = 168.000 lecturas).

25 La **FIG. 39** representa una tabla del número esperado y real de secuencias para cada muestra con código de barras y el porcentaje de las lecturas totales de secuencia por muestra con código de barras.

30 La **FIG. 40** representa una gráfica de cuantificación del número de copias que muestra la proporción de moléculas únicas capturadas por gen. La proporción de lecturas únicas (filtradas por UID) para un gen dado se comparó para genes en cromosomas autosómicos frente al cromosoma X. Se representa la proporción de lecturas entre un paciente de referencia masculino y tres pacientes de prueba. Esto demuestra la capacidad cuantitativa del uso del análisis de UID para la secuenciación dirigida.

35 La **FIG. 41** representa un esquema de un procedimiento basado en ARN ejemplar para la secuenciación dirigida de extensión de cebador con una capacidad demostrada para amplificar productos de más de 700 pb de longitud.

La **FIG. 42A** representa los productos de un procedimiento de secuenciación dirigida de ARN ejemplar después de los ciclos de PCR indicados usando las cantidades de entrada de ARN indicadas en un gel de agarosa junto con una escalera de 100 pares de bases (pb).

40 La **FIG. 42B** representa una lista ejemplar de dianas a las que se han aplicado con éxito los procedimientos ejemplares de secuenciación dirigida de ARN descritos en el presente documento.

45 La **FIG. 43** representa los productos de un procedimiento de secuenciación dirigida ejemplar en un gel de agarosa junto con una escalera de 100 pares de bases (pb) realizado como repeticiones técnicas. Esto demuestra la reproducibilidad de los procedimientos descritos en el presente documento.

La **FIG. 44** representa un esquema de software de diseño de cebadores ejemplar desarrollado para producir paneles de cebadores.

50 La **FIG. 45A** representa una curva del porcentaje de amplicones mayor que el 20 % de la media sobre el intervalo límite *in silico* indicado que muestra la uniformidad de cobertura de un procedimiento ejemplar para la secuenciación dirigida descrita en el presente documento usando un subpanel de cebadores en las coberturas de lectura de pliegue indicadas normalizadas a un promedio de 100x.

55 La **FIG. 45B** representa una gráfica que compara la uniformidad de cobertura del panel de cebadores indicado de aproximadamente 350 cebadores y un subpanel generado a partir del mismo usando los criterios de exclusión mostrados en la FIG. 15 y otros procedimientos descritos en el presente documento.

60 La **FIG. 46** representa un gráfico que resume las medidas de calidad de los procedimientos descritos en el presente documento.

65 La **FIG. 47** representa los productos del procedimiento de secuenciación dirigida de ADN en un gel de agarosa al 2 % junto con una escalera de 100 pares de bases (pb). Se muestran muestras de 2 pacientes (B1 y B2) después de los ciclos de amplificación por PCR indicados.

La **FIG. 48** representa productos de un procedimiento de secuenciación dirigida de ARN en un gel de agarosa al 2 % junto

con una escalera de 100 pb. Se muestran muestras de 2 pacientes (B1 y B2) después de los ciclos de amplificación por PCR indicados. Para cada paciente, se realizó una titulación del material de entrada de ARN de partida desde 1000 ng hasta 1 ng.

5 La **FIG. 49** representa histogramas de resultados usando un proceso de filtrado de datos posterior a la secuenciación de próxima generación (NGS) usando un procedimiento de secuenciación dirigida. El histograma en la **FIG. 49B** es una versión a escala logarítmica de 49A. La NGS se realizó usando un enfoque de lectura de extremo emparejado (R1 y R2), proporcionando un total de ~ 6 millones de lecturas para la muestra mostrada. Se analizaron además las secuencias con una puntuación Q phred de 30 o superior (calidad aprobada R1 y R2). Los datos de secuencia se consultaron a continuación para determinar la presencia de un panel de cebadores esperado usado en el protocolo de la biblioteca dirigida de ADN (cebador aprobado R1 y R2). Se descartó cualquier lectura de secuencia que no comenzara con una de las secuencias de cebador esperadas. Para cada lectura con una secuencia de cebador conocida o esperada en R1, se calificó el cebador esperado en R2 (emparejado R1 y R2). Por lo tanto, cuando un cebador R1 conocido no coincide con un cebador diana diferente en R2 (o viceversa), corresponde a un producto de amplificación no específico (se muestra en gris claro). Si un cebador R1 conocido no coincide con un cebador diana diferente en R2 (o viceversa), corresponde a un producto de amplificación específico (mostrado en gris oscuro).

20 Las **FIG. 50A-50C** representan gráficas de recuentos de lectura de secuenciación de paneles dirigidos de ADN. Cada gen indicado estaba dirigido por un par de cebadores específico usado en la preparación de la muestra de ADN. La **FIG. 50A** muestra gráficas de recuentos de lectura de secuenciación de paneles dirigidos de ADN usando un primer par de cebadores (BC3) sin una UID (izquierda) y con una UID (derecha). La **FIG. 50B** muestra gráficas de recuentos de lectura de secuenciación de paneles dirigidos de ADN usando un segundo par de cebadores (BC1) sin una UID (izquierda) y con una UID (derecha) usando un procedimiento de secuenciación dirigida. La **FIG. 50C** muestra una gráfica de recuentos de lectura de secuenciación de paneles dirigidos de ADN usando un procedimiento de secuenciación dirigida con filtrado de UID posterior.

30 Las **FIG. 51A-51B** representan el recuento de lectura de secuenciación de paneles dirigidos de ARN usando un procedimiento de secuenciación dirigida con filtrado de UID. Cada transcripción de gen indicada fue dirigida por un par de cebadores específico usado en la preparación de la muestra de ARN. La **FIG. 51A** representa una gráfica de recuentos de lectura de secuenciación de paneles dirigidos de ARN (izquierda) y una gráfica de las frecuencias de lectura de secuenciación (derecha). La **FIG. 51B** es una versión a escala logarítmica de la gráfica mostrada en la **FIG. 51A** (izquierda) de las frecuencias de lectura de secuenciación. Los datos que se muestran aquí representan el filtrado posterior al recuento de lectura/expresión.

35 La **FIG. 52** representa un gráfico de resultados de un análisis de especificidad diana para las dianas y condiciones indicadas usando un procedimiento de secuenciación dirigida. Se probaron diversas condiciones de protocolo (número de ciclos, tampones, condiciones de hibridación, etc.). Como se muestra, se logró el 99,2 % de especificidad diana en algunas condiciones. (por ejemplo, el 99,2 % de las lecturas de secuenciación fueron la diana deseada con una amplificación no específica mínima).

40 La **FIG. 53** representa un gráfico de distribución de UID para las dianas y condiciones indicadas usando un procedimiento de secuenciación dirigida con filtrado de UID. Se probaron diversas condiciones de protocolo (número de ciclos, tampones, condiciones de hibridación, etc.). El número de secuencias sin procesar por UID puede variar dependiendo de las condiciones usadas.

45 La **FIG. 54** representa un gráfico del aumento supuesto en la puntuación de phred de exactitud de secuenciación (Q) en relación con el número de lecturas por secuencia de UID usando un procedimiento de secuenciación dirigida con filtrado de UID.

50 La **FIG. 55** representa un gráfico que muestra la mejora de la exactitud de cada diana indicada usando un procedimiento de secuenciación dirigida cuando se aplica el filtrado de UID.

55 La **FIG. 56** representa un gráfico de análisis de consenso de UID y exactitud del análisis de genotipificación de SNP usando un procedimiento de secuenciación dirigida con filtrado de UID. Se probaron diversas regiones diana de ADN (eje y) frente a diversas condiciones experimentales (eje x). La secuencia consenso para cada diana indicada se muestra en gris, y la mutación/SNP se muestran en blanco. Los genes homocigóticos están dominados por el gris. Los genes heterocigóticos se indican mediante aproximadamente ~ 50 % de sus secuencias que muestran una secuencia común en blanco. Las mutaciones e indel causadas por PCR o los errores de secuenciación se muestran en negro.

60 La **FIG. 57** representa un análisis de secuencia del gen GBA usando un procedimiento de secuenciación dirigida con filtrado de UID. Ambos alelos del gen GBA de una muestra de paciente se alinearon usando Clustal W. El paciente muestra heterocigosis posterior al filtrado de UID. Ambos alelos se compararon con la referencia del genoma humano Ensembl. La falta de un "*" denota una alineación de emparejamiento erróneo entre una de las 3 secuencias. El gen GBA del paciente presentado aquí tiene un alelo idéntico al genoma de referencia humano, y un segundo alelo con 6 polimorfismos/mutaciones de secuencia observados. La figura divulga las SEQ ID NO: 98-100, respectivamente, en orden de aparición.

La **FIG. 58** representa el análisis de genotipificación de SNP del gen GBA en el mismo paciente analizado en la **FIG. 12**. Los datos presentados aquí muestran la localización de un SNP patógeno presente en la diana del gen GBA (rs1064644) descubierto usando los procedimientos de secuenciación dirigida con el filtrado de UID descrito en el presente documento.

La **FIG. 59** es un diagrama de bloques que ilustra una primera arquitectura ejemplar de un sistema informático que se puede usar en relación con modos de realización ejemplares de la presente invención.

La **FIG. 60** es un diagrama que ilustra una red informática que se puede usar en relación con modos de realización ejemplares de la presente invención.

DESCRIPCIÓN DETALLADA

Como se usa en el presente documento, la amplificación comprende realizar una reacción de amplificación. Un producto de una reacción de extensión del cebador puede comprender la secuencia del cebador junto con el complemento del molde producido durante la extensión del cebador. En algunos modos de realización, las reacciones de amplificación comprenden la extensión de dos cebadores, cada uno hibridado con una cadena complementaria de un polinucleótido. La amplificación de polinucleótidos se puede realizar mediante cualquier medio conocido en la técnica. Los polinucleótidos se pueden amplificar por reacción en cadena de la polimerasa (PCR) o por amplificación isotérmica de ADN.

Una reacción de amplificación puede comprender uno o más aditivos. En algunos modos de realización, los uno o más aditivos son dimetilsulfóxido (DMSO), glicerol, betaina (mono)hidrato (*N,N,N*-trimetilglicina = [caroximetil] trimetilamonio), trehalosa, 7-Deaza-2'-desoxiguanosina trifosfato (dC7GTP o 7-deaza-2'-dGTP), BSA (albúmina de suero bovino), formamida (metanamida), sulfato de amonio, cloruro de magnesio, cloruro de tetrametilamonio (TMAC), otros derivados de tetraalquilamonio (por ejemplo, cloruro de tetraetilamonio (TEA-Cl) y cloruro de tetrapropilamonio (TPra-Cl), detergente no iónico (por ejemplo, Triton X-100, Tween 20, Nonidet P-40 (NP-40)) o PREXCEL-Q. En algunos modos de realización, una reacción de amplificación puede comprender 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10 aditivos diferentes. En otros casos, una reacción de amplificación puede comprender al menos 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10 aditivos diferentes. En algunos modos de realización, una reacción de extensión, transcripción inversa o amplificación que comprende uno o más aditivos se puede caracterizar por un aumento.

Como se usa en el presente documento, una reacción en cadena de la polimerasa (PCR) comprende una reacción de amplificación *in vitro* de secuencias de polinucleótidos específicos mediante la extensión simultánea de cebadores de cadenas complementarias de un polinucleótido bicatenario. Las reacciones de PCR producen copias de un polinucleótido molde flanqueado por sitios de unión de cebador. El resultado, con dos cebadores, es un aumento exponencial en el número de copias del polinucleótido molde de ambas cadenas con cada ciclo, porque con cada ciclo ambas cadenas se replican. El dúplex polinucleotídico tiene extremos correspondientes a los extremos de los cebadores usados. La PCR puede comprender una o más repeticiones de desnaturar un polinucleótido molde, hibridar cebadores en sitios de unión de cebadores y extender los cebadores mediante una ADN o ARN polimerasa en presencia de nucleótidos. Las temperaturas particulares, las duraciones en cada etapa y las tasas de cambio entre etapas dependen de muchos factores bien conocidos por los expertos en la técnica. (McPherson y col., IRL Press, Oxford (1991 y 1995)). Por ejemplo, en una PCR convencional que usa la ADN polimerasa Taq, un polinucleótido molde bicatenario se puede desnaturar a una temperatura > 90 °C, los cebadores se pueden hibridar a una temperatura en el intervalo de 50-75 °C, y los cebadores se pueden extender a una temperatura en el intervalo de 72-78 °C. En algunos modos de realización, la PCR comprende PCR de transcripción inversa (RT-PCR), PCR en tiempo real, PCR con cebadores internos, PCR cuantitativa, PCR multiplexada o similares. En algunos modos de realización, la PCR no comprende RT-PCR. (Patentes de los Estados Unidos n.º 5.168.038, 5.210.015, 6.174.670, 6.569.627 y 5.925.517; Mackay y col., Nucleic Acids Research, 30:1292-1305 (2002)). La RT-PCR comprende una reacción de PCR precedida por una reacción de transcripción inversa y se amplifica un ADNc resultante, la PCR con cebadores internos comprende una PCR de dos fases en la que un amplicón de una primera reacción de PCR usando un primer conjunto de cebadores se convierte en la muestra para una segunda reacción de PCR usando un segundo conjunto de cebadores, al menos uno de los cuales se une a una localización interior de un amplicón de una primera reacción de PCR. La PCR multiplexada comprende una reacción de PCR, en la que una pluralidad de secuencias de polinucleótidos se somete a la PCR en la misma mezcla de reacción simultáneamente. Los volúmenes de reacción de PCR pueden situarse en cualquier punto de 0,2 nL-1000 µL. La PCR cuantitativa comprende una reacción de PCR diseñada para medir una cantidad, abundancia o concentración absoluta o relativa de una o más secuencias en una muestra. Las mediciones cuantitativas pueden incluir comparar una o más secuencias de referencia o estándares con una secuencia de polinucleótidos de interés. (Freeman y col., Biotechniques, 26: 112-126 (1999); Becker-Andre y col., Nucleic Acids Research, 17: 9437-9447 (1989); Zimmerman y col., Biotechniques, 21: 268-279 (1996); Diviacco y col., Gene, 122: 3013-3020 (1992); Becker-Andre y col., Nucleic Acids Research, 17: 9437-9446 (1989)).

Como se usa en el presente documento, un alelo puede ser una secuencia genética específica dentro de una célula, individuo o población que difiere de otras secuencias del mismo gen en la secuencia de al menos un sitio variante dentro de la secuencia del gen. Las secuencias de sitios variantes que difieren entre alelos diferentes pueden ser variantes, tales como polimorfismos o mutaciones. Las variantes pueden comprender mutaciones puntuales, polimorfismos, polimorfismos de un solo nucleótido (SNPs), variaciones de un solo nucleótido (SNV), translocaciones, inserciones, deleciones, amplificaciones, inversiones, deleciones intersticiales, variaciones del número de copias (CNV), pérdida de heterocigosis

o cualquier combinación de los mismos. Una muestra es "heterocigótica" en un locus cromosómico si tiene dos alelos diferentes en ese locus. Una muestra es "homocigótica" en un locus cromosómico si tiene dos alelos idénticos en ese locus.

5 En algunos modos de realización, las variantes pueden incluir cambios que afectan a un polipéptido, tales como un cambio en el nivel de expresión, secuencia, función, localización, ligandos o cualquier combinación de los mismos. En algunos modos de realización, una variación genética puede ser una mutación con desplazamiento del marco de lectura, mutación finalizadora, mutación de aminoácido, mutación neutra o mutación silenciosa. Por ejemplo, las diferencias de secuencia, cuando se comparan con una secuencia de nucleótidos de referencia, pueden incluir la inserción o delección de un solo nucleótido, o de más de un nucleótido, dando como resultado un desplazamiento del marco de lectura; el cambio de al menos un nucleótido, dando como resultado un cambio en el aminoácido codificado; el cambio de al menos un nucleótido, dando como resultado la generación de un codón de terminación prematuro; la delección de varios nucleótidos, dando como resultado una delección de uno o más aminoácidos codificados por los nucleótidos; la inserción de uno o varios nucleótidos, tal como por recombinación desigual o conversión génica, dando como resultado una interrupción de la secuencia codificante de un marco de lectura; duplicación de todo o parte de una secuencia; transposición; o un reordenamiento de una secuencia de nucleótidos. Dichos cambios de secuencia pueden alterar el polipéptido codificado por el ácido nucleico, por ejemplo, si el cambio en la secuencia de ácido nucleico provoca un desplazamiento del marco de lectura, el desplazamiento del marco de lectura puede dar como resultado un cambio en los aminoácidos codificados y/o puede dar como resultado la generación de un codón de terminación prematuro, que provoca la generación de un polipéptido truncado. En algunos modos de realización, una variante puede ser un cambio sinónimo en uno o más nucleótidos, por ejemplo, un cambio que no da como resultado un cambio en la secuencia de aminoácidos. Dicho polimorfismo puede, por ejemplo, alterar sitios de empalme, afectar la estabilidad o el transporte de ARNm, o afectar de otro modo la transcripción o traducción de un polipéptido codificado. En algunos modos de realización, una mutación sinónima puede dar como resultado que el producto de polipéptido tenga una estructura alterada debido al uso de codones raros que afecta el plegamiento del polipéptido durante la traducción, que en algunos casos puede alterar su función y/o propiedades de unión al fármaco si es una diana farmacológica. En algunos modos de realización, los cambios que pueden alterar el ADN aumentan la posibilidad de que se produzcan cambios estructurales, tales como amplificaciones o delecciones, a nivel somático.

30 Como se usa en el presente documento, un polimorfismo puede ser una aparición de dos o más secuencias o alelos alternativos determinados genéticamente en una población. Un polimórfico o sitio comprende el locus en el que se produce la divergencia. En algunos modos de realización, los polimorfismos se producen a una frecuencia inferior al 0,5 %, 1 %, 2 % o 5 %. En algunos modos de realización, los polimorfismos se producen a una frecuencia superior al 1 %, 5 %, 10 %, 20 % o 30 %. En algunos modos de realización, los biomarcadores tienen al menos dos alelos, cada uno de los cuales se produce a una frecuencia superior al 1 %, 5 %, 10 % o 20 % en una población seleccionada. En algunos modos de realización, los polimorfismos comprenden secuencias virales o bacterianas y se producen a una frecuencia inferior al 0,5 %, 1 %, 2 % o 5 % en una población seleccionada. Un polimorfismo puede incluir una o más variantes que incluyen cambios de base, inserciones, repeticiones o delecciones de una o más bases. Los polimorfismos pueden incluir polimorfismos de un solo nucleótido (SNP). Las variantes de número de copia (CNV), las transversiones y otras reorganizaciones también son formas de variantes. Los polimorfismos incluyen polimorfismos de longitud de fragmentos de restricción, número variable de repeticiones en tándem (VNTR), regiones hipervariables, minisatélites, repeticiones de dinucleótidos, repeticiones de trinucleótidos, repeticiones de tetranucleótidos, repeticiones de secuencias simples y elementos de inserción. La secuencia de alelos más frecuente de una población seleccionada puede ser el alelo natural. Los organismos diploides pueden ser homocigóticos o heterocigóticos para los alelos.

45 Como se usa en el presente documento, la genotipificación comprende determinar la secuencia genética de un sujeto en una o más posiciones genómicas. Por ejemplo, la genotipificación puede incluir determinar qué alelo o alelos tiene un sujeto para un solo SNP o dos o más SNP. Un sujeto diploide puede ser homocigótico para cada uno de los dos alelos posibles o heterocigóticos. Las células normales heterocigóticas en uno o más loci pueden dar lugar a células tumorales homocigóticas en esos loci. Esta pérdida de heterocigosis (LOH) puede ser el resultado de la delección de genes normales, la pérdida del cromosoma que porta el gen normal, la recombinación mitótica o la pérdida de un cromosoma con un gen normal y la duplicación de un cromosoma con un gen delecionado o inactivado. La LOH puede ser neutra a la copia o puede ser el resultado de una delección o amplificación.

55 Como se usa en el presente documento, un sujeto, un individuo y un paciente incluyen organismos vivos tales como mamíferos. Los ejemplos de sujetos y huéspedes incluyen, pero no se limitan a, caballos, vacas, camellos, ovejas, cerdos, cabras, perros, gatos, conejos, cobayas, ratas, ratones (por ejemplo, ratones humanizados), jerbos, primates no humanos (por ejemplo, macacos), seres humanos y similares, no mamíferos, incluyendo, por ejemplo, vertebrados no mamíferos, tales como aves (por ejemplo, pollos o patos), peces (por ejemplo, tiburones) o ranas (por ejemplo, *Xenopus*) e invertebrados no mamíferos, así como especies transgénicas de los mismos. En determinados aspectos, un sujeto se refiere a un solo organismo (por ejemplo, ser humano). En determinados aspectos, o se proporciona un grupo de individuos que componen una pequeña cohorte que tiene un factor inmune común al estudio y/o enfermedad, y/o una cohorte de individuos sin la enfermedad (por ejemplo, control negativo/normal). A un sujeto del que se obtienen muestras se le puede provocar una enfermedad y/o trastorno (por ejemplo, una o más alergias, infecciones, cánceres o trastornos autoinmunitarios o similares) y se puede comparar con un sujeto de control negativo que no se ve afectado por la enfermedad.

PROCEDIMIENTOS DE SECUENCIACIÓN DIRIGIDA EN GENERAL

5 Los procedimientos descritos aquí se pueden usar para generar una biblioteca de polinucleótidos para secuenciación. La
 10 secuencia determinada para un polinucleótido en una muestra se puede determinar con alta exactitud y confianza en
 15 secuencias nucleotídicas procedentes de la secuenciación. Los procedimientos pueden comprender específicamente,
 dirigirse a, codificar, modificar, amplificar, secuenciar y/o cuantificar de manera única secuencias de ADN o ARN presentes
 en la muestra. Estos procedimientos permiten la adición de secuencias que pueden formatear una biblioteca de
 amplicones de polinucleótidos para secuenciación u otros análisis moleculares. La biblioteca de secuenciación producida
 mediante estos procedimientos puede incorporar una UID que puede permitir la agrupación de lecturas de secuencia
 derivadas de la misma molécula inicial de ARN o ADN en la muestra. Estos procedimientos pueden permitir determinar si
 una variante de secuencia observada encontrada en una población de moléculas de ARN o ADN es un verdadero
 polimorfismo o mutación, o la variante de secuencia observada como resultado de un artefacto de amplificación, tal como
 un error o sesgo de amplificación. En cualquiera de los procedimientos descritos en el presente documento, se contempla
 que la UID es opcional. Por tanto, cualquier mención de "UID" se refiere a una UID opcional.

Estos incluyen procedimientos para preparar una biblioteca de polinucleótidos generados usando cebadores específicos
 de diana para ser secuenciados en una plataforma de NGS. Muchas dianas biológicas, tales como las de una muestra de
 paciente biológica, se pueden analizar a partir de la biblioteca compatible con NGS después de la secuenciación. Los
 20 procedimientos permiten la identificación de frecuencias diana (por ejemplo, expresión génica o distribución alélica). Los
 procedimientos también permiten la identificación y mutaciones o SNP en un genoma o transcriptoma, tales como las de
 un sujeto enfermo o no enfermo, del que se puede derivar información exacta de la secuencia. Los procedimientos también
 permiten determinar la presencia o ausencia de contaminación o infecciones en una muestra biológica de un sujeto, tal
 como mediante el uso de cebadores específicos de diana para organismos o virus exógenos, tales como una bacteria o
 un hongo.

Los procedimientos descritos en el presente documento ofrecen un equilibrio ventajoso de sensibilidad y especificidad y
 ventajas conferidas por reacciones de extensión de cebador lineal y/o marcado de UID. En algunos modos de realización,
 los procedimientos están diseñados para tamaños de panel más pequeños, tales como paneles de interés clínico. Estos
 30 procedimientos pueden tener costes iniciales muy bajos, se pueden realizar rápidamente y se pueden modificar para
 dianas de ARN o ADN. Además, el diseño de cebadores para su uso en estos procedimientos no es engorroso y es similar
 a la facilidad de diseño de cebadores para reacciones de PCR estándar. Los procedimientos se pueden usar para
 formatear bibliotecas de polinucleótidos para una variedad de secuenciación y otros análisis moleculares. Adicionalmente,
 se pueden realizar diversas aplicaciones de forma individual o simultánea. Por ejemplo, la secuenciación de dianas
 35 requeridas para el perfil de mutación del cáncer, el análisis de SNP y mutaciones, las pruebas a portadores, la detección
 de infecciones, el diagnóstico de enfermedades y el análisis de la expresión génica se pueden realizar de forma individual
 o simultánea.

40 **Direccionamiento inicial: formación de polinucleótidos marcados con UID complementarios a los polinucleótidos diana**

Dependiendo del tipo de polinucleótido diana que se va a analizar, los procedimientos pueden utilizar la transcripción
 inversa (RT) o la extensión de cebador (PE). Una reacción de extensión de cebador puede ser una sola etapa de extensión
 de cebador. Una reacción de extensión de cebador puede comprender extender uno o más cebadores individuales una
 45 vez. Una reacción de extensión de cebador puede comprender extender uno o más cebadores individuales en una sola
 etapa. En algunos modos de realización, se pueden generar polinucleótidos complementarios a dianas de ADN realizando
 reacciones de extensión de cebador. Por ejemplo, los polinucleótidos marcados con UID complementarios a las dianas
 de ADN se pueden generar realizando reacciones de extensión de cebador. En algunos modos de realización, las
 50 secuencias del complemento de polinucleótidos diana, tales como polinucleótidos marcados con UID complementarios a
 las dianas de ARN, se pueden generar realizando reacciones de transcripción inversa. Secuencias del complemento de
 polinucleótidos diana, tales como polinucleótidos marcados con UID complementarios a las dianas de ARN, se pueden
 generar realizando reacciones de transcripción inversa. Un polinucleótido diana incluye polinucleótidos presentes en una
 muestra inicialmente.

55 Como se usa en el presente documento, una "secuencia del complemento de polinucleótido diana" es un polinucleótido
 que comprende una secuencia complementaria a una secuencia diana o un complemento de la misma (complemento de
 una secuencia complementaria a una secuencia diana). En algunos modos de realización, una secuencia del
 complemento de polinucleótidos diana comprende una primera secuencia del complemento. Una "primera secuencia del
 60 complemento" es un polinucleótido transcrito inversamente a partir de un polinucleótido diana o formado a partir de una
 reacción de extensión de cebador en un polinucleótido diana. En algunos modos de realización, una secuencia del
 complemento de polinucleótido diana comprende una secuencia del complemento modificada. Una "secuencia del
 complemento modificada" es un polinucleótido transcrito inversamente a partir de un polinucleótido diana o formado a
 partir de una reacción de extensión de cebador en un polinucleótido diana, que comprende un adaptador. En algunos
 65 modos de realización, una secuencia del complemento de polinucleótido diana comprende una segunda secuencia del
 complemento. Una "segunda secuencia del complemento" es un polinucleótido que comprende una secuencia
 complementaria a una primera secuencia del complemento o secuencia del complemento modificada. En algunos modos

de realización, una secuencia del complemento de polinucleótido diana comprende una UID. Por ejemplo, una primera secuencia del complemento puede comprender una UID. Por ejemplo, una secuencia del complemento modificada puede comprender una UID. Por ejemplo, una segunda secuencia del complemento puede comprender una UID. Por ejemplo, una segunda secuencia del complemento puede comprender una secuencia complementaria a una UID de una primera secuencia del complemento o secuencia del complemento modificada. En algunos modos de realización, una secuencia del complemento de polinucleótido diana no comprende una UID. Por ejemplo, una primera secuencia del complemento puede no comprender una UID. Por ejemplo, una secuencia del complemento modificada puede no comprender una UID. Por ejemplo, una segunda secuencia del complemento puede no comprender una UID.

Los procedimientos pueden comprender una reacción de RT o PE en una primera etapa. Los procedimientos pueden comprender una reacción de extensión de cebador lineal en una etapa posterior. Una reacción de extensión de cebador lineal puede dar como resultado una amplificación lineal en lugar de una amplificación exponencial. Para la secuenciación dirigida de muchos polinucleótidos, cada cebador específico de diana individual puede tener cierto grado de variación de eficacia causada por variaciones en la extensión por diversas enzimas, o diferencias en la eficacia de hibridación a sus dianas respectivas. Esto puede crear un sesgo que se puede extender exponencialmente mediante PCR. Los procedimientos descritos en el presente documento pueden utilizar la extensión de cebador lineal para reducir o evitar este sesgo, lo que da como resultado una reducción o evitación de la frecuencia de variación de las dianas una con respecto a otra y puede proporcionar una mayor confianza y frecuencia, o análisis y exactitud de secuencias nucleotídicas procedentes de la secuenciación. Se ha descubierto que los procedimientos descritos en el presente documento evitan estos problemas de sesgo y pueden mantener una representación de frecuencia real del grupo inicial de dianas. En algunos modos de realización, la única reacción de amplificación exponencial, tal como una reacción de PCR, realizada en los procedimientos está en una fase final de generación de biblioteca y puede utilizar un conjunto de cebadores universal. En estos modos de realización, todas las dianas se pueden amplificar uniformemente durante una etapa de amplificación exponencial sin la introducción de variación o sesgo específico del gen.

Transcripción inversa (RT de polinucleótidos diana para formar polinucleótidos complementarios marcados con UID)

Usando cebadores descritos en el presente documento, los polinucleótidos de ARN se pueden transcribir inversamente usando reactivos adecuados conocidos en la técnica. El ARN puede comprender ARNm.

En algunos modos de realización, un procedimiento comprende la transcripción inversa de un polinucleótido de ARN diana para formar ADNc usando uno o más cebadores (cebadores de RT). En algunos modos de realización, un cebador de RT comprende un cebador oligo-dT o un cebador específico de secuencia. En algunos modos de realización, una pluralidad de cebadores de RT comprende uno o más cebadores oligo-dT o uno o más cebadores específicos de secuencia. En algunos modos de realización, una reacción de transcripción inversa es la primera etapa para generar una biblioteca de polinucleótidos a partir de una muestra que contiene un polinucleótido diana. En algunos modos de realización, un polinucleótido diana no se somete a una RT-PCR. En algunos modos de realización, un polinucleótido diana no se somete a una amplificación exponencial. En algunos modos de realización, la amplificación exponencial no se realiza en la siguiente etapa después de la transcripción inversa. En algunos modos de realización, la amplificación exponencial no se realiza en las siguientes 2 etapas después de la transcripción inversa. En algunos modos de realización, la amplificación exponencial no se realiza en las siguientes 3 etapas después de la transcripción inversa. En algunos modos de realización, el ADNc del polinucleótido diana producido a partir de la etapa de transcripción inversa no se amplifica más durante esta etapa. En algunos modos de realización, el procedimiento comprende solo un ciclo de transcripción inversa. En otros modos de realización, el procedimiento comprende la transcripción inversa repetida de la molécula de ARN diana para producir múltiples moléculas de ADNc, tales como una primera secuencia del complemento que puede contener una UID.

Un cebador de RT puede comprender además una región que no es complementaria a una región del ARN. En algunos modos de realización, los cebadores de RT pueden comprender además una UID. Por ejemplo, cada cebador de RT de una pluralidad de cebadores de RT puede comprender una UID diferente. Esto puede permitir un código de barras único para cada uno de los ADNc copiados de las moléculas de ARN que se transcriben inversamente. En algunos modos de realización, la región de un cebador de RT que no es complementaria a una región del ARN diana puede comprender una UID. En algunos modos de realización, la región de cada cebador de RT de una pluralidad de cebadores de RT que no es complementaria a una región del ARN diana puede comprender una UID. En algunos modos de realización, los cebadores de RT pueden comprender además una secuencia conocida, tal como un sitio de unión de cebador universal o una secuencia complementaria a un sitio de cebado universal. En algunos modos de realización, los cebadores de RT pueden comprender además un extremo 5' fosforilado. En algunos modos de realización, los cebadores de RT pueden comprender además una secuencia conocida, tal como un sitio de unión de cebador universal o una secuencia complementaria a un sitio de cebado universal, en el extremo 5'. En algunos modos de realización, la región que no es complementaria a una región del ARN es 5' a una región del cebador que es complementaria al ARN. En algunos modos de realización, la región que no es complementaria a una región del ARN es una región saliente 5'. En algunos modos de realización, la región que no es complementaria a una región del ARN diana comprende un sitio de cebado para amplificación y/o una reacción de secuenciación.

En algunos modos de realización, un cebador de RT puede comprender una secuencia de ligadura universal. En algunos modos de realización, la secuencia de ligadura universal es 5' de la UID. En algunos modos de realización, la secuencia de ligadura universal es 5' a la región específica de diana. En algunos modos de realización, la secuencia de ligadura

universal es 5' de la UID y 5' de la región específica de diana. En algunos modos de realización, la secuencia de ligadura universal está en el extremo 5' del cebador de RT. En algunos modos de realización, una pluralidad de cebadores de RT puede comprender un primer cebador de RT con una primera secuencia de ligadura universal y uno o más segundos cebadores de RT que comprenden al menos una segunda secuencia de cebador universal.

5

Extensión de cebador de polinucleótidos diana de ADN monocatenario o bicatenario para formar polinucleótidos complementarios marcados con UID

10

15

20

25

30

Usando cebadores descritos en el presente documento, se pueden hibridar polinucleótidos de ADN con un cebador y la extensión de cebador (gPE o PE) se puede realizar usando reactivos adecuados conocidos en la técnica. En algunos modos de realización, la extensión de cebador comprende una única extensión de un cebador. En algunos modos de realización, la extensión de cebador no comprende múltiples extensiones de un cebador. En algunos modos de realización, la extensión de cebador no comprende una única extensión de un cebador. En algunos modos de realización, la extensión de cebador comprende múltiples extensiones de un cebador. En algunos modos de realización, un procedimiento comprende realizar una extensión de cebador en un polinucleótido de ADN diana para formar una secuencia del complemento de polinucleótido diana, tal como una primera secuencia del complemento, usando uno o más cebadores (cebadores de PE). En algunos modos de realización, un cebador de PE comprende un cebador específico de secuencia. En algunos modos de realización, una pluralidad de cebadores de PE comprende uno o más cebadores específicos de secuencia. En algunos modos de realización, una reacción de extensión de cebador es la primera etapa para generar una biblioteca de polinucleótidos a partir de una muestra que contiene un polinucleótido diana. En algunos modos de realización, un polinucleótido diana no se somete a una PCR. En algunos modos de realización, un polinucleótido diana no se somete a una amplificación exponencial. En algunos modos de realización, la amplificación exponencial no se realiza en la siguiente etapa después de la extensión de cebador. En algunos modos de realización, la amplificación exponencial no se realiza en las siguientes 2 etapas después de la extensión de cebador. En algunos modos de realización, la amplificación exponencial no se realiza en las siguientes 3 etapas después de la extensión de cebador. En algunos modos de realización, el polinucleótido complementario del polinucleótido diana producido a partir de la etapa de extensión de cebador no se amplifica más durante esta etapa. En algunos modos de realización, el procedimiento comprende solo un ciclo de extensión de cebador. En otros modos de realización, el procedimiento comprende la extensión repetida o la amplificación lineal de un cebador hibridado con una molécula de ADN diana para producir múltiples copias de las moléculas de ADN, tal como la secuencia del complemento del polinucleótido diana que puede contener una UID.

35

40

Los uno o más cebadores de PE pueden comprender una región complementaria a una región o secuencia de un ADN diana, tal como una región específica de diana que se hibrida con un polinucleótido diana, tal como un biomarcador. Los uno o más cebadores de PE pueden comprender una región complementaria o sustancialmente complementaria a una región del ADN diana. En algunos modos de realización, los uno o más cebadores de PE pueden comprender un primer cebador de PE con una región complementaria a una secuencia de un primer polinucleótido diana, y un segundo cebador de PE con una región complementaria a la secuencia de un segundo polinucleótido diana. Por ejemplo, el primer polinucleótido diana puede ser una primera molécula de ADN y el segundo polinucleótido diana puede ser una segunda molécula de ADN. En algunos modos de realización, los uno o más cebadores de PE pueden comprender un primer cebador de PE con una región complementaria a una secuencia de un primer ADN, y uno o más segundos cebadores de PE, cada uno con una región complementaria a una secuencia de uno o más segundos ADN. En algunos modos de realización, las primera y segunda secuencias diana son iguales. En algunos modos de realización, las primera y segunda secuencias diana son diferentes.

45

50

55

60

Un cebador de PE puede comprender además una región que no es complementaria a una región del ADN. Los cebadores de PE pueden comprender además una UID. Por ejemplo, cada cebador de PE de una pluralidad de cebadores de PE puede comprender una UID diferente. Esto puede permitir un código de barras único para cada uno de los ADN complementarios copiados de las moléculas de ADN que se están sometiendo a una reacción de extensión de cebador. En algunos modos de realización, la región de un cebador de PE que no es complementaria a una región del ADN diana puede comprender una UID. En algunos modos de realización, la región de cada cebador de PE de una pluralidad de cebadores de PE que no es complementaria a una región del ADN diana puede comprender una UID. En algunos modos de realización, los cebadores de PE pueden comprender además una secuencia conocida, tal como un sitio de unión de cebador universal o una secuencia complementaria a un sitio de cebado universal. En algunos modos de realización, los cebadores de PE pueden comprender además un extremo 5' fosforilado. En algunos modos de realización, los cebadores de PE pueden comprender además una secuencia conocida, tal como un sitio de unión de cebador universal o una secuencia complementaria a un sitio de cebado universal, en el extremo 5'. En algunos modos de realización, la región que no es complementaria a una región del ADN es 5' a una región del cebador que es complementaria al ADN. En algunos modos de realización, la región que no es complementaria a una región del ADN es una región saliente 5'. En algunos modos de realización, la región que no es complementaria a una región del ADN diana comprende un sitio de cebado para amplificación y/o una reacción de secuenciación.

En algunos modos de realización, se puede usar una biblioteca de cebadores de PE durante la etapa de extensión de cebador.

65

En algunos modos de realización, un cebador de PE puede comprender una secuencia de ligadura universal. En algunos modos de realización, la secuencia de ligadura universal es 5' de la UID. En algunos modos de realización, la secuencia

de ligadura universal es 5' a la región específica de diana. En algunos modos de realización, la secuencia de ligadura universal es 5' de la UID y 5' de la región específica de diana. En algunos modos de realización, la secuencia de ligadura universal está en el extremo 5' del cebador de PE. En algunos modos de realización, una pluralidad de cebadores de PE puede comprender un primer cebador de PE con una primera secuencia de ligadura universal y uno o más segundos cebadores de PE que comprenden al menos una segunda secuencia de cebador universal.

En algunos modos de realización, se usa una temperatura de hibridación de 55 °C para adaptarse a temperaturas de fusión de cebador más bajas. En algunos modos de realización, se usa una etapa de retención a 68 °C para la etapa inicial de PE. En algunos modos de realización, la concentración global de los cebadores se fija a una concentración. En algunos modos de realización, se usa cloruro de magnesio, sulfato de amonio, D-(+)-trehalosa, betaina o una combinación de los mismos durante la etapa de extensión de cebador.

Formateo parcial de polinucleótidos marcados con UID complementarios a las dianas

Después de generar secuencias del complemento de polinucleótidos diana, por ejemplo, primeras secuencias del complemento, se puede añadir una secuencia adaptadora de polinucleótido a las primeras secuencias del complemento. Una secuencia del complemento de polinucleótidos diana, tal como la primera secuencia del complemento que puede contener una UID, a la que se ha añadido una secuencia adaptadora, puede ser una secuencia del complemento modificada (MCS). En algunos modos de realización, se puede añadir una secuencia adaptadora de polinucleótidos a las secuencias del complemento de polinucleótidos diana, tales como las primeras secuencias del complemento que pueden contener una UID, en la siguiente etapa después de generar secuencias del complemento de polinucleótidos diana. En algunos modos de realización, se puede añadir una secuencia adaptadora de polinucleótidos a las secuencias del complemento de polinucleótidos diana, tales como las primeras secuencias del complemento que pueden contener una UID, en la segunda etapa después de generar secuencias del complemento de polinucleótidos diana que contienen UID. En algunos modos de realización, se puede añadir una secuencia adaptadora de polinucleótidos a las secuencias del complemento de polinucleótidos diana, tales como las primeras secuencias del complemento que pueden contener una UID, en la tercera etapa después de generar secuencias del complemento de polinucleótidos diana que contienen UID. En algunos modos de realización, una secuencia adaptadora de polinucleótidos no contiene una UID.

En algunos modos de realización, se puede añadir una secuencia adaptadora de polinucleótidos a las secuencias del complemento de polinucleótidos diana, tales como las primeras secuencias del complemento que pueden contener una UID mediante ligadura (patentes de los Estados Unidos n.º 4.883.750, 5.476.930, 5.593.826, 5.426.180, 5.871.921; y publicación de patente de los Estados Unidos n.º 2004/0110213). Las técnicas de ligadura pueden comprender la ligadura de extremo romo y la ligadura de extremo adhesivo. Las reacciones de ligadura pueden incluir ADN ligasas tales como ADN ligasa I, ADN ligasa III, ADN ligasa IV y ADN ligasa T4. Las reacciones de ligadura pueden incluir ARN ligasas tales como ARN ligasa I T4 y ARN ligasa II T4. Los procedimientos incluyen el uso de la ADN ligasa T4 que cataliza la formación de un enlace fosfodiéster entre los extremos 5' del fosfato y 3' del hidroxilo yuxtapuestos en ADN o ARN bicatenario con extremos romos y adhesivos; la ADN ligasa Taq que cataliza la formación de un enlace fosfodiéster entre los extremos 5' del fosfato y 3' del hidroxilo yuxtapuestos de dos oligonucleótidos adyacentes que se hibridan con un ADN diana complementario; la ADN ligasa de *E. coli* que cataliza la formación de un enlace fosfodiéster entre los extremos 5'-fosfato y 3'-hidroxilo yuxtapuestos en los extremos cohesivos que contienen ADN bicatenario; y la ARN ligasa T4 que cataliza la ligadura de un donante de ácido nucleico terminado en fosforilo 5' a un aceptador de ácido nucleico terminado en hidroxilo 3' a través de la formación de un enlace fosfodiéster de 3' a 5', los sustratos incluyen también ARN y ADN monocatenario, así como pirofosfatos de dinucleósidos.

En algunos modos de realización, una secuencia adaptadora de polinucleótidos no se añade a secuencias del complemento de polinucleótidos diana, tales como primeras secuencias del complemento que pueden contener una UID, mediante ligadura. En algunos modos de realización, se puede añadir una secuencia adaptadora de polinucleótidos a las secuencias del complemento de polinucleótidos diana, tales como primeras secuencias del complemento que pueden contener una UID, mediante una reacción de amplificación. En algunos modos de realización, se puede añadir una secuencia adaptadora de polinucleótidos a las secuencias del complemento de polinucleótidos diana, tales como primeras secuencias del complemento que pueden contener una UID, mediante una reacción de amplificación con uno o más cebadores que contienen la secuencia adaptadora. En algunos modos de realización, una secuencia adaptadora de polinucleótidos no se añade a secuencias del complemento de polinucleótidos diana, tales como primeras secuencias del complemento que pueden contener una UID, mediante una reacción de amplificación. En algunos modos de realización, no se añade una secuencia adaptadora de polinucleótidos a las secuencias del complemento de polinucleótidos diana, tales como primeras secuencias del complemento que pueden contener una UID, mediante una reacción de amplificación con uno o más cebadores que contienen la secuencia adaptadora. En algunos modos de realización, se puede añadir una secuencia adaptadora de polinucleótidos a las secuencias del complemento de polinucleótidos diana, tales como segundas secuencias del complemento que pueden contener una UID, durante una etapa de enriquecimiento por PCR como se describe a continuación.

En algunos modos de realización, se puede añadir una secuencia adaptadora de polinucleótidos a las secuencias del complemento de polinucleótido diana, tales como primeras secuencias del complemento que pueden contener una UID, mediante ligadura en la siguiente etapa después de generar secuencias del complemento de polinucleótido diana. En algunos modos de realización, un adaptador puede ser un polinucleótido monocatenario. En algunos modos de realización,

un adaptador puede ser un polinucleótido bicatenario. En algunos modos de realización, un adaptador puede ser un polinucleótido puente que contiene una región bicatenaria y una región monocatenaria, tal como una región saliente. En algunos modos de realización, un adaptador puede ser un polinucleótido puente que contiene una región bicatenaria y una región monocatenaria, en el que la cadena que contiene la región monocatenaria no está ligada a las secuencias del complemento de polinucleótidos diana, tales como las primeras secuencias del complemento que pueden contener una UID. En algunos modos de realización, un adaptador puede ser un polinucleótido puente que contiene una región bicatenaria y una región monocatenaria, en el que la cadena que no contiene la región monocatenaria está ligada a las secuencias del complemento de polinucleótidos diana, tales como primeras secuencias del complemento que pueden contener una UID. En algunos modos de realización, un adaptador puede ser un polinucleótido puente que contiene una región bicatenaria y una región monocatenaria, en el que la cadena que no contiene una región complementaria a las secuencias del complemento de polinucleótidos diana, tales como primeras secuencias del complemento que pueden contener una UID, está ligada a las secuencias del complemento de polinucleótidos diana. En algunos modos de realización, un adaptador puede ser un polinucleótido puente que contiene una región bicatenaria y una región monocatenaria, en el que la cadena que contiene una región complementaria a las secuencias del complemento de polinucleótidos diana, tales como primeras secuencias del complemento que pueden contener una UID, no está ligada a las secuencias del complemento de polinucleótidos diana. En algunos modos de realización, un adaptador puede ser un polinucleótido puente que contiene una región bicatenaria y una región monocatenaria, en el que la cadena que contiene una región complementaria a las secuencias del complemento de polinucleótidos diana, tales como primeras secuencias del complemento que pueden contener una UID, se hibrida con las secuencias del complemento de polinucleótidos diana. En algunos modos de realización, un adaptador puede ser un polinucleótido puente que contiene una región bicatenaria y una región monocatenaria, en el que la cadena que no contiene una región complementaria a las secuencias del complemento de polinucleótidos diana, tales como primeras secuencias del complemento que pueden contener una UID, no se hibrida con las secuencias del complemento de polinucleótidos diana.

En algunos modos de realización, la región saliente 5' puede ser complementaria a una o más secuencias del complemento de polinucleótidos diana, tales como las que contienen UID. En algunos modos de realización, la región saliente 5' puede ser complementaria a una región 5' de una o más secuencias del complemento de polinucleótidos, tales como las que contienen UID. En algunos modos de realización, la región saliente 5' puede comprender una secuencia complementaria a una secuencia de ligadura universal, tal como una secuencia de ligadura universal de un cebador de RT o un cebador de PE. En algunos modos de realización, la región saliente 5' puede ser complementaria a una región 5' de una o más secuencias del complemento de polinucleótidos diana, tales como las que contienen UID, en la que la región 5' es 5' a la UID. En algunos modos de realización, un adaptador puede ser un polinucleótido puente que contiene una región bicatenaria y una región monocatenaria, tal como una región o extremo saliente 5'. En algunos modos de realización, un adaptador puede ser un polinucleótido puente que contiene una región bicatenaria y una región monocatenaria, tal como una región o extremo saliente 3'. En algunos modos de realización, un adaptador puede ser un polinucleótido puente que contiene una región bicatenaria y dos regiones monocatenarias, tales como una región o extremo saliente 3' y una región o extremo saliente 5'. En algunos modos de realización, la región saliente 5' puede ser complementaria a una región 5' de una o más secuencias del complemento de polinucleótidos diana, tales como las que contienen UID, en la que el adaptador puede estar ligado a las una o más secuencias del complemento de polinucleótidos diana, tales como las que contienen UID cuando se hibridan. En algunos modos de realización, la región saliente 5' puede ser complementaria a una región 5' de una o más secuencias del complemento de polinucleótidos diana, tales como las que contienen UID, en la que el adaptador puede estar muy cerca o al lado del extremo 5' de una o más secuencias del complemento de polinucleótidos diana que contienen UID cuando se hibridan. En algunos modos de realización, la región saliente 5' puede ser complementaria a una región 5' de una o más secuencias del complemento de polinucleótidos diana, tales como las que contienen UID, en la que el adaptador puede estar muy cerca o al lado del extremo fosfato 5' de una o más secuencias del complemento de polinucleótidos diana, tales como las que contienen UID, cuando se hibridan. En algunos modos de realización, la región saliente 5' puede tener la misma longitud, o sustancialmente la misma longitud, que la secuencia a la que es complementaria en las una o más secuencias del complemento de polinucleótidos diana, tales como las que contienen UID.

En algunos modos de realización, se puede añadir una secuencia adaptadora de polinucleótidos que comprende un sitio de unión de cebador o un complemento de un sitio de unión de cebador, a las secuencias del complemento de polinucleótidos diana, tales como las primeras secuencias del complemento que pueden contener una UID. En algunos modos de realización, una secuencia del complemento de polinucleótidos diana, tal como primeras secuencias del complemento que pueden contener una UID, que contiene un primer sitio de unión de cebador de un conjunto de unión de cebador, tal como para amplificación o secuenciación exponencial, puede ser una secuencia del complemento de polinucleótidos diana parcialmente formateada, tal como una secuencia del complemento modificada que puede contener una UID. En algunos modos de realización, una secuencia del complemento de polinucleótidos diana, tal como una primera secuencia del complemento, que contiene un primer sitio de unión de cebador de un primer conjunto de cebador y un primer sitio de unión de cebador de un segundo conjunto de unión de cebador, tal como para amplificación o secuenciación exponencial, puede ser una secuencia del complemento de polinucleótidos diana completamente formateada, tal como una secuencia del complemento modificada que puede contener una UID. En algunos modos de realización, el sitio de unión de cebador o el complemento del mismo se añade a cada una de una pluralidad de secuencias del complemento de polinucleótidos diana, tales como primeras secuencias del complemento que pueden contener una UID. En algunos modos de realización, el sitio de unión de cebador o el complemento del mismo que se añade a cada una de una pluralidad de secuencias del complemento de polinucleótidos diana, tales como primeras secuencias del

complemento que pueden contener una UID, es la misma secuencia. En algunos modos de realización, el sitio de unión de cebador o el complemento del mismo que se añade a cada una de una pluralidad de secuencias del complemento de polinucleótidos diana, tales como primeras secuencias del complemento que pueden contener una UID, es una secuencia diferente. En algunos modos de realización, el sitio de unión de cebador o el complemento del mismo que se añade a cada una de una pluralidad de secuencias del complemento de polinucleótidos diana en un primer amplicón o conjunto de amplicón es la misma secuencia que un sitio de unión de cebador o complemento del mismo que se añade a cada una de una pluralidad de secuencias del complemento de polinucleótidos diana, en un segundo amplicón o conjunto de amplicones. Como se usa en el presente documento, un amplicón comprende un producto polinucleotídico de una reacción de amplificación. Un conjunto de amplicones comprende una población clonal de polinucleótidos producidos a partir de una reacción de amplificación. En algunos modos de realización, los conjuntos de amplicones se forman mediante la amplificación de una única secuencia de partida. En algunos modos de realización, un conjunto de amplicones comprende una población de polinucleótidos derivados de un único polinucleótido en una reacción de amplificación. En algunos modos de realización, un conjunto de amplicones comprende una población de polinucleótidos derivados de un único polinucleótido o amplicones de ese polinucleótido en una reacción de amplificación. Los amplicones pueden ser producidos por una variedad de reacciones de amplificación. Los amplicones pueden comprender copias de uno o más ácidos nucleicos. En algunos modos de realización, los amplicones o conjuntos de amplicones se producen mediante PCR. En algunos modos de realización, los amplicones o conjuntos de amplicones no se producen mediante PCR.

En algunos modos de realización, el sitio de unión de cebador o el complemento del mismo que se añade a cada una de una pluralidad de secuencias del complemento de polinucleótidos diana en un primer amplicón o conjunto de amplicones es una secuencia diferente a un sitio de unión de cebador o complemento del mismo que se añade a cada una de una pluralidad de secuencias del complemento de polinucleótidos diana, en un segundo amplicón o conjunto de amplicones. En algunos modos de realización, el sitio de unión de cebador o complemento del mismo que se añade a cada uno de una pluralidad de polinucleótidos que contienen UID a partir de una primera muestra es una secuencia diferente a un sitio de unión de cebador o complemento del mismo que se añade a cada una de una pluralidad de secuencias del complemento de polinucleótidos diana, a partir de una segunda muestra. En algunos modos de realización, el sitio de unión de cebador o complemento del mismo que se añade a cada una de una pluralidad de secuencias del complemento de polinucleótidos diana, a partir de una primera muestra es la misma secuencia que un sitio de unión de cebador o complemento del mismo que se añade a cada una de una pluralidad de secuencias del complemento de polinucleótidos diana, a partir de una segunda muestra. En algunos modos de realización, el sitio de unión de cebador o complemento del mismo comprende una secuencia conocida. En algunos modos de realización, el sitio de unión de cebador o complemento del mismo comprende un sitio de unión de cebador para amplificación. En algunos modos de realización, el sitio de unión de cebador o complemento del mismo comprende una secuencia de cebado universal. En algunos modos de realización, el sitio de unión de cebador o complemento del mismo comprende una primera unión de cebador para un primer cebador de un conjunto de cebadores. En algunos modos de realización, el sitio de unión de cebador o complemento del mismo comprende una primera unión de cebador para realizar una reacción de amplificación exponencial, tal como PCR, por ejemplo, para su uso en una etapa de enriquecimiento por PCR como se describe a continuación. En algunos modos de realización, el sitio de unión de cebador o complemento del mismo comprende una primera unión de cebador para realizar una reacción de amplificación no exponencial. En algunos modos de realización, el sitio de unión de cebador o complemento del mismo comprende un sitio de unión de cebador para secuenciación. En algunos modos de realización, el sitio de unión de cebador o complemento del mismo comprende un sitio de unión de cebador para análisis.

En algunos modos de realización, una secuencia adaptadora de polinucleótidos comprende además una secuencia de código de barras de muestra (SBC). En los procedimientos descritos, el código de barras de muestra en una secuencia adaptadora genérica puede eliminar la necesidad de múltiples conjuntos de sondas para cada UID empleada. Como se usa en el presente documento, un código de barras de muestra (SBC) en un polinucleótido comprende una secuencia que se puede usar para identificar una fuente de la que se deriva un polinucleótido. Por ejemplo, una muestra de ácido nucleico puede ser un grupo de polinucleótidos derivados de una pluralidad de muestras diferentes (por ejemplo, polinucleótidos derivados de diferentes individuos, diferentes tejidos o células, o polinucleótidos aislados en diferentes momentos), donde los polinucleótidos de cada muestra diferente de la pluralidad están marcados con un SBC único. Por tanto, un SBC proporciona una correlación entre un polinucleótido y su fuente. (Patentes de Estados Unidos n.º 7.537.897, 7.544.473 y 7.393.665). En algunos modos de realización, se puede usar el mismo SBC para marcar una muestra diferente que se está procesando en un experimento diferente. En algunos modos de realización, se puede usar un SBC diferente para marcar cada muestra diferente o un subconjunto de muestras que se están procesando en un experimento. Por ejemplo, las muestras de uno o más sujetos con una enfermedad o afección pueden tener un primer SBC y las muestras de uno o más sujetos sin una enfermedad o afección pueden tener un segundo SBC diferente. Por ejemplo, diferentes muestras derivadas de la misma muestra se pueden marcar con diferentes SBC.

En algunos modos de realización, una secuencia adaptadora de polinucleótidos comprende además un SBC o complemento del mismo que se encuentra entre una secuencia del sitio de unión de cebador o complemento del mismo del adaptador, y una región del adaptador, tal como una región saliente 5' que es complementaria a una secuencia de las una o más secuencias del complemento de polinucleótidos diana. En algunos modos de realización, una secuencia adaptadora de polinucleótidos comprende además un SBC, en la que el SBC está dentro de una región duplicada del adaptador. En algunos modos de realización, una secuencia adaptadora de polinucleótidos comprende además un SBC, en la que el SBC no está dentro de una región duplicada del adaptador. En algunos modos de realización, una secuencia

adaptadora de polinucleótidos comprende además un SBC, en la que el SBC está dentro de una región monocatenaria del adaptador. En algunos modos de realización, una secuencia adaptadora de polinucleótidos comprende además un SBC, en la que el SBC está en una cadena diferente a la cadena que contiene una región de complementariedad a las una o más secuencias del complemento de polinucleótidos diana, tal como una región saliente 5'. En algunos modos de realización, una secuencia adaptadora de polinucleótidos comprende además un SBC, en la que el SBC está en la misma cadena que la cadena que contiene una región de complementariedad, tal como una región saliente 5', a una o más secuencias del complemento de polinucleótidos diana, tales como primeras secuencias del complemento. En algunos modos de realización, una secuencia adaptadora de polinucleótidos comprende además un SBC, en la que el SBC está en la cadena que no contiene una región de complementariedad, tal como una región saliente 5', a una o más secuencias del complemento de polinucleótidos diana, tales como primeras secuencias del complemento. En algunos modos de realización, el sitio de unión de cebador o complemento del mismo que se añade a una pluralidad de secuencias del complemento de polinucleótidos diana, tales como primeras secuencias del complemento, es 5' a una secuencia de SBC del adaptador. En algunos modos de realización, el sitio de unión de cebador o complemento del mismo que se añade a una pluralidad de secuencias del complemento de polinucleótidos diana, tales como primeras secuencias del complemento, es 3' a una secuencia de SBC del adaptador.

Un procedimiento puede comprender además combinar una primera y una segunda muestra antes de llevar a cabo cualquiera de las una o más reacciones. En algunos modos de realización, un procedimiento comprende además combinar polinucleótidos generados a partir de una primera y una segunda muestra. En algunos modos de realización, un procedimiento comprende además combinar polinucleótidos generados a partir de una primera y una segunda muestra después de realizar una reacción de extensión de cebador. En algunos modos de realización, un procedimiento comprende además combinar polinucleótidos generados a partir de una primera y una segunda muestra después de unir un adaptador a polinucleótidos en la primera o segunda muestra. En algunos modos de realización, un procedimiento comprende además combinar polinucleótidos generados a partir de una primera y una segunda muestra después de unir un adaptador que comprende un SBC a polinucleótidos en la primera o segunda muestra. En algunos modos de realización, un procedimiento comprende además combinar secuencias del complemento de polinucleótidos diana, generadas a partir de una primera y una segunda muestra. En algunos modos de realización, un procedimiento comprende además combinar polinucleótidos generados a partir de una primera y una segunda muestra que comprende uno o más sitios de unión de cebador, tales como uno o más sitios de unión de cebador universales. En algunos modos de realización, un procedimiento comprende además combinar polinucleótidos generados a partir de una primera y una segunda muestra después de realizar una amplificación exponencial de los polinucleótidos en la primera y/o segunda muestra. En algunos modos de realización, el origen de la muestra de los polinucleótidos que se originan a partir de una primera muestra y una segunda muestra se puede determinar usando un SBC. En algunos modos de realización, el origen de la muestra de los polinucleótidos que se originan a partir de una primera muestra y una segunda muestra se puede determinar usando una UID. El origen de la muestra de los polinucleótidos que se originan a partir de una primera muestra y una segunda muestra se puede determinar usando una secuencia del sitio de unión de cebador. El origen de la muestra de los polinucleótidos que se originan a partir de una primera muestra y una segunda muestra se puede determinar usando una secuencia específica de diana.

40 Limpieza opcional

En algunos modos de realización, un procedimiento comprende además purificar opcionalmente uno o más de los polinucleótidos marcados con adaptador, tales como secuencias del complemento modificadas que pueden contener una UID. En algunos modos de realización, el adaptador añadido a una pluralidad de secuencias del complemento de polinucleótidos diana, tales como primeras secuencias del complemento, comprende una marca de afinidad. Una marca de afinidad puede unirse a un ligando y las moléculas que no se unen al ligando (por ejemplo, las moléculas sin la marca de afinidad) se pueden eliminar por lavado, o las moléculas marcadas por afinidad se pueden aislar de las moléculas sin una marca de afinidad. En algunos modos de realización, una marca de afinidad puede ser una primera molécula que se une específicamente a una segunda molécula. En algunos modos de realización, la marca de afinidad puede ser una secuencia de nucleótidos conocida. En algunos modos de realización, la marca de afinidad puede ser un resto químico. En algunos modos de realización, la marca de afinidad puede ser biotina o estreptavidina. En algunos modos de realización, la marca de afinidad puede ser un péptido o una proteína, tal como un anticuerpo. Por tanto, el adaptador puede comprender un complejo de proteína-ácido nucleico. Se puede usar cualquier marca de afinidad conocida en la técnica. En algunos modos de realización, la marca de afinidad se puede usar para purificar las secuencias del complemento de polinucleótidos diana modificadas por adaptador (por ejemplo, ligadas o amplificadas), tales como secuencias del complemento modificadas que pueden contener una UID, a partir de uno o más de otros polinucleótidos. Se puede usar un soporte o superficie que contiene una o más moléculas polinucleotídicas, químicas o proteicas inmovilizadas que se unen a una marca de afinidad. Por ejemplo, la marca de afinidad se puede usar para purificar las secuencias del complemento de polinucleótidos diana del adaptador, tales como secuencias del complemento modificadas que pueden contener una UID, a partir de uno o más de otros polinucleótidos al unir una biotina de las secuencias del complemento de polinucleótidos diana modificadas por adaptador a una superficie o sustrato que comprende un resto de estreptavidina. Como se usa en el presente documento, la inmovilización comprende la unión directa o indirecta a un soporte sólido a través de uno o más enlaces covalentes o no covalentes. En algunos modos de realización, la inmovilización comprende la unión directa o indirecta a un soporte sólido por hibridación. En algunos modos de realización, la marca de afinidad se puede usar para purificar las secuencias del complemento de polinucleótidos diana del adaptador, tales como secuencias del complemento modificadas que pueden contener una UID, a partir de una o más

secuencias de polinucleótidos que no son de interés, tales como un polinucleótido no diana. En algunos modos de realización, la marca de afinidad se puede usar para purificar las secuencias del complemento de polinucleótidos diana del adaptador, tales como secuencias del complemento modificadas que pueden contener una UID, a partir de uno o más cebadores usados en una reacción o etapa de procedimiento previa. En algunos modos de realización, la marca de afinidad se puede usar para purificar las secuencias del complemento de polinucleótidos diana del adaptador, tales como secuencias del complemento modificadas que pueden contener una UID, a partir de uno o más cebadores usados en una reacción o etapa de procedimiento previa, o a partir de una o más secuencias de polinucleótidos que no son de interés, tales como un polinucleótido no diana. En algunos modos de realización, una marca de afinidad no se usa en los procedimientos descritos. Por ejemplo, en algunos modos de realización, un adaptador no comprende una marca de afinidad. Por ejemplo, en algunos modos de realización, una marca de afinidad no se usa en los procedimientos descritos cuando la molécula diana es ARN.

Extensión de cebador lineal/amplificación lineal

Un procedimiento puede comprender además realizar una segunda ronda única de extensión de cebador o extensión de cebador lineal (también llamada amplificación lineal). En algunos modos de realización, uno o más cebadores usados para la extensión/amplificación lineal se aíslan en una o más reacciones separadas a partir de los uno o más cebadores de RT o PE usados en la etapa de transcripción inversa o extensión de cebador. Al separar los pares de cebadores de esta manera, se pueden reducir las interacciones de cebadores no deseadas. Como se usa en el presente documento, la amplificación lineal o la extensión de cebador lineal se refiere a un proceso de extensión no exponencial del número de copias de producto. En algunos modos de realización, solo la cadena molde se replica durante cada ciclo de una amplificación lineal. En algunos modos de realización, la extensión de cebador por sí misma no se copia durante la amplificación lineal. Cuando se usa un único cebador no emparejado en lugar de dos cebadores, el resultado es un crecimiento lineal en el número de copias de producto de extensión en lugar de un crecimiento exponencial de ambas cadenas como en la PCR.

