

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 597 032**

51 Int. Cl.:

C12Q 1/68 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **26.04.2012 PCT/US2012/035260**

87 Fecha y número de publicación internacional: **08.11.2012 WO12151111**

96 Fecha de presentación y número de la solicitud europea: **26.04.2012 E 12779794 (2)**

97 Fecha y número de publicación de la concesión europea: **24.08.2016 EP 2705165**

54 Título: **Mejoras en un ensayo cuantitativo de protección contra la nucleasa (qNPA) y secuenciación (qNPS)**

30 Prioridad:

04.05.2011 US 201161482486 P
21.09.2011 US 201161537492 P
15.12.2011 US 201161576143 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
13.01.2017

73 Titular/es:

HTG MOLECULAR DIAGNOSTICS, INC (100.0%)
3430 E. Global Loop
Tuscon, AZ 85706, US

72 Inventor/es:

SELIGMANN, BRUCE;
THOMPSON, DEBRAH;
VASICEK, TOM y
GORDON, DEBRA A.

74 Agente/Representante:

PONS ARIÑO, Ángel

ES 2 597 032 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

Mejoras en un ensayo cuantitativo de protección contra la nucleasa (qNPA) y secuenciación (qNPS)

5 CAMPO

La presente descripción proporciona procedimientos mejorados de ensayo cuantitativo de protección contra la nucleasa (qNPA) y secuenciación cuantitativa de protección contra la nucleasa (qNPS). Dichos procedimientos pueden usarse en la identificación, detección y/o secuenciación de ácidos nucleicos diana.

10

ANTECEDENTES

Aunque los procedimientos de detección y secuenciación de moléculas de ácidos nucleicos son conocidos, existe todavía la necesidad de procedimientos que permitan el análisis de múltiples muestras o múltiples secuencias de forma simultánea o contemporánea. Por ejemplo, el documento US-2008/007.121 divulga un procedimiento para el análisis de un ácido nucleico diana que combina la hibridación con la digestión de nucleasa seguido por la detección de la diana usando sondas de oligonucleótidos. El documento WO-2008/121.927 divulga la detección de un ácido nucleico diana usando una sonda de protección contra la nucleasa y un conector de detección que es complementario a parte de la sonda de protección contra la nucleasa. Los procedimientos de multiplexación de reacciones de detección o secuenciación de moléculas de ácido nucleico no se han realizado en los niveles más deseados de rendimiento o sencillez.

RESUMEN

Se proporcionan procedimientos que mejoran enormemente los procedimientos anteriores de ensayo cuantitativo de protección contra la nucleasa (qNPA) y secuenciación cuantitativa de protección contra la nucleasa (qNPS) y representan una mejora con respecto a los procedimientos actuales de detección y secuenciación de ácidos nucleicos. Estos procedimientos pueden usarse en la identificación, detección y/o secuenciación de dianas de moléculas de ácidos nucleicos. Los procedimientos usan una sonda de protección contra la nucleasa que incluye una o más secuencias flanqueantes (NPPF). Las NPPF incluyen una secuencia que es complementaria a parte o la totalidad de la molécula de ácido nucleico diana, permitiendo así una unión o hibridación específica entre la molécula de ácido nucleico diana y la NPPF. Por ejemplo, la región de la NPPF que es complementaria a una región de la molécula de ácido nucleico diana se une o se hibrida con esa región de la molécula de ácido nucleico diana con alta especificidad (y en algunos ejemplos puede unirse también a una región de un conector bifuncional). Las NPPF incluyen además una o más secuencias flanqueantes en el extremo 5' y/o el extremo 3' de la NPPF. Así, la una o más secuencias flanqueantes están situadas en 5', 3' o en ambos, con respecto a la secuencia complementaria a la molécula de ácido nucleico diana. Si la NPPF incluye una secuencia flanqueante en el extremo 5' y en el extremo 3', en algunos ejemplos la secuencia de cada NPPF es diferente y no complementaria. La o las secuencias flanqueantes incluyen varios nucleótidos contiguos que tienen una secuencia (por ejemplo, una secuencia de al menos 12 nucleótidos) no presente en una molécula de ácido nucleico presente en la muestra, y proporcionan una secuencia de amplificación y/o hibridación universal. Esta secuencia de amplificación y/o hibridación universal, cuando tiene una secuencia complementaria a al menos una parte de un cebador de amplificación, permite la multiplexación, ya que pueden usarse los mismos cebadores de amplificación para amplificar NPPF específicas para diferentes moléculas de ácidos nucleicos diana. También proporciona una secuencia de hibridación universal para todas las NPPF, que puede usarse para añadir una etiqueta detectable a la NPPF o para capturar y concentrar las NPPF. Por ejemplo, si está presente la misma secuencia flanqueante en NPPF específicas para diferentes moléculas de ácidos nucleicos diana, puede usarse el mismo cebador para amplificar cualquier NPPF que tenga la misma secuencia flanqueante, incluso si la NPPF se dirige a una molécula de ácido nucleico diferente. Por ejemplo, la secuencia flanqueante puede usarse para capturar NPPF, por ejemplo en una superficie. La secuencia flanqueante puede contener una secuencia variable, tal como una secuencia que es específica para cada NPPF específica y puede usarse para capturar esa NPPF en una superficie o para otros fines, tales como identificar la NPPF. Así, en algunos ejemplos, los procedimientos divulgados se usan para detectar o secuenciar varias moléculas de ácidos nucleicos diana diferentes en una muestra usando una pluralidad de NPPF, donde cada NPPF se une específicamente a una molécula de ácido nucleico diana en particular. En un ejemplo, los procedimientos divulgados se usan para detectar o secuenciar al menos una molécula de ácido nucleico diana en una pluralidad de muestras simultáneamente.

La descripción proporciona procedimientos para la detección o la determinación de una secuencia de al menos una molécula de ácido nucleico diana en una muestra. Los procedimientos pueden incluir la puesta en contacto de la

muestra (por ejemplo, una que haya sido calentada para desnaturalizar las moléculas de ácidos nucleicos en la muestra) con al menos una NPPF en condiciones suficientes para que la NPPF se una específicamente a la molécula de ácido nucleico diana. La molécula de NPPF incluye una secuencia complementaria a parte o la totalidad de la molécula de ácido nucleico diana. Esto permite una unión o hibridación específica entre la NPPF y la molécula de ácido nucleico diana. El procedimiento incluye además la puesta en contacto de la muestra con una o más moléculas de ácidos nucleicos que tienen una secuencia que es complementaria a parte o la totalidad de una secuencia flanqueante (dicha molécula se refiere en la presente memoria descriptiva como CFS) en condiciones suficientes para que la secuencia flanqueante se una o se hibride específicamente con la CFS. Puede usarse más de una CFS para hibridarse con una secuencia flanqueante completa (por ejemplo, múltiples CFS individuales pueden hibridarse con una única secuencia flanqueante, de manera que se cubra la secuencia flanqueante completa). Se produce así la generación de moléculas de NPPF que se han unido (hibridado) con la molécula de ácido nucleico diana, así como la o las CFS, generando así una molécula bicatenaria, que puede incluir al menos cuatro secuencias de oligonucleótidos contiguas, donde todas las bases intervienen en la hibridación con una base complementaria

Después de permitir que la molécula de ácido nucleico diana y la o las CFS se unan con las NPPF, el procedimiento puede incluir además la puesta en contacto de la muestra con una nucleasa específica para moléculas de ácidos nucleicos monocatenarias (ss) (o regiones ss de una molécula de ácido nucleico) en condiciones suficientes para eliminar las bases de ácidos nucleicos que no se hibridan con una base complementaria. Así, por ejemplo, las NPPF que no se han unido a moléculas de ácidos nucleicos diana o CFS, así como las moléculas de ácidos nucleicos diana no unidas, aparte de moléculas de ácidos nucleicos ss en la muestra, y las CFS no unidas, se degradarán. Se genera así una muestra digerida que incluye NPPF intactas presentes como aductos bicatenarios con CFS y moléculas de ácido nucleico diana. En algunos ejemplos, el procedimiento incluye además el incremento del pH de la muestra y/o su calentamiento, para disociar o eliminar las moléculas de ácidos nucleicos diana y las CFS que están unidas a las NPPF.

Las NPPF que estaban unidas a la molécula de ácido nucleico diana y las CFS, y sobrevivieron así al tratamiento con la nucleasa, pueden ser amplificadas y/o etiquetadas. Las NPPF en la muestra digerida pueden ser amplificadas usando uno o más cebadores de amplificación, generando así amplicones de NPPF. Al menos un cebador de amplificación incluye una región que es complementaria a parte o la totalidad de la secuencia flanqueante de la NPPF. En algunos ejemplos, la NPPF incluye una secuencia flanqueante en el extremo 5' y en el extremo 3', y se usan dos cebadores de amplificación, donde un cebador de amplificación tiene una región que es complementaria a la secuencia flanqueante del extremo 5' y el otro cebador de amplificación tiene una región que es complementaria a la secuencia flanqueante del extremo 3'.

Alternativamente, en lugar de usar las NPPF que sobrevivieron al tratamiento con la nucleasa, puede usarse la cadena del ácido nucleico diana que se hibridó con la NPPF (como una cadena de ADN) directamente, por ejemplo se amplificó, etiquetó, detectó, secuenció, o combinaciones de lo anterior. Por ejemplo, la cadena del ácido nucleico diana puede amplificarse usando uno o más cebadores de amplificación, generando así amplicones diana, que pueden ser detectados y/o secuenciados. Así, aunque en la presente memoria descriptiva se hace referencia a amplicones de NPPF, se apreciará que los amplicones diana pueden estar sustituidos.

A continuación, los amplicones resultantes (o parte de los mismos, por ejemplo, la parte 3') pueden ser secuenciados o detectados. En un ejemplo, los amplicones se fijan a un sustrato. Por ejemplo, el sustrato puede incluir al menos una sonda de captura que tiene una secuencia complementaria a parte o la totalidad de una secuencia flanqueante en el amplicón de NPPF, permitiendo así la captura de los amplicones de NPPF que tienen la secuencia flanqueante complementaria. Alternativamente, el sustrato puede incluir al menos un anclaje en asociación con un conector bifuncional, donde el conector bifuncional incluye una primera parte que se une específicamente al anclaje y una segunda parte que se une específicamente a una parte de uno de los amplicones de NPPF. A continuación, los amplicones de NPPF capturados pueden ser secuenciados o detectados, determinando así la secuencia de, o detectando, la al menos una molécula de ácido nucleico diana en la muestra.

En otros ejemplos, los amplicones de NPPF son detectados o secuenciados sin captura en una matriz. Por ejemplo, los amplicones de NPPF pueden ser transferidos a una plataforma de secuenciación.

La NPPF puede ser etiquetada con una etiqueta detectable, por ejemplo durante la amplificación, o como una etapa sin amplificación. Alternativamente, una o las dos regiones flanqueantes pueden usarse para hibridar una etiqueta detectable a la NPPF.

Los objetos y características precedentes de la descripción se harán más evidentes a partir la siguiente descripción detallada, que se ofrece con referencia a las figuras adjuntas.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

5

La **FIG. 1** es un diagrama esquemático que muestra una sonda de protección contra la nucleasa de ejemplo que tiene secuencias flanqueantes (NPPF), (100). La NPPF (100) incluye una región (102) que tiene una secuencia que se une específicamente a la secuencia de ácidos nucleicos diana y a un conector bifuncional (o de programación). La NPPF incluye también una secuencia flanqueante en 5' (104) y una secuencia flanqueante en 3' (106).

10

La **FIG. 2** es un diagrama esquemático que muestra las etapas iniciales de un procedimiento de uso de las NPPF (202) para detectar o secuenciar una molécula de ácido nucleico usando los procedimientos divulgados. Las barras en trazo discontinuo representan una NPPF específica para una primera diana (en algunos ejemplos la NPPF se etiqueta con biotina (B)), las barras grises en trazo continuo representan una NPPF específica para una segunda diana (en algunos ejemplos la NPPF se etiqueta con B), las barras verdes en trazo discontinuo representan moléculas de ácidos nucleicos que son complementaria a las secuencias flanqueantes (CFS) (204) de la NPPF, y las barras negras en trazo continuo representan el ácido nucleico diana (200) (por ejemplo, ADN o ARN). La biotina puede añadirse durante la amplificación usando un cebador que es biotina (o digoxina) etiquetada. Alternativamente, el cebador puede ser etiquetado con otra etiqueta (por ejemplo, un fluoróforo), que produce una NPPF que está etiquetada. (1) Una muestra (por ejemplo, células o tejido FFPE) se pone en contacto con tampón de desorganización de muestra (por ejemplo, para permitir la lisis de células y tejidos en la muestra) y se incuba con las NPPF y las CFS. (2) Se digiere el ácido nucleico no unido (por ejemplo, monocatenario) con una nucleasa específica para moléculas de ácidos nucleicos ss (como la nucleasa S1). (3) La nucleasa puede ser inactivada y las NPPF disociadas de las moléculas diana unidas y las CFS unidas, por ejemplo por adición de base y calentamiento. (4) Las NPPF restantes son amplificadas, por ejemplo usando PCR con cebadores (208) apropiados. En algunos ejemplos, los cebadores (208) incluyen una etiqueta detectable, para permitir el etiquetado de los amplicones resultantes (210). Los amplicones resultantes (210) pueden ser detectados (FIG. 3) o secuenciados (FIG. 4).

La **FIG. 3** es un diagrama esquemático que muestra el modo en que los amplicones de NPPF (210) pueden ser (5) capturados en una matriz (212) que incluye conectores bifuncionales (216) asociados con anclajes (214) o que incluye moléculas de captura de ácidos nucleicos (220). El conector bifuncional (216) incluye una región que es complementaria a una región de los amplicones de NPPF (210) (por ejemplo, complementaria a una secuencia que se había hibridado con el ácido nucleico diana), y una región que es complementaria a una parte del anclaje. Las moléculas de captura de ácidos nucleicos (220) incluyen una región que es complementaria a una región de los amplicones de NPPF (210) (por ejemplo, a una secuencia flanqueante o una parte de la misma). (6) En un ejemplo, se usa avidina-peroxidasa de rábano picante (HRP) para detectar las NPPF unidas y (7) se obtiene una imagen de la matriz después de la adición de sustrato. La localización de la señal en la matriz permite la identificación de la señal generada por una molécula de ácido nucleico diana.

La **FIG. 4** es un diagrama esquemático que muestra que los amplicones de NPPF (210) pueden ser (5) secuenciados.

Las **FIGS. 5A-B** son diagramas esquemáticos que muestran detalles de las moléculas de ácidos nucleicos cuando son procesadas durante las etapas de un procedimiento de uso de las NPPF (402) para detectar o secuenciar una molécula de ácido nucleico usando los procedimientos divulgados. Las barras de color de trazo continuo más largas representan moléculas de ácidos nucleicos diana (400), las barras con colores más claros y más oscuros en sus extremos son NPPF (402) específicas para una diana, donde los diferentes extremos coloreados (404) representan las secuencias flanqueantes. El color de la diana se corresponde con el color de su NPPF correspondiente. Las barras de color de trazo continuo más cortas representan moléculas de ácidos nucleicos que son complementarias a las secuencias flanqueantes (CFS) (406) de la NPPF.

Las **FIGS. 6A-F** son dibujos esquemáticos que muestran realizaciones de ejemplo de moléculas de NPPF, incluyendo realizaciones con (A y B) una secuencia flanqueante sólo en un extremo de la NPPF o (C-F) con secuencias flanqueantes en los dos extremos de la NPPF.

55

La **FIG. 7** es un gráfico de barras que muestra el número de amplicones detectados para cada una de siete NPPF especiales. Las barras de error representan una desviación típica con respecto a la media.

La **FIG. 8** es un gráfico de barras que compara las relaciones observadas para cada una de las 7 NPPF especiales

con las relaciones esperadas basándose en la cantidad de NPPF añadidas a la reacción de PCR original.

La **FIG. 9** es un gráfico de barras y tablas que comparan las sondas NPPF detectadas sin (normal) o con amplificación (sensibilidad extrema, ES). Cada experimento se realizó por triplicado. Las reacciones amplificadas por 5 PCR se diluyeron antes de la captura y la medida.

La **FIG. 10** es una tabla que muestra el material de entrada, los tipos de NPPF y las etiquetas experimentales usados para secuenciar lisados de líneas celulares. Cada experimento se realizó por duplicado o triplicado.

10 La **FIG. 11** es un gráfico de barras que muestra los recuentos de secuenciación de cuarenta y seis NPPF a partir de un experimento de secuenciación por triplicado usando los lisados celulares THP1. Las barras de error representan 1 desviación típica con respecto a la media.

Las **FIGS. 12A y 12B** son gráficos lineales de recuentos de secuenciación de NPPF a partir de secuenciación de valoración. Este experimento buscó la linealidad de salida en un intervalo de entrada, así como el intervalo y los límites de detección del procedimiento de recuento qNPS. Como material de entrada se usaron cuatro concentraciones de lisado celular THP1. (A) muestra las ocho NPPF con recuentos mínimos, (B) muestra las cuatro NPPF con recuentos máximos. Este experimento se realizó por triplicado; este gráfico muestra el resultado sólo de un replicado.

20 La **FIG. 13** es un gráfico lineal de recuentos de secuenciación de NPPF para medir miRNA. Como material de entrada se usaron tres concentraciones de lisado celular HepG2. Se muestran los recuentos para cinco NPPF representativas. Este experimento se realizó por triplicado; este gráfico muestra el resultado sólo de un replicado.

25 La **FIG. 14** es un gráfico de barras de recuentos de secuenciación de NPPF que fueron amplificadas usando un intervalo de números de ciclos de PCR y cantidades de entrada. Cada barra representa un experimento qNPA que usa una de tres diferentes concentraciones de entrada y uno de tres números de ciclos de PCR. Todos los experimentos se normalizaron de manera que el número total de lecturas se definió igual, con el fin de facilitar la comparación de los resultados.

30 Las **FIGS. 15A y 15B** son gráficos de barras que representan las NPPF en reacciones por triplicado que se dividieron y (A) se hibridaron en una matriz o bien (B) se secuenciaron y se contaron. Las reacciones por triplicado se promediaron y las barras de error representan una desviación típica con respecto a la media.

35 LISTADO DE SECUENCIAS

Las secuencias de ácidos nucleicos recogidas en la presente memoria descriptiva se muestran usando abreviaturas de letras estándar para las bases de nucleótidos, tal como se define en 37 C.F.R. 1.822. Sólo se muestra una cadena de cada secuencia de ácidos nucleicos, aunque se entiende que la cadena complementaria está incluida en cualquier referencia a la cadena mostrada. En las secuencias proporcionadas:

Las **SEQ ID NO: 1-16** proporcionan secuencias de anclaje de ejemplo de ácidos nucleicos que pueden usarse con los procedimientos divulgados.

45 Las **SEQ ID NO: 17 y 18** proporcionan secuencias flanqueantes en 5' y 3', respectivamente, de ejemplo que pueden usarse con una NPPF.

Las **SEQ ID NO: 19 y 20** proporcionan cebadores de PCR de ejemplo.

50 Las **SEQ ID NO 21-44** proporcionan cebadores de ejemplo que contienen secuencias de códigos de barras presentes en los nucleótidos 25-30.

DESCRIPCIÓN DETALLADA

55 I. Visión general

La presente descripción proporciona procedimientos mejorados de detección o secuenciación de una molécula de ácido nucleico diana, que permite la multiplexación. Los procedimientos divulgados proporcionan varias mejoras con respecto a los procedimientos de secuenciación y detección disponibles. Por ejemplo, dado que los procedimientos

requieren menos procesamiento de las moléculas de ácidos nucleicos diana, los sesgos introducidos por dicho procesamiento pueden reducirse o eliminarse. Por ejemplo, en los procedimientos actuales, por ejemplo cuando la diana es un ARN, los procedimientos emplean normalmente etapas para aislar o extraer el ARN de una muestra, someterlo a RT-PCR, ligar el ARN, o combinaciones de lo anterior. En los procedimientos divulgados, dichas etapas no son necesarias. Como consecuencia, los procedimientos permiten analizar un intervalo de tipos de muestra no asequible por otros medios para secuenciación de detección. Además, se produce así menos pérdida de ARN de la muestra, lo que proporciona un resultado más preciso. Además reduce el sesgo enzimático. Los procedimientos divulgados proporcionan también detección y secuenciación dirigidas de una molécula de ácido nucleico deseada. Así se simplifica enormemente el análisis de datos. Los procedimientos actuales de secuenciación del genoma completo se enfrentan al problema de la gran cantidad de datos generados, y a la necesidad de una bioinformática complicada. Aunque los costes de secuenciación han disminuido, la capacidad de determinar secuencias es muy superior a la capacidad de los investigadores de almacenar, transmitir y analizar los datos. Como consecuencia, en general se generan más datos de los que pueden analizarse en una cantidad razonable de tiempo. Dado que los procedimientos divulgados están direccionados, pueden superarse estos obstáculos. Por ejemplo, la cantidad de datos generados se simplifica, ya que sólo es preciso detectar o secuenciar una parte de la diana. No se necesitan largas lecturas de nucleótidos, ni la alineación de fragmentos de secuencias a una secuencia de referencia. Además, los resultados pueden simplemente contarse, sin necesidad de un análisis de bioinformática complicado.