Usando cebadores descritos en el presente documento, los polinucleótidos de ADN producidos a partir de uno o más de los procedimientos o etapas de procedimiento anteriores se pueden hibridar con un cebador (cebador de LPE) y se puede realizar una extensión de cebador lineal usando reactivos adecuados conocidos en la técnica. Por ejemplo, una o más secuencias del complemento de polinucleótidos diana, tales como secuencias del complemento modificadas que pueden contener una UID, se pueden hibridar con un cebador de LPE y se puede realizar una extensión de cebador lineal. Por ejemplo, una o más secuencias del complemento de polinucleótidos diana, tales como primeras secuencias del complemento que pueden contener una UID, a las que se ha añadido un adaptador, tal como por ligadura o amplificación, se pueden hibridar con un cebador de LPE y se puede realizar la extensión de cebador lineal. En algunos modos de realización, una LPE comprende una UID. En algunos modos de realización, una LPE comprende una UID, y un cebador de RT o PE no contiene una UID. En algunos modos de realización, una LPE comprende una UID, y un cebador de RT o PE comprende una UID. En algunos modos de realización, una LPE y un cebador de RT comprenden una UID, y un cebador de PE no contiene una UID. En algunos modos de realización, una LPE y un cebador de PE comprenden una UID, y un cebador de RT no contiene una UID.

En algunos modos de realización, la extensión de cebador lineal comprende múltiples extensiones de un cebador de LPE. En algunos modos de realización, la extensión de cebador lineal comprende múltiples extensiones de cada cebador de LPE en una pluralidad de cebadores de LPE. En algunos modos de realización, la extensión de cebador lineal comprende múltiples extensiones de cada cebador de LPE en una pluralidad de cebadores de LPE, en la que cada cebador de LPE en la pluralidad se dirige a un polinucleótido diferente. En algunos modos de realización, la extensión de cebador lineal comprende múltiples extensiones de cada cebador de LPE en una pluralidad de cebadores de LPE, en la que cada cebador de LPE en la pluralidad se dirige al mismo polinucleótido. En algunos modos de realización, una segunda ronda de extensión de cebador comprende una única extensión de un cebador de LPE. En algunos modos de realización, la extensión de cebador lineal no comprende múltiples extensiones de un cebador. En algunos modos de realización, un procedimiento comprende realizar una extensión de cebador lineal en una o más secuencias del complemento de polinucleótidos diana, tales como secuencias del complemento modificadas que pueden contener una UID, que comprende un adaptador, para formar un polinucleótido complementario, tal como ADN, usando uno o más cebadores (cebadores de LPE). En algunos modos de realización, un procedimiento comprende realizar una extensión de cebador lineal en una o más secuencias del complemento de polinucleótidos diana, en el que las una o más secuencias del complemento de polinucleótidos diana no comprenden un adaptador, tales como primeras secuencias del complemento que pueden contener una UID. En algunos modos de realización, un cebador de LPE comprende un cebador específico de secuencia. En algunos modos de realización, una pluralidad de cebadores de LPE comprende uno o más cebadores específicos de secuencia. En algunos modos de realización, una reacción de extensión de cebador lineal es la primera, segunda, tercera o cuarta etapa para generar una biblioteca de polinucleótidos a partir de una muestra que contiene un polinucleótido diana. En algunos modos de realización, una reacción de extensión de cebador lineal es la tercera etapa para generar una biblioteca de polinucleótidos a partir de una muestra que contiene un polinucleótido diana. En algunos modos de realización, una reacción de extensión de cebador lineal es la cuarta etapa para generar una biblioteca de polinucleótidos a partir de una muestra que contiene un polinucleótido diana. En algunos modos de realización, se realiza una reacción de extensión de cebador lineal después de una reacción de RT o PE. En algunos modos de realización, se realiza una reacción de extensión de cebador lineal después de una reacción que añade un adaptador a una secuencia del complemento de polinucleótidos diana, tal como una primera secuencia del complemento que puede contener una

UID. En algunos modos de realización, se realiza una reacción de extensión de cebador lineal después de una reacción de RT o PE y después de una reacción que añade un adaptador a una secuencia del complemento de polinucleótidos diana, tal como una primera secuencia del complemento que puede contener una UID. En algunos modos de realización, se realiza una reacción de extensión de cebador lineal antes de realizar una reacción de amplificación exponencial, tal como PCR. En algunos modos de realización, la amplificación exponencial se realiza en la siguiente etapa después de la extensión de cebador lineal. En algunos modos de realización, la amplificación exponencial no se realiza en la siguiente etapa después de la extensión de cebador lineal. En algunos modos de realización, la amplificación exponencial se realiza en las siguientes 2 etapas después de la extensión de cebador lineal. En algunos modos de realización, la amplificación exponencial no se realiza en las siguientes 3 etapas después de la extensión de cebador lineal. En algunos modos de realización, un polinucleótido complementario de la secuencia del complemento de polinucleótidos diana, tal como una segunda secuencia del complemento que puede contener una UID, producida a partir de la etapa de extensión de cebador lineal no se amplifica más después de esta etapa. En algunos modos de realización, el procedimiento comprende solo un ciclo de extensión de cebador lineal. En otros modos de realización, el procedimiento comprende extender repetidamente un cebador hibridado a una secuencia del complemento de polinucleótidos diana para producir múltiples copias de las secuencias del complemento de polinucleótidos diana, tales como segundas secuencias del complemento que pueden contener una UID. Los procedimientos pueden comprender llevar a cabo al menos aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 o 20 reacciones de extensión de cebador lineal o ciclos de extensión de cebador lineal. En algunos modos de realización, se puede usar menos entrada de muestra en los procedimientos que usan amplificación/extensión lineal como se describe en el presente documento que en un procedimiento similar que emplea una etapa de amplificación no lineal. En algunos modos de realización, se pueden usar menos ciclos de PCR en los procedimientos que usan amplificación/extensión lineal como se describe en el presente documento que en un procedimiento similar que emplea una etapa de amplificación no lineal. Por ejemplo, 20 ciclos de PCR pueden ser suficientes para los procedimientos que usan amplificación/extensión lineal, mientras que se pueden requerir 24 ciclos de PCR para un procedimiento similar que emplea una etapa de amplificación no lineal.

Los uno o más cebadores de LPE pueden comprender una secuencia complementaria a una secuencia o secuencia del complemento de una secuencia del complemento de polinucleótidos diana, tal como una primera secuencia del complemento o secuencia del complemento modificada. Por ejemplo, los uno o más cebadores de LPE pueden comprender una secuencia complementaria a una secuencia o secuencia del complemento de una secuencia del complemento de polinucleótidos diana, tal como una primera secuencia del complemento o secuencia del complemento modificada o un polinucleótido diana en una muestra inicial. Por ejemplo, los uno o más cebadores de LPE pueden comprender una secuencia complementaria a una secuencia o secuencia del complemento de una secuencia del complemento de polinucleótidos diana, tal como una primera secuencia del complemento o secuencia del complemento modificada que es un producto de una reacción de amplificación, reacción de ligadura, extensión de cebador o combinaciones de las mismas.

En algunos modos de realización, los uno o más cebadores de LPE comprenden una secuencia complementaria a una secuencia del complemento de un polinucleótido diana. En algunos modos de realización, los uno o más cebadores de LPE pueden comprender una secuencia complementaria a una secuencia de una secuencia del complemento de polinucleótidos diana, tal como una primera secuencia del complemento o secuencia del complemento modificada. En algunos modos de realización, los uno o más cebadores de LPE comprenden una primera secuencia complementaria a una secuencia del complemento de un polinucleótido diana y una segunda secuencia complementaria a una secuencia de una secuencia del complemento de polinucleótidos diana, tal como una primera secuencia del complemento o secuencia del complemento modificada. En algunos modos de realización, las primera y segunda secuencias son la misma secuencia. En algunos modos de realización, las primera y segunda secuencias son secuencias diferentes. En algunos modos de realización, la secuencia complementaria a una secuencia del complemento de polinucleótidos diana, tal como una primera secuencia del complemento o secuencia del complemento modificada, de uno o más cebadores de LPE no es complementaria a una secuencia diana. En algunos modos de realización, la secuencia complementaria a un polinucleótido que contiene UID de uno o más cebadores de LPE no es complementaria a ningún polinucleótido que no contenga una UID. En algunos modos de realización, las secuencias complementarias a una secuencia del complemento de polinucleótidos diana, tal como una primera secuencia del complemento o secuencia del complemento modificada, de uno o más cebadores de LPE no son complementarias a ningún otro polinucleótido en una muestra.

En algunos modos de realización, la secuencia del complemento de polinucleótidos diana es un polinucleótido monocatenario. En algunos modos de realización, la secuencia del complemento de polinucleótidos diana es un polinucleótido bicatenario. En algunos modos de realización, la secuencia del complemento de polinucleótidos diana, tal como una primera secuencia del complemento, es un producto de extensión de una reacción de PE o RT. En algunos modos de realización, la secuencia del complemento de polinucleótidos diana comprende además una secuencia adaptadora, tal como una secuencia adaptadora ligada o una secuencia del complemento modificada. En algunos modos de realización, la secuencia del complemento de polinucleótidos diana, es un producto de extensión de una reacción de PE o RT que comprende además una secuencia adaptadora, tal como una secuencia del complemento modificada. En algunos modos de realización, la secuencia del complemento de polinucleótidos diana es un producto de extensión de una reacción de PE o RT que comprende además un primer sitio de cebador, tal como un sitio de secuenciación por PCR o de cebado universal. En algunos modos de realización, la secuencia del complemento de polinucleótidos diana, tal como una secuencia del complemento primera, segunda o modificada, se inmoviliza en un sustrato o superficie. En algunos modos de realización, la secuencia del complemento de polinucleótidos diana, tal como una secuencia del complemento

primera o modificada, comprende un SBC.

5 En algunos modos de realización, la secuencia complementaria a una secuencia del complemento de polinucleótidos diana, tal como una secuencia del complemento primera o modificada de uno o más cebadores de LPE no es una secuencia complementaria a una primera cadena de cualquier polinucleótido diana. En algunos modos de realización, la secuencia complementaria a una secuencia del complemento de polinucleótidos diana, tal como una secuencia del complemento primera o modificada, de uno o más cebadores de LPE es complementaria a una secuencia generada durante una reacción de RT o PE. En algunos modos de realización, la secuencia complementaria a una secuencia del complemento de polinucleótidos diana, tal como una secuencia del complemento primera o modificada, de uno o más cebadores de LPE es complementaria a una secuencia del complemento de un polinucleótido diana que se puede hibridar con una secuencia del polinucleótido diana que es 5' a la secuencia del polinucleótido diana complementaria a un cebador de RT o PE. En algunos modos de realización, la secuencia complementaria a una secuencia del complemento de polinucleótidos diana de uno o más cebadores de LPE es complementaria a una secuencia del complemento de un polinucleótido diana que se hibrida con una secuencia de la diana 3' a la secuencia del polinucleótido diana complementaria a un cebador de RT o PE. En algunos modos de realización, una secuencia de un polinucleótido diana que contiene una variante o una región para su análisis por cualquiera de los procedimientos descritos en el presente documento puede estar entre la secuencia del polinucleótido diana complementaria a uno o más cebadores de RT o PE y la secuencia del polinucleótido diana cuyo complemento es complementario a uno o más cebadores de LPE.

20 En algunos modos de realización, la secuencia complementaria a una secuencia del complemento de polinucleótidos diana, tal como una secuencia del complemento primera o modificada, de uno o más cebadores de LPE no es una secuencia complementaria a una secuencia de uno o más cebadores de PE o RT. En algunos modos de realización, la secuencia complementaria a una secuencia del complemento de polinucleótidos diana, una secuencia del complemento primera o modificada, de uno o más cebadores de LPE no es una secuencia complementaria a una secuencia específica de diana de uno o más cebadores de PE o RT.

30 En algunos modos de realización, los uno o más cebadores de LPE comprenden un primer cebador de LPE con una región complementaria a una secuencia de un primer polinucleótido molde y un segundo cebador de LPE con una región complementaria a una secuencia de un segundo polinucleótido molde. Por ejemplo, el primer polinucleótido molde puede ser una primera molécula de ADN y el segundo polinucleótido molde puede ser una segunda molécula de ADN. Por ejemplo, el primer polinucleótido molde puede ser una primera molécula de ADN derivada de un primer polinucleótido diana en una muestra y el segundo polinucleótido molde puede ser una segunda molécula de ADN derivada de un segundo polinucleótido diana en una muestra. En algunos modos de realización, los uno o más cebadores de LPE comprenden un primer cebador de LPE con una región complementaria a una secuencia de un primer ADN, y uno o más segundos cebadores de LPE, cada uno con una región complementaria a una secuencia de uno o más segundos ADN. En algunos modos de realización, las secuencias de los primer y segundo ADN son las mismas. En algunos modos de realización, las secuencias de los primer y segundo ADN son diferentes. En algunos modos de realización, las primera y segunda secuencias molde son iguales. En algunos modos de realización, las primera y segunda secuencias molde son diferentes. En algunos modos de realización, las primera y segunda secuencias diana son iguales. En algunos modos de realización, las primera y segunda secuencias diana son diferentes.

45 Un cebador de LPE puede comprender además una región que no es complementaria a una región del molde. En algunos modos de realización, los cebadores de LPE pueden comprender además una secuencia conocida, tal como un sitio de unión de cebador universal o una secuencia complementaria a un sitio de cebado universal. En algunos modos de realización, los cebadores de LPE pueden comprender además una secuencia conocida, tal como un sitio de unión de cebador universal o una secuencia complementaria a un sitio de cebado universal, en el extremo 5'. En algunos modos de realización, la región que no es complementaria a una región del molde es 5' a una región del cebador que es complementaria al molde. En algunos modos de realización, la región que no es complementaria a una región del molde es una región saliente 5' o una región saliente 3'. En algunos modos de realización, la región que no es complementaria a una región del molde comprende un sitio de cebado para amplificación y/o una reacción de secuenciación. En algunos modos de realización, la región que no es complementaria a una región del molde comprende un sitio de cebado para un segundo cebador de un conjunto de cebadores para amplificación y/o una reacción de secuenciación, tal como una reacción PCR o etapa de enriquecimiento por PCR. Opcionalmente, la región que no es complementaria a una región del molde comprende una secuencia universal para la agrupación en una plataforma de secuenciación de alto rendimiento. En algunos modos de realización, la región que no es complementaria a una región del molde comprende un sitio de cebado para un segundo cebador de un conjunto de cebadores para amplificación y/o una reacción de secuenciación, en la que el sitio de cebado para un primer cebador del conjunto de cebadores está contenido dentro del molde de LPE. En algunos modos de realización, el sitio de cebado para un primer cebador del conjunto de cebadores contenido dentro del molde de LPE se añade en una reacción previa de RT, PE, LPE o de adición de adaptador (por ejemplo, ligadura). En algunos modos de realización, se realiza una reacción de LPE usando una ADN polimerasa.

65 En algunos modos de realización, los cebadores de LPE pueden comprender además una segunda UID. Por ejemplo, cada cebador de LPE de una pluralidad de cebadores de LPE puede comprender una segunda UID diferente. Esto puede permitir el código de barras de cada uno de los ADN copiados de las moléculas de ADN que se están sometiendo a una reacción de extensión de cebador lineal con una segunda UID. En algunos modos de realización, la segunda UID es la misma que la UID en las moléculas de ADN que se están sometiendo a una reacción de extensión de cebador lineal. En

5 algunos modos de realización, la segunda UID es diferente a la UID en las moléculas de ADN que se están sometiendo a una reacción de extensión de cebador lineal. En algunos modos de realización, la región de un cebador de LPE que no es complementaria a una región del molde comprende una segunda UID. En algunos modos de realización, la región de cada cebador de LPE de una pluralidad de cebadores de LPE que no es complementaria a una región del ADN diana comprende una segunda UID.

10 En algunos modos de realización, se usan velocidades de rampa lentas para la etapa de extensión/amplificación lineal. En algunos modos de realización, los cebadores de extensión/amplificación lineal se usan a una concentración global fija. En algunos modos de realización, se usa cloruro de magnesio, sulfato de amonio, D-(+)-trehalosa, betaína o una combinación de los mismos durante la etapa de amplificación/extensión lineal.

Enriquecimiento por PCR

15 Un procedimiento puede comprender además realizar una reacción de amplificación exponencial. En algunos modos de realización, un procedimiento puede comprender además realizar PCR. Por ejemplo, una reacción de amplificación exponencial puede utilizar una pluralidad de cebadores directos/inversos y un cebador inverso. En algunos modos de realización, una reacción de amplificación exponencial puede comprender dos o más amplificaciones exponenciales. En algunos modos de realización, una primera y/o segunda reacción de PCR puede utilizar una pluralidad de cebadores directos/inversos y una pluralidad de cebadores inversos. Un primer y/o segundo cebador de una pluralidad de cebadores directos/inversos puede ser un cebador directo/inverso que contiene una región complementaria a los polinucleótidos molde, tal como moléculas de ADN o ADNc. En algunos modos de realización, una pluralidad de cebadores directos/inversos comprende uno o más cebadores directos/inversos en los que cada uno de los cebadores directos/inversos en la pluralidad de cebadores directos/inversos comprende una región complementaria a uno o más sitios de unión de cebador anteriores o posteriores, tales como sitios de unión de cebador universales.

25 En algunos modos de realización, una reacción de amplificación exponencial no se realiza antes de una extensión de cebador o una reacción de transcripción inversa. En algunos modos de realización, una reacción de amplificación exponencial no se realiza antes o se realiza después de generar una secuencia del complemento de polinucleótidos diana, tal como una primera secuencia del complemento.

30 En algunos modos de realización, una reacción de amplificación exponencial no se realiza antes, o se realiza después, de unir un adaptador a un polinucleótido molde. En algunos modos de realización, una reacción de amplificación exponencial no se realiza antes, o se realiza después, de unir un adaptador mediante ligadura a un polinucleótido molde. En algunos modos de realización, una reacción de amplificación exponencial no se realiza antes, o se realiza después, de unir un adaptador a una secuencia del complemento de polinucleótidos diana, tal como una primera secuencia del complemento. En algunos modos de realización, una reacción de amplificación exponencial no se realiza antes, o se realiza después, de unir un adaptador mediante ligadura a una secuencia del complemento de polinucleótidos diana, tal como una primera secuencia del complemento.

40 En algunos modos de realización, una reacción de amplificación exponencial no se realiza antes, o se realiza después, de unir un primer sitio de cebado a una secuencia molde o complemento de la misma, para la amplificación exponencial. Por ejemplo, una reacción de amplificación exponencial puede no realizarse antes, o puede realizarse después, de unir un sitio de cebado para un primer cebador de un conjunto de cebadores a una secuencia molde o complemento de la misma. En algunos modos de realización, una reacción de amplificación exponencial no se realiza antes, o se realiza después, de unir un primer sitio de cebado a una secuencia del complemento de polinucleótidos diana, tal como una primera secuencia del complemento. En algunos modos de realización, una reacción de amplificación exponencial no se realiza antes, o se realiza después, de unir un primer sitio de cebado mediante ligadura a un polinucleótido que comprende una secuencia del complemento de polinucleótidos diana, tal como una primera secuencia del complemento.

50 En algunos modos de realización, una reacción de amplificación exponencial no se realiza antes, o se realiza después, de unir un SBC a una secuencia molde o complemento de la misma. En algunos modos de realización, una reacción de amplificación exponencial no se realiza antes, o se realiza después, de unir un SBC a una secuencia del complemento de polinucleótidos diana, tal como una primera secuencia del complemento. En algunos modos de realización, una reacción de amplificación exponencial se realiza mientras se introduce un SBC mediante amplificación en una secuencia molde o complemento de la misma. En algunos modos de realización, una reacción de amplificación exponencial se realiza mientras se introduce un SBC mediante amplificación en una secuencia del complemento de polinucleótidos diana, tal como una segunda secuencia del complemento.

60 En algunos modos de realización, una reacción de amplificación exponencial no se realiza antes, o se realiza después, de unir una secuencia de cebado universal a una secuencia molde o complemento de la misma. En algunos modos de realización, una reacción de amplificación exponencial no se realiza antes, o se realiza después de, unir una secuencia de cebado universal a una secuencia del complemento de polinucleótidos diana. En algunos modos de realización, una reacción de amplificación exponencial no se realiza antes, o se realiza después, de unir una secuencia de cebado universal mediante ligadura a una secuencia molde o complemento de la misma. En algunos modos de realización, una reacción de amplificación exponencial no se realiza antes, o se realiza después, de unir una secuencia de cebado universal mediante ligadura a una secuencia del complemento de polinucleótidos diana.

5 En algunos modos de realización, una amplificación exponencial no se realiza antes, o se realiza después, de una reacción de amplificación lineal. En algunos modos de realización, una reacción de amplificación exponencial no se realiza antes, o se realiza después, de unir un adaptador a una secuencia molde de amplificación lineal. En algunos modos de realización, una reacción de amplificación exponencial no se realiza antes, o se realiza después, de unir un adaptador mediante ligadura a una secuencia molde de amplificación lineal. En algunos modos de realización, una reacción de amplificación exponencial no se realiza antes, o se realiza después, de unir un primer sitio de cebado a una secuencia molde de amplificación lineal. En algunos modos de realización, una reacción de amplificación exponencial no se realiza antes, o se realiza después, de unir un SBC a una secuencia molde de amplificación lineal. En algunos modos de realización, una reacción de amplificación exponencial no se realiza antes, o se realiza después, de unir una secuencia de cebado universal a una secuencia molde de amplificación lineal.

15 Por ejemplo, una reacción de amplificación exponencial no se puede realizar antes de una reacción de extensión de cebador lineal. Por ejemplo, una reacción de amplificación exponencial se puede realizar después de una reacción de extensión de cebador lineal. En algunos modos de realización, una reacción de amplificación exponencial no se realiza antes, o se realiza después, de generar una o más copias de una secuencia del complemento de polinucleótidos diana, tal como una segunda secuencia del complemento. En algunos modos de realización, una reacción de amplificación exponencial no se realiza antes, o se realiza después, de generar una o más copias de una secuencia del complemento de polinucleótidos diana, tal como una segunda secuencia del complemento, usando un cebador de LPE. En algunos modos de realización, una reacción de amplificación exponencial no se realiza antes, o se realiza después, de generar una o más copias de una pluralidad de secuencias del complemento de polinucleótidos diana, tales como segundas secuencias del complemento, usando una pluralidad de cebadores de LPE. En algunos modos de realización, una reacción de amplificación exponencial no se realiza antes, o se realiza después, de unir un primer y un segundo sitio de cebado para la amplificación exponencial. Por ejemplo, una reacción de amplificación exponencial puede no realizarse antes, o puede realizarse después, de unir un primer sitio de cebado para un primer cebador de un conjunto de cebadores y un segundo sitio de cebado para un segundo cebador del conjunto de cebadores. En algunos modos de realización, una reacción de amplificación exponencial no se realiza antes, o se realiza después, de unir un primer sitio de cebado mediante ligadura y un segundo sitio de cebado para la amplificación exponencial. En algunos modos de realización, una reacción de amplificación exponencial no se realiza antes, o se realiza después, de unir un primer sitio de cebado y un segundo sitio de cebado mediante una reacción de extensión de cebador lineal para la amplificación exponencial. En algunos modos de realización, una reacción de amplificación exponencial no se realiza antes, o se realiza después, de unir un primer sitio de cebado o complemento del mismo mediante ligadura y un segundo sitio de cebado mediante una reacción de extensión de cebador lineal para la amplificación exponencial. Por ejemplo, los primer y segundo sitios de cebado pueden ser sitios de cebado para un par de cebadores usados para la reacción de amplificación exponencial. Por ejemplo, los primer y segundo sitios de cebado pueden ser sitios de cebado universales. Por ejemplo, los primer y segundo sitios de cebado pueden ser sitios de cebado para secuenciación.

40 En algunos modos de realización, una reacción de amplificación exponencial no se realiza antes, o se realiza después, de inmovilizar un polinucleótido en una superficie o soporte. En algunos modos de realización, se realiza una reacción de amplificación exponencial en una copia de un polinucleótido inmovilizado en una superficie o soporte. En algunos modos de realización, se realiza una reacción de amplificación exponencial en una copia de un polinucleótido inmovilizado en una superficie o soporte generado a partir de una reacción de extensión de cebador lineal.

45 En algunos modos de realización, una reacción de amplificación exponencial no se realiza antes, o se realiza después, de inmovilizar una o más secuencias del complemento de polinucleótidos diana en una superficie o soporte. En algunos modos de realización, una reacción de amplificación exponencial no se realiza antes, o se realiza después, de realizar una reacción de cebador lineal en una o más secuencias del complemento de polinucleótidos diana inmovilizadas. En algunos modos de realización, se realiza una reacción de amplificación exponencial en un polinucleótido copiado de un polinucleótido unido a una superficie o soporte sólido. En algunos modos de realización, se realiza una reacción de amplificación exponencial en una secuencia del complemento de polinucleótidos, tal como una segunda secuencia del complemento, copiada de una secuencia del complemento de polinucleótidos diana, tal como una primera secuencia del complemento o secuencia del complemento modificada que puede contener una UID, unida a una superficie o soporte sólido.

55 En algunos modos de realización, se realiza una reacción de amplificación exponencial en un polinucleótido que contiene SBC copiado de un polinucleótido que contiene UID unido a una superficie o soporte sólido. En algunos modos de realización, se realiza una reacción de amplificación exponencial en un polinucleótido que contiene un primer sitio de unión de cebador, un segundo sitio de unión de cebador, o ambos, que se copió de un polinucleótido que contiene UID unido a una superficie o soporte sólido. En algunos modos de realización, se realiza una reacción de amplificación exponencial en un polinucleótido que contiene un primer sitio de unión de cebador, un segundo sitio de unión de cebador, un primer sitio de cebado universal, un segundo sitio de cebado universal, o cualquier combinación de los mismos, que se copió de un polinucleótido que contiene UID, unido a una superficie o soporte sólido.

65 En algunos modos de realización, se realiza una reacción de amplificación exponencial en un polinucleótido que contiene SBC copiado de un polinucleótido que contiene SBC unido a una superficie o soporte sólido. En algunos modos de realización, se realiza una reacción de amplificación exponencial en un polinucleótido que contiene UID copiado de un

5 polinucleótido que contiene SBC unido a una superficie o soporte sólido. En algunos modos de realización, se realiza una reacción de amplificación exponencial en un polinucleótido que contiene un primer sitio de unión de cebador, un segundo sitio de unión de cebador, o ambos, que se copió de un polinucleótido que contiene SBC unido a una superficie o soporte sólido. En algunos modos de realización, se realiza una reacción de amplificación exponencial en un polinucleótido que contiene un primer sitio de unión de cebador, un segundo sitio de unión de cebador, un primer sitio de cebado universal, un segundo sitio de cebado universal, o cualquier combinación de los mismos, que se copió de un polinucleótido que contiene SBC, unido a una superficie o soporte sólido.

10 En algunos modos de realización, se realiza una reacción de amplificación exponencial en un polinucleótido que contiene un primer y/o segundo sitio de cebador copiado de un polinucleótido que contiene un primer y/o segundo sitio de cebador unido a una superficie o soporte sólido. En algunos modos de realización, se realiza una reacción de amplificación exponencial en un polinucleótido que contiene SBC copiado de un polinucleótido que contiene un primer y/o segundo sitio de cebador unido a una superficie o soporte sólido. En algunos modos de realización, se realiza una reacción de amplificación exponencial en un polinucleótido que contiene UID copiado de un polinucleótido que contiene un primer y/o segundo sitio de cebador unido a una superficie o soporte sólido. En algunos modos de realización, se realiza una reacción de amplificación exponencial en un polinucleótido que contiene un primer sitio de unión de cebador universal, un segundo sitio de unión de cebador universal, o ambos, que se copió de un polinucleótido que contiene un primer y/o segundo sitio de cebador unido a una superficie o soporte sólido. En algunos modos de realización, se realiza una reacción de amplificación exponencial en un polinucleótido que contiene un primer sitio de unión de cebador, un segundo sitio de unión de cebador, un primer sitio de cebado universal, un segundo sitio de cebado universal, o cualquier combinación de los mismos, que se copió de un polinucleótido que contiene un primer y/o segundo sitio de cebador unido a una superficie o soporte sólido.

25 Usando cebadores descritos en el presente documento, los polinucleótidos de ADN producidos a partir de uno o más de los procedimientos o etapas de procedimiento anteriores se pueden hibridar con un conjunto de cebadores (por ejemplo, un conjunto de cebadores de PCR o un conjunto de cebadores de amplificación exponencial) y la amplificación exponencial se puede realizar usando reactivos adecuados conocidos en la técnica. Por ejemplo, una o más segundas secuencias del complemento se pueden hibridar con el primer cebador de un conjunto de cebadores (tal como un cebador inverso) y se puede realizar la extensión de cebador; un segundo cebador de un conjunto de cebadores (tal como un cebador directo) se puede hibridar a continuación con un producto de la reacción de extensión y se puede realizar la extensión de cebador.

35 En algunos modos de realización, la amplificación exponencial comprende múltiples ciclos. En algunos modos de realización, los mismos primer y segundo cebadores de un conjunto de cebadores se usan para la reacción de amplificación exponencial de múltiples polinucleótidos molde. En algunos modos de realización, uno o más de los cebadores de amplificación exponencial no son cebadores específicos de diana. En algunos modos de realización, ambos cebadores de un conjunto de cebadores de amplificación exponencial no son cebadores específicos de diana. En algunos modos de realización, los mismos primer y segundo cebadores de un conjunto de cebadores se usan para la reacción de amplificación exponencial de múltiples polinucleótidos molde en el mismo recipiente de reacción. En algunos modos de realización, los mismos primer y segundo cebadores de un conjunto de cebadores se usan para la reacción de amplificación exponencial de múltiples polinucleótidos molde en la misma reacción. En algunos modos de realización, los mismos primer y segundo cebadores de un conjunto de cebadores se usan para la reacción de amplificación exponencial de múltiples polinucleótidos molde simultáneamente. Por ejemplo, los mismos primer y segundo cebadores de un conjunto de cebadores se pueden usar para amplificar exponencialmente una pluralidad de secuencias del complemento de polinucleótidos diana, tales como una pluralidad de segundas secuencias del complemento derivadas de una secuencia diana diferente. Por ejemplo, los mismos primer y segundo cebadores de un conjunto de cebadores se pueden usar para amplificar exponencialmente una pluralidad de secuencias del complemento de polinucleótidos diana, tales como una pluralidad de segundas secuencias del complemento derivadas de una secuencia diana diferente. Por ejemplo, los mismos primer y segundo cebadores de un conjunto de cebadores se pueden usar para amplificar exponencialmente una pluralidad de secuencias del complemento de polinucleótidos diana, tales como una pluralidad de segundas secuencias del complemento de polinucleótidos diana, tales como una pluralidad de segundas secuencias del complemento de la misma. Por ejemplo, los mismos primer y segundo cebadores de un conjunto de cebadores se pueden usar para amplificar exponencialmente una pluralidad de secuencias del complemento de polinucleótidos diana de un amplicón. Por ejemplo, los mismos primer y segundo cebadores de un conjunto de cebadores se pueden usar para amplificar exponencialmente una pluralidad de secuencias del complemento de polinucleótidos diana, tales como una pluralidad de segundas secuencias del complemento de un conjunto de amplicones. Por ejemplo, los mismos primer y segundo cebadores de un conjunto de cebadores se pueden usar para amplificar exponencialmente cada una de una pluralidad de secuencias del complemento de polinucleótidos diana generadas usando cualquiera de los procedimientos descritos en el presente documento. Por ejemplo, los mismos primer y segundo cebadores de un conjunto de cebadores se pueden usar para amplificar exponencialmente cada una de una pluralidad de secuencias del complemento de polinucleótidos diana que contienen una secuencia adaptadora. Por ejemplo, los mismos primer y segundo cebadores de un conjunto de cebadores se pueden usar para amplificar exponencialmente cada una de una pluralidad de secuencias del complemento de polinucleótidos diana que contienen un SBC. Por ejemplo, los mismos primer y segundo cebadores de un conjunto de cebadores se pueden usar para amplificar exponencialmente cada una de una pluralidad de secuencias del complemento de polinucleótidos diana que contienen un primer y un segundo sitio de cebado universal.

5 En algunos modos de realización, los primer y segundo cebadores de un conjunto de cebadores se pueden usar para amplificar exponencialmente una UID, un SBC, una región diana, cualquier complemento de los mismos o cualquier combinación de los mismos. Por ejemplo, los primer y segundo sitios de unión de cebador se pueden hibridar 5' y 3', respectivamente, con una UID, un SBC, una región diana, cualquier complemento de los mismos o cualquier combinación de los mismos.

10 En algunos modos de realización, una reacción de amplificación exponencial es la segunda, tercera, cuarta o quinta etapa para generar una biblioteca de polinucleótidos a partir de una muestra que contiene un polinucleótido diana. En algunos modos de realización, una reacción de amplificación exponencial no es la segunda etapa para generar una biblioteca de polinucleótidos a partir de una muestra que contiene un polinucleótido diana. En algunos modos de realización, una reacción de amplificación exponencial no es la primera reacción de amplificación realizada en un procedimiento para generar una biblioteca de polinucleótidos a partir de una muestra que contiene un polinucleótido diana. En algunos modos de realización, una reacción de amplificación exponencial es la tercera etapa para generar una biblioteca de polinucleótidos a partir de una muestra que contiene un polinucleótido diana. En algunos modos de realización, una reacción de amplificación exponencial es la cuarta etapa para generar una biblioteca de polinucleótidos a partir de una muestra que contiene un polinucleótido diana. En algunos modos de realización, una reacción de amplificación exponencial es la quinta etapa para generar una biblioteca de polinucleótidos a partir de una muestra que contiene un polinucleótido diana. En algunos modos de realización, se realiza una reacción de amplificación exponencial después de una reacción de RT o PE. En algunos modos de realización, se realiza una reacción de amplificación exponencial después de una reacción que añade un adaptador a una secuencia del complemento de polinucleótidos diana, tal como una primera secuencia del complemento. En algunos modos de realización, se realiza una reacción de amplificación exponencial después de una reacción de RT o PE y después de una reacción que añade un adaptador a una secuencia del complemento de polinucleótidos diana, tal como una primera secuencia del complemento. En algunos modos de realización, se realiza una reacción de amplificación exponencial antes de realizar una segunda reacción de amplificación exponencial, tal como PCR. En algunos modos de realización, la amplificación exponencial se realiza en la siguiente etapa después de la extensión de cebador lineal. En algunos modos de realización, la amplificación exponencial no se realiza en la siguiente etapa después de la extensión de cebador lineal. En algunos modos de realización, la amplificación exponencial no se realiza en la siguiente etapa después de una reacción de RT o PE. En algunos modos de realización, la amplificación exponencial no se realiza en las siguientes 2 etapas después de una reacción de RT o PE. En algunos modos de realización, la amplificación exponencial no se realiza en las siguientes 3 etapas después de una reacción de RT o PE. En algunos modos de realización, una biblioteca de secuencias de polinucleótidos, que puede contener una UID, producida a partir de una etapa de amplificación exponencial, no se amplifica más después de esta etapa. En algunos modos de realización, el procedimiento comprende solo un ciclo de amplificación exponencial. En algunos modos de realización, el procedimiento comprende extender repetidamente ambos cebadores de un conjunto de cebadores para producir múltiples copias de las secuencias de polinucleótidos que pueden contener una UID.

40 Los cebadores de amplificación exponencial pueden comprender una secuencia complementaria a una secuencia, o secuencia del complemento de una secuencia del complemento de polinucleótidos diana. Por ejemplo, los uno o más cebadores de amplificación exponencial pueden comprender una secuencia complementaria a una secuencia o secuencia del complemento de una secuencia del complemento de polinucleótidos diana o un polinucleótido diana en una muestra inicial. Por ejemplo, los uno o más cebadores de amplificación exponencial pueden comprender una secuencia complementaria a una secuencia o secuencia del complemento de una secuencia del complemento de polinucleótidos diana que es un producto de una reacción de amplificación, reacción de ligadura, extensión de cebador, extensión de cebador lineal o combinaciones de las mismas. Por ejemplo, los uno o más cebadores de amplificación exponencial pueden comprender una secuencia complementaria a una secuencia o secuencia del complemento de una secuencia primera o segunda o modificada.

50 En algunos modos de realización, los uno o más cebadores de amplificación exponencial no comprenden una secuencia complementaria a una secuencia o secuencia del complemento de un polinucleótido diana. En algunos modos de realización, los uno o más cebadores de amplificación exponencial no comprenden una secuencia complementaria a una secuencia o secuencia del complemento de una secuencia del complemento de polinucleótidos diana. En algunos modos de realización, los uno o más cebadores de amplificación exponencial no comprenden una secuencia que es complementaria a una secuencia o secuencia del complemento de un polinucleótido diana y no comprenden una secuencia que es complementaria a una secuencia o secuencia del complemento de una secuencia del complemento de polinucleótidos diana.

60 En algunos modos de realización, los uno o más cebadores de amplificación exponencial comprenden una secuencia complementaria a una secuencia o secuencia del complemento de un polinucleótido diana. En algunos modos de realización, los uno o más cebadores de amplificación exponencial comprenden una secuencia complementaria a una secuencia o secuencia del complemento de un polinucleótido que contiene UID. En algunos modos de realización, los uno o más cebadores de amplificación exponencial comprenden una secuencia que es complementaria a una secuencia o secuencia del complemento de un polinucleótido diana y comprenden una secuencia que es complementaria a una secuencia o secuencia del complemento de un polinucleótido que contiene UID.

65 En algunos modos de realización, la secuencia complementaria a un polinucleótido que contiene UID de uno o más cebadores de amplificación exponencial no es complementaria a una secuencia diana. En algunos modos de realización,

la secuencia complementaria a un polinucleótido que contiene UID de uno o más cebadores de amplificación exponencial no es complementaria a ningún polinucleótido que no contenga una UID. En algunos modos de realización, las secuencias complementarias a un polinucleótido que contiene UID de uno o más cebadores de amplificación exponencial no son complementarias a ningún otro polinucleótido en una muestra.

5 En algunos modos de realización, la secuencia del complemento de polinucleótidos diana amplificada exponencialmente es un polinucleótido monocatenario. En algunos modos de realización, la secuencia del complemento de polinucleótidos diana amplificada exponencialmente es un polinucleótido bicatenario. En algunos modos de realización, la secuencia del complemento de polinucleótidos diana amplificada exponencialmente es una copia de un producto de extensión de una
10 reacción de PE o RT. En algunos modos de realización, la secuencia del complemento de polinucleótidos diana amplificada exponencialmente comprende además una secuencia adaptadora, tal como una secuencia adaptadora ligada. En algunos modos de realización, la secuencia del complemento de polinucleótidos diana amplificada exponencialmente es un complemento de un producto de extensión de una reacción de PE o RT que comprende además una secuencia adaptadora. En algunos modos de realización, la secuencia del complemento de polinucleótidos diana amplificada exponencialmente es un complemento de una secuencia del complemento de un producto de extensión de una reacción
15 de PE o RT que comprende además un primer y/o segundo sitio de unión de cebador, tal como un sitio de PCR, secuenciación o cebado universal. En algunos modos de realización, la secuencia del complemento de polinucleótidos diana amplificada exponencialmente se inmoviliza en un sustrato o superficie. En algunos modos de realización, la secuencia del complemento de polinucleótidos diana amplificada exponencialmente comprende un SBC.

20 En algunos modos de realización, la secuencia complementaria a una secuencia del complemento de polinucleótidos diana amplificada exponencialmente de uno o más cebadores de amplificación exponencial no es una secuencia en un polinucleótido diana. En algunos modos de realización, la secuencia complementaria a una secuencia del complemento de polinucleótidos diana amplificada exponencialmente de uno o más cebadores de amplificación exponencial es complementaria a una secuencia del complemento de una secuencia generada durante una reacción de RT o PE. En algunos modos de realización, la secuencia complementaria a una secuencia del complemento de polinucleótidos diana
25 amplificada exponencialmente de uno o más cebadores de amplificación exponencial es complementaria a una secuencia de un polinucleótido diana que se hibrida con una secuencia de la diana 5' a la secuencia del polinucleótido diana complementaria a un cebador de RT o PE. En algunos modos de realización, la secuencia complementaria a una secuencia del complemento de polinucleótidos diana amplificada exponencialmente de uno o más cebadores de amplificación exponencial es complementaria a una secuencia de un polinucleótido diana que se hibrida con una
30 secuencia de la diana 3' a la secuencia del polinucleótido diana complementaria a un cebador de RT o PE. En algunos modos de realización, una secuencia de un polinucleótido diana que contiene una variante o una región para su análisis por cualquiera de los procedimientos descritos en el presente documento puede estar entre la secuencia del polinucleótido diana complementaria a uno o más cebadores de RT o PE y la secuencia del polinucleótido diana complementaria a uno o más cebadores de amplificación exponencial.

35 En algunos modos de realización, la secuencia complementaria a una secuencia del complemento de polinucleótidos diana amplificada exponencialmente de uno o más cebadores de amplificación exponencial no es una secuencia complementaria a una secuencia de uno o más cebadores de PE o RT. En algunos modos de realización, la secuencia complementaria a una secuencia del complemento de polinucleótidos diana amplificada exponencialmente de uno o más cebadores de amplificación exponencial no es una secuencia complementaria a una secuencia específica de diana de uno o más cebadores de PE o RT.

40 En algunos modos de realización, los uno o más cebadores de amplificación exponencial comprenden un primer cebador de amplificación exponencial con una región complementaria a una secuencia de un primer polinucleótido molde, y un segundo cebador de amplificación exponencial con una región complementaria a una secuencia de un segundo polinucleótido molde. Por ejemplo, el primer polinucleótido molde puede ser una primera molécula de ADN y el segundo polinucleótido molde puede ser una segunda molécula de ADN. Por ejemplo, el primer polinucleótido molde puede ser una primera molécula de ADN derivada de un primer polinucleótido diana en una muestra y el segundo polinucleótido molde puede ser una segunda molécula de ADN derivada de un segundo polinucleótido diana en una muestra. En algunos
45 modos de realización, los uno o más cebadores de amplificación exponencial comprenden un primer cebador de amplificación exponencial con una región complementaria a una secuencia de un primer ADN, y uno o más segundos cebadores de amplificación exponencial, cada uno con una región complementaria a una secuencia de uno o más segundos ADN. En algunos modos de realización, las secuencias de los primer y segundo ADN son las mismas. En algunos modos de realización, las secuencias de los primer y segundo ADN son diferentes. En algunos modos de realización, las primera y segunda secuencias molde son iguales. En algunos modos de realización, las primera y segunda secuencias molde son diferentes. En algunos modos de realización, las primera y segunda secuencias diana son iguales. En algunos modos de realización, las primera y segunda secuencias diana son diferentes.

60 **Secuenciación**

Después de realizar uno o más de los procedimientos o etapas de procedimiento descritos en el presente documento, se puede secuenciar una biblioteca de polinucleótidos generados.

65 La secuenciación se puede realizar mediante cualquier procedimiento de secuenciación conocido en la técnica. En

algunos modos de realización, la secuenciación se puede realizar con alto rendimiento. Las tecnologías de secuenciación de próxima generación adecuadas incluyen la plataforma 454 Life Sciences (Roche, Branford, CT) (Margulies y col., Nature, 437, 376-380 (2005)); Analizador de genoma de Illumina, ensayo de metilación GoldenGate o ensayos de metilación Infinium, es decir, matriz de metilación Infinium HumanMethylation 27K BeadArray o VeraCode GoldenGate (Illumina, San Diego, CA; Bibkova y col., Genome Res. 16, 383-393 (2006); y patentes de Estados Unidos n.º 6.306.597, 7.598.035, 7.232.656) o secuenciación de ADN por ligadura, sistema SOLiD (Applied Biosystems/Life Technologies; patentes de Estados Unidos n.º 6.797.470, 7.083.917, 7.166.434, 7.320.865, 7.332.285, 7.364.858 y 7.429.453); o la tecnología de secuenciación de ADN de molécula única Helico True (Harris y col., Science, 320, 106-109 (2008); y las patentes de Estados Unidos n.º 7.037.687, 7.645.596, 7.169.560 y 7.769.400), la tecnología de molécula única, en tiempo real (SMRTTm) de Pacific Biosciences, y secuenciación (Soni y col., Clin. Chem. 53, 1996-2001 (2007)). Un procedimiento puede comprender además secuenciar uno o más polinucleótidos en la biblioteca. Un procedimiento puede comprender además alinear una o más secuencias de polinucleótidos, lecturas de secuencia, secuencias de amplicón o secuencias de conjuntos de amplicones en la biblioteca entre sí.

Como se usa en el presente documento, la alineación comprende comparar una secuencia de prueba, tal como una secuencia leída, con una o más secuencias de prueba, secuencias de referencia o una combinación de las mismas. En algunos modos de realización, la alineación se puede usar para determinar una secuencia consenso a partir de una pluralidad de secuencias o secuencias alineadas. En algunos modos de realización, la alineación comprende determinar una secuencia consenso a partir de una pluralidad de secuencias en las que cada una tiene una UID idéntica. En algunos modos de realización, la longitud de una secuencia alineada para propósitos de comparación es al menos el 30 %, al menos el 40 %, al menos el 50 %, al menos el 60 %, al menos el 70 %, al menos el 80 %, al menos el 90 % o al menos el 95 % de la longitud de una secuencia de referencia. La comparación real de las dos o más secuencias se puede lograr mediante procedimientos bien conocidos, por ejemplo, usando un algoritmo matemático. Un ejemplo no limitante de dicho algoritmo matemático se describe en Karlin, S. y Altschul, S., Proc. Natl. Acad. Sci. Estados Unidos, 90-5873-5877 (1993). Dicho algoritmo se incorpora en los programas NBLAST y XBLAST (versión 2.0), como se describe en Altschul, S. y col., Nucleic Acids Res., 25:3389-3402 (1997). Cuando se utilizan los programas BLAST y Gapped BLAST, se puede usar cualquier parámetro relevante de los respectivos programas (por ejemplo, NBLAST). Por ejemplo, los parámetros para la comparación de secuencias se pueden establecer en puntuación = 100, longitud de palabra = 12, o se pueden variar (por ejemplo, W = 5 o W = 20). Otros ejemplos incluyen el algoritmo de Myers y Miller, CABIOS (1989), ADVANCE, ADAM, BLAT y FASTA. En algunos modos de realización, el porcentaje de identidad entre dos secuencias de aminoácidos se puede lograr usando, por ejemplo, el programa GAP en el paquete de software GCG (Accelrys, Cambridge, Reino Unido).

En algunos aspectos, determinar el número de polinucleótidos, amplicones o conjuntos de amplicones con diferentes secuencias puede comprender determinar las secuencias de los polinucleótidos, amplicones o conjuntos de amplicones. En algunos aspectos, determinar el número de polinucleótidos, amplicones o conjuntos de amplicones que contienen UID diferentes puede comprender determinar la secuencia de los polinucleótidos, amplicones o conjuntos de amplicones que contienen UID. Determinar la secuencia de un polinucleótido puede comprender llevar a cabo una reacción de secuenciación para determinar la secuencia de al menos una parte de la región diana, UID, SBC, al menos una parte del polinucleótido, un complemento del mismo, un complemento inverso del mismo o cualquier combinación de los mismos. En algunos modos de realización, solo se secuencian la UID o una parte de la UID. En algunos modos de realización, solo se secuencian el SBC o una parte del SBC. En algunos modos de realización, solo se secuencian la región diana o una parte de la región diana. En algunos modos de realización, se puede producir una reacción de secuenciación en un soporte como se describe en el presente documento, en un seguimiento continuo, en una dilución o en uno o más volúmenes físicamente separados.

La secuenciación puede comprender al menos aproximadamente 200, 300, 400, 500, 600, 700, 800, 900, 1000 o más lecturas de secuenciación por ejecución. Como se usa en el presente documento, una lectura de secuencia comprende una secuencia de nucleótidos determinada a partir de una secuencia o flujo de datos generados por una técnica de secuenciación. En algunos modos de realización, la secuenciación comprende secuenciar al menos aproximadamente 1500, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10.000 o más lecturas de secuenciación por ejecución. La secuenciación puede comprender más de, menos de o igual a aproximadamente 1.000.000.000 de lecturas de secuenciación por ejecución. La secuenciación puede comprender más de, menos de o igual a aproximadamente 200.000.000 de lecturas por ejecución.

Un procedimiento puede comprender determinar una secuencia de un polinucleótido diana determinando una secuencia consenso a partir de dos o más lecturas de secuencia. En algunos modos de realización, un promedio de 5-50 o 20-30 lecturas sin procesar por UID proporciona un equilibrio deseado de exactitud de secuencia consenso y suficiente profundidad de secuenciación (los recuentos de lectura sin procesar más altos pueden necesitar una mayor profundidad de secuenciación). En algunos modos de realización, la exactitud (por ejemplo, la distribución normal agregada) se puede mejorar al alinear y colapsar las lecturas de secuencia en secuencias consenso usando información de UID. Un rasgo característico de la exactitud del consenso de UID es la capacidad mejorada para determinar con exactitud la presencia o ausencia de una mutación o SNP en un segundo alelo, lo que da como resultado una lectura exacta de la heterocigosis de un paciente con un SNP detectado.

Un procedimiento puede comprender generar una secuencia consenso a partir de una o más alineaciones, tales como una o más alineaciones de una o más secuencias de polinucleótidos, lecturas de secuencia, secuencias de amplicón o

secuencias de conjuntos de amplicones en la biblioteca entre sí. Una secuencia consenso determinada usando los procedimientos y bibliotecas producidos, como se describe en el presente documento, puede mejorar la exactitud de las secuencias nucleotídicas procedentes de una secuenciación. Por ejemplo, una secuencia consenso determinada puede tener una puntuación de calidad mejorada en comparación con otros procedimientos en la técnica. Como se usa en el presente documento, una puntuación de calidad comprende una medida de la probabilidad de que una asignación de base en una localización de secuencia particular sea correcta. Por tanto, un valor de puntuación de calidad se puede relacionar con una probabilidad de una lectura automática de nucleótidos. Los procedimientos descritos en el presente documento se pueden usar para determinar una secuencia de polinucleótidos diana con una puntuación de calidad de aproximadamente, o al menos aproximadamente 10. Los procedimientos descritos en el presente documento pueden disminuir o usar un número inferior de lecturas de secuencia para lograr la misma o mayor confianza en la exactitud de la secuencia. En algunos modos de realización, se usan menos lecturas de secuencia en un procedimiento descrito en el presente documento que emplea el uso de UID que un procedimiento similar sin el uso de UID para determinar una secuencia con una confianza igual o similar con respecto a la exactitud de la lectura automática de nucleótidos.

En algunos modos de realización, las lecturas de secuencia sin ambos sitios de cebado de amplificación exponencial o complementos de los mismos, una secuencia adaptadora, un SBC, una UID opcional, dos secuencias de cebado universales, o cualquier combinación de los mismos, pueden ser lecturas erróneas. Un procedimiento puede comprender secuenciar lecturas erróneas. Un procedimiento puede comprender determinar el número de lecturas erróneas, tal como para determinar una condición de reacción o diseñar secuencias de cebador. La comparación del número de lecturas erróneas generadas en una o más primeras condiciones o conjuntos de condiciones se puede usar para determinar una condición o conjunto de condiciones preferente. Por ejemplo, un primer procedimiento se puede llevar a cabo a una alta concentración de sal durante una reacción de PCR, y un segundo procedimiento se puede llevar a cabo a una baja concentración de sal durante una reacción de PCR, en el que el primer y segundo procedimiento se llevan a cabo sustancialmente de la misma manera independientemente de la diferencia de concentración de sal. Si el primer procedimiento da como resultado un mayor número de lecturas erróneas, tal como un mayor número de lecturas erróneas para una secuencia o cebador de polinucleótidos diana particular, se puede determinar que se prefiere una condición de reacción de sal inferior para esa secuencia o cebador de polinucleótidos diana particular.

En algunos modos de realización, solo las lecturas de secuencia con ambos sitios de cebado de amplificación exponencial o complementos de los mismos, una secuencia adaptadora, un SBC, una UID opcional, dos secuencias de cebado universales, o cualquier combinación de los mismos, se usan para alinear o determinar una secuencia consenso. En algunos modos de realización, una o más lecturas de secuencia sin ambos sitios de cebado de amplificación exponencial o complementos de los mismos, una secuencia adaptadora, un SBC, una UID opcional, dos secuencias de cebado universales, o cualquier combinación de los mismos, no se usan para alinear o determinar una secuencia consenso.

En algunos modos de realización, una o más lecturas de secuencia sin ambos sitios de cebado de amplificación exponencial o complementos de los mismos no se usan para alinear o determinar una secuencia consenso. En algunos modos de realización, una o más lecturas de secuencia sin un único sitio de cebado de amplificación exponencial (por ejemplo, sitio de cebado de PCR) o complemento del mismo no se usan para alinear o determinar una secuencia consenso. En algunos modos de realización, una o más lecturas de secuencia que comprenden dos sitios de cebado de amplificación exponencial o complementos de los mismos no se usan para alinear o determinar una secuencia consenso, cuando los dos sitios de cebado de amplificación exponencial no son sitios de cebado de amplificación exponencial correspondientes para un par de cebadores usados, tales como un par de cebadores usados en una reacción de PCR.

En algunos modos de realización, solo las lecturas de secuencia con ambos sitios de cebado de amplificación exponencial o complementos de los mismos se usan para alinear o determinar una secuencia consenso. En algunos modos de realización, solo las lecturas de secuencia con dos sitios de cebado de amplificación exponencial o complementos de los mismos que corresponden a sitios de cebado de amplificación exponencial para un par de cebadores usados, tal como un par de cebadores usados en una reacción de PCR, se usan para alinear o determinar una secuencia consenso. En algunos modos de realización, una o más lecturas de secuencia sin un SBC no se usan para alinear o determinar una secuencia consenso. En algunos modos de realización, solo las lecturas de secuencia con un SBC se usan para alinear o determinar una secuencia consenso. En la mayoría de los modos de realización, una o más lecturas de secuencia sin una UID no se usan para alinear o determinar una secuencia consenso. En la mayoría de los modos de realización, solo las lecturas de secuencia con una UID se usan para alinear o determinar una secuencia consenso. En algunos modos de realización, una o más lecturas de secuencia sin una secuencia adaptadora no se usan para alinear o determinar una secuencia consenso. En algunos modos de realización, solo las lecturas de secuencia con una secuencia adaptadora se usan para alinear o determinar una secuencia consenso. En algunos modos de realización, una o más lecturas de secuencia sin dos secuencias de cebado universales no se usan para alinear o determinar una secuencia consenso. En algunos modos de realización, solo las lecturas de secuencia con dos secuencias de cebado universales se usan para alinear o determinar una secuencia consenso.

En algunos modos de realización, una secuencia se puede determinar como exacta cuando está presente al menos el 5 % de las secuencias que contienen la misma UID, las secuencias en un amplicón o las secuencias en un conjunto de amplicones. Por ejemplo, una secuencia se puede determinar como exacta cuando están presentes al menos el 10 %, 15 %, 20 %, 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 92 %, 95 %, 97 %, 98 %, 99 % o más de las secuencias que contienen la misma UID, las secuencias en un amplicón o las secuencias en un

conjunto de amplicones. Por ejemplo, una secuencia se puede determinar como exacta cuando están presentes al menos aproximadamente el 75 % a aproximadamente el 99 % de las secuencias que contienen la misma UID, las secuencias en un amplicón o las secuencias en un conjunto de amplicones. Por ejemplo, una secuencia se puede determinar como exacta cuando están presentes al menos aproximadamente el 85% a aproximadamente el 99 % de las secuencias que contienen la misma UID, las secuencias en un amplicón o las secuencias en un conjunto de amplicones. Por ejemplo, una secuencia se puede determinar como exacta cuando están presentes al menos aproximadamente el 92% a aproximadamente el 99 % de las secuencias que contienen la misma UID, las secuencias en un amplicón o las secuencias en un conjunto de amplicones.

En algunos modos de realización, se emplean químicas de secuenciación que tienen tasas de error relativamente altas. En dichos modos de realización, las puntuaciones de calidad promedio producidas por dichas químicas son funciones que disminuyen monótonamente de las longitudes de lectura de secuencia. En un modo de realización, dicha disminución corresponde al 0,5 por ciento de las lecturas de secuencia que tienen al menos un error en las posiciones 1-75; el 1 por ciento de las lecturas de secuencia tienen al menos un error en las posiciones 76-100; y el 2 por ciento de las lecturas de secuencia tienen al menos un error en las posiciones 101-125.

Polinucleótidos diana

Los procedimientos descritos en el presente documento se pueden usar para generar una biblioteca de polinucleótidos a partir de uno o más polinucleótidos diana para secuenciación. Los polinucleótidos diana incluyen cualquier polinucleótido de interés que no sea producto de una reacción de amplificación. Por ejemplo, un polinucleótido diana puede incluir un polinucleótido en una muestra biológica. Por ejemplo, los polinucleótidos diana no incluyen productos de una reacción de PCR. Por ejemplo, los polinucleótidos diana pueden incluir un molde de polinucleótidos usado para generar productos de una reacción de amplificación, pero no incluyen los productos de amplificación en sí. Por ejemplo, los polinucleótidos diana incluyen polinucleótidos de interés que se pueden someter a una reacción de transcripción inversa o una reacción de extensión de cebador. Por ejemplo, los polinucleótidos diana incluyen ARN o ADN. En algunos modos de realización, los polinucleótidos de ARN diana son ARNm. En algunos modos de realización, los polinucleótidos de ARN diana están poliadenilados. En algunos modos de realización, los polinucleótidos de ARN no están poliadenilados. En algunos modos de realización, los polinucleótidos diana son polinucleótidos de ADN. Los polinucleótidos de ADN pueden ser ADN genómico. Los polinucleótidos de ADN pueden comprender exones, intrones, regiones no traducidas o cualquier combinación de los mismos.

En algunos modos de realización, las bibliotecas se pueden generar a partir de dos o más regiones de un polinucleótido diana. En algunos modos de realización, se pueden generar bibliotecas de procedimientos a partir de dos o más polinucleótidos diana. En algunos modos de realización, los polinucleótidos diana son ácidos nucleicos genómicos o ADN derivado de cromosomas. En algunos modos de realización, los polinucleótidos diana incluyen secuencias que comprenden una variante, tal como un polimorfismo o mutación. En algunos modos de realización, los polinucleótidos diana incluyen ADN y no ARN. En algunos modos de realización, los polinucleótidos diana incluyen ARN y no ADN. En algunos modos de realización, los polinucleótidos diana incluyen ADN y ARN. En algunos modos de realización, un polinucleótido diana es una molécula de ARNm. En algunos modos de realización, un polinucleótido diana es una molécula de ADN. En algunos modos de realización, un polinucleótido diana es un polinucleótido monocatenario. En algunos modos de realización, un polinucleótido diana es un polinucleótido bicatenario. En algunos modos de realización, un polinucleótido diana es una cadena sencilla de un polinucleótido bicatenario.

Los polinucleótidos diana se pueden obtener a partir de cualquier muestra biológica y preparar usando procedimientos conocidos en la técnica. En algunos modos de realización, los polinucleótidos diana se aíslan directamente sin amplificación. Son conocidos en la técnica procedimientos para el aislamiento directo. Los ejemplos no limitantes incluyen la extracción de ADN genómico o ARNm a partir de una muestra biológica, organismo o célula.

En algunos modos de realización, uno o más polinucleótidos diana se purifican a partir de una muestra biológica. En algunos modos de realización, un polinucleótido diana no se purifica a partir de la muestra biológica en la que está contenido. En algunos modos de realización, un polinucleótido diana se aísla a partir de una muestra biológica. En algunos modos de realización, un polinucleótido diana no se aísla a partir de la muestra biológica en la que está contenido. Por ejemplo, en algunos modos de realización, un polinucleótido diana no se extrae o purifica a partir de la muestra. Por ejemplo, en algunos modos de realización, un ARNm diana no se purifica a partir de una muestra, tal como a través de un procedimiento de purificación de poli-A. En algunos modos de realización, un polinucleótido diana puede ser un ácido nucleico libre de células. En algunos modos de realización, un polinucleótido diana puede ser un ácido nucleico transcrito. En algunos modos de realización, un polinucleótido diana es un polinucleótido modificado. En algunos modos de realización, un polinucleótido diana es un polinucleótido no modificado.

En algunos modos de realización, un polinucleótido diana es un polinucleótido de una sola célula. En algunos modos de realización, los polinucleótidos diana son de células individuales. En algunos modos de realización, un polinucleótido diana es un polinucleótido de una muestra que contiene una pluralidad de células.

En algunos modos de realización, un polinucleótido diana codifica una secuencia de biomarcador. En algunos modos de

pluralidad de polinucleótidos diana que comprenden una primera diana que es ADN y una segunda diana que es ADN. Por ejemplo, un panel de polinucleótidos diana puede comprender una pluralidad de polinucleótidos diana que comprenden una primera diana que es ADN genómico y una segunda diana que es ADN genómico. Por ejemplo, un panel de polinucleótidos diana puede comprender una pluralidad de polinucleótidos diana que comprenden una primera diana que es ADN celular y una segunda diana que es ADN circulante.

En algunos modos de realización, los tipos de biomarcadores de dos o más polinucleótidos diana en un panel de polinucleótido diana son diferentes. Por ejemplo, un panel de polinucleótidos diana puede comprender una pluralidad de biomarcadores que comprenden un primer biomarcador para un locus genético, un segundo biomarcador para una secuencia variante. Por ejemplo, un panel de polinucleótidos diana puede comprender una pluralidad de biomarcadores que comprenden un primer biomarcador para un SNP y un segundo biomarcador para una mutación. En algunos modos de realización, los tipos de biomarcadores de dos o más polinucleótidos diana en un panel de polinucleótido diana son los mismos. Por ejemplo, un panel de polinucleótidos diana puede comprender una pluralidad de biomarcadores que comprenden un primer biomarcador de un locus genético, un segundo biomarcador para otro locus genético. Por ejemplo, un panel de polinucleótidos diana puede comprender una pluralidad de biomarcadores que comprenden un primer biomarcador para un SNP, un segundo biomarcador para un SNP.

En algunos modos de realización, un procedimiento puede comprender además diagnosticar o pronosticar a un sujeto una enfermedad, trastorno, síntoma y/o afección con al menos un 50 % de confianza. En algunos modos de realización, la presencia, ausencia, nivel, secuencia o cualquier combinación de las mismas, de un polinucleótido diana en el sujeto, tal como un biomarcador, se puede determinar con al menos un 50 % de confianza. En algunos modos de realización, la presencia, ausencia, nivel, secuencia o cualquier combinación de las mismas, de un polinucleótido diana en el sujeto se puede determinar con un 50 % - 100 % de confianza.

Muestras

Como se usa en el presente documento, una muestra comprende una fuente o muestra biológica, ambiental, médica o del paciente que contiene un polinucleótido, tal como un polinucleótido diana. Cualquier muestra biológica que contenga polinucleótidos se puede usar en los procedimientos descritos en el presente documento. Por ejemplo, una muestra puede ser una muestra biológica de un sujeto que contiene ARN o ADN. Los polinucleótidos se pueden extraer de la muestra biológica, o la muestra se puede someter directamente a los procedimientos sin extracción de los polinucleótidos. La muestra puede ser ADN o ARN extraído o aislado. Una muestra también puede ser ARN o ADN total extraído de una muestra biológica, una biblioteca de ADNc, ADN vírico o genómico. En un modo de realización, los polinucleótidos se aíslan de una muestra biológica que contiene una variedad de otros componentes, tales como proteínas, lípidos y ácidos nucleicos no molde. Las moléculas molde de ácido nucleico se pueden obtener de cualquier material celular, obtenido de un animal, planta, bacteria, hongo o cualquier otro organismo celular. En determinados modos de realización, los polinucleótidos se obtienen de una sola célula. Los polinucleótidos se pueden obtener directamente de un organismo o de una muestra biológica obtenida de un organismo. Cualquier muestra de tejido o fluido corporal se puede usar como una fuente de ácido nucleico para su uso en la invención. Los polinucleótidos también se pueden aislar de células cultivadas, tales como un cultivo celular primario o una línea celular. Las células o tejidos a partir de los cuales se obtienen los ácidos nucleicos molde se pueden infectar con un virus u otro patógeno intracelular.