Por ejemplo, el procedimiento puede usarse para detectar ADN o ARN, mutaciones tales como fusiones, inserciones o deleciones de genes, repeticiones en tándem, polimorfismos de nucleótidos únicos (SNP) y metilación de ADN. El procedimiento usa una sonda, referida en la presente memoria descriptiva como sonda de protección contra la nucleasa, que comprende una secuencia flanqueante (NPPF). El uso de NPPF permite la multiplexación, y conserva la estequiometría de la molécula de ácido nucleico diana detectada o secuenciada, dado que las secuencias flanqueantes en la sonda permiten sitios de unión a cebadores universales para amplificación y permiten la adición de adaptadores de secuenciación y etiquetas experimentales (en el extremo 3' o el extremo 5', o en los dos extremos, por ejemplo para aumentar la multiplexación), sin destruir la estequiometría. Dado que los sitios flanqueantes pueden ser universales, pueden usarse los mismos cebadores para amplificar cualquier NPPF para cualquier secuencia diana, con lo que se permite la multiplexación y la conservación de la estequiometría. En un ejemplo, amplificando los dos extremos de la NPPF, los procedimientos divulgados proporcionan mayor especificidad que los procedimientos qNPA y qNPS anteriores. Sólo las NPPF con secuencias flanqueantes en 3' y 5' intactas se amplificarán exponencialmente, mientras que las NPPF escindidas por la nucleasa no se amplificarán suficientemente para poder ser secuenciadas o detectadas.

Además, los cebadores permiten la adición de etiquetas (por ejemplo, etiquetas experimentales) para permitir la identificación de la diana sin tener que secuenciar toda la NPPF en sí o para permitir combinar muestras de diferentes pacientes en una única ejecución, ya sea en el extremo 3' o en el extremo 5', o en los dos extremos por ejemplo para aumentar la multiplexación, así como adaptadores de secuenciación para permitir la fijación de una secuencia necesaria para una plataforma de secuenciación en particular y la formación de colonias para algunas plataformas de secuenciación). El uso de NPPF también simplifica la complejidad de la muestra que se analiza (por ejemplo, por secuenciación), ya que reduce la muestra que contiene por ejemplo genes completos a las NPPF (o NPPF o amplicones diana). La secuenciación de las NPPF (o la diana hibridada con la NPPF) simplifica el análisis de datos comparado con el necesario para otros procedimientos de secuenciación, reduciendo el algoritmo para simplemente contar las correspondencias con las NPPF que se añadieron a la muestra, en lugar de tener que establecer correspondencias de las secuencias con el genoma y realizar la desconvolución de las múltiples secuencias/genes que se obtienen de los procedimientos de secuenciación estándar. En algunos ejemplos, los procedimientos divulgados aumentan la señal obtenida en comparación con los procedimientos qNPA y qNPS anteriores, tal como un aumento de al menos 10 veces, al menos 100 veces, al menos 125 veces, al menos 150 veces o al menos 200 veces sin dilución sustancial del producto de NPPF antes de realizar la amplificación.

En un ejemplo, la descripción proporciona procedimientos para detectar al menos una molécula de ácido nucleico diana en una muestra (por ejemplo, al menos 2, al menos 3, al menos 4, al menos 5, al menos 10, al menos 20, al menos 30, al menos 40, al menos 50 o al menos 100 moléculas de ácidos nucleicos diana) o para determinar una secuencia de al menos una molécula de ácido nucleico diana en una muestra. En algunos ejemplos, la muestra se calienta para desnaturalizar moléculas de ácidos nucleicos en la muestra, por ejemplo para permitir la posterior hibridación entre la NPPF y las moléculas de ácidos nucleicos diana en la muestra. En algunos ejemplos, la muestra es una muestra lisada. En algunos ejemplos, la muestra es una muestra fija (por ejemplo, una muestra fijada en formalina integrada en parafina (FFPE), tejidos teñidos con hematoxilina y eosina, o tejidos fijados con glutaraldehído). Por ejemplo, las moléculas de ácidos nucleicos diana pueden ser fijas, reticuladas o insolubles.

Los procedimientos pueden incluir la puesta en contacto de la muestra con al menos una sonda de protección contra la nucleasa que comprende una secuencia flanqueante (NPPF) en condiciones suficientes para que la NPPF se una específicamente a la molécula de ácido nucleico diana. En algunos ejemplos, los procedimientos divulgados secuencian o detectan al menos una molécula de ácido nucleico diana en una pluralidad de muestras de forma simultánea o contemporánea. En algunos ejemplos, los procedimientos divulgados secuencian o detectan dos o más moléculas de ácidos nucleicos diana en una muestra (por ejemplo, de forma simultánea o contemporánea). En dicho ejemplo, la muestra se pone en contacto con una pluralidad de NPPF, donde cada NPPF se une específicamente a una molécula de ácido nucleico diana en particular. Por ejemplo, si existen 10 moléculas de ácidos nucleicos diana, la muestra puede ponerse en contacto con 10 NPPF diferentes cada una de ellas específica para una de las 10 dianas. En algunos ejemplos, al menos 10 NPPF diferentes se incuban con la muestra. Sin embargo, se observa que en algunos ejemplos puede usarse más de una NPPF (por ejemplo, 2, 3, 4, 5, 10, 20, o más) específica para una única molécula de ácido nucleico diana, por ejemplo una población de NPPF que son específicas para diferentes regiones de la diana, o una población de NPPF que pueden unirse a la diana y variaciones de las mismas (por ejemplo, las que tienen mutaciones o polimorfismos).

La molécula de NPPF incluye un extremo 5' y un extremo 3', así como una secuencia intermedia que es complementaria a parte o la totalidad de la molécula de ácido nucleico diana. Esto permite una unión o hibridación específica entre la NPPF y la molécula de ácido nucleico diana. Por ejemplo, la región de la NPPF que es complementaria a una región de la molécula de ácido nucleico diana se une o se hibrida con esa región de la molécula de ácido nucleico diana con alta especificidad. La NPPF puede ser complementaria a la totalidad, o una parte, de la secuencia de ácidos nucleicos diana. La molécula de NPPF incluye además una o más secuencias flanqueantes, que están en el extremo 5' y/o el extremo 3' de la NPPF. Así, la una o más secuencias flanqueantes están situadas en 5', 3' o en ambos, con respecto a la secuencia complementaria a la molécula de ácido nucleico diana. Cada secuencia flanqueante incluye varios nucleótidos contiguos, que generan una secuencia que no se encuentra en una molécula de ácido nucleico presente en la muestra (por ejemplo, una secuencia de al menos 12 nucleótidos contiguos). Si la NPPF incluye una secuencia flanqueante en el extremo 5' y en el extremo 3', en algunos ejemplos la secuencia de cada NPPF es diferente y no complementaria a otra.

La o las secuencias flanqueantes proporcionan una secuencia de hibridación/amplificación universal, que es complementaria a al menos una parte de un cebador de amplificación. En algunos ejemplos, la secuencia flanqueante puede incluir (o permitir la adición de) una etiqueta experimental, un adaptador de secuenciación o combinaciones de los mismos. Por ejemplo, la etiqueta experimental puede ser una secuencia complementaria a una sonda de captura que permite la captura de NPPF, por ejemplo en una superficie (por ejemplo, en un lugar específico de la superficie, o en una perla específica). En algunos ejemplos, la etiqueta experimental puede ser una secuencia que identifica una NPPF, tal como una etiqueta específica para un paciente o una secuencia diana en particular, por ejemplo para permitir distinguir una NPPF etiquetada o un grupo de ellas. En algunos ejemplos, se usará el adaptador de secuenciación de una secuencia que permite usar un amplificador de NPPF con una plataforma de secuenciación en particular.

La NPPF puede ser cualquier molécula de ácido nucleico, tal como una molécula de ADN o ARN, y puede incluir bases no naturales. En algunos ejemplos la NPPF es de al menos 35 nucleótidos, por ejemplo de 40 a 80 o de 50 a 150 nucleótidos. La parte de la NPPF que es complementaria a una región de la molécula de ácido nucleico diana puede tener al menos 6 nucleótidos de longitud, por ejemplo al menos 10, al menos 25, o al menos 60, por ejemplo 6 a 60 nucleótidos de longitud. La o las secuencias flanqueantes de la NPPF pueden tener al menos 6 nucleótidos, al menos 12 nucleótidos o al menos 25 nucleótidos, por ejemplo de 12 a 50 nucleótidos de longitud. En algunos ejemplos, la NPPF incluye dos secuencias flanqueantes: una en el extremo 5' y la otra en el extremo 3'. En algunos ejemplos, la secuencia flanqueante en el extremo 5' difiere de la secuencia flanqueante en el extremo 3'. Además, si la NPPF incluye dos secuencias flanqueantes, idealmente las dos secuencias flanqueantes tienen una temperatura de fusión (T_f) similar, por ejemplo una T_f de +/- 5°C.

El procedimiento incluye además la puesta en contacto de la muestra con una molécula de ácido nucleico que tiene una secuencia que es complementaria a la secuencia flanqueante (dicha molécula se refiere en la presente memoria descriptiva como CFS) en condiciones suficientes para que la secuencia flanqueante se una o se hibride específicamente con la CFS. Un experto en la materia observará que en lugar de usar una única CFS para proteger una secuencia flanqueante, pueden usarse múltiples CFS para proteger una secuencia flanqueante. Se produce así la generación de moléculas de NPPF que se han unido a la molécula de ácido nucleico diana, así como la CFS, generando así una molécula bicatenaria que incluye al menos tres secuencias de oligonucleótidos contiguas, donde todas las bases intervienen en la hibridación con una base complementaria, donde las bases de la NPPF y las CFS pueden incluir bases no naturales. La CFS se hibrida y así protege su secuencia flanqueante correspondiente de la

digestión con la nucleasa en etapas posteriores. En algunos ejemplos, cada CFS es la longitud exacta de su secuencia flanqueante correspondiente. En algunos ejemplos, la CFS es completamente complementaria a su secuencia flanqueante correspondiente. Sin embargo, un experto en la materia observará que el extremo 3' de una CFS que protege una secuencia flanqueante del extremo 5' o el extremo 5' de una CFS que protege la secuencia flanqueante del extremo 3' puede haber tenido una diferencia, tal como un nucleótido en cada una de estas posiciones.

Después de permitir que la molécula de ácido nucleico diana, así como la o las CFS, se una a las NPPF, el procedimiento puede incluir además la puesta en contacto de la muestra con una nucleasa específica para moléculas de ácidos nucleicos monocatenarias (ss) o regiones ss de una molécula de ácido nucleico, tal como nucleasa S1, en condiciones suficientes para eliminar las bases de ácidos nucleicos que no se han hibridado con una base complementaria. Así, por ejemplo, las NPPF que no se han unido a la molécula de ácido nucleico diana o las CFS, así como las moléculas de ácidos nucleicos diana no unidas, otras moléculas de ácidos nucleicos ss en la muestra, y las CFS no unidas, se degradarán. Así se genera una muestra digerida que incluye NPPF intactas presentes como aductos bicatenarios hibridados con las CFS y los ácidos nucleicos diana. En algunos ejemplos, por ejemplo si la NPPF está compuesta por ADN, la nucleasa puede incluir una exonucleasa, una endonucleasa o una combinación de las mismas.

En algunos ejemplos, el procedimiento incluye además el aumento del pH de la muestra y/o su calentamiento, por ejemplo para inactivar la nucleasa, para eliminar moléculas de ácidos nucleicos diana y CFS que están unidas a las NPPF, o combinaciones de lo anterior. En algunos ejemplos, el procedimiento incluye la liberación del ácido nucleico diana (por ejemplo, un ADN) de la NPPF, y el análisis posterior de la diana liberada (por ejemplo, detección o secuenciación de la diana). En algunos ejemplos el ácido nucleico diana es ADN, y el ADN se amplifica antes de su detección o secuenciación.

Las NPPF que se unieron a la molécula de ácido nucleico diana y las CFS y sobrevivieron así al tratamiento con la nucleasa pueden ser amplificadas, por ejemplo, usando amplificación PCR. Las NPPF en la muestra digerida pueden amplificarse usando uno o más cebadores de amplificación, generando así amplicones de NPPF. Al menos un cebador de amplificación incluye una región que es complementaria a una secuencia flanqueante de NPPF. En algunos ejemplos, la NPPF incluye una secuencia flanqueante en el extremo 5' y en el extremo 3', y se usan dos cebadores de amplificación, donde un cebador de amplificación tiene una región que es complementaria a la secuencia flanqueante del extremo 5' y el otro cebador de amplificación tiene una región que es complementaria a la secuencia flanqueante del extremo 3'. Uno o los dos cebadores de amplificación pueden incluir una secuencia que permite la fijación de una etiqueta experimental o un adaptador de secuenciación al amplicón de NPPF durante la amplificación, y uno o los dos cebadores pueden etiquetarse para permitir etiquetar el amplicón de NPPF. En algunos ejemplos, se añade una etiqueta experimental y un adaptador de secuenciación, por ejemplo en extremos opuestos del amplicón de NPPF. Por ejemplo, el uso de dichos cebadores puede generar una etiqueta experimental o una etiqueta de secuencias que se extiende desde el extremo 5' o desde el extremo 3' del amplicón de NPPF, o tanto desde el extremo 3' como del extremo 5' para aumentar el grado de multiplexación posible. La etiqueta experimental puede incluir una única secuencia de ácidos nucleicos que permite la identificación de una muestra, sujeto o secuencia de ácidos nucleicos diana. En algunos ejemplos, el cebador de amplificación contiene una etiqueta experimental que permite la captura del amplicón de NPPF en un sustrato (por ejemplo, por hibridación con una sonda en el sustrato que tiene una secuencia complementaria a la secuencia de captura en el amplicón de NPPF). El adaptador de secuenciación puede incluir una secuencia de ácidos nucleicos que permite la captura de la NPPF resultante en una plataforma de secuenciación. Por ejemplo, el cebador de amplificación puede incluir una secuencia que permite la fijación de una etiqueta de secuencias poli A o poli T que puede facilitar la amplificación una vez capturada en el chip de secuenciación. En algunos ejemplos, el cebador de amplificación se usa para etiquetar el amplicón de NPPF. En otros ejemplos, se usa una o las dos regiones flanqueantes para hibridar una etiqueta detectable a la NPPF, por ejemplo con una sonda etiquetada (por ejemplo, sin amplificación).

Los amplicones de NPPF (o diana) resultantes (o parte de los mismos, por ejemplo, una parte en 3') pueden ser secuenciados o detectados a continuación, determinando así la secuenciación, o detección, de la al menos una molécula de ácido nucleico diana en la muestra.

En un ejemplo, los amplicones de NPPF (o parte de los mismos) son secuenciados. Puede usarse cualquier procedimiento para secuenciar los amplicones de NPPF, y la descripción no se limita a procedimientos de secuenciación en particular. En algunos ejemplos, el procedimiento de secuenciación usado es secuenciación Solexa®, secuenciación 454®, secuenciación de terminación de cadena, secuenciación de terminación con colorante o pirosecuenciación. En algunos ejemplos, se usa secuenciación de moléculas individuales. En algunos

ejemplos en que los amplicones de NPPF son secuenciados, el procedimiento incluye también la comparación de la secuencia de NPPF obtenida con una base de datos de secuencias de referencia; y la determinación del número de cada secuencia de NPPF identificada.

5 En algunos ejemplos, los amplicones de NPPF son objeto de detección. En dichos ejemplos, el procedimiento puede incluir la puesta en contacto de los amplicones de NPPF con una superficie, por ejemplo, una que tenga múltiples regiones espacialmente discretas. En un ejemplo, los amplicones de NPPF son capturados por una o más moléculas de captura de ácidos nucleicos en la superficie, donde las secuencias de las moléculas de captura de ácidos nucleicos en la superficie son complementarias a al menos una parte de una secuencia flanqueante en el amplicón de NPPF. Esta complementariedad permite la hibridación y la unión de los amplicones de NPPF a las moléculas de captura en la superficie. Dichas moléculas de captura pueden conjugarse directamente con la superficie. Los amplicones de NPPF se incuban o se ponen en contacto con la superficie en condiciones suficientes para que los amplicones de NPPF se unan específicamente a las moléculas de captura en la superficie. En algunos ejemplos, los amplicones de NPPF se ponen en contacto con una población de superficies, donde la población de superficies incluye subpoblaciones de superficies (por ejemplo, una población de perlas), y donde cada subpoblación de superficies comprende al menos una molécula de captura de ácidos nucleicos complementaria a al menos una parte de una secuencia flanqueante en el amplicón de NPPF. Así, esto permite la captura de todas las NPPF que tienen una secuencia complementaria a las moléculas de captura en la superficie, con independencia de la secuencia direccionada por la NPPF. A continuación, los amplicones de NPPF unidos pueden ser detectados. En algunos ejemplos, esta etapa se usa para purificar o concentrar los amplicones de NPPF (por ejemplo, a partir de una mezcla que contiene cebadores), y los amplicones de NPPF pueden liberarse posteriormente de la superficie, por ejemplo invirtiendo la hibridación (por ejemplo, aumentando la temperatura para fundir las NPPF capturadas o cambiando el pH y la temperatura), y los amplicones de NPPF analizados.

25 En otro ejemplo, los amplicones de NPPF son capturados en una superficie usando anclajes y conectores bifuncionales. La superficie puede incluir una pluralidad de regiones, incluyendo cada región al menos un anclaje en asociación con un conector bifuncional. El conector bifuncional incluye una primera parte que se une específicamente al anclaje y una segunda parte que se une específicamente a o se hibrida con al menos una parte de uno de los amplicones de NPPF. Los amplicones de NPPF se incuban o se ponen en contacto con la superficie en condiciones suficientes para que los amplicones de NPPF se unan específicamente a la segunda parte del conector bifuncional. En algunos ejemplos, los amplicones de NPPF se ponen en contacto con una población de superficies, donde la población de superficies incluye subpoblaciones de superficies (por ejemplo, una población de perlas), y donde cada subpoblación de superficies comprende al menos un anclaje en asociación con un conector bifuncional. A continuación los amplicones de NPPF unidos pueden ser detectados.

35 Además, el amplicón de NPPF puede incluir una etiqueta detectable que permite así su detección. En algunos ejemplos, dicha etiqueta se introduce durante la amplificación. En ejemplos específicos, la etiqueta detectable es un hapteno, una molécula fluorescente, una enzima, o un radioisótopo. Por ejemplo, la biotina presente en un amplicón de NPPF puede ser detectada por puesta en contacto de los amplicones de NPPF con avidina o estreptavidina conjugadas con peroxidasa de rábano picante o fosfatasa alcalina

II. Términos

A no ser que se indique lo contrario, los términos técnicos se usan de acuerdo con el uso convencional. Las definiciones de los términos comunes en biología molecular pueden encontrarse en Benjamin Lewin, *Genes VII*, publicado por Oxford University Press, 2000 (ISBN 019879276X); Kendrew y col. (eds.), *The Encyclopedia of Molecular Biology*, publicado por Blackwell Publishers, 1994 (ISBN 0632021829); Robert A. Meyers (ed.), *Molecular Biology and Biotechnology: a Comprehensive Desk Reference*, publicado por Wiley, John & Sons, Inc., 1995 (ISBN 0471186341); y George P. Rédei, *Encyclopedic Dictionary of Genetics, Genomics, and Proteomics*, 2ª edición, 2003 (ISBN: 0-471-26821-6).

Las siguientes explicaciones de términos y procedimientos se proporcionan para describir mejor la presente descripción y para orientar a los expertos en la materia en la práctica de la presente descripción. Las formas en singular "un", "una" y "el/la" se refieren a uno o a más de uno, salvo que el contexto indique claramente lo contrario. Por ejemplo, el término "que comprende una célula" incluye células individuales o varias células y se considera equivalente a la frase "que comprende al menos una célula". El término "o" se refiere a un único elemento de elementos alternativos indicados o a una combinación de dos o más elementos, a no ser que el contexto indique claramente lo contrario. Tal como se usa en la presente memoria descriptiva, "comprende" significa "incluye". Así, "que comprende A o B", significa "que incluye A, B, o A y B", sin excluir elementos adicionales.