Los procedimientos de extracción de ADN son bien conocidos en la técnica. Un protocolo de aislamiento de ADN clásico se basa en la extracción usando disolventes orgánicos tales como una mezcla de fenol y cloroformo, seguido de precipitación con etanol (J. Sambrook y col., "Molecular Cloning: A Laboratory Manual," 1989, 2.^a ed., Cold Spring Harbour Laboratory Press: Nueva York, N.Y.). Otros procedimientos incluyen: extracción de ADN por precipitación salina (P. Sunnucks y col., Genetics, 1996, 144:747-756; S. M. Aljanabi e I. Martínez, Nucl. Acids Res. 1997, 25:4692-4693), extracción de ADN mediante sales de bromuro de trimetilamonio (S. Gustincich y col., BioTechniques, 1991, 11:298-302) y extracción de ADN mediante tiocianato de guanidinio (J. B. W. Hammond y col., Biochemistry, 1996, 240:298-300). Una variedad de kits está disponible comercialmente para extraer ADN de muestras biológicas (por ejemplo, BD Biosciences Clontech (Palo Alto, CA); Epicenter Technologies (Madison, WI); Genra Systems, Inc. (Minneapolis, MN); MicroProbe Corp. (Bothell, WA); Organon Teknika (Durham, NC); y Qiagen Inc. (Valencia, CA)).

Los procedimientos de extracción de ARN también son bien conocidos en la técnica (véase, por ejemplo, J. Sambrook y col., "Molecular Cloning: A Laboratory Manual" 1989, 211^a Ed., Cold Spring Harbour Laboratory Press: Nueva York) y varios kits para la extracción de ARN de fluidos corporales están disponibles comercialmente (por ejemplo, Ambion, Inc. (Austin, TX); Amersham Biosciences (Piscataway, NJ); BD Biosciences Clontech (Palo Alto, CA); BioRad Laboratories (Hercules, CA); Dynal Biotech Inc. (Lake Success, NY); Epicenter Technologies (Madison, WI); Genra Systems, Inc. (Minneapolis, MN); GIBCO BRL (Gaithersburg, MD); Invitrogen Life Technologies (Carlsbad, CA); MicroProbe Corp. (Bothell, WA); Organon Teknika (Durham, NC); Promega, Inc. (Madison, WI); y Qiagen Inc. (Valencia, CA)).

Una o más muestras pueden ser de una o más fuentes. Una o más de las muestras pueden ser de dos o más fuentes. Una o más de las muestras pueden ser de uno o más sujetos. Una o más de las muestras pueden ser de dos o más sujetos. Una o más de las muestras pueden ser del mismo sujeto. Uno o más sujetos pueden ser de la misma especie. Uno o más sujetos pueden ser de diferentes especies. Uno o más sujetos pueden estar sanos. Uno o más sujetos pueden padecer una enfermedad, trastorno o afección.

En algunos modos de realización, una muestra es un fluido, tal como sangre, saliva, linfa, orina, líquido cefalorraquídeo, líquido seminal, esputo, heces u homogeneizados de tejidos.

5 Se puede tomar una muestra de un sujeto con una afección. En algunos modos de realización, el sujeto del que se toma una muestra puede ser un paciente, por ejemplo, un paciente con cáncer o un paciente sospechoso de tener cáncer. El sujeto puede ser un mamífero, por ejemplo, un ser humano, y puede ser hombre o mujer. En algunos modos de realización, la mujer está embarazada. La muestra puede ser una biopsia tumoral. La biopsia puede ser realizada, por ejemplo, por un proveedor de atención médica, incluyendo un médico, asistente médico, enfermera, veterinario, dentista, quiropráctico, paramédico, dermatólogo, oncólogo, gastroenterólogo o cirujano.

En algunos modos de realización, la enfermedad o afección es una infección patógena. Los polinucleótidos diana pueden ser de un patógeno. El patógeno puede ser un virus, bacteria, hongo o protozoo. En algunos modos de realización, el patógeno puede ser un protozoo, tal como Acanthamoeba (por ejemplo, *A. astronyxis*, *A. castellanii*, *A. culbertsoni*, *A. hatchetti*, *A. polyphaga*, *A. rhysodes*, *A. healyi*, *A. divionensis*), Brachiola (por ejemplo, *B. connori*, *B. vesicularum*), Cryptosporidium (por ejemplo, *C. parvum*), Cyclospora (por ejemplo, *C. cayetanensis*), Encephalitozoon (por ejemplo, *E. cuniculi*, *E. hellem*, *E. intestinalis*), Entamoeba (por ejemplo, *E. histolytica*), Enterocytozoon (por ejemplo, *E. bieneusi*), Giardia (por ejemplo, *G. lamblia*), Isospora (por ejemplo, *I. belli*), Microsporidium (por ejemplo, *M. africanum*, *M. ceylonensis*), Naegleria (por ejemplo, *N. fowleri*), Nosema (por ejemplo, *N. algerae*, *N. ocularum*), Pleistophora, Trachipleistophora (por ejemplo, *T. anthropophthera*, *T. hominis*) y Vittaforma (por ejemplo, *V. corneae*). El patógeno puede ser un hongo, tal como Cándida, Aspergillus, Criptococo, Histoplasma, Pneumocystis y Stachybotrys. El patógeno puede ser una bacteria. Las bacterias ejemplares incluyen, pero no se limitan a, Bordetella, Borrelia, Brucella, Campylobacter, Clamidia, Chlamydothyla, Clostridium, Corynebacterium, Enterococo, Escherichia, Francisella, Haemophilus, Helicobacter, Legionella, Leptospira, Listeria, Micobacteria, micoplasma, Neisseria, Pseudomonas, Rickettsia, Salmonella, Shigella, Estafilococo, Estreptococo, Treponema, Vibrio o Yersinia. El virus puede ser un virus de transcripción inversa. Los ejemplos de virus de transcripción inversa incluyen, pero no se limitan a, virus de ARN-RT monocatenario (ARNmc-RT) y virus de ADN-RT bicatenario (ADNbc-RT). Ejemplos no limitantes de virus de ARNmc-RT incluyen retrovirus, alfarretrovirus, betarretrovirus, gammaretrovirus, deltarretrovirus, épsilonretrovirus, lentivirus, virus espumoso, metavirus y pseudovirus. Ejemplos no limitantes de virus de ADNbc-RT incluyen hepadenovirus y caulimovirus. El virus puede ser un virus de ADN. El virus puede ser un virus de ARN. El virus de ADN puede ser un virus de ADN bicatenario (ADNbc). En algunos modos de realización, el virus ADNbc es un adenovirus, virus del herpes o virus de la viruela. Los ejemplos de adenovirus incluyen, pero no se limitan a, adenovirus y virus de la hepatitis canina infecciosa. Los ejemplos de virus del herpes incluyen, pero no se limitan a, el virus del herpes simple, el virus de la varicela-zóster, el citomegalovirus y el virus de Epstein-Barr. Una lista no limitante de virus de la viruela incluye el virus de la viruela, el virus de la viruela de las vacas, el virus de la viruela de las ovejas, el virus de la viruela del mono y el virus de la variolovacuna. El virus de ADN puede ser un virus de ADN monocatenario (ADNmc). El virus de ADNmc puede ser un parvovirus. Los ejemplos de parvovirus incluyen, pero no se limitan a, parvovirus B19, parvovirus canino, parvovirus de ratón, parvovirus porcino, panleucopenia felina y virus de la enteritis de visón.

40 El virus puede ser un virus de ARN. El virus de ARN puede ser un virus de ARN bicatenario (ARNbc), un virus de ARN monocatenario de sentido (+) (ARNmc(+)), o un virus monocatenario de sentido (-) (ARNmc(-)). Una lista no limitante de virus de ARNmc incluye reovirus, ortorreovirus, cipovirus, rotavirus, virus de la lengua azul y fitorreovirus. Los ejemplos de virus de ARNmc (+) incluyen, pero no se limitan a, picornavirus y togavirus. Los ejemplos de picornavirus incluyen, pero no se limitan a, enterovirus, rinovirus, hepatovirus, cardiovirus, aftovirus, poliovirus, parechovirus, erbovirus, kobuvirus, teschovirus y coxsackie. En algunos modos de realización, el togavirus es un virus de la rubéola, el virus Sindbis, el virus de la encefalitis equina oriental, el virus de la encefalitis equina occidental, el virus de la encefalitis equina venezolana, el virus del río Ross, el virus O'nyong'nyong, Chikungunya o el virus del bosque Semliki. Una lista no limitante de virus de ARNmc (-) incluye ortomixovirus y rabdovirus. Los ejemplos de ortomixovirus incluyen, pero no se limitan a, virus de la gripe A, virus de la gripe B, virus de la gripe C, isavirus y thogotovirus. Los ejemplos de rabdovirus incluyen, pero no se limitan a, citorrabdovirus, dicorrabdovirus, efemerovirus, lissavirus, novirrabdovirus y vesiculovirus.

Una muestra puede ser una muestra biológica de cualquier organismo o virus. Las muestras para su uso en la presente invención incluyen partículas o preparaciones virales. En algunos modos de realización, el material de partida puede ser una muestra que contiene ácidos nucleicos, de cualquier organismo, a partir del cual se puede obtener material genético. Una o más de las muestras pueden ser de un mamífero, bacteria, virus, hongo o planta. Una o más muestras pueden ser de un ser humano, caballo, vaca, pollo, cerdo, rata, ratón, mono, conejo, cobaya, oveja, cabra, perro, gato, pájaro, pez, rana y mosca de la fruta.

En algunos modos de realización, los polinucleótidos están unidos a otras moléculas diana tales como proteínas, enzimas, sustratos, anticuerpos, agentes de unión, microesferas, moléculas pequeñas, péptidos o cualquier otra molécula. En general, el ácido nucleico se puede extraer de una muestra biológica mediante una variedad de técnicas (Sambrook y col., Molecular Cloning: A Laboratory Manual, 3.ª edición, Cold Spring Harbor, N.Y. (2001)).

En algunos modos de realización, la muestra es saliva. En algunos modos de realización, la muestra es sangre completa. En algunos modos de realización, para obtener una cantidad suficiente de polinucleótidos para la prueba, se extrae un volumen de sangre de al menos aproximadamente 0,001, 0,005, 0,01, 0,05, 0,1, 0,5, 1, 2, 3, 4, 5, 10, 20, 25, 30, 35, 40,

45 o 50 ml. En algunos modos de realización, la sangre se puede recoger en un aparato que contiene un quelante de magnesio que incluye, pero no se limita a EDTA, y se almacena a 4 °C. Opcionalmente, se puede añadir un quelante de calcio, que incluye, pero no se limita a EGTA.

5 En algunos modos de realización, se añade un inhibidor de lisis celular a la sangre que incluye, pero no se limita a formaldehído, derivados de formaldehído, formalina, glutaraldehído, derivados de glutaraldehído, un reticulante de proteínas, un reticulante de ácidos nucleicos, un reticulante de proteínas y ácidos nucleicos, reticulantes reactivos con amina primaria, reticulantes reactivos con sulfhidrilo, adición de sulfhidrilo o reducción de disulfuro, reticulantes reactivos con carbohidrato, reticulantes reactivos con carboxilo, reticulantes fotorreactivos o reticulantes escindibles. En algunos
10 modos de realización, los materiales no de ácido nucleico se pueden eliminar del material de partida usando tratamientos enzimáticos (tales como digestión con proteasas).

En algunos modos de realización, el material de partida puede ser una muestra de tejido que comprende un tejido, con ejemplos no limitantes que incluyen cerebro, hígado, pulmón, riñón, próstata, ovario, bazo, ganglio linfático (incluyendo la amígdala), tiroides, páncreas, corazón, músculo esquelético, intestino, laringe, esófago, estómago, hueso, corazón, timo, arteria, vaso sanguíneo, pulmón, músculo, estómago, intestino, hígado, páncreas, bazo, riñón, vesícula biliar, glándula tiroides, glándula suprarrenal, glándula mamaria, ovario, glándula prostática, testículo, piel, tejido adiposo, ojo o cerebro. En otros casos, el material de partida puede ser células que contienen ácidos nucleicos. El tejido puede comprender un tejido infectado, tejido enfermo, tejido maligno, tejido calcificado o tejido sano. Una muestra puede comprender al menos
15 una célula de uno o más tejidos biológicos. Por ejemplo, una muestra puede comprender una o más células malignas.

Las una o más células malignas pueden derivarse de un tumor, carcinoma, sarcoma o leucemia. Los sarcomas son cánceres del hueso, cartílago, grasa, músculo, vasos sanguíneos u otro tejido conectivo o de soporte. Los sarcomas incluyen, pero no se limitan a, cáncer de hueso, fibrosarcoma, condrosarcoma, sarcoma de Ewing, hemangioendotelio
25 sarcoma maligno, schwannoma maligno, schwannoma vestibular bilateral, osteosarcoma, sarcomas de partes blandas (por ejemplo, sarcoma de partes blandas alveolar, angiosarcoma, cistosarcoma filoide, dermatofibrosarcoma, tumor desmoide, sarcoma epiteloide, osteosarcoma extraesquelético, fibrosarcoma, hemangiopericitoma, hemangiosarcoma, sarcoma de Kaposi, leiomiomasarcoma, liposarcoma, linfangiosarcoma, linfosarcoma, histiocitoma fibroso maligno, neurofibrosarcoma, rabdomiosarcoma y sarcoma sinovial). Los carcinomas son cánceres que comienzan en las células epiteliales, que son células que cubren la superficie del cuerpo, producen hormonas y forman glándulas. A modo de ejemplo no limitante, los carcinomas incluyen cáncer de mama, cáncer de páncreas, cáncer de pulmón, cáncer de colon, cáncer colorrectal, cáncer de recto, cáncer de riñón, cáncer de vejiga, cáncer de estómago, cáncer de próstata, cáncer de hígado, cáncer de ovario, cáncer de cerebro, cáncer vaginal, cáncer vulvar, cáncer uterino, cáncer bucal, cáncer de pene, cáncer testicular, cáncer esofágico, cáncer de piel, cáncer de las trompas de Falopio, cáncer de cabeza y cuello, cáncer estromal gastrointestinal, adenocarcinoma, melanoma cutáneo o intraocular, cáncer de la región anal, cáncer del intestino delgado, cáncer del sistema endocrino, cáncer de la glándula tiroides, cáncer de la glándula paratiroides, cáncer de la glándula suprarrenal, cáncer de la uretra, cáncer de la pelvis renal, cáncer del uréter, cáncer del endometrio, cáncer de cuello uterino, cáncer de la glándula pituitaria, neoplasias del sistema nervioso central (SNC), linfoma del SNC primario, glioma del tronco encefálico y tumores del eje espinal. En algunos modos de realización, el cáncer es un cáncer de piel, tal como un carcinoma basocelular, carcinoma de células escamosas, melanoma, no melanoma o queratosis actínica (solar). En algunos modos de realización, el cáncer es un cáncer de pulmón. El cáncer de pulmón puede comenzar en las vías respiratorias que se ramifican desde la tráquea para abastecer a los pulmones (bronquios) o a los pequeños sacos de aire del pulmón (los alvéolos). Los cánceres de pulmón incluyen carcinoma de pulmón no microcítico (CPNM), carcinoma de pulmón microcítico y mesotelioma. Los ejemplos de CPNM incluyen carcinoma de células escamosas, adenocarcinoma y carcinoma macrocítico. El mesotelioma puede ser un tumor canceroso del revestimiento del pulmón y la cavidad torácica (pleura) o de la mucosa abdominal (peritoneo). El mesotelioma puede deberse a la exposición al asbesto. El cáncer puede ser un cáncer cerebral, tal como un glioblastoma. En algunos modos de realización, el cáncer puede ser un tumor del sistema nervioso central (SNC). Los tumores del SNC se pueden clasificar como gliomas o no gliomas. El glioma puede ser un glioma maligno, glioma de alto grado, glioma pontino intrínseco difuso. Los ejemplos de gliomas incluyen astrocitomas, oligodendrogliomas (o mezclas de oligodendroglioma y elementos de astrocitoma) y ependimomas. Los astrocitomas incluyen, pero no se limitan a, astrocitomas de bajo grado, astrocitomas anaplásicos, glioblastoma multiforme, astrocitoma pilocítico, xantastrocitoma pleomórfico y astrocitoma gigantocelular subependimario. Los oligodendrogliomas incluyen oligodendrogliomas de bajo grado (u oligoastrocitomas) y oligodendriogliomas anaplásicos. Los no gliomas incluyen meningiomas, adenomas pituitarios, linfomas del SNC primario y meduloblastomas. En algunos modos de realización, el cáncer es un meningioma. La leucemia puede ser una leucemia linfocítica aguda, leucemia mielocítica aguda, leucemia linfocítica crónica o leucemia mielocítica crónica. Los tipos adicionales de leucemias incluyen leucemia de células pilosas, leucemia mielomonocítica crónica y leucemia mielomonocítica juvenil. Los linfomas son cánceres de los linfocitos y se pueden desarrollar a partir de linfocitos B o T. Los dos tipos principales de linfoma son el linfoma de Hodgkin, previamente conocido como enfermedad de Hodgkin, y el linfoma no hodgkiniano. El linfoma de Hodgkin está marcado por la presencia de la célula de Reed-Sternberg. Los linfomas no hodgkinianos son todos linfomas que no son linfomas de Hodgkin. Los linfomas no hodgkinianos pueden ser linfomas indolentes y linfomas agresivos. Los linfomas no hodgkinianos incluyen, pero no se limitan a, linfoma difuso de linfocitos B grandes, linfoma folicular, linfoma de tejido linfático asociado a la mucosa (MALT), linfoma linfocítico de células pequeñas, linfoma de células del manto, linfoma de Burkitt, linfoma mediastínico de linfocitos B grandes, macroglobulinemia de Waldenström, linfoma ganglionar de la zona marginal de linfocitos B (NMZL), linfoma de la zona marginal esplénica (SMZL), linfoma extranodal de la zona marginal de linfocitos B, linfoma intravascular de linfocitos B
40
45
50
55
60
65

grandes, linfoma de efusión primaria y granulomatosis linfomatoide.

Una pluralidad de muestras puede comprender al menos 5, 10, 20, 30, 40, 50, 60, 70, 80, 90 o 100 o más muestras. La pluralidad de muestras puede comprender al menos aproximadamente 100, 200, 300, 400, 500, 600, 700, 800, 900 o 1000 o más muestras. La pluralidad de muestras puede comprender al menos aproximadamente 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000 muestras, 9000 o 10.000 muestras o 100.000 muestras o 1.000.000 o más muestras. La pluralidad de muestras puede comprender al menos aproximadamente 10.000 muestras.

Los uno o más polinucleótidos en una primera muestra pueden ser diferentes de uno o más polinucleótidos en una segunda muestra. Los uno o más polinucleótidos en una primera muestra pueden ser diferentes de uno o más polinucleótidos en una pluralidad de muestras. Uno o más polinucleótidos en una muestra pueden comprender al menos aproximadamente el 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % o 100 % de identidad de secuencia. En algunos modos de realización, uno o más polinucleótidos en una muestra pueden diferir en menos de aproximadamente 100, 90, 80, 70, 60, 50, 40, 30, 25, 20, 25, 10, 9, 8, 7, 6, 5, 4, 3, 2 o 1 nucleótido(s) o par(es) de bases. Una pluralidad de polinucleótidos en una o más muestras de la pluralidad de muestras puede comprender dos o más secuencias idénticas. Al menos aproximadamente el 1 %, 2 %, 3 %, 4 %, 5 %, 6 %, 7 %, 8 %, 9 %, 10 %, 15 %, 20 %, 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 97 % o 100 % de los polinucleótidos totales en una o más de la pluralidad de muestras pueden comprender la misma secuencia. Una pluralidad de polinucleótidos en una o más muestras de la pluralidad de muestras puede comprender al menos dos secuencias diferentes. Al menos aproximadamente el 5 %, 10 %, 15 %, 20 %, 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 %, 100 % de los polinucleótidos totales en una o más de la pluralidad de muestras pueden comprender al menos dos secuencias diferentes. En algunos modos de realización, uno o más polinucleótidos son variantes entre sí. Por ejemplo, uno o más polinucleótidos pueden contener polimorfismos de un solo nucleótido u otros tipos de mutaciones. En otro ejemplo, uno o más polinucleótidos son variantes de empalme.

Una primera muestra puede comprender una o más células y la segunda muestra puede comprender una o más células. Las una o más células de la primera muestra pueden ser del mismo tipo de célula que las una o más células de la segunda muestra. Las una o más células de la primera muestra pueden ser de un tipo de célula diferente como una o más células diferentes de la pluralidad de muestras.

La pluralidad de muestras se puede obtener simultáneamente. Se puede obtener una pluralidad de muestras al mismo tiempo. La pluralidad de muestras se puede obtener secuencialmente. Se puede obtener una pluralidad de muestras en el transcurso de años, 100 años, 10 años, 5 años, 4 años, 3 años, 2 años o 1 año después de obtener una o más muestras diferentes. Se pueden obtener una o más muestras dentro de aproximadamente un año después de obtener una o más muestras diferentes. Se pueden obtener una o más muestras dentro de los 12 meses, 11 meses, 10 meses, 9 meses, 8 meses, 7 meses, 6 meses, 4 meses, 3 meses, 2 meses o 1 mes después de obtener una o más muestras diferentes. Se pueden obtener una o más muestras dentro de los 30 días, 28 días, 26 días, 24 días, 21 días, 20 días, 18 días, 17 días, 16 días, 15 días, 14 días, 13 días, 12 días, 11 días, 10 días, 9 días, 8 días, 7 días, 6 días, 5 días, 4 días, 3 días, 2 días o un día después de obtener una o más muestras diferentes. Se pueden obtener una o más muestras en aproximadamente 24 horas, 22 horas, 20 horas, 18 horas, 16 horas, 14 horas, 12 horas, 10 horas, 8 horas, 6 horas, 4 horas, 2 horas o 1 hora después de obtener una o más muestras diferentes. Se pueden obtener una o más muestras dentro de aproximadamente 60 segundos, 45 segundos, 30 segundos, 20 segundos, 10 segundos, 5 segundos, 2 segundos o 1 segundo después de obtener una o más muestras diferentes. Se pueden obtener una o más muestras en menos de un segundo después de obtener una o más muestras diferentes.

Los diferentes polinucleótidos de una muestra pueden estar presentes en la muestra a diferentes concentraciones o cantidades. Por ejemplo, la concentración o cantidad de un polinucleótido puede ser mayor que la concentración o cantidad de otro polinucleótido en la muestra. En algunos modos de realización, la concentración o cantidad de al menos un polinucleótido en la muestra es al menos aproximadamente 1,5, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, o más veces mayor que la concentración o cantidad de al menos otro polinucleótido en la muestra. En otro ejemplo, la concentración o cantidad de un polinucleótido es menor que la concentración o cantidad de otro polinucleótido en la muestra. La concentración o cantidad de al menos un polinucleótido en la muestra puede ser al menos aproximadamente 1,5, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000 o más veces menor que la concentración o cantidad de al menos otro polinucleótido en la muestra.

En algunos modos de realización, dos o más muestras pueden contener diferentes cantidades o concentraciones de los polinucleótidos. En algunos modos de realización, la concentración o cantidad de un polinucleótido en una muestra puede ser mayor que la concentración o cantidad del mismo polinucleótido en una muestra diferente. Por ejemplo, una muestra de sangre puede contener una cantidad mayor de un polinucleótido particular que una muestra de orina. De forma alternativa, una sola muestra se puede dividir en dos o más submuestras. Las submuestras pueden contener diferentes cantidades o concentraciones del mismo polinucleótido. La concentración o cantidad de al menos un polinucleótido en una muestra puede ser al menos aproximadamente 1,5, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000 o más veces mayor que la concentración o cantidad del mismo polinucleótido en otra muestra. De forma alternativa, la concentración o cantidad de un polinucleótido en una

muestra puede ser menor que la concentración o cantidad del mismo polinucleótido en una muestra diferente. Por ejemplo, la concentración o cantidad de al menos un polinucleótido en una muestra puede ser al menos aproximadamente 1,5, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000 o más veces menor que la concentración o cantidad del mismo polinucleótido en otra muestra.

5

Muestras de sangre completa

En algunos modos de realización, la muestra es sangre completa. En algunos modos de realización, el porcentaje de amplicones, que contienen 10 o más UID, generados a partir de una muestra de sangre completa es igual al porcentaje de amplicones, que contienen 10 o más UID, generados a partir de una muestra de polinucleótidos purificados. En algunos modos de realización, el porcentaje de amplicones, que contienen 10 o más UID, generados a partir de una muestra de sangre completa es solo inferior a aproximadamente el 1 %, 2 %, 3 %, 4 %, 5 %, 6 %, 7 %, 8 %, 9 %, o como máximo el 10 % menos que el porcentaje de amplicones, que contienen 10 o más UID, generados a partir de una muestra de polinucleótidos purificados. En algunos modos de realización, la especificidad en la diana observada a partir de una muestra de sangre completa es igual a la especificidad en la diana observada a partir de una muestra de polinucleótido purificado. En algunos modos de realización, la especificidad en la diana observada a partir de una muestra de sangre completa es solo inferior a aproximadamente el 1 %, 2 %, 3 %, 4 %, 5 %, 6 %, 7 %, 8 %, 9 % o como máximo el 10 % menos que la especificidad en la diana observada a partir de una muestra de polinucleótido purificado. En algunos modos de realización, la uniformidad de cobertura observada a partir de una muestra de sangre completa es igual a la uniformidad de cobertura observada a partir de una muestra de polinucleótido purificado. En algunos modos de realización, la uniformidad de cobertura observada en una muestra de sangre completa es solo inferior a aproximadamente el 1 %, 2 %, 3 %, 4 %, 5 %, 6 %, 7 %, 8 %, 9 %, o como máximo el 10 % menos que la uniformidad de cobertura observada a partir de una muestra de polinucleótido purificado.

10

15

20

25

Muestras de FFPE

En algunos modos de realización, la muestra es una muestra incluida en parafina fijada con formol (FFPE). En algunos modos de realización, el porcentaje de amplicones, que contienen 10 o más UID, generados a partir de una muestra de FFPE es igual al porcentaje de amplicones, que contienen 10 o más UID, generados a partir de una muestra de polinucleótidos purificados. En algunos modos de realización, el porcentaje de amplicones, que contienen 10 o más UID, generados a partir de una muestra de FFPE es solo inferior a aproximadamente el 1 %, 2 %, 3 %, 4 %, 5 %, 6 %, 7 %, 8 %, 9 %, o como máximo el 10 % menos que el porcentaje de amplicones, que contienen 10 o más UID, generados a partir de una muestra de polinucleótidos purificados. En algunos modos de realización, la especificidad en la diana observada a partir de una muestra de FFPE es igual a la especificidad en la diana observada a partir de una muestra de polinucleótido purificado. En algunos modos de realización, la especificidad en la diana observada a partir de una muestra de FFPE es solo inferior a aproximadamente el 1 %, 2 %, 3 %, 4 %, 5 %, 6 %, 7 %, 8 %, 9 % o como máximo el 10 % menos que la especificidad en la diana observada a partir de una muestra de polinucleótido purificado. En algunos modos de realización, la uniformidad de cobertura observada a partir de una muestra de FFPE es igual a la uniformidad de cobertura observada a partir de una muestra de polinucleótido purificado. En algunos modos de realización, la uniformidad de cobertura observada en una muestra de FFPE es solo inferior a aproximadamente el 1 %, 2 %, 3 %, 4 %, 5 %, 6 %, 7 %, 8 %, 9 %, o como máximo el 10 % menos que la uniformidad de cobertura observada a partir de una muestra de polinucleótido purificado.

30

35

40

Bibliotecas

Las bibliotecas divulgadas en el presente documento se pueden usar en una variedad de aplicaciones. Como se usa en el presente documento, una biblioteca comprende una pluralidad de moléculas. En algunos modos de realización, una biblioteca comprende una pluralidad de polinucleótidos. En algunos modos de realización, una biblioteca comprende una pluralidad de cebadores. En algunos modos de realización, una biblioteca comprende una pluralidad de cebadores de RT. En algunos modos de realización, una biblioteca comprende una pluralidad de cebadores de PE. En algunos modos de realización, una biblioteca comprende una pluralidad de cebadores de extensión de cebador lineal (LPE). En algunos modos de realización, una biblioteca comprende una pluralidad de adaptadores. En algunos modos de realización, una biblioteca comprende una pluralidad de cebadores para amplificación no exponencial, tal como amplificación lineal. En algunos modos de realización, una biblioteca comprende una pluralidad de cebadores para amplificación exponencial, tal como PCR. En algunos modos de realización, una biblioteca comprende una pluralidad de polinucleótidos para secuenciación. Por ejemplo, la biblioteca se podría usar para secuenciar aplicaciones. En algunos modos de realización, una biblioteca comprende una pluralidad de lecturas de secuencia de uno o más polinucleótidos, amplicones o conjuntos de amplicones. Una biblioteca se puede almacenar y usar varias veces para generar muestras para su análisis. Algunas aplicaciones incluyen, por ejemplo, la genotipificación de polimorfismos, el estudio del procesamiento de ARN y la selección de representantes clonales para realizar la secuenciación de acuerdo con los procedimientos proporcionados en el presente documento. Se pueden generar bibliotecas que comprenden una pluralidad de polinucleótidos, tales como cebadores o bibliotecas para secuenciación o amplificación, en las que una pluralidad de polinucleótidos comprende al menos aproximadamente 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800 o 900 UID o polinucleótidos únicos. En algunos modos de realización, las bibliotecas de polinucleótidos comprenden una pluralidad de al menos aproximadamente 1000, 1500, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10.000, 11.000, 12.000, 13.000, 14.000, 15.000, 16.000, 17.000, 18.000, 19.000, 20.000, 30.000,

45

50

55

60

65

40.000, 50.000, 60.000, 70.000, 80.000, 90.000, 100.000, 200.000, 300.000, 400.000, 500.000, 600.000, 700.000, 800.000, 900.000, 1.000.000, 50.000.000, 100.000.000 o más polinucleótidos únicos, en las que cada polinucleótido único comprende una UID. En algunos modos de realización, las bibliotecas de polinucleótidos comprenden una pluralidad de conjuntos de amplicones, en las que cada conjunto de amplicones comprende una pluralidad de polinucleótidos con la misma UID. En algunos modos de realización, las bibliotecas de polinucleótidos comprenden una pluralidad de al menos aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 100 200, 300, 40, 500, 600, 700, 800, 900, 1000, 1500, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10.000, 11.000, 12.000, 13.000, 14.000, 15.000, 16.000, 17.000, 18.000, 19.000, 20.000, 30.000, 40.000, 50.000, 60.000, 70.000, 80.000, 90.000, 100.000 o más amplicones, en las que cada polinucleótido en los uno o más amplicones comprende una pluralidad de polinucleótidos con la misma UID. En algunos modos de realización, las bibliotecas de polinucleótidos comprenden una pluralidad de al menos aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1500, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10.000, 11.000, 12.000, 13.000, 14.000, 15.000, 16.000, 17.000, 18.000, 19.000, 20.000, 30.000, 40.000, 50.000, 60.000, 70.000, 80.000, 90.000, 100.000 o más conjuntos de amplicones, en las que cada conjunto de amplicones comprende una pluralidad de polinucleótidos o amplicones con la misma UID. En algunos modos de realización, las bibliotecas de polinucleótidos comprenden una pluralidad de al menos aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1500, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10.000, 11.000, 12.000, 13.000, 14.000, 15.000, 16.000, 17.000, 18.000, 19.000, 20.000, 30.000, 40.000, 50.000, 60.000, 70.000, 80.000, 90.000, 100.000 o más polinucleótidos, amplicones o conjuntos de amplicones, en las que cada polinucleótido, amplicón o conjunto de amplicones comprende una pluralidad de polinucleótidos, amplicones o conjuntos de amplicones con la misma secuencia molde o parte de la misma. En algunos modos de realización, las bibliotecas de polinucleótidos comprenden una pluralidad de al menos aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1500, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10.000, 11.000, 12.000, 13.000, 14.000, 15.000, 16.000, 17.000, 18.000, 19.000, 20.000, 30.000, 40.000, 50.000, 60.000, 70.000, 80.000, 90.000, 100.000 o más polinucleótidos, amplicones o conjuntos de amplicones, en las que cada polinucleótido, amplicón o conjunto de amplicones comprende una pluralidad de polinucleótidos, amplicones o conjuntos de amplicones con una secuencia molde o parte de la misma que difiere de uno o más de otros polinucleótidos, amplicones o conjuntos de amplicones por una o más bases causadas por un error o sesgo de amplificación o secuenciación.

30 Cebadores

La realización de las una o más reacciones de los procedimientos divulgados en el presente documento puede comprender el uso de uno o más cebadores. Como se usa en el presente documento, un **cebador** comprende un oligonucleótido bicatenario, monocatenario o parcialmente monocatenario que es suficientemente complementario para hibridarse con un polinucleótido molde. Un cebador puede ser un ADN monocatenario antes de unirse a un polinucleótido molde. En algunos modos de realización, el cebador comprende inicialmente una secuencia bicatenaria. Un sitio de cebador incluye el área del molde a la que se hibrida un cebador. En algunos modos de realización, los cebadores son capaces de actuar como un punto de inicio para la síntesis de ácido nucleico dirigida por molde. Por ejemplo, los cebadores pueden iniciar la síntesis de ácido nucleico dirigida por molde cuando cuatro nucleótidos diferentes y un agente de polimerización o enzima, tal como ADN o ARN polimerasa o retrotranscriptasa. Un par de cebadores incluye 2 cebadores: un primer cebador con una región anterior de 5' que se hibrida con un extremo 5' de una secuencia molde, y un segundo cebador con una región posterior de 3' que se hibrida con el complemento del extremo 3' de la secuencia molde. En algunos modos de realización, un cebador comprende una secuencia específica de diana y una secuencia de UID. En algunos modos de realización, un cebador comprende una secuencia de código de barras. En algunos modos de realización, un cebador comprende una secuencia de UID. En algunos modos de realización, un cebador comprende una secuencia de código de barras de muestra. En algunos modos de realización, un cebador comprende una secuencia de cebado universal. En algunos modos de realización, un cebador comprende una secuencia de cebado de PCR. En algunos modos de realización, un cebador comprende una secuencia de cebado de PCR usada para iniciar la amplificación de un polinucleótido. (Dieffenbach, PCR Primer: A Laboratory Manual, 2.^a edición (Cold Spring Harbor Press, Nueva York (2003)). El sitio o secuencia de unión del cebador universal permite la unión de un cebador universal a un polinucleótido y/o amplicón. Los cebadores universales son bien conocidos en la técnica e incluyen, pero no se limitan a, -47F (M13F), alfaMF, AOX3', AOX5', BGHr, CMV-30, CMV-50, CVMf, LACrmt, lambda gt10F, lambda gt 10R, lambda gt11F, lambda gt11R, M13 rev, M13Forward(-20), M13Reverse, male, p10SEQPpQE, pA-120, pet4, pGAP Forward, pGLRVpr3, pGLpr2R, pKLAC14, pQEFS, pQERS, pucU1, pucU2, reversA, seqIRESam, seqIRESzpet, seqori, seqPCR, seqpIRES-, seqpIRES+, seqpSecTag, seqpSecTag+, seqretro+PSI, SP6, T3-prom, T7-prom, y T7-terminInv. Como se usa en el presente documento, unión se puede referir tanto a interacciones covalentes como a interacciones no covalentes. La unión del cebador universal al sitio de cebador universal se puede usar para la amplificación, detección y/o secuenciación del polinucleótido y/o amplicón. El sitio de unión de cebador universal puede comprender al menos aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900 o 1000 nucleótidos o pares de bases. En otro ejemplo, el sitio de unión de cebador universal comprende al menos aproximadamente 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000, 5500, 6000, 6500, 7000, 7500, 8000, 8500, 9000, 9500 o 10000 nucleótidos o pares de bases. En algunos modos de realización, el sitio de unión de cebador universal comprende 1-10, 10-20, 10-30 o 10-100 nucleótidos o pares de bases. En algunos modos de realización, el sitio de unión de cebador universal comprende de aproximadamente 1-90, 1-80, 1-70, 1-60, 1-50, 1-40, 1-30, 1-20, 1-10, 2-90, 2-80, 2-70, 2-60, 2-50, 2-40, 2-30, 2-20, 2-10, 1-900, 1-800, 1-700, 1-600, 1-500, 1-400, 1-300, 1-200, 1-100, 2-900, 2-800, 2-700, 2-600, 2-500, 2-400, 2-300, 2-200, 2-100, 5-90, 5-80, 5-70, 5-60, 5-50, 5-40, 5-30, 5-20,

5-10, 10-90, 10-80, 10-70, 10-60, 10-50, 10-40, 10-30, 10-20, 10-10, 5-900, 5-800, 5-700, 5-600, 5-500, 5-400, 5-300, 5-200, 5-100, 10-900, 10-800, 10-700, 10-600, 10-500, 10-400, 10-300, 10-200, 10-100, 25-900, 25-800, 25-700, 25-600, 25-500, 25-400, 25-300, 25-200, 25-100, 100-1000, 100-900, 100-800, 100-700, 100-600, 100-500, 100-400, 100-300, 100-200, 200-1000, 200-900, 200-800, 200-700, 200-600, 200-500, 200-400, 200-300, 300-1000, 300-900, 300-800, 300-700, 300-600, 300-500, 300-400, 400-1000, 400-900, 400-800, 400-700, 400-600, 400-500, 500-1000, 500-900, 500-800, 500-700, 500-600, 600-1000, 600-900, 600-800, 600-700, 700-1000, 700-900, 700-800, 800-1000, 800-900 o 900-1000 nucleótidos o pares de bases.

Los cebadores pueden tener una longitud compatible con su uso en la síntesis de productos de extensión de cebador. Un cebador puede ser un polinucleótido que tiene de 8 a 200 nucleótidos de longitud. La longitud de un cebador puede depender de la secuencia del polinucleótido molde y el locus molde. Por ejemplo, se puede optimizar la longitud y/o temperatura de fusión (T_m) de un cebador o conjunto de cebadores. En algunos casos, un cebador puede ser de aproximadamente, de más de aproximadamente o de menos de aproximadamente 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59 o 60 nucleótidos de longitud. En algunos modos de realización, los cebadores tienen aproximadamente 8-100 nucleótidos de longitud, por ejemplo, 10-75, 15-60, 15-40, 18-30, 20-40, 21-50, 22-45, 25-40, 7-9, 12-15, 15-20, 15-25, 15-30, 15-45, 15-50, 15-55, 15-60, 20-25, 20-30, 20-35, 20-45, 20-50, 20-55 o 20-60 nucleótidos de longitud y cualquier longitud entre ellos. En algunos modos de realización, los cebadores tienen como máximo aproximadamente 10, 12, 15, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95 o 100 nucleótidos de longitud.

En general, se pueden usar uno o más pares de cebadores en una reacción de amplificación exponencial; un cebador de un par de cebadores puede ser un cebador directo y un cebador de un par de cebadores puede ser un cebador inverso. En algunos modos de realización, se puede usar un primer par de cebadores en la reacción de amplificación exponencial; un cebador del primer par puede ser un cebador directo complementario a una secuencia de una primera molécula de polinucleótido molde y un cebador del primer par puede ser un cebador inverso complementario a una segunda secuencia de la primera molécula de polinucleótido molde, y un primer locus molde puede residir entre la primera secuencia y la segunda secuencia. En algunos modos de realización, se puede usar un segundo par de cebadores en la reacción de amplificación; un cebador del segundo par puede ser un cebador directo complementario a una primera secuencia de una segunda molécula de polinucleótido diana y un cebador del segundo par puede ser un cebador inverso complementario a una segunda secuencia de la segunda molécula de polinucleótido diana, y un segundo locus diana puede residir entre la primera secuencia y la segunda secuencia. En algunos modos de realización, el segundo locus diana comprende una secuencia de anticuerpo de cadena ligera variable. En algunos modos de realización, se puede usar un tercer par de cebadores en la reacción de amplificación exponencial; un cebador del tercer par puede ser un cebador directo complementario a una secuencia de una tercera molécula de polinucleótido molde y un cebador del tercer par puede ser un cebador inverso complementario a una segunda secuencia de la tercera molécula de polinucleótido molde, y un tercer locus molde puede residir entre la primera secuencia y la segunda secuencia. En algunos modos de realización, un primer, segundo o tercer locus molde comprende un código de barras, tal como una UID.

Los uno o más cebadores se pueden hibridar con al menos una parte de una pluralidad de polinucleótidos molde. Los uno o más cebadores se pueden hibridar con el extremo 3' y/o el extremo 5' de la pluralidad de polinucleótidos molde. Los uno o más cebadores se pueden hibridar con una región interna de la pluralidad de polinucleótidos molde. La región interna puede ser de al menos aproximadamente 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 100, 150, 200, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 650, 700, 750, 800, 850, 900 o 1000 nucleótidos de los extremos 3' o de los extremos 5' de la pluralidad de polinucleótidos molde. Los uno o más cebadores pueden comprender un panel fijo de cebadores. Los uno o más cebadores pueden comprender al menos uno o más cebadores personalizados. Los uno o más cebadores pueden comprender al menos uno o más cebadores de control. Los uno o más cebadores pueden comprender un cebador universal. El cebador universal se puede hibridar con un sitio de unión de cebador universal. En algunos modos de realización, los uno o más cebadores personalizados no se hibridan con una UID. En algunos modos de realización, los uno o más cebadores personalizados se hibridan con un SBC, una región específica de diana, complementos de los mismos o cualquier combinación de los mismos. Los uno o más cebadores pueden comprender un cebador universal y un cebador que contiene UID. Los uno o más cebadores se pueden diseñar para amplificar o realizar extensión de cebadores, transcripción inversa, extensión lineal, amplificación no exponencial, amplificación exponencial, PCR o cualquier otro procedimiento de amplificación de uno o más polinucleótidos diana o molde.

La región específica de diana puede comprender al menos aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 100, 150, 200, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 650, 700, 750, 800, 850, 900 o 1000 nucleótidos o pares de bases. En otro ejemplo, la región específica de diana comprende al menos aproximadamente 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000, 5500, 6000, 6500, 7000, 7500, 8000, 8500, 9000, 9500 o 10000 nucleótidos o pares de bases. En algunos modos de realización, la región específica de diana comprende aproximadamente de 5-10, 10-15, 10-20, 10-30, 15-30, 10-75, 15-60, 15-40, 18-30, 20-40, 21-50, 22-45, 25-40, 7-9, 12-

15, 15-20, 15-25, 15-30, 15-45, 15-50, 15-55, 15-60, 20-25, 20-30, 20-35, 20-45, 20-50, 20-55, 20-60, 2-900, 2-800, 2-700, 2-600, 2-500, 2-400, 2-300, 2-200, 2-100, 25-900, 25-800, 25-700, 25-600, 25-500, 25-400, 25-300, 25-200, 25-100, 100-1000, 100-900, 100-800, 100-700, 100-600, 100-500, 100-400, 100-300, 100-200, 200-1000, 200-900, 200-800, 200-700, 200-600, 200-500, 200-400, 200-300, 300-1000, 300-900, 300-800, 300-700, 300-600, 300-500, 300-400, 400-1000, 400-900, 400-800, 400-700, 400-600, 400-500, 500-1000, 500-900, 500-800, 500-700, 500-600, 600-1000, 600-900, 600-800, 600-700, 700-1000, 700-900, 700-800, 800-1000, 800-900 o 900-1000 nucleótidos o pares de bases.

Los cebadores se pueden diseñar de acuerdo con parámetros conocidos para evitar estructuras secundarias y autohibridación. En algunos modos de realización, diferentes pares de cebadores se pueden hibridar y fundir a aproximadamente las mismas temperaturas, por ejemplo, dentro de 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10 °C de otro par de cebadores. En algunos modos de realización, uno o más cebadores en una pluralidad de cebadores se pueden hibridar y fundir a aproximadamente las mismas temperaturas, por ejemplo, dentro de 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10 °C de otro cebador en la pluralidad de cebadores. En algunos modos de realización, uno o más cebadores en una pluralidad de cebadores se pueden hibridar y fundir a diferentes temperaturas con respecto a otro cebador en la pluralidad de cebadores.

Una pluralidad de cebadores para una o más etapas de los procedimientos descritos en el presente documento puede comprender una pluralidad de cebadores que comprenden aproximadamente, como máximo aproximadamente, o al menos aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1500, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10.000, 11.000, 12.000, 13.000, 14.000, 15.000, 16.000, 17.000, 18.000, 19.000, 20.000, 30.000, 40.000, 50.000, 60.000, 70.000, 80.000, 90.000, 100.000, 200.000, 300.000, 400.000, 500.000, 600.000, 700.000, 800.000, 900.000, 1.000.000, 50.000.000, 100.000.000 cebadores diferentes. Por ejemplo, cada cebador en una pluralidad de cebadores puede comprender una UID. Por ejemplo, cada cebador en una pluralidad de cebadores puede comprender una región o secuencia diana o molde específica diferente. Por ejemplo, cada cebador en una pluralidad de cebadores puede comprender una UID diferente y una región o secuencia diana o molde específica diferente. Por ejemplo, cada cebador en una pluralidad de cebadores puede comprender una UID diferente y la misma región o secuencia diana o molde específica.

Paneles de cebadores

En algunos modos de realización, los paneles de cebadores usados para los procedimientos descritos en el presente documento comprenden o consisten en cebadores con un intervalo de temperatura de fusión de entre 60 °C y 68 °C. En algunos modos de realización, los paneles de cebadores usados para los procedimientos descritos en el presente documento comprenden o consisten en cebadores con una longitud de entre 21 y 32 nucleótidos. En algunos modos de realización, los paneles de cebadores usados para los procedimientos descritos en el presente documento comprenden o consisten en cebadores que no contienen 4 o más pirimidinas en los últimos 5 nucleótidos en el extremo 3'. En algunos modos de realización, los paneles de cebadores usados para los procedimientos descritos en el presente documento comprenden o consisten en cebadores diseñados para producir un amplicón que contiene entre el 30 % y el 70 % de contenido de GC. En algunos modos de realización, los paneles de cebadores usados para los procedimientos descritos en el presente documento comprenden o consisten en cebadores diseñados para producir amplicones con una longitud de entre 225 y 300 pares de bases. En algunos modos de realización, los paneles de cebadores usados para los procedimientos descritos en el presente documento comprenden o consisten en cebadores de un panel inicial que excluye los cebadores con el mayor número de lecturas erróneas (causadas por el cebado erróneo) durante la etapa de RT/PE inicial o la etapa de extensión/amplificación lineal. En algunos modos de realización, los paneles de cebadores usados para los procedimientos descritos en el presente documento comprenden o consisten en cebadores de un panel inicial que excluye cebadores prevalentes en dímeros. En algunos modos de realización, los paneles de cebadores usados para los procedimientos descritos en el presente documento comprenden o consisten en cebadores de un panel inicial que excluye cebadores que son responsables de generar una o más del mayor número de lecturas totales para una diana (sobreamplificadores). Se puede usar cualquiera o una combinación de las medidas anteriores para generar paneles de cebadores para su uso en los procedimientos descritos.

UID

En algunos modos de realización, los códigos de barras, tales como un SBC o UID, pueden tener una longitud dentro de un intervalo de 4 a 36 nucleótidos, o de 6 a 30 nucleótidos, o de 8 a 20 nucleótidos. En determinados aspectos, las temperaturas de fusión de los códigos de barras dentro de un conjunto están dentro de 10 °C uno del otro, dentro de 5 °C uno del otro, o dentro de 2 °C uno del otro. En otros aspectos, los códigos de barras son miembros de un conjunto de hibridación mínimamente cruzada. Por ejemplo, la secuencia de nucleótidos de cada miembro de dicho conjunto puede ser suficientemente diferente de la de cada otro miembro del conjunto para que ningún miembro pueda formar un dúplex estable con el complemento de cualquier otro miembro en condiciones de hibridación rigurosas. En algunos modos de realización, la secuencia de nucleótidos de cada miembro de un conjunto de hibridación mínimamente cruzada difiere de las de cualquier otro miembro en al menos dos nucleótidos. Las tecnologías de código de barras se describen en Winzeler y col. (1999) Science 285:901; Brenner (2000) Genome Biol. 1:1 Kumar y col. (2001) Nature Rev. 2:302; Giaever y col. (2004) Proc. Natl. Acad. Sci. Estados Unidos 101:793; Eason y col. (2004) Proc. Natl. Acad. Sci. Estados Unidos 101:11046; y Brenner (2004) Genome Biol. 5:240.

Como se usa en el presente documento, una marca de identificación única (UID) comprende información que es única

para una sola molécula, o dos o más moléculas de una pluralidad o biblioteca de moléculas. Un código de barras puede ser una UID. En algunos modos de realización, la información única comprende una secuencia única de nucleótidos. Por ejemplo, la secuencia de la UID se puede determinar determinando la identidad y el orden de la secuencia única o aleatoria de nucleótidos que comprende la UID. En algunos modos de realización, la información única no se puede usar para identificar la secuencia de un polinucleótido diana. En algunos modos de realización, la información única no es una secuencia conocida vinculada a la identidad de la secuencia de un polinucleótido diana. Por ejemplo, una UID puede estar unida a uno o más polinucleótidos diana, pero la UID no se puede usar para determinar a cuál de los uno o más polinucleótidos diana está unida. En algunos modos de realización, la información única comprende una secuencia aleatoria de nucleótidos. En algunos modos de realización, la información única comprende una o más secuencias únicas de nucleótidos en un polinucleótido. En algunos modos de realización, la información única comprende una secuencia de nucleótidos degenerada o un código de barras degenerado. Un código de barras degenerado puede comprender una composición o secuencia base de nucleótidos variable. Por ejemplo, un código de barras degenerado puede ser una secuencia aleatoria. En algunos modos de realización, una secuencia del complemento de una UID es también una secuencia de UID.

Una UID puede comprender cualquier longitud de nucleótidos. Por ejemplo, una UID puede comprender al menos aproximadamente 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 60, 70, 80, 90, 100, 200, 500 o 1000 nucleótidos. Por ejemplo, una UID puede comprender como máximo aproximadamente 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 60, 70, 80, 90, 100, 200, 500 o 1000 nucleótidos. En algunos modos de realización, una UID tiene una longitud particular de nucleótidos. Por ejemplo, una UID puede ser de aproximadamente 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 60, 70, 80, 90, 100, 200, 500 o 1000 nucleótidos de longitud.

En algunos modos de realización, cada UID en una pluralidad de UID tiene al menos aproximadamente 2 nucleótidos. Por ejemplo, cada UID en una pluralidad de UID puede ser de al menos aproximadamente 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 60, 70, 80, 90, 100, 200, 500 o 1000 nucleótidos de longitud. En algunos modos de realización, cada UID en una pluralidad de UID tiene como máximo aproximadamente 1000 nucleótidos. Por ejemplo, cada UID en una pluralidad de UID puede ser como máximo de aproximadamente 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 60, 70, 80, 90, 100, 200, 500 o 1000 nucleótidos de longitud. En algunos modos de realización, cada UID en una pluralidad de UID tiene la misma longitud de nucleótidos. Por ejemplo, cada UID en una pluralidad de UID puede ser de 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 60, 70, 80, 90, 100, 200, 500 o 1000 nucleótidos de longitud. En algunos modos de realización, una o más UID en una pluralidad de UID tienen una longitud de nucleótidos diferente. Por ejemplo, una o más primeras UID en una pluralidad de UID pueden tener aproximadamente, o al menos aproximadamente 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 60, 70, 80, 90, 100, 200, 500 o 1000 nucleótidos y una o más segundas UID en una pluralidad de UID pueden tener aproximadamente 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 60, 70, 80, 90, 100, 200, 500 o 1000 nucleótidos, en las que el número de nucleótidos de las una o más primeras UID es diferente de las una o más segundas UID.

El número de UID puede exceder el número de moléculas que se van a marcar. En algunos modos de realización, el número de UID es al menos aproximadamente 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90 o 100 veces mayor que el número de moléculas que se van a marcar.

El número de UID diferentes puede exceder el número de moléculas diferentes que se van a marcar. En algunos modos de realización, el número de UID diferentes es al menos aproximadamente 1, 1,5, 2, 2,5, 3, 3,5, 4, 4,5, 5, 6, 7, 8, 9, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90 o 100 veces mayor que el número de moléculas diferentes que se van a marcar.

En algunos modos de realización, al menos aproximadamente el 1 %, 2 %, 3 %, 4 %, 5 %, 6 %, 7 %, 8 %, 9 %, 10 %, 15 %, 20 %, 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 97 % o el 100 % de las UID diferentes tienen la misma concentración. En algunos modos de realización, al menos aproximadamente el 1 %, 2 %, 3 %, 4 %, 5 %, 6 %, 7 %, 8 %, 9 %, 10 %, 15 %, 20 %, 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 97 % o el 100% de las UID diferentes tienen una concentración diferente.

Las UID en una población de UID pueden tener al menos 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000 o más secuencias diferentes. Por ejemplo, las UID en una población pueden tener al menos 2.000, 3.000, 4.000, 5.000, 6.000, 7.000, 8.000, 9.000, 10.000, 15.000, 20.000, 25.000, 30.000, 35.000, 40.000, 45.000, 50.000, 60.000, 70.000, 80.000, 90.000, 100.000, 200.000, 300.000, 400.000, 500.000, 600.000, 700.000, 800.000, 900.000, 1.000.000 o más secuencias diferentes. Por tanto, se puede usar una pluralidad de UID para generar al menos 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000 o más secuencias diferentes a partir de uno o más polinucleótidos, tales como polinucleótidos diana. Por ejemplo, se puede usar

una pluralidad de UID para generar al menos 2.000, 3.000, 4.000, 5.000, 6.000, 7.000, 8.000, 9.000, 10.000, 15.000, 20.000, 25.000, 30.000, 35.000, 40.000, 45.000, 50.000, 60.000, 70.000, 80.000, 90.000, 100.000, 200.000, 300.000, 400.000, 500.000, 600.000, 700.000, 800.000, 900.000, 1×10^6 , 2×10^6 , 3×10^6 , 4×10^6 , 5×10^6 , 6×10^6 , 7×10^6 , 8×10^6 , 9×10^6 , 1×10^7 , 2×10^7 , 3×10^7 , 4×10^7 , 5×10^7 , 6×10^7 , 7×10^7 , 8×10^7 , 9×10^7 , 1×10^8 , 2×10^8 , 3×10^8 , 4×10^8 , 5×10^8 , 6×10^8 , 7×10^8 , 8×10^8 , 9×10^8 , 1×10^9 , 2×10^9 , 3×10^9 , 4×10^9 , 5×10^9 , 6×10^9 , 7×10^9 , 8×10^9 , 9×10^9 , 1×10^{10} , 2×10^{10} , 3×10^{10} , 4×10^{10} , 5×10^{10} , 6×10^{10} , 7×10^{10} , 8×10^{10} , 9×10^{10} , 1×10^{11} , 2×10^{11} , 3×10^{11} , 4×10^{11} , 5×10^{11} , 6×10^{11} , 7×10^{11} , 8×10^{11} , 9×10^{11} , 1×10^{12} , 2×10^{12} , 3×10^{12} , 4×10^{12} , 5×10^{12} , 6×10^{12} , 7×10^{12} , 8×10^{12} , 9×10^{12} o más secuencias diferentes a partir de uno o más polinucleótidos, tales como polinucleótidos diana. Por ejemplo, se puede usar una pluralidad de UID para generar al menos aproximadamente 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10.000, 15.000, 20.000, 25.000, 30.000, 35.000, 40.000, 45.000, 50.000, 60.000, 70.000, 80.000, 90.000, 100.000, 200.000, 300.000, 400.000, 500.000, 600.000, 700.000, 800.000, 900.000, 1×10^6 , 2×10^6 , 3×10^6 , 4×10^6 , 5×10^6 , 6×10^6 , 7×10^6 , 8×10^6 , 9×10^6 , 1×10^7 , 2×10^7 , 3×10^7 , 4×10^7 , 5×10^7 , 6×10^7 , 7×10^7 , 8×10^7 , 9×10^7 , 1×10^8 , 2×10^8 , 3×10^8 , 4×10^8 , 5×10^8 , 6×10^8 , 7×10^8 , 8×10^8 , 9×10^8 , 1×10^9 , 2×10^9 , 3×10^9 , 4×10^9 , 5×10^9 , 6×10^9 , 7×10^9 , 8×10^9 , 9×10^9 , 1×10^{10} , 2×10^{10} , 3×10^{10} , 4×10^{10} , 5×10^{10} , 6×10^{10} , 7×10^{10} , 8×10^{10} , 9×10^{10} , 1×10^{11} , 2×10^{11} , 3×10^{11} , 4×10^{11} , 5×10^{11} , 6×10^{11} , 7×10^{11} , 8×10^{11} , 9×10^{11} , 1×10^{12} , 2×10^{12} , 3×10^{12} , 4×10^{12} , 5×10^{12} , 6×10^{12} , 7×10^{12} , 8×10^{12} , 9×10^{12} o más secuencias diferentes a partir de al menos aproximadamente 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10.000, 15.000, 20.000, 25.000, 30.000, 35.000, 40.000, 45.000, 50.000, 60.000, 70.000, 80.000, 90.000, 100.000, 200.000, 300.000, 400.000, 500.000, 600.000, 700.000, 800.000, 900.000, 1×10^6 , 2×10^6 , 3×10^6 , 4×10^6 , 5×10^6 , 6×10^6 , 7×10^6 , 8×10^6 , 9×10^6 , 1×10^7 , 2×10^7 , 3×10^7 , 4×10^7 , 5×10^7 , 6×10^7 , 7×10^7 , 8×10^7 , 9×10^7 , 1×10^8 , 2×10^8 , 3×10^8 , 4×10^8 , 5×10^8 , 6×10^8 , 7×10^8 , 8×10^8 , 9×10^8 , 1×10^9 , 2×10^9 , 3×10^9 , 4×10^9 , 5×10^9 , 6×10^9 , 7×10^9 , 8×10^9 , 9×10^9 , 1×10^{10} , 2×10^{10} , 3×10^{10} , 4×10^{10} , 5×10^{10} , 6×10^{10} , 7×10^{10} , 8×10^{10} , 9×10^{10} , 1×10^{11} , 2×10^{11} , 3×10^{11} , 4×10^{11} , 5×10^{11} , 6×10^{11} , 7×10^{11} , 8×10^{11} , 9×10^{11} , 1×10^{12} , 2×10^{12} , 3×10^{12} , 4×10^{12} , 5×10^{12} , 6×10^{12} , 7×10^{12} , 8×10^{12} , 9×10^{12} o más polinucleótidos diana.

En algunos modos de realización, se usan una o más UID para agrupar o contener secuencias. En algunos modos de realización, se usan una o más UID para agrupar o contener secuencias, en las que las secuencias en cada contenedor contienen la misma UID. En algunos modos de realización, se usan una o más UID para agrupar o contener secuencias, en las que las secuencias en cada contenedor comprenden un conjunto de amplicones. En algunos modos de realización, se usan una o más UID para agrupar o contener secuencias, en las que las secuencias en cada contenedor comprenden una pluralidad de secuencias en las que los polinucleótidos a partir de los cuales se generaron la pluralidad de secuencias se derivaron del mismo polinucleótido en una reacción de amplificación. Por ejemplo, se pueden usar una o más UID para agrupar o contener secuencias en un amplicón o un conjunto de amplicones, o en ambos. En algunos modos de realización, una o más UID no se usan para alinear secuencias.

En algunos modos de realización, una o más UID no se usan para alinear secuencias. En algunos modos de realización, una o más UID no se usan para alinear secuencias y se usan para agrupar o contener secuencias. En algunos modos de realización, una o más UID no se usan para alinear secuencias y una región específica de diana se usa para alinear secuencias. En algunos modos de realización, se usan una o más UID para agrupar o contener secuencias y se usa una región específica de diana para alinear secuencias. En algunos modos de realización, una o más UID no se usan para alinear secuencias, una o más UID se usan para agrupar o contener secuencias, y una región específica de diana se usa para alinear secuencias.

En algunos modos de realización, se usan una o más UID para alinear secuencias. En algunos modos de realización, se usan una o más UID para alinear secuencias, en las que las secuencias alineadas contienen la misma UID. En algunos modos de realización, se usan una o más UID para alinear secuencias, en las que las secuencias alineadas comprenden dos o más secuencias de un conjunto de amplicones. En algunos modos de realización, se usan una o más UID para alinear secuencias, en las que las secuencias alineadas comprenden una pluralidad de secuencias en las que los polinucleótidos a partir de los cuales se generaron la pluralidad de secuencias se derivaron del mismo polinucleótido en una reacción de amplificación.

Enzimas

Los procedimientos y kits divulgados en el presente documento pueden comprender una o más enzimas. Los ejemplos de enzimas incluyen, pero no se limitan a, ligasas, retrotranscriptasas, polimerasas y nucleasas de restricción.

En algunos modos de realización, la unión de un adaptador a polinucleótidos comprende el uso de una o más ligasas. Los ejemplos de ligasas incluyen, pero no se limitan a, ADN ligasas tales como ADN ligasa I, ADN ligasa III, ADN ligasa IV y ADN ligasa T4, y ARN ligasas tales como ARN ligasa I T4 y ARN ligasa II T4.

Los procedimientos y kits divulgados en el presente documento pueden comprender además el uso de una o más retrotranscriptasas. En algunos modos de realización, la retrotranscriptasa es una retrotranscriptasa de VIH-1, retrotranscriptasa de M-MLV, retrotranscriptasa de AMV y retrotranscriptasa de telomerasa. En algunos modos de realización, la retrotranscriptasa es la retrotranscriptasa de M-MLV.

En algunos modos de realización, los procedimientos y kits divulgados en el presente documento comprenden el uso de

una o más polimerasas. Los ejemplos de polimerasas incluyen, pero no se limitan a, ADN polimerasas y ARN polimerasas. En algunos modos de realización, la ADN polimerasa es una ADN polimerasa I, ADN polimerasa II, ADN polimerasa III holoenzima y ADN polimerasa IV. Las ADN polimerasas disponibles comercialmente incluyen, pero no se limitan a, ADN polimerasa Bst 2.0, ADN polimerasa Bst 2.0 WarmStart™, ADN polimerasa Bst, ADN polimerasa Sulfolobus IV, ADN polimerasa Taq, ADN polimerasa 9°N™m, (exo-) ADN polimerasa Deep VentR™, ADN polimerasa Deep VentR™, Hemo KlenTaq™, ADN polimerasa Taq LongAmp®, ADN polimerasa OneTaq®, ADN polimerasa Phusion®, ADN polimerasa de alta fidelidad Q5™, ADN polimerasa Thermo Terminator™ γ, ADN polimerasa Thermo Terminator™, ADN polimerasa Thermo Terminator™ II, ADN polimerasa Thermo Terminator™ III, ADN polimerasa VentR®, (exo-) ADN polimerasa VentR®, ADN polimerasa Bsu, ADN polimerasa phi29, ADN polimerasa T4, ADN polimerasa T7, transferasa terminal, polimerasa Titanium® Taq, ADN polimerasa KAPA Taq y ADN polimerasa de arranque en caliente KAPA Taq.

En algunos modos de realización, la polimerasa es una ARN polimerasa tal como ARN polimerasa I, ARN polimerasa II, ARN polimerasa III, Poli(A) polimerasa de E. coli, ARN polimerasa phi6 (RdRP), Poli(U) polimerasa, ARN polimerasa SP6 y ARN polimerasa T7.