Para facilitar la revisión de las diversas realizaciones de la presente descripción, se proporcionan las siguientes explicaciones de términos específicos:

- 5 **Extremo 3'**: El extremo de una molécula de ácido nucleico que no tiene un nucleótido unido a su 3' del residuo terminal.
- Extremo 5'**: El extremo de una secuencia de ácidos nucleicos donde la posición 5' de los residuos terminales no está unida por un nucleótido.
- Amplificación de una molécula de ácido nucleico**: Para aumentar el número de copias de una molécula de ácido nucleico, tal como una NPPF o parte de los mismos. Los productos resultantes se denominan productos de amplificación o amplicones. Un ejemplo de amplificación *in vitro* es la reacción en cadena de la polimerasa (PCR), en la que una muestra (por ejemplo, una muestra que contiene NPPF) se pone en contacto con un par de cebadores de oligonucleótidos, en condiciones que permiten la hibridación de los cebadores con una molécula de ácido nucleico en la muestra. Los cebadores se extienden en condiciones adecuadas, se disocian de la plantilla y después se vuelven a hibridarse, se extienden y se disocian para amplificar el número de copias de la molécula de ácido nucleico.
- 10 **Unión o unión estable (de un ácido nucleico)**: Una primera molécula de ácido nucleico (por ejemplo, una NPPF) se une o se une de forma estable con otra molécula de ácido nucleico (por ejemplo, una molécula de ácido nucleico diana) si una cantidad suficiente de la primera molécula de ácido nucleico forma pares de bases o se hibrida con la otra molécula de ácido nucleico, por ejemplo la unión de una NPPF con su secuencia de ácidos nucleicos diana complementaria. La unión puede detectarse por sus propiedades físicas o funcionales.
- La unión entre moléculas de ácidos nucleicos puede detectarse mediante cualquier procedimiento conocido para un experto en la materia, que incluye ensayos de unión funcionales (por ejemplo, reducción en la expresión y/o la actividad) y físicos.
- 25 **Complementario**: Capacidad para formar pares de bases entre ácidos nucleicos. Los oligonucleótidos y sus análogos se hibridan por unión de hidrógeno, lo que incluye unión de hidrógeno de Watson-Crick, de Hoogsteen o de Hoogsteen invertida, entre base complementarias. Generalmente, las moléculas de ácidos nucleicos consisten en bases nitrogenadas que son pirimidinas (citosina (C), uracilo (U), y timina (T)) o purinas (adenina (A) y guanina (G)). Estas bases nitrogenadas forman enlaces de hidrógeno entre una pirimidina y una purina, y la unión de la pirimidina con la purina se refiere como "apareamiento de bases". Más específicamente, A formará enlaces de hidrógeno con T o U, y G se unirá con C. "Complementario" se refiere al apareamiento de bases que tiene lugar entre dos ácidos nucleicos distintos o dos regiones distintas del mismo ácido nucleico. "Específicamente hibridable" y "específicamente complementario" son términos que indican un grado suficiente de complementariedad de manera que tenga lugar una unión estable y específica entre la sonda (por ejemplo, una NPPF) o su análogo y el ácido nucleico diana (por ejemplo, ADN o ARN diana). La sonda o análogo no ha de tener una complementariedad del 100% con su secuencia diana para ser específicamente hibridable. Una sonda o análogo es específicamente hibridable cuando existe un grado de complementariedad suficiente para evitar una unión no específica de la sonda o análogo con secuencias no diana en condiciones donde se desea unión específica, por ejemplo, en los procedimientos divulgados en la presente memoria descriptiva.
- 30 **Condiciones suficientes para**: Cualquier entorno que permita la actividad deseada, por ejemplo, que permita la unión o hibridación específica entre dos moléculas de ácidos nucleicos (por ejemplo, una NPPF y un ácido nucleico diana, una NPPF y una CFS, o entre una NPPF y un conector bifuncional) o que permita que una nucleasa elimine (o digiera) ácidos nucleicos no unidos.
- Contacto**: Puesta en asociación física directa; incluye forma sólida y líquida. Por ejemplo, la puesta en contacto puede tener lugar *in vitro* con una sonda de ácido nucleico (por ejemplo, una NPPF) y una muestra biológica en solución.
- 45 **Detectar**: Determinar si un agente (por ejemplo, una señal, un nucleótido en particular, un aminoácido, una molécula de ácido nucleico y/o un organismo) está presente o ausente. En algunos ejemplos, esto puede incluir además una cuantificación. Por ejemplo, el uso de los procedimientos divulgados permite la detección de moléculas de ácidos nucleicos diana en una muestra.
- 50 **Etiqueta detectable**: Un compuesto o composición que se conjuga directa o indirectamente con otra molécula (por ejemplo, una molécula de ácido nucleico, por ejemplo una NPPF o un cebador de amplificación/sonda) para facilitar la detección de esa molécula. Los ejemplos específicos no limitativos de etiquetas incluyen fracciones fluorescentes y fluorógenas, fracciones cromógenas, haptenos, etiquetas de afinidad e isótopos radiactivos. La etiqueta puede ser detectable directamente (por ejemplo, detectable ópticamente) o detectable indirectamente (por ejemplo, por interacción con una o más adicional moléculas que a su vez son detectables). A continuación se describen etiquetas de ejemplo en el contexto de las sondas divulgadas en la presente memoria descriptiva. Los procedimientos para etiquetar ácidos nucleicos, y orientación sobre la elección de etiquetas útiles para diversos fines se abordan, por ejemplo, en Sambrook y Russell, en *Molecular Cloning: A Laboratory Manual*, 3ª Ed., Cold Spring Harbor Laboratory

Press (2001) y Ausubel y col., en Current Protocols in Molecular Biology, Greene Publishing Associates y Wiley-Intersciences (1987, y que incluye actualizaciones).

Hibridación: Capacidad de ADN, ARN o híbridos ADN/ARN monocatenarios complementarios de formar una molécula dúplex (también referida como complejo de hibridación). Las técnicas de hibridación de ácidos nucleicos pueden usarse para formar complejos de hibridación entre una sonda de ácidos nucleicos, y el gen se diseña como objetivo.

"Específicamente hibridable" y "específicamente complementario" son términos que indican un grado suficiente de complementariedad de manera que tiene lugar una unión específica y estable entre una primera molécula de ácido nucleico (o su análogo) y una segunda molécula de ácido nucleico (por ejemplo, un ácido nucleico diana, por ejemplo, un ADN o ARN diana). No es necesario que las moléculas de ácidos nucleicos sean complementarias al 100% para ser específicamente hibridables. La hibridación específica se refiere también en la presente memoria descriptiva como "unión específica".

Las condiciones de hibridación que producen grados especiales de astringencia variarán dependiendo de la naturaleza del procedimiento de hibridación y de la composición y la longitud de las secuencias de hibridación de ácidos nucleicos. Generalmente, la temperatura de hibridación y la fuerza iónica (por ejemplo, la concentración Na^+) del tampón de hibridación determinarán la astringencia de hibridación. Los cálculos relativos a las condiciones de hibridación para alcanzar grados determinados de astringencia se exponen en Sambrook y col., (1989) Molecular Cloning, segunda edición, Cold Spring Harbor Laboratory, Plainview, NY (capítulos 9 y 11).

Nucleasa: Una enzima que escinde un enlace de fosfodiéster. Una endonucleasa es una enzima que escinde un enlace de fosfodiéster interno en una cadena de nucleótidos (en contraste con las exonucleasas, que escinden un enlace de fosfodiéster en el extremo de una cadena de nucleótidos). Las endonucleasas incluyen endonucleasas de restricción u otras endonucleasas específicas del sitio (que escinden ADN en sitios específicos de secuencias), ADNasa I, Bal 31 nucleasa, nucleasa S1, nucleasa de Mung, Ribonucleasa A, Ribonucleasa T1, ARNasa I, ARNasa PhyM, ARNasa U2, ARNasa CLB, nucleasa microcócica y endonucleasas apurínica/apirimidínica. Las exonucleasas incluyen exonucleasa III y exonucleasa VII. En ejemplos en particular, una nucleasa es específica para ácidos nucleicos monocatenarios, tal como nucleasa S1, nucleasa de Mung, Ribonucleasa A, o Ribonucleasa T1.

Ácido nucleico: Un polímero de desoxirribonucleótido o ribonucleótido en forma monocatenaria o bicatenaria, y salvo que esté limitado de otro modo, que comprende análogos de nucleótidos naturales que se hibridan con ácidos nucleicos de una forma similar a nucleótidos de ocurrencia natural. El término "nucleótido" incluye, pero no se limita a, un monómero que incluye una base (por ejemplo, pirimidina, purina o análogos sintéticos de las mismas) ligada a un azúcar (por ejemplo, ribosa, desoxirribosa o análogos sintéticos de las mismas), o una base unida a un aminoácido, como en un ácido nucleico peptídico (PNA). Un nucleótido es un monómero en un polinucleótido. Una secuencia de nucleótidos se refiere a la secuencia de bases en un polinucleótido.

Un **ácido nucleico diana** (por ejemplo, un ADN o ARN diana) es una molécula de ácido nucleico cuya detección, cantidad o secuencia pretende determinarse (por ejemplo, de una forma cuantitativa o cualitativa). En un ejemplo, la diana es una región definida o una parte particular de una molécula de ácido nucleico, por ejemplo un ADN o ARN de interés. En un ejemplo donde la secuencia de ácidos nucleicos diana es un ADN diana o un ARN diana, por ejemplo una diana puede definirse por su secuencia o función específica; por su nombre de gen o proteína; o por cualquier otro medio que lo identifique de forma única de entre otros ácidos nucleicos.

En algunos ejemplos, las alteraciones de una secuencia de ácidos nucleicos diana (por ejemplo, un ADN o ARN) están "asociadas con" una enfermedad o dolencia. Es decir, la detección de ácidos nucleicos diana puede usarse para inferir el estado de una muestra con respecto a la enfermedad o dolencia. Por ejemplo, la secuencia de ácidos nucleicos diana puede existir en dos (o más) formas diferenciables, de manera que una primera forma esté correlacionada con la ausencia de una enfermedad o dolencia y una segunda forma (o diferente) esté correlacionada con la presencia de la enfermedad o dolencia. Las dos formas diferentes pueden distinguirse cualitativamente, por ejemplo por polimorfismos o mutación de nucleótidos, y/o las dos formas diferentes pueden distinguirse cuantitativamente, por ejemplo por el número de copias de la secuencia de ácidos nucleicos diana que están presentes en una muestra.

Nucleótido: La unidad fundamental de moléculas de ácidos nucleicos. Un nucleótido incluye una base que contiene nitrógeno unida a un monosacárido de pentosa con uno, dos o tres grupos fosfato ligados por uniones estéricas a la fracción de sacáridos.

Los principales nucleótidos de ADN son desoxiadenosina 5'-trifosfato (dATP o A), desoxiguanosina 5'-trifosfato (dGTP o G), desoxicitidina 5'-trifosfato (dCTP o C) y desoxitimidina 5'-trifosfato (dTTP o T). Los principales nucleótidos de ARN son adenosina 5'-trifosfato (ATP o A), guanosina 5'-trifosfato (GTP o G), citidina 5'-trifosfato (CTP o C) y uridina 5'-trifosfato (UTP o U).

Los nucleótidos incluyen aquellos nucleótidos que contienen bases modificadas, fracciones de azúcares modificadas y estructuras de fosfato modificadas, por ejemplo tal como se describe en la patente de EE.UU. n° 5.866.336 para Nazarenko y col. (incorporada en la presente memoria descriptiva como referencia). Incluye nucleótidos que contienen otras modificaciones, por ejemplo presentes en ácidos nucleicos bloqueados (LNA). Así, las NPPF, los

cebadores, las CFS, los conectores bifuncionales y los anclajes divulgados en la presente memoria descriptiva pueden incluir bases naturales y no naturales.

Los ejemplos de fracciones de bases modificadas que pueden usarse para modificar nucleótidos en cualquier posición en su estructura incluyen, pero no se limitan a: 5-fluorouracilo, 5-bromouracilo, 5-clorouracilo, 5-yodouracilo, 5 hipoxantina, xantina, acetilcitosina, 5-(carboxihidroximetil)uracilo, 5-carboximetilaminometil-2-tiouridina, 5-carboximetilaminometiluracilo, dihidouracilo, beta-D-galactosilqueosina, inosina, N-6-sopenteniladenina, 1-metilguanina, 1-metilinosina, 2,2-dimetilguanina, 2-metiladenina, 2-metilguanina, 3-metilcitosina, 5-metilcitosina, N6-adenina, 7-metilguanina, 5-metilaminometiluracilo, metoxiaminometil-2-tiouracilo, beta-D-mannosilqueosina, 5'-metoxycarboximetiluracilo, 5-metoxiuracilo, 2-metiltio-N6-isopenteniladenina, ácido uracilo-5-oxiacético, seoudouracilo, 10 queosina, 2-tiocitosina, 5-metil-2-tiouracilo, 2-tiouracilo, 4-tiouracilo, 5-metiluracilo, éster metílico del ácido uracilo-5-oxiacético, ácido uracilo-S-oxiacético, 5-metil-2-tiouracilo, 3-(3-amino-3-N-2-carboxipropil)uracilo y 2,6-diaminopurina.

Los ejemplos de fracciones de azúcares modificados que pueden usarse para modificar nucleótidos en cualquier posición en su estructura incluyen, pero no se limitan a: arabinosa, 2-fluoroarabinosa, xilosa y hexosa, o un componente modificado de la estructura principal de fosfato, tal como fosforotioato, fosforoditioato, fosforamidotioato, fosforamidoato, fosfordiamidato, metilfosfonato, fosfotriéster alquílico o formacetal o un análogo de los mismos.

Cebador. Una molécula corta de ácido nucleico, tal como un oligonucleótido de ADN de 9 nucleótidos o más de longitud, que en algunos ejemplos se usa para iniciar la síntesis de una secuencia de ácidos nucleicos más larga. Los cebadores largos pueden tener aproximadamente 10, 12, 15, 20, 25, 30 ó 50 nucleótidos o más de longitud. Los cebadores pueden hibridarse con una cadena complementaria de ácidos nucleicos por hibridación de ácidos nucleicos para formar un híbrido entre el cebador y la cadena de complemento, y a continuación el cebador extendido a lo largo de la cadena de complemento por una enzima polimerasa. Los pares de cebadores pueden usarse para amplificación de una secuencia de ácidos nucleicos, por ejemplo por PCR u otros procedimientos de amplificación de ácidos nucleicos.

En un ejemplo, a cebador incluye una etiqueta, que puede referirse como una sonda.

Sonda: Una molécula de ácido nucleico capaz de hibridarse con una molécula de ácido nucleico diana (por ejemplo, un ADN o ARN diana) y, cuando se hibrida con la diana, es susceptible de ser detectada directa o indirectamente. Así, las sondas permiten la detección, y en algunos ejemplos la cuantificación, de una molécula de ácido nucleico diana, tal como un ADN o ARN. En algunos ejemplos, una sonda incluye una etiqueta detectable.

Sonda de protección contra la nucleasa (NPP): Una molécula de ácido nucleico que tiene una secuencia que es complementaria a un ADN o ARN diana y es capaz de hibridarse con el ADN o ARN diana. La NPP protege la molécula de ácido nucleico complementaria de ADN o ARN diana de su escisión por una nucleasa, tal como una nucleasa específica para ácidos nucleicos monocatenarios. Una **sonda de protección contra la nucleasa que comprende una secuencia flanqueante (NPPF)** es una NPP que incluye además una o más secuencias 35 flanqueantes en el extremo 5', el extremo 3' o ambos, donde la secuencia flanqueante incluye una secuencia de nucleótidos contiguos no presentes en una molécula de ácido nucleico presente en la muestra, y que puede proporcionar un punto de secuencia de amplificación universal que puede usarse como punto de fijación para un cebador de amplificación. En un ejemplo la secuencia flanqueante se usa para capturar la NPPF a un sustrato, donde una secuencia de captura de ácidos nucleicos en el sustrato y al menos una parte de la secuencia 40 flanqueante son complementarios entre sí, lo que permite así la captura de la NPPF en el sustrato.

Muestra: Un espécimen biológico que contiene ADN (por ejemplo, ADN genómico o cDNA), ARN (que incluye mRNA o miRNA), proteínas o combinaciones de los mismos, obtenido de un sujeto (por ejemplo, un ser humano u otro mamífero). Los ejemplos incluyen, pero no se limitan a células, lisados celulares, preparaciones cromosómicas, sangre periférica o fracciones de las mismas, orina, saliva, biopsia de tejidos (por ejemplo, una biopsia de un tumor o una biopsia de ganglios linfáticos), muestra quirúrgica, médula ósea, muestras de amniocentesis, aspirados con 45 aguja fina, células de tumores circulantes y material de autopsia. En un ejemplo, una muestra incluye ARN o ADN. En ejemplos particulares, las muestras se usan directamente (por ejemplo, frescas o congeladas), o pueden manipularse antes de su uso, por ejemplo, por fijación (por ejemplo, usando formalina) y/o incrustación en cera (por ejemplo, muestras de tejidos FFPE).

Identidad/ semejanza de secuencias: La identidad/ semejanza entre dos o más secuencias de ácidos nucleicos, o dos o más secuencias de aminoácidos, se expresa en términos de identidad o semejanza entre las secuencias. La identidad de secuencias puede medirse en términos de identidad porcentual; cuanto más elevado es el porcentaje, más idénticas son las secuencias. Los procedimientos de alineación de secuencias para comparación son bien conocidos en la técnica. Se describen varios programas y algoritmos de alineación en: Smith & Waterman, Adv. 50 Appl. Math. 2:482, 1981; Needleman & Wunsch, J. Mol. Biol. 48:443, 1970; Pearson & Lipman, Proc. Natl. Acad. Sci. USA 85:2444, 1988; Higgins & Sharp, Gene, 73:237-44, 1988; Higgins & Sharp, Compact. Appl. Biosci. 5:151-3, 1989; Corpet y col., Nucl. Acids Res. 16:10881-90, 1988; Huang y col. Comput. Appl. Biosci. 8, 155-65, 1992; y Pearson y col., Meth. Mol. Bio. 24:307-31, 1994. Altschul y col., J. Mol. Biol. 215:403-10, 1990, presenta una consideración detallada de procedimientos de alineación de secuencias y cálculos de homología.

La herramienta NCBI Basic Local Alignment Search Tool (BLAST) (Altschul y col., J. Mol. Biol. 215:403-10, 1990) está disponible en varias fuentes, que incluyen el National Center for Biological Information (NCBI, National Library of Medicine, Building 38A, Room 8N805, Bethesda, MD 20894) y en Internet, para su uso en relación con los programas de análisis de secuencias blastp, blastn, blastx, tblastn y tblastx. Blastn se usa para comparar secuencias de ácidos nucleicos, mientras que blastp se usa para comparar secuencias de aminoácidos. Puede encontrarse información adicional en la página web de NCBI. Una vez alineado, el número de correspondencias se determina contando el número de posiciones donde está presente un número idéntico de residuos de nucleótidos o aminoácidos en las dos secuencias. El porcentaje de identidad de secuencias se determina dividiendo el número de correspondencias por la longitud de la secuencia indicada en la secuencia identificada o mediante una longitud articulada (por ejemplo, 100 nucleótidos o residuos de aminoácidos consecutivos a partir de una secuencia expuesta en una secuencia identificada), seguido por la multiplicación del valor resultante por 100.

Una indicación de que dos moléculas de ácidos nucleicos están estrechamente relacionadas es que las dos moléculas se hibridan entre sí en condiciones de astringencia. Las condiciones de astringencia son dependientes de la secuencia y son diferentes según distintos parámetros ambientales. Las sondas de ácidos nucleicos divulgadas en la presente memoria descriptiva no se limitan a las secuencias exactas mostradas, ya que los expertos en la materia observarán que pueden realizarse cambios en una secuencia, sin que afecten sustancialmente a la capacidad de una sonda de actuar como se desea. Por ejemplo, en la presente memoria descriptiva se proporcionan secuencias que tienen al menos el 80%, al menos el 85%, al menos el 90%, al menos el 91%, al menos el 92%, al menos el 93%, al menos el 94%, al menos el 95%, al menos el 96%, al menos el 97%, al menos el 98% o al menos el 99%, por ejemplo el 100% de identidad de secuencias con las sondas divulgadas (por ejemplo, SEQ ID NO: 1-16). Un experto en la materia observará que estos intervalos de identidad de secuencias se proporcionan sólo como orientación; es posible que puedan usarse sondas situadas fuera de estos intervalos.

Secuenciación: Para determinar la estructura primaria (o secuencia primaria) de un biopolímero no ramificado. La secuenciación produce una representación lineal simbólica conocida como una secuencia que resume sucintamente buena parte de la estructura a nivel atómico de la molécula secuenciada, por ejemplo, un polinucleótido. Cuando la molécula es un polinucleótido, por ejemplo, ARN o ADN, la secuenciación puede usarse para obtener información acerca de la molécula en el nivel del nucleótido, que a continuación puede usarse para descifrar diversas informaciones secundarias sobre la molécula en sí y/o el polipéptido codificado. La secuenciación de ADN es el proceso de determinación del orden de nucleótidos de una molécula de ADN dada y la secuenciación de ARN es el proceso de determinación del orden de nucleótidos de una molécula de ARN dada. En algunos ejemplos, la secuenciación de una molécula de ácido nucleico se realiza indirectamente, por ejemplo determinando la secuencia de al menos una parte de una sonda de protección contra la nucleasa que comprende una secuencia flanqueante (NPPF), que se une a la molécula de ácido nucleico diana.

Simultáneo: Que tiene lugar al mismo tiempo o sustancialmente al mismo tiempo y/o que ocurre en la misma muestra o en la misma reacción (por ejemplo, contemporáneo). En algunos ejemplos, los acontecimientos tienen lugar en el plazo de 1 microsegundo a 120 segundos de otro (por ejemplo, en el plazo de 0,5 a 120 segundos, de 1 a 60 segundos, o de 1 a 30 segundos, o de 1 a 10 segundos).

Sujeto: Cualquier organismo vertebrado multicelular, por ejemplo un ser humano y un mamífero no humano (por ejemplo, sujetos veterinarios). En un ejemplo, se sabe o se sospecha que un sujeto tiene un tumor o una infección.

Superficie (o sustrato): Cualquier soporte o material sólido que es insoluble, o puede hacerse insoluble por una reacción posterior. En la técnica se conocen numerosos y variados soportes sólidos que incluyen, sin limitación, nitrocelulosa, las paredes de los pocillos de una bandeja de reacción, placas multipocillo, tubos de ensayo, perlas de poliestireno, perlas magnéticas, membranas y micropartículas (por ejemplo, partículas de látex). En este término se contempla cualquier material poroso adecuado con porosidad suficiente para permitir el acceso de reactivos de detector y una afinidad de superficie adecuada para inmovilizar reactivos de captura (por ejemplo, oligonucleótidos). Por ejemplo, la estructura porosa de la nitrocelulosa tiene unas cualidades excelentes de absorción y adsorción para una amplia variedad de reactivos, por ejemplo, reactivos de captura. El nailon posee características similares y también es adecuado. Las estructuras microporosas son útiles, ya que son materiales con estructura en gel en estado hidratado.

Los ejemplos adicionales de soportes sólidos útiles incluyen hidratos de carbono poliméricos naturales y sus derivados modificados, reticulados o sustituidos sintéticamente, tales como agar, agarosa, ácido alginico reticulado, gomas guar sustituidas y reticuladas, ésteres de celulosa, especialmente con ácido nítrico y ácidos carboxílicos, ésteres de celulosa mixtos, y éteres de celulosa; polímeros naturales que contienen nitrógeno, tal como proteínas y derivados, que incluyen gelatinas reticuladas o modificadas; polímeros de hidrocarburos naturales, tal como látex y caucho; polímeros sintéticos que pueden prepararse con estructuras adecuadamente porosas, tales como polímeros de vinilo, que incluyen polietileno, polipropileno, poliestireno, policloruro de vinilo, poliacetato de vinilo y sus derivados parcialmente hidrolizados, poliacrilamidas, polimetacrilatos, copolímeros y terc-polímeros de los policondensados anteriores, tal como poliésteres, poliamidas y otros polímeros, tales como poliuretanos o

poliepóxidos; materiales inorgánicos porosos tales como sulfatos o carbonatos de metales alcalinotérreos y magnesio, que incluyen sulfato de bario, sulfato de calcio, carbonato de calcio, silicatos de metales alcalinos y alcalinotérreos, aluminio y magnesio; y óxidos o hidratos de aluminio o silicio, tales como arcillas, alúmina, talco, caolín, zeolita, gel de sílice o vidrio (estos materiales pueden usarse como filtros con los materiales poliméricos anteriores); y mezclas o copolímeros de las clases anteriores, tales como copolímeros de injerto obtenidos iniciando la polimerización de polímeros sintéticos en un polímero natural preexistente.