Reactivos adicionales

Los procedimientos y kits divulgados en el presente documento pueden comprender el uso de uno o más reactivos. Los ejemplos de reactivos incluyen, pero no se limitan a, reactivos de PCR, reactivos de ligadura, reactivos de transcripción inversa, reactivos enzimáticos, reactivos de hibridación, reactivos de preparación de muestras, reactivos de captura por afinidad, soportes sólidos tales como microesferas y reactivos para la purificación y/o aislamiento de ácidos nucleicos.

Un soporte sólido puede comprender prácticamente cualquier material insoluble o sólido, y a menudo se selecciona una composición de soporte sólido que es insoluble en agua. Por ejemplo, un soporte sólido puede comprender o consistir esencialmente en gel de sílice, vidrio (por ejemplo, vidrio de poro controlado (CPG)), nailon, Sephadex®, Sepharose®, celulosa, una superficie metálica (por ejemplo, acero, oro, plata, aluminio, silicio y cobre), un material magnético, un material plástico (por ejemplo, polietileno, polipropileno, poliamida, poliéster, difluoruro de polivinilideno (PVDF)) y similares. Los ejemplos de microesferas para su uso de acuerdo con los modos de realización pueden incluir un resto de afinidad que permite que la microesfera interactúe con una molécula de ácido nucleico. Una fase sólida (por ejemplo, una microesfera) puede comprender un miembro de un par de unión (por ejemplo, avidina, estreptavidina o un derivado de la misma). Por ejemplo, la microesfera puede ser una microesfera recubierta con estreptavidina y una molécula de ácido nucleico para la inmovilización en la microesfera puede incluir un resto de biotina. En algunos casos, cada molécula de polinucleótido puede incluir dos restos de afinidad, tales como biotina, para estabilizar aún más el polinucleótido. Las microesferas pueden incluir rasgos característicos adicionales para su uso en la inmovilización de ácidos nucleicos o que se pueden usar en procesos de cribado o selección posteriores. Por ejemplo, la microesfera puede incluir un resto de unión, un marcador fluorescente o un supresor fluorescente. En algunos casos, la microesfera puede ser magnética. En algunos casos, el soporte sólido es una microesfera. Los ejemplos de microesferas incluyen, pero no se limitan a, microesferas de estreptavidina, microesferas de agarosa, microesferas magnéticas, microesferas Dynabeads®, microesferas MACS®, microesferas conjugadas con anticuerpos (por ejemplo, microesfera antiinmunoglobulina), microesferas conjugadas con proteína A, microesferas conjugadas con proteína G, microesferas conjugadas con proteína A/G, microesferas conjugadas con proteína L, microesferas conjugadas oligo-dT, microesferas de sílice, microesferas similares a sílice, microesferas antibiotina, microesferas antil fluorocromo y microesferas magnéticas terminadas en carboxi BcMag™. Las microesferas o partículas pueden ser hinchables (por ejemplo, microesferas poliméricas tales como resina Wang) o no hinchables (por ejemplo, CPG). En algunos modos de realización, una fase sólida es sustancialmente hidrófila. En algunos modos de realización, una fase sólida (por ejemplo, una microesfera) es sustancialmente hidrófoba. En algunos modos de realización, una fase sólida comprende un miembro de un par de unión (por ejemplo, avidina, estreptavidina o derivado de la misma) y es sustancialmente hidrófoba o sustancialmente hidrófila. En algunos modos de realización, una fase sólida comprende un miembro de un par de unión (por ejemplo, avidina, estreptavidina o derivado de la misma) y tiene una capacidad de unión mayor de aproximadamente 1350 pmoles de agente de captura libre (por ejemplo, biotina libre) por mg de soporte sólido. En algunos modos de realización, la capacidad de unión de la fase sólida que comprende un miembro de un par de unión es mayor de 800, 900, 1000, 1100, 1200, 1250, 1300, 1350, 1400, 1450, 1500, 1600, 1800, 2000 pmoles de agente de captura libre por mg de soporte sólido. Otros ejemplos de microesferas que son adecuadas para la invención son coloides de oro o microesferas tales como microesferas de poliestireno o microesferas de sílice. Sustancialmente se puede usar cualquier radio de microesfera. Los ejemplos de microesferas pueden incluir microesferas que tienen un radio que varía de 150 nanómetros a 10 micras. También se pueden usar otros tamaños.

Los procedimientos y kits divulgados en el presente documento pueden comprender el uso de uno o más tampones. Los ejemplos de tampones incluyen, pero no se limitan a, tampones de lavado, tampones de ligadura, tampones de hibridación, tampones de amplificación y tampones de transcripción inversa. En algunos modos de realización, el tampón de hibridación es un tampón disponible comercialmente, tal como solución TMAC Hyb, solución de hibridación SSPE y tampón de hibridación ECONOTM. Los tampones divulgados en el presente documento pueden comprender uno o más detergentes.

Los procedimientos y kits divulgados en el presente documento pueden comprender el uso de uno o más vehículos. Los vehículos pueden potenciar o mejorar la eficacia de una o más reacciones divulgadas en el presente documento (por

ejemplo, reacción de ligadura, transcripción inversa, amplificación, hibridación). Los vehículos pueden disminuir o prevenir la pérdida no específica de las moléculas o cualquier producto de las mismas (por ejemplo, un polinucleótido y/o amplicón). Por ejemplo, el vehículo puede disminuir la pérdida no específica de un polinucleótido a través de la absorción en las superficies. El vehículo puede disminuir la afinidad de un polinucleótido por una superficie o sustrato (por ejemplo, recipiente, tubo Eppendorf, punta de pipeta). De forma alternativa, el vehículo puede aumentar la afinidad de un polinucleótido por una superficie o sustrato (por ejemplo, microesfera, matriz, vidrio, portaobjetos o chip). Los vehículos pueden proteger el polinucleótido de la degradación. Por ejemplo, los vehículos pueden proteger una molécula de ARN de las ribonucleasas. De forma alternativa, los vehículos pueden proteger una molécula de ADN de una DNasa. Los ejemplos de vehículos incluyen, pero no se limitan a, polinucleótidos tales como ADN y/o ARN o polipéptidos. Los ejemplos de vehículos de ADN incluyen plásmidos, vectores, ADN poliadenilado y oligonucleótidos de ADN. Los ejemplos de vehículos de ARN incluyen ARN poliadenilado, ARN de fago, ARN de fago MS2, ARN de *E. coli*, ARN de levadura, ARNt de levadura, ARN de mamífero, ARNt de mamífero, ribonucleótidos sintéticos poliadenilados cortos y oligonucleótidos de ARN. El vehículo de ARN puede ser un ARN poliadenilado. De forma alternativa, el vehículo de ARN puede ser un ARN no poliadenilado. En algunos modos de realización, el vehículo es de una bacteria, levadura o virus. Por ejemplo, el vehículo puede ser un polinucleótido o un polipéptido derivado de una bacteria, levadura o virus. Por ejemplo, el vehículo es una proteína de *Bacillus subtilis*. En otro ejemplo, el vehículo es un polinucleótido de *Escherichia coli*. De forma alternativa, el vehículo es un polinucleótido o péptido de un mamífero (por ejemplo, ser humano, ratón, cabra, rata, vaca, oveja, cerdo, perro o conejo), aviar, anfibio o reptil.

Los procedimientos y kits divulgados en el presente documento pueden comprender el uso de uno o más agentes de control. Los agentes de control pueden incluir polinucleótidos de control, enzimas inactivas y competidores no específicos. De forma alternativa, los agentes de control comprenden hibridación brillante, controles de sonda brillante, moldes de ácido nucleico, controles de adición, controles de amplificación por PCR. Los controles de amplificación por PCR pueden ser controles positivos. En otros casos, los controles de amplificación por PCR son controles negativos. Los controles molde de ácido nucleico pueden ser de concentraciones conocidas. Los agentes de control pueden comprender uno o más marcadores.

Los controles de adición pueden ser moldes que se añaden a una reacción o muestra. Por ejemplo, se puede añadir un molde de adición a una reacción de amplificación. El molde de adición se puede añadir a la reacción de amplificación en cualquier momento después del primer ciclo de amplificación. En algunos modos de realización, el molde de adición se añade a una reacción de amplificación después del número de ciclo 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25, 30, 35, 40, 45 o 50. El molde de adición se puede añadir a la reacción de amplificación en cualquier momento antes del último ciclo de amplificación. El molde de adición puede comprender uno o más nucleótidos o pares de bases de ácido nucleico. El molde de adición puede comprender ADN, ARN o cualquier combinación de los mismos. El molde de adición puede comprender uno o más marcadores.

Aspectos implementados por ordenador

Como entienden los expertos en la técnica, los procedimientos y la información descritos en el presente documento se pueden implementar, en su totalidad o en parte, como instrucciones ejecutables por ordenador en medios legibles por ordenador conocidos. Por ejemplo, los procedimientos descritos en el presente documento se pueden implementar en hardware. De forma alternativa, los procedimientos se pueden implementar en software almacenado en, por ejemplo, una o más memorias u otro medio legible por ordenador e implementar en uno o más procesadores. Como es conocido, los procesadores se pueden asociar con uno o más controladores, unidades de cálculo y/u otras unidades de un sistema informático, o implantar en firmware según se desee. Si se implementan en software, las rutinas se pueden almacenar en cualquier memoria legible por ordenador, tal como en RAM, ROM, memoria flash, un disco magnético, un disco láser u otro medio de almacenamiento, como también se conoce. Asimismo, este software se puede suministrar a un dispositivo informático por medio de cualquier procedimiento de suministro conocido, incluyendo, por ejemplo, a través de un canal de comunicación tal como una línea telefónica, Internet, una conexión inalámbrica, etc., o por medio de un medio transportable, tal como un disco legible por ordenador, unidad flash, etc.

De forma más general, y como entienden los expertos en la técnica, las diversas etapas descritas anteriormente se pueden implementar como diversos bloques, operaciones, herramientas, módulos y técnicas que, a su vez, se pueden implementar en hardware, firmware, software o cualquier combinación de hardware, firmware y/o software. Cuando se implementan en hardware, algunos o todos los bloques, operaciones, técnicas, etc., se pueden implementar en, por ejemplo, un circuito integrado personalizado (IC), un circuito integrado específico de la aplicación (ASIC), una matriz lógica programable in situ (FPGA), una matriz lógica programable (PLA), etc.

Los resultados de la secuenciación de datos se pueden almacenar en una unidad de almacenamiento de datos, tal como un soporte de datos, incluyendo bases de datos informáticas, discos de almacenamiento de datos u otros medios convenientes de almacenamiento de datos. En determinados modos de realización, la base de datos informática es una base de datos de objetos, una base de datos relacional o una base de datos post-relacional. Los datos se pueden recuperar de la unidad de almacenamiento de datos usando cualquier procedimiento conveniente de consulta de datos.

Cuando se implementa en software, el software se puede almacenar en cualquier medio legible por ordenador conocido, tal como un disco magnético, un disco óptico u otro medio de almacenamiento, en una RAM o ROM o memoria flash de

un ordenador, procesador, unidad de disco duro, unidad de disco óptico, unidad de cinta, etc. Asimismo, el software se puede suministrar a un usuario o un sistema informático por medio de cualquier procedimiento de suministro conocido, incluyendo, por ejemplo, en un disco legible por ordenador u otro mecanismo de almacenamiento informático transportable.

5 Las etapas de los procedimientos reivindicados pueden ser operativas con muchos otros entornos o configuraciones de sistemas informáticos de uso general o de uso especial. Los ejemplos de sistemas informáticos, entornos y/o configuraciones bien conocidos que pueden ser adecuados para su uso con los procedimientos o sistemas de las reivindicaciones incluyen, pero no se limitan a, ordenadores personales, servidores, dispositivos portátiles o de mano, sistemas multiprocesadores, sistemas basados en microprocesador, decodificadores, electrónica de consumo
10 programable, ordenadores en red, miniordenadores, ordenadores centrales, entornos informáticos distribuidos que incluyen cualquiera de los sistemas o dispositivos anteriores, y similares.

Las etapas de los procedimientos reivindicados se pueden describir en el contexto general de las instrucciones ejecutables por ordenador, tales como módulos de programa, que son ejecutadas por un ordenador. En general, los módulos de programa incluyen rutinas, programas, objetos, componentes y/o estructuras de datos que realizan tareas particulares o implementan tipos de datos abstractos particulares. Los procedimientos también se pueden poner en práctica en entornos informáticos distribuidos donde las tareas son realizadas por dispositivos de procesamiento remoto que están vinculados a través de una red de comunicaciones. En entornos informáticos integrados y distribuidos, se pueden localizar módulos de programa en medios de almacenamiento informáticos tanto locales como remotos incluyendo los dispositivos de
15 almacenamiento de memoria. Se podrían implementar numerosos modos de realización alternativos, usando la tecnología actual o la tecnología desarrollada después de la fecha de presentación de esta solicitud, que todavía estaría dentro del alcance de las reivindicaciones que definen la divulgación.

Si bien los procedimientos y otros elementos se han descrito como implementados preferentemente en software, se pueden implementar en hardware, firmware, etc., y pueden ser implementados por cualquier otro procesador. Por tanto, los elementos descritos en el presente documento se pueden implementar en una CPU multiuso estándar o en un hardware o firmware específicamente diseñado, tal como un circuito integrado específico de la aplicación (ASIC) u otro dispositivo de cableado permanente según se desee. Cuando se implementa en software, la rutina de software se puede almacenar en cualquier memoria legible por ordenador, tal como en un disco magnético, un disco láser u otro medio de almacenamiento, en una RAM o ROM de un ordenador o procesador, en cualquier base de datos, etc. Asimismo, este software se puede suministrar a un usuario o a un sistema de cribado por medio de cualquier procedimiento de suministro conocido o deseado, incluyendo, por ejemplo, en un disco legible por ordenador u otro mecanismo de almacenamiento informático transportable o a través de un canal de comunicación, por ejemplo, una línea telefónica, Internet o comunicación inalámbrica. Se pueden hacer modificaciones y variaciones en las técnicas y estructuras descritas e
25 ilustradas en el presente documento sin apartarse del espíritu y el alcance de la presente divulgación.

La **FIG. 58** es un diagrama de bloques que ilustra una primera arquitectura ejemplar de un sistema informático **100** que se puede usar en relación con modos de realización ejemplares de la presente invención. Como se representa en la **FIG. 58**, el sistema informático ejemplar puede incluir un procesador **102** para procesar instrucciones. Los ejemplos no limitantes de procesadores incluyen: un procesador Intel Xeon™, procesador AMD Opteron™, procesador Samsung 32-bit RISC ARM1176JZ(F)-S v1.0™, procesador ARM Cortex-A8 Samsung S5PC100™, procesador ARM Cortex-A8 Apple A4™, procesador Marvell PXA930™ o un procesador funcionalmente equivalente. Se pueden usar múltiples hilos de ejecución para el procesamiento paralelo. En algunos modos de realización, también se pueden usar múltiples procesadores o procesadores con múltiples núcleos, ya sea en un solo sistema informático, en un grupo o distribuidos a través de sistemas a través de una red que comprende una pluralidad de ordenadores, teléfonos móviles y/o dispositivos asistentes de datos personales.
40

Como se ilustra en la **FIG. 59**, una memoria caché de alta velocidad **104** se puede conectar a, o incorporar en, el procesador **102** para proporcionar una memoria de alta velocidad para instrucciones o datos que han sido usados recientemente, o que son usados con frecuencia, por el procesador **102**. El procesador **102** está conectado a un puente norte **106** mediante un bus de procesador **108**. El puente norte **106** está conectado a la memoria de acceso aleatorio (RAM) **110** mediante un bus de memoria **112** y gestiona el acceso a la RAM **110** mediante el procesador **102**. El puente norte **106** también está conectado a un puente sur **114** mediante un bus de un conjunto de chips **116**. El puente sur **114** está, a su vez, conectado a un bus periférico **118**. El bus periférico puede ser, por ejemplo, PCI, PCI-X, PCI Express u otro bus periférico. El puente norte y el puente sur a menudo se denominan como conjunto de chips de procesador y gestionan la transferencia de datos entre el procesador, la RAM y los componentes periféricos en el bus periférico **118**. En algunas arquitecturas alternativas, la funcionalidad del puente norte se puede incorporar al procesador en lugar de usar un chip de puente norte separado.
50

En algunos modos de realización, el sistema **100** puede incluir una tarjeta aceleradora **122** unida al bus periférico **118**. El acelerador puede incluir matrices de compuerta programable in situ (FPGA) u otro hardware para acelerar determinado procesamiento. Por ejemplo, se puede usar un acelerador para la reestructuración adaptativa de datos o para evaluar expresiones algebraicas usadas en el procesamiento de conjuntos extendidos.
55

El software y los datos se almacenan en el almacenamiento externo **124** y se pueden cargar en la RAM **110** y/o la caché **104** para su uso por el procesador. El sistema **100** incluye un sistema operativo para gestionar recursos del sistema; los
60

ejemplos no limitantes de sistemas operativos incluyen: Linux, Windows™, MACOS™, BlackBerry OS™, iOS™ y otros sistemas operativos funcionalmente equivalentes, así como software de aplicación que se ejecuta en la parte superior del sistema operativo para gestionar el almacenamiento y la optimización de datos de acuerdo con modos de realización ejemplares de la presente invención.

5

En este ejemplo, el sistema **100** también incluye tarjetas de interfaz de red (NIC) **120** y **121** conectadas al bus periférico para proporcionar interfaces de red para almacenamiento externo, tal como almacenamiento conectado a la red (NAS) y otros sistemas informáticos que se pueden usar para procesamiento paralelo distribuido.

10

La **FIG. 59** es un diagrama que muestra una red **200** con una pluralidad de sistemas informáticos **202a** y **202b**, una pluralidad de teléfonos móviles y asistentes de datos personales **202c**, y almacenamiento conectado a la red (NAS) **204a** y **204b**. En modos de realización ejemplares, los sistemas **202a**, **202b** y **202c** pueden gestionar el almacenamiento de datos y optimizar el acceso de datos a los datos almacenados en el almacenamiento conectado a la red (NAS) **204a** y **204b**. Se puede usar un modelo matemático para los datos y evaluarlo usando el procesamiento paralelo distribuido a través de los sistemas informáticos **202a** y **202b**, y los sistemas de teléfono móvil y asistente de datos personales **202c**. Los sistemas informáticos **202a** y **202b**, y los sistemas de teléfono móvil y asistente de datos personales **202c** también pueden proporcionar procesamiento paralelo para la reestructuración adaptativa de los datos almacenados en el almacenamiento conectado a la red (NAS) **204a** y **204b**. La **FIG. 59** ilustra solo un ejemplo, y se puede usar una amplia variedad de otras arquitecturas y sistemas informáticos junto con los diversos modos de realización de la presente invención. Por ejemplo, se puede usar un servidor de tarjeta única para proporcionar procesamiento en paralelo. Las tarjetas únicas de procesador se pueden conectar a través de un plano posterior para proporcionar procesamiento en paralelo. El almacenamiento también se puede conectar al plano posterior o como almacenamiento conectado a la red (NAS) a través de una interfaz de red separada.

15

20

25

En algunos modos de realización ejemplares, los procesadores pueden mantener espacios de memoria separados y transmitir datos a través de interfaces de red, plano posterior u otros conectores para procesamiento en paralelo por otros procesadores. En otros modos de realización, algunos o todos los procesadores pueden usar un espacio de memoria de dirección virtual compartido.

30

La **FIG. 60** es un diagrama de bloques de un sistema informático multiprocesador **300** que usa un espacio de memoria de dirección virtual compartido de acuerdo con un modo de realización ejemplar. El sistema incluye una pluralidad de procesadores **302a-f** que pueden acceder a un subsistema de memoria compartida **304**. El sistema incorpora una pluralidad de procesadores de algoritmos de memoria de hardware programables (MAP) **306a-f** en el subsistema de memoria **304**. Cada MAP **306a-f** puede comprender una memoria **308a-f** y una o más matrices de compuertas programables in situ (FPGA) **310a-f**. El MAP proporciona una unidad funcional configurable y se pueden proporcionar algoritmos o partes de algoritmos particulares a las FPGA **310a-f** para su procesamiento en estrecha coordinación con un procesador respectivo. Por ejemplo, los MAP se pueden usar para evaluar expresiones algebraicas con respecto al modelo de datos y para realizar una reestructuración de datos adaptativa en modos de realización ejemplares. En este ejemplo, todos los procesadores tienen acceso global a cada MAP para estos propósitos. En una configuración, cada MAP puede usar el acceso directo a la memoria (DMA) para acceder a una memoria asociada **308a-f**, lo que le permite ejecutar tareas independientemente de, y de forma asíncrona desde, el microprocesador respectivo **302a-f**. En esta configuración, un MAP puede enviar resultados directamente a otro MAP para la canalización y la ejecución en paralelo de algoritmos.

35

40

45

Las arquitecturas y sistemas informáticos anteriores son solo ejemplos, y se puede usar una amplia variedad de otras arquitecturas y sistemas informáticos de ordenador, teléfono móvil y asistente de datos personales en relación con modos de realización ejemplares, incluyendo los sistemas que usan cualquier combinación de procesadores generales, coprocesadores, FPGA y otros dispositivos lógicos programables, sistema en chips (SOC), circuitos integrados específicos de la aplicación (ASIC) y otros elementos lógicos y de procesamiento. En algunos modos de realización, todo o parte del sistema de gestión y optimización de datos se puede implementar en software o hardware y se puede usar cualquier variedad de medios de almacenamiento de datos en relación con modos de realización ejemplares, incluyendo memoria de acceso aleatorio, discos duros, memoria flash, unidades de cinta, matrices de discos, almacenamiento conectado a la red (NAS) y otros dispositivos y sistemas de almacenamiento de datos locales o distribuidos.

50

55

En modos de realización ejemplares, el sistema de gestión y optimización de datos se puede implementar usando módulos de software que se ejecutan en cualquiera de las arquitecturas y sistemas informáticos anteriores u otros. En otros modos de realización, las funciones del sistema se pueden implementar parcial o completamente en firmware, dispositivos lógicos programables tales como matrices de compuertas programables in situ (FPGA), sistema en chips (SOC), circuitos integrados específicos de la aplicación (ASIC) u otros elementos lógicos y de procesamiento. Por ejemplo, el conjunto de procesador y optimizador se puede implementar con aceleración de hardware a través del uso de una tarjeta aceleradora de hardware, tal como una tarjeta aceleradora.

60

Un experto en la técnica apreciará que, aunque solo uno de cada uno de los componentes identificados anteriormente se representa en las figuras, se puede proporcionar cualquier número de cualquiera de estos componentes. Además, un experto en la técnica reconocerá que uno o más componentes de cualquiera de los sistemas divulgados se pueden combinar o incorporar en otro componente mostrado en las figuras. Uno o más de los componentes representados en las

65

figuras se pueden implementar en software en uno o más sistemas informáticos. Por ejemplo, pueden comprender una o más aplicaciones, que pueden comprender una o más unidades informáticas de instrucciones legibles por ordenador que, cuando son ejecutadas por un procesador, hacen que un ordenador realice las etapas de un procedimiento. Las instrucciones legibles por ordenador se pueden almacenar en un medio legible por ordenador, tal como una memoria o un disco. Dichos medios típicamente proporcionan almacenamiento no transitorio. De forma alternativa, uno o más de los componentes representados en las figuras pueden ser componentes de hardware o combinaciones de hardware y software tales como, por ejemplo, ordenadores de uso especial u ordenadores de uso general. Un ordenador o sistema informático también puede comprender una base de datos interna o externa. Los componentes de un ordenador o sistema informático se pueden conectar a través de una interfaz de bus local.

Un experto en la técnica apreciará que las etapas descritas anteriormente se pueden incorporar en distintos módulos de software. Aunque los componentes divulgados se han descrito anteriormente como unidades separadas, un experto en la técnica reconocerá que las funcionalidades proporcionadas por una o más unidades se pueden combinar. Como apreciará un experto en la técnica, una o más de las unidades pueden ser opcionales y se pueden omitir de las implementaciones en determinados modos de realización.

Kits

Los kits útiles en los procedimientos de la divulgación comprenden componentes útiles en cualquiera de los procedimientos descritos en el presente documento, que incluyen, por ejemplo, cebadores para amplificación de ácido nucleico, sondas de hibridación para detectar variación genética u otra detección de marcadores, enzimas de restricción, sondas de ácido nucleico, opcionalmente marcados con marcadores adecuados, oligonucleótidos específicos de alelo, anticuerpos que se unen a un polipéptido alterado codificado por un ácido nucleico de la divulgación como se describe en el presente documento o a un polipéptido natural codificado por un ácido nucleico de la divulgación como se describe en el presente documento, medios para la amplificación de variaciones genéticas o fragmentos de las mismas, medios para analizar la secuencia de ácido nucleico de ácidos nucleicos que comprenden variaciones genéticas como se describe en el presente documento, medios para analizar la secuencia de aminoácidos de un polipéptido codificado por una variación genética, o un ácido nucleico asociado con una variación genética, etc. Los kits pueden, por ejemplo, incluir tampones necesarios, cebadores de ácido nucleico para amplificar ácidos nucleicos, soportes sólidos y reactivos para la detección específica de alelo de los fragmentos amplificados usando dichos cebadores y enzimas necesarias (por ejemplo, ADN polimerasa), tales como cualquiera de los descritos en el presente documento. Adicionalmente, los kits pueden proporcionar reactivos para ensayos que se usarán en combinación con los procedimientos de la presente divulgación, por ejemplo, reactivos para su uso con otros ensayos de cribado para una enfermedad o afección.

En algunos modos de realización, la divulgación se refiere a un kit para analizar una muestra de ácido nucleico de un sujeto para detectar la presencia de una variación genética, en el que el kit comprende reactivos necesarios para detectar selectivamente al menos una variación genética particular en el genoma del individuo. En algunos modos de realización, la divulgación se refiere a un kit para analizar una muestra de ácido nucleico de un sujeto para detectar la presencia de al menos un alelo particular de al menos un polimorfismo asociado con una variación genética en el genoma del sujeto. En algunos modos de realización, los reactivos comprenden al menos un oligonucleótido contiguo que se hibrida con un fragmento del genoma del individuo que comprende al menos variación genética. En algunos modos de realización, los reactivos comprenden al menos un par de oligonucleótidos que se hibridan con cadenas opuestas de un segmento genómico obtenido de un sujeto, en el que cada par de cebadores de oligonucleótidos está diseñado para amplificar selectivamente un fragmento del genoma del individuo que incluye al menos una variación genética o un fragmento de una variación genética. Dichos oligonucleótidos o ácidos nucleicos se pueden diseñar usando los procedimientos descritos en el presente documento. En algunos modos de realización, el kit comprende uno o más ácidos nucleicos marcados capaces de detección específica de alelo de uno o más marcadores polimórficos o haplotipos específicos con una variación genética, y reactivos para la detección del marcador. En algunos modos de realización, un kit para detectar marcadores de SNP puede comprender una sonda de detección de oligonucleótidos, que se hibrida con un segmento de ADN molde que contiene un polimorfismo de SNP que se va a detectar, una sonda de oligonucleótidos potenciadores, una sonda de detección, un cebador y/o una endonucleasa, por ejemplo, como se describe por Kutyaev y col. (Nucleic Acid Res. 34:e128 (2006)).

En algunos modos de realización, el molde de ADN se amplifica por cualquier medio de la presente divulgación, antes de la evaluación de la presencia de variaciones genéticas específicas como se describe en el presente documento. Se pueden utilizar procedimientos estándar bien conocidos por el experto en la técnica para realizar estos procedimientos y están dentro del alcance de la divulgación. En uno de dichos modos de realización, los reactivos para realizar estos procedimientos se pueden incluir en el kit de reactivos.

En un aspecto adicional de la presente divulgación, se proporciona un paquete farmacéutico (kit), comprendiendo el paquete un agente terapéutico y un conjunto de instrucciones para la administración del agente terapéutico a seres humanos cribados para una o más variantes de la presente divulgación, como se divulga en el presente documento. El agente terapéutico puede ser un fármaco de molécula pequeña, un anticuerpo, un péptido, una molécula antisentido o de ARNi u otras moléculas terapéuticas como se describe en el presente documento. En algunos modos de realización, un individuo identificado como un portador de al menos una variante de la presente divulgación recibe instrucciones de tomar una dosis prescrita del agente terapéutico. En uno de dichos modos de realización, un individuo identificado como un

portador de al menos una variante de la presente divulgación recibe instrucciones de tomar una dosis prescrita del agente terapéutico. En algunos modos de realización, un individuo identificado como un no portador de al menos una variante de la presente divulgación recibe instrucciones de tomar una dosis prescrita del agente terapéutico.

5 También se proporcionan en el presente documento artículos de fabricación, que comprenden una sonda que se hibrida con una región del cromosoma humano como se describe en el presente documento y se puede usar para detectar un polimorfismo descrito en el presente documento. Por ejemplo, cualquiera de las sondas para detectar polimorfismos
10 descritas en el presente documento se puede combinar con material de embalaje para generar artículos de fabricación o kits. El kit puede incluir uno o más de otros elementos que incluyen: instrucciones de uso; y otros reactivos tales como un marcador o un agente útil para unir un marcador a la sonda. Las instrucciones de uso pueden incluir instrucciones para
15 las aplicaciones de cribado de la sonda para hacer un diagnóstico, pronóstico o teranosis a una enfermedad o afección en un procedimiento descrito en el presente documento. Otras instrucciones pueden incluir instrucciones para unir un marcador a la sonda, instrucciones para realizar análisis *in situ* con la sonda y/o instrucciones para obtener una muestra de ácido nucleico que se va a analizar de un sujeto. En algunos casos, el kit puede incluir una sonda marcada que se hibrida con una región del cromosoma humano como se describe en el presente documento.

El kit también puede incluir una o más sondas de referencia o control adicionales que se hibridan con el mismo cromosoma u otro cromosoma o parte del mismo que puede tener una anomalía asociada con un endofenotipo particular. Un kit que
20 incluye sondas adicionales puede incluir además marcadores, por ejemplo, uno o más marcadores iguales o diferentes para las sondas. En otros modos de realización, la sonda o sondas adicionales proporcionadas con el kit pueden ser una sonda o sondas marcadas. Cuando el kit incluye además una o más sondas o sondas adicionales, el kit puede proporcionar además instrucciones para el uso de la sonda o sondas adicionales. También se pueden proporcionar kits para su uso en autodiagnóstico. Dichos kits de prueba pueden incluir dispositivos e instrucciones que un sujeto puede usar para obtener una muestra de ácido nucleico (por ejemplo, células bucales, sangre) sin la ayuda de un proveedor de
25 atención médica. Por ejemplo, las células bucales se pueden obtener usando un hisopo o cepillo bucal o usando un enjuague bucal.

Los kits que se proporcionan en el presente documento también pueden incluir un sobre prepagado (por ejemplo, un sobre o paquete de envío con franqueo pagado) que se puede usar para devolver la muestra de ácido nucleico para su análisis,
30 por ejemplo, a un laboratorio. El kit puede incluir uno o más recipientes para la muestra de ácido nucleico, o la muestra de ácido nucleico puede estar en un vial de recolección de sangre estándar. El kit también puede incluir uno o más formularios de consentimiento informado, un formulario de solicitud de prueba e instrucciones sobre cómo usar el kit en un procedimiento descrito en el presente documento. Los procedimientos para usar dichos kits también se incluyen en el presente documento. Uno o más de los formularios (por ejemplo, el formulario de solicitud de prueba) y el recipiente que
35 contiene la muestra de ácido nucleico se pueden codificar, por ejemplo, con un código de barras para identificar al sujeto que proporcionó la muestra de ácido nucleico.

En algunos modos de realización, una prueba de cribado *in vitro* puede comprender uno o más dispositivos, herramientas y equipos configurados para recolectar una muestra de ácido nucleico de un individuo. En algunos modos de realización
40 de una prueba de cribado *in vitro*, las herramientas para recolectar una muestra de ácido nucleico pueden incluir uno o más de un hisopo, un bisturí, una jeringa, un raspador, un recipiente y otros dispositivos y reactivos diseñados para facilitar la recolección, el almacenamiento y transporte de una muestra de ácido nucleico. En algunos modos de realización, una prueba de cribado *in vitro* puede incluir reactivos o soluciones para recolectar, estabilizar, almacenar y procesar una muestra de ácido nucleico.
45

Dichos reactivos y soluciones para la recolección, estabilización, almacenamiento y procesamiento de nucleótidos son bien conocidos por los expertos en la técnica y se pueden indicar mediante procedimientos específicos usados mediante una prueba de cribado *in vitro* como se describe en el presente documento. En algunos modos de realización, una prueba de cribado *in vitro* como se divulga en el presente documento, puede comprender un aparato de micromatrices y reactivos,
50 un aparato de células de flujo y reactivos, un secuenciador de nucleótidos de multiplexación y reactivos, y hardware y software adicionales necesarios para analizar una muestra de ácido nucleico para determinados marcadores genéticos y para detectar y visualizar determinados marcadores genéticos.

EJEMPLOS

55 **Ejemplo 1:** protocolo de secuenciación dirigida por ARN

Síntesis de ADNc

60 Se combinaron 1 ng hasta 1000 ng de ARN con 5 µl de la siguiente mezcla de cebadores que contenía 5 pmoles de cada cebador (SEQ ID NO: 3-7, respectivamente, en orden de aparición):

ACTB250A RT6p7_UID
/5FosCGATCTNNNNWNNNNAACCGACTGCTGTCACCTTC
ACTB250B RT6p7_UID

```

/5Fos/CGATCTNNNNWNNNNCCAGGGAGACCAAAGCCTT
RB2M250A RT6p7_UID
/5Fos/CGATCTNNNNWNNNNACCAGATTAACCACAACCATGC
GAPDH250A RT6p7_UID
/5Fos/CGATCTNNNNWNNNNATGGTTCACACCCATGACGAAC
GAPDH250B RT6p7_UID
/5Fos/CGATCTNNNNWNNNNGTTTTTCTAGACGGCAGGTCAG

```

La reacción de 12 µl se calentó durante 1 minuto a 95 °C, seguido de 65 °C durante 1 minuto y una retención a 4 °C. 4 µl de tampón 5x de primera cadena (Life Technologies, Carlsbad, CA.), 1 µl de dNTP 10 mM, 1 µl de DTT 0,1 M, 1 µl de inhibidor de ARNasa (Enzymatics, Beverly, MA.) y a continuación, se añadieron 1 µl de Superscript III (Life Technologies, Carlsbad, CA.) a la reacción. Esta reacción se incubó durante 45 minutos a 55 °C seguido de otros 5 minutos adicionales a 85 °C. La reacción se incubó a continuación a 37 °C después de la adición de 1 µl de ARNasa H (Enzymatics, Beverly, MA). La reacción se purificó con Ampure (Beckman Coulter Genomics, Danvers, MA).

Ligadura de adaptador

Se combinaron 3 µl de ADNc con 2 µl de adaptador P7/C7 10 µM, 1 µl de ADN ligasa T4 (Enzymatics, MA), 2 µl de tampón de ligasa rápida, y 2 µl de dH₂O libre de nucleasa. Las reacciones se incubaron durante 1 hora a temperatura ambiente. La reacción se inactivó con calor a continuación incubando durante 10 minutos a 65 °C, y a continuación, se purificó con Ampure XP (Beckman Coulter Genomics, Danvers, MA).

Secuencias adaptadoras (SEQ ID NO: 8-19, respectivamente, en orden de aparición)

P7 Cadena superior BC-1

5'-AGATCGGAAGAGCACACGTCTGAACTCCAGTCACATCACGATCTCGTATGCCGTCTTCTGCTTG

P7 Cadena superior BC-2

5'-AGATCGGAAGAGCACACGTCTGAACTCCAGTCACCGATGTATCTCGTATGCCGTCTTCTGCTTG

P7 Cadena superior BC-3

5'-AGATCGGAAGAGCACACGTCTGAACTCCAGTCACTTAGGCATCTCGTATGCCGTCTTCTGCTTG

P7 Cadena superior BC-4

5'-AGATCGGAAGAGCACACGTCTGAACTCCAGTCACTGACCAATCTCGTATGCCGTCTTCTGCTTG

P7 Cadena superior BC-5

5'-AGATCGGAAGAGCACACGTCTGAACTCCAGTCACACAGTGATCTCGTATGCCGTCTTCTGCTTG

P7 Cadena superior BC-6

5'-AGATCGGAAGAGCACACGTCTGAACTCCAGTCACGCCAATATCTCGTATGCCGTCTTCTGCTTG

C7/P7 Cadena inferior BC-1

/5BiotinTEG/CAAGCAGAAGACGGCATAACGAGATCGTGATGTGACTGGAGTTCAGACGTGTGCTCTTC

C7/P7 Cadena inferior BC-2

/5BiotinTEG/CAAGCAGAAGACGGCATAACGAGATACATCGGTGACTGGAGTTCAGACGTGTGCTCTTC

C7/P7 Cadena inferior BC-3

/5BiotinTEG/CAAGCAGAAGACGGCATAACGAGATGCCTAAGTGACTGGAGTTCAGACGTGTGCTCTTC

C7/P7 Cadena inferior BC-4

/5BiotinTEG/CAAGCAGAAGACGGCATAACGAGATTGGTTCAGTGACTGGAGTTCAGACGTGTGCTCTTC

C7/P7 Cadena inferior BC-5

/5BiotinTEG/CAAGCAGAAGACGGCATAACGAGATCACTGTGTGACTGGAGTTCAGACGTGTGCTCTTC

C7/P7 Cadena inferior BC-6

/5BiotinTEG/CAAGCAGAAGACGGCATAACGAGATATTGGCGTGACTGGAGTTCAGACGTGTGCTCTTC

Reacción de extensión de cebador

Se añadieron 10 µl de ADN ligado al adaptador a 8,4 µl de dH₂O, 0,3 µl de dNTP 10 mM, 5 µl de tampón Phusion HF, 0,3 µl de polimerasa Phusion Hotstart II (Thermo Fischer, Chicago, IL) y 0,5 pmoles de cada uno de los siguientes cebadores en un volumen de 1 µl:

Cebadores de control de patógenos (SEQ ID NO: 20-22, respectivamente, en orden de aparición)

HCV-1 A 250

C5/P5/PEAATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATCTCTCC
GAGCGGTCGCAAC

```

EBV A 250
C5/P5/PEAATGATACGGCGACCACCGAGATCTACTCTTTCCCTACACGACGCTCTTCCGATCTCCTGC
GCTCCATGAACATG
CMV A
C5/P5/PEAATGATACGGCGACCACCGAGATCTACTCTTTCCCTACACGACGCTCTTCCGATCTTAGAA
AAGTGACACACACGGATC
Cebadores diana (SEQ ID NO: 23-27, respectivamente, en orden de aparición)
ACTB250A
C5/P5/PEAATGATACGGCGACCACCGAGATCTACTCTTTCCCTACACGACGCTCTTCCGATCTTCCAG
CAGATGTGGATCAGCA
ACTB250B
C5/P5/PEAATGATACGGCGACCACCGAGATCTACTCTTTCCCTACACGACGCTCTTCCGATCTACAGG
AAGTCCCTTGCCATC
RB2M250A
C5/P5/PEAATGATACGGCGACCACCGAGATCTACTCTTTCCCTACACGACGCTCTTCCGATCTTCCAA
CATCAACATCTTGGTCAG
GAPDH250A
C5/P5/PEAATGATACGGCGACCACCGAGATCTACTCTTTCCCTACACGACGCTCTTCCGATCTCAAAT
TCCATGGCACCGTCAAG
GAPDH250B
C5/P5/PEAATGATACGGCGACCACCGAGATCTACTCTTTCCCTACACGACGCTCTTCCGATCTGGTAT
CGTGGAAGGACTCATG
    
```

La reacción se incubó durante 1 minuto a 98 °C, seguido de 5 ciclos de 98 °C, 20 segundos a 60 °C, 30 segundos a 72 °C seguido de una retención a 4 °C. La reacción se purificó, a continuación, con Ampure.

5 Amplificación por PCR

Se combinaron 5 µl de producto de extensión de cebador purificado con 10 µl de tampón 5x Phusion Hotstart, 0,6 µl de dNTP 10 mM, 2 µl de cebador de PCR C5 12,5 µM (AATGATACGGCGACCACCGAGATCT) (SEQ ID NO: 28), 2 µl de cebador de PCR C7 12,5 µM (CAAGCAGAAGACGGCATAACGAGAT) (SEQ ID NO: 29) 29,8 µl de dH₂O y 0,6 µl de polimerasa Phusion Hotstart II. La reacción se incubó durante 1 minuto a 98 °C seguido de 25 ciclos de 98 °C durante 10 segundos, 60 °C durante 20 segundos y 72 °C durante 30 segundos.

Reacciones agrupadas

Los productos de PCR se separaron en un gel de agarosa. Las bandas de gel se escindieron y purificaron con el kit de purificación de gel Minelute de Qiagen. Las muestras purificadas se analizaron mediante el análisis Agilent TapeStation, se diluyeron y se agruparon por cuantos de banda de biblioteca antes de la secuenciación en la plataforma Illumina MiSeq.

Ejemplo 2: preparación de secuenciación dirigida por ADN

Extensión de cebador genómico

Se combinaron 4 µg de ADN genómico humano, extraído de la sangre del paciente, con 0,6 µl de dNTP 10 mM, 1 µl de polimerasa BST 2.0 (New England Biolabs, Ipswich, MA), 5 µl de tampón de amplificación isotérmica (NEB) 10x y 1 µl de cebador CS-30 0,5 µM que contiene las secuencias a continuación.

```

CS_30 PE-1 (SEQ ID NO: 30-59, respectivamente, en orden de aparición)
SCA_1_UID_DPE1
/5Fos/CGATCTNNNNNWNNNNNWNNNNNACCTGTCTTGTAACCTTGATACC
SCA_2_UID_DPE1
/5Fos/CGATCTNNNNNWNNNNNWNNNNNGGGTATAAGTCTCTCTCGTATGTGATG
SCA_3_UID_DPE1
/5Fos/CGATCTNNNNNWNNNNNWNNNNNTCCCAAACAGCTTG
AATCACT
SCA_4_UID_DPE1
/5Fos/CGATCTNNNNNWNNNNNWNNNNNTCCCAAAGTGCTG
GGATTAC
    
```

ES 2 819 277 T3

SCA_5_UID_DPE1	/5Fos/CGATCTNNNNNWNNNNWNNNNNCATTTGCCATTCAA
ACAGAAGC	
SCA_6_UID_DPE1	/5Fos/CGATCTNNNNNWNNNNWNNNNNAGCAGGCTGGTAA
GAAATGG	
SCA_7_UID_DPE1	/5Fos/CGATCTNNNNNWNNNNWNNNNNGATCGCGCCACTGT
ACTC	
SCA_8_UID_DPE1	/5Fos/CGATCTNNNNNWNNNNWNNNNNGGAGAACACAGGA
ATGGGATG	
MSU_1_UID_DPE1	/5Fos/CGATCTNNNNNWNNNNWNNNNNCAGGGTTTGATTGT
CCCTAATG	
MSU_2_UID_DPE1	/5Fos/CGATCTNNNNNWNNNNWNNNNNTGATTCTCTGGGCAA
TGGG	
SNM1_1_DPE1_UID	/5Fos/CGATCTNNNNNWNNNNWNNNNNATACTTAGGGACA
ATGCAAGAGT	
SNM1_2_DPE1_UID	/5Fos/CGATCTNNNNNWNNNNWNNNNNTTATACTTAGGGACAATGCAAGAG
SNM1_3_DPE1_UID	/5Fos/CGATCTNNNNNWNNNNWNNNNNTTGCTCCTCTCTAT
TTCCATATCC	
SNM1_4_DPE1_UID	/5Fos/CGATCTNNNNNWNNNNWNNNNNACCTTAAATGAAG
CCACAGC	
CFTR_1_DPE1_UID	/5Fos/CGATCTNNNNNWNNNNWNNNNNTCCTTGCTTGAGA
GAAACC	
CFTR_2_DPE1_UID	/5Fos/CGATCTNNNNNWNNNNWNNNNNTGTTCCCACTGTGC
TATTAAG	
APOE_1_DPE1_UID	/5Fos/CGATCTNNNNNWNNNNWNNNNNCCTGCACCTGCTCA
GAC	
FMR1_1_DPE1_UID	/5Fos/CGATCTNNNNNWNNNNWNNNNNTGCCATGGGACAT
CAACAC	
G6PD_1_DPE1_UID	/5Fos/CGATCTNNNNNWNNNNWNNNNNACCACCCACCTTGA
AGAAG	
APOE_2_DPE1_UID	/5Fos/CGATCTNNNNNWNNNNWNNNNNGCTTCTGCAGGTCA
TCGG	
HexA_1_DPE1_UID	/5Fos/CGATCTNNNNNWNNNNWNNNNN

```

GGGATATGCCACTTCCATGAG
HexA_2_DPE1_UID      /5Fos/CGATCTNNNNNWNNNNNWNNNNN
CCCAAAGTGTTGGGATTACAG
SMPD1_1_DPE1_UID    /5Fos/CGATCTNNNNNWNNNNNWNNNNN
GGTCCTGACGAGTCTGGTG
CFTR_3_DPE1_UID
/5Fos/CGATCTNNNNNWNNNNNWNNNNNNTAGTTTCTTACCTCTTCTAGTTGGC
ASPA_1_DPE1_UID
                        /5Fos/CGATCTNNNNNWNNNNNWNNNNNNAGAAATTTGCTTAG
ATGCCTACC
ASPA_2_DPE1_UID
/5Fos/CGATCTNNNNNWNNNNNWNNNNNNTGTAAGACACCGTGTAAGATGTAAG
ASPA_3_DPE1_UID
                        /5Fos/CGATCTNNNNNWNNNNNWNNNNNGTACAGTCTCCGCC
CAGTG
CDH23_1_DPE1_UID
                        /5Fos/CGATCTNNNNNWNNNNNWNNNNNNCATGATCACGTCGC
GAAGTTTG
GBA_1_DPE1_UID
                        /5Fos/CGATCTNNNNNWNNNNNWNNNNNNAGGCCAGTCCTGA
TCCC
GBA_2_DPE1_UID
                        /5Fos/CGATCTNNNNNWNNNNNWNNNNNACAGGGCAAGGAT
GTTGAG

```

Ligadura de adaptador

5 Se combinaron 20 µl de la reacción de extensión del cebador eluido con 1,5 µl de adaptador P7/C7 5 µM (dúplex hibridado de oligo de 1 cadena superior y 1 cadena inferior descrito anteriormente con el emparejamiento de código de barras correcto), 1 µl de ADN ligasa T4, 6 µl de tampón de ligasa rápida 5x (New England Biolabs, Ipswich, MA) y 1,5 µl de dH₂O libre de nucleasa. Las reacciones se incubaron durante 1 hora a temperatura ambiente. La reacción se inactivó con calor incubando durante 10 minutos a 68 °C. La reacción se purificó a continuación con Ampure XP (Beckman).

10 Captura de microesferas

15 Se lavaron 180 µl de microesferas my One C1 SA (Dyna, Lifetech) con 1 ml de B&W 1x. Las microesferas se lavaron con 2 lavados de B&W 1x adicionales a 200 µl cada uno. El volumen de elución total de la ligadura del adaptador purificado con Ampure fue de 65 µl. Se añadió un volumen igual (65 µl) de B&W 2x y 100 µl adicionales de B&W 1x para un volumen total de 230 µl por unión. La reacción se colocó en el agitador de incubadora durante 20 minutos. Después de la unión de la muestra, las microesferas se lavaron con 200 µl de NSX y se eliminó el líquido.

20 Las muestras se volvieron a suspender a continuación en 200 µl de NaOH 0,1 N y se rotaron durante 20 minutos a temperatura ambiente. Se eliminó el NaOH y se realizó un segundo lavado con 200 µl adicionales de NaOH 0,1 N. Las microesferas se lavaron 2 veces con 600 µl de TE seguido de la eliminación de NaOH. Las microesferas se lavaron 2 veces con NSX. Las microesferas se colocaron en 100 µl de Tex (TE con Triton X al 0,01 %) y se almacenaron durante la noche a 4 C. Antes de la extensión de cebador, las microesferas se lavaron 2 veces con 200 µl de Phusion HF 1x (con Triton X al 0,01 %) y una vez con HF 1x sin Triton X.

25 Reacción de extensión de cebador

30 La mezcla de microesferas se volvió a suspender en 21,1 µl de dH₂O, 0,6 µl de dNTP 10 mM, 6 µl de tampón Phusion HF, 0,3 µl de polimerasa Phusion Hotstart II (Thermo Fischer, Chicago, IL), y 0,5 pmoles de cada uno de los siguientes cebadores en un volumen de 2 µl:

Cebadores de extensión de cebador 2 (SEQ ID NO: 60-89, respectivamente, en orden de aparición)

HexA_1_DPE2_P5/C5
AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATCTAACCTGAAGGGTG
TCTTGTG

HexA_2_DPE2_P5/C5
AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATCTATCAACAAGACTG
AGATTGAGG

SMPD1_1_DPE2_P5/C5
AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATCTCTGGGATCATGAC
TACCTGGAG

CFTR_3_DPE2_P5/C5
AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATCT
CTGAGCGTGATTTGATAATGACC

ASPA_1_DPE2_P5/C5
AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATCTCCCGTGTTTGTGA
ATGAGG

ASPA_2_DPE2_P5/C5
AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATCT
TTGTTTCCCTGAGAGGATCAAGAC

ASPA_3_DPE2_P5/C5
AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATCTATGTCAGCGCAGT
CAGATCAC

CDH23_1_DPE2_P5/C5
AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATCT
AGGGTAGCCTGCGCTTC

GBA_1_DPE2_P5/C5
AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATCT
ACTCTGGGTGCTTCTCTTTC

GBA_2_DPE2_P5/C5
AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATCT
CCCATCCAGGCTAATCACAC

SCA_1_C5/P5_DPE2
AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATCTGGGTTGGCCAATC
TACTCC

SCA_2_C5/P5_DPE2
AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATCTAATGACAGGGAGC
TTATAATTTAGCC

SCA_3_C5/P5_DPE2
AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATCTACATATTCAGCTG
GCACAGTTA

SCA_4_C5/P5_DPE2
AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATCTTGAAACACACCTG
AATACCTACAG

SCA_5_C5/P5_DPE2
AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATCTACAGGGCAGGCAT
GTTATC

SCA_6_C5/P5_DPE2
AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATCTGTGGTTTGGATCG
ACGTCTC

SCA_7_C5/P5_DPE2
AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATCTGGCCTTCAAAGAG
CACCTG

SCA_8_C5/P5_DPE2

```

AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCTACCCAGCTGCT
      CATGC
      MSU_1_C5/P5_DPE2
AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCAGGGAACAAATG
      CCAAGTG
      MSU_2_C5/P5_DPE2
AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGAAGGGAAGGAA
      GGAAGGG
      SNM1_1
AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGATTCTTTGATGA
      TGCTGATGC
      SNM1_2
AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTTCTTTGATGATG
      CTGATGC
      SNM1_3
AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCCTTCCAAATCTCT
      ACCCTCTATC
      SNM1_4
AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTAGTAAAGTCACAT
      AACCTCTAACC
      CFTR_1
AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTAGAGTTGGTAAGG
      AGGAGAATG
      CFTR_2
AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCTGTGGTATCTGA
      ACTATCTTCTC
      APOE_1
AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCATTTGTGGAGCA
      CCTTCTG
      FMR1_1
AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTAAGGATAGTTTGG
      AACTGAGAGAC
      G6PD_1
AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTGACCTGGCCAAG
      AAGAAG
      APOE_2
      AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT
      AATCGGAAGTGGAGGAACAAC
    
```

La reacción se incubó durante 1 minuto a 98 °C, seguido de 5 ciclos de 98 °C, 20 segundos a 60 °C, 30 segundos a 72 °C seguido de una retención a 4 °C. La reacción se purificó, a continuación, con Ampure.

5 Amplificación por PCR

Se combinaron 5 µl de producto de extensión de cebador purificado con 10 µl de tampón 5x Phusion Hotstart (HF), 0,6 µl de dNTP 10 mM, 2 µl de cebador de PCR C5 12,5 µM (AATGATACGGCGACCACCGAGATCT) (SEQ ID NO: 28), 2 µl de cebador de PCR C7 12,5 µM (CAAGCAGAAGACGGCATAACGAGAT) (SEQ ID NO: 29) 29,8 µl de dH₂O y 0,6 µl de polimerasa Phusion Hotstart II. La reacción se incubó durante 1 minuto a 98 °C seguido de 25 ciclos de 98 °C durante 10 segundos, 60 °C durante 20 segundos y 72 °C durante 30 segundos.

Las bandas de gel se escindieron y purificaron con el kit de purificación de gel Minelute de Qiagen de acuerdo con las instrucciones del fabricante. Las muestras purificadas se analizaron mediante el análisis Agilent TapeStation, y se diluyeron y se agruparon por cuantos de banda de biblioteca antes de la secuenciación en la Illumina MiSeq.

Ejemplo 3: creación mejorada del panel de cebadores - análisis de formación de dímero de cebador

Para crear paneles de cebadores para su uso en los procedimientos de secuenciación dirigida descritos, se evaluó la

estabilidad y la robustez de las dianas amplificadas. Adicionalmente, se evaluó la uniformidad de la cobertura y la exactitud de la secuencia para crear los paneles de cebadores y mejorar el rendimiento del ensayo.

Para mejorar estos parámetros, se evaluaron varias medidas, incluyendo la calidad de las dianas finales amplificadas, los requisitos de ciclos de amplificación, la limpieza de los productos amplificadas y el rendimiento de los productos amplificadas. También se realizó el análisis de secuencia de productos amplificadas para mejorar la especificidad en la diana, la uniformidad de cobertura, la profundidad de secuenciación y la exactitud de las lecturas de SNP. Se usaron ciclos iterativos de modificación del protocolo, análisis de la formación de producto y calidad de secuencia para mejorar el rendimiento del ensayo.

Utilizando el análisis de secuencia, se determinó que un producto no deseado de 75 pb estaba relacionado con los cebadores usados durante la etapa de extensión/amplificación lineal. Se determinó que un producto doblete o triplete más grande de 125-200 pb estaba relacionado con el adaptador C7/P7 y los cebadores usados durante la etapa de extensión/amplificación lineal. Se determinó que los productos de dímero más grandes > 150 pb estaban relacionados con los cebadores usados durante la etapa inicial de RT/PE.

Las principales longitudes de producto de dímero detectadas con el análisis de secuencia fueron 143, 155 y 160 y correspondieron a productos de dímero. El análisis de secuencia reveló que el producto de 143 pb estaba asociado con el cebador MCOLN1_11_1_f_PE2_5, que se produjo 132 veces, y el cebador GAA_14_1_0_PE2_7, que se produjo 660 veces. El análisis de secuencia reveló que el producto de 155 pb estaba asociado con el cebador GAA_14_1_0_PE2_7, que se produjo 1146 veces. El análisis de secuencia reveló que el producto de 160 pb estaba asociado con el cebador IKBKAP_32_1_CPE2_6, que se produjo 464 veces. Como resultado de este análisis, estos cebadores se eliminaron del panel de cebadores.

A partir de estos análisis, se descubrió que la formación de dímeros no deseada se ve facilitada por cebadores con altas temperaturas de fusión (por ejemplo, T_M de 70 °C) y bajas temperaturas de hibridación (por ejemplo, 60 °C), cebadores con alto contenido de GC al interactuar con las regiones de cebador/UID, y la exoactividad de 3' de algunas ADN polimerasas (por ejemplo, Phusion). Como resultado de estos análisis y conclusiones, se han creado paneles de cebadores con cebadores que no tienen un alto porcentaje de GC en sus últimos 5 nucleótidos en su extremo 3'. Como resultado de estos análisis y conclusiones, la formación de producto de dímero se ha reducido considerablemente en comparación con el uso de paneles de cebadores iniciales y las mejoras obviaron la necesidad de purificación en gel del producto diana.

Se crearon varios criterios de exclusión de cebadores a partir de los experimentos anteriores y se usaron para generar subpaneles a partir del panel CS-350. Los subpaneles se crearon usando uno o una combinación de estos parámetros de exclusión. En primer lugar, los cebadores con el mayor número de lecturas erróneas (causadas por errores de cebado) durante la etapa inicial de RT/PE o la etapa de extensión/amplificación lineal. En segundo lugar, los cebadores prevalentes en dímeros como se definen mediante análisis de secuencia se excluyeron de los subpaneles. En tercer lugar, los cebadores que fueron responsables de generar uno o más del mayor número de lecturas totales para una diana (sobreamplificadores) se excluyeron de los subpaneles.

Ejemplo 4: paneles de cebadores mejorados - análisis del % de contenido de GC en el amplicón y temperaturas de fusión del cebador

Para crear paneles de cebadores para su uso en los procedimientos de secuenciación dirigida descritos, se evaluó la estabilidad y la robustez de las dianas amplificadas en comparación con el % de contenido de GC de los amplicones y las temperaturas de fusión de los cebadores usados. El número de lecturas generadas para un cebador particular se usó como medida para el rendimiento del cebador. Adicionalmente, se evaluó la uniformidad de la cobertura y la exactitud de la secuencia para crear los paneles de cebadores y mejorar el rendimiento del ensayo.

La mayoría de los de bajo rendimiento (menor número de lecturas) tenían un cebador de extensión/amplificación lineal con una $T_M < 60$ °C y se derivaron de amplicones ricos en AT. Un segundo grupo con bajo rendimiento estaba compuesto por amplicones con porcentajes de GC más altos y cebadores con altas temperaturas de fusión. Como resultado de estos experimentos y análisis, se crearon varios criterios para los amplicones y cebadores. En primer lugar, el intervalo de temperatura de fusión de los cebadores que se van a usar debe estar entre 60 °C - 68 °C. En segundo lugar, los cebadores pueden tener una longitud de entre 21 y 32 nucleótidos. En tercer lugar, los cebadores no deben contener 4 o más pirimidinas en los últimos 5 nucleótidos en el extremo 3'. En cuarto lugar, el amplicón debe contener entre el 30 % y el 70 % de contenido de GC. Finalmente, la longitud del amplicón debe ser de entre 225 y 300 pares de bases de longitud.

Ejemplo 5: condiciones de reacción mejoradas

Para mejorar las condiciones de reacción para su uso en los procedimientos de secuenciación dirigida descritos, se evaluó la estabilidad y la robustez de las dianas amplificadas. Adicionalmente, se evaluó la uniformidad de la cobertura y la exactitud de la secuencia para mejorar las condiciones de reacción y mejorar el rendimiento del ensayo.

Para mejorar estos parámetros, se evaluaron varias medidas, incluyendo la calidad de las dianas finales amplificadas, los

requisitos de ciclos de amplificación, la limpieza de los productos amplificados y el rendimiento de los productos amplificados. Se usaron ciclos iterativos de modificación del protocolo, análisis de la formación de producto y calidad de secuencia para mejorar el rendimiento del ensayo.

5 Los experimentos iniciales de titulación de cebadores no fueron suficientes para permitir la producción diana con las condiciones de rampa e hibridación de amplificación existentes. Para grupos de cebadores altamente complejos se planteó la hipótesis de que se requerirían condiciones de rampa más rigurosas basadas en la evaluación de los parámetros y medidas anteriores.

10 Usando condiciones de rampa originales para el panel de cebadores CS-30, 30 dianas no funcionaron con paneles de cebadores más complejos. La rigurosidad se incrementó disminuyendo las velocidades de rampa para la etapa de extensión/amplificación lineal (PE2) y añadiendo una retención a 68 °C para la etapa inicial de RT/PE. La retención de temperatura mínima de hibridación se redujo a 55 °C para adaptarse a temperaturas de fusión del cebador más bajas. La fijación de la concentración global de los grupos de cebadores mostró una mejor formación del producto con tamaños de panel que varían de 24 a 346 amplicones. Una combinación de las rigurosas condiciones de RT/PE y de rampa de extensión/amplificación lineal con los grupos de cebadores globales fijos mostró mejoras con respecto a los mismos procedimientos que emplean diferentes condiciones.

15 Adicionalmente, se realizaron otros experimentos que emplearon diversos aditivos durante las etapas de RT/PE y extensión/amplificación lineal para mejorar la formación de producto. Se probaron varias condiciones de aditivos y se evaluó su impacto en la formación de producto. Los datos mostraron mejoras en la cobertura de lectura con condiciones de reacción optimizadas. El sulfato de amonio y el MgCl₂ adicional tuvieron el impacto más significativo en la profundidad de lectura. Estos experimentos se realizaron con el panel CS-350 completo antes de la optimización del panel. Estos experimentos se realizaron para ayudar a definir el mecanismo de formación de dímero e identificar los cebadores involucrados.

Ejemplo 5: protocolo de secuenciación dirigida

20 Los procedimientos descritos aquí se han usado para dirigir, amplificar, secuenciar y/o cuantificar específicamente secuencias de ADN o ARN presentes en una muestra. Estos procedimientos han permitido la adición de secuencias adicionales que formatearán las secuencias diana para la secuenciación u otros análisis moleculares. Los procedimientos se han usado para añadir una secuencia de identificación única (UID) que permite contener las lecturas derivadas de la misma molécula de ARN o ADN, lo que permite determinar si se encontraron determinados polimorfismos de secuencia en una población de moléculas de ARN o ADN, o resultaron de un artefacto de amplificación. Se ha usado ARN o ADN como el material molde/de partida. La muestra puede ser de cualquier organismo o virus. Los procedimientos se han usado para formatear moléculas diana para una variedad de dispositivos de secuenciación y otros dispositivos de análisis molecular.

25 Se usó un protocolo de preparación de biblioteca con el propósito de secuenciación dirigida que se va a secuenciar en la plataforma de NGS. En este ensayo, muchas dianas biológicas específicos (de una a muchos miles), de una muestra biológica del paciente, se convirtieron en una biblioteca compatible con NGS y se secuenciaron. Esto permitió la identificación de frecuencias diana (expresión génica) y de mutaciones o SNP en el genoma o transcriptoma del paciente, de los cuales se ha obtenido información clínica. Este ensayo también se usó para identificar la presencia o ausencia y la frecuencia de diversas infecciones dirigidas a ARN o ADN de virus, bacterias u hongos en muestras de pacientes.

30 Se han realizado diversas aplicaciones de forma individual o simultánea mediante la secuenciación de dianas necesarias para el perfil de mutación del cáncer, SNP y análisis de mutación, pruebas de vehículos, diagnósticos infecciosos y análisis de expresión génica, por ejemplo.

35 Para el ARN, la transcripción inversa (RT) se realizó usando enzima retrotranscriptasa, para generar un complemento de ADNc para las dianas o interés. Para el ADN, se realizó la extensión de cebador (PE), usando ADN polimerasa para generar ADN como complemento de las dianas o interés. En ambos casos, el oligo usado para realizar dicha RT o PE estaba compuesto por un cebador específico del gen dirigido contra la diana de interés, una marca de identificador único (UID) (un código de barras largo degenerado total o parcialmente compuesto por 15 o más bases degeneradas);
 40 NNNNNNNNNNNNNNNN (SEQ ID NO: 1) o NNNNNWNNNNWNNNNN (SEQ ID NO: 2), y la marca universal de una secuencia conocida (denominada cebador directo P7: P7f), con un extremo 5' fosforilado. La UID se usó para el código de barras de una sola molécula de cualquier molécula de ARN o ADN y se ha usado en la fase de análisis de secuencia para identificar el número absoluto de moléculas de partida en las muestras biológicas, deconvolucionar las secuencias consenso de la diana y eliminar todos los errores de PCR o secuenciación, aumentando, por lo tanto, la exactitud de secuenciación. Para capturar muchos genes diferentes, se usó un grupo compuesto por muchos de dichos oligo, donde las partes específicas del gen correspondientes del oligo eran un complemento para cada diana a capturar.