Todas las publicaciones, solicitudes de patentes, patentes y otras referencias mencionadas en la presente memoria descriptiva se incorporan en su totalidad como referencia para todos los fines. Todas las secuencias asociadas con los números de acceso de GenBank mencionados en la presente memoria descriptiva se incorporan en su totalidad como referencia tal como se presentaron el 15 de diciembre de 2011, en la medida permisible por las reglas y/o leyes aplicables. En caso de conflicto prevalecerá la presente memoria descriptiva, que incluye explicaciones de los términos.

Aunque para la aplicación o prueba de la tecnología divulgada pueden usarse procedimientos y materiales similares o equivalentes a los descritos en la presente memoria descriptiva, a continuación se describen procedimientos y materiales adecuados. Los materiales, procedimientos, y ejemplos son sólo ilustrativos y no pretenden ser limitativos.

20 III. Procedimientos de detección de o secuenciación de moléculas de ácidos nucleicos

En la presente memoria descriptiva se divulgan procedimientos de detección y/o secuenciación de moléculas de ácidos nucleicos presentes en una muestra. En algunos ejemplos, se detectan al menos dos moléculas de ácidos nucleicos diferentes en la misma muestra o el mismo ensayo (por ejemplo, en el mismo pocillo de una placa o matriz de ensayo). En algunos ejemplos, se detecta la misma molécula o moléculas de ácido nucleico en al menos dos muestras o ensayos diferentes (por ejemplo, en muestras de diferentes pacientes).

Los procedimientos divulgados proporcionan mejoras a un ensayo cuantitativo de protección contra la nucleasa (qNPA), por ejemplo tal como se describe en las solicitudes de patentes internacionales WO-99/032.663; WO-00/037.683; WO-00/037.684; WO-00/079.008; WO-03/002.750; y WO-08/121.927; y en las patentes de EE.UU. nº 6.232.066, 6.238.869; 6.458.533; y 7.659.063. Véase también, Martel y col., Assay and Drug Development Technologies, 2002, 1 (1-1):61-71; Martel y col., Progress in Biomedical Optics and Imaging, 2002, 3:35-43; Martel y col., Gene Cloning and Expresión Technologies, Q. Lu y M. Weiner, Eds., Eaton Publishing, Natick (2002); Seligmann Pharmacogenomics, 2003, 3:36-43; Martel y col., "Array Formats" en "Microarray Technologies and Applications", U.R. Muller y D. Nicolau, Eds, Springer-Verlag, Heidelberg (2005); Sawada y col., Toxicology in Vitro, 20:1506-1513, 2006; Bakir, y col., Bioorg. & Med. Chem Lett, 17:3473-3479, 2007; Kris y col., Plant Physiol. 144:1256-1266, 2007; Roberts y col., Laboratory Investigation, 87:979-997, 2007; Rimsza y col., Blood, 2008 Oct 15, 112 (8):3425-3433; Pechhold y col., Nature Biotechnology, 27:1038-1042, 2009. Por ejemplo, los procedimientos qNPA divulgados han mejorado la sensibilidad en comparación con los procedimientos qNPA anteriores, por ejemplo un aumento en la señal detectable de al menos 10 veces, al menos 25 veces, al menos 100 veces, al menos 125 veces, al menos 150 veces, al menos 170 veces o al menos 200 veces. Es decir, se requiere al menos 10 veces, o incluso hasta 200 veces menos muestra, o por el contrario, con los procedimientos divulgados son detectables genes raros que están 10 veces por debajo de la sensibilidad, o incluso hasta 20 veces por debajo de la sensibilidad de los procedimientos disponibles en la actualidad. En consecuencia, pueden someterse a ensayo tipos de muestra como, por ejemplo, aspirados con aguja fina que proporcionan cantidades muy pequeñas de FFPE, o células de tumores circulantes, donde pueden recuperarse incluso 10, 50 ó 100 células de un paciente, y detectarse genes raros usando los procedimientos divulgados.

Además, los procedimientos divulgados proporcionan mejoras a un procedimiento de secuenciación cuantitativa de protección contra la nucleasa (qNPS), por ejemplo tal como se describe en la solicitud de patente de EE.UU. nº US-2011-0.104.693. qNPS es un procedimiento de secuenciación que usa qNPA para convertir moléculas de ácidos nucleicos diana presentes en una muestra, incluso en estado reticulado, en ácidos nucleicos diana monocatenarios estables (sondas de protección contra la nucleasa, NPP) que pueden recuperarse en solución sin captura o separación, mediante el uso de la etapa de protección contra la nucleasa y (en caso necesario) tratamiento con base para disociar las sondas de protección contra la nucleasa de la protección de moléculas diana, y en el caso de ARN, hidrolizar el ARN diana. Las cantidades de las NPP que quedan después de la hidrólisis de nucleasa se determinan entonces por secuenciación que puede incluir la secuenciación de las sondas en sí. Los procedimientos divulgados mejorados en la presente memoria descriptiva usan una variación de una NPP, una sonda de protección contra la nucleasa que comprende una secuencia flanqueante (NPPF). El uso de NPPF permite la multiplexación, así como la

- conservación de la estequiometría de la molécula de ácido nucleico diana detectada o secuenciada, ya que las secuencias flanqueantes en la sonda permiten sitios de cebadores universales para la amplificación. Dado que los sitios de unión de cebadores son universales, pueden usarse los mismos cebadores para amplificar cualquier NPPF para cualquier secuencia diana, permitiendo así la multiplexación y la conservación de estequiometría. En un ejemplo, la amplificación a partir de secuencias flanqueantes en los dos extremos de la NPPF proporciona una especificidad inesperada y mayor que los procedimientos anteriores de qNPA y qNPS. Las NPPF con secuencias flanqueantes en 3' y 5' intactas se amplificarán exponencialmente; las NPPF no escindidas con nucleasa no se amplificarán suficientemente para ser secuenciadas o detectadas. En cambio, las NPPF procesadas usando procedimientos qNPA anteriores pueden escindirse parcialmente en el extremo de la secuencia que interviene en la captura en la matriz o en el extremo de la secuencia que interviene en la detección en la matriz, o en los dos debido a la hibridación débil o incorrecta para ácidos nucleicos diana incorrectos, y aún así ser capturadas o detectadas, lo que conduce a una pérdida de especificidad para el ácido nucleico diana correcto. No sucede así con las sondas de NPPF divulgadas. Los procedimientos divulgados conservan la estequiometría original de moléculas de ácidos nucleicos de manera que las moléculas de ácidos nucleicos detectadas o secuenciadas conservan las mismas cantidades relativas de moléculas de ácidos nucleicos que en la muestra de prueba, por ejemplo, una variación de no más del 20%, no más del 15%, no más del 10%, no más del 9%, no más del 8%, no más del 7%, no más del 6%, no más del 5%, no más del 4%, no más del 3%, no más del 2%, no más del 1%, no más del 0,5% o no más del 0,1%, tal como el 0,001%-5%, el 0,01%-5%, el 0,1%-5% o el 0,1%-1%.
- 20 Los procedimientos divulgados permiten también experimentos de multiplexación, tales como múltiples reacciones dentro del mismo ensayo (por ejemplo, múltiples muestras de diferentes pacientes en el mismo pocillo de reacción), y múltiples reacciones analizadas en la misma ejecución/canal del secuenciador.

Específicamente, a diferencia de los procedimientos qNPA y qNPS anteriores, los procedimientos divulgados usan sondas de protección de ácidos nucleicos (NPP) modificadas, que incluyen secuencias flanqueantes en uno o en los dos extremos de las NPP. Estas NPP modificadas con secuencias flanqueantes en el extremo 5' y/o el extremo 3' se refieren en la presente memoria descriptiva como sondas de protección de ácidos nucleicos con secuencias flanqueantes (NPPF). La presencia de una o las dos secuencias flanqueantes, que sirven como puntos de cebadores universales para hibridación y/o amplificación (y pueden usarse para otros fines que incluyen la captura o etiquetado de NPPF), conserva la estequiometría original de los ácidos nucleicos en la muestra dado que las secuencias flanqueantes son parte de la NPPF. Además, así se elimina la necesidad de unión para añadir sitios de cebado, etiquetas, y similares a las NPPF, que pueden incorporar artefactos que degradan la estequiometría de ácidos nucleicos en la muestra, y proporcionan una fuente adicional de variabilidad. La eliminación de la necesidad de unión elimina los posibles artefactos que degradan la estequiometría y perjudican la reproducibilidad.

La FIG. 1 es un diagrama esquemático que muestra una NPPF de ejemplo. La sonda de protección contra la nucleasa que tiene al menos una secuencia flanqueante (NPPF) (100) incluye una región (102) que incluye una secuencia que se une específicamente a la secuencia de ácidos nucleicos diana (y puede unirse también específicamente a un conector bifuncional). La secuencia de ácidos nucleicos diana puede ser ADN (por ejemplo, ADN genómico o cDNA) o ARN (por ejemplo, mRNA, miRNA, tRNA, siRNA), o ambos. La NPPF incluye una o más secuencias flanqueantes (104) y (106). La FIG. 1 muestra una NPPF (100) con una secuencia flanqueante en 5' (104) y una secuencia flanqueante en 3' (106). Sin embargo, las NPPF pueden tener en algunos ejemplos sólo una secuencia flanqueante.

La FIG. 2 es un diagrama esquemático que muestra las etapas iniciales de un procedimiento de ejemplo de uso de las NPPF para detectar o secuenciar una molécula de ácido nucleico usando los procedimientos divulgados. Tal como se muestra en la etapa 1, una muestra (por ejemplo, una en la que se sabe o se sospecha que contiene un ácido nucleico diana, (200)) que ha sido tratada con un tampón de desorganización de muestra (por ejemplo, lisado o tratado por otros medios para hacer accesibles los ácidos nucleicos) se pone en contacto o se incuba con una pluralidad de sondas de protección contra la nucleasa que tienen una o más secuencias flanqueantes (NPPF) (202) que incluyen al menos una NPPF que se une específicamente a un primer ácido nucleico diana (por ejemplo, un ADN o ARN diana). La reacción puede incluir también otras NPPF que se unen específicamente a un segundo ácido nucleico diana, y así sucesivamente. Por ejemplo, el procedimiento puede usar una o más NPPF diferentes diseñadas para ser específicas para cada molécula de ácido nucleico diana individual. Así, la medida de 100 genes requiere el uso de al menos 100 NPPF diferentes, con al menos una NPPF específica por gen (por ejemplo, varias NPPF diferentes/gen). Así, por ejemplo, el procedimiento puede usar al menos 2 NPPF diferentes, al menos 3, al menos 4, al menos 5, al menos 10, al menos 25, al menos 50, al menos 75, al menos 100 o incluso al menos 200 NPPF diferentes (por ejemplo, de 2 a 500, de 2 a 100, de 5 a 10, de 2 a 10 o de 2 a 20 NPPF diferentes). Sin embargo, se observará que en algunos ejemplos, la pluralidad de NPPF puede incluir más de una (por ejemplo, 2, 3,

4, 5, 10, 20, 50 o más) NPPF específicas para una molécula de ácido nucleico diana individual. Las barras en trazo discontinuo en la FIG. 2 representan una NPPF específica para una primera diana y las barras grises en trazo continuo representan una NPPF específica par una segunda diana. En algunos ejemplos, las NPPF incluyen una etiqueta detectable, tal como biotina (B), pero un experto en la materia observará que una etiqueta puede añadirse en otras etapas, tal como durante la amplificación. Así, la biotina mostrada en la FIG. 2 es opcional, y pueden usarse otras etiquetas. La reacción incluye también moléculas de ácidos nucleicos que son complementarias a las secuencias flanqueantes (CFS), (204), que son específicas para las secuencias flanqueantes de la NPPF (202). La FIG. 2 muestra las barras verdes en trazo discontinuo (204) como las CFS específicas para una secuencia o secuencias flanqueantes de la NPP. Un experto en la materia observará que la secuencia de las CFS variará dependiendo de la secuencia flanqueante presente. Además, puede usarse más de una CFS para asegurar que se protege una región flanqueante (por ejemplo, pueden usarse al menos dos CFS que se unen a diferentes regiones de una secuencia flanqueante individual). La CFS puede incluir bases naturales o no naturales. Aunque la FIG. 2 muestra NPPF con secuencias flanqueantes en los dos extremos de la NPPF, un experto en la materia observará que puede usarse una sola secuencia flanqueante. La muestra, las NPPF y las CFS se incuban en condiciones suficientes para que las NPPF se unan específicamente a su molécula de ácido nucleico diana respectiva, y para que las CFS se unan a su secuencia complementaria en la secuencia flanqueante de NPPF. En algunos ejemplos, las CFS (204) se añaden en exceso con respecto a las NPPF (202), por ejemplo al menos 5 veces más CFS que NPPF (exceso molar), por ejemplo al menos 6 veces, al menos 7 veces, al menos 8 veces, al menos 9 veces, al menos 10 veces, al menos 20 veces, al menos 40 veces, al menos 50 veces o al menos 100 veces más CFS que NPPF. En algunos ejemplos, las NPPF (202) se añaden en exceso con respecto a las moléculas de ácidos nucleicos totales en la muestra, por ejemplo al menos 50 veces más NPPF que moléculas de ácidos nucleicos totales en la muestra (exceso molar), por ejemplo al menos 75 veces, al menos 100 veces, al menos 200 veces, al menos 500 veces o al menos 1.000 veces más NPPF que las moléculas de ácidos nucleicos totales en la muestra. Por comodidad experimental puede incluirse una concentración similar de cada NPPF para preparar un cóctel, de manera que para el ácido nucleico diana medido más abundante existirán al menos 50 veces más NPPF para ese ácido nucleico diana, por ejemplo un exceso de al menos 100 veces. El exceso real y la cantidad total de todas las NPPF usadas están limitados sólo por la capacidad de la nucleasa (por ejemplo, nucleasa S1) de destruir todas las NPPF que no están hibridadas con ácidos nucleicos diana. En algunos ejemplos la reacción se calienta, por ejemplo se incuba durante toda la noche a 50°C.

30 Tal como se muestra en la etapa 2 en la FIG. 2, después de permitir que se produzcan las reacciones de unión/hibridación, la muestra se pone en contacto con una nucleasa específica para moléculas de ácidos nucleicos monocatenarias (ss) en condiciones suficientes para eliminar (o digerir) moléculas de ácidos nucleicos ss, por ejemplo moléculas de ácidos nucleicos no unidas (por ejemplo, NPPF, CFS y moléculas de ácidos nucleicos diana no unidas, o porciones de dichas moléculas que siguen siendo monocatenarias). Tal como se muestra en la FIG. 2, la incubación de la muestra con una nucleasa específica para moléculas de ácidos nucleicos ss produce la degradación de cualquier molécula de ácidos nucleicos ss, lo que deja intactas las moléculas bicatenarias de ácidos nucleicos, que incluyen NPPF que se han unido a las mismas y CFS y molécula de ácido nucleico diana. Por ejemplo, la reacción puede incubarse a 50°C durante 1,5 horas con nucleasa S1 (aunque la hidrólisis puede producirse a otras temperaturas y efectuarse durante otros periodos de tiempo, y en parte el tiempo y la temperatura necesarios dependerán de la cantidad de nucleasa y de la cantidad de ácido nucleico requerida para la hidrólisis, así como la Tf de la región bicatenaria para su protección).

45 Después de esta reacción, las muestras pueden tratarse opcionalmente para eliminar o separar por otros medios material no hibridado y/o para inactivar o eliminar enzimas residuales (por ejemplo, por calor, extracción con fenol, precipitación, filtración en columna, etc.). Por ejemplo, tal como se muestra en la etapa 3 el pH de la reacción puede incrementarse para inactivar la nucleasa, y la reacción calentarse para destruir la nucleasa. Además, al calentar la reacción se disociará además el ácido nucleico diana (por ejemplo, ADN diana o ARN diana) y las CFS de las regiones complementarias en la NPPF. Esto deja tras de sí las NPPF intactas que se unieron anteriormente a las moléculas de ácidos nucleicos diana y las CFS, donde las NPPF intactas están en proporción directa con la cantidad de NPPF que se han hibridado con la diana. En algunos ejemplos, el ácido nucleico diana y las CFS hibridados pueden ser degradados, por ejemplo, por nucleasas o por tratamientos químicos. Alternativamente, la muestra puede tratarse de manera que abandone la parte hibridada (monocatenaria) de las moléculas de ácidos nucleicos diana, o el dúplex formado por las moléculas de ácidos nucleicos diana y las CPS hibridadas con la NPPF, para su análisis adicional (por ejemplo, la diana hibridada con la NPPF puede ser secuenciada). En un ejemplo, el pH aumentó a aproximadamente pH 8, y la reacción se incubó a 95°C durante 10 minutos y el ácido nucleico diana y las CFS se disociaron (y si el ácido nucleico diana es ARN, hidrólisis de dichos ácidos nucleicos diana).

Tal como se muestra en la etapa 4 en la FIG. 2, después de la etapa 2 o la etapa 3, las NPPF se amplifican, por

ejemplo usando PCR. La FIG. 2 muestra los cebadores de PCR o sondas (208) en forma de flechas. Los cebadores o sondas PCR pueden incluir una etiqueta, tal como biotina, lo que ocasiona la producción de amplicones que son etiquetados. Al menos una parte de los cebadores/sondas de PCR (208) es específica para las secuencias flanqueantes de las NPPF (202). A continuación pueden detectarse los amplicones (210) resultantes, por ejemplo
5 por unión con una matriz (véase FIG. 3) o secuenciación (véase FIG. 4). En algunos ejemplos, la concentración de los cebadores (208) es superior a la de las CPS (204), por ejemplo superior a al menos 10.000 veces, al menos 50.000 veces, al menos 100.000 veces, al menos 150.000 veces, al menos 200.000 veces o al menos 400.000 veces. En algunos ejemplos, la concentración de cebadores (208) en la reacción es de al menos 200 nM (por ejemplo, al menos 400 nM, al menos 500 nM, o al menos 1.000 nM), y la concentración de CPS (204) en la reacción
10 es menor que 1 pM, menor que 0,5 pM o menor que 0,1 pM.

Tal como se muestra en la etapa 5 en la FIG. 3, los amplicones (210), que son las NPPF amplificadas, pueden ponerse en contacto con una superficie (212) que incluye múltiples regiones espacialmente discretas. Se muestran dos versiones diferentes. En un ejemplo (arriba), la superficie incluye al menos un anclaje (214) en asociación con
15 un conector bifuncional (216). En algunos ejemplos los amplicones (210) se añaden a un tampón 2X antes de la puesta en contacto con la superficie (212). El conector bifuncional (216) incluye una primera parte que se une específicamente al anclaje y una segunda parte que se une específicamente a uno de la pluralidad de amplicones de NPPF (210). Los amplicones (210) se incuban con la superficie (212) en condiciones suficientes para cada uno de la pluralidad de amplicones de NPPF (210) para unirse específicamente a la segunda parte de un conector bifuncional
20 (216). Tal como se muestra en la FIG. 1, la región de la NPPF (102) que se une específicamente a un conector bifuncional es complementaria en secuencia al conector bifuncional (y es también complementaria a la secuencia de ácidos nucleicos diana). Los amplicones de NPPF (210) unidos a la segunda parte del conector bifuncional (216) se detectan usando la etiqueta detectable incluida en los amplicones de NPPF (210), con lo que se detecta el ácido nucleico diana en la muestra.

25 En otro ejemplo (abajo), la superficie incluye al menos una molécula de captura de ácido nucleico (220), que puede unirse directamente con la superficie a través de un enlace covalente. En algunos ejemplos los amplicones (210) se añaden a un tampón 2X antes de la puesta en contacto con la superficie (212). La molécula de captura de ácido nucleico (220) incluye una secuencia que es complementaria a al menos una parte de una de la pluralidad de
30 amplicones de NPPF (210), por ejemplo al menos una parte de una región de secuencia flanqueante de la NPPF (o una región añadida a la secuencia flanqueante durante la amplificación por ejemplo). Los amplicones (210) se incuban con la superficie (212) en condiciones suficientes para que cada uno de la pluralidad de amplicones de NPPF (210) se una específicamente a la molécula de captura de ácido nucleico (220).

35 Los amplicones de NPPF (210) unidos a la molécula de captura de ácido nucleico (220) son detectados usando la etiqueta detectable incluida en los amplicones de NPPF (210), detectando así el ácido nucleico diana en la muestra. Por ejemplo, los amplicones de NPPF pueden incubarse con la superficie durante toda la noche a 50°C para permitir la unión de los amplicones de NPPF a la molécula de captura de ácido nucleico (220). En un ejemplo, los amplicones de NPPF se etiquetan con biotina. Tal como se muestra en la etapa 6 de la FIG. 3, la biotina puede
40 detectarse usando avidina-HRP (218) (por ejemplo, incubando con la avidina-HRP durante 1 hora a 37°C). Tal como se muestra en la etapa 7 de la FIG. 3, el exceso de avidina-HRP (218) no unida se elimina, se añade un sustrato apropiado y se obtiene una imagen de la superficie para detectar las NPPF unidas. Aunque la biotina se muestra a modo de ejemplo, un experto en la materia observará que pueden usarse otros procedimientos de detección, por ejemplo detectando un fluoróforo o anticuerpo en los amplicones de NPPF.

45 En algunos ejemplos, si los amplicones de NPPF (210) no se etiquetan (por ejemplo, no se añade ninguna etiqueta durante amplificación en la etapa 4 de la FIG. 2), los amplicones de NPPF (210) pueden incluir una región (por ejemplo, la secuencia flanqueante o parte de la misma) que es complementaria a la secuencia de una sonda etiquetada (donde esta región no es complementaria al conector bifuncional (216)). Esta sonda complementaria
50 puede hibridarse a continuación con los amplicones de NPPF (210) antes de unirlos a un sustrato tal como se muestra en la etapa 5 de la FIG. 3.

En algunos ejemplos, los amplicones de NPPF se ponen en contacto con una pluralidad de superficies (por ejemplo, una población de perlas u otras partículas). En un ejemplo, cada superficie (por ejemplo, cada perla o subpoblación
55 de perlas en una población de perlas mixta) incluye al menos un anclaje en asociación con un conector bifuncional que incluye una primera parte que se une específicamente al anclaje y una segunda parte que se une específicamente a uno de la pluralidad de amplicones de NPPF, en condiciones suficientes para que cada uno de la pluralidad de amplicones de NPPF se una específicamente a la segunda parte de un conector bifuncional. Los amplicones de NPPF unidos a la segunda parte del conector bifuncional pueden detectarse usando la etiqueta

detectable que se asocia con los amplicones de NPPF, detectando así la molécula de ácido nucleico diana en la muestra. En otro ejemplo, cada superficie (por ejemplo, cada perla o subpoblación de perlas dentro de una población de perlas mixta) incluye al menos una molécula de captura de ácido nucleico que tiene una secuencia complementaria a al menos una parte de los amplicones de NPPF (por ejemplo, una secuencia flanqueante o parte de la misma), en condiciones suficientes para que cada uno de la pluralidad de amplicones de NPPF se una específicamente a la molécula de captura de ácido nucleico. Los amplicones de NPPF unidos a la molécula de captura de ácido nucleico pueden detectarse usando la etiqueta detectable que se asocia con los amplicones de NPPF, detectando así la molécula de ácido nucleico diana en la muestra.

- 10 Tal como se muestra en la etapa 5 en la FIG. 4, los amplicones (210), que son las NPPF amplificadas, pueden ser secuenciados. Por ejemplo, una o las dos secuencias flanqueantes de las NPPF amplificadas pueden incluir (o tener añadido) un adaptador de secuencias, o un cebador que es complementario a y se hibrida con la secuencia flanqueante, puede incluir un adaptador de secuencias, que es complementario para capturar secuencias en el chip de secuenciación y permitir secuenciación de la NPPF usando una plataforma de secuenciación en particular. En algunos ejemplos, se secuencia una pluralidad de NPPF en paralelo, por ejemplo de forma simultánea o contemporánea. Este procedimiento puede usarse así para secuenciar una pluralidad de secuencias de NPPF.

Las FIG. 5A y 5B son diagramas esquemáticos que proporcionan un resumen adicional del procedimiento, con más detalles de las moléculas de ácidos nucleicos. Tal como se muestra en el panel izquierdo de la FIG. 5A, los ácidos nucleicos diana (400) en una muestra (por ejemplo, una muestra que ha sido tratada con un tampón de desorganización de muestra) se pone en contacto o se incuba con una pluralidad de sondas de protección contra la nucleasa que tiene una o más secuencias flanqueantes (NPPF) (402) (donde cada NPPF es específica para un ácido nucleico diana (400) en particular), y con moléculas de ácidos nucleicos que son complementarias a las secuencias flanqueantes (CFS) (406), que son específicas para las secuencias flanqueantes (404) en los extremos de las NPPF. Se muestran los tres ácidos nucleicos diana (400) diferentes: una copia de la diana 1 (verde), dos copias de la diana 2 (rojo) y tres copias de la diana 3 (azul). Este ejemplo muestra que se añaden cantidades iguales de cada NPPF (402). Aunque la FIG. 5A muestra NPPF con secuencias flanqueantes en los dos extremos de la NPP; un experto en la materia observará que puede usarse una única secuencia flanqueante. El panel central de la FIG. 5A muestra los productos de reacción después de permitir que se produzcan las reacciones de unión/hibridación entre los ácidos nucleicos diana (400), las NPPF (402) y las CFS (406). Los ácidos nucleicos diana (400) se hibridan con una región central de las NPPF, y las CFS (406) se hibridan con las secuencias flanqueantes 3' y 5' (404). El panel derecho de la FIG. 5A muestra los productos de reacción después de que la muestra se pone en contacto con una nucleasa específica para moléculas de ácidos nucleicos monocatenarias (ss) en condiciones suficientes para eliminar (o digerir) moléculas de ácidos nucleicos ss. Tal como se muestra, las regiones de los ácidos nucleicos diana que no se hibridaron con una NPPF (408) son digeridas, dado que son regiones de NPPF ss que no se unieron con un ácido nucleico diana o una CFS (por ejemplo, (410)). De esta forma quedan intactas las moléculas bicatenarias de ácidos nucleicos, que incluyen NPPF que se han unido a las mismas y CFS y moléculas de ácido nucleico diana (por ejemplo, (412)) así como regiones de la NPPF que se hibridan sólo con la diana (pero no con CFS), o que se hibridan sólo con CFS (pero no con la diana) (por ejemplo, (414)).

El panel izquierdo de la FIG. 5B muestra los productos de reacción después de separar las moléculas bicatenarias de ácidos nucleicos (por ejemplo, usando calor y aumentando el pH). Las NPPF resultantes que sobreviven, que están en proporción directa con las moléculas de ácidos nucleicos diana que las protegieron durante la etapa de la nucleasa, pueden amplificarse a continuación. El panel central de la FIG. 5B muestra los productos de reacción después de que sean amplificados. El panel derecho de la FIG. 5B muestra que después de la amplificación, los amplicones de NPPF resultantes pueden ser detectados o secuenciados (por ejemplo, véase FIG. 2-4).

En algunas realizaciones, los procedimientos pueden incluir la puesta en contacto de una muestra de un sujeto (por ejemplo, una muestra que incluye ácidos nucleicos, tales como ADN o ARN) con una pluralidad de NPPF que incluye al menos una NPPF que se une específicamente a una primera diana (por ejemplo, un primer ARN) y opcionalmente al menos una NPPF que se une específicamente a una segunda diana (por ejemplo, un segundo ARN). En algunos ejemplos, la pluralidad de NPPF incluye más de una (por ejemplo, 2, 3, 4, 5, o más) NPPF específica para una única molécula de ácido nucleico diana. Por ejemplo, la pluralidad de NPPF puede incluir al menos una NPPF (por ejemplo, al menos 1, 2, 3, 4, 5, 10, 15, 20, 25, 50, 75, 100, 200, 300, 500, 1.000, 2.000, 3.000, o más), donde cada NPPF se une específicamente a una única molécula de ácido nucleico diana. En otro ejemplo adicional, la pluralidad de NPPF incluye al menos dos poblaciones de NPPF diferentes (por ejemplo, 2, 3, 4, 5, 10, 20 ó 50 secuencias de NPPF diferentes), donde cada población (o secuencia) de NPPF se une específicamente a una molécula de ácido nucleico diana diferente.

En algunos ejemplos, varias NPPF se hibridan con diferentes porciones del mismo ácido nucleico diana, y el número de NPPF que se hibrida con diferentes porciones de cada ácido nucleico diana puede ser el mismo o diferente. Por ejemplo, un ácido nucleico diana con baja expresión puede tener más NPPF que se hibridan con respecto a un ácido nucleico diana expresado en un nivel superior, tal como cuatro NPPF que se hibridan con un ácido nucleico diana de baja expresión y una única NPPF que se hibrida con un ácido nucleico diana de alta expresión. En algunos ejemplos, algunas de las NPPF específicas para algunos ácidos nucleicos diana pueden no tener secuencias flanqueantes (por ejemplo, NPP), y así pueden no ser amplificadas o etiquetadas, o tener los adaptadores apropiados unidos, y así esta parte de NPPF no será detectada o secuenciada. Al usar dicha mezcla, que puede tener aproximadamente de 1 a 5, o aproximadamente de 1 a 10, o aproximadamente de 1 a 100, o aproximadamente de 1 a 1.000 NPPF con secuencia flanqueante con NPP sin secuencia flanqueante, la señal medida, o el número de NPPF secuenciadas, puede ser "atenuado", de manera que si existen 10.000 copias de ácido nucleico diana, y se usa una relación de 1 a 5, después de la amplificación sólo 1/5 parte del número de NPPF será secuenciada ya que si se secuencian todas las NPPF que contienen secuencias flanqueantes.

En algunos ejemplos, la pluralidad de NPPF incluye al menos 2, al menos 5, al menos 10, al menos 20, al menos 100 o al menos 1.000 (por ejemplo, de 2 a 5.000, de 2 a 3.000, de 10 a 1.000, de 50 a 500, de 25 a 300, de 50 a 300, de 10 a 100 o de 50 a 100) secuencias de NPPF únicas. La pluralidad de NPPF puede incluir cualquier combinación de NPPF específicas para una o más moléculas de ácidos nucleicos diana. La pluralidad de NPPF, junto con las CFS, se incuban con la muestra en condiciones suficientes para que las NPPF se hibriden específicamente con sus ácidos nucleicos diana respectivos y sus CFS respectivas. En algunos ejemplos, se añaden más CFS que NPPF, por ejemplo al menos 2 veces, al menos 3 veces, al menos 4 veces, al menos 5 veces o al menos 10 veces más exceso molar de CFS con respecto a NPPF. En algunos ejemplos, se añaden más NPPF que las moléculas de ácidos nucleicos en la muestra, por ejemplo al menos 10 veces, al menos 50 veces, al menos 75 veces, al menos 100 veces, al menos 250 veces, al menos 1.000 veces, al menos 10.000 veces o al menos 100.000 veces un exceso molar o más de NPPF con respecto a las moléculas de ácidos nucleicos en la muestra. Se observará que si la NPPF para un ácido nucleico diana muy abundante tiene un exceso de 1.000 veces, y la misma concentración de cada NPPF diferente es la misma, entonces el exceso de NPPF para un gen poco abundante puede ser muchas veces mayor, por ejemplo 1.000 veces mayor para un gen que tiene una abundancia 1.000 veces menor que el ácido nucleico diana de alta abundancia.

La muestra hibridada puede ponerse en contacto a continuación con una nucleasa específica para ácidos nucleicos monocatenarios (por ejemplo, nucleasa S1). Las NPPF resultantes que sobreviven, que están en proporción directa con las moléculas de ácidos nucleicos diana que las protegieron durante la etapa de la nucleasa, pueden amplificarse a continuación. Por ejemplo, pueden usarse los cebadores de amplificación que incluyen una secuencia complementaria a la secuencia flanqueante de la NPPF. Los amplicones de NPPF resultantes pueden detectarse a continuación mediante procedimientos conocidos en la técnica, por ejemplo por unión de los mismos a una matriz u otro sustrato, o bien secuenciarse. La o las moléculas molécula de ácidos nucleicos diana se identifican como presentes en la muestra cuando su NPPF respectiva es detectada o secuenciada.

40 A. Condiciones de hibridación de ejemplo

En la presente memoria descriptiva se divulgan condiciones suficientes para que una pluralidad de NPPF se hibride específicamente con molécula o moléculas de ácidos nucleicos diana, tales como ADN y ARN presentes en una muestra de un sujeto, así como se hibriden específicamente con la CFS complementaria a la o las secuencias flanqueantes. Por ejemplo, las características (por ejemplo, longitud, composición de bases y grado de complementariedad) que permitirá que un ácido nucleico (por ejemplo, una NPPF) se hibride con otro ácido nucleico (por ejemplo, un ADN diana o ARN diana o CFS) en condiciones de astringencia seleccionada, a la vez que se reduce al mínimo la hibridación no específica con otras sustancias o moléculas, puede determinarse basándose en la presente descripción. Las características de las NPPF se exponen en más detalle en la Sección IV, más adelante. Normalmente, una región de una NPPF tendrá una secuencia de ácidos nucleicos (por ejemplo, FIG. 1, (102)) que es de complementariedad suficiente con su molécula de ácido nucleico diana correspondiente para permitir que se hibride en condiciones seleccionadas de astringencia de hibridación, así como una región (por ejemplo, FIG. 1, (104, 106)) que es de complementariedad suficiente con su CFS correspondiente para permitir que se hibride en condiciones de astringencia de hibridación seleccionadas. Las condiciones de hibridación de ejemplo incluyen hibridación a aproximadamente 37°C o más (por ejemplo, aproximadamente 37°C, 42°C, 50°C, 55°C, 60°C, 65°C, 70°C, 75°C, o más). Entre los parámetros de la reacción de hibridación que pueden modificarse están la concentración salina, el tampón, el pH, la temperatura, el tiempo de incubación, la cantidad y el tipo de desnaturante como, por ejemplo, formamida. Por ejemplo, el ácido nucleico (por ejemplo, una pluralidad de NPPF) puede añadirse a una muestra a una concentración comprendida entre aproximadamente 10 pM y

aproximadamente 10 nM (por ejemplo, de aproximadamente 30 pM a 5 nM, de aproximadamente 100 pM a aproximadamente 1 nM), en un tampón (por ejemplo, uno que contiene NaCl, KCl, H₂PO₄, EDTA, Triton X-100 al 0,05%, o combinaciones de los mismos), por ejemplo un tampón de lisis.

- 5 En un ejemplo, cada NPPF se añade a la muestra a una concentración final de al menos 10 pM, por ejemplo al menos 20 pM, al menos 30 pM, al menos 50 pM, al menos 100 pM, al menos 150 pM, al menos 200 pM, al menos 500 pM, al menos 1 nM, o al menos 10 nM. En un ejemplo, cada NPPF se añade a la muestra a una concentración final de aproximadamente 30 pM. En otro ejemplo, cada NPPF se añade a la muestra a una concentración final de aproximadamente 167 pM. En un ejemplo adicional, cada NPPF se añade a la muestra a una concentración final de
- 10 aproximadamente 1 nM. En un ejemplo, cada CFS se añade a la muestra a una concentración final de aproximadamente al menos 6 veces la cantidad de sonda, por ejemplo al menos 10 veces o al menos 20 veces la cantidad de sonda (por ejemplo, de 6 a 20 veces la cantidad de sonda). En un ejemplo, cada CFS se añade a al menos 1 nM, al menos 5 nM, al menos 10 nM, al menos 50 nM, al menos 100 nM, o al menos 200 nm, por ejemplo de 1 a 100, de 5 a 100 o de 5 a 50 nM. Por ejemplo si existen seis sondas, cada una a 166 pM, cada CFS puede
- 15 añadirse a entre 5 y 50 nM.

- Los ácidos nucleicos en la muestra son desnaturalizados, para hacerlos monocatenarios y disponibles para hibridación (por ejemplo, a entre aproximadamente 95°C y aproximadamente 105°C durante aproximadamente 5-15 minutos). Usando diferentes soluciones de desnaturalización, esta temperatura de desnaturalización puede
- 20 modificarse, en la medida en que la combinación de temperatura y composición del tampón lleva a la formación de ADN o ARN diana monocatenario o ambos. A continuación los ácidos nucleicos en la muestra y las CFS se hibridan con la pluralidad de NPPF durante entre aproximadamente 10 minutos y aproximadamente 72 horas (por ejemplo, al menos de aproximadamente 1 hora a 48 horas, de aproximadamente 6 horas a 24 horas, de aproximadamente 12 horas a 18 horas, o durante toda la noche) a una temperatura comprendida entre aproximadamente 4°C y
- 25 aproximadamente 70°C (por ejemplo, aproximadamente 37°C y aproximadamente 65°C, aproximadamente 42°C y aproximadamente 60°C, o aproximadamente 50°C y aproximadamente 60°C). Naturalmente las condiciones de hibridación variarán dependiendo de las NPPF y las CFS usadas en particular, pero se ajustan de manera que aseguren la hibridación de las NPPF con las moléculas diana y las CFS. En algunos ejemplos, la pluralidad de NPPF y CFS se incuba con la muestra a una temperatura de al menos aproximadamente 37°C, al menos
- 30 aproximadamente 40°C, al menos aproximadamente 45°C, al menos aproximadamente 50°C, al menos aproximadamente 55°C, al menos aproximadamente 60°C, al menos aproximadamente 65°C, o al menos aproximadamente 70°C. En un ejemplo, la pluralidad de NPPF y CFS se incuba con la muestra a aproximadamente 37°C, a aproximadamente 42°C o a aproximadamente 50°C.

- 35 En algunas realizaciones, los procedimientos no incluyen purificación de ácidos nucleicos (por ejemplo, la purificación de ácidos nucleicos no se realiza antes de la puesta en contacto de la muestra con las NPPF y/o la purificación de ácidos nucleicos no se realiza después de la puesta en contacto de la muestra con las NPPF). En algunos ejemplos, no se requiere procesamiento previo de la muestra excepto para lisis celular. En algunos ejemplos, la lisis celular y la puesta en contacto de la muestra con la pluralidad de NPPF y CFS tienen lugar en
- 40 secuencia. En otros ejemplos, la lisis celular y la puesta en contacto de la muestra con la pluralidad de NPPF y CFS tienen lugar de forma concurrente, en algunos ejemplos no limitativos sin ninguna etapa intermedia.

- Cuando las NPPF se someten posteriormente a PCR (por ejemplo, amplificación universal o amplificación específica de NPPF, tal como para PCR en tiempo real), los tampones y reactivos usados para la lisis, la hibridación de NPPF
- 45 en sus ácidos nucleicos diana, la digestión de nucleasa y la hidrólisis de bases pueden ser compatibles con la polimerasa usada para la amplificación.

B. Tratamiento con nucleasa

- 50 Después de la hibridación de las NPPF con los ácidos nucleicos diana en la muestra y en las CFS, se somete la muestra a un procedimiento de protección contra la nucleasa. Las NPPF que se han hibridado con una molécula de ácido nucleico diana y (cuando se usa) CFS (una o dos CFS, dependiendo de si existen las secuencias flanqueantes 5' y 3' en la NPPF o sólo una, o ninguna CFS cuando no se requieran secuencias flanqueantes para la amplificación o la medida) no son hidrolizadas por la nucleasa y pueden amplificarse posteriormente, y después detectarse o
- 55 secuenciarse (o ambos).

El tratamiento con una o más nucleasas destruirá todas las moléculas de ácidos nucleicos ss (lo que incluye ARN y ADN en la muestra que no se han hibridado con (y así no están protegidas por) NPPF, las NPPF que no se hibridan en ácido nucleico diana, y (cuando se usan) las CFS no hibridadas con una NPPF), pero no destruirá las moléculas

de ácidos nucleicos ds tales como NPPF que se han hibridado con CFS y una molécula de ácido nucleico diana presente en la muestra. Por ejemplo, si la muestra incluye un extracto o lisado celular, los ácidos nucleicos no deseados, tales como ADN genómico, tRNA, rRNA, mRNA, miRNA no diana, y porciones de la molécula o moléculas de ácidos nucleicos diana que no se hibridan con una secuencia complementaria de NPPF (por ejemplo, 5 protuberantes), que en el caso de ácidos nucleicos de mRNA o ADN diana constituirán la mayor parte de la secuencia de ácidos nucleicos diana, puede destruirse sustancialmente en esta etapa. Esto deja una cantidad estequiométrica de dúplex de ácido nucleico diana/CFS/NPPF. Si la molécula diana se reticula en el tejido que tiene lugar a partir de la fijación, las NPPF se hibridan con la molécula diana reticulada sin necesidad de revertir la reticulación, o de liberar por otros medios el ácido nucleico diana del tejido con el que está reticulado.

10

Pueden seleccionarse condiciones de manera que diferencias en nucleótidos individuales conduzcan a la no escisión de una base no apareada, o puede usarse una nucleasa que sólo escinda bases no apareadas en los extremos de la sonda de protección contra la nucleasa hibridada, por ejemplo, una exonucleasa. Pueden seleccionarse también condiciones que hidrolizarán la secuencia de NPPF en el punto de una única base no 15 apareada, e hidrolizar análogamente el ácido nucleico diana en esa posición.

Los ejemplos de nucleasas incluyen endonucleasas, exonucleasas y combinaciones de las mismas. Puede usarse cualquiera de una diversidad de nucleasas, entre ellas, ADNasa, ARNasa pancreática, nucleasa de Mung, nucleasa S1, ARNasa A, ribonucleasa T1, exonucleasa III, exonucleasa VII, ARNasa CLB, ARNasa PhyM, ARNasa U2, o 20 similares, dependiendo de la naturaleza de los complejos hibridados y del resto de ácidos nucleicos y secuencias no de ácidos nucleicos diana presentes en la muestra. Un experto en la materia puede seleccionar una nucleasa apropiada. En un ejemplo en particular, la nucleasa es específica para ácidos nucleicos monocatenarios (ss), por ejemplo, nucleasa S1. Una ventaja de usar una nucleasa específica para ácidos nucleicos ss, además de hidrolizar el exceso de NPPF y conferir la estequiometría de ácido nucleico diana a las NPPF, es eliminar dichas moléculas 25 monocatenarias ("pegajosas") de las posteriores etapas de reacción donde pueden llevar a un sustrato o una reactividad cruzada no deseables. Sin embargo, un experto en la materia observará que si el ácido nucleico diana debe secuenciarse, esto no es necesario, ya que sólo las NPPF con los adaptadores de secuenciación apropiados se hibridarán con los chips de secuenciación, punto en el cual las moléculas ss de la muestra pueden ser lavadas y eliminadas. La nucleasa S1 está disponible comercialmente, por ejemplo, en Promega, Madison, WI (nº cat. M5761); 30 Life Technologies/Invitrogen, Carlsbad, CA (nº cat. 18001-016); Fermentas, Glen Burnie, MD (nº cat. EN0321), y otros. Las condiciones de reacción para estas enzimas son bien conocidas en la técnica y pueden optimizarse empíricamente.