45 **Formateo/ligadura de adaptador:** en esta etapa, se añadió una secuencia adicional requerida para la amplificación/el análisis al ácido nucleico recién sintetizado. Esta secuencia adicional se puede añadir mediante ligadura (enfoque preferente), usando un oligo monocatenario o un oligo puente. Esta secuencia se ha añadido mediante amplificación en etapas posteriores. Esta secuencia se ha usado como una secuencia de cebado genérica para la amplificación de una

5 gran población de secuencias formateadas. Esta secuencia ha contenido un código de barras para la identificación de la muestra. Esta secuencia también ha contenido una marca de purificación tal como biotina. En el enfoque, un adaptador usado para la ligadura estaba compuesto por una cadena superior que servía como un oligo puente complementario a la región P7f, y un oligo de cadena inferior que estaba ligado al producto generado durante la etapa de RT o PE. El producto resultante añadió el resto de la región P7 (para la secuenciación), así como un código de barras de muestra (SBC), requerido si muchas muestras de pacientes se procesan en paralelo, y opcionalmente, la región C7, para agruparse en una plataforma de NGS.

10 **Captura de microesferas (opcional):** en esta etapa, se capturó un ácido nucleico parcialmente formateado por medio de una marca o secuencia de afinidad añadida anteriormente. Esta captura se usó para separar las secuencias diana de las secuencias molde/de muestra que no son de interés.

15 **Extensión de cebador/amplificación lineal:** se realizó amplificación lineal o extensión de cebador lineal (LPE) usando una ADN polimerasa y usando un grupo de oligos compuestos por una región específica del gen para cada una de las dianas a capturar, una marca de cebador de secuenciación (P5), y una marca universal (C5) para agrupar en una plataforma de NGS. Se usó un grupo de oligos para realizar LPE de muchas dianas a la vez en una sola reacción. Esta extensión se produjo en solución o con el molde unido a una microesfera o una matriz. La LPE se ha realizado como un solo ciclo o muchos ciclos (hasta cientos), evitando el sesgo de amplificación por PCR que se generaría en la PCR estándar.

20 **Enriquecimiento por PCR:** las dianas de interés se amplificaron simultáneamente mediante PCR usando los siguientes oligos: un cebador directo compuesto por cualquier parte del oligo de LPE, preferentemente compuesto por C5 (u opcionalmente P5C5, o solo P5), y un cebador inverso complementario a cualquier parte del adaptador universal, pero preferentemente complementario a C7 (u opcionalmente P7-BC-C7 o solo P7).

25 **Biblioteca final:** la biblioteca final estaba compuesta por un grupo de todas las dianas capturadas con las marcas.

30 Se pretende que las siguientes reivindicaciones definan el alcance de los procedimientos, composiciones y kits descritos en el presente documento. Las reivindicaciones se pueden redactar de modo que excluyan cualquier elemento opcional. Como tal, esta declaración tiene la intención de servir como base antecedente para el uso de terminología exclusiva tal como "únicamente", "solo" y similares en relación con la mención reiterada de elementos de la reivindicación o el uso de una limitación "negativa".

LISTADO DE SECUENCIAS

<110> F. HOFFMANN-LA ROCHE AG

5 <120> SECUENCIACIÓN DIRIGIDA Y FILTRADO DE UID

<130> 32510-WO

<140>

10 <141>

<150> 62/031.405

<151> 31/07/2014

15 <150> 61/938.227

<151> 11/02/2014

<160> 100

20 <170> PatentIn versión 3.5

<210> 1

<211> 15

<212> ADN

25 <213> Secuencia artificial

<220>

<221> fuente

<223> /nota="Descripción de secuencia artificial: oligonucleótido sintético"

30 <220>

<221> base_modificada

<222> (1). . (15)

<223> a, c, t, g, desconocida u otra

35 <400> 1

nnnnnnnnnnnn nnnnn 15

<210> 2

40 <211> 17

<212> ADN

<213> Secuencia artificial

<220>

45 <221> fuente

<223> /nota="Descripción de secuencia artificial: oligonucleótido sintético"

<220>

50 <221> base_modificada

<222> (1). . (5)

<223> a, c, t, g, desconocida u otra

<220>

55 <221> base_modificada

<222> (7). . (11)

<223> a, c, t, g, desconocida u otra

<220>

60 <221> base_modificada

<222> (13). . (17)

<223> a, c, t, g, desconocida u otra

<400> 2

nnnnnnwnnnnn nwnnnnnn 17

65 <210> 3

<211> 35
 <212> ADN
 <213> Secuencia artificial

5 <220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"

10 <220>
 <221> fuente
 <223> /nota="5' -Fos modificado"

15 <220>
 <221> base_modificada
 <222> (7). . (10)
 <223> a, c, t, g, desconocida u otra

20 <220>
 <221> base_modificada
 <222> (12). . (15)
 <223> a, c, t, g, desconocida u otra

25 <400> 3
cgatctnnnn wnnnaaccg actgctgtca ccttc 35

30 <210> 4
 <211> 35
 <212> ADN
 <213> Secuencia artificial

35 <220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"

40 <220>
 <221> base_modificada
 <222> (7). . (10)
 <223> a, c, t, g, desconocida u otra

45 <220>
 <221> base_modificada
 <222> (12). . (15)
 <223> a, c, t, g, desconocida u otra

50 <400> 4
cgatctnnnn wnnnccagg gagaccaaaa gcctt 35

55 <210> 5
 <211> 37
 <212> ADN
 <213> Secuencia artificial

60 <220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"

65 <220>
 <221> fuente
 <223> /nota="5' -Fos modificado"

65 <220>
 <221> base_modificada

<222> (7). . (10)
 <223> a, c, t, g, desconocida u otra

 <220>
 5 <221> base_modificada
 <222> (12). . (15)
 <223> a, c, t, g, desconocida u otra

 <400> 5
 10 **cgatctnnnn wnnnnaccag attaaccaca accatgc** 37

 <210> 6
 <211> 37
 <212> ADN
 15 <213> Secuencia artificial

 <220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"
 20
 <220>
 <221> fuente
 <223> /nota="5' -Fos modificado"
 25
 <220>
 <221> base_modificada
 <222> (7). . (10)
 <223> a, c, t, g, desconocida u otra
 30
 <220>
 <221> base_modificada
 <222> (12). . (15)
 <223> a, c, t, g, desconocida u otra
 35
 <400> 6
cgatctnnnn wnnnnatggt tcacacccat gacgaac 37

 <210> 7
 <211> 37
 40 <212> ADN
 <213> Secuencia artificial

 <220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"
 45
 <220>
 <221> fuente
 <223> /nota="5' -Fos modificado"
 50
 <220>
 <221> base_modificada
 <222> (7). . (10)
 <223> a, c, t, g, desconocida u otra
 55
 <220>
 <221> base_modificada
 <222> (12). . (15)
 <223> a, c, t, g, desconocida u otra
 60
 <400> 7
cgatctnnnn wnnnngtttt tctagacggc aggtcag 37

 <210> 8
 65 <211> 64

ES 2 819 277 T3

<212> ADN
 <213> Secuencia artificial

<220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: oligonucleótido sintético"

5 <400> 8
agatcggaag agcacacgtc tgaactccag tcacatcacg atctcgtatg ccgtcttctg 60
cttg 64

10 <210> 9
 <211> 64
 <212> ADN
 <213> Secuencia artificial

15 <220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: oligonucleótido sintético"

20 <400> 9
agatcggaag agcacacgtc tgaactccag tcaccgatgt atctcgtatg ccgtcttctg 60
cttg 64

25 <210> 10
 <211> 64
 <212> ADN
 <213> Secuencia artificial

30 <220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: oligonucleótido sintético"

<400> 10
agatcggaag agcacacgtc tgaactccag tcacttaggc atctcgtatg ccgtcttctg 60
cttg 64

35 <210> 11
 <211> 64
 <212> ADN
 <213> Secuencia artificial

40 <220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: oligonucleótido sintético"

<400> 11
agatcggaag agcacacgtc tgaactccag tcactgacca atctcgtatg ccgtcttctg 60
cttg 64

45 <210> 12
 <211> 64
 <212> ADN
 <213> Secuencia artificial

50 <220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: oligonucleótido sintético"

55 <400> 12

	agatcgggaag agcacacgtc tgaactccag tcacacagtg atctcgtatg ccgtcttctg	60
	cttg	64
5	<210> 13 <211> 64 <212> ADN <213> Secuencia artificial	
10	<220> <221> fuente <223> /nota="Descripción de secuencia artificial: oligonucleótido sintético"	
	<400> 13 agatcgggaag agcacacgtc tgaactccag tcacgccaat atctcgtatg ccgtcttctg	60
	cttg	64
15	<210> 14 <211> 58 <212> ADN <213> Secuencia artificial	
20	<220> <221> fuente <223> /nota="Descripción de secuencia artificial: oligonucleótido sintético"	
25	<220> <221> fuente <223> /nota="5'-BiotinTEG modificado"	
30	<400> 14 caagcagaag acggcatacg agatcgtgat gtgactggag ttcagacgtg tgctcttc	58
35	<210> 15 <211> 58 <212> ADN <213> Secuencia artificial	
40	<220> <221> fuente <223> /nota="5'-BiotinTEG modificado"	
45	<400> 15 caagcagaag acggcatacg agatacatcg gtgactggag ttcagacgtg tgctcttc	58
50	<210> 16 <211> 58 <212> ADN <213> Secuencia artificial	
55	<220> <221> fuente <223> /nota="Descripción de secuencia artificial: oligonucleótido sintético"	
60	<220> <221> fuente <223> /nota="5'-BiotinTEG modificado"	
	<400> 16	

	caagcagaag acggcatacg agatgcctaa gtgactggag ttcagacgtg tgctcttc	58
	<210> 17	
	<211> 58	
5	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<221> fuente	
10	<223> /nota="Descripción de secuencia artificial: oligonucleótido sintético"	
	<220>	
	<221> fuente	
15	<223> /nota="5'-BiotinTEG modificado"	
	<400> 17	
	caagcagaag acggcatacg agattggtca gtgactggag ttcagacgtg tgctcttc	58
	<210> 18	
20	<211> 58	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
25	<221> fuente	
	<223> /nota="Descripción de secuencia artificial: oligonucleótido sintético"	
	<220>	
	<221> fuente	
30	<223> /nota="5'-BiotinTEG modificado"	
	<400> 18	
	caagcagaag acggcatacg agatcactgt gtgactggag ttcagacgtg tgctcttc	58
35	<210> 19	
	<211> 58	
	<212> ADN	
	<213> Secuencia artificial	
40	<220>	
	<221> fuente	
	<223> /nota="Descripción de secuencia artificial: oligonucleótido sintético"	
45	<220>	
	<221> fuente	
	<223> /nota="5'-BiotinTEG modificado"	
	<400> 19	
	caagcagaag acggcatacg agatattggc gtgactggag ttcagacgtg tgctcttc	58
50	<210> 20	
	<211> 76	
	<212> ADN	
	<213> Secuencia artificial	
55	<220>	
	<221> fuente	
	<223> /nota="Descripción de secuencia artificial: cebador sintético"	
60	<400> 20	
	aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatctct	60
	tccgagcggg cgcaac	76

ES 2 819 277 T3

5	<p><210> 21 <211> 77 <212> ADN <213> Secuencia artificial</p>	
	<p><220> <221> fuente <223> /nota="Descripción de secuencia artificial: cebador sintético"</p>	
10	<p><400> 21 aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatctcc</p>	60
	<p>tgcgctccat gaacatg</p>	77
15	<p><210> 22 <211> 81 <212> ADN <213> Secuencia artificial</p>	
	<p><220> <221> fuente <223> /nota="Descripción de secuencia artificial: cebador sintético"</p>	
20	<p><400> 22 aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatctta</p>	60
	<p>gaaaagtgac acacacggat c</p>	81
25	<p><210> 23 <211> 79 <212> ADN <213> Secuencia artificial</p>	
	<p><220> <221> fuente <223> /nota="Descripción de secuencia artificial: cebador sintético"</p>	
30	<p><400> 23 aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatcttc</p>	60
35	<p>cagcagatgt ggatcagca</p>	79
40	<p><210> 24 <211> 78 <212> ADN <213> Secuencia artificial</p>	
	<p><220> <221> fuente <223> /nota="Descripción de secuencia artificial: cebador sintético"</p>	
45	<p><400> 24 aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatctac</p>	60
	<p>aggaagtccc ttgccatc</p>	78
50	<p><210> 25 <211> 81 <212> ADN <213> Secuencia artificial</p>	
	<p><220> <221> fuente <223> /nota="Descripción de secuencia artificial: cebador sintético"</p>	
55	<p><400> 25 aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatctac</p>	60
	<p>aggaagtccc ttgccatc</p>	78

ES 2 819 277 T3

	<400> 25		
	aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatcttc		60
	caacatcaac atcttggtca g		81
5	<210> 26 <211> 80 <212> ADN <213> Secuencia artificial		
10	<220> <221> fuente <223> /nota="Descripción de secuencia artificial: cebador sintético"		
	<400> 26		
	aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatctca		60
15	aattccatgg caccgtcaag		80
	<210> 27 <211> 79 <212> ADN <213> Secuencia artificial		
20	<220> <221> fuente <223> /nota="Descripción de secuencia artificial: cebador sintético"		
25	<400> 27		
	aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatctgg		60
	tatcgtggaa ggactcatg		79
30	<210> 28 <211> 25 <212> ADN <213> Secuencia artificial		
35	<220> <221> fuente <223> /nota="Descripción de secuencia artificial: cebador sintético"		
40	<400> 28		
	aatgatacgg cgaccaccga gatct		25
45	<210> 29 <211> 24 <212> ADN <213> Secuencia artificial		
50	<220> <221> fuente <223> /nota="Descripción de secuencia artificial: cebador sintético"		
55	<400> 29		
	caagcagaag acggcatacg agat		24
	<210> 30 <211> 46 <212> ADN <213> Secuencia artificial		
	<220>		

<221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"

<220>
 5 <221> fuente
 <223> /nota="5'-Fos modificado"

<220>
 10 <221> base_modificada
 <222> (7). . (11)
 <223> a, c, t, g, desconocida u otra

<220>
 15 <221> base_modificada
 <222> (13). . (17)
 <223> a, c, t, g, desconocida u otra

<220>
 20 <221> base_modificada
 <222> (19). . (23)
 <223> a, c, t, g, desconocida u otra

<400> 30
 25 **cgatctnnnn nwnnnnnwnn nnnacctgtc ttgtaacctt gatacc** 46

<210> 31
 <211> 50
 <212> ADN
 <213> Secuencia artificial

30 <220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"

35 <220>
 <221> fuente
 <223> /nota="5'-Fos modificado"

40 <220>
 <221> base_modificada
 <222> (7). . (11)
 <223> a, c, t, g, desconocida u otra

45 <220>
 <221> base_modificada
 <222> (13). . (17)
 <223> a, c, t, g, desconocida u otra

50 <220>
 <221> base_modificada
 <222> (19). . (23)
 <223> a, c, t, g, desconocida u otra

<400> 31
 55 **cgatctnnnn nwnnnnnwnn nnnnggtata agtctctctc gtatgtgatg** 50

<210> 32
 <211> 44
 <212> ADN
 60 <213> Secuencia artificial

<220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"

65

<220>
 <221> fuente
 <223> /nota="5'-Fos modificado"

5 <220>
 <221> base_modificada
 <222> (7). . (11)
 <223> a, c, t, g, desconocida u otra

10 <220>
 <221> base_modificada
 <222> (13). . (17)
 <223> a, c, t, g, desconocida u otra

15 <220>
 <221> base_modificada
 <222> (19). . (23)
 <223> a, c, t, g, desconocida u otra

20 <400> 32
 cgatctnnnn nwnnnnnwnn nnntcccaaa cagcttgaat cact **44**

<210> 33
 <211> 43

25 <212> ADN
 <213> Secuencia artificial

<220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"

30 <220>
 <221> fuente
 <223> /nota="5'-Fos modificado"

35 <220>
 <221> base_modificada
 <222> (7). . (11)
 <223> a, c, t, g, desconocida u otra

40 <220>
 <221> base_modificada
 <222> (13). . (17)
 <223> a, c, t, g, desconocida u otra

45 <220>
 <221> base_modificada
 <222> (19). . (23)
 <223> a, c, t, g, desconocida u otra

50 <400> 33
 cgatctnnnn nwnnnnnwnn nnntcccaaa gtgctgggat tac **43**

<210> 34
 <211> 45

55 <212> ADN
 <213> Secuencia artificial

<220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"

60 <220>
 <221> fuente
 <223> /nota="5'-Fos modificado"

65 <220>
 <221> fuente
 <223> /nota="5'-Fos modificado"

<220>
 <221> base_modificada
 <222> (7). . (11)
 <223> a, c, t, g, desconocida u otra
 5

<220>
 <221> base_modificada
 <222> (13). . (17)
 <223> a, c, t, g, desconocida u otra
 10

<220>
 <221> base_modificada
 <222> (19). . (23)
 <223> a, c, t, g, desconocida u otra
 15

<400> 34
cgatctnnnn nwnnnnnwnn nnncatttgc cattcaaaca gaagc 45

<210> 35
 <211> 43
 <212> ADN
 <213> Secuencia artificial

<220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"
 25

<220>
 <221> fuente
 <223> /nota="5'-Fos modificado"
 30

<220>
 <221> base_modificada
 <222> (7). . (11)
 <223> a, c, t, g, desconocida u otra
 35

<220>
 <221> base_modificada
 <222> (13). . (17)
 <223> a, c, t, g, desconocida u otra
 40

<220>
 <221> base_modificada
 <222> (19). . (23)
 <223> a, c, t, g, desconocida u otra
 45

<400> 35
cgatctnnnn nwnnnnnwnn nnnagcaggc tggtagaataa tgg 43

<210> 36
 <211> 41
 <212> ADN
 <213> Secuencia artificial

<220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"
 55

<220>
 <221> fuente
 <223> /nota="5'-Fos modificado"
 60

<220>
 <221> base_modificada
 <222> (7). . (11)
 65

<223> a, c, t, g, desconocida u otra

<220>
 <221> base_modificada
 5 <222> (13). . (17)
 <223> a, c, t, g, desconocida u otra

<220>
 <221> base_modificada
 10 <222> (19). . (23)
 <223> a, c, t, g, desconocida u otra

<400> 36
 15 **cgatctnnnn nwnnnnnwnn nnnngatcgcg ccactgtact c** 41

<210> 37
 <211> 44
 <212> ADN
 <213> Secuencia artificial

20 <220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"

25 <220>
 <221> fuente
 <223> /nota="5'-Fos modificado"

30 <220>
 <221> base_modificada
 <222> (7). . (11)
 <223> a, c, t, g, desconocida u otra

35 <220>
 <221> base_modificada
 <222> (13). . (17)
 <223> a, c, t, g, desconocida u otra

40 <220>
 <221> base_modificada
 <222> (19). . (23)
 <223> a, c, t, g, desconocida u otra

<400> 37
 45 **cgatctnnnn nwnnnnnwnn nnnngagaac acaggaatgg gatg** 44

<210> 38
 <211> 45
 <212> ADN
 <213> Secuencia artificial

50 <220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"

55 <220>
 <221> fuente
 <223> /nota="5'-Fos modificado"

60 <220>
 <221> base_modificada
 <222> (7). . (11)
 <223> a, c, t, g, desconocida u otra

65 <220>

<221> base_modificada
 <222> (13). . (17)
 <223> a, c, t, g, desconocida u otra

5

<220>
 <221> base_modificada
 <222> (19). . (23)
 <223> a, c, t, g, desconocida u otra

10

<400> 38
cgatctnnnn nwnnnnnwnn nncagggtt tgattgtccc taatg
45

<210> 39
 <211> 41
 <212> ADN
 <213> Secuencia artificial

15

<220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"

20

<220>
 <221> fuente
 <223> /nota="5'-Fos modificado"

25

<220>
 <221> base_modificada
 <222> (7). . (11)
 <223> a, c, t, g, desconocida u otra

30

<220>
 <221> base_modificada
 <222> (13). . (17)
 <223> a, c, t, g, desconocida u otra

35

<220>
 <221> base_modificada
 <222> (19). . (23)
 <223> a, c, t, g, desconocida u otra

40

<400> 39
cgatctnnnn nwnnnnnwnn nnntgattcc tgggcaatgg g
41

45

<210> 40
 <211> 46
 <212> ADN
 <213> Secuencia artificial

50

<220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"

55

<220>
 <221> fuente
 <223> /nota="5'-Fos modificado"

60

<220>
 <221> base_modificada
 <222> (7). . (11)
 <223> a, c, t, g, desconocida u otra

65

<220>
 <221> base_modificada
 <222> (13). . (17)
 <223> a, c, t, g, desconocida u otra

<220>
 <221> base_modificada
 <222> (19). . (23)
 5 <223> a, c, t, g, desconocida u otra

 <400> 40
cgatctnnnn nwnnnnnwnn nnnataactta gggacaatgc aagagt 46

10 <210> 41
 <211> 47
 <212> ADN
 <213> Secuencia artificial

15 <220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"

20 <220>
 <221> fuente
 <223> /nota="5'-Fos modificado"

25 <220>
 <221> base_modificada
 <222> (7). . (11)
 <223> a, c, t, g, desconocida u otra

30 <220>
 <221> base_modificada
 <222> (13). . (17)
 <223> a, c, t, g, desconocida u otra

35 <220>
 <221> base_modificada
 <222> (19). . (23)
 <223> a, c, t, g, desconocida u otra

40 <400> 41
cgatctnnnn nwnnnnnwnn nnnttataact tagggacaat gcaagag 47

45 <210> 42
 <211> 47
 <212> ADN
 <213> Secuencia artificial

50 <220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"

55 <220>
 <221> base_modificada
 <222> (7). . (11)
 <223> a, c, t, g, desconocida u otra

60 <220>
 <221> base_modificada
 <222> (13). . (17)
 <223> a, c, t, g, desconocida u otra

65 <220>
 <221> base_modificada

<222> (19). . (23)
 <223> a, c, t, g, desconocida u otra

<400> 42
 5 **cgatctnnnn nwnnnnnwnn nnnttgctcc tctctatttc catatcc** 47

<210> 43
 <211> 43
 <212> ADN
 10 <213> Secuencia artificial

<220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"
 15

<220>
 <221> fuente
 <223> /nota="5'-Fos modificado"

<220>
 <221> base_modificada
 <222> (7). . (11)
 <223> a, c, t, g, desconocida u otra
 20

<220>
 <221> base_modificada
 <222> (13). . (17)
 <223> a, c, t, g, desconocida u otra
 25

<220>
 <221> base_modificada
 <222> (19). . (23)
 <223> a, c, t, g, desconocida u otra
 30

<400> 43
 35 **cgatctnnnn nwnnnnnwnn nnnaccttaa atgaagccac agc** 43

<210> 44
 <211> 43
 40 <212> ADN
 <213> Secuencia artificial

<220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"
 45

<220>
 <221> fuente
 <223> /nota="5'-Fos modificado"
 50

<220>
 <221> base_modificada
 <222> (7). . (11)
 <223> a, c, t, g, desconocida u otra
 55

<220>
 <221> base_modificada
 <222> (13). . (17)
 <223> a, c, t, g, desconocida u otra
 60

<220>
 <221> base_modificada
 <222> (19). . (23)
 <223> a, c, t, g, desconocida u otra
 65

	<400> 44 cgatctnnnn nwnnnnnwnn nnntccttgg cttgagagaa acc	43
5	<210> 45 <211> 44 <212> ADN <213> Secuencia artificial	
10	<220> <221> fuente <223> /nota="Descripción de secuencia artificial: cebador sintético"	
15	<220> <221> fuente <223> /nota="5'-Fos modificado"	
20	<220> <221> base_modificada <222> (7). . (11) <223> a, c, t, g, desconocida u otra	
25	<220> <221> base_modificada <222> (13). . (17) <223> a, c, t, g, desconocida u otra	
30	<220> <221> base_modificada <222> (19). . (23) <223> a, c, t, g, desconocida u otra	
	<400> 45 cgatctnnnn nwnnnnnwnn nnntggtccc actgtgctat taag	44
35	<210> 46 <211> 40 <212> ADN <213> Secuencia artificial	
40	<220> <221> fuente <223> /nota="Descripción de secuencia artificial: cebador sintético"	
45	<220> <221> fuente <223> /nota="5'-Fos modificado"	
50	<220> <221> base_modificada <222> (7). . (11) <223> a, c, t, g, desconocida u otra	
55	<220> <221> base_modificada <222> (13). . (17) <223> a, c, t, g, desconocida u otra	
60	<220> <221> base_modificada <222> (19). . (23) <223> a, c, t, g, desconocida u otra	
65	<400> 46 cgatctnnnn nwnnnnnwnn nnnctgcac ctgctcagac	40

<210> 47
 <211> 42
 <212> ADN
 <213> Secuencia artificial
 5
 <220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"
 10
 <220>
 <221> fuente
 <223> /nota="5'-Fos modificado"
 15
 <220>
 <221> base_modificada
 <222> (7). . (11)
 <223> a, c, t, g, desconocida u otra
 20
 <220>
 <221> base_modificada
 <222> (13). . (17)
 <223> a, c, t, g, desconocida u otra
 25
 <220>
 <221> base_modificada
 <222> (19). . (23)
 <223> a, c, t, g, desconocida u otra
 30
 <400> 47
cgatctnnnn nwnnnnnwnn nnntgccatg ggacatcaac ac 42
 <210> 48
 <211> 42
 <212> ADN
 <213> Secuencia artificial
 35
 <220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"
 40
 <220>
 <221> fuente
 <223> /nota="5'-Fos modificado"
 45
 <220>
 <221> base_modificada
 <222> (7). . (11)
 <223> a, c, t, g, desconocida u otra
 50
 <220>
 <221> base_modificada
 <222> (13). . (17)
 <223> a, c, t, g, desconocida u otra
 55
 <220>
 <221> base_modificada
 <222> (19). . (23)
 <223> a, c, t, g, desconocida u otra
 60
 <400> 48
cgatctnnnn nwnnnnnwnn nnnaccaccc accttgaaga ag 42
 <210> 49
 <211> 41
 <212> ADN
 65

<213> Secuencia artificial

<220>
 <221> fuente
 5 <223> /nota="Descripción de secuencia artificial: cebador sintético"

<220>
 <221> fuente
 10 <223> /nota="5'-Fos modificado"

<220>
 <221> base_modificada
 <222> (7). . (11)
 15 <223> a, c, t, g, desconocida u otra

<220>
 <221> base_modificada
 <222> (13). . (17)
 20 <223> a, c, t, g, desconocida u otra

<220>
 <221> base_modificada
 <222> (19). . (23)
 25 <223> a, c, t, g, desconocida u otra

<400> 49
cgatctnnnn nwnnnnnwnn nnnngcttctg caggtcacg g 41

<210> 50
 30 <211> 44
 <212> ADN
 <213> Secuencia artificial

<220>
 <221> fuente
 35 <223> /nota="Descripción de secuencia artificial: cebador sintético"

<220>
 <221> fuente
 40 <223> /nota="5'-Fos modificado"

<220>
 <221> base_modificada
 <222> (7). . (11)
 45 <223> a, c, t, g, desconocida u otra

<220>
 <221> base_modificada
 <222> (13). . (17)
 50 <223> a, c, t, g, desconocida u otra

<220>
 <221> base_modificada
 <222> (19). . (23)
 55 <223> a, c, t, g, desconocida u otra

<400> 50
cgatctnnnn nwnnnnnwnn nnnnggatat gccacttcca tgag 44

<210> 51
 60 <211> 44
 <212> ADN
 <213> Secuencia artificial

65 <220>

<221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"

<220>
 5 <221> fuente
 <223> /nota="5'-Fos modificado"

<220>
 10 <221> base_modificada
 <222> (7). . (11)
 <223> a, c, t, g, desconocida u otra

<220>
 15 <221> base_modificada
 <222> (13). . (17)
 <223> a, c, t, g, desconocida u otra

<220>
 20 <221> base_modificada
 <222> (19). . (23)
 <223> a, c, t, g, desconocida u otra

<400> 51
 25 **cgatctnnnn nwnnnnnwnn nnncccaaag tgttgggatt acag** **44**

<210> 52
 <211> 42
 <212> ADN
 <213> Secuencia artificial

30 <220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"

35 <220>
 <221> fuente
 <223> /nota="5'-Fos modificado"

40 <220>
 <221> base_modificada
 <222> (7). . (11)
 <223> a, c, t, g, desconocida u otra

45 <220>
 <221> base_modificada
 <222> (13). . (17)
 <223> a, c, t, g, desconocida u otra

50 <220>
 <221> base_modificada
 <222> (19). . (23)
 <223> a, c, t, g, desconocida u otra

<400> 52
 55 **cgatctnnnn nwnnnnnwnn nnnnggtcctg acgagtctgg tg** **42**

<210> 53
 <211> 48
 <212> ADN
 60 <213> Secuencia artificial

<220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"

65

<220>
 <221> fuente
 <223> /nota="5'-Fos modificado"

5 <220>
 <221> base_modificada
 <222> (7). . (11)
 <223> a, c, t, g, desconocida u otra

10 <220>
 <221> base_modificada
 <222> (13). . (17)
 <223> a, c, t, g, desconocida u otra

15 <220>
 <221> base_modificada
 <222> (19). . (23)
 <223> a, c, t, g, desconocida u otra

20 <400> 53
 cgatctnnnn nwnnnnnwnn nnntagtttc ttacctcttc tagttggc **48**

<210> 54
 <211> 46
 <212> ADN
 <213> Secuencia artificial

25 <220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"

<220>
 <221> fuente
 <223> /nota="5'-Fos modificado"

35 <220>
 <221> base_modificada
 <222> (7). . (11)
 <223> a, c, t, g, desconocida u otra

40 <220>
 <221> base_modificada
 <222> (13). . (17)
 <223> a, c, t, g, desconocida u otra

45 <220>
 <221> base_modificada
 <222> (19). . (23)
 <223> a, c, t, g, desconocida u otra

50 <400> 54
 cgatctnnnn nwnnnnnwnn nnnagaaatt tgcttagatg cctacc **46**

<210> 55
 <211> 48
 <212> ADN
 <213> Secuencia artificial

55 <220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"

<220>
 <221> fuente
 <223> /nota="5'-Fos modificado"

65 <220>
 <221> fuente
 <223> /nota="5'-Fos modificado"

<220>
 <221> base_modificada
 <222> (7). . (11)
 <223> a, c, t, g, desconocida u otra
 5

<220>
 <221> base_modificada
 <222> (13). . (17)
 <223> a, c, t, g, desconocida u otra
 10

<220>
 <221> base_modificada
 <222> (19). . (23)
 <223> a, c, t, g, desconocida u otra
 15

<400> 55
cgatctnnnn nwnnnnnwnn nntgtaaga caccgtgtaa gatgtaag 48

<210> 56
 <211> 42
 <212> ADN
 <213> Secuencia artificial

<220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"

<220>
 <221> fuente
 <223> /nota="5'-Fos modificado"

<220>
 <221> base_modificada
 <222> (7). . (11)
 <223> a, c, t, g, desconocida u otra
 35

<220>
 <221> base_modificada
 <222> (13). . (17)
 <223> a, c, t, g, desconocida u otra
 40

<220>
 <221> base_modificada
 <222> (19). . (23)
 <223> a, c, t, g, desconocida u otra
 45

<400> 56
cgatctnnnn nwnnnnnwnn nngtacagt ctccgccag tg 42

<210> 57
 <211> 45
 <212> ADN
 <213> Secuencia artificial

<220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"

<220>
 <221> fuente
 <223> /nota="5'-Fos modificado"

<220>
 <221> base_modificada
 <222> (7). . (11)
 65

<223> a, c, t, g, desconocida u otra

<220>
 <221> base_modificada
 5 <222> (13). . (17)
 <223> a, c, t, g, desconocida u otra

<220>
 <221> base_modificada
 10 <222> (19). . (23)
 <223> a, c, t, g, desconocida u otra

<400> 57
 15 **cgatctnnnn nwnnnnnwnn nnnatgatc acgtcgcgaa gtttg** 45

<210> 58
 <211> 40
 <212> ADN
 <213> Secuencia artificial

20 <220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"

25 <220>
 <221> fuente
 <223> /nota="5'-Fos modificado"

30 <220>
 <221> base_modificada
 <222> (7). . (11)
 <223> a, c, t, g, desconocida u otra

35 <220>
 <221> base_modificada
 <222> (13). . (17)
 <223> a, c, t, g, desconocida u otra

40 <220>
 <221> base_modificada
 <222> (19). . (23)
 <223> a, c, t, g, desconocida u otra

<400> 58
 45 **cgatctnnnn nwnnnnnwnn nnnaggccag tcctgatccc** 40

<210> 59
 <211> 42
 <212> ADN
 <213> Secuencia artificial

50 <220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"

55 <220>
 <221> fuente
 <223> /nota="5'-Fos modificado"

60 <220>
 <221> base_modificada
 <222> (7). . (11)
 <223> a, c, t, g, desconocida u otra

65 <220>
 <221> base_modificada

ES 2 819 277 T3

	<222> (13). . (17)		
	<223> a, c, t, g, desconocida u otra		
	<220>		
5	<221> base_modificada		
	<222> (19). . (23)		
	<223> a, c, t, g, desconocida u otra		
	<400> 59		
10	cgatctnnnn nwnnnnnwnn nnnacagggc aaggatggtg ag		42
	<210> 60		
	<211> 78		
	<212> ADN		
15	<213> Secuencia artificial		
	<220>		
	<221> fuente		
20	<223> /nota="Descripción de secuencia artificial: cebador sintético"		
	<400> 60		
	aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatctaa		60
	cctgaagggt gtcttgtg		78
	<210> 61		
25	<211> 80		
	<212> ADN		
	<213> Secuencia artificial		
	<220>		
30	<221> fuente		
	<223> /nota="Descripción de secuencia artificial: cebador sintético"		
	<400> 61		
	aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatctat		60
	caacaagact gagattgagg		80
35			
	<210> 62		
	<211> 80		
	<212> ADN		
40	<213> Secuencia artificial		
	<220>		
	<221> fuente		
	<223> /nota="Descripción de secuencia artificial: cebador sintético"		
45	<400> 62		
	aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatctct		60
	gggatcatga ctacctggag		80
	<210> 63		
50	<211> 81		
	<212> ADN		
	<213> Secuencia artificial		
	<220>		
	<221> fuente		
55	<223> /nota="Descripción de secuencia artificial: cebador sintético"		
	<400> 63		

ES 2 819 277 T3

	aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatctct	60
	gagcgtgatt tgataatgac c	81
5	<210> 64 <211> 77 <212> ADN <213> Secuencia artificial	
10	<220> <221> fuente <223> /nota="Descripción de secuencia artificial: cebador sintético"	
	<400> 64 aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatctcc	60
	cgtgtttgtg aatgagg	77
15	<210> 65 <211> 81 <212> ADN <213> Secuencia artificial	
20	<220> <221> fuente <223> /nota="Descripción de secuencia artificial: cebador sintético"	
	<400> 65 aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatcttt	60
25	gtttcctgag aggatcaaga c	81
30	<210> 66 <211> 79 <212> ADN <213> Secuencia artificial	
35	<220> <221> fuente <223> /nota="Descripción de secuencia artificial: cebador sintético"	
	<400> 66 aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatctat	60
	gtcagcgcag tcagatcac	79
40	<210> 67 <211> 75 <212> ADN <213> Secuencia artificial	
45	<220> <221> fuente <223> /nota="Descripción de secuencia artificial: cebador sintético"	
	<400> 67 aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatctag	60
50	ggtagcctgc gcttc	75
	<210> 68 <211> 79 <212> ADN <213> Secuencia artificial	

<220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"
 5
 <400> 68
aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatctac 60
tctgggtgct tctctcttc 79
 <210> 69
 <211> 78
 <212> ADN
 <213> Secuencia artificial
 10
 <220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"
 15
 <400> 69
aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatctcc 60
catccaggct aatcacac 78
 20
 <210> 70
 <211> 77
 <212> ADN
 <213> Secuencia artificial
 25
 <220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"
 30
 <400> 70
aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatctgg 60
gttggccaat ctactcc 77
 <210> 71
 <211> 84
 <212> ADN
 <213> Secuencia artificial
 35
 <220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"
 40
 <400> 71
aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatctaa 60
tgacagggag cttataattt agcc 84
 45
 <210> 72
 <211> 80
 <212> ADN
 <213> Secuencia artificial
 50
 <220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"
 <400> 72

ES 2 819 277 T3

	aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatctac	60
	atattcagct ggcacagtta	80
5	<210> 73 <211> 82 <212> ADN <213> Secuencia artificial	
10	<220> <221> fuente <223> /nota="Descripción de secuencia artificial: cebador sintético"	
	<400> 73 aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatcttg	60
	aaacacacct gaatacctac ag	82
15	<210> 74 <211> 77 <212> ADN <213> Secuencia artificial	
20	<220> <221> fuente <223> /nota="Descripción de secuencia artificial: cebador sintético"	
	<400> 74 aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatctac	60
25	agggcaggca tggtatc	77
30	<210> 75 <211> 78 <212> ADN <213> Secuencia artificial	
35	<220> <221> fuente <223> /nota="Descripción de secuencia artificial: cebador sintético"	
	<400> 75 aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatctgt	60
	ggtttggatc gacgtctc	78
40	<210> 76 <211> 77 <212> ADN <213> Secuencia artificial	
45	<220> <221> fuente <223> /nota="Descripción de secuencia artificial: cebador sintético"	
	<400> 76 aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatctgg	60
50	ccttcaaaga gcacctg	77
	<210> 77 <211> 76 <212> ADN <213> Secuencia artificial	

ES 2 819 277 T3

<220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"
 5
 <400> 77
aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatctct 60
accagctgc tcatgc 76
 <210> 78
 <211> 78
 <212> ADN
 <213> Secuencia artificial
 10
 <220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"
 15
 <400> 78
aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatctca 60
ggaacaaat gccaaagtg 78
 20
 <210> 79
 <211> 77
 <212> ADN
 <213> Secuencia artificial
 25
 <220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"
 30
 <400> 79
aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatctga 60
aggaaggaa ggaagg 77
 <210> 80
 <211> 81
 <212> ADN
 <213> Secuencia artificial
 35
 <220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"
 40
 <400> 80
aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatctga 60
ttctcttgat gatgctgatg c 81
 45
 <210> 81
 <211> 79
 <212> ADN
 <213> Secuencia artificial
 50
 <220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"
 <400> 81

	aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatcttt	60
	ctcttgatga tgctgatgc	79
5	<210> 82 <211> 82 <212> ADN <213> Secuencia artificial	
10	<220> <221> fuente <223> /nota="Descripción de secuencia artificial: cebador sintético"	
	<400> 82 aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatctcc	60
	ttccaaatct ctaccctcta tc	82
15	<210> 83 <211> 82 <212> ADN <213> Secuencia artificial	
20	<220> <221> fuente <223> /nota="Descripción de secuencia artificial: cebador sintético"	
	<400> 83 aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatctag	60
25	taaagtcaca taacctctaa cc	82
30	<210> 84 <211> 80 <212> ADN <213> Secuencia artificial	
35	<220> <221> fuente <223> /nota="Descripción de secuencia artificial: cebador sintético"	
	<400> 84 aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatctag	60
	agttggtaag gaggagaatg	80
40	<210> 85 <211> 82 <212> ADN <213> Secuencia artificial	
45	<220> <221> fuente <223> /nota="Descripción de secuencia artificial: cebador sintético"	
	<400> 85 aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatctct	60
50	gtggtatctg aactatcttc tc	82
	<210> 86 <211> 78 <212> ADN <213> Secuencia artificial	

	<220>		
	<221> fuente		
	<223> /nota="Descripción de secuencia artificial: cebador sintético"		
5	<400> 86		
	aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatctca		60
	tttgtggagc accttctg		78
	<210> 87		
10	<211> 82		
	<212> ADN		
	<213> Secuencia artificial		
	<220>		
15	<221> fuente		
	<223> /nota="Descripción de secuencia artificial: cebador sintético"		
	<400> 87		
	aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatctaa		60
	ggatagtttg gaactgagag ac		82
20	<210> 88		
	<211> 77		
	<212> ADN		
	<213> Secuencia artificial		
25	<220>		
	<221> fuente		
	<223> /nota="Descripción de secuencia artificial: cebador sintético"		
30	<400> 88		
	aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatcttg		60
	acctggccaa gaagaag		77
	<210> 89		
35	<211> 79		
	<212> ADN		
	<213> Secuencia artificial		
	<220>		
40	<221> fuente		
	<223> /nota="Descripción de secuencia artificial: cebador sintético"		
	<400> 89		
	aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatctaa		60
45	tcggaactgg aggaacaac		79
	<210> 90		
	<211> 21		
	<212> ADN		
50	<213> Secuencia artificial		
	<220>		
	<221> fuente		
	<223> /nota="Descripción de secuencia artificial: cebador sintético"		
55	<400> 90		
	catgccgcca cccccaccct a		21

<210> 91
 <211> 41
 <212> ADN
 <213> Secuencia artificial
 5
 <220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"
 10
 <220>
 <221> base_modificada
 <222> (1). . (5)
 <223> a, c, t, g, desconocida u otra
 15
 <220>
 <221> base_modificada
 <222> (7). . (11)
 <223> a, c, t, g, desconocida u otra
 20
 <220>
 <221> base_modificada
 <222> (13). . (17)
 <223> a, c, t, g, desconocida u otra
 25
 <400> 91
nnnnnwnnnn nwnnnntgg gggtggctcc aggaaaattg g 41
 <210> 92
 <211> 33
 <212> ADN
 <213> Secuencia artificial
 30
 <220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"
 35
 <220>
 <221> base_modificada
 <222> (20). . (21)
 <223> a, c, t, g, desconocida u otra
 40
 <220>
 <221> base_modificada
 <222> (23). . (27)
 <223> a, c, t, g, desconocida u otra
 45
 <220>
 <221> base_modificada
 <222> (29). . (33)
 <223> a, c, t, g, desconocida u otra
 50
 <400> 92
catgccgccca cccccaccn nwnnnnnwnn nnn 33
 55
 <210> 93
 <211> 73
 <212> ADN
 <213> Secuencia artificial
 60
 <220>
 <221> fuente
 <223> /nota="Descripción de secuencia artificial: cebador sintético"
 <400> 93

	catgccgcca cccccacccg atcagtc AAC cggagatcgg aagagcacac gtctgaactc	60
	cagtcacccg tcc	73
5	<210> 94 <211> 19 <212> ADN <213> Secuencia artificial	
10	<220> <221> fuente <223> /nota="Descripción de secuencia artificial: oligonucleótido sintético"	
	<400> 94 cccgcccgcc gcccgccgc	19
15	<210> 95 <211> 26 <212> ADN <213> Secuencia artificial	
20	<220> <221> fuente <223> /nota="Descripción de secuencia artificial: oligonucleótido sintético"	
25	<400> 95 cccgcccgcc gcccgccgcc cgccgc	26
30	<210> 96 <211> 24 <212> ADN <213> Secuencia artificial	
35	<220> <221> fuente <223> /nota="Descripción de secuencia artificial: oligonucleótido sintético"	
	<400> 96 atgtgtgtgt gtgtgtgtgt gtgt	24
40	<210> 97 <211> 11 <212> ADN <213> Secuencia artificial	
45	<220> <221> fuente <223> /nota="Descripción de secuencia artificial: oligonucleótido sintético"	
50	<400> 97 gccagcacca g	11
55	<210> 98 <211> 211 <212> ADN <213> Homo sapiens	
	<400> 98 actacctttg tctctagata cccctgattc accgagccct gcagttggcc cagcgtcccg	60
	tttactcct tgccagcccc tggacatcac ccaactggct caagaccaat ggagcgggtga	120
	atgggaaggg gtcactcaag ggacagcccg gagacatcta ccaccagacc tgggccagat	180
	actttgtgaa gtaagggatc agcaaggatg t	211

ES 2 819 277 T3

<210> 99
 <211> 211
 <212> ADN
 5 <213> Homo sapiens

 <400> 99
 actacctttg tctctagata cccctgattc accgagccct gcagttggcc cagcgtcccg 60
 tttcactcct tgccagcccc tggacatcac ccacttggct caagaccaat ggagcgggga 120
 atgggaaggg gtcactcaag ggacagcccg gagacatcta ccaccagacc tgggcccagat 180
 actttgtgaa gtaagggatc agcaaggatg t 211

 10 <210> 100
 <211> 211
 <212> ADN
 <213> Homo sapiens

 15 <400> 100
 actacctttg tctctagata cccctgattc accgagccct gcagttggcc cagcgtcccg 60
 tttcactcct tgccagcccc tggacatcac ccactcggct caagaccaag ggagcgggga 120
 atgggaaggg gccactcaag ggacagccca gagacatcta ccaccagacc tgggcccagat 180
 acattgtgaa gtaagggatc agcaaggatg t 211

REIVINDICACIONES

1. Un procedimiento para generar una biblioteca de polinucleótidos que comprende:

5 (a) generar una primera secuencia del complemento (CS) de un polinucleótido diana a partir de una muestra usando un primer cebador, comprendiendo el primer cebador una secuencia específica de diana;

10 (b) unir por medio de ligadura a la primera CS un adaptador que comprende una primera secuencia de unión de cebador (PBS) o una parte de la misma, formando de este modo una secuencia del complemento modificada (MCS);

(c) extender un segundo cebador hibridado a la MCS, formando de este modo una segunda CS, en el que el segundo cebador comprende:

15 (i) una región específica de diana, y

(ii) una segunda PBS; y

20 (d) amplificar la segunda CS usando cebadores que se hibridan con la primera PBS y la segunda PBS respectivamente, en el que el primer o el segundo cebador comprende una secuencia de identificación única (UID).

2. El procedimiento de una cualquiera de las reivindicaciones 1, en el que el primer cebador comprende una secuencia de ligadura universal (ULS).

25 3. El procedimiento de una cualquiera de las reivindicaciones 1-2, en el que el segundo cebador comprende además una secuencia de cebado universal (UPS).

30 4. El procedimiento de una cualquiera de las reivindicaciones 1-3, en el que el adaptador comprende además una secuencia de código de barras de muestra (SBC).

5. El procedimiento de una cualquiera de las reivindicaciones 1-4, en el que la MCS comprende además una molécula de afinidad o secuencia de captura.

35 6. El procedimiento de una cualquiera de las reivindicaciones 1-5, en el que la UID comprende la secuencia NNNNNNNNNNNNNNNN (SEQ ID NO: 1), en la que N es cualquier residuo de ácido nucleico.

40 7. El procedimiento de una cualquiera de las reivindicaciones 1-5, en el que la UID comprende la secuencia NNNNNWNNNNNNWNNNNN (SEQ ID NO: 2), en la que N es cualquier residuo de ácido nucleico y W es adenina o timina.

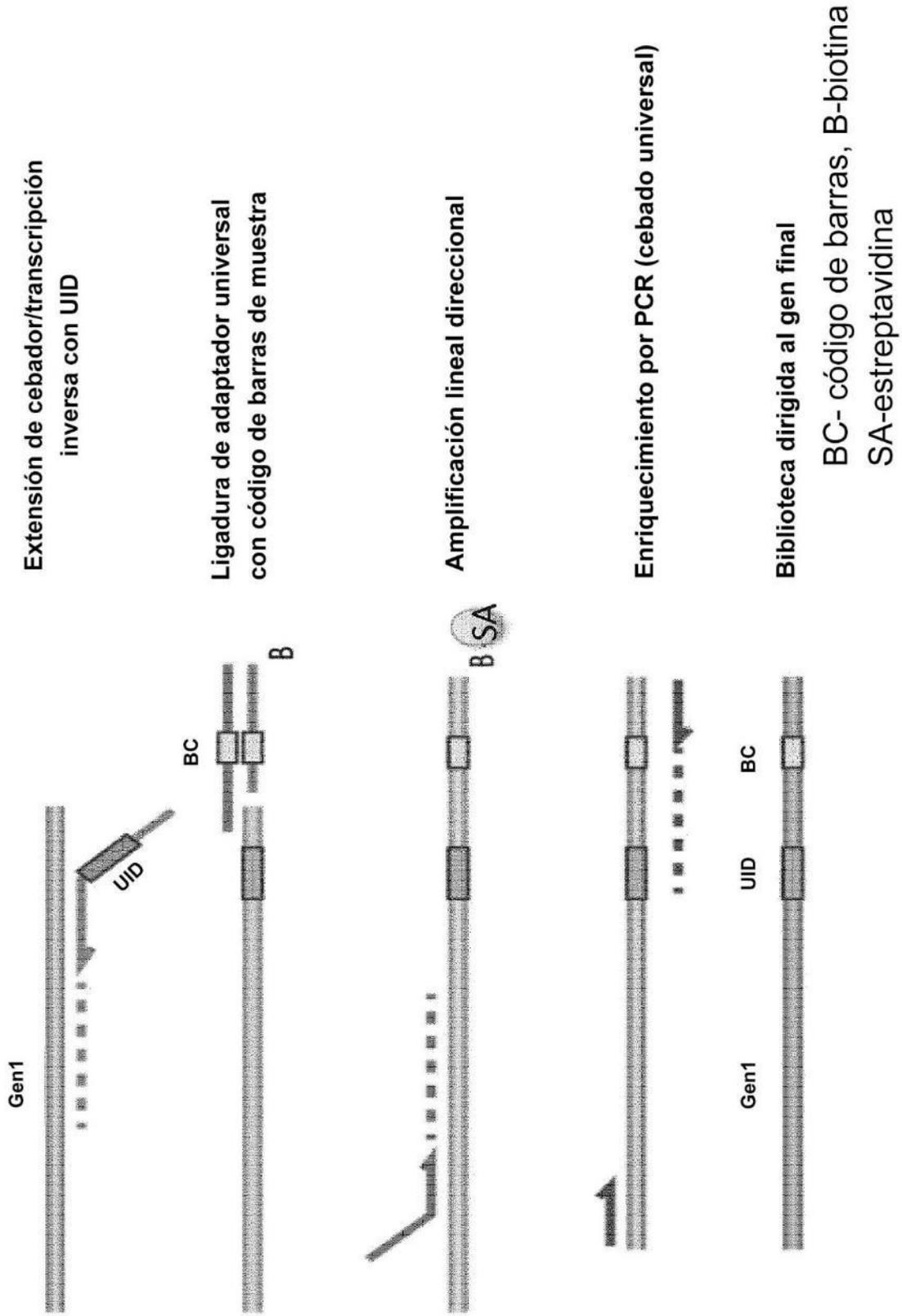


Fig. 1

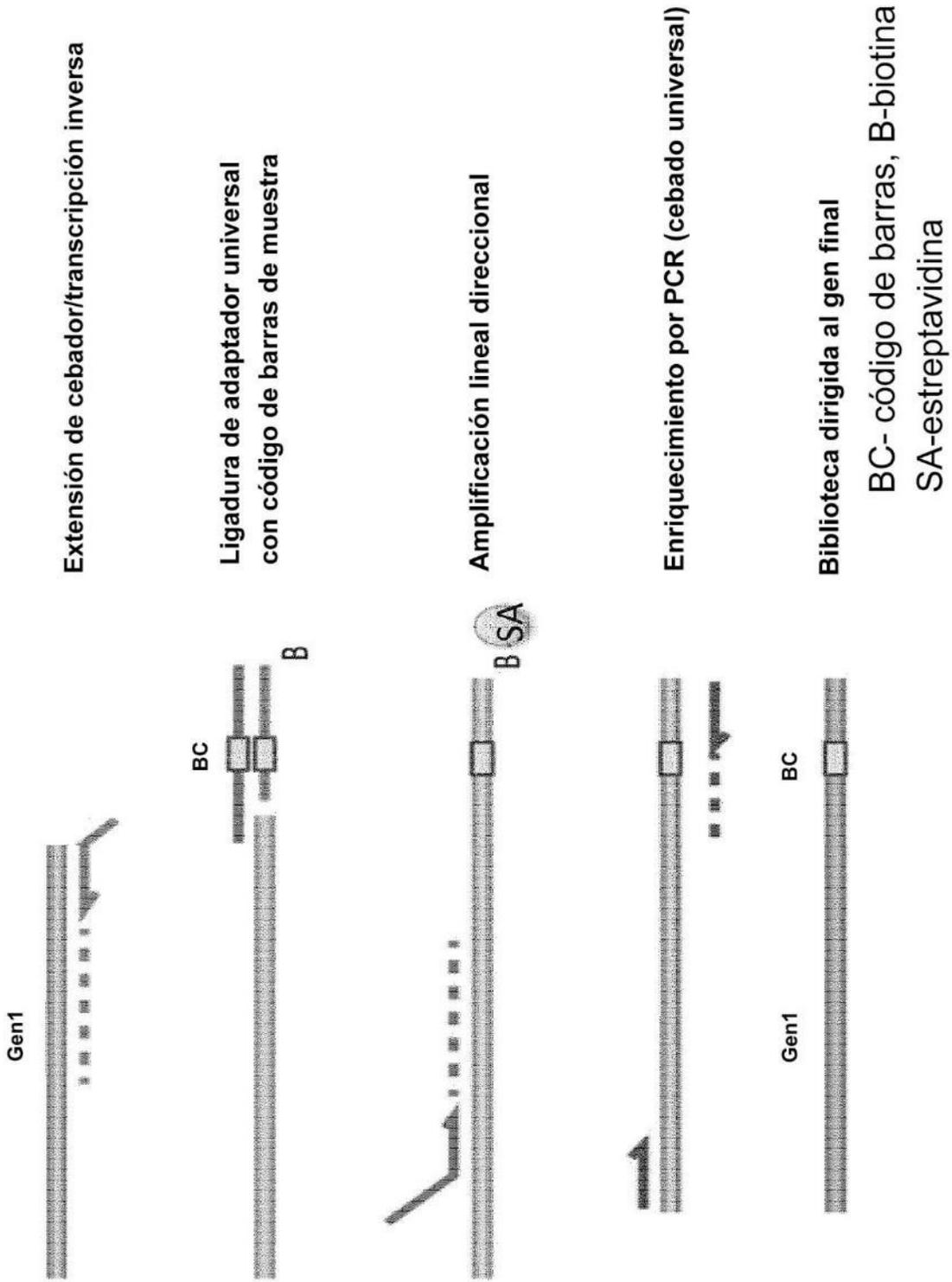
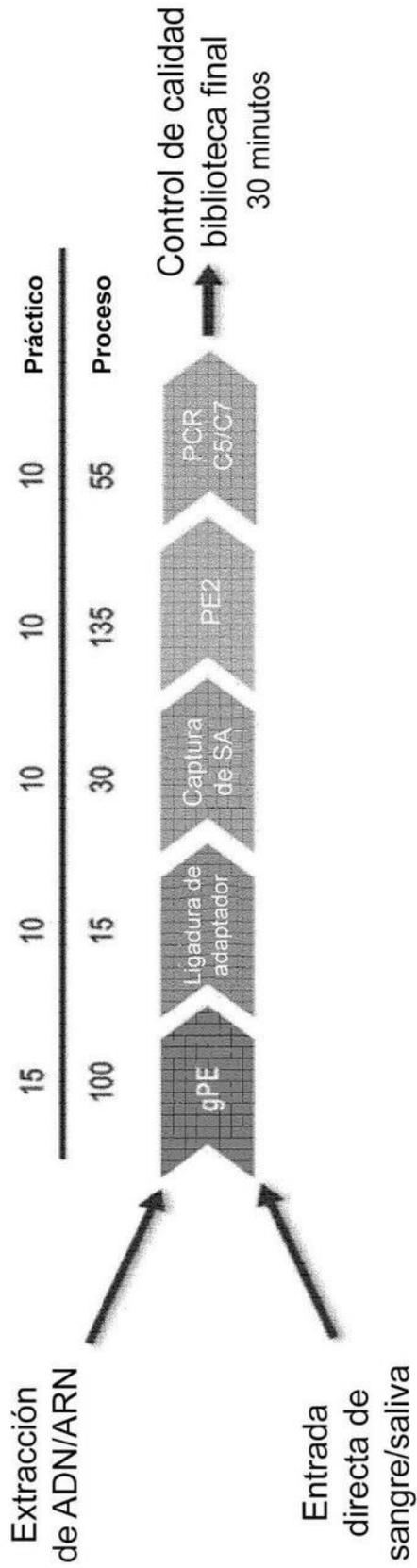


Fig. 2



Proveedor	Producto	Tiempo de práctica (horas)	Tiempo de proceso
Life Technologies	Ion Ampliseq™	2,5	6
Illumina	Amplicon Truseq™	2,5	7
Agilent	Haloplex™	8	12-24
AbViro	Direccionamiento de extensión de cebador	1	6

Fig. 3

Panel	Ampliciones	% en la diana
CS-350*	346	92,20 %
Tex-1	336	97,10 %
CS-30*	30	94,20 %
CS-23	23	99,20 %
Otro N.º 1	207	90 %
Otro N.º 2	220	93 %

Fig. 4

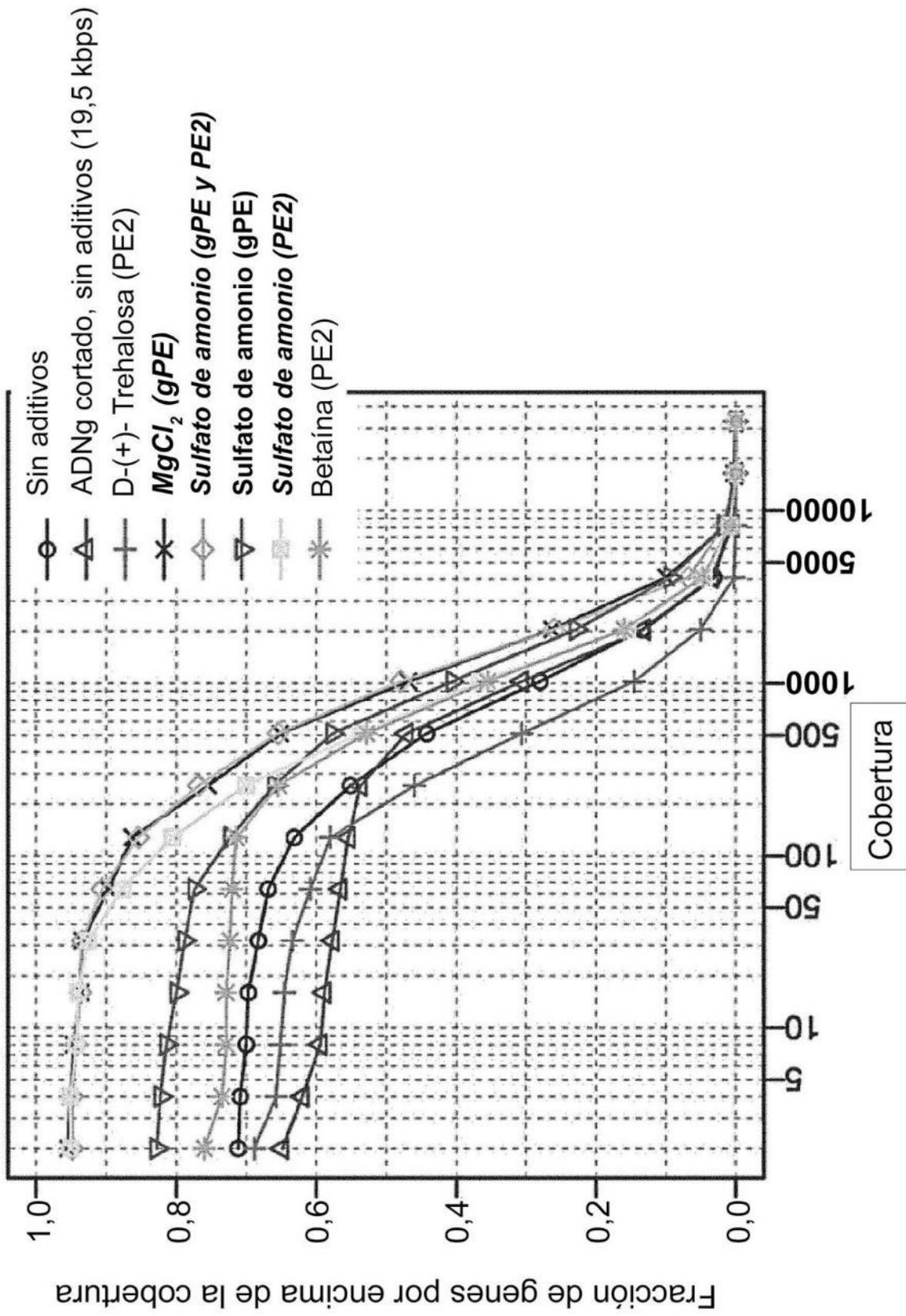


Fig. 5

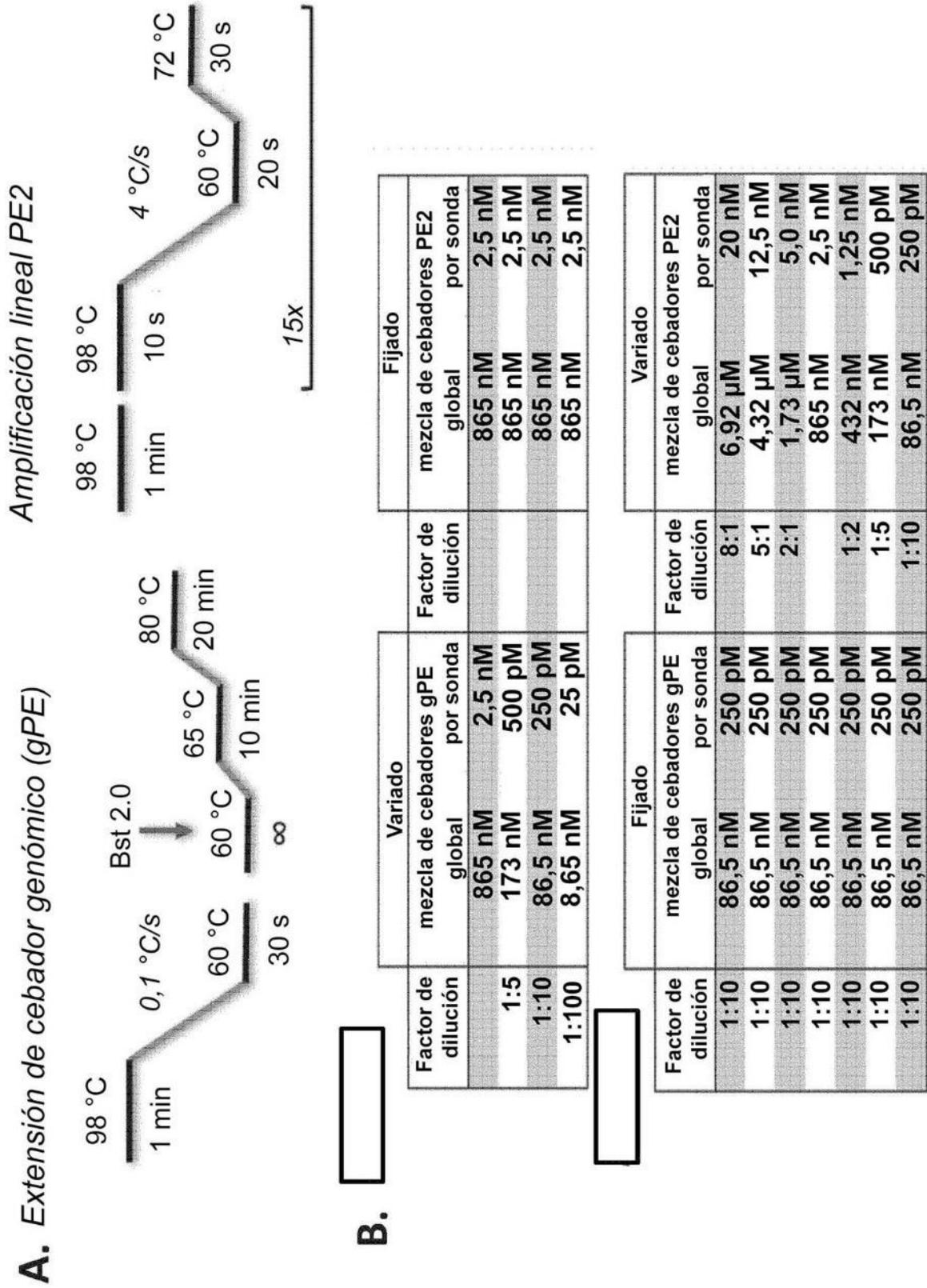
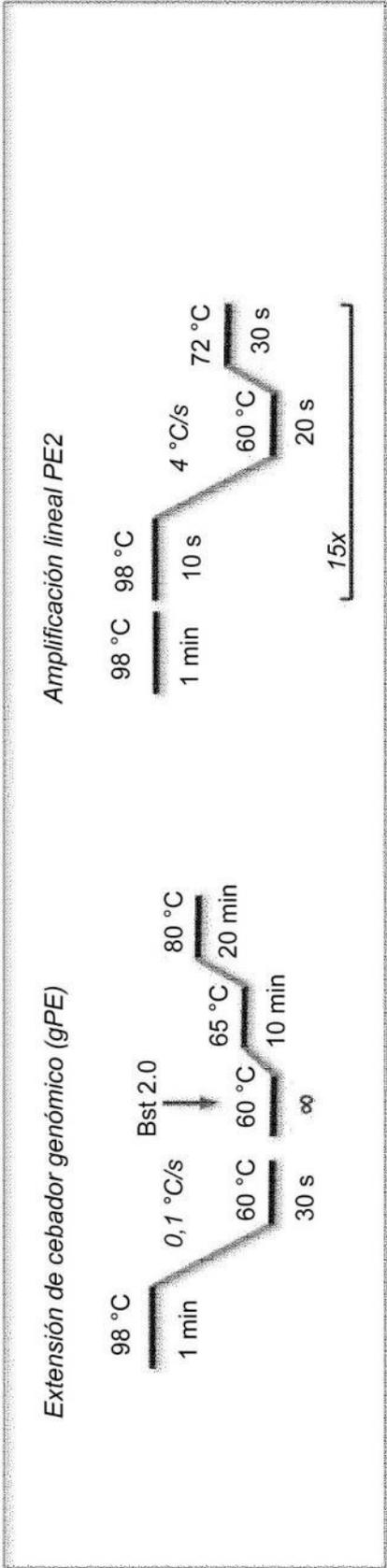