En algunos ejemplos, la nucleasa S1 diluida en un tampón (por ejemplo, que contiene acetato de sodio, NaCl, KCl, 35 ZnSO₄, KATHON, o combinaciones de los mismos) se añade a la mezcla de sonda/muestra hibridada y se incuba a entre aproximadamente 37°C y aproximadamente 60°C (por ejemplo, aproximadamente 50°C) durante 10-120 minutos (por ejemplo, 10-30 minutos, 30-60 minutos, 60-90 minutos o 120 minutos) para digerir ácido nucleico no hibridado de la muestra y NPPF no hibridadas.

40 Las muestras pueden tratarse opcionalmente para eliminar por otros medios el material no hibridado y/o para inactivar o eliminar las enzimas residuales (por ejemplo, por calentamiento, extracción con fenol, precipitación, filtración por columna, adición de proteinasa k, adición de un inhibidor de nucleasa, cationes divalentes de quelación requeridos por la nucleasa para actividad, o combinaciones de los mismos). En algunos ejemplos, las muestras son tratadas opcionalmente para disociar el ácido nucleico diana y la o las CFS de su NPPF complementaria (por 45 ejemplo, usando hidrólisis de bases y calor). En algunos ejemplos, después de la hibridación y el tratamiento con nucleasas, la molécula de ARN diana hibridada con la NPPF puede ser degradada, por ejemplo, disociando el dúplex con NPPF en la base y después destruyendo el ARN mediante nucleasas o por tratamientos químicos/físicos, tales como hidrólisis de bases a temperatura elevada, dejando la NPPF en proporción directa con la cantidad que se haya hibridado con el ácido nucleico diana. Alternativamente, la muestra puede ser tratada de manera que deje la 50 parte hibridada (monocatenaria) del ácido nucleico diana, o el dúplex formado por el ácido nucleico diana hibridado y la sonda, para su análisis posterior.

En algunos ejemplos después de la incubación con una nucleasa, se añade una base (por ejemplo, NaOH o KOH) para aumentar el pH a aproximadamente de 9 a 12 y se calienta la muestra (por ejemplo, a 95°C durante 10 55 minutos). Así se disocian los dímeros molécula/CFS/NPPF diana, dejando la NPPF en un estado monocatenario, y en el caso de ARN, hidroliza las moléculas de ARN diana. Esta etapa puede también neutralizar o desactivar la nucleasa, por ejemplo, elevando el pH por encima de aproximadamente 6.

En algunos ejemplos la muestra se trata para ajustar el pH a entre aproximadamente 7 y aproximadamente 8, por

ejemplo por adición de ácido (por ejemplo, HCl). En algunos ejemplos el pH se eleva a entre aproximadamente 7 y aproximadamente 8 en tampón Tris. La elevación del pH puede evitar la despurinación del ADN y evita también el funcionamiento pleno de muchas nucleasas específicas ss (por ejemplo, S1).

- 5 En algunos ejemplos, la muestra se purifica o se separa para eliminar ácido nucleico u otras moléculas no deseados, antes de la amplificación, por ejemplo por purificación en gel u otro procedimiento de separación.

C. Amplificación

- 10 Las moléculas de NPPF resultantes (o las moléculas de ácidos nucleicos diana resultantes que se han separado de la NPPF), que están en proporción directa con la cantidad de moléculas de ácidos nucleicos diana presentes en la muestra sometida a ensayo, pueden amplificarse, usando por ejemplo procedimientos rutinarios tales como PCR u otras formas de amplificación enzimática o procedimientos de amplificación basados en el ligamiento.
- 15 Los ejemplos de procedimientos de amplificación *in vitro* que pueden usarse incluyen, pero no se limitan a, PCR cuantitativa en tiempo real, amplificación por desplazamiento de cadena (véase documento USPN 5.744.311); amplificación isotérmica sin transcripción (véase USPN 6.033.881); amplificación por reacción en cadena de reparación (véase el documento WO-90/01.069); amplificación por reacción en cadena de ligasa (véase documento EP-A-320.308); amplificación por reacción en cadena de ligasa para llenado de huecos (véase documento USPN-5.427.930); detección de ligasa y PCR acopladas (véase documento USPN 6.027.889); y amplificación sin transcripción de ARN NASBA™ (véase documento USPN 6.025.134). En un ejemplo, se usa un procedimiento de amplificación basado en ligamiento, donde los cebadores son específicos de la NPPF y se acoplan juntos de manera que pueden ligarse conjuntamente, fusionarse, y después los cebadores nuevos ligarse conjuntamente durante una serie de ciclos. El ligamiento puede ser enzimático o no enzimático. Si las secuencias de NPPF flanqueantes se usan para hibridación de los cebadores, la amplificación puede ser universal.

La PCR cuantitativa en tiempo real es otra forma de amplificación *in vitro* de moléculas de ácidos nucleicos, habilitada por Applied Biosystems (TaqMan PCR). El ensayo de 5' nucleasa proporciona un procedimiento en tiempo real para detectar sólo productos de amplificación específicos. Durante la amplificación, la hibridación de la sonda con su secuencia diana genera un sustrato que es escindido por la actividad de nucleasa en 5' de la ADN polimerasa Taq cuando la enzima se extiende desde el cebador en la dirección 5' en la región de la sonda. Esta dependencia de la polimerización garantiza que tenga lugar la escisión de la sonda sólo si la secuencia diana se está amplificando. El uso de sondas fluorógenas hace posible eliminar el procesamiento posterior a la PCR para el análisis de degradación de la sonda. La sonda es un oligonucleótido con un colorante fluorescente indicador y un colorante de inactivación unido. Mientras la sonda está intacta, la proximidad del agente de inactivación reduce enormemente la fluorescencia emitida por el colorante indicador mediante transferencia de energía de resonancia de Förster (FRET) a través del espacio. Para PCR en tiempo real, la muestra de NPPF puede dividirse en pocillos separados o lugares de reacción, y se añade un conjunto específico de NPPF diferente de cebadores a cada pocillo o lugar de reacción. Usando sondas (cada una con una etiqueta diferente que permite la multiplexación de PCR en tiempo real con el fin de medir múltiples NPPF diferentes en un único pocillo, o lugar de reacción).

Durante la amplificación de la NPPF, puede incorporarse una etiqueta experimental, y/o adaptador de secuenciación, por ejemplo, como parte del cebador y de las construcciones de extensión, por ejemplo en el extremo 3' o 5' o en los dos extremos. Por ejemplo, un cebador de amplificación, que incluye una primera parte que es complementaria a parte o la totalidad de la secuencia flanqueante de NPPF puede incluir una segunda parte que es complementaria a una etiqueta experimental y/o adaptador de secuenciación deseados. Un experto en la materia observará que pueden añadirse diferentes combinaciones de etiquetas experimentales y/o adaptadores de secuenciación a cualquier extremo de la NPPF. En un ejemplo, la NPPF es amplificada usando un primer cebador de amplificación que incluye una primera parte complementaria a parte o la totalidad de la secuencia flanqueante en 3' de NPPF y una segunda parte complementaria a (o que comprende) un adaptador de secuenciación deseado, y el segundo cebador de amplificación incluye una primera parte complementaria a parte o la totalidad de la secuencia flanqueante en 5' de NPPF y una segunda parte complementaria a (o que comprende) una etiqueta experimental deseada. En otro ejemplo, la NPPF se amplifica usando un primer cebador de amplificación que incluye una parte o la totalidad de una primera parte complementaria a la secuencia flanqueante en 3' de NPPF y una segunda parte complementaria a (o que comprende) un adaptador de secuenciación deseado y una etiqueta experimental deseada, y el segundo cebador de amplificación incluye una primera parte complementaria a parte o la totalidad de la secuencia flanqueante en 5' de NPPF y una segunda parte complementaria a (o que comprende) una etiqueta experimental deseada.

Se observará que los cebadores específicos de NPPF pueden usarse para añadir adaptadores de secuenciación, etiquetas experimentales (que incluyen etiquetas que permiten la captura de una NPPF por un sustrato) y etiquetas NPPF. Las muestras de NPPF pueden separarse en pocillos separados o lugares que contienen uno o más cebadores específicos de NPPF diferentes, amplificarse y después secuenciarse por separado o combinarse para
5 secuenciación (o detección).

La amplificación puede usarse también para introducir una etiqueta detectable en los amplicones de NPPF generados (por ejemplo, si la NPPF originalmente no estaba etiquetada o si se deseaba un etiquetado adicional), u otra molécula que permite la detección o la inactivación. Por ejemplo, el cebador de amplificación puede incluir una
10 etiqueta detectable, haptano o agente de inactivación que se incorpora en la NPPF durante la amplificación. Dicha una etiqueta, haptano o agente de inactivación puede introducirse en cualquier extremo del amplicón de NPPF (o los dos extremos), o en cualquier lugar intermedio.

En algunos ejemplos, los amplicones de NPPF resultantes son limpiados antes de su detección o secuenciación. Por
15 ejemplo, la mezcla de reacción de amplificación puede limpiarse antes de la detección o secuenciación usando procedimientos bien conocidos en la técnica (por ejemplo, purificación en gel, captura y liberación de biotina/avidina, electroforesis capilar). En un ejemplo, los amplicones de NPPF son biotinilados (o incluyen otro haptano) y se capturan en avidina o una perla o superficie recubierta con anti-haptano, se lavan y después se liberan para detección o secuenciación. Análogamente, los amplicones de NPPF pueden ser capturados en un oligonucleótido
20 complementario (por ejemplo, unido a una superficie), lavados y después liberados para su detección o secuenciación. La captura de amplicones no tiene que ser particularmente específica, ya que los procedimientos divulgados eliminan la mayor parte del genoma o transcriptoma, dejando la NPPF que se ha hibridado con la molécula de ácido nucleico diana. Si se desea pueden usarse otros procedimientos para limpiar el producto amplificado.

Los productos amplificados también pueden limpiarse después de la última etapa de amplificación, siendo todavía
25 bicatenarios, por un procedimiento que usa una nucleasa que hidroliza los oligonucleótidos monocatenarios (por ejemplo, exonucleasa I), nucleasa que a su vez puede inactivarse antes de avanzar a la siguiente etapa, por ejemplo, hibridación a una superficie.

30

D. Detección de amplicones NPPF

En algunos ejemplos, los amplicones resultantes son detectados por cualquier medio adecuado, por ejemplo basado en la etiqueta detectable presente en los amplicones de NPPF. En un ejemplo específico no limitativo, los
35 amplicones de NPPF incluyen una etiqueta de biotina. En este ejemplo, los amplicones de NPPF pueden ser detectados incubando los amplicones (por ejemplo, en un soporte, por ejemplo, matriz o perla, que contiene los amplicones de NPPF) con avidina-HRP, estreptavidina-HRP o un conjugado con otra enzima adecuada tal como fosfatasa alcalina, y después la puesta en contacto del soporte con sustrato generador cromógeno, quimioluminiscencia o fluorescencia. En un ejemplo no limitativo, el sustrato es TMA-3 (Lumigen, Southfield, MI).
40 Existen sustratos quimioluminiscentes adicionales disponibles comercialmente, tales como LumiGlo® (KPL, Gaithersburg, MD), SuperSignal® (Pierce, Rockford, IL) y ECL™ (Amersham/GE Healthcare, Piscataway, NJ). La señal producida por el sustrato es detectada, por ejemplo, usando un dispositivo de imagen de micromatriz (por ejemplo, un dispositivo de imagen OMIX, OMIX HD, CAPELLA, o SUPERCAPELLA, HTG Molecular Diagnostics, Tucson, Arizona), o visualmente tal como en un dispositivo de flujo lateral. Puede usarse luminiscencia basada en
45 europio, así como electroluminiscencia o dispersión luminosa, o eléctrica (por ejemplo, conductividad o resistencia). En otro ejemplo, las NPPF incluyen una etiqueta fluorescente, tal como Cy-3 o Cy-5. Los amplicones de NPPF pueden ser detectados usando un dispositivo de imagen de micromatriz estándar (por ejemplo, un dispositivo de imagen Typhoon™ (GE Life Sciences, Piscataway, NJ), un explorador de micromatriz GenePix® (Molecular Devices, Sunnyvale, CA), un explorador GeneChip® (Affymetrix, Santa Clara, CA), procedimientos de citometría de flujo o
50 procedimientos de microscopía fluorescente. Un experto en la materia puede seleccionar procedimientos de detección y reactivos adecuados para estas u otras etiquetas detectables.

E. Detección de NPPF usando moléculas de captura

55 En algunas realizaciones, después de hibridación, tratamiento con nucleasa y amplificación, la muestra que contiene amplicones de NPPF se pone en contacto con una superficie que incluye múltiples regiones espacialmente discretas, cada una de las cuales incluye una molécula de captura, o se pone en contacto con una pluralidad de superficies, cada una de las cuales incluye una molécula de captura. Por ejemplo, la superficie puede ser una población de perlas, donde las subpoblaciones de las perlas incluyen cada una al menos una molécula de captura.

Por ejemplo una primera subpoblación podría incluir al menos una molécula de captura, mientras que una segunda subpoblación podría incluir al menos una molécula de captura que tiene una secuencia diferente a la primera, y así sucesivamente. En algunos ejemplos, la molécula de captura incluye al menos un anclaje asociado con un conector bifuncional (también referido como "conector de programación"). Alternativamente, la molécula de captura incluye una sonda de captura de ácidos nucleicos, que tiene una secuencia que es complementaria a al menos una parte de un amplicón de NPPF, por ejemplo, complementaria a parte o la totalidad de una región flanqueante de un amplicón de NPPF.

En un ejemplo donde la molécula de captura incluye al menos un anclaje asociado con un conector bifuncional, el anclaje y el conector bifuncional se asocian por hibridación, apareamiento, unión covalente u otra unión. El conector bifuncional incluye una primera parte que se une específicamente (por ejemplo, es complementaria) al anclaje y una segunda parte que se une específicamente (por ejemplo, es complementaria) a uno de la pluralidad de amplicones de NPPF (por ejemplo, complementaria a parte o la totalidad de la región (102) de la NPPF (100) mostrada en la FIG. 1).

En algunas realizaciones, los procedimientos divulgados incluyen un anclaje en una superficie (por ejemplo, en una matriz), que se asocia con un conector bifuncional que se usa para capturar los amplicones de NPPF después de la etapa de amplificación. En algunos ejemplos, un anclaje es un oligonucleótido de aproximadamente 8 a 150 nucleótidos de longitud (por ejemplo, aproximadamente 8 a 100, 15 a 100, 20 a 80, 25 a 75, ó de 25 a 50, por ejemplo aproximadamente 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140 ó 150 nucleótidos). En un ejemplo no limitativo, el anclaje tiene aproximadamente 25 nucleótidos de longitud. En algunos ejemplos, el anclaje incluye una primera parte que se une específicamente a la primera parte del conector bifuncional y una segunda parte que actúa como un espaciador entre la superficie y la primera parte del anclaje. En algunos ejemplos, la segunda parte del anclaje tiene aproximadamente 6 a 60 átomos de carbono o nucleótidos de longitud (por ejemplo, aproximadamente 6, 12, 24, 30, 36, 42, 48, 54 ó 60 átomos de carbono o nucleótidos). En otros ejemplos, la segunda parte del anclaje tiene aproximadamente 5 a 100 átomos de carbono o nucleótidos de longitud (por ejemplo, aproximadamente de 10 a 50, de 15 a 40, de 20 a 30, o aproximadamente 25 átomos de carbono o nucleótidos).

La composición de bases para anclajes para los procedimientos divulgados es tal que la estabilidad termodinámica del apareamiento entre anclaje y conector bifuncional es alta. En algunos ejemplos, la composición de bases porcentual para los anclajes es de aproximadamente el 30-40% de G, el 30-40% de C, el 10-20% de A y el 10-20% de T. En algunos ejemplos, la frecuencia del vecino más cercano en los anclajes minimiza los vecinos más cercanos G-G o C-C para reducir las reacciones laterales mediadas por la formación de cuarteto G. En otros ejemplos pueden incorporarse bases no naturales o ácidos nucleicos peptídicos en el anclaje o el conector bifuncional para modificar sus propiedades.

Los procedimientos de diseño y síntesis de anclajes de uso en los procedimientos divulgados se describen, por ejemplo, en la publicación PCT nº WO-98/24.098. En algunos ejemplos es conveniente un conjunto de anclajes que son sustancialmente diferentes entre sí. Un algoritmo de ejemplo para obtener un conjunto de anclajes diferentes es el siguiente:

- 1) Se define el tamaño del conjunto. En algunas realizaciones, 16, 24, 36, 48, 49, 64, 81, 96 y 100 constituyen tamaños útiles.
- 2) Se define la estructura global de la secuencia del conjunto de anclaje. Se usa la longitud y la composición de bases tal como se describe anteriormente para definir dichos parámetros. En general, el número de bases G y bases C se mantiene igual, así como el número de bases A y bases T. Esta igualdad optimiza la diversidad de configuración de los conjuntos finales. Así, dichos conjuntos se describirán mediante la ecuación $G_n C_n A_m T_m$.
- 3) Para una estructura de conjunto definida por m y n, se emplea un generador de números aleatorios para producir un conjunto de isómeros de secuencias aleatorias.
- 4) Se selecciona un miembro del conjunto de secuencias aleatorias para su uso como elemento #1 del conjunto.
- 5) Se define la semejanza máxima admisible entre miembros del conjunto. Se define la semejanza en términos de comparación local de bases por pares. Por ejemplo, cuando se alinean dos cadenas de oligómeros de longitud idéntica n de manera que los extremos 5' y 3' están en registro, la ausencia de fallos de correspondencia se refiere a la situación en que en todas las posiciones l-n, las bases en las dos cadenas son idénticas. Un fallo de correspondencia completo se refiere a la situación en que en todas las posiciones l-n, las bases en las dos cadenas son diferentes. Por ejemplo, una semejanza máxima útil podría ser 10 o más fallos de correspondencia en un conjunto de 16, sondas de captura de 16 mer.
- 6) Se selecciona un segundo miembro del conjunto de secuencias aleatorias y se determina su semejanza con el

elemento #1. Si el elemento #2 posee una semejanza admisible inferior al máximo con el elemento #1, se mantendrá en el conjunto. Si el elemento #2 posee una semejanza admisible superior al máximo, se descarta y se elige una nueva secuencia para comparación. Se repite este proceso hasta que se ha determinado un segundo elemento.

7) De una forma secuencial, se eligen miembros adicionales del conjunto de secuencias aleatorias que satisfacen las 5 constantes de no semejanza con respecto a todos los elementos seleccionados anteriormente.

En la Tabla 1 se muestra un ejemplo no limitativo de un conjunto de 16 anclajes que puede usarse en los procedimientos divulgados.

10

Tabla 1. Secuencias de anclaje de ejemplo

Secuencia de anclaje (5'→3')	SEQ ID NO:
TGATTCAGACCGGCCG	1
CCCGGGGCGTCTTAAC	2
GGACGCCATATGCGCT	3
TGAGGGCTCCGCCATA	4
AACCCGTGACGTGTGC	5
AGCATCGCCGGTCCTG	6
CCTGCAAGGCTGACGT	7
CAGTTGTCGACCCCGG	8
CGGCGCGTCCAATTCG	9
ATCGATCTGAGGGCCC	10
GTACATGCGGCCTGCA	11
TAGCCGCTCGCTAGAG	12
CCTAGTGATGACCGGC	13
GTCTGAGGGCAACCTC	14
CTAGCTGGCTACGCAG	15
GCCATCCGCTTGGAGC	16

En otros ejemplos la molécula de captura incluye al menos una sonda de captura de ácidos nucleicos, que tiene una secuencia que es complementaria a al menos una parte de un amplicón de NPPF, por ejemplo complementaria a parte o la totalidad de una región flanqueante de un amplicón de NPPF. Por ejemplo, la sonda de captura de ácidos nucleicos puede incluir una región que es complementaria al amplicón de NPPF, y puede incluir una región que no lo es (por ejemplo, una región que permite la fijación de la sonda a una superficie). La sonda de captura de ácidos nucleicos puede estar unida directamente a una superficie. Por ejemplo, la sonda de captura de ácidos nucleicos puede incluir una amina para fijación covalente a una superficie. En algunos ejemplos, una sonda de captura de ácidos nucleicos es un oligonucleótido de al menos 8 nucleótidos de longitud, por ejemplo al menos 10, al menos 15, al menos 20, al menos 30, al menos 50 o al menos 100 nucleótidos de longitud (por ejemplo, aproximadamente de 8 a 100, de 15 a 100, de 20 a 80, de 25 a 75 o de 25 a 50, por ejemplo aproximadamente 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140 ó 150 nucleótidos). Un experto en la materia observará que no es preciso que la región de la sonda de captura de ácidos nucleicos complementaria a una región del amplicón de NPPF tenga una complementariedad del 100%, siempre y cuando puede producirse una hibridación entre la sonda de captura de ácidos nucleicos y amplicones de NPPF apropiados. En algunos ejemplos, la región de la sonda de captura de ácidos nucleicos complementaria a una región del amplicón de NPPF tiene al menos 8 nucleótidos de longitud, por ejemplo al menos 8, al menos 10, al menos 15, al menos 20, al menos 30, al menos 50 o al menos 100 nucleótidos de longitud (por ejemplo, aproximadamente de 8 a 100, de 15 a 100, de 20 a 80, de 25 a 75 o de 25 a 50, por ejemplo aproximadamente 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140 ó 150 nucleótidos de longitud).

En algunos ejemplos, la muestra que contiene amplicones de NPPF se desnaturaliza antes de la puesta en contacto con la superficie de la matriz (por ejemplo, por calentamiento a 95°C durante 5 minutos y enfriamiento rápido de la muestra en hielo). En algunos ejemplos, la muestra que contiene NPPF se ajusta antes de la puesta en contacto con la superficie (por ejemplo, para ajustar la concentración de sal o formamida). La muestra que contiene amplicones de NPPF se incuba con la superficie (por ejemplo, una matriz o perlas) durante un periodo de tiempo suficiente para que los amplicones de NPPF se unan específicamente (por ejemplo, se hibriden) con la molécula de captura. En algunos ejemplos, la incubación de la muestra con la superficie a entre aproximadamente 37°C y aproximadamente 65°C (por ejemplo, entre aproximadamente 45°C y aproximadamente 60°C, o aproximadamente 50°C y aproximadamente 60°C, por ejemplo 50°C) durante al menos 1 hora (por ejemplo, de 1 a 8 horas, de 1 a 36 horas, de 12 a 24 horas o de 16 a 24 horas, o durante toda la noche) para permitir que los amplicones de NPPF se hibriden

con la molécula de captura ("captura de NPPF"). El tiempo de captura puede acortarse, por ejemplo si se usan dispositivos microfluidicos o macrofluidicos, dispositivos de flujo lateral o reduciendo la difusión y usando flujo o mezclado activo.