Fig. 6

Menos riguroso



Más riguroso

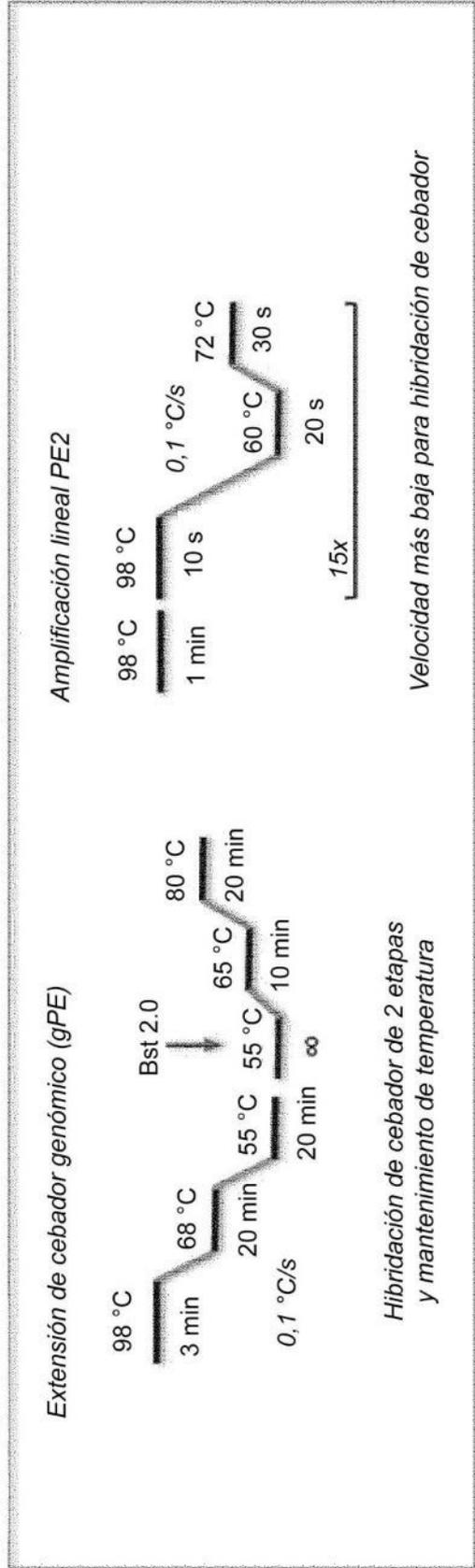


Fig. 7

Tamaño del panel	Amplicones	Grupo 1				Grupo 2			
		Fijado por concentración de sonda				Concentración global fija			
		Concentración final*		RESULTADO		Concentración final*		RESULTADO	
global	por sonda	¿Productos?		global	por sonda	¿Productos?			
Completo	346	86,5 nM	250 pM	NO	480 nM	1,38 nM	Sí, leve		
Mitad	173	43,2 nM	250 pM	NO	480 nM	2,77 nM	Sí, leve		
Cuarto	86	21,5 nM	250 pM	Sí, leve	480 nM	5,58 nM	Sí, fuerte		
Pequeño	24	6,0 nM	250 pM	Sí, leve	480 nM	20 nM	Sí, fuerte		

* Se refiere a concentraciones de cebador tanto en gPE como en PE2

Fig. 8

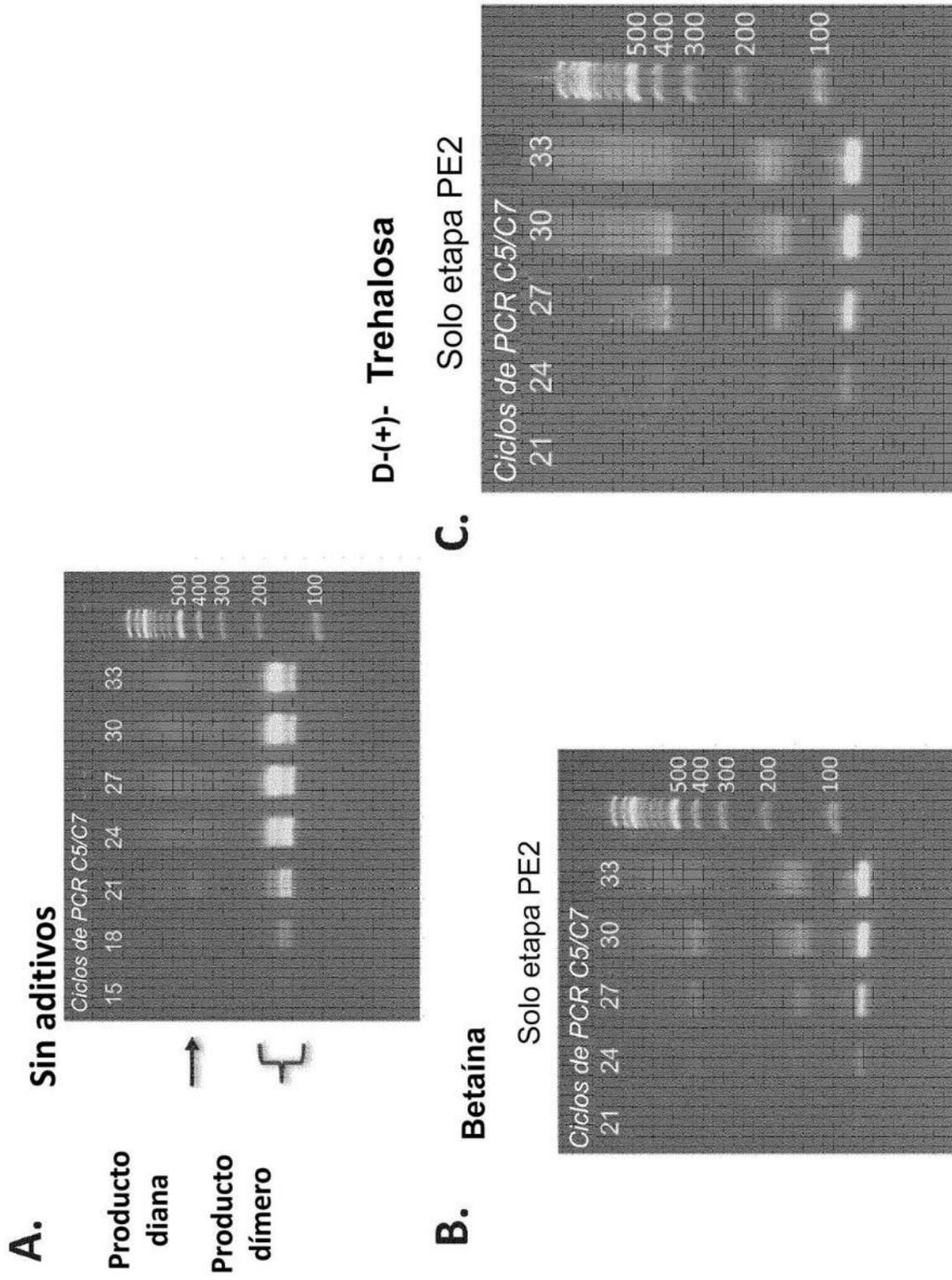
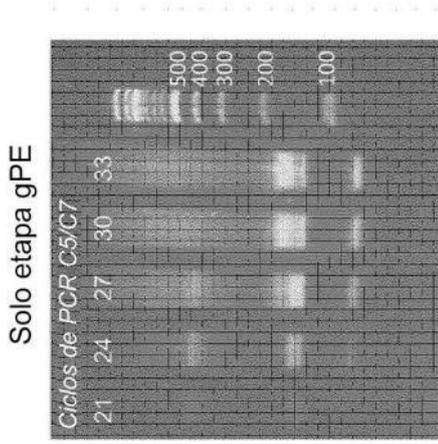


Fig. 9

D.

Cloruro de magnesio, $MgCl_2$



E.

Sulfato de amonio, $(NH_4)_2SO_4$

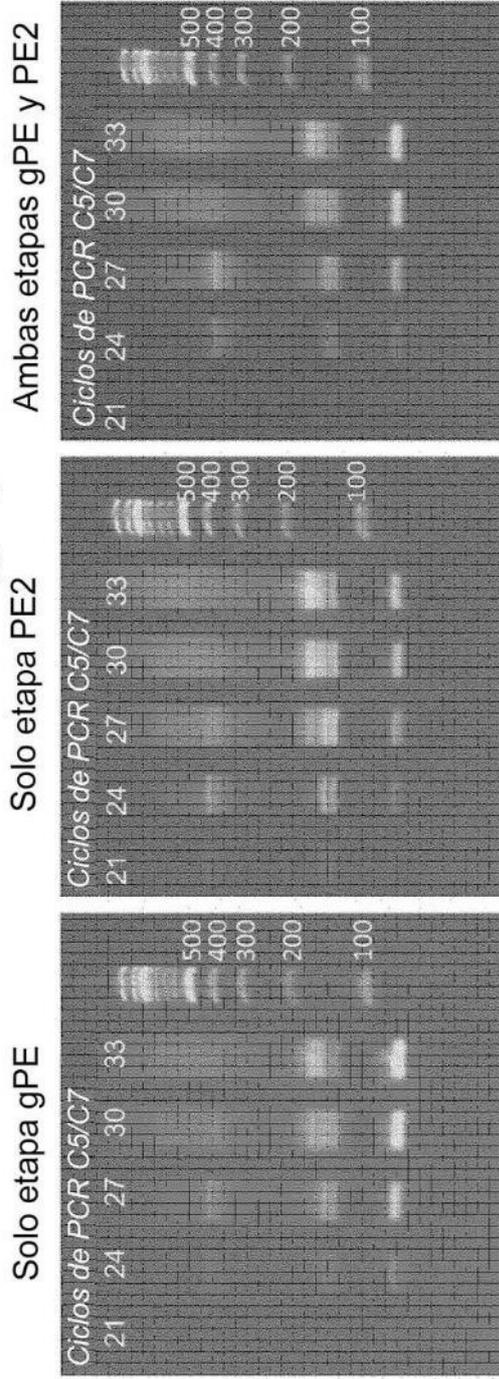


Fig. 9 (Cont)

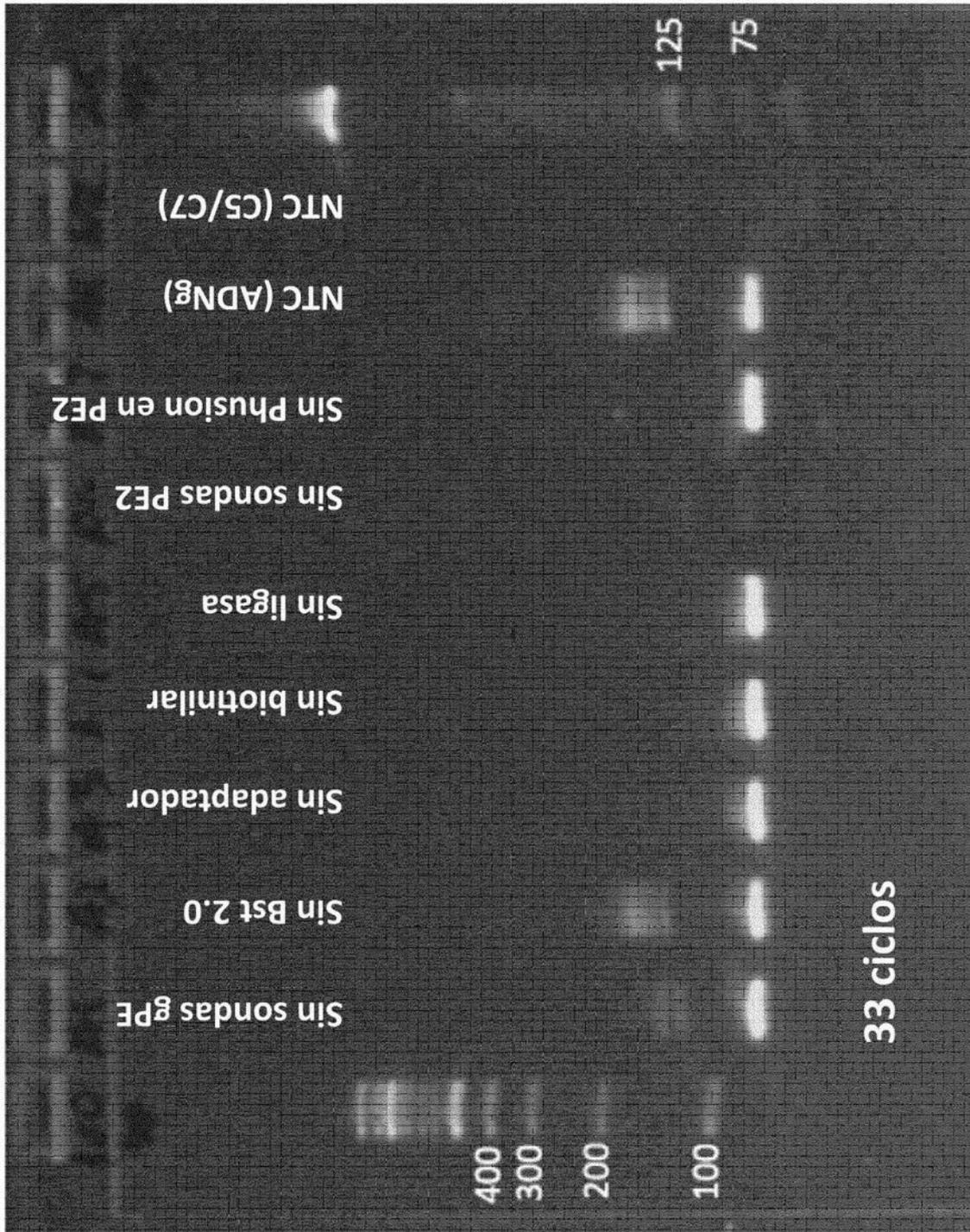


Fig. 10

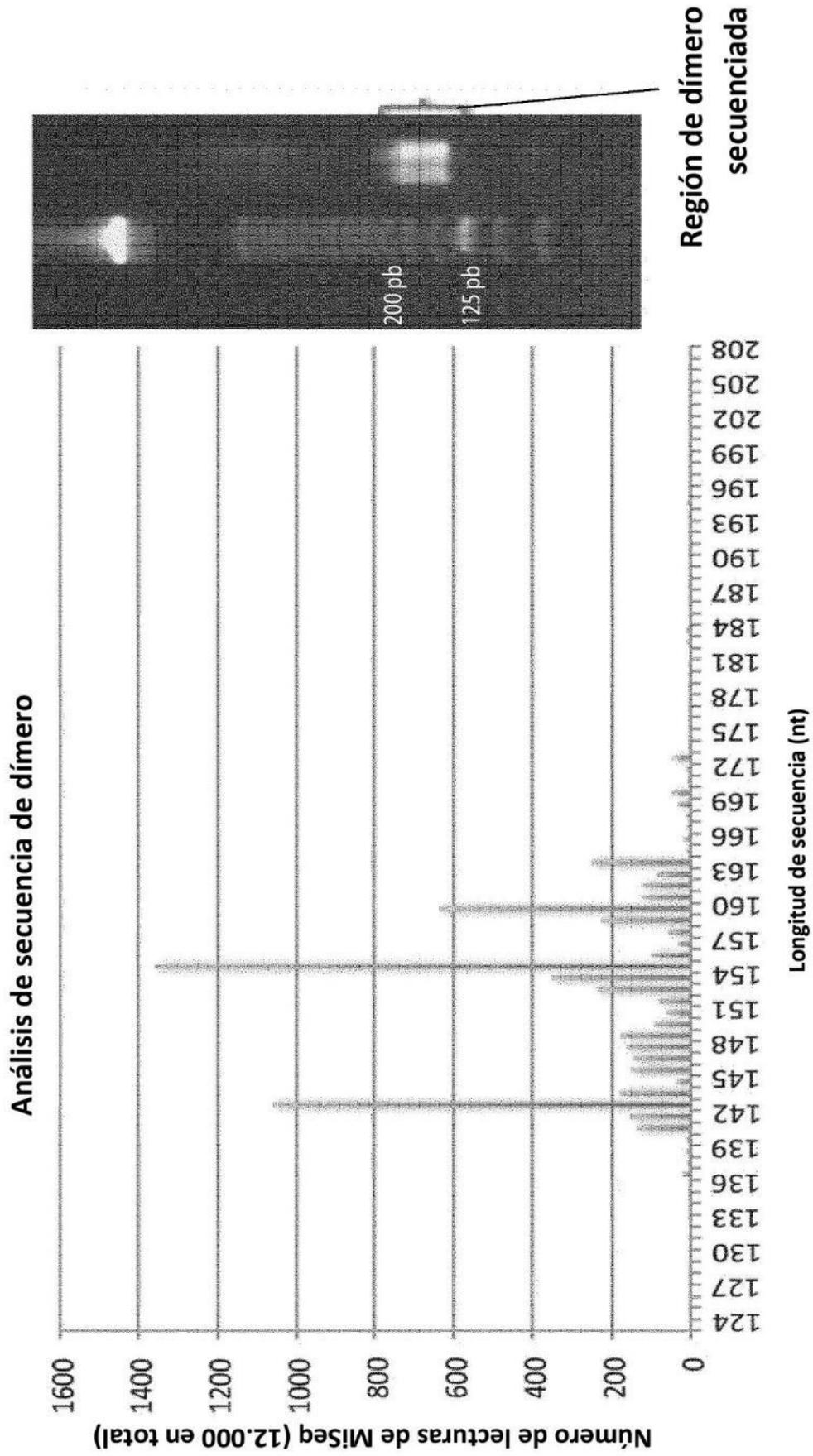


Fig. 11

Clave de color
GAA_14_1_o_PE2_7
POMGNT1_21_-1_o_gPE_5

Exactitud 3'

C5 – P5 – CATGCCGCACCCACCCCTA
 GGTAAAAGGACCTCGGTGGGGTNNNNNNNNNNNNNNNN
 (GGG; 1/64 UID tendrán GGG en esta posición)

Extensión

C5 – P5 – CATGCCGCACCCACCCNNNNNNNNNNNN
 GGTAAAAGGACCTCGGTGGGGTNNNNNNNNNNNNNNNN

Dímero no predicho por software: combinación de dímero de cebador/complementariedad de UID

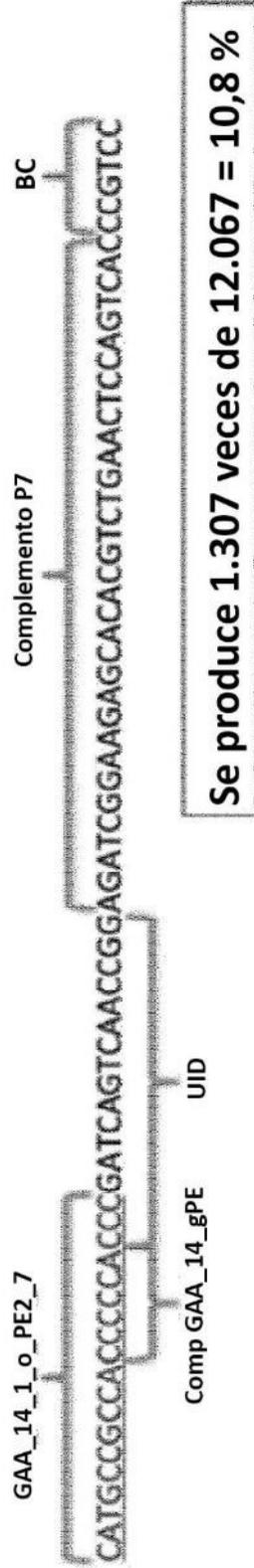


Fig. 12

Enfermedad	Gen	Exones	Conjuntos de sondas
Alfa-talasemia	HBA1	3	3
Alfa-talasemia	HBA2	3	2
Beta-talasemia	HBB	3	3
Síndrome de Bloom	BLM	22	24
Enfermedad de Canavan	ASPA	6	6
Fibrosis quística	CFTR	27	31
Deficiencia de dihidropirimidina deshidrogenasa	DLD	14	12
Disautonomía familiar	IKBKAP	37	39
Anemia de Fanconi tipo C	FANCC	15	16
Síndrome X frágil (sin ngs)	FMR1	17	17
Enfermedad de Gaucher	GBA	11	12
Enfermedad de almacenamiento de glucógeno tipo 1a	G6PC	5	6
Síndrome de Joubert 2	TMEM237	13	13
Enfermedad de Krabbe	GALC	17	18
Enfermedad de la orina con olor a jarabe de arce tipo 1B	BCKDHB	10	11
Mucopolipidosis IV	MCOLN1	14	13
Enfermedad muscular-ocular-cerebral (síndrome de Walker-Warburg)	POMGNT1	22	22
Enfermedad de Niemann Pick tipo A/B	SMPD1	6	11
Enfermedad de Pompe	GAA	20	22
Atrofia muscular espinal	SMN1	9	8
Enfermedad de Tay-Sachs	HEXA	14	17
Síndrome de Usher tipo 1F	PCDH15	33	35
Síndrome de Usher tipo 3	CLRN1	3	5
22	23	324	346

Exones sin cebador

- SMN1_04
- HBB_03
- ASPA_06
- DLD_01-01
- DLD_05
- DLD_14
- BLM_03-02
- BLM_16
- BLM_22
- MCOLN1_03
- HEXA_14
- POMGNT1_06
- PCDH15_33
- CFTR_04
- CFTR_08-01
- CFTR_14-01

Fig. 13

Criterios de exclusión

Subpaneles CS-350

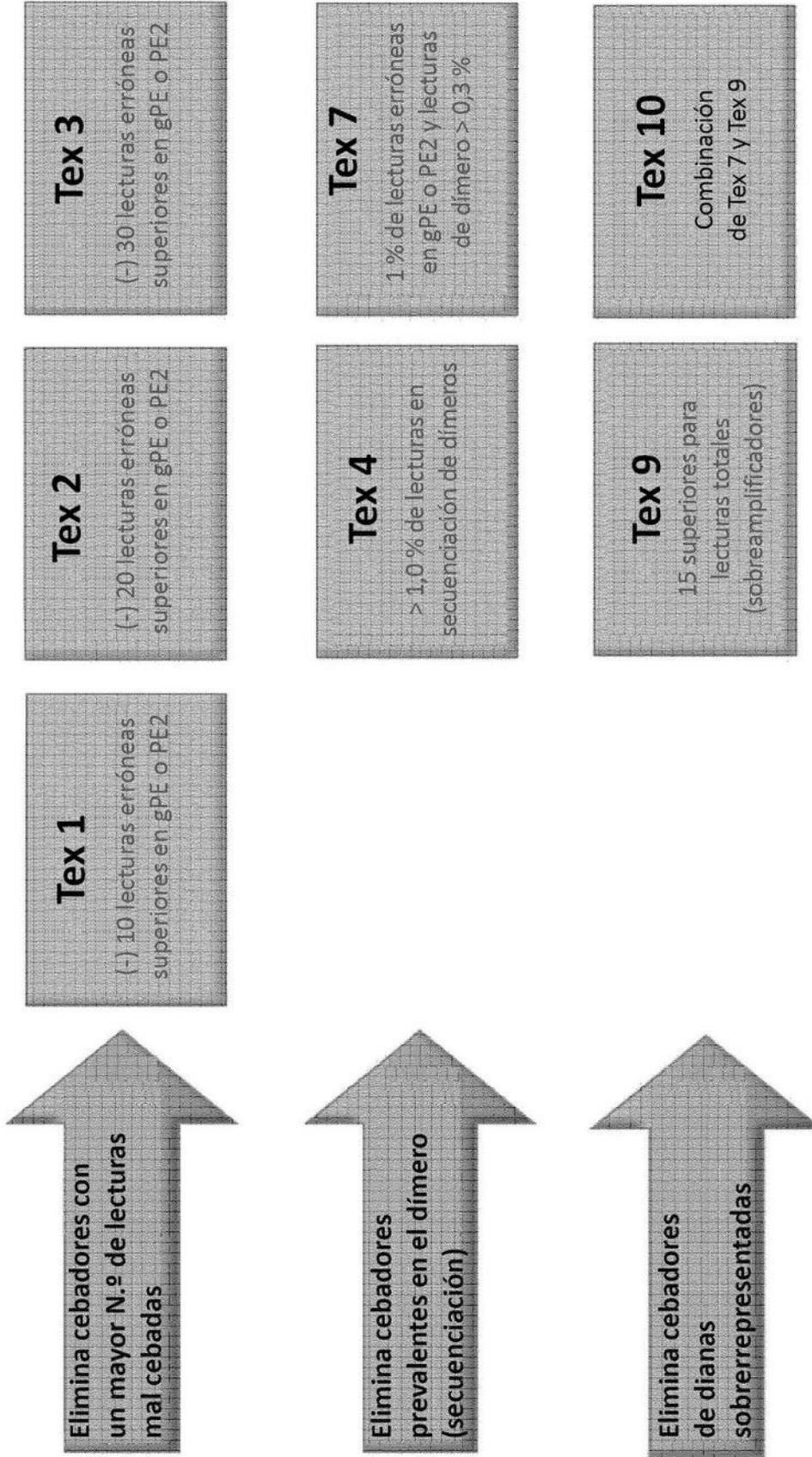


Fig. 14

Especificidad en la diana y uniformidad de cobertura

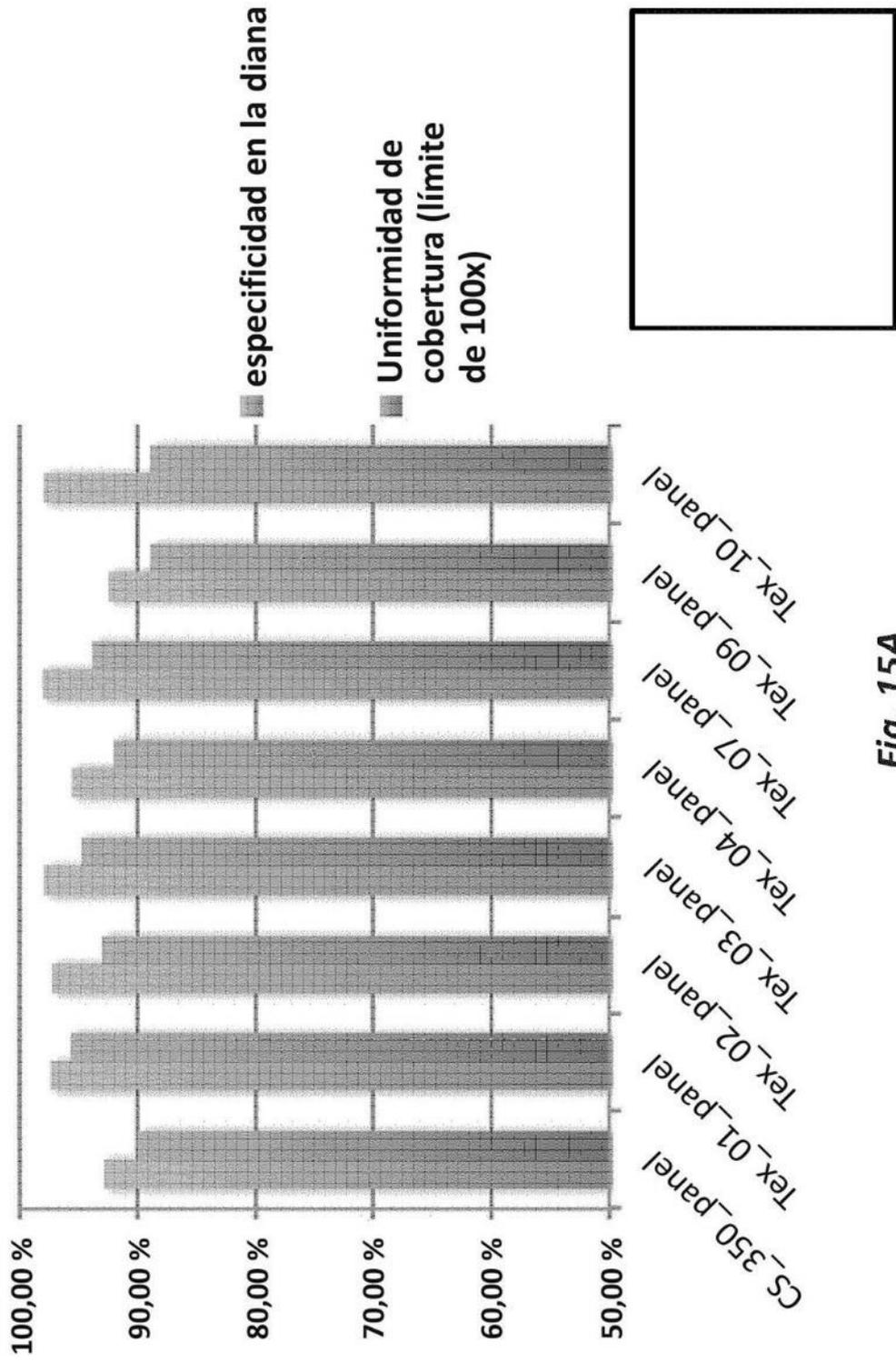


Fig. 15A

panel	N.º de amplicones	profundidad de lectura media por amplicón	especificidad en la diana	Uniformidad de cobertura (límite de 100x)
CS_350_panel	346	1339	92,69 %	89,88 %
Tex_01_panel	336	1404	97,20 %	95,54 %
Tex_02_panel	326	1403	97,11 %	92,94 %
Tex_03_panel	316	1412	97,78 %	94,62 %
Tex_04_panel	335	1378	95,42 %	91,94 %
Tex_07_panel	303	1415	97,92 %	93,73 %
Tex_09_panel	332	1333	92,31 %	88,86 %
Tex_10_panel	295	1414	97,85 %	88,81 %

Fig. 15B

Uniformidad de cobertura: profundidad de lectura sin procesar en subpaneles de cebador

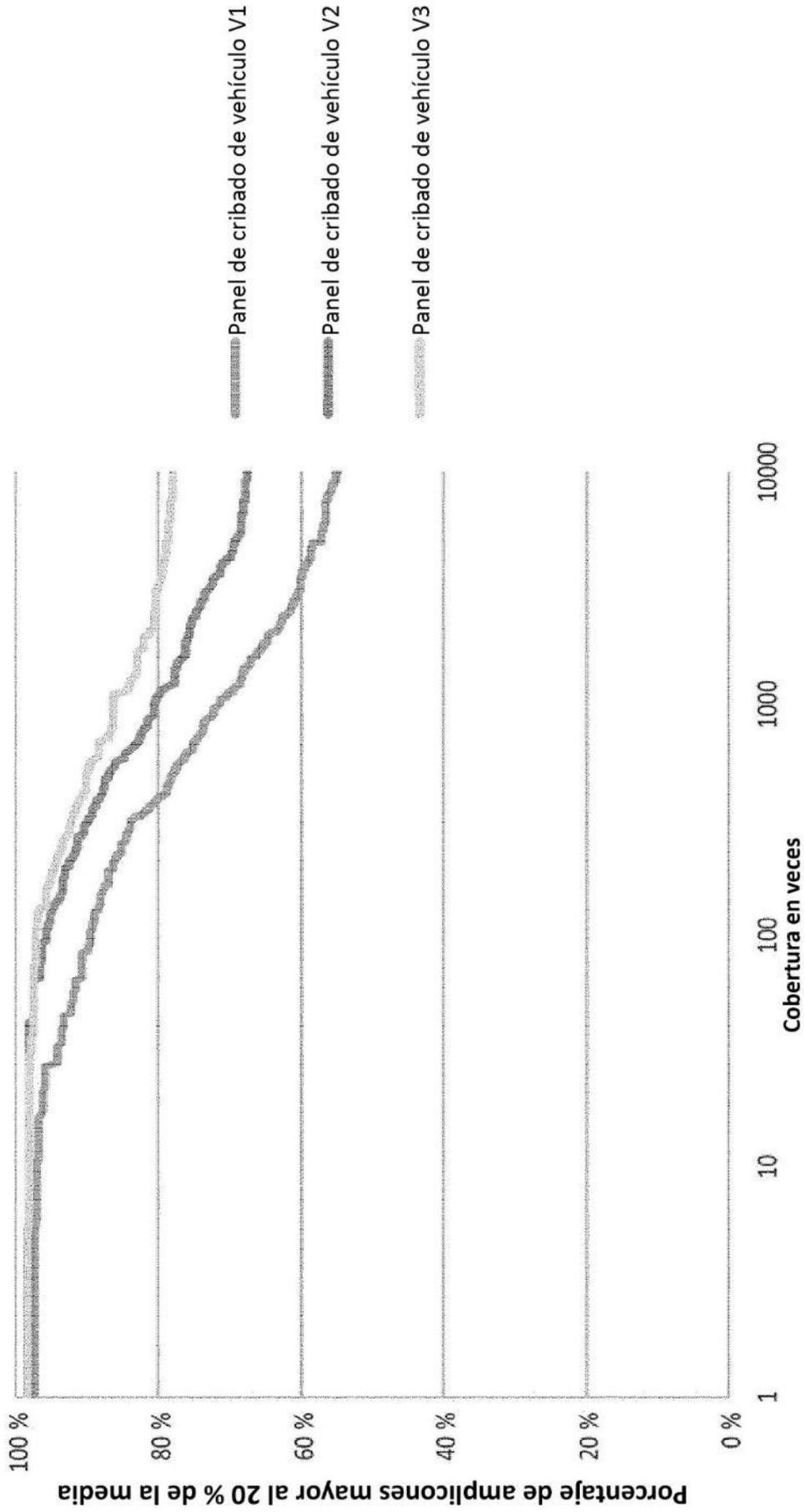
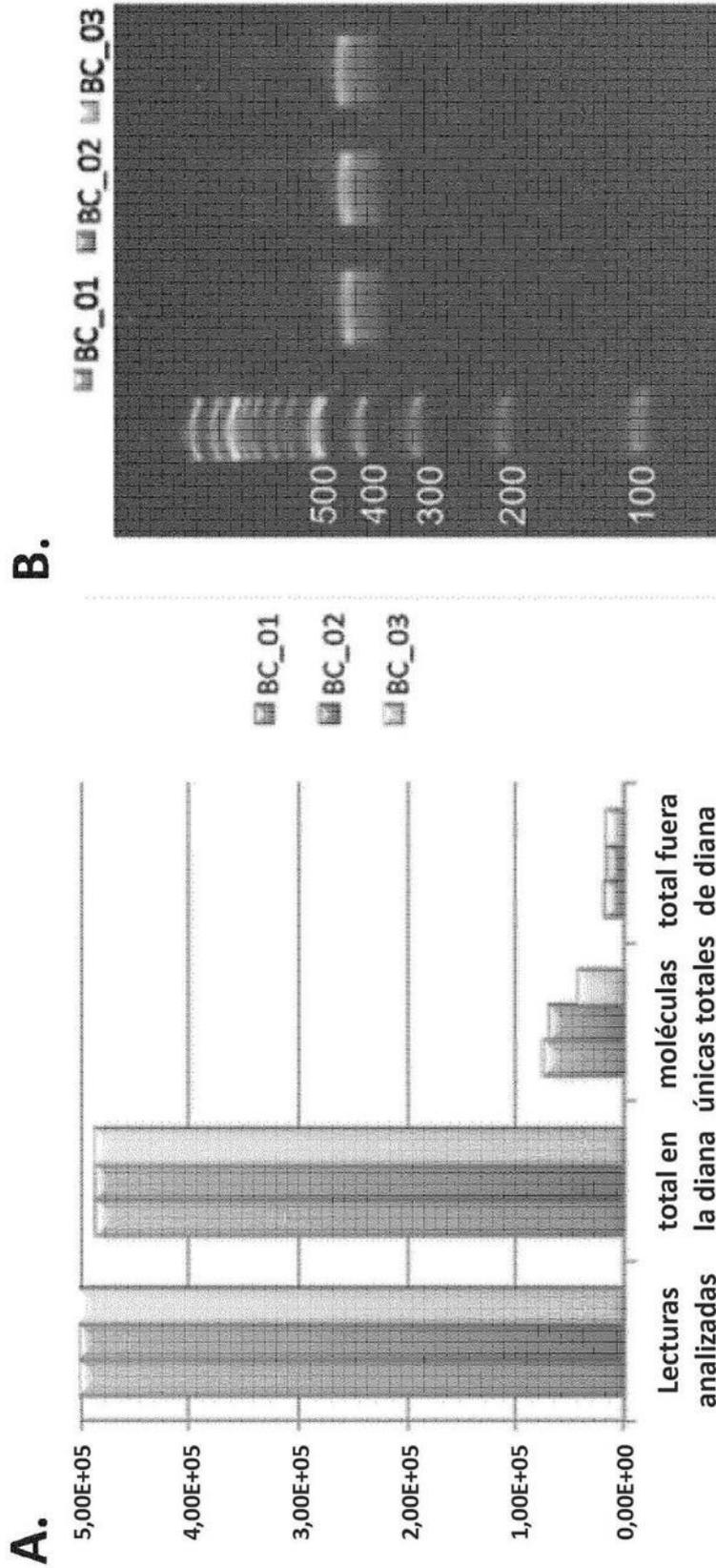


Fig. 16



C.

	BC_01	BC_02	BC_03	Media	Descripción del análisis
Lecturas analizadas	5,00E+05	5,00E+05	5,00E+05	5,00E+05	clasificación aleatoria in silico y selección de 500.000 secuencias
total en la diana	4,84E+05	4,85E+05	4,85E+05	4,85E+05	lecturas totales con cebadores gPE y PE2 coincidentes
moléculas únicas totales	7,29E+04	6,86E+04	4,17E+04	6,11E+04	moléculas únicas totales en el grupo de secuencias emparejadas correctamente
total fuera de diana	1,63E+04	1,48E+04	1,47E+04	1,53E+04	lecturas totales con cebadores gPE y PE2 no coincidentes
tasa en la diana	96,74 %	97,04 %	97,05 %	96,94 %	total en la diana/(total en la diana + total fuera de diana)

Fig. 17

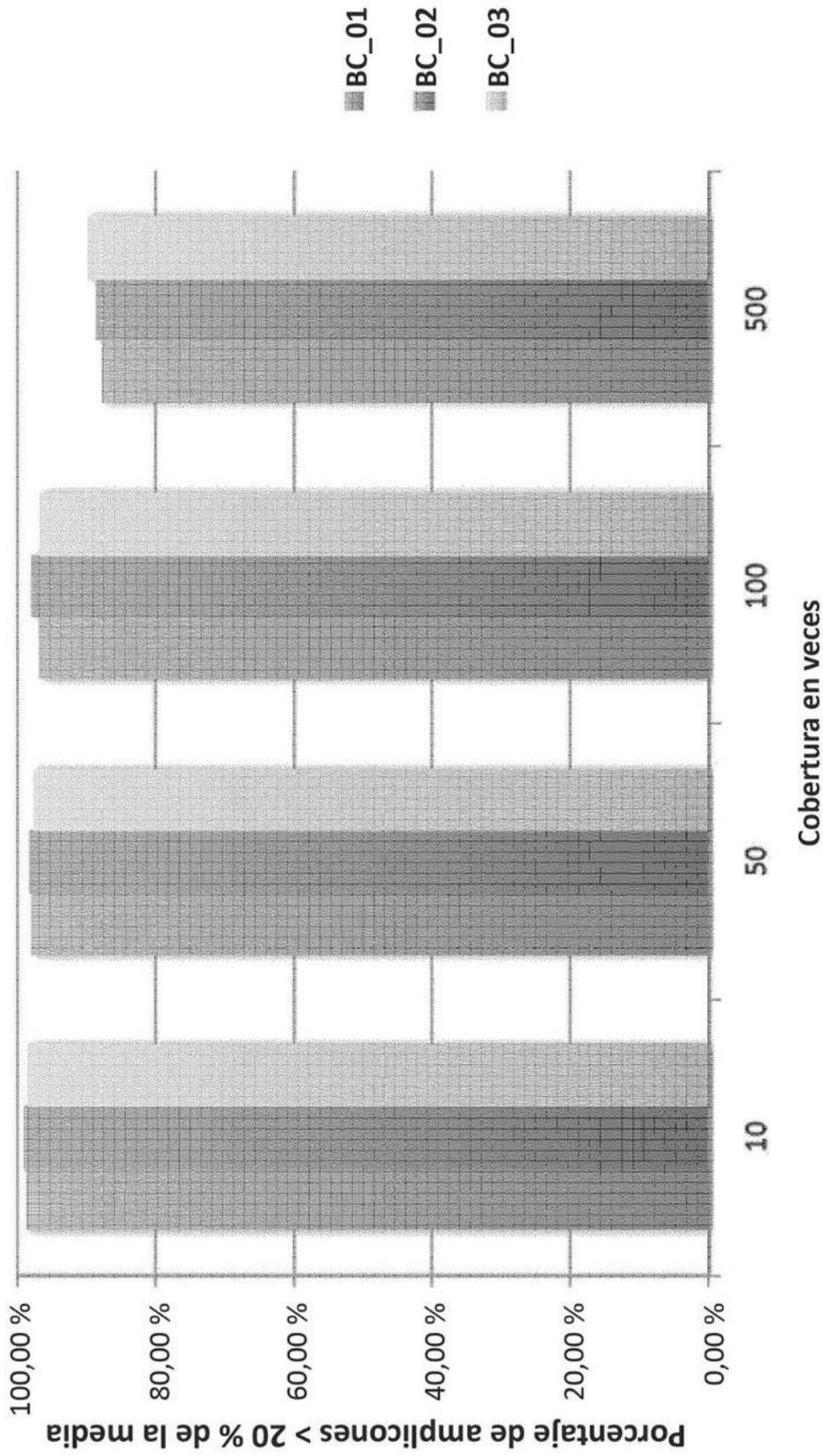


Fig. 18

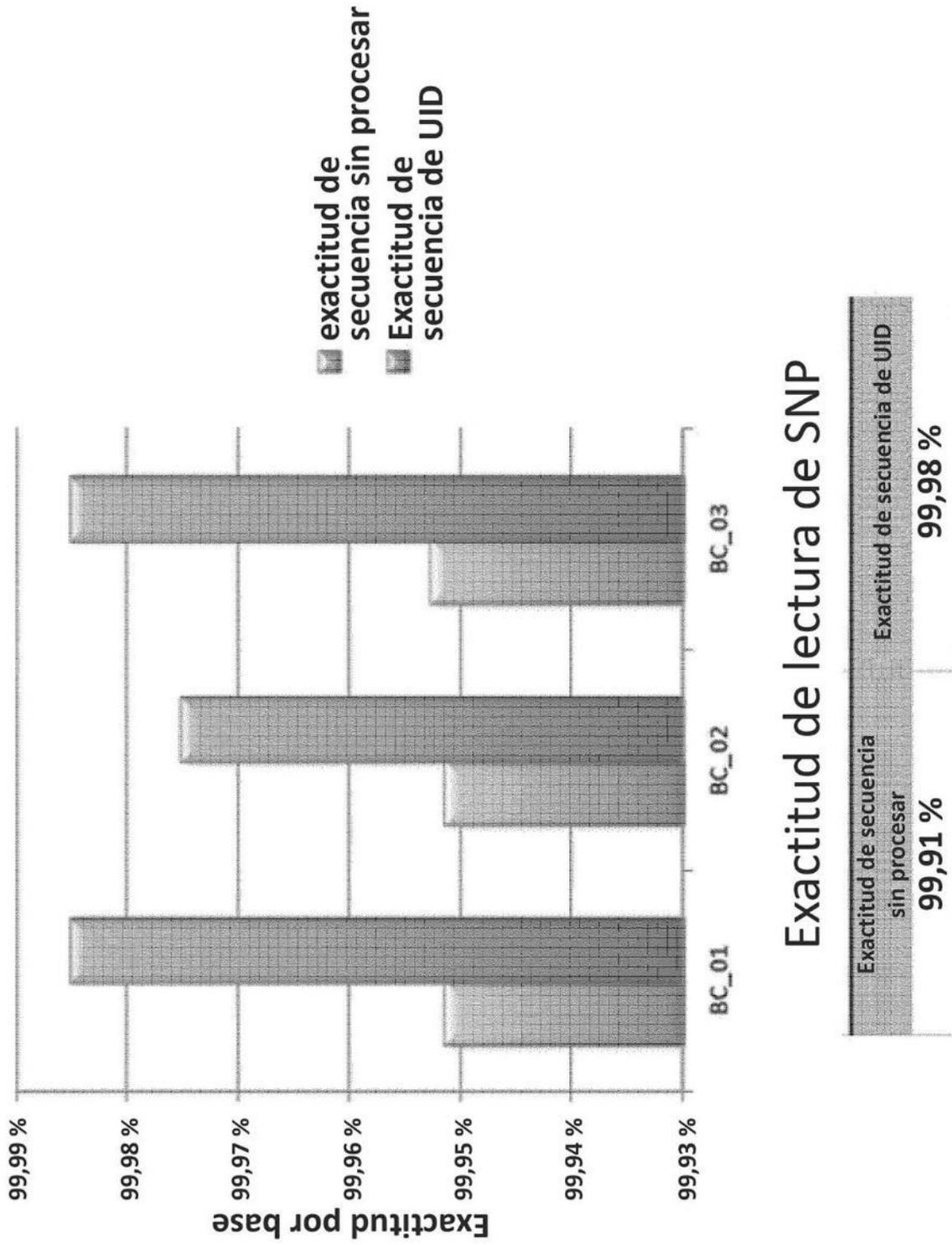


Fig. 19

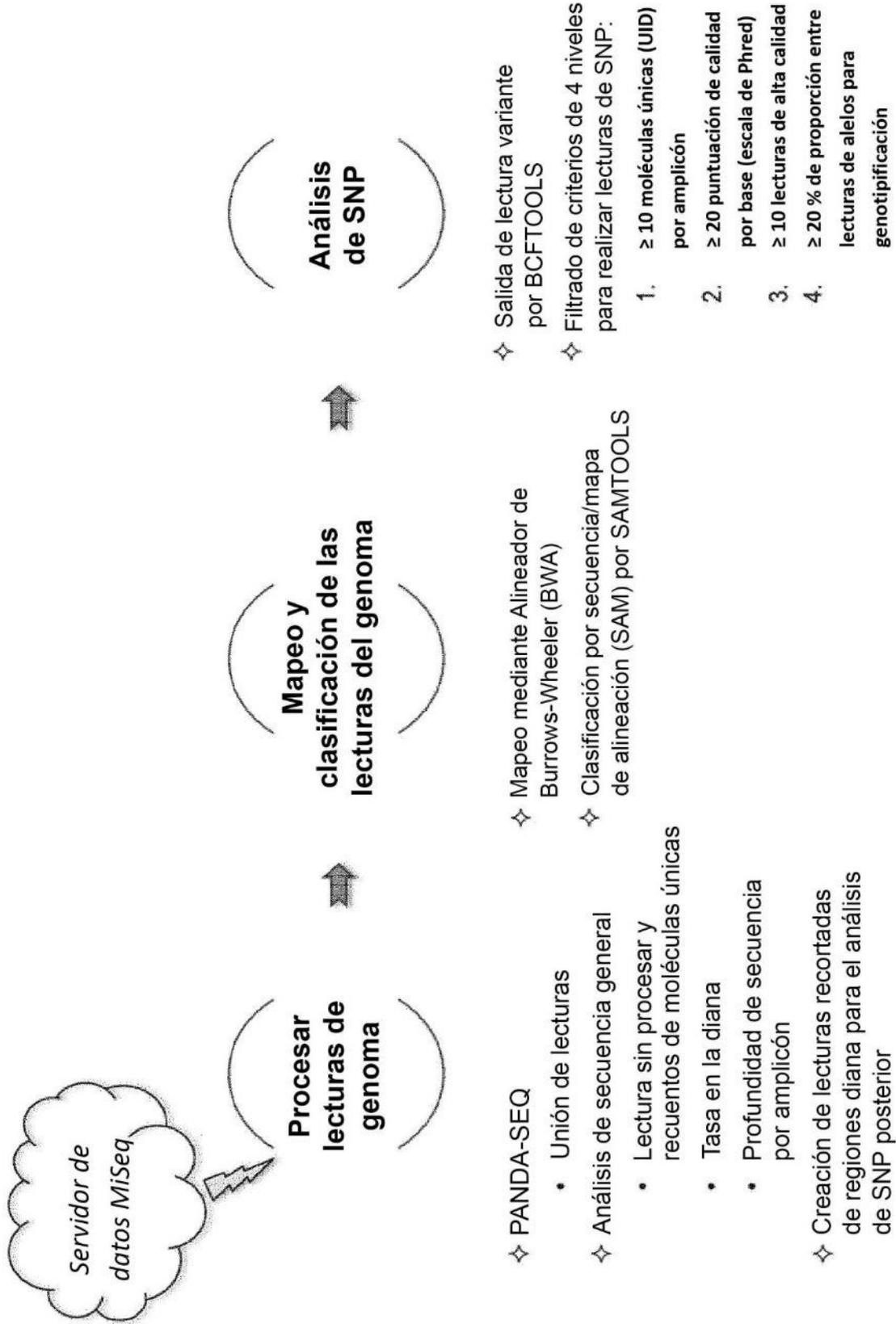
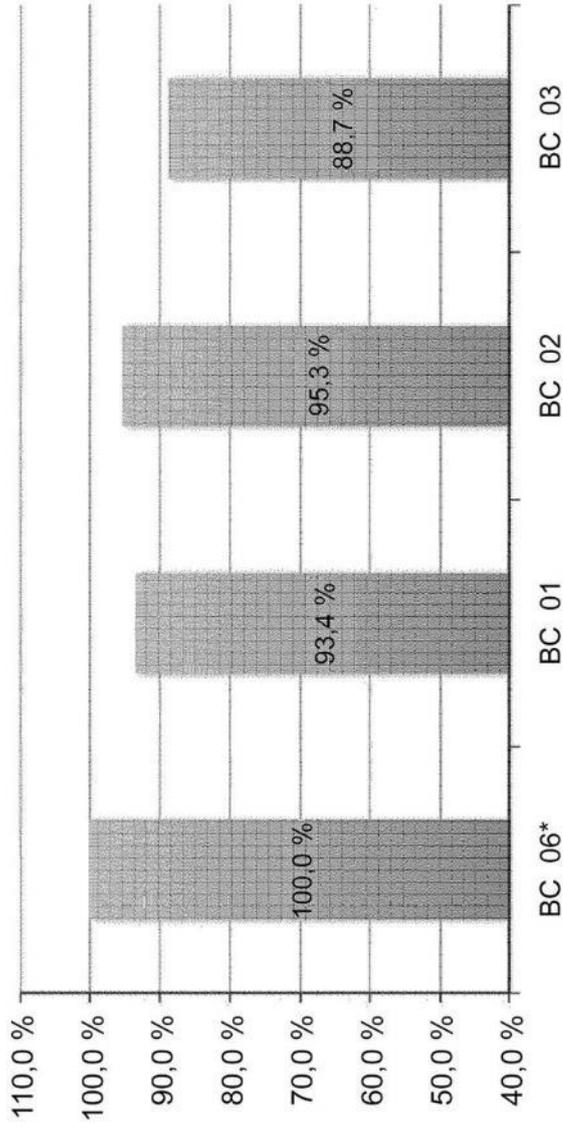


Fig. 20

Porcentaje relativo de lecturas de SNP capturadas



Muestra de ADNg-FV	Ciclos de PCR C5/C7	UID totales		Amplificones ≥ 10 UID [^]	% de ampli-cones ≥ 10 UID	Filtro de 4 niveles		Coincidencia con BC_06 (ref)	% de coincidencia relativa
		UID totales	UID ≥ 10 UID [^]			Sin filtro	Filtro de 4 niveles		
BC_06*	20	136.687	326	326	97,0 %	169	106	106	100,0 %
BC_01	24	72.951	305	305	90,8 %	150	100	99	93,4 %
BC_02	24	68.528	305	305	90,8 %	154	102	101	95,3 %
BC_03	24	41.671	289	289	86,0 %	137	96	94	88,7 %

* BC_06 es la muestra de referencia

[^] El panel optimizado CS-350 Tex01 contiene 336 amplificones.

Fig. 21

Panel completo CS-350: Análisis de rendimiento por % de contenido de GC del amplicón

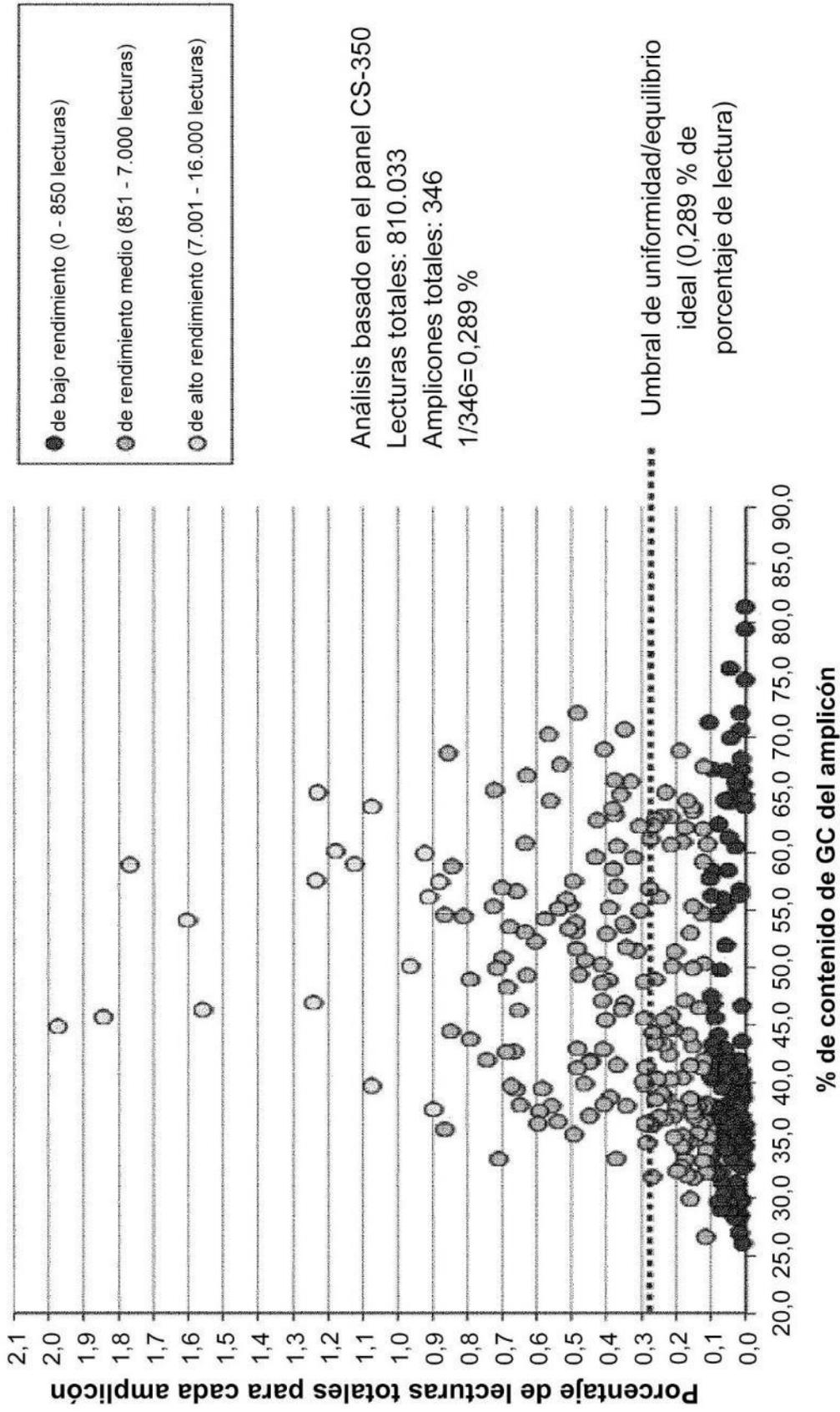


Fig. 22A

Panel completo CS-350: Análisis de rendimiento por % de contenido de GC del amplicón