- 5 Algunas de las superficies (o sustratos) que pueden usarse en los procedimientos divulgados están disponibles fácilmente en proveedores comerciales. En algunas realizaciones, la superficie es una placa de microvaloración de 96, 384 ó 1.536 pocillos, por ejemplo, placas modificadas comercializadas por Corning Costar. En otras realizaciones, un sustrato incluye una o más perlas (por ejemplo, una población de perlas que pueden diferenciarse por tamaño o color, por ejemplo por citometría de flujo). Alternativamente, puede formarse una superficie que
 10 comprende pocillos que, a su vez, comprenden endentados o "depresiones" por micromecanizado de una sustancia tal como aluminio o acero para preparar un molde, y después por microinyección de plástico u otro material similar en el molde con el fin de formar una estructura. Alternativamente, puede ensamblarse una estructura formada por vidrio, plástico, cerámica o similares. El espaciador puede ser, por ejemplo, una pieza de material, por ejemplo, silicona, con orificios separados en toda su longitud, de manera que cada orificio formará las paredes de un pocillo
 15 de prueba cuando se unan las tres piezas. El subdivisor puede ser, por ejemplo, una pieza delgada de material, por ejemplo, silicona, modelada en forma de un tamiz o una malla fina. El divisor en la superficie que separa diferentes reacciones puede ser también una superficie recubierta a la que no se adherirán las soluciones, o una nanoestructura, o simplemente ser gotas individuales, o capilares o canales o lugares microfluidicos. En algunos ejemplos, la base es una pieza plana de material (por ejemplo, vidrio o plástico), por ejemplo, en forma de la parte
 20 inferior de una microplaca típica usada para un ensayo bioquímico. La superficie superior de la base puede ser plana, o puede formarse con endentados que se alinearán con la forma del subdivisor para proporcionar subdivisiones completas, o pocillos, en cada pocillo de muestra. Las tres piezas pueden unirse mediante procedimientos estándar, por ejemplo, los procedimientos usados en el ensamblaje de obleas de silicio.
- 25 Los materiales adecuados para la superficie incluyen, pero no se limitan a: vidrio, sílice, oro, plata, un gel o un polímero, nitrocelulosa, polipropileno, polietileno, polibutileno, poliisobutileno, polibutadieno, poliisopreno, polivinilpirrolidina, politetrafluoroetileno, difluoruro de polivinilideno, polifluoroetileno-propileno, alcohol polietilenevinílico, polimetilpenteno, policolorotrifluoroetileno, polisulfonas, polipropileno orientado biaxialmente hidroxilado, polipropileno orientado biaxialmente aminado, polipropileno orientado biaxialmente tiolado, ácido
 30 etilenoacrílico, ácido tilenometacrílico, y mezclas de copolímeros de los mismos (véase la patente de EE.UU. nº 5.985.567), o formado por nanomateriales que incluyen carbono.

En general, las características adecuadas del material que puede usarse para formar la superficie incluyen: ser asequible para activación superficial de manera que tras la activación, la superficie del soporte sea capaz de unirse
 35 de forma covalente a una biomolécula tal como un oligonucleótido (por ejemplo, anclaje); asequibilidad para síntesis "in situ" de biomoléculas; condición de ser inerte químicamente de manera que las áreas del soporte no ocupadas por oligonucleótidos o proteínas no son accesibles para unión no específica, o cuando se produce unión no específica, dichos materiales pueden retirarse fácilmente de la superficie sin eliminar los oligonucleótidos o proteínas. Las superficies pueden ser permeables, parcialmente permeables o impermeables.

40 Puede emplearse una amplia variedad de formatos de matriz para la configuración de los anclajes de acuerdo con la presente descripción. Un formato adecuado incluye un patrón bidimensional de células discretas (por ejemplo, 4.096 cuadrados en una matriz de 64 por 64). Tal como observarán los expertos en la materia, otros formatos de matriz que incluyen, pero no se limitan a, matrices de rendija (rectangulares) y circulares son también adecuados para su
 45 uso (véase la patente de EE.UU. nº 5.981.185). En algunos ejemplos, la matriz es una placa con múltiples pocillos.

Los anclajes de oligonucleótido, los conectores bifuncionales y otras moléculas de captura (así como sondas/cebadores de PCR, NPPF y CFS) pueden sintetizarse mediante tecnología convencional, por ejemplo, con un sintetizador comercial de oligonucleótidos y/o por ligamiento de subfragmentos que han sido sintetizados así. Los
 50 ácidos nucleicos que son demasiado largos para ser sintetizados de manera fiable por dichos procedimientos pueden generarse por procedimientos de amplificación, usando procedimientos convencionales.

En una realización, anclajes preformados de ácidos nucleicos (por ejemplo, anclajes de oligonucleótidos) o sondas de captura de ácidos nucleicos que tienen una secuencia complementaria a al menos una parte de un amplicón de
 55 NPPF (por ejemplo, sondas de oligonucleótidos), pueden situarse sobre o en la superficie de una región de prueba mediante cualquiera de una diversidad de técnicas convencionales, que incluyen unión química fotolitográfica o de serigrafía, disposición por tecnología de inyección de tinta, capilar, chip de cribado o canal de fluidos, conformación electroquímica mediante matrices de electrodos, puesta en contacto con una clavija o manguito o desnaturalización seguida por horneado o irradiación UV en filtros (véase, por ejemplo, Rava y col. (1996). Patente de EE.UU. nº

5.545.531; Fodor y col. (1996). Patente de EE.UU. nº 5.510.270; Zanzucchi y col. (1997). Patente de EE.UU. nº 5.643.738; Brennan (1995). Patente de EE.UU. nº 5.474.796; documento PCT WO-92/10092; documento PCT WO-90/15070). Los anclajes o sondas de oligonucleótidos pueden colocarse en la parte superior de la superficie de una región de prueba o, por ejemplo en el caso de un lecho de gel de poliacrilamida gel, pueden integrarse en la superficie de tal manera que parte del anclaje o sonda sobresale de la estructura del gel en porciones acuosas en el gel y la superficie de gel y está disponible para interacciones con un conector o NPPF. Esto se cumple para superficies permeables y superficies parcialmente permeables, tales como una superficie donde la primera parte, por ejemplo el área de la superficie en contacto con las soluciones que contienen conectores bifuncionales o NPPF es permeable pero una segunda parte, por ejemplo a cierta distancia en la superficie, no es permeable. En una realización, los anclajes o sondas preformados de oligonucleótidos se derivan en el extremo 5' con un grupo amino libre; se disuelven a una concentración determinada de forma rutinaria empíricamente (por ejemplo, aproximadamente 1 μ M) en un tampón tal como tampón de fosfato 50 mM, pH 8,5 y EDTA 1 mM; y se distribuyen con un dispensador de nanoinyección Pixus (Cartesian Technologies) en gotitas de aproximadamente 10,4 nanolitros en lugares específicos dentro de un pocillo de prueba cuya superficie superior es la de una placa DNA Bind seca y nueva (Corning Costar). Dependiendo de la velocidad relativa de fijación y evaporación de oligonucleótidos, puede ser necesario controlar la humedad en los pocillos durante la preparación. En otra realización, los anclajes o sondas de oligonucleótidos pueden sintetizarse directamente en la superficie de una región de prueba, usando procedimientos convencionales como, por ejemplo, desprotección activada por luz de cadenas de oligonucleótidos en crecimiento (por ejemplo, en conjunción con el uso de una "máscara" de direccionamiento del sitio) o por dispensación pautada de gotitas de nanolitros de compuesto de desactivación usando un dispensador de nanoinyección. Por ejemplo, puede realizarse la desprotección de todos los oligonucleótidos en crecimiento que se usan para recibir un nucleótido individual, y a continuación se añade el nucleótido a través de la superficie. En otra realización, los anclajes o sondas de oligonucleótidos se unen a la superficie a través de los extremos 3' de los oligonucleótidos, usando la metodología convencional.

F. Detección de NPP usando procedimientos alternativos

En algunas realizaciones, después de hibridación, tratamiento con nucleasa y amplificación, los amplicones de NPPF se detectan usando procedimientos alternativos, por ejemplo plataformas de alto rendimiento. En algunos ejemplos, los amplicones de NPPF se detectan usando electroforesis en gel, cromatografía, espectrometría de masas, secuenciación, análisis convencional de micromatrices, detectado durante la etapa de amplificación por PCR, o captura híbrida. En algunas realizaciones, los amplicones de NPPF no incluyen una etiqueta detectable y se usan procedimientos de detección indirecta. Dichos procedimientos son conocidos por los expertos en la materia e incluyen, pero no se limitan a, los descritos en la presente memoria descriptiva.

En un ejemplo, los amplicones de NPPF se detectan usando un ensayo basado en perlas, tal como una matriz de perlas. Un ejemplo de un ensayo basado en perlas usa perlas X-MAP® (Luminex, Austin, TX), por ejemplo un ensayo QBEAD. En algunos ejemplos, las NPP son capturadas en perlas X-MAP® u otras perlas por hibridación en un oligonucleótido asociado con las perlas (por ejemplo, aproximadamente 1 hora a aproximadamente 50°C). La etiqueta detectable incluida en los amplicones de NPPF puede detectarse, por ejemplo, por citometría de flujo (por ejemplo, usando un instrumento Luminex 200, Flexmap 3D, u otro instrumento adecuado).

En otro ejemplo, los amplicones de NPPF se detectan usando una micromatriz estándar. Un ejemplo de dicha matriz es una micromatriz Nimblegen (Nimblegen, Madison, WI). En algunos ejemplos, los amplicones de NPPF son hibridados en una matriz que incluye oligonucleótidos que se unen específicamente a los amplicones de NPPF. La etiqueta detectable incluida en los amplicones de NPPF puede ser detectada.

En algunos ejemplos, los amplicones de NPPF son detectados con un ensayo de "código de barras". Un ejemplo de dicho ensayo es nCounter® Analysis System (Nanostring Technologies, Seattle, WA). En algunos ejemplos, los amplicones de NPPF se hibridan en una sonda que incluye una o más etiquetas con códigos de color (un "código de barras"). La detección de las etiquetas con códigos de color proporciona la identificación del amplicón de NPPF. Véanse, por ejemplo, los documentos WO-07/0.761.282; WO-07/076.129; WO-07/139.766.

F. Secuenciación de amplicones

En algunos ejemplos, los amplicones de NPPF resultantes son secuenciados, por ejemplo por secuenciación del amplicón de NPPF complejo o de una parte del mismo (por ejemplo, una cantidad suficiente para permitir la identificación de la molécula de ácido nucleico diana). La descripción no se limita a un procedimiento de secuenciación en particular. En algunos ejemplos, se secuencian múltiples amplicones de NPPF diferentes en una

sola reacción. En un ejemplo, puede secuenciarse una etiqueta experimental del amplicón de NPPF, que puede diseñarse para que se corresponda con una secuencia diana en particular. Así, si el extremo 3' del amplicón de NPPF tiene una secuencia en el terminal de 2 a 25 nucleótidos (por ejemplo, el terminal 2 a 5 ó 2 a 7, por ejemplo el terminal de 2, 3, 4, 5, 6, 7, 8, 9 ó 10 nucleótidos) que representan una secuencia individual para cada diana medida, entonces todo el amplicón de NPPF debe ser secuenciado para identificar la diana, y contando el número de dichas etiquetas experimentales secuenciadas puede determinarse la cantidad de cada diana en la muestra.

En un ejemplo, los amplicones de NPPF resultantes, por ejemplo uno compuesto por ADN, son secuenciados usando el procedimiento de terminación de cadena. Esta técnica usa terminación específica de cadena de una reacción de síntesis de ADN usando sustratos de nucleótidos modificados. En la secuenciación de terminador de cadena, la extensión se inicia en un sitio específico en el amplicón de NPPF usando un cebador de oligonucleótidos complementario a una parte del amplicón de NPPF. El cebador de oligonucleótidos se extiende usando una polimerasa, por ejemplo, ARN o ADN polimerasa. Con el cebador y la polimerasa se incluyen las cuatro bases de desoxinucleótidos (o ribonucleótidos), junto con una baja concentración de un nucleótido de terminación de cadena (comúnmente un didesoxinucleótido). La incorporación limitada del nucleótido de terminación de cadena por la polimerasa produce una serie de fragmentos de ácidos nucleicos relacionados que terminan sólo en posiciones en las que se usa un nucleótido en particular. A continuación, los fragmentos son separados por tamaño, por ejemplo por electroforesis en un gel de poliacrilamida en bloques, o en un tubo de vidrio estrecho (capilar) relleno con un polímero viscoso.

Un procedimiento alternativo es la secuenciación con terminador de colorante. Este enfoque permite la secuenciación completa en una sola reacción, en vez de en las cuatro necesarias con el procedimiento de terminación de cadena. Se consigue etiquetando cada uno de los terminadores de cadena de didesoxinucleótidos con un colorante fluorescente separado, que emite fluorescencia en una longitud de onda diferente.

En otro ejemplo se usa pirosecuenciación, por ejemplo con los procedimientos comercializados por Biotage (para secuenciación de bajo rendimiento) y 454 Life Sciences (para secuenciación de alto rendimiento). En el procedimiento basado en matrices (por ejemplo, 454 Life Sciences), se aparea un ácido nucleico monocatenario (por ejemplo, ADN) con perlas y se amplifica mediante EmPCR. Estas perlas unidas a ácidos nucleicos se colocan a continuación en pocillos en un chip de fibra óptica junto con enzimas que producen luz en presencia de ATP. Cuando se lavan nucleótidos libres en este chip, se produce luz a medida que se genera ATP cuando los nucleótidos se unen con sus pares de bases complementarias. La adición de uno o más nucleótidos produce una reacción que genera una señal luminosa que se registra, por ejemplo, mediante una cámara CCD. La intensidad de la señal es proporcional al número de nucleótidos, por ejemplo, extensiones de homopolímeros, incorporados en un único flujo de nucleótidos.

En otro ejemplo, los amplicones de NPPF son secuenciados usando Illumina® (por ejemplo, HiSeq) o Ion Torrent®, 454®, Helicos, PacBio®, Solid® (Applied Vioasystems) o cualquier número de otros sistemas de secuenciación comerciales. Para la captura se usan adaptadores de secuenciación (por ejemplo, colas poli-A o poli-T presentes en los amplicones de NPPF, por ejemplo, introducidos mediante PCR). La secuenciación por 454® o Illumina® implica normalmente preparación de bibliotecas, realizada por fragmentación de ADN aleatoria, seguido por ligamiento *in vitro* de secuencias de adaptadores comunes. Para los procedimientos divulgados, puede eliminarse la etapa de fragmentación aleatoria del ácido nucleico que se secuenciará, y el ligamiento *in vitro* de secuencias de adaptadores puede realizarse con los amplicones de NPPF, tales como una etiqueta experimental presente en los amplicones de NPPF, aunque también pueden incorporarse mediante el uso de las regiones flanqueantes y PCR, eludiendo la necesidad de ligamiento. Una vez realizada la captura a través de adaptadores de secuenciación en el chip/perla de secuenciación, se realiza amplificación mediante puentes para formar colonias de cada sonda para secuenciación. En estos procedimientos, los amplicones de NPPF terminan por agruparse espacialmente, ya sea en una única localización en un sustrato plano (Illumina®, colonias *in situ*, PCR tipo puente) o en la superficie de perlas de escala micrométrica (454®, PCR de emulsión), que pueden recuperarse y disponerse en matrices (PCR de emulsión). El procedimiento de secuenciación incluye ciclos alternos de bioquímica activada por enzimas y adquisición de datos basada en imágenes. Estas plataformas se basan en la secuenciación por síntesis, es decir, la extensión en serie de plantillas cebadas. Se usan iteraciones sucesivas de interrogación enzimática y técnicas de imagen para construir una lectura de secuenciación contigua para cada característica de matriz. Los datos se adquieren mediante imagen de la matriz completa en cada ciclo (por ejemplo, de nucleótidos etiquetados por fluorescencia incorporados por una polimerasa). Puede usarse más de un cebador de secuenciación en las colonias formadas en la célula de flujo, lo que permite secuenciación por extremo doble o secuenciación de una o más partes del amplicón, por ejemplo, un código de barras o una etiqueta de índice, o una etiqueta experimental.

- Para 454®, se hibrida un cebador de secuenciación con la NPPF después de amplificación en el amplicón del chip/perla de secuenciación. La secuenciación se realiza por pirosecuenciación. Las perlas portadoras de amplicones son preincubadas con polimerasa de *Bacillus stearothermophilus* (Bst) y proteína de unión monocatenaria, y después se depositan en una matriz microfabricada de pocillos de escala de picolitros, una perla por pocillo, lo que hace esta bioquímica compatible con secuenciación basada en matriz. También se añaden perlas más pequeñas, que llevan enzimas inmovilizadas necesarias también para la pirosecuenciación (ATP sulfurilasa y luciferasa). Durante la secuenciación, un lado de la matriz semiordenada actúa como una célula de flujo para introducir y eliminar reactivos de secuenciación. El otro lado está unido a un haz de fibra óptica para detección de señales basada en CCD. En cada ciclo, se introduce una única especie de nucleótido no etiquetado. Para 10 secuencias donde esta introducción produce una incorporación, el pirofosfato es liberado por ATP sulfurilasa y luciferasa, generando una ráfaga de luz detectada por el CCD para coordenadas de matriz específicas. A través de múltiples ciclos, el patrón de episodios de incorporación detectados revela la secuencia de plantillas representada por perlas individuales.
- 15 En procedimientos que usan PCR tipo puente (por ejemplo, Illumina®), las características de secuenciación amplificadas son generadas por PCR tipo puente. Se fijan cebadores de PCR directos e inversos a un sustrato sólido mediante un conector flexible, de manera que todos los amplicones que proceden de cualquier molécula de plantilla individual durante la amplificación permanecen inmovilizados y agrupados en una única posición física en una matriz. En algunos ejemplos, la PCR tipo puente usa ciclos de extensión alternos con polimerasa Bst y 20 desnaturalización (por ejemplo, con formamida). Los 'grupos' resultantes consisten cada uno aproximadamente en 1.000 amplicones clonales. Pueden amplificarse varios millones de grupos para localizaciones diferenciables en cada una de las ocho 'calles' independientes que están en una única célula de flujo (de manera que pueden secuenciarse en paralelo ocho experimentos independientes durante la ejecución del mismo instrumento). Después de la generación de grupos, los amplicones son linealizados y se hibrida un cebador de secuenciación con una 25 secuencia de adaptador universal que flanquea la región de interés. Cada ciclo de interrogación de secuencia consiste en una extensión de base única con una ADN polimerasa modificada y una mezcla de cuatro nucleótidos. Estos nucleótidos son 'terminadores reversibles', en el sentido de que una fracción escindible químicamente en la posición hidroxilo en 3' permite que tenga lugar sólo una incorporación de base única en cada ciclo, y una de cuatro etiquetas fluorescentes, también químicamente escindibles, corresponde a la identidad de cada nucleótido. Después 30 de una extensión de base única y de la adquisición de imágenes en cuatro canales, se establece la escisión química de los dos grupos para el siguiente ciclo. En la actualidad se realizan rutinariamente lecturas de longitudes de hasta 36 pb.

- En un ejemplo, se usa el procedimiento de secuenciación de molécula individual Helicos® o PacBio®.
- 35 Se observará que la NPPF puede diseñarse para secuenciación mediante cualquier procedimiento, en cualquier secuenciador desarrollado en la actualidad o en el futuro. La NPPF en sí no limita el procedimiento de secuenciación usado, ni la enzima que se use. Existen o se desarrollarán otros procedimientos de secuenciación, y un experto en la materia puede observar que los amplicones de NPPF generados (o ADN hibridado con la NPPF) serán adecuados 40 para secuenciación en estos sistemas.

G. Controles

- En algunas realizaciones, el procedimiento incluye el uso de uno o más controles positivos y/o negativos sujeto a las 45 mismas condiciones de reacción que las NPPF experimentales reales. El uso de retardo permite usar en la práctica diferentes muestras como controles pero procesadas para secuenciación y ejecución en la misma calle de secuenciación que las muestras de prueba. El ADN puede medirse como control para el número de células cuando se mide el ARN diana.
- 50 En algunos ejemplos, un "control positivo" incluye un control de normalización interno para variables como el número de células lisadas para cada muestra, la recuperación de ADN o ARN o la eficiencia de hibridación, por ejemplo una o más NPPF, CFS, conectores correspondientes, y similares, que son específicos para uno o más genes de mantenimiento de nivel basal I o constitutivos, tales como genes estructurales (por ejemplo, actina, tubulina u otros) o proteínas de unión a ADN (por ejemplo, factores de regulación de transcripción, u otros). En algunos ejemplos, un 55 control positivo incluye gliceraldehído-3-fosfato deshidrogenasa (GAPDH), peptidilprooil-isomerasa A (PPIA), proteína ribosómica grande (RPLP0), proteína ribosómica L 19 (RPL19) u otros genes de mantenimiento expuestos más adelante. En la muestra pueden aparecer picos de otros controles positivos para controlar el proceso de ensayo, con independencia de la muestra.

En otros ejemplos, un control positivo incluye una NPPF específica para un ADN o ARN del que se sabe que está presente en la muestra (por ejemplo, una secuencia de ácidos nucleicos que estará presente probablemente en la especie sujeta a ensayo, tal como un gen de mantenimiento). Por ejemplo, la NPPF de control positivo correspondiente puede añadirse a la muestra antes o durante la hibridación con la pluralidad de NPPF de prueba.