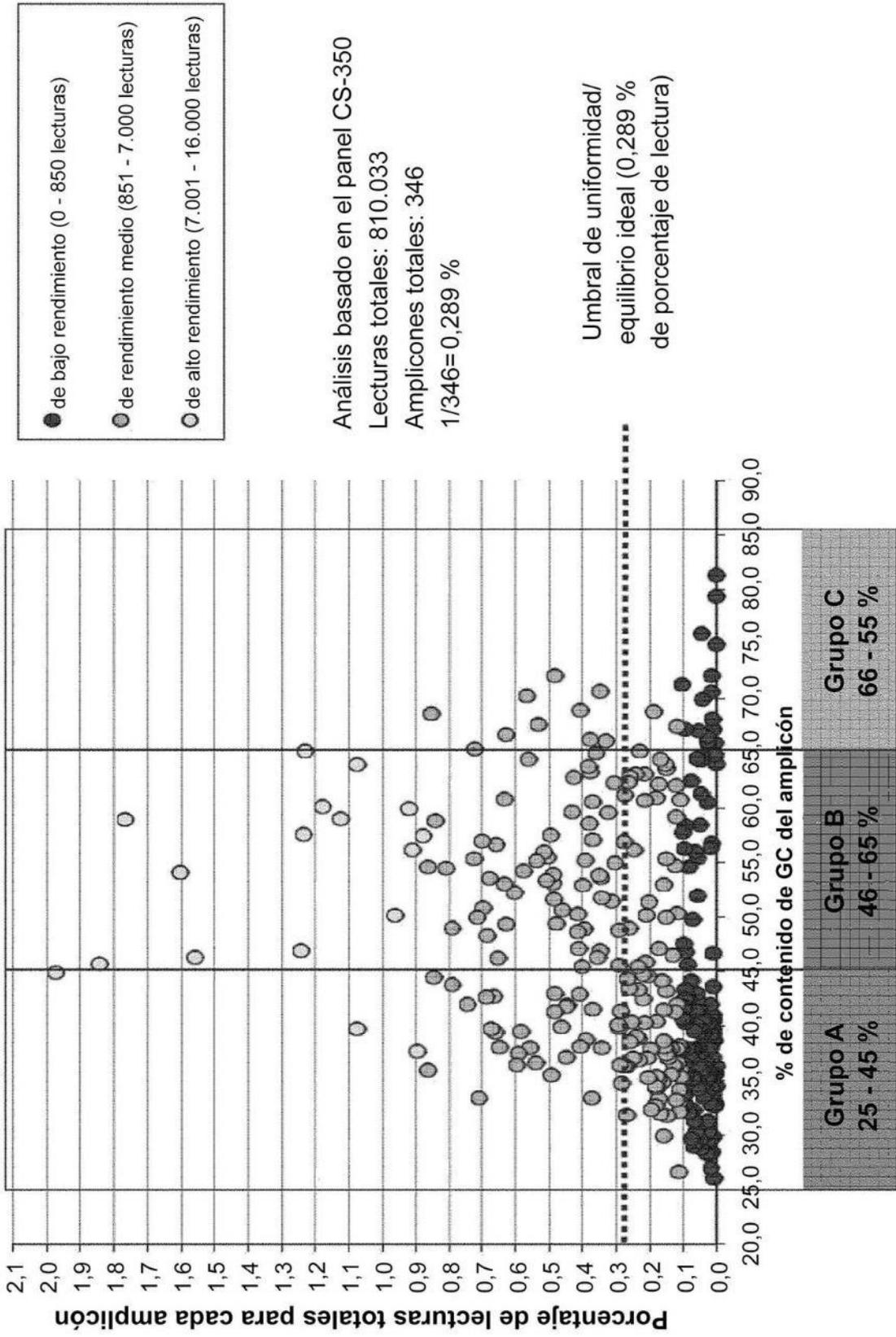


Fig. 22B

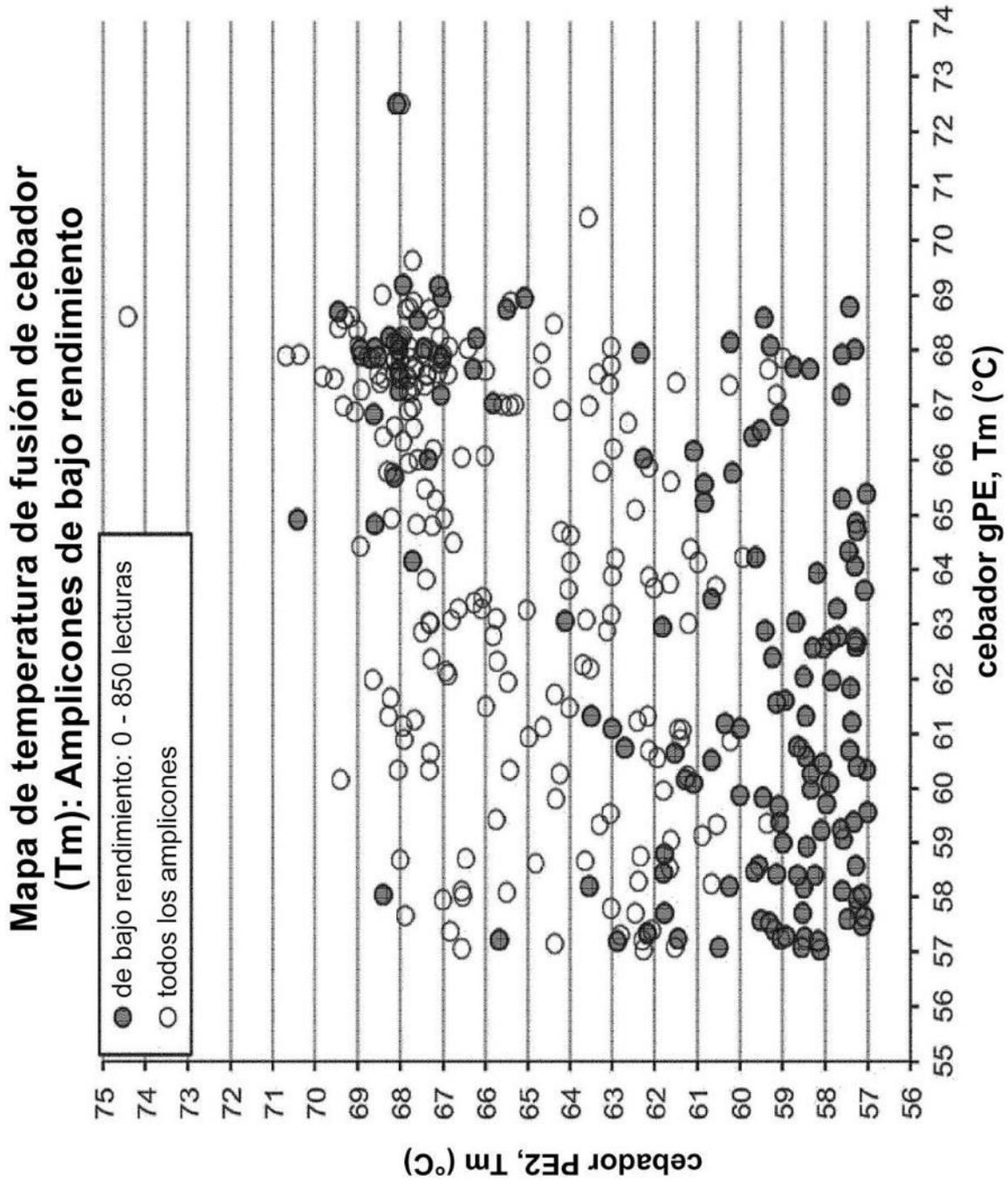


Fig. 23A

Amplicón % de contenido de GC de amplicones de bajo rendimiento

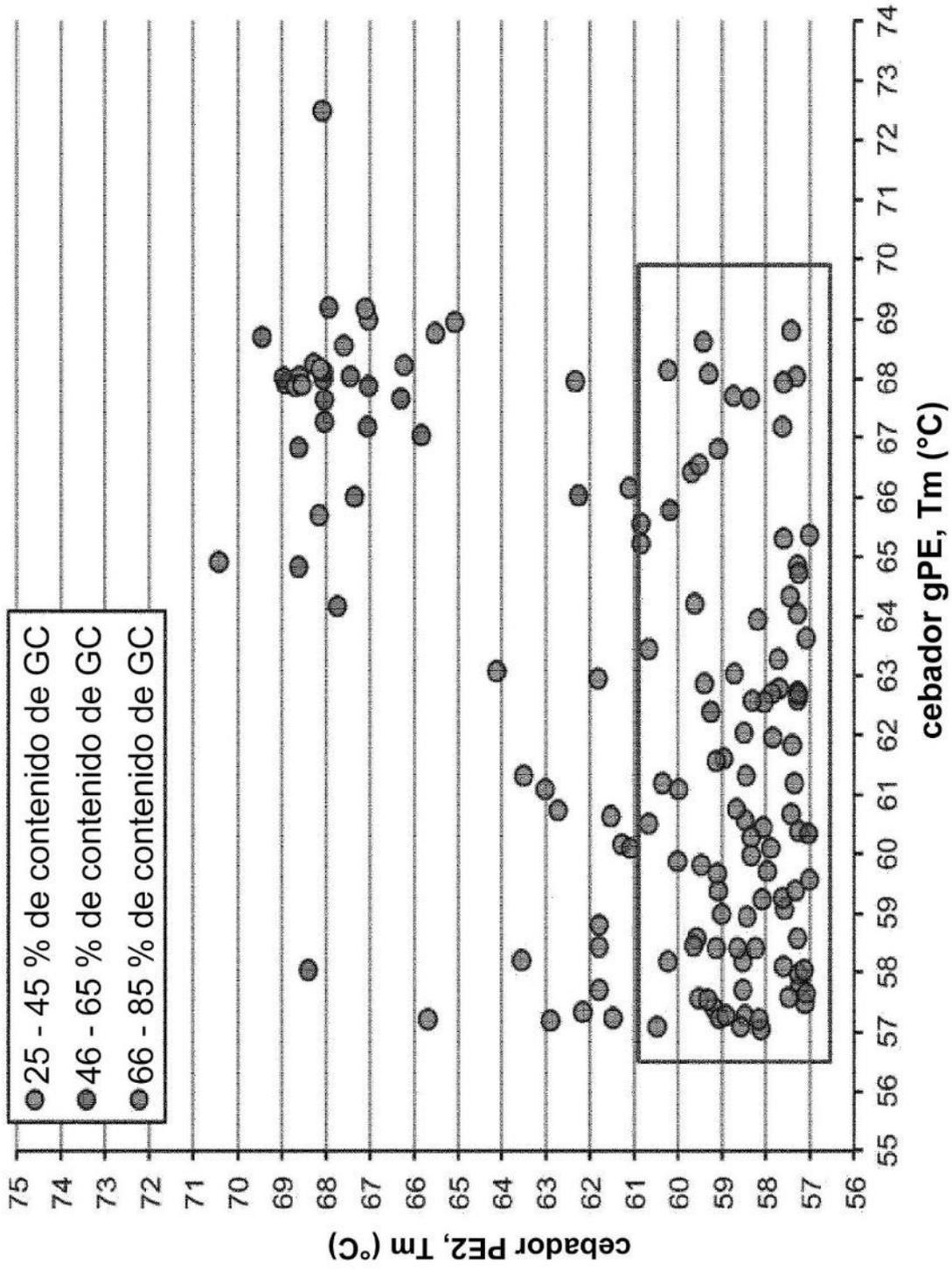


Fig. 23B

Mapa de temperatura de fusión de cebador (Tm): Resumen de rendimiento de los amplicones CS-350

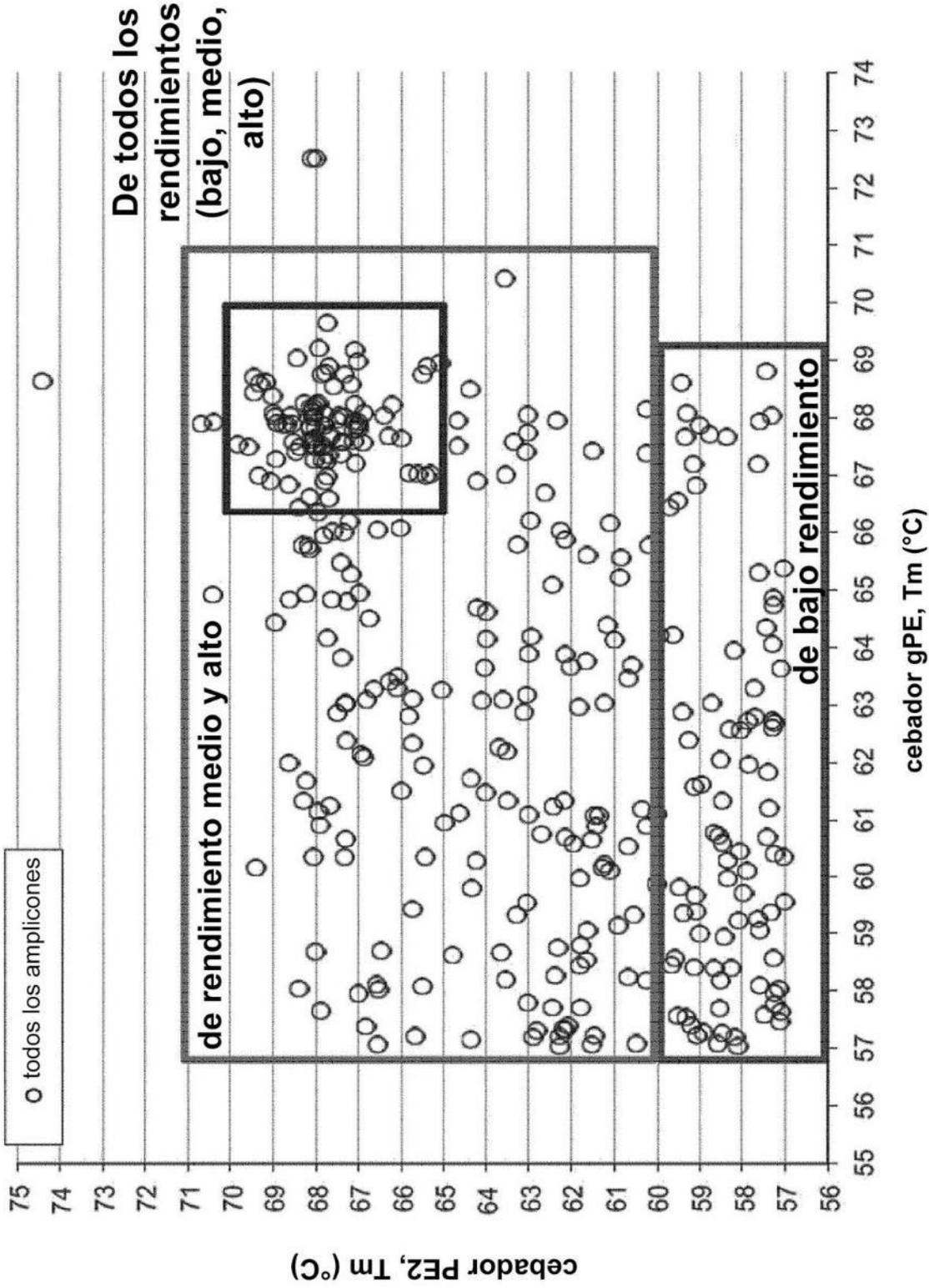
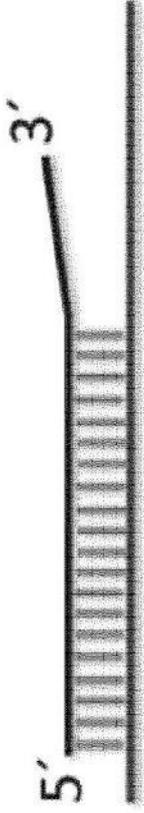


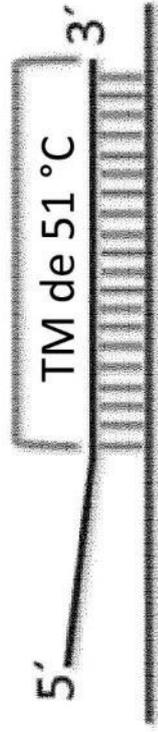
Fig. 24

- ◆ Ajustar el intervalo de temperatura de fusión: 60 - 68 °C; óptima 64 °C
- ◆ Reducir la longitud: 21 - 32 nt; óptimo 25 nt
- ◆ Evitar GC alto en el extremo 3': ≤ 3 GC en los últimos 5 nt
- ◆ % de GC: 30 - 70 %; óptimo 50 %
- ◆ Longitud del amplicón: 225 - 300 pb

Fig. 25



No se cuenta porque no hay especificidad de 3'



Se cuenta debido a la especificidad de 3' y $T_M > 50\text{ }^\circ\text{C}$

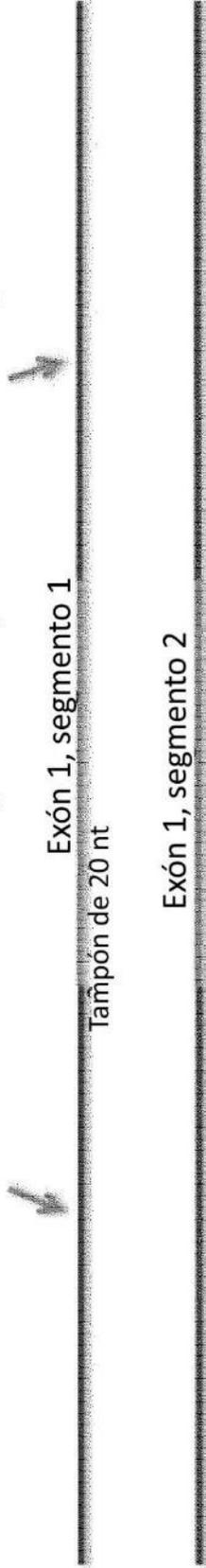
Fig. 26

exón de 315 pb + 40 pb de tampón intrónico = segmento de 355 pb

Cortar los exones más grandes en segmentos (150 pb o menos)



Permitir el diseño del cebador en la región de 200 pb a cada lado del segmento diana



Región de diseño de 200 pb Región diana Región de diseño de 200 pb

Fig. 27

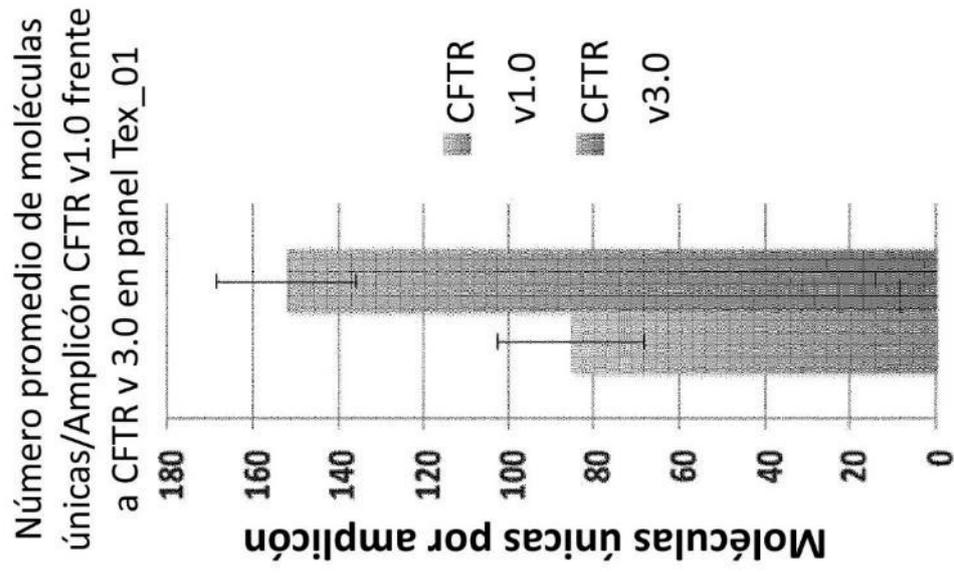


Fig. 28

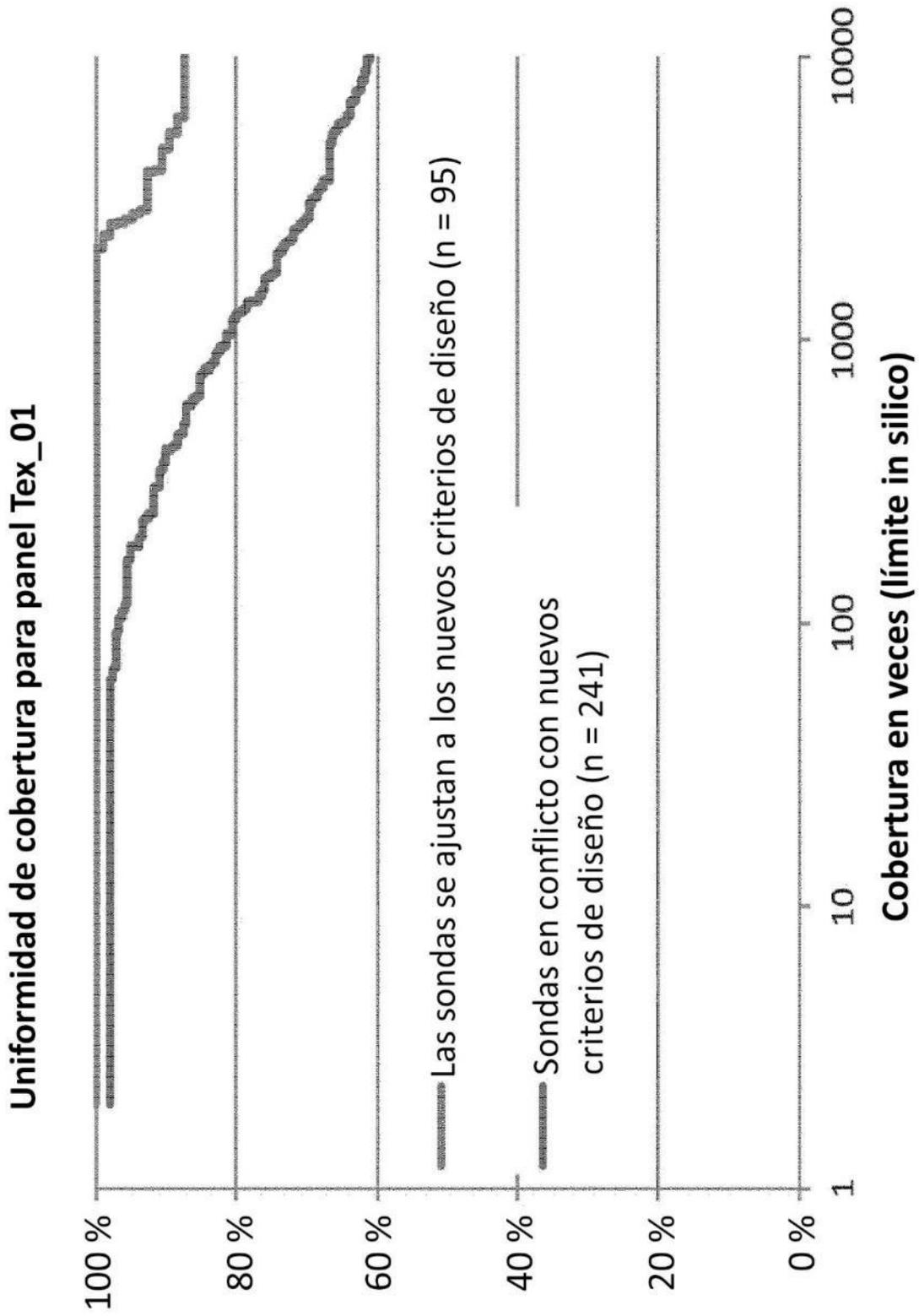


Fig. 29

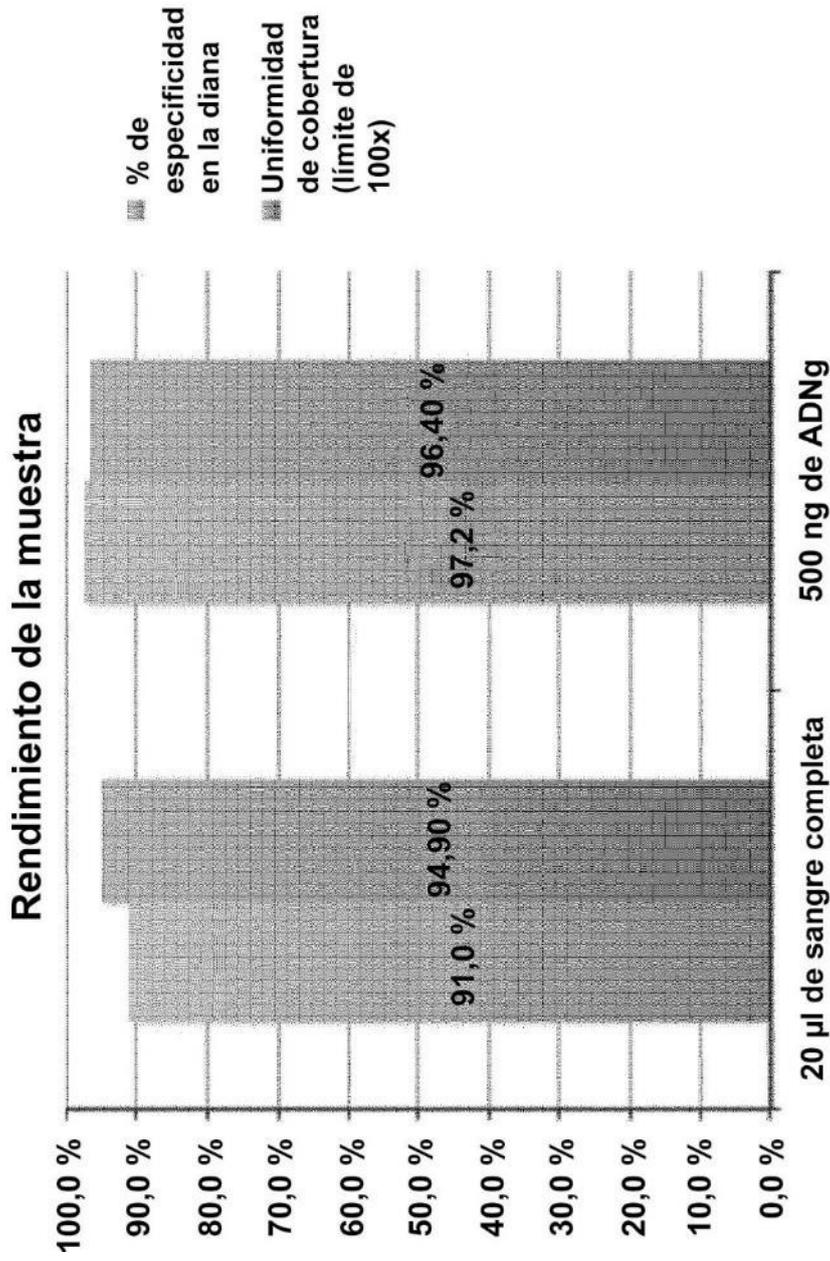


Fig. 30A

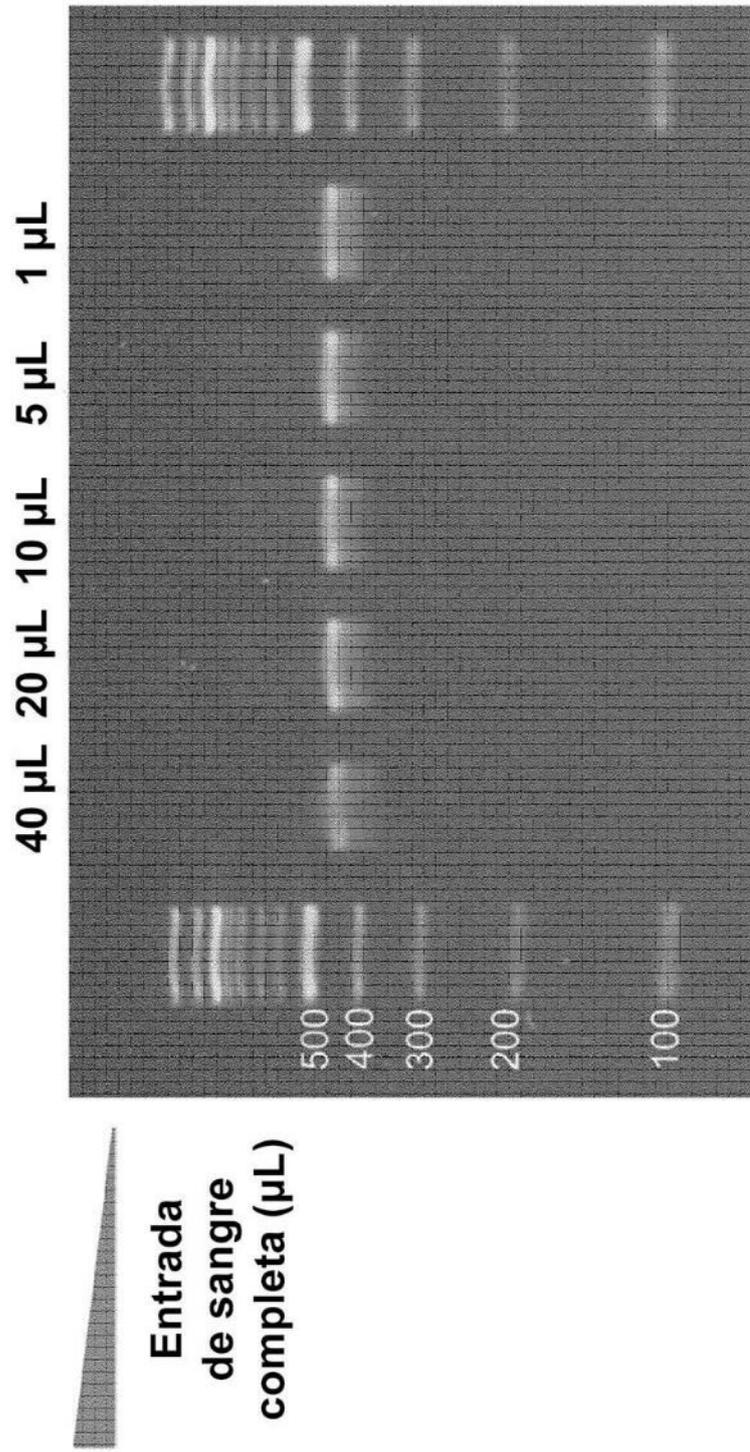
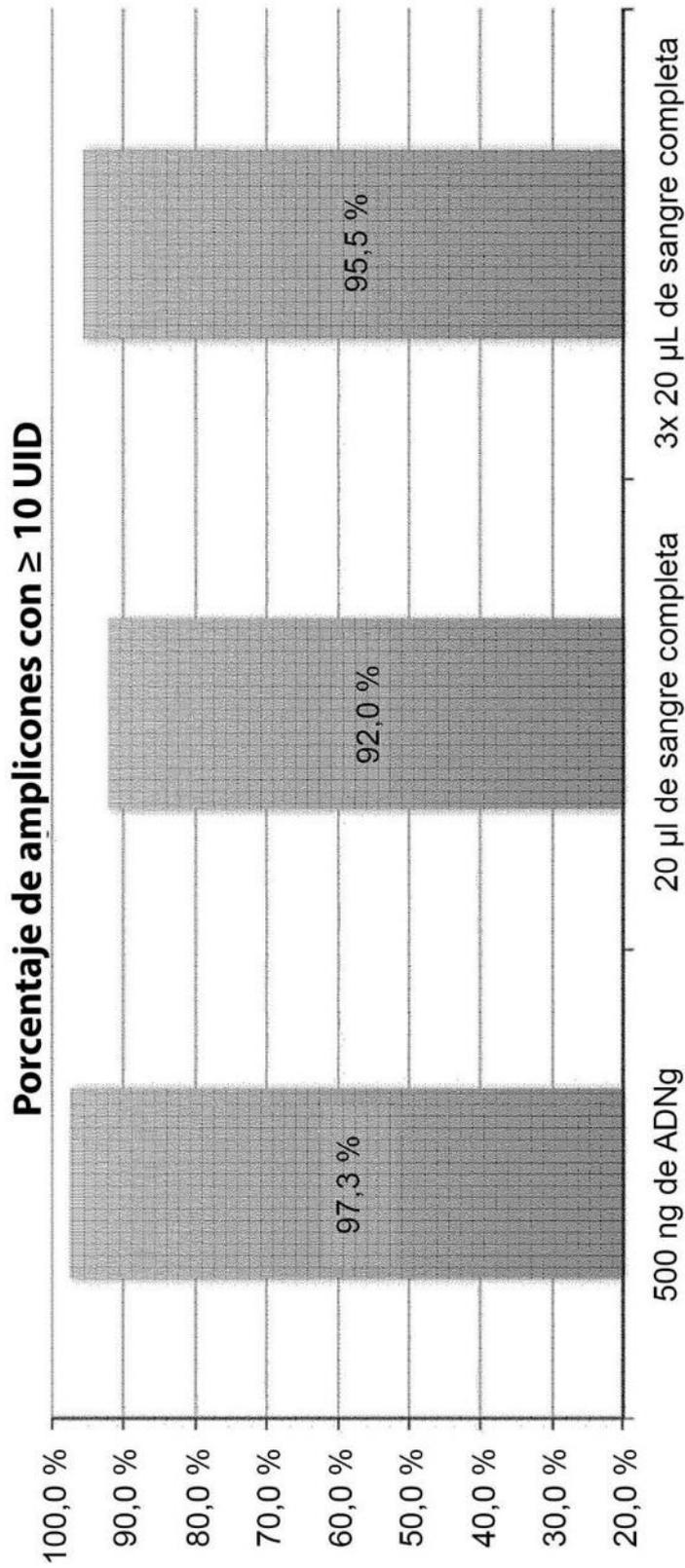


Fig. 30B



Muestra	Total de UID [^]	Amplicones ≥ 10 UID	% de amplicones ≥ 10 UID	Total de lecturas de SNP	Lecturas de SNP filtradas
500 ng de ADNg	186.243	327	97,3 %	115	102
20 μ l de sangre completa	80.676	309	92,0 %	163	99
3x 20 μ L de sangre completa	182.118	321	95,5 %	166	103

[^] Basado en el análisis de 500.000 lecturas sin procesar La sangre y el ADNg son del mismo paciente

Fig. 31

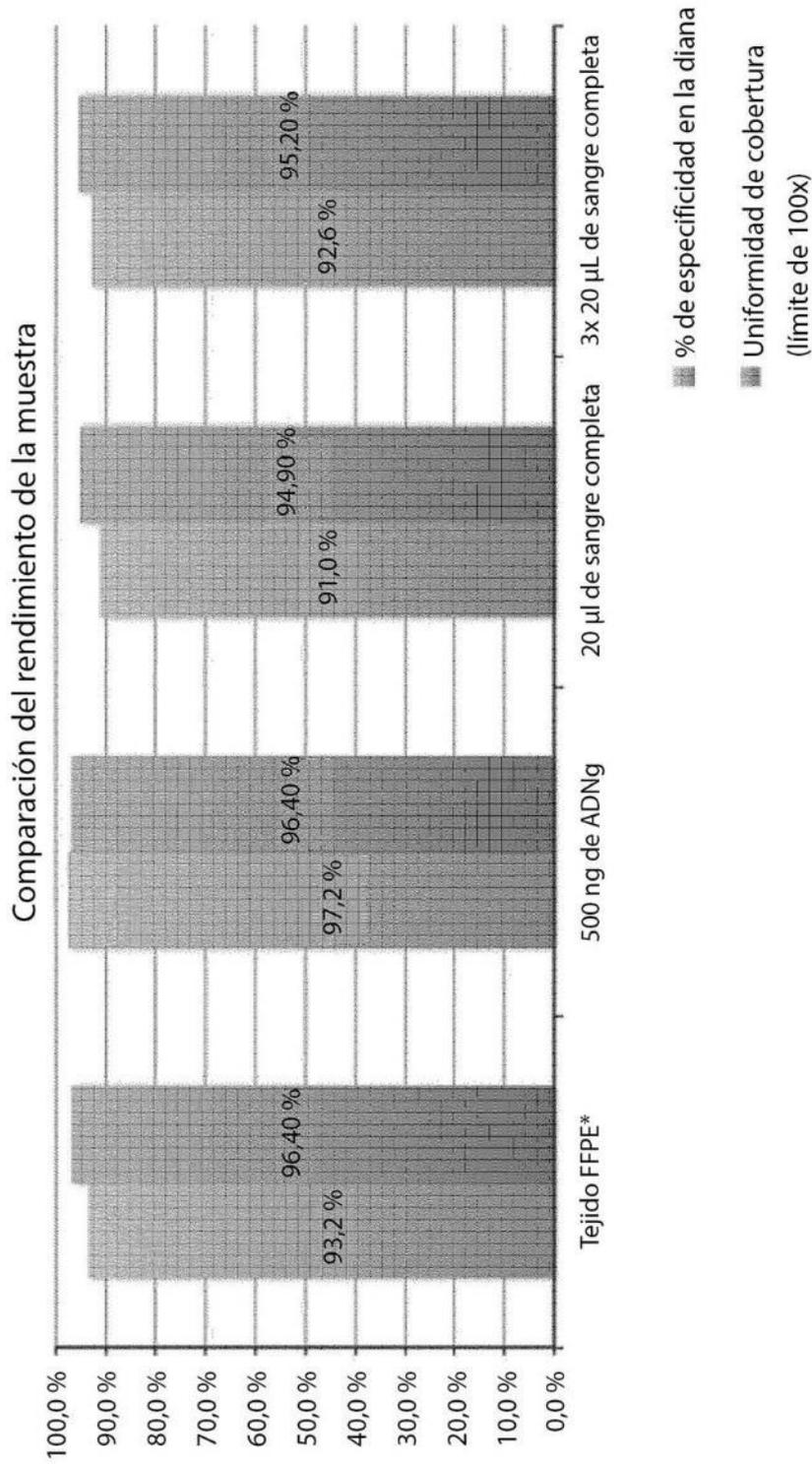


Fig. 33

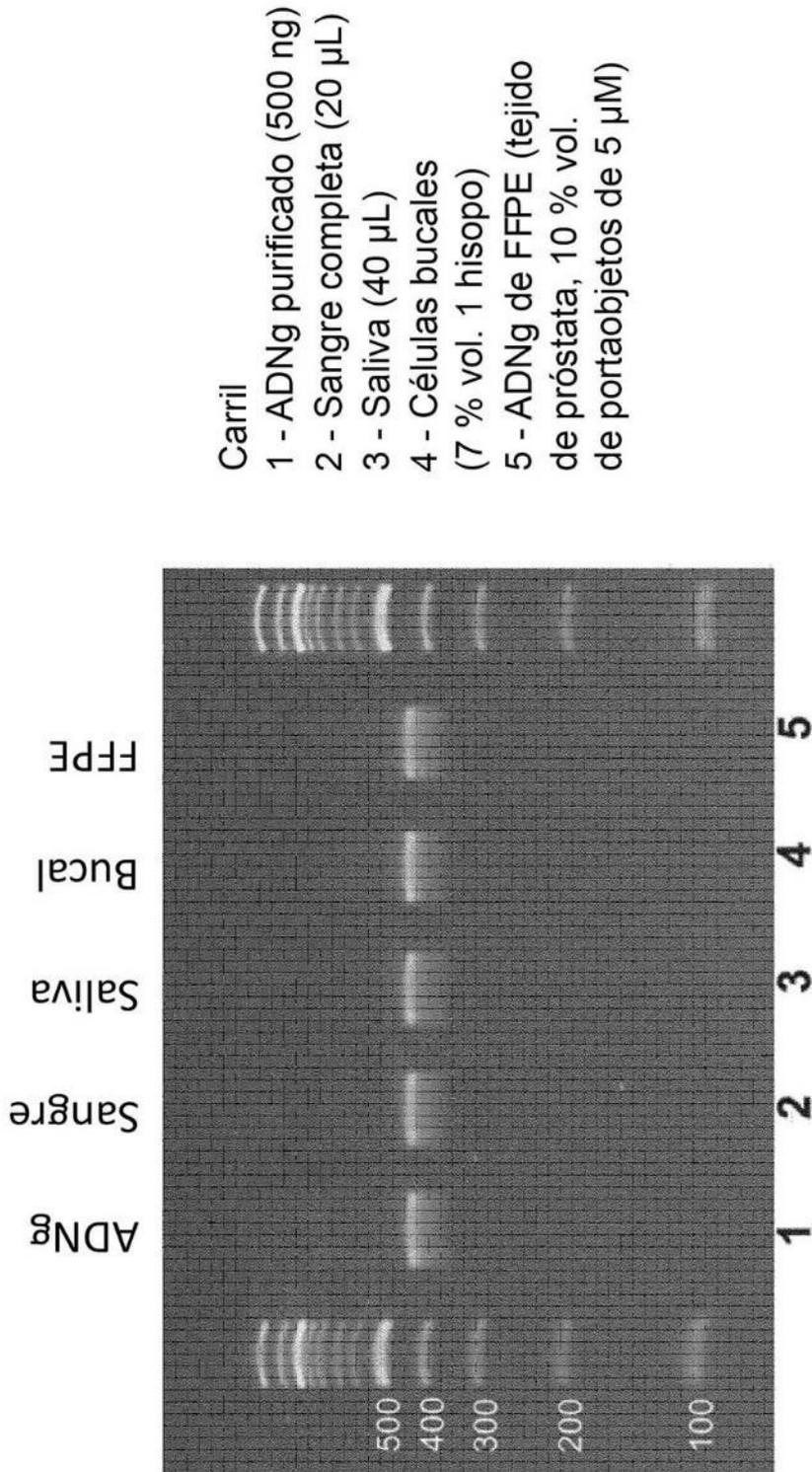
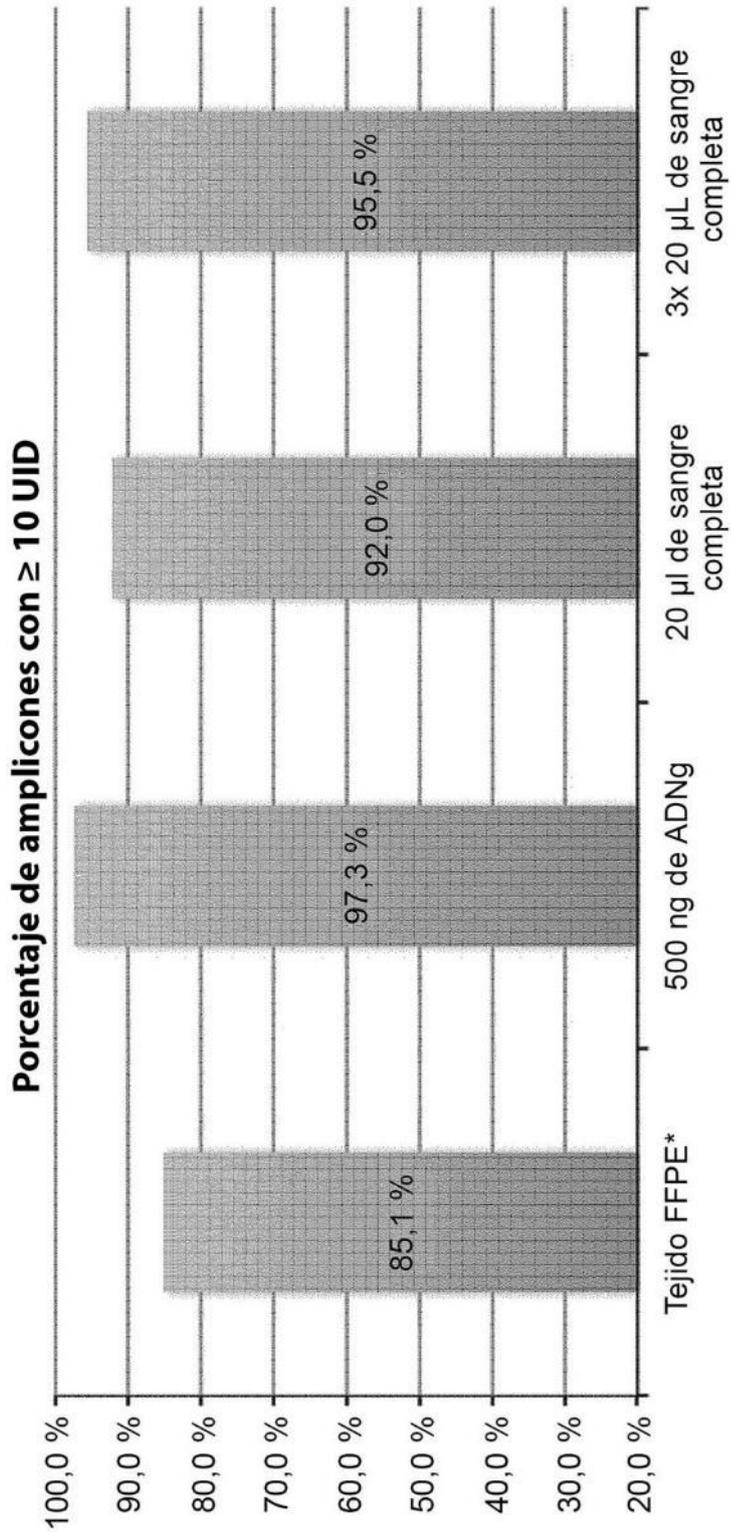


Fig. 34



Muestra	Total de UID [^]	Amplicones ≥ 10 UID	% de amplicones ≥ 10 UID	Total de lecturas de SNP	Lecturas de SNP filtradas
Tejido FFPE*	24.378	286	85,1 %	158	97
500 ng de ADNg	186.243	327	97,3 %	115	102
20 μ l de sangre completa	80.676	309	92,0 %	163	99
3x 20 μ L de sangre completa	182.118	321	95,5 %	166	103

* Fijado en formalina e incluido en parafina; tejido prostático de adulto normal

[^] Basado en el análisis de 500.000 lecturas sin procesar

Fig. 35

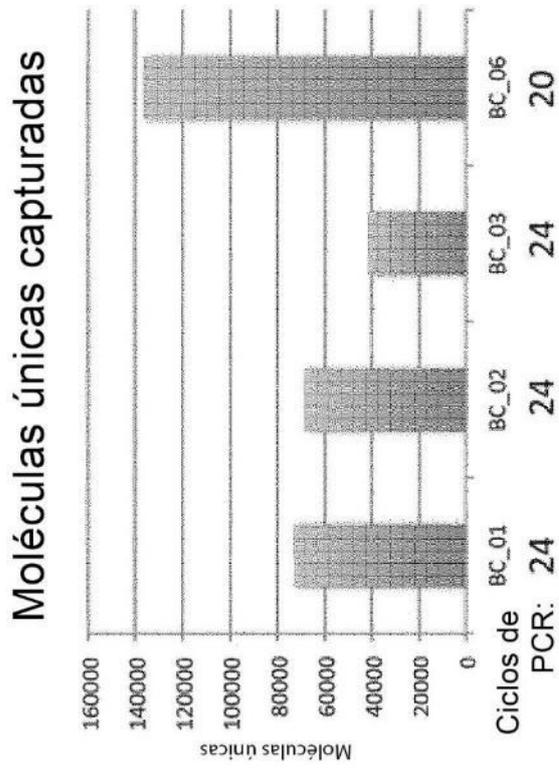


Fig. 36

Calidad de los datos

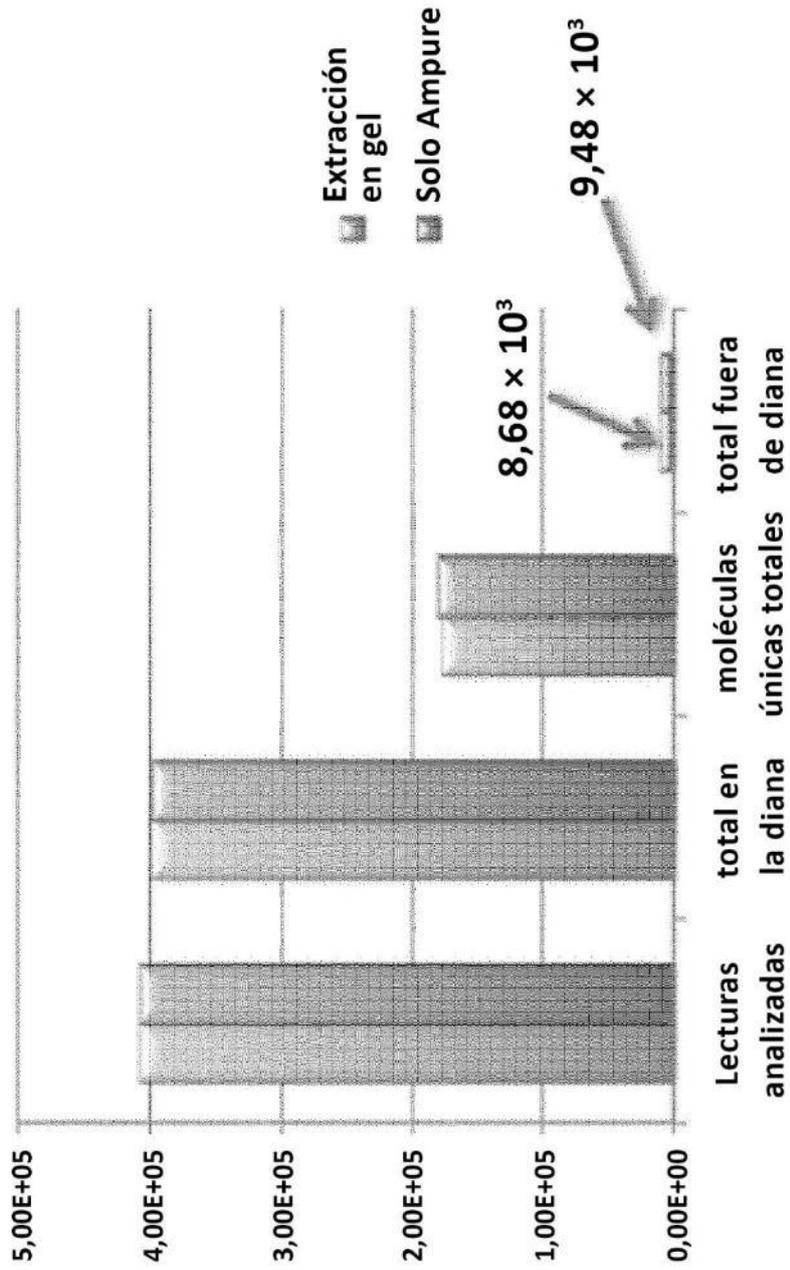


Fig. 37

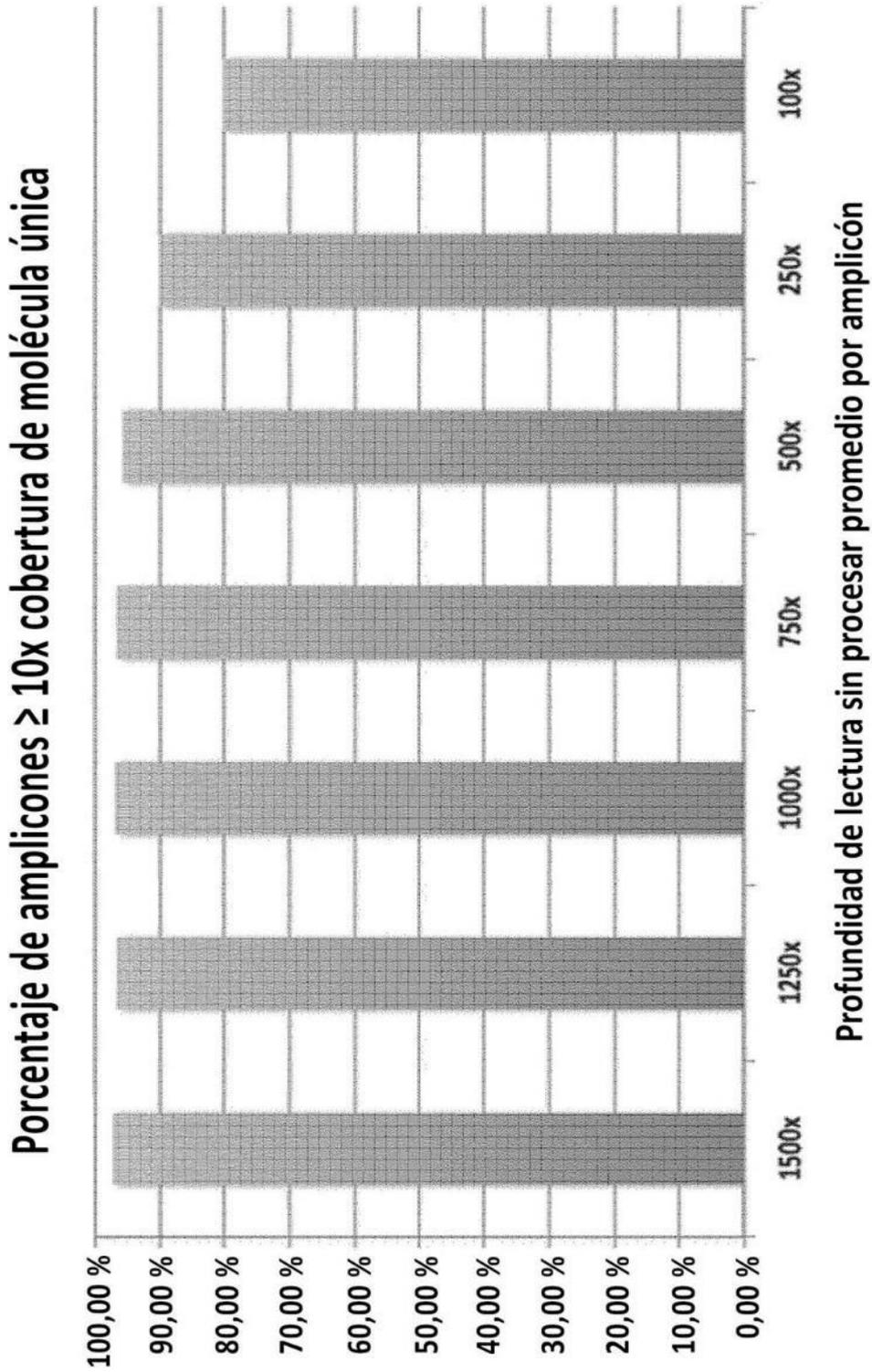


Fig. 38

N.º	Números esperados		Números reales	
	Secuencias	% de ejecución	Secuencias totales	% de ejecución
BC_05	115.000	0,75	175.120	0,95
BC_11	200.000	1,30	249.559	1,35
BC_17	200.000	1,30	439.386	2,38
BC_01	400.000	2,59	415.986	2,25
BC_02	400.000	2,59	381.609	2,07
BC_03	400.000	2,59	279.127	1,51
BC_04	400.000	2,59	518.732	2,81
BC_08	500.000	3,24	641.863	3,48
BC_10	500.000	3,24	602.764	3,27
BC_14	500.000	3,24	630.015	3,41
BC_16	500.000	3,24	618.677	3,35
BC_06	500.000	3,24	661.402	3,58
BC_21	500.000	3,24	611.911	3,32
BC_36	500.000	3,24	732.291	3,97
BC_39	500.000	3,24	600.183	3,25
BC_26	500.000	3,24	578.438	3,13
BC_29	500.000	3,24	637.069	3,45
BC_19	500.000	3,24	612.868	3,32
BC_31	500.000	3,24	660.459	3,58
BC_34	500.000	3,24	509.304	2,76
BC_30	500.000	3,24	489.322	2,65
BC_32	500.000	3,24	546.824	2,96
BC_33	500.000	3,24	497.437	2,70
BC_35	500.000	3,24	547.810	2,97
BC_43	800.000	5,19	932.081	5,05
BC_44	800.000	5,19	1.052.359	5,70
BC_45	800.000	5,19	998.955	5,41
BC_46	800.000	5,19	1.100.682	5,96
BC_47	800.000	5,19	914.388	4,96
BC_48	800.000	5,19	815.854	4,42
Total	15.415.000	100,00	18.452.475	100,00

Fig. 39

Proporción de moléculas únicas capturadas por gen

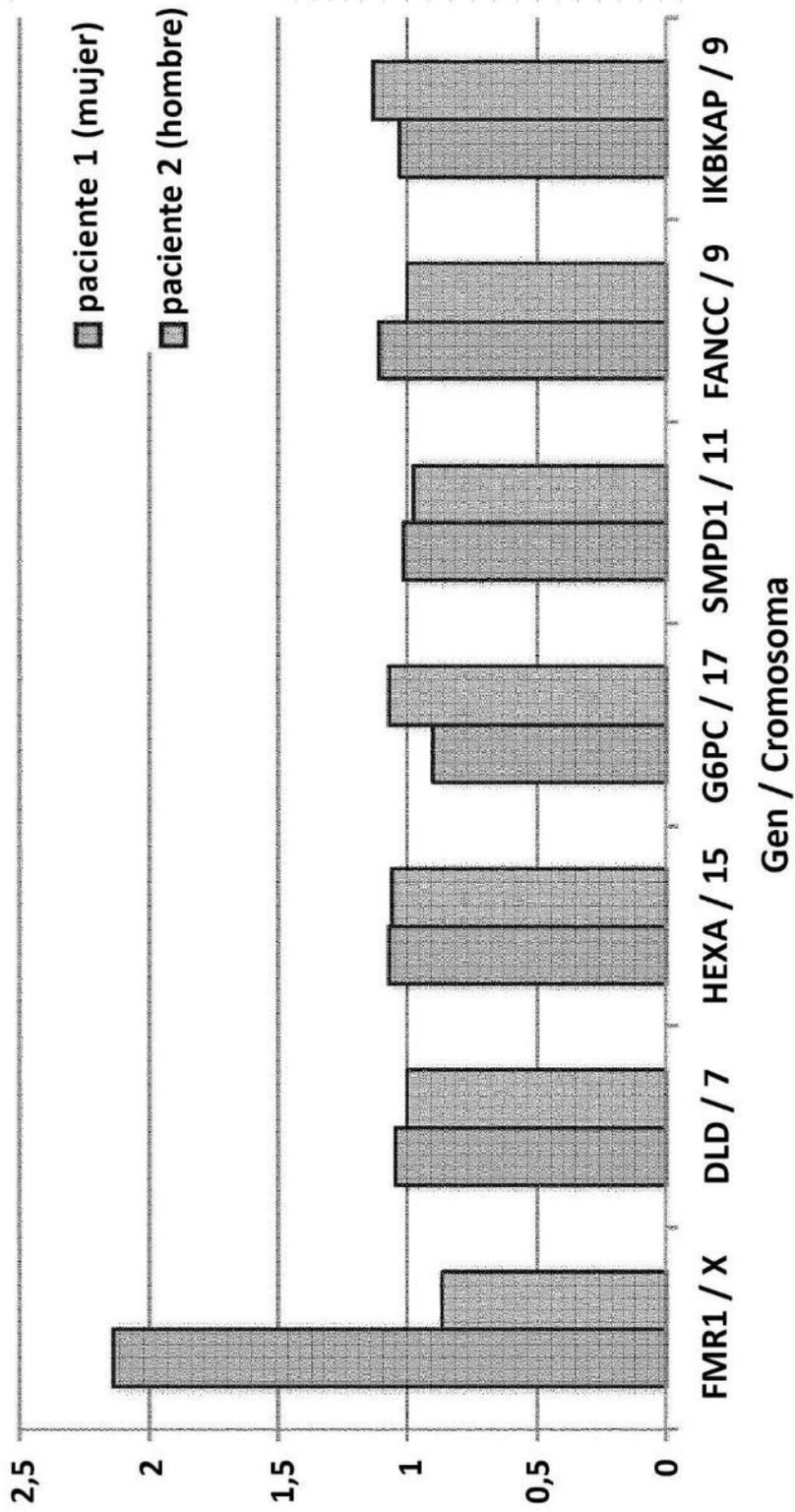


Fig. 40

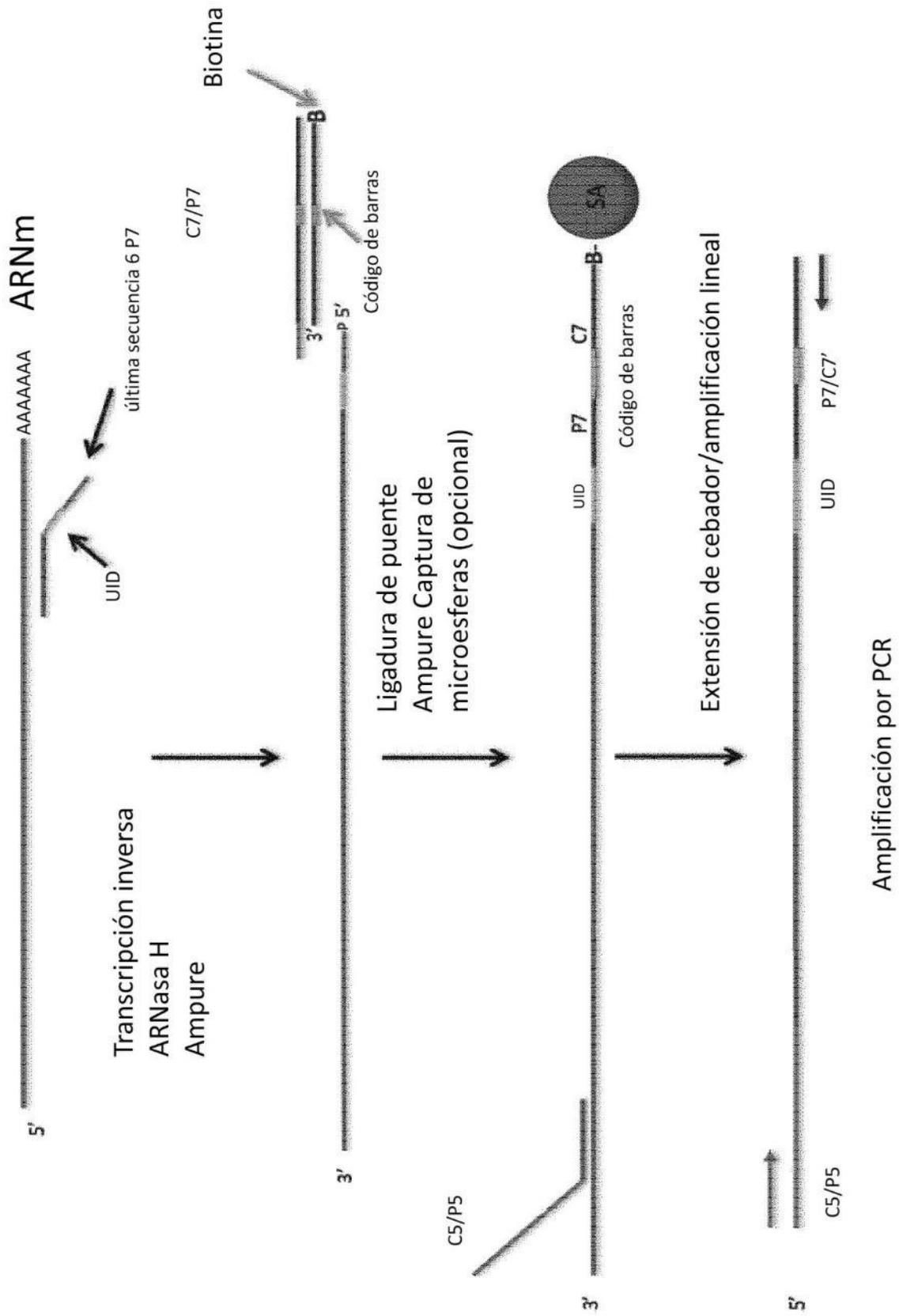
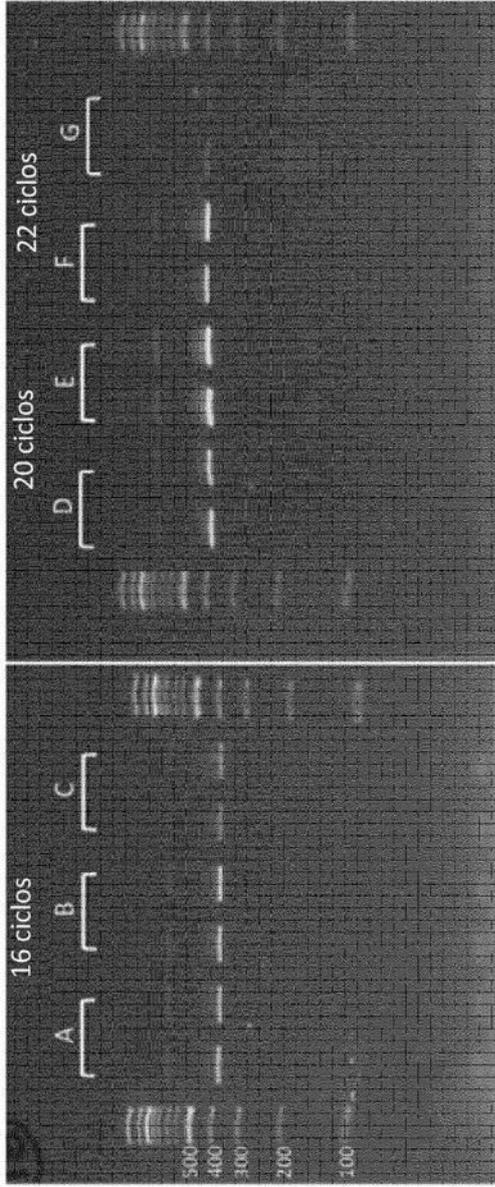


Fig. 41



A.

Entrada: 1000 ng 500 ng 100 ng

50 ng 25 ng 10 ng 1 ng

Titulación de entrada de ARN

Entrada de muestra: ARN total extraído de sangre

15 ciclos de amplificación lineal

B.

3 sondas de patógenos (V3)

- HCV
- EBV
- CMV

5 genes constitutivos (HK5)

- ACTB (exón 6)
- ACTB (exón 6)
- RB2M (exón 4)
- GADPH (exón 5-6)
- GADPH (exón 9)

11 dianas oncológicas (OP-11)

- BRCA1
- ABL1_1
- ABL1_2
- JAK1_1
- JAK1_2
- JAK2_1
- EGFR_1
- CDH1_1
- CDH1_2
- CDH1_3
- CDKN2A_1

Fig. 42

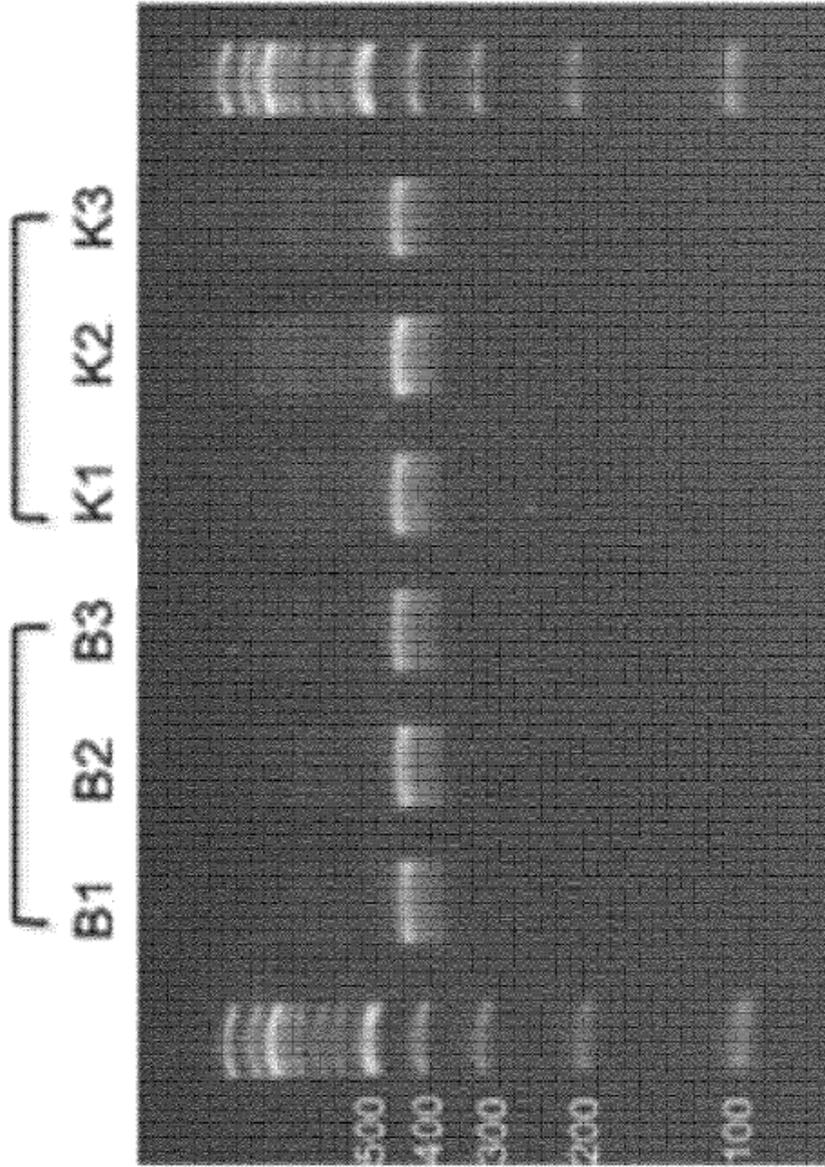


Fig. 43

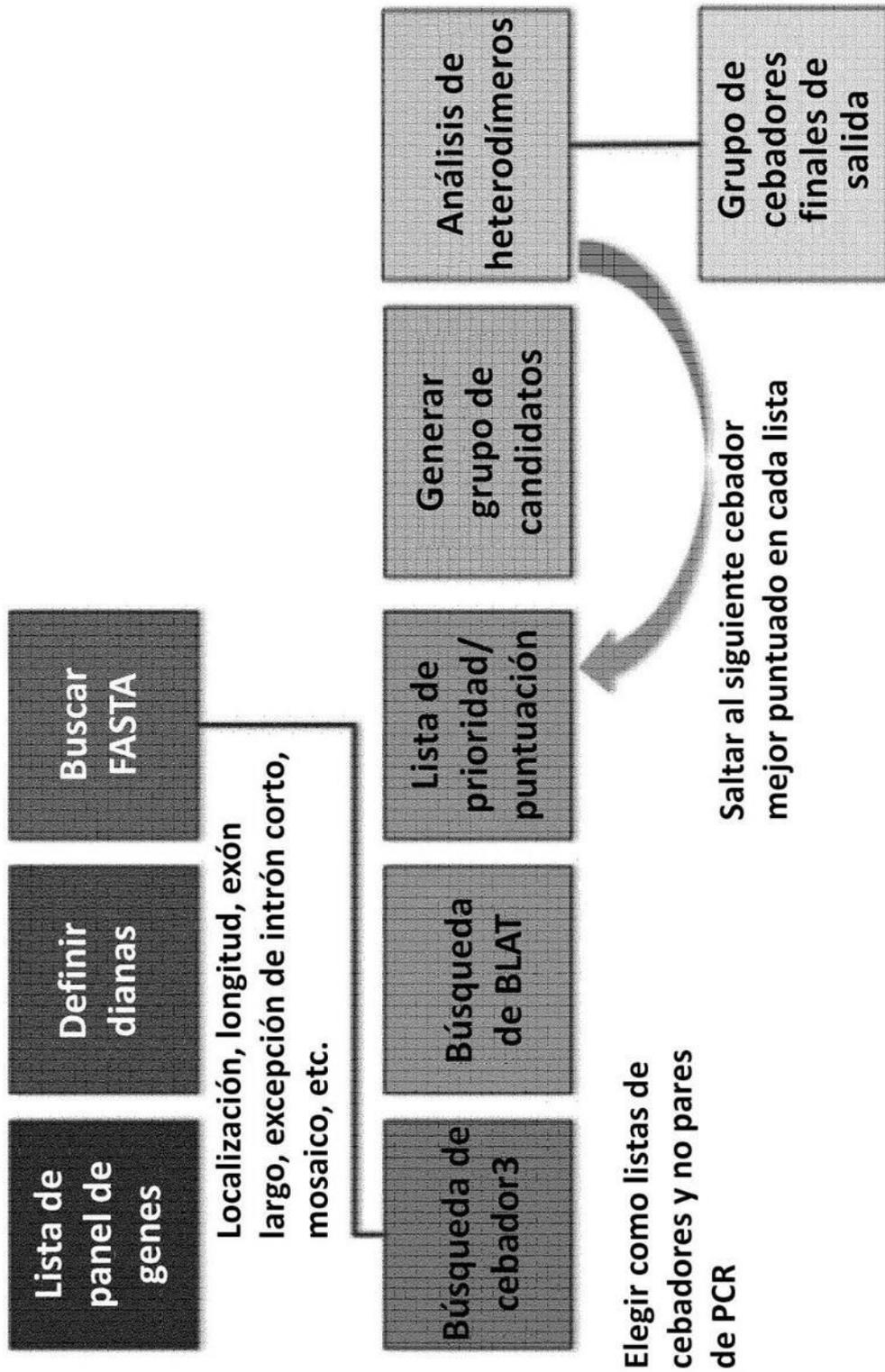
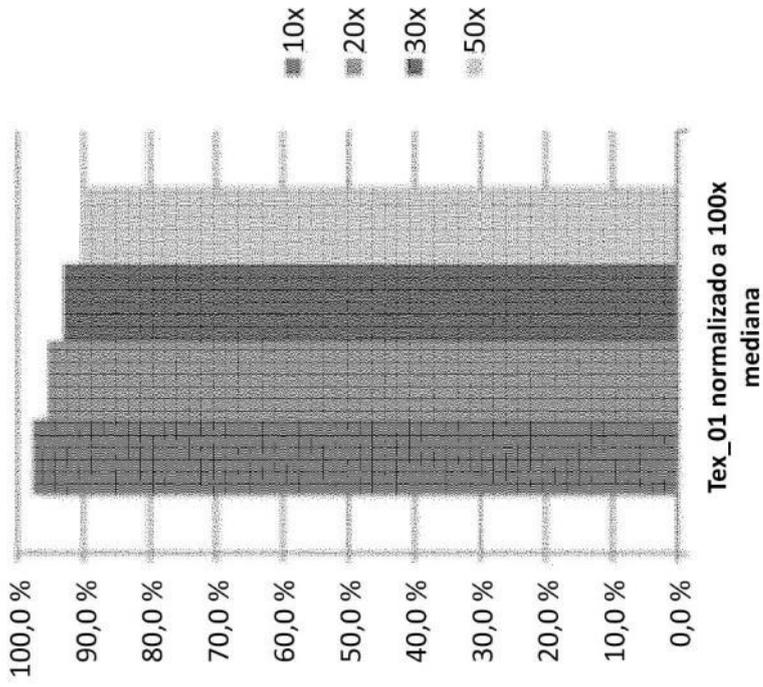


Fig. 44

Uniformidad de cobertura



***Uniformidad de cobertura: bases cubiertas en > 20 % de la cobertura media.**

Fig. 45A

Uniformidad de cobertura: Tex_01 frente a panel completo CS-350

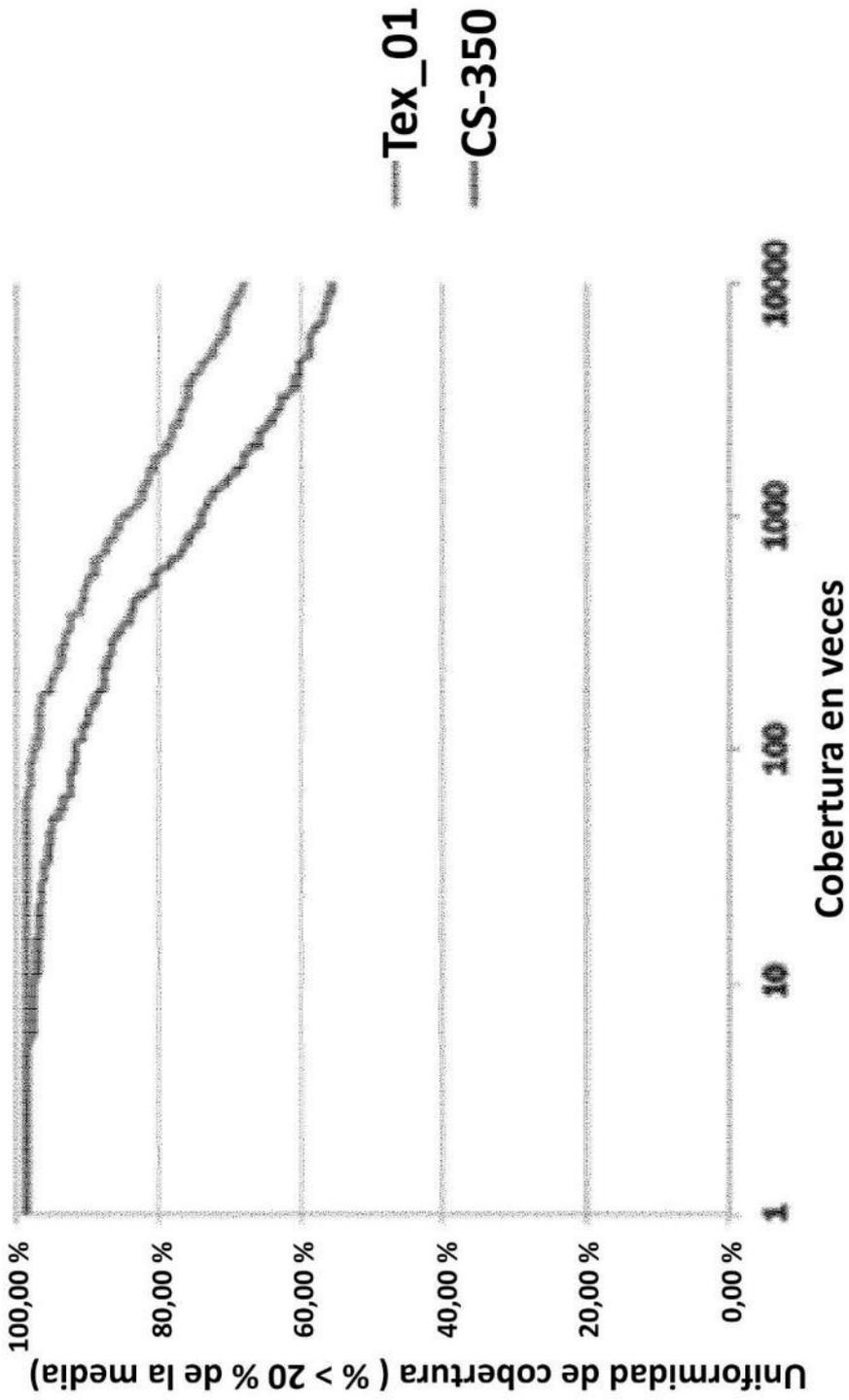


Fig. 45B

Resumen de medidas de calidad	
Especificidad en la diana ¹	97,1 %
Uniformidad de cobertura ²	> 90 %
Exactitud de SNP ³	99,98%

1- Lecturas asignadas en lecturas diana/totales asignadas

2- Uniformidad de cobertura: bases cubiertas en > 20 % de la cobertura media.

3- Total de secuencias nucleotídicas correctas/Total de secuencias nucleotídicas

Fig. 46

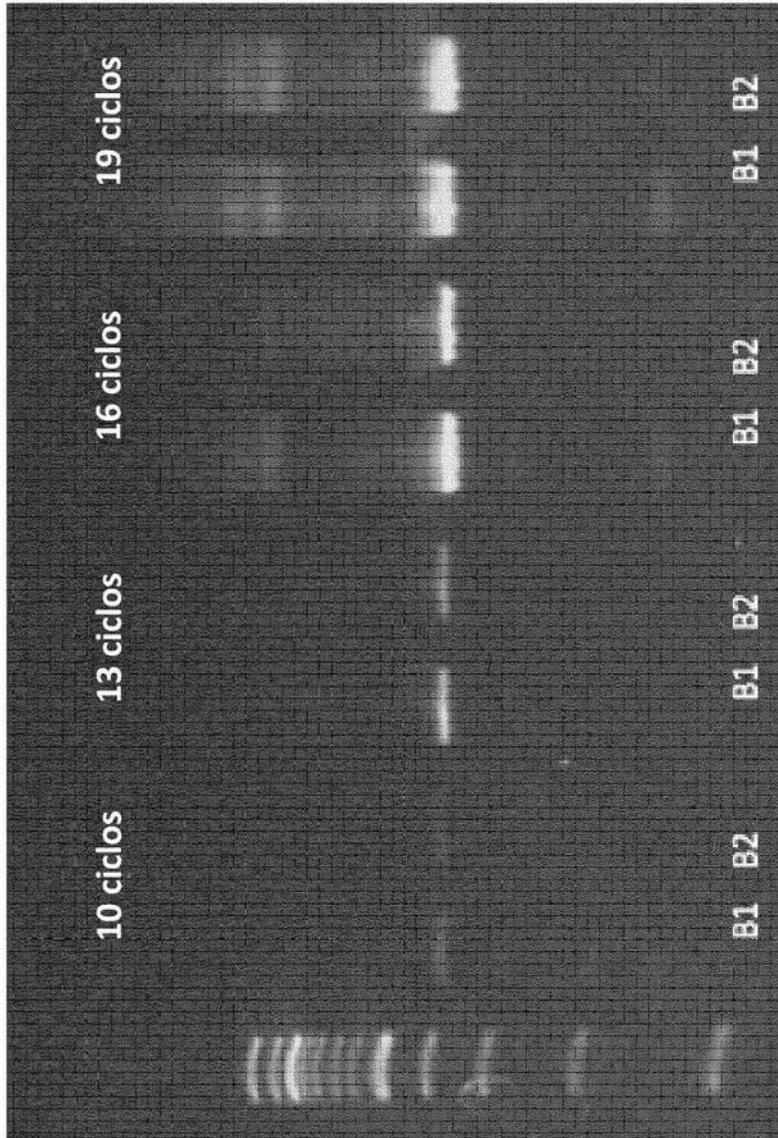
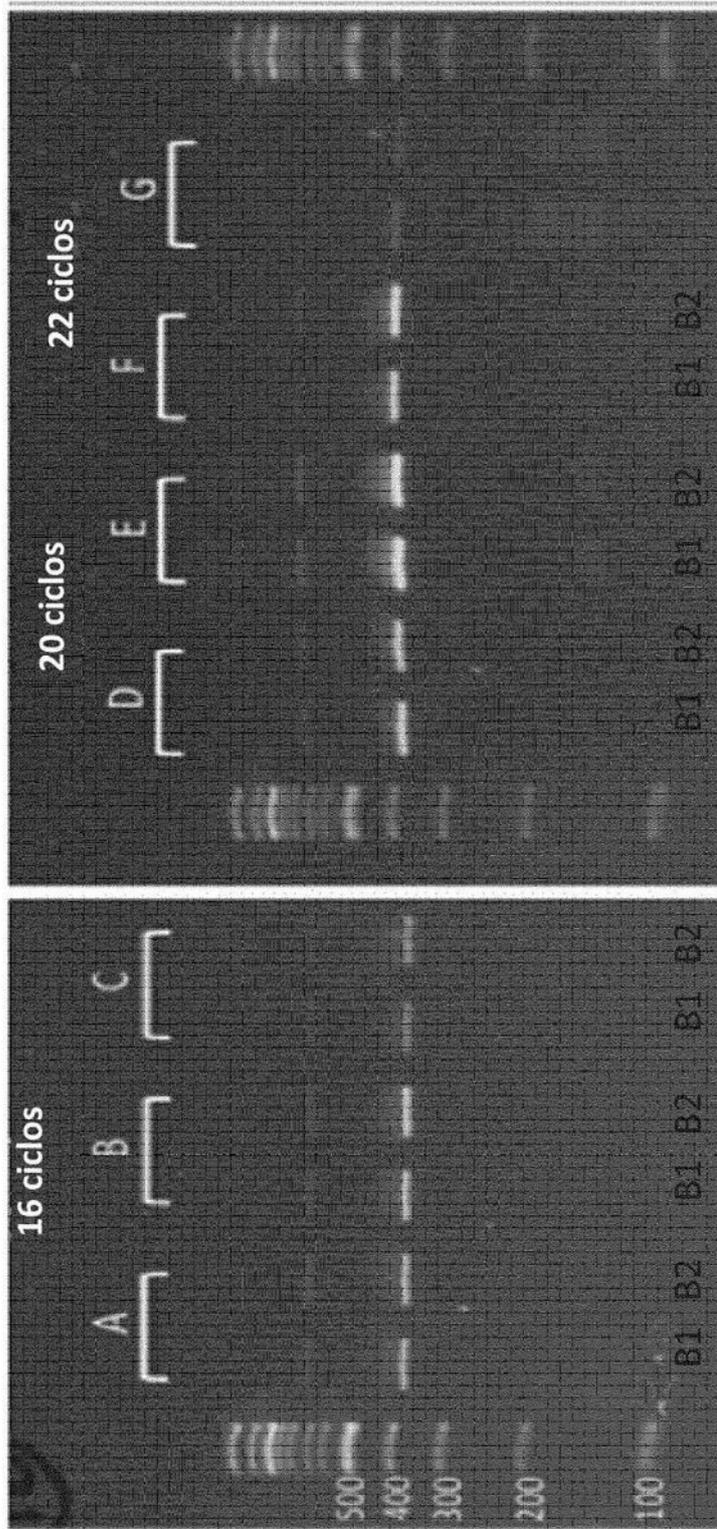


Fig. 47



Entrada de 1000 500 100 1
ARN (ng) 50 25 10 1

Fig. 48

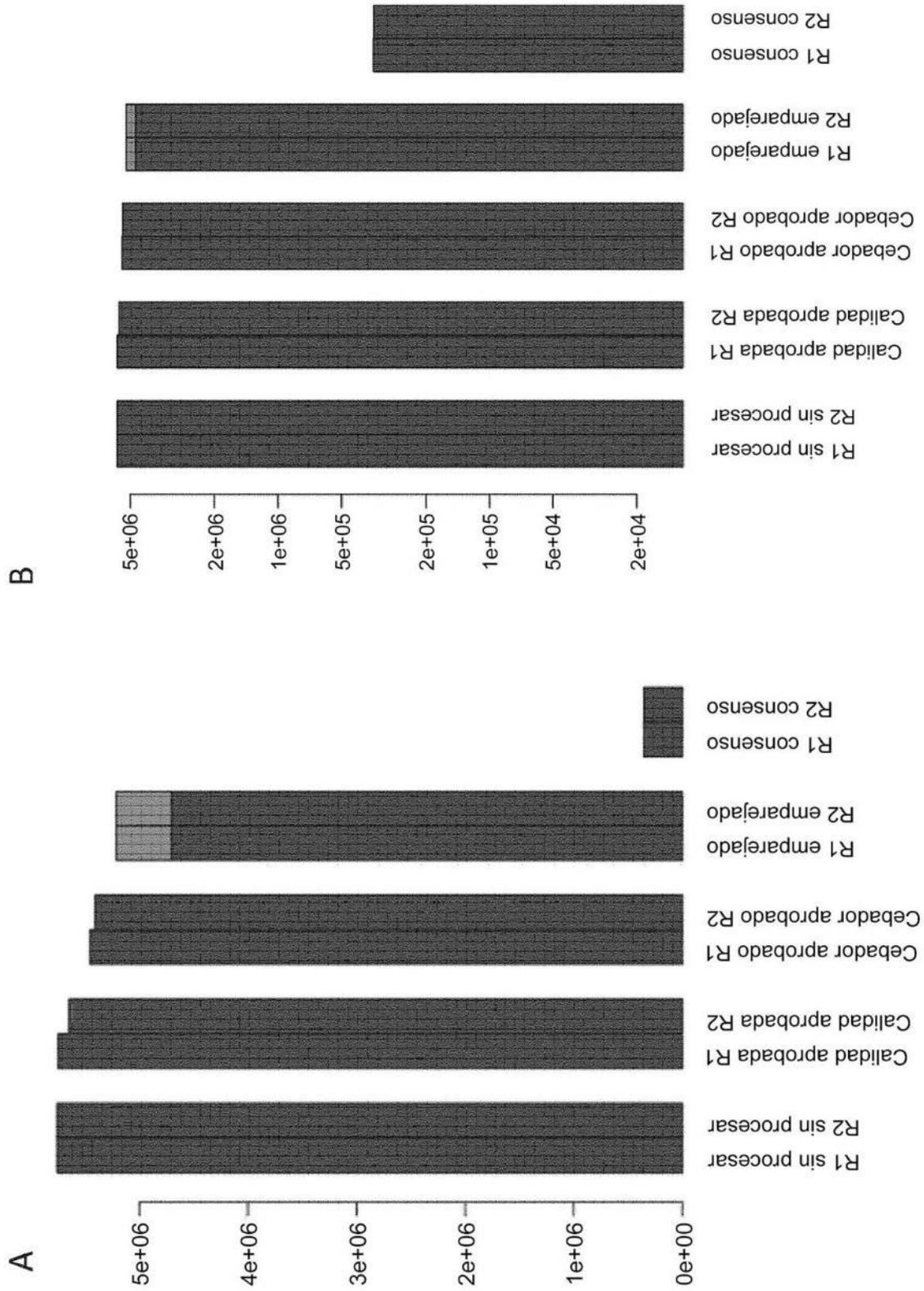


Fig. 49

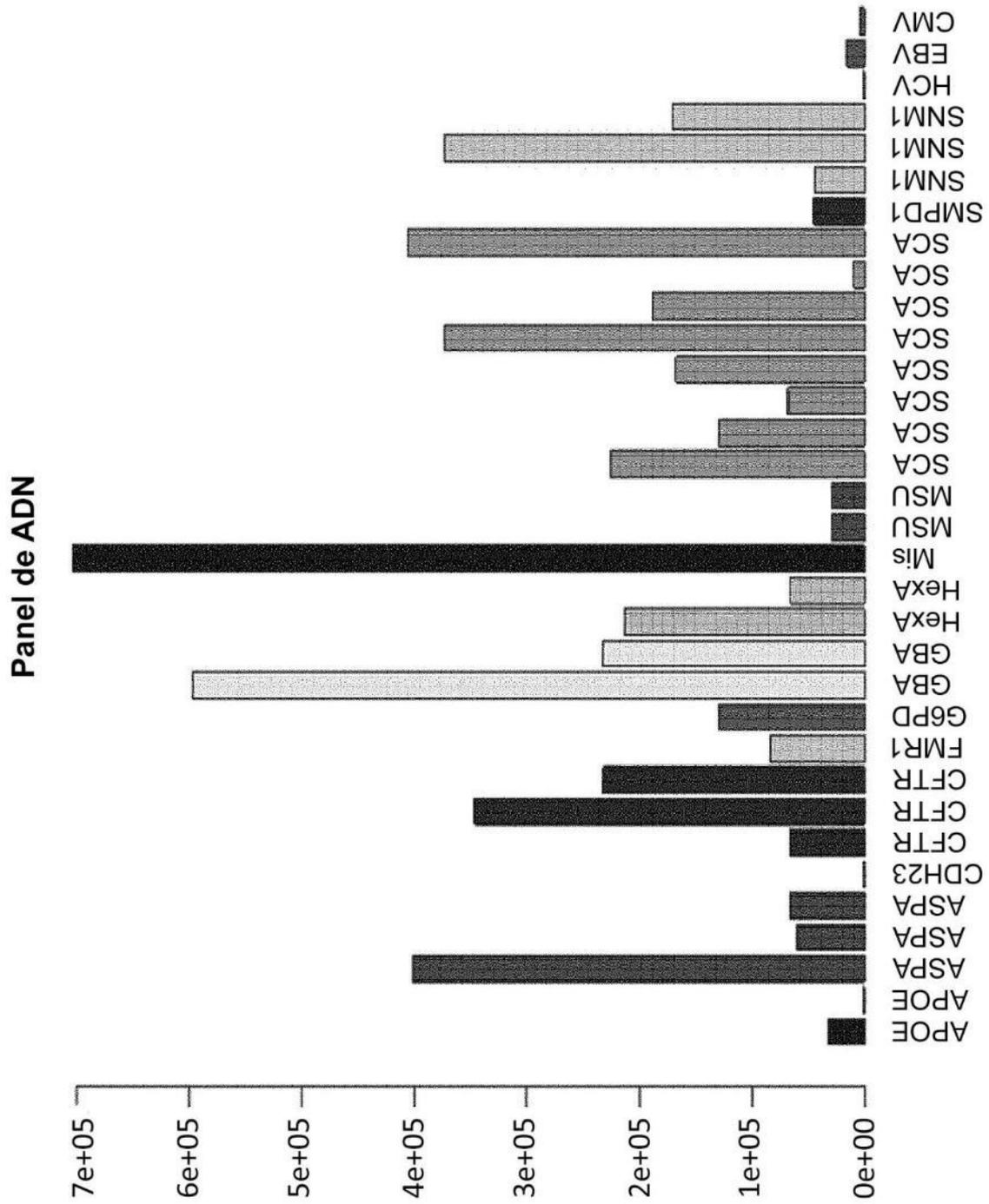


Fig. 50C

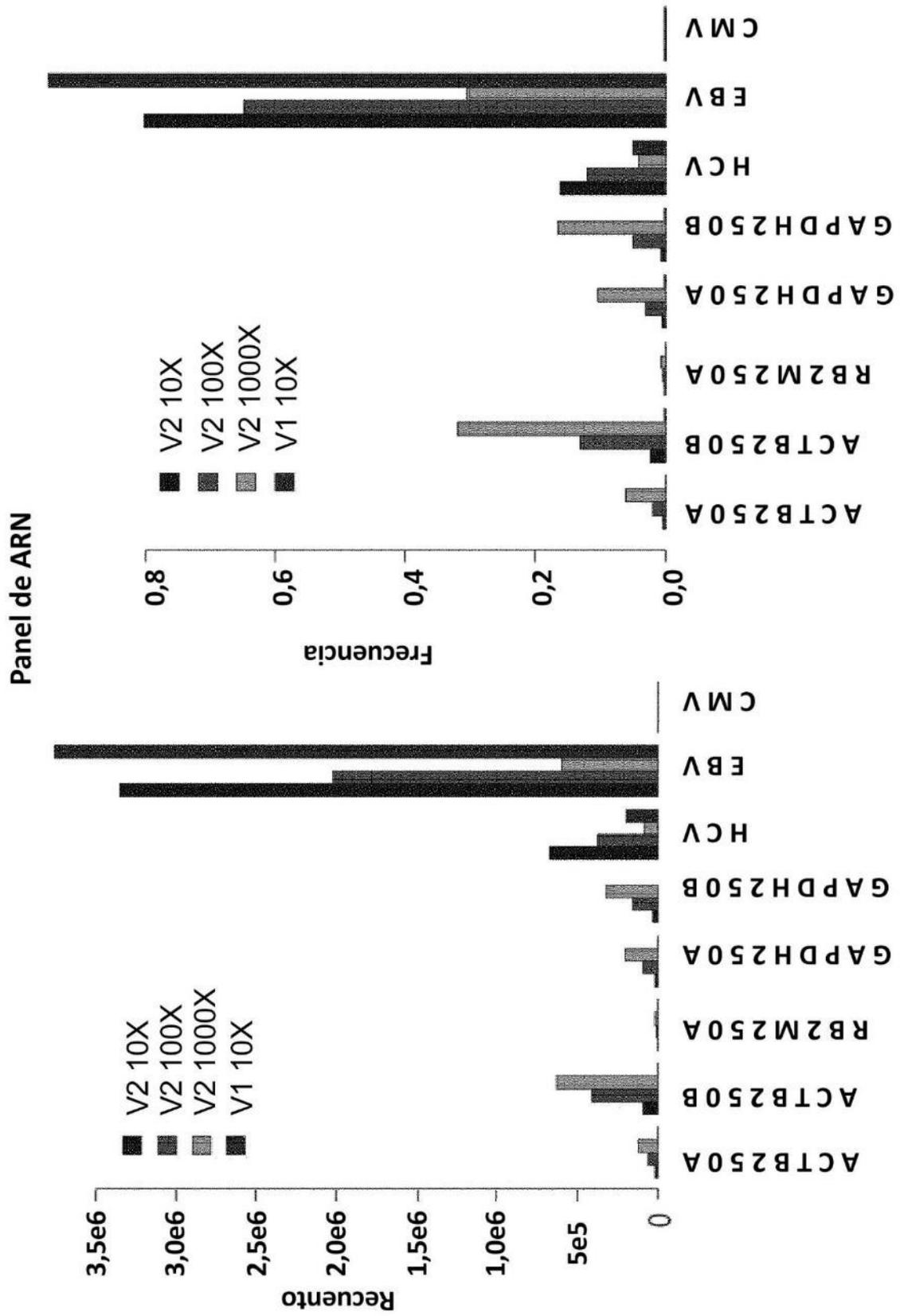


Fig. 51A

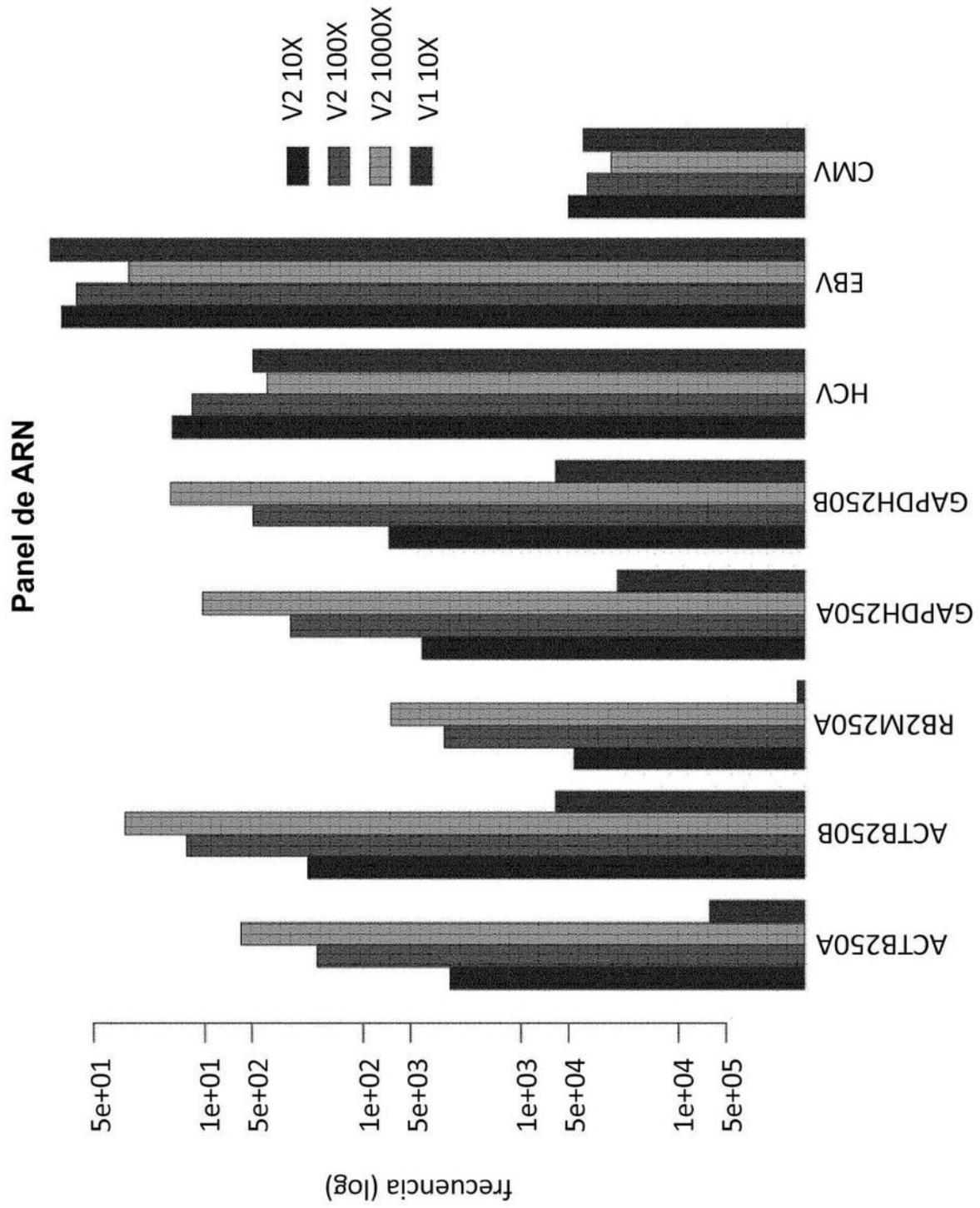


Fig. 51B

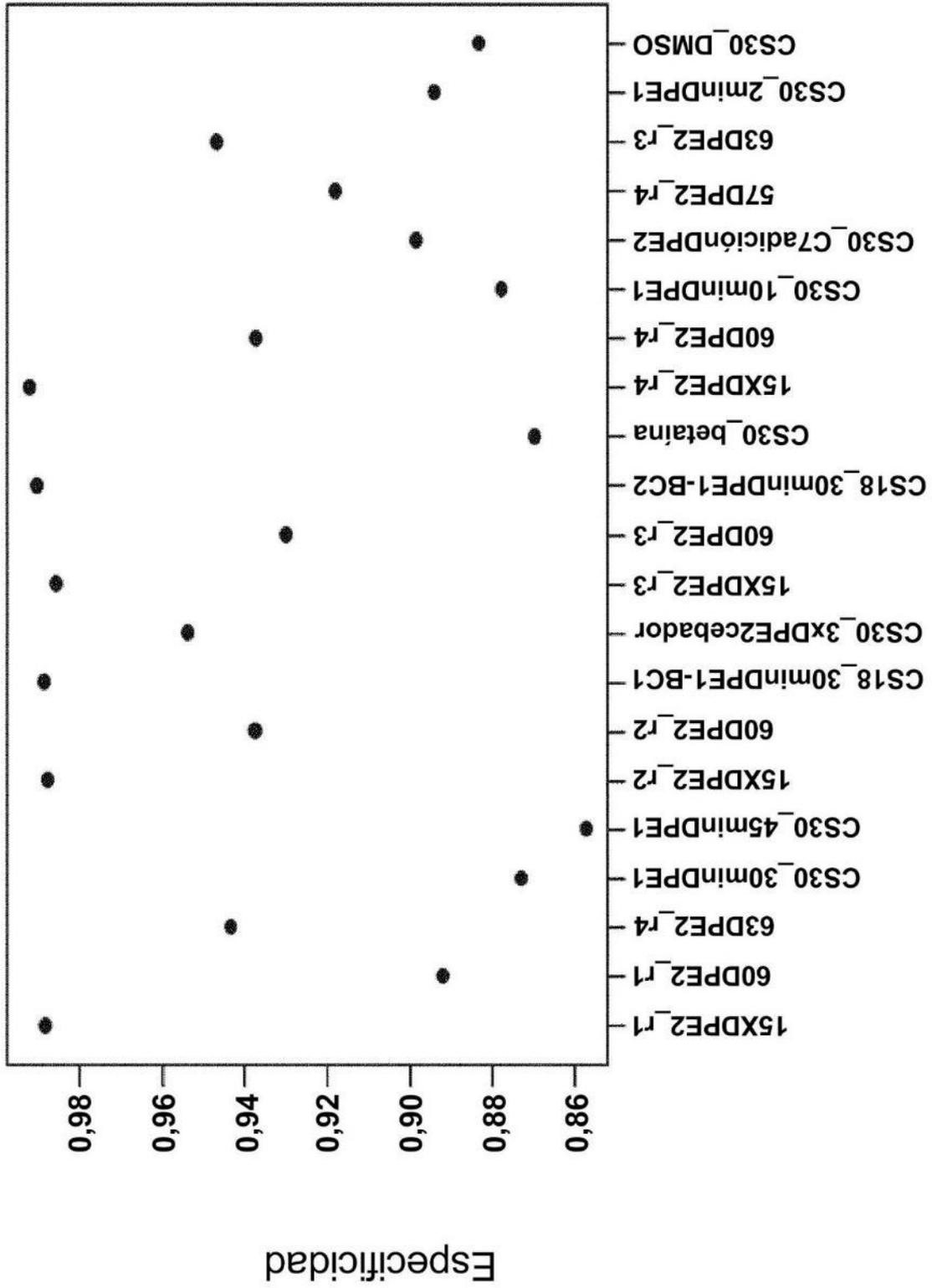


Fig. 52

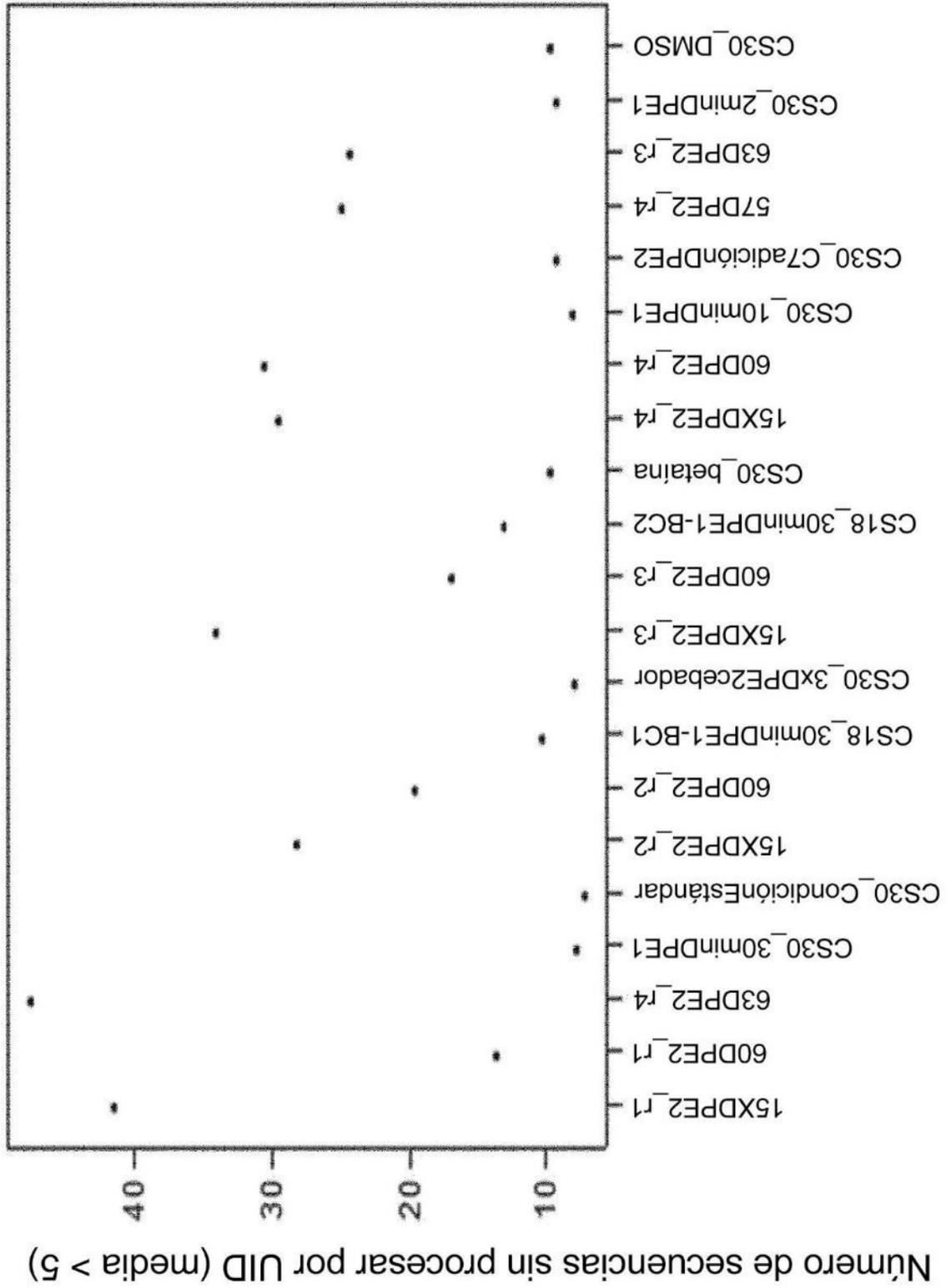


Fig. 53

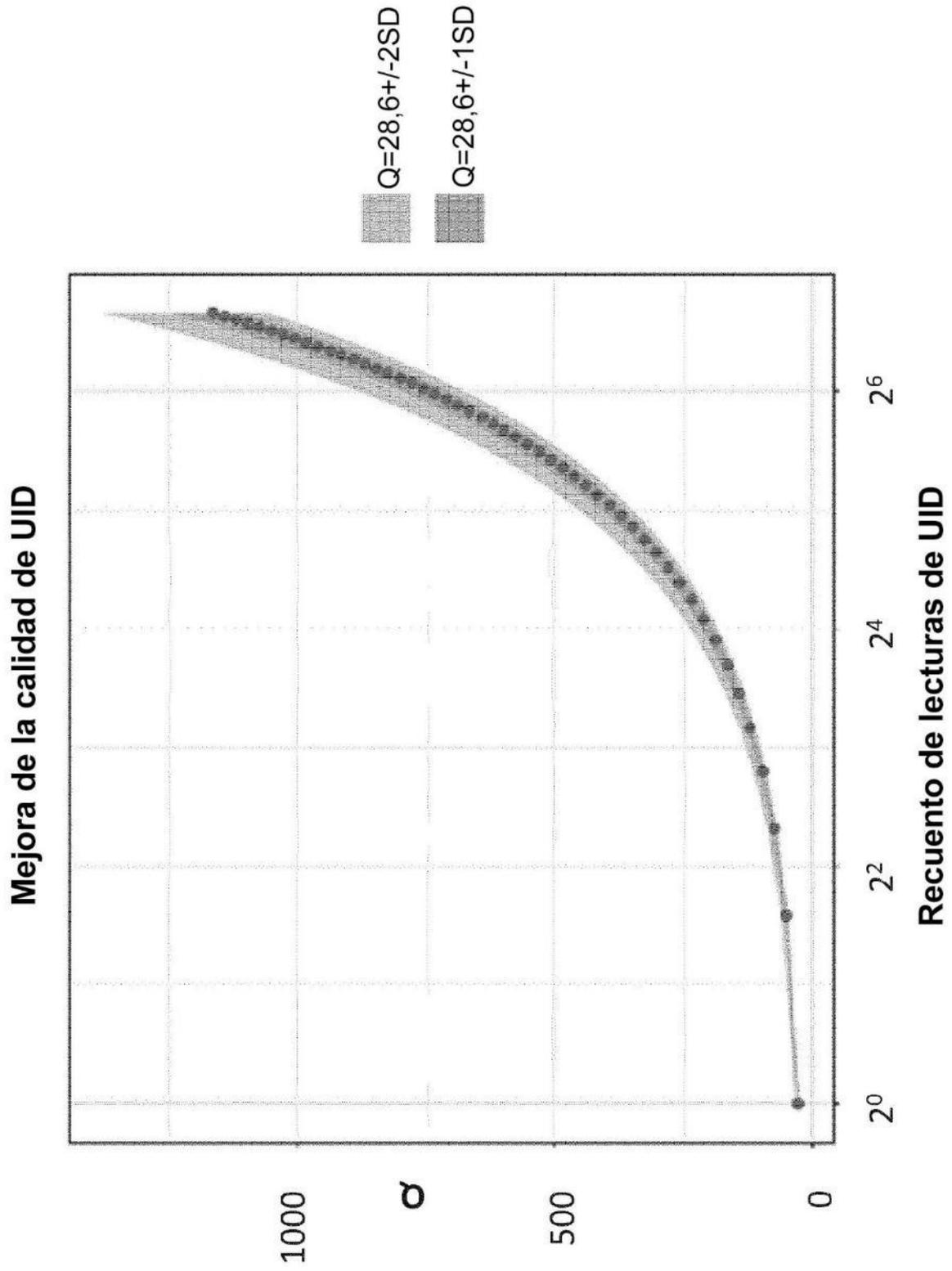


Fig. 54

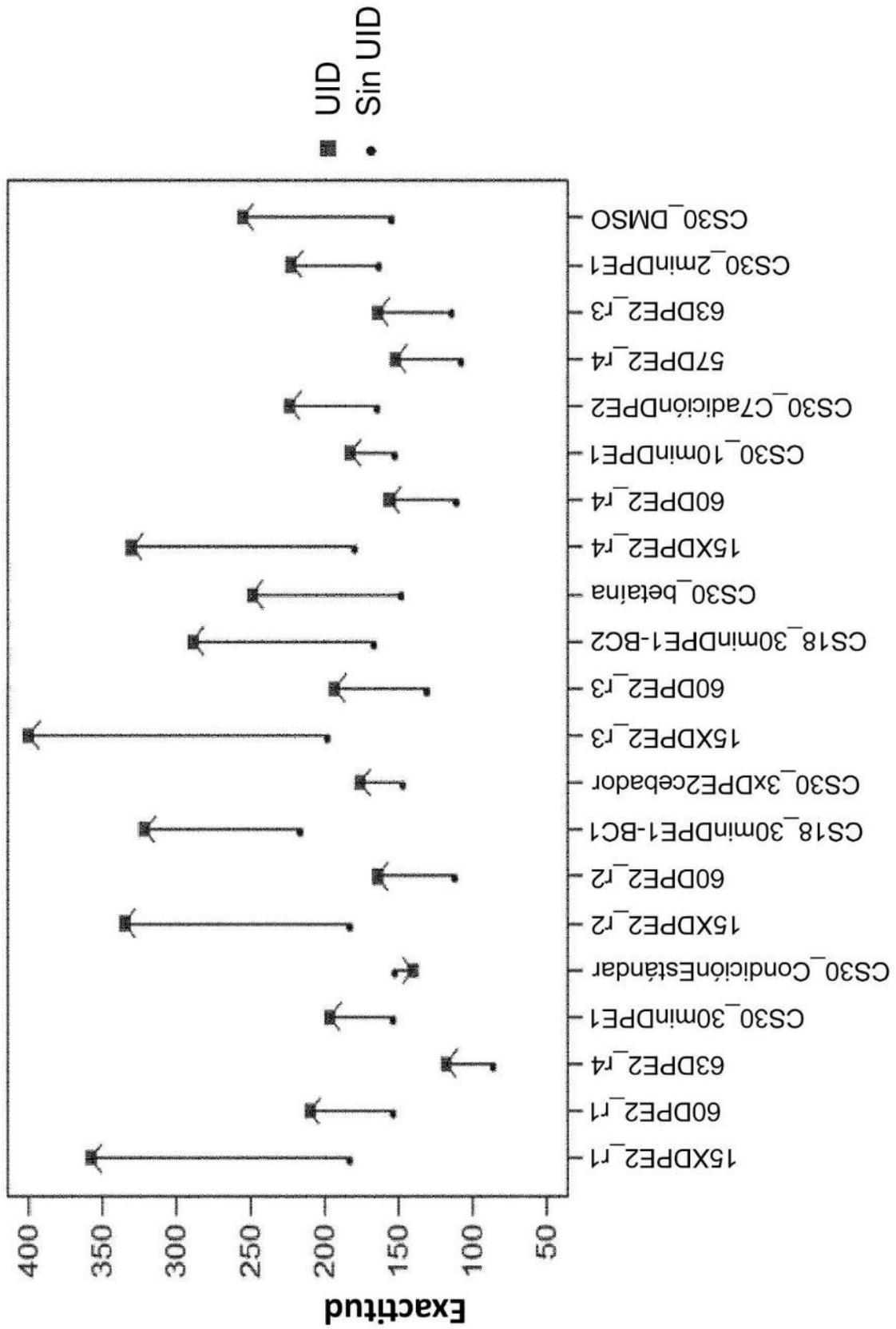


Fig. 55

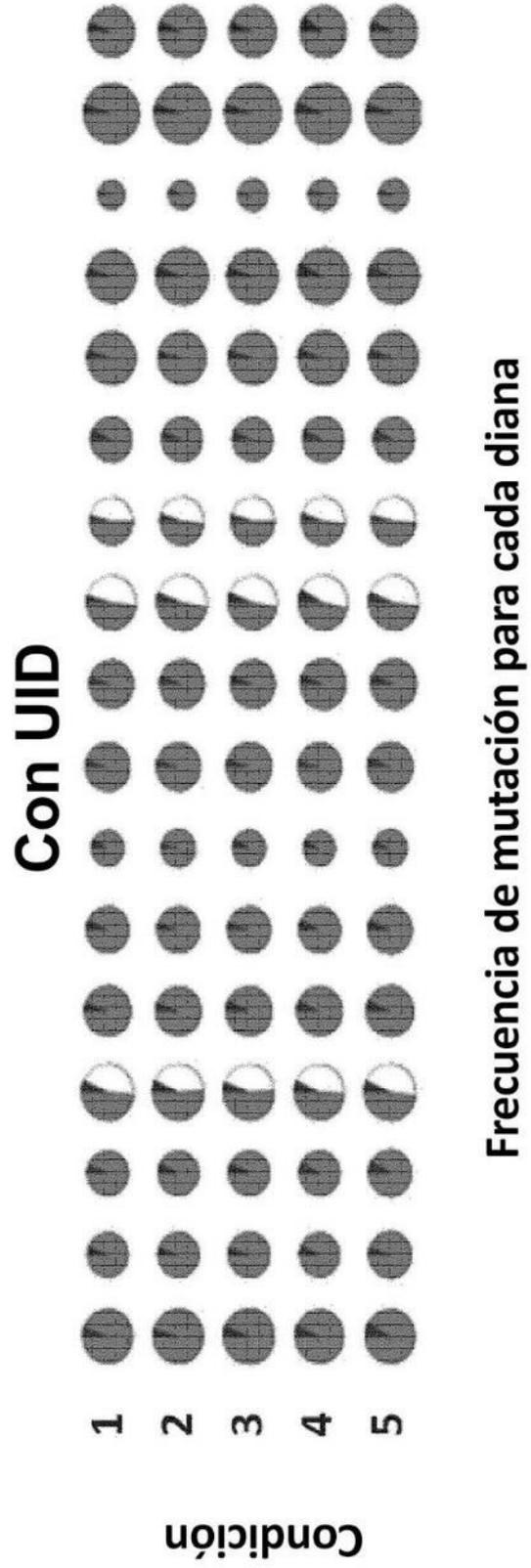
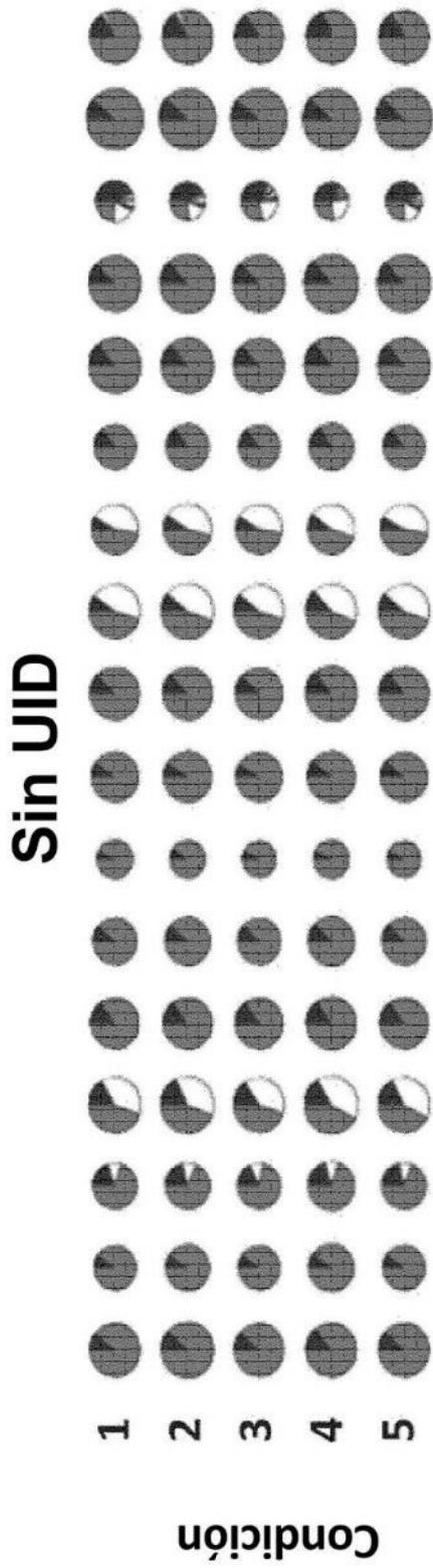


Fig. 56

GBA_alelo2
Genoma_referencia
GBA_alelo1
 ACTACCTTGTCTCTAGATACCCCTGATTACCCGGAGCCCTGCAGTTGGCCAGCGTCCCG 60
 ACTACCTTGTCTCTAGATACCCCTGATTACCCGGAGCCCTGCAGTTGGCCAGCGTCCCG 60
 ACTACCTTGTCTCTAGATACCCCTGATTACCCGGAGCCCTGCAGTTGGCCAGCGTCCCG 60

GBA_alelo2
Genoma_referencia
GBA_alelo1
 TTTCACCTCTGGCAGCCCTGGACATCACCCACTTGGCTCAAGACCAATGGAGCGGTGA 120
 TTTCACCTCTGGCAGCCCTGGACATCACCCACTTGGCTCAAGACCAATGGAGCGGTGA 120
 TTTCACCTCTGGCAGCCCTGGACATCACCCACTTGGCTCAAGACCAAGGGAGCGGGGA 120

GBA_alelo2
Genoma_referencia
GBA_alelo1
 ATGGGAAGGGGTCACTCAAGGGACAGCCCGGAGACATCTACCACAGACCTGGGCCAGAT 180
 ATGGGAAGGGGTCACTCAAGGGACAGCCCGGAGACATCTACCACAGACCTGGGCCAGAT 180
 ATGGGAAGGGGTCACTCAAGGGACAGCCCGGAGACATCTACCACAGACCTGGGCCAGAT 180

GBA_alelo2
Genoma_referencia
GBA_alelo1
 ACTTTGTGAAGTAAGGGATCAGCAAGGATGT 211
 ACTTTGTGAAGTAAGGGATCAGCAAGGATGT 211
 ACATTGTGAAGTAAGGGATCAGCAAGGATGT 211
 ** *****

Fig. 57

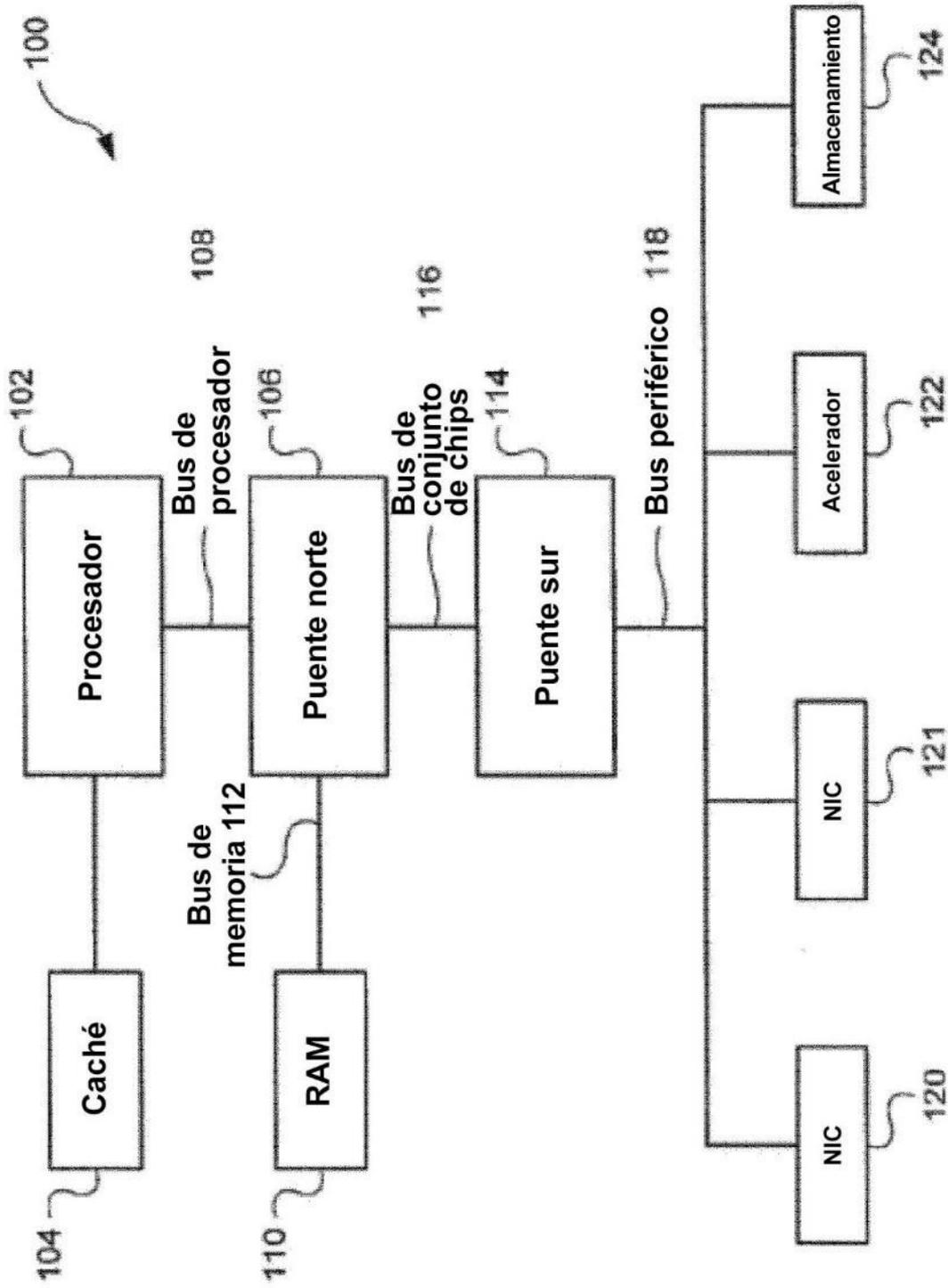


Fig. 58

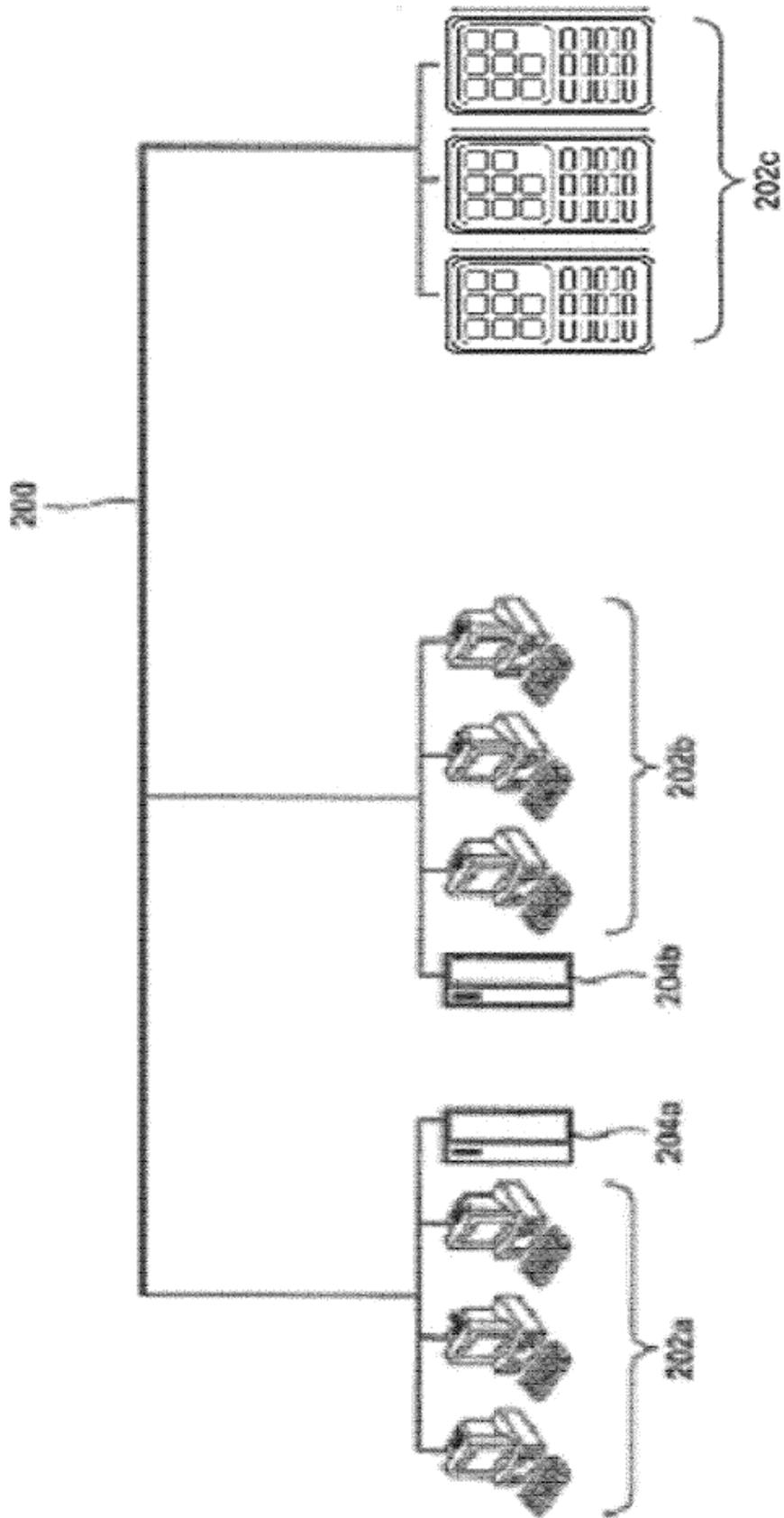


Fig. 59

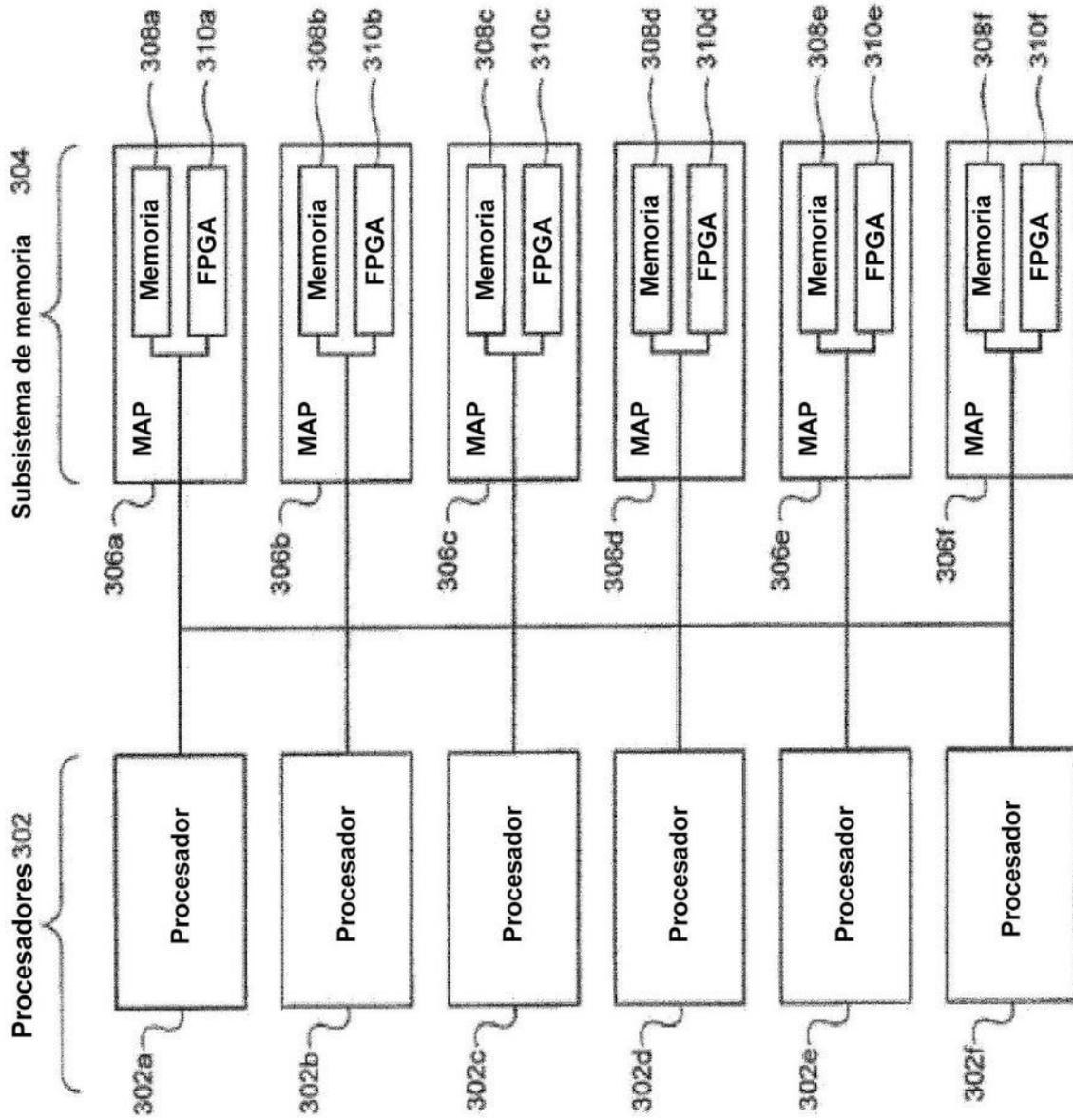


Fig. 60