5 Alternativamente, la NPPF de control positivo se añade a la muestra después de tratamiento con nucleasa.

En algunos ejemplos, un control positivo incluye una molécula de ácido nucleico de la que se sabe que está presente en la muestra (por ejemplo, una secuencia de ácidos nucleicos presente probablemente en la especie sujeta a ensayo, tal como un gen de mantenimiento). La molécula de ácido nucleico de control positivo correspondiente (por

10 ejemplo, ácido nucleico transcrito *in vitro* o ácido nucleico aislado a partir de una muestra no relacionada) puede añadirse a la muestra antes o durante hibridación con la pluralidad de NPPF.

En algunos ejemplos, un "control negativo" incluye uno o más NPPF, CFS, conectores correspondientes, o similares, cuyo complemento se sabe que está presente en la muestra, por ejemplo como un control para especificidad de

15 hibridación, tal como una secuencia de ácidos nucleicos de una especie distinta de la sujeta a ensayo, por ejemplo, una secuencia de ácidos nucleicos de una planta cuando se analizan ácidos nucleicos humanos (por ejemplo, factor de transcripción que responde al etileno del tipo *Arabidopsis thaliana* AP2 (ANT)), o una secuencia de ácidos nucleicos no presentes en la naturaleza.

20 En algunas realizaciones, la señal de cada amplicón de NPPF se normaliza con la señal de al menos una molécula de ácido nucleico de mantenimiento, por ejemplo para tener en cuenta las diferencias en la celularidad entre muestras. Los genes de mantenimiento de ejemplo incluyen uno o más de GAPDH (gliceraldehído 3-fosfato deshidrogenasa), SDHA (succinato deshidrogenasa), HPRT1 (hipoxantina fosforribosil-transferasa 1), HBS1L (proteína de tipo HBS1), β -actina (ACTB), β -2 microglobulina (B2m) y AHSP (proteína de estabilización de alfa-

25 hemoglobina). Un experto en la materia puede seleccionar genes de mantenimiento adicionales para su uso en la normalización de señales en los procedimientos divulgados, lo que incluye, pero no se limita a proteína ribosómica S13 (RPS13), proteína ribosómica S20 (RPS20), proteína ribosómica L27 (RPL27), proteína ribosómica L37 (RPL37), proteína ribosómica 38 (RPL38), antienzima ornitina-descarboxilasa 1 (OAZ1), polimerasa (ARN) II (dirigida a ADN) polipéptido A, (220) kDa (POLR2A), proteína asociada sí 1 (YAP1), esterasa D (ESD), proteasoma

30 (prosoma, macrodolor) subunidad 26S, ATPasa, 1 (PSMC1), factor de inicio de traducción eucariótico 3, subunidad A (EIF3A), o rRNA 18S (véase, por ejemplo, de Jonge y col., *PLoS One* 2:e898, 2007; Saviozzi y col., *BMC Cancer* 6:200, 2006; Kouadjo y col., *BMC Genomics* 8:127, 2007; cada uno de los cuales se incorpora en la presente memoria descriptiva como referencia). Los valores normalizados pueden compararse directamente entre muestras o ensayos (por ejemplo, entre dos muestras diferentes en un mismo ensayo o entre la misma muestra sometida a

35 prueba en dos ensayos separados).

IV. Sondas de protección contra la nucleasa con secuencias flanqueantes (NPPF)

Los procedimientos divulgados permiten la detección y/o secuenciación de una o más moléculas de ácidos nucleicos

40 diana, por ejemplo de forma simultánea o contemporánea. Basándose en el ácido nucleico diana, las NPPF pueden diseñarse para su uso en los procedimientos divulgados usando los criterios expuestos en la presente memoria descriptiva en combinación con el conocimiento de un experto en la materia. En algunos ejemplos, los procedimientos divulgados incluyen la generación de una o más NPPF apropiadas para la detección de moléculas de ácidos nucleicos diana en particular. La NPPF, en diversas condiciones (conocidas o determinadas

45 empíricamente), se une específicamente (o es capaz de unirse específicamente) a un ácido nucleico diana o parte del mismo, si dicha diana está presente en la muestra.

Las NPPF incluyen una región que es complementaria a una molécula de ácido nucleico diana, de manera que para cada secuencia de ácidos nucleicos diana en particular existe al menos una NPPF en la reacción que es específica

50 para la secuencia de ácidos nucleicos diana. Por ejemplo, si existen 2, 3, 4, 5, 6, 7, 8, 9 ó 10 secuencias de ácidos nucleicos diana diferentes para su detección o secuenciación, el procedimiento usará de forma correspondiente al menos 2, 3, 4, 5, 6, 7, 8, 9 ó 10 NPPF diferentes (donde cada NPPF corresponde a una diana en particular). Así en algunos ejemplos, los procedimientos usan al menos dos NPPF, donde cada NPPF es específica para una molécula de ácido nucleico diana diferente. Sin embargo, se observará que pueden generarse varios NPPF diferentes para

55 una molécula de ácido nucleico diana en particular, por ejemplo muchas regiones diferentes de una única secuencia de ácidos nucleicos diana. En un ejemplo, una NPPF incluye una región que es complementaria a una secuencia encontrada sólo en un único gen en el transcriptoma. Las NPPF están diseñadas para ser específicas para una molécula de ácido nucleico diana y para tener Tf similares (si deben usarse en la misma reacción).

Así, una muestra individual puede ponerse en contacto con una o más NPPF. Un conjunto de NPPF es una colección de dos o más NPPF cada una específica para una diana diferente y/o una parte diferente de una misma diana. Un conjunto de NPPF puede incluir al menos, hasta o exactamente, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 15, 20, 25, 30, 50, 100, 500, 1.000, 2.000, 3.000, 5.000 ó 10.000 NPPF diferentes. En algunos ejemplos, una muestra se pone en contacto con una cantidad suficiente de NPPF para que esté en exceso con respecto a la diana para dicha NPPF, por ejemplo un exceso de 100 veces, 500 veces, 1.000 veces, 10.000 veces, 100.000 veces o 10^6 veces. En algunos ejemplos, si se usa un conjunto de NPPF, cada NPPF del conjunto puede proporcionarse en exceso con su diana respectiva (o parte de una diana) en la muestra. La NPPF en exceso puede facilitar la cuantificación de la cantidad de NPPF que se une con una diana en particular. Algunas realizaciones del procedimiento implican una pluralidad de muestras (por ejemplo, al menos, hasta o exactamente, 10, 25, 50, 75, 100, 500, 1.000, 2.000, 3.000, 5.000 ó 10.000 muestras diferentes) de forma simultánea o contemporánea en contacto con la misma NPPF o conjunto de NPPF.

La FIG. 1 muestra una NPPF de ejemplo (100) que tiene una región (102) que incluye una secuencia que se une específicamente a o se hibrida con la secuencia de ácidos nucleicos diana, así como secuencias flanqueantes (104, 106) en los extremos 5' y 3' de la NPPF, donde las secuencias flanqueantes se unen o se hibridan con sus secuencias complementarias (referidas en la presente memoria descriptiva como CFS). Las NPPF (así como las CFS que se unen a las NPPF) pueden estar compuestas por nucleótidos naturales (por ejemplo, ribonucleótidos (ARN), o desoxirribonucleótidos (ADN)) o no naturales (por ejemplo, ácidos nucleicos bloqueados (LNA, véase, por ejemplo, patente de EE.UU. nº 6.794.499), ácidos nucleicos peptídicos (PNA)), y similares. Las NPPF pueden ser monocatenarias o bicatenarias. En algunos ejemplos, las NPPF incluyen una o más bases sintéticas o bases alternativas (por ejemplo, inosina). En las NPPF pueden usarse nucleótidos modificados, nucleótidos no naturales, sintéticos o alternativos en una o más posiciones (por ejemplo, 1, 2, 3, 4, 5 posiciones o más). En algunos ejemplos, el uso de uno o más nucleótidos modificados o no naturales en la NPPF puede incrementar la T_f de la NPPF con respecto a la T_f de una NPPF de la misma longitud y composición que no incluye el ácido nucleico modificado. Un experto en la materia puede diseñar sondas que incluyen dichos nucleótidos modificados para obtener una sonda con una T_f deseada. En un ejemplo, una NPPF está compuesta por ADN o ARN, por ejemplo ADN monocatenario (ssDNA) o ADN ramificado (bdNA). En un ejemplo, una NPPF es un aptámero.

Los procedimientos para la determinación empírica del tamaño apropiado de una NPPF para su uso con dianas o muestras en particular (por ejemplo, muestras fijas o reticuladas) son rutinarios. En realizaciones específicas, una NPPF puede tener hasta 500 nucleótidos de longitud, por ejemplo hasta 400, hasta 250, hasta 100 o hasta 75 nucleótidos de longitud, incluido, por ejemplo, el intervalo de 20-500, 20-250, 25-200, 25-100, 25-75 ó 25-50 nucleótidos de longitud. En un ejemplo no limitativo, una NPPF tiene al menos 35 nucleótidos de longitud, por ejemplo al menos 40, al menos 45, al menos 50, al menos 75, al menos 100, al menos 150 o al menos 200 nucleótidos de longitud, por ejemplo de 50 a 200, de 50 a 100 o de 75 a 200, o 36, 72 ó 100 nucleótidos de longitud. Las realizaciones de NPPF en particular pueden ser más largas o más cortas dependiendo de la funcionalidad deseada. En algunos ejemplos, la NPPF está dimensionada apropiadamente (por ejemplo, suficientemente pequeña) para penetrar en las muestras fijas y/o reticuladas. Las muestras fijas o reticuladas pueden variar en el grado de fijación o reticulación; así, un experto en la materia puede determinar un tamaño apropiado de NPPF para una condición o tipo de muestra en particular, por ejemplo, realizando una serie de experimentos que usan muestras con concentraciones de diana fijas conocidas y comparando el tamaño de NPPF con la intensidad de la señal diana. A medida que aumenta la longitud de NPPF, en dicho experimento, en algún punto la intensidad de la señal diana debería empezar a disminuir. En algunos ejemplos, la muestra (y, por tanto, al menos una proporción de diana) es fija o reticulada, y la NPPF es suficientemente pequeña para que la intensidad de la señal siga siendo alta y no varíe sustancialmente en función del tamaño de NPPF.

La secuencia (102) que se une específicamente a la secuencia de ácidos nucleicos diana es complementaria en secuencia a la secuencia de ácidos nucleicos diana que debe detectarse o secuenciarse. Un experto en la materia observará que no es necesario que la secuencia (102) sea complementaria a un ácido nucleico diana completo (por ejemplo, si la diana es un gen de 100.000 nucleótidos, la secuencia (102) puede ser una parte del mismo, por ejemplo al menos 10, al menos 15, al menos 20, al menos 25, al menos 30, al menos 40, al menos 50, al menos 100 o más nucleótidos consecutivos complementarios a una molécula de ácido nucleico diana en particular). La especificidad de una sonda aumenta con la longitud. Así, por ejemplo, una secuencia (102) que se une específicamente a la secuencia de ácidos nucleicos diana que incluye 25 nucleótidos consecutivos se hibridará con la secuencia diana con mayor especificidad que una secuencia correspondiente de sólo 15 nucleótidos. Así, las NPPF divulgadas en la presente memoria descriptiva pueden tener una secuencia (102) que se une específicamente a la secuencia de ácidos nucleicos diana que incluye al menos 10, al menos 15, al menos 20, al menos 25, al menos 30, al menos 40, al menos 50, al menos 100, o más nucleótidos consecutivos complementarios a una molécula de ácido nucleico diana en particular (por ejemplo, aproximadamente de 6 a 50, de 10 a 40, de 10 a 60, de 15 a 30, de

18 a 23, de 19 a 22 o de 20 a 25 nucleótidos consecutivos complementarios a un ADN diana o un ARN diana). Las longitudes de secuencia (102) en particular que se unen específicamente a la secuencia de ácidos nucleicos diana que puede formar parte de las NPPF usadas para aplicar los procedimientos de la presente descripción incluyen 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, 5 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 ó 50 nucleótidos contiguos complementarios a una molécula de ácido nucleico diana. En algunos ejemplos donde la molécula de ácido nucleico diana es un miRNA (o siRNA), la longitud de la secuencia (102) que se une específicamente a la secuencia de ácidos nucleicos diana puede ser más corta, por ejemplo de 20-30 nucleótidos de longitud (por ejemplo, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 ó 30 nucleótidos) para corresponderse con la longitud de miRNA (o siRNA). Sin embargo, un experto en la materia observará que no es necesario que la secuencia (102) que se une específicamente a la diana sea complementaria al 100% con la molécula de ácido nucleico diana. Dependiendo de las condiciones de reacción y de la selectividad correspondiente de la nucleasa usada, puede necesitarse más de un fallo de correspondencia (por ejemplo, al menos dos fallos de correspondencia adyacentes) para que tenga lugar la digestión de nucleasa. En algunos ejemplos, la NPPF está degenerada en una o más posiciones (por ejemplo, 1, 2, 3, 4, 5 o más posiciones), por ejemplo, una mezcla de 15 nucleótidos (por ejemplo, 2, 3 ó 4 nucleótidos) en una posición específica en la secuencia (104) que se une específicamente a la diana.

La secuencia (102) también se une específicamente a un conector de programación o bifuncional (donde una región del conector bifuncional es complementaria a la secuencia (102)). En algunas realizaciones, después de la hibridación y el tratamiento con nucleasa, la muestra se pone en contacto con una superficie (por ejemplo, una que incluye múltiples regiones espacialmente discretas), que incluye al menos molécula de captura, tal como un anclaje asociado con un conector de programación o una sonda de captura de ácidos nucleicos que incluye una secuencia complementaria a una parte del amplicón de NPPF (por ejemplo, una secuencia flanqueante o parte de la misma). Tal como se muestra en la FIG. 3, el conector bifuncional (216) incluye una primera parte que se une específicamente (por ejemplo, es complementaria) al anclaje (214) y una segunda parte que se une específicamente (por ejemplo, es complementaria) a una región del amplicón de NPPF (210). En algunos ejemplos, el amplicón de NPPF tiene al menos 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 ó 50 nucleótidos contiguos complementarios al conector bifuncional.

La secuencia de la secuencia flanqueante (104, 106) puede proporcionar un punto de amplificación universal que es complementaria a al menos una parte de un cebador de amplificación. La secuencia flanqueante permite así la multiplexación, ya que pueden usarse los mismos cebadores de amplificación para amplificar NPPF específicas para diferentes moléculas de ácidos nucleicos diana. La secuencia flanqueante no es similar a una secuencia encontrada en el genoma diana. Por ejemplo, si el ácido nucleico diana es una secuencia humana, la secuencia de la secuencia flanqueante no es similar a una secuencia encontrada en el genoma diana. Esto ayuda a reducir la unión de secuencias no diana que pueden estar presentes en el genoma diana a partir de la unión a las NPPF. Los procedimientos de análisis de una secuencia en cuanto a su semejanza con un genoma son bien conocidos en la técnica.

La secuencia flanqueante (104, 106) puede usarse también para permitir la captura de un amplicón de NPPF, por ejemplo captura en un sustrato. Por ejemplo, una NPPF que contiene una secuencia flanqueante que incluye una secuencia complementaria a una sonda de captura de ácidos nucleicos presente en una superficie (por ejemplo, conjugada directamente con una superficie), puede hibridarse con la sonda de captura de ácidos nucleicos que permite la captura o unión del amplicón de NPPF a la superficie. Así, en algunos ejemplos, la secuencia flanqueante incluye (o permite la adición, por ejemplo, durante la amplificación) de una etiqueta experimental, por ejemplo una que permite la captura del amplicón de NPPF. Se observará que pueden usarse otras etiquetas experimentales, tales como las usadas para identificar de forma única una NPPF o poblaciones de NPPF, y que dichas etiquetas experimentales pueden formar parte de la NPPF, o añadirse posteriormente, por ejemplo usando un cebador complementario a la secuencia flanqueante y que incluye también una secuencia complementaria a la etiqueta para añadirse al amplicón resultante. Las secuencias flanqueantes permiten también el etiquetado de la NPPF, por ejemplo durante la amplificación de la NPPF, o usando una sonda etiquetada que es complementaria a la secuencia flanqueante, y permitiendo que la sonda se una a la NPPF. En algunos ejemplos, la secuencia flanqueante incluye (o permite la adición, por ejemplo durante amplificación) de un adaptador de secuenciación, tal como una secuencia 55 poli-A o poli-T necesaria para algunas plataformas de secuenciación.

Se observará que una NPPF puede incluir una o dos secuencias flanqueantes (por ejemplo, una en el extremo 5', una en el extremo 3', o ambas cosas), y que las secuencias flanqueantes pueden ser la misma o diferente. Tal como se ilustra en las FIG. 6A y B, la NPPF puede incluir una única secuencia flanqueante. Las FIG. 6A y 6B muestran la

secuencia flanqueante en el extremo 5', pero se observará que también puede estar en su lugar en el extremo 3'. La FIG. 6A muestra un ejemplo donde todas las NPPF en la reacción tienen la misma secuencia flanqueante F1. Podría usarse la amplificación con un cebador específico de F1 (por ejemplo, un cebador etiquetado) para añadir la misma etiqueta en 5' o 3' (por ejemplo, adaptador de secuenciación o etiqueta experimental) a cada NPPF. Por ejemplo, 5 podría añadirse el mismo adaptador de secuenciación a todas las NPPF, lo que permite la secuenciación de las NPPF en la misma plataforma de secuenciación. La FIG. 6B muestra un ejemplo donde cada NPPF (o cada subpoblación de NPPF) en la reacción tiene una secuencia flanqueante diferente, de F1 a F3. Por ejemplo, F1, F2 y F3 podrían ser complementarias a una sonda de ácidos nucleicos de captura 1, 2 y 3, respectivamente, en una superficie. Así se elimina la necesidad de conectores bifuncionales (por ejemplo, véase parte inferior de la FIG. 3). 10 En otro ejemplo, puede usarse la amplificación con cebadores específicos de T1-F1, T2-F2 y T3-F3 para añadir una etiqueta experimental diferente a cada NPPF diferente (o poblaciones de NPPF).

Tal como se ilustra en las FIG. 6C-6F, la NPPF puede incluir en algunos ejemplos dos secuencias flanqueantes, una en el extremo 5' y la otra en el extremo 3' de la NPPF. La FIG. 6C muestra un ejemplo donde todas las NPPF en la 15 reacción tienen la misma secuencia flanqueante, F1, en los dos extremos. La FIG. 6D muestra un ejemplo donde todas las secuencias flanqueantes en el extremo 5' son iguales (por ejemplo, F1), y todas las secuencias flanqueantes en el extremo 3' son iguales (por ejemplo, F(a)), pero las secuencias flanqueantes en el extremo 5' y el extremo 3' son diferentes. En dicho ejemplo, esto permite la inclusión por ejemplo de la misma etiqueta experimental en un extremo de las NPPF, y la inclusión por ejemplo del mismo adaptador de secuenciación en el otro lado de las 20 NPPF. Siempre que no exista sesgo de hibridación de cebador cada NPPF debe etiquetarse con la misma fidelidad. La FIG. 6E muestra un ejemplo donde todas las secuencias flanqueantes en un extremo son iguales (por ejemplo, F1 en el extremo 5'), pero todas las secuencias flanqueantes en el otro extremo difieren entre sí (por ejemplo, F(a), F(b) y F(c)). En dicho ejemplo, esto permite el uso de una única sonda de captura para capturar todas las NPPF (por ejemplo, usando una sonda de captura que tiene al menos una parte de su secuencia complementaria a F1). 25 Podrían usarse las secuencias flanqueantes en el otro extremo, F(a), F(b) y F(c), por ejemplo para etiquetar de forma diferencial cada NPPF (por ejemplo, usando diferentes etiquetas experimentales). Alternativamente, F(a), F(b) y F(c) podrían ser complementarias para capturar las sondas 1, 2 y 3, respectivamente, y F1 podría usarse para etiquetar todas las NPPF de la misma forma. La FIG. 6F muestra un ejemplo donde todas las secuencias flanqueantes son diferentes, con independencia de su posición (por ejemplo, F(a), F(b), F(c), F1, F2 y F3). En este 30 ejemplo, cada secuencia flanqueante puede usarse para una etiqueta experimental diferente o para combinaciones de etiquetas experimentales diferentes y adaptadores de secuenciación diferentes.

Así, una secuencia de NPPF puede representarse por 1-2-3 donde 1 y 3 son secuencias flanqueantes a cada lado de la secuencia 2 (que es complementaria al ácido nucleico diana). Cada una de estas regiones puede hibridarse en 35 algún punto del procedimiento con su secuencia complementaria. Por ejemplo, A puede ser complementaria a la secuencia flanqueante 1 de la NPPF (por ejemplo, A puede ser una CFS complementaria a la secuencia 1), B puede ser complementaria a la secuencia 2 de la NPPF (por ejemplo, una secuencia diana complementaria a la secuencia 2) y C puede ser complementaria a la secuencia flanqueante 3 de la NPPF (por ejemplo, C puede ser una CFS complementaria a la secuencia 3). Esto es lo que sucede durante la hibridación de las moléculas de ácidos nucleicos 40 diana y CFS, con su NPPF correspondiente. Por ejemplo:

1-2-3
A-B-C

45 En algunos ejemplos, las etiquetas experimentales (por ejemplo, las que distinguen experimentos o pacientes entre sí) y adaptadores de secuenciación, representada por D y E respectivamente, se añaden usando las secuencias flanqueantes, por ejemplo durante la amplificación (de manera que el cebador de amplificación es complementario a la secuencia flanqueante e incluye una secuencia complementaria a la etiqueta o adaptador para añadirse al amplicón de NPPF resultante). Por ejemplo, la amplificación de la NPPF con dichos cebadores daría lugar a una 50 secuencia como la siguiente: E-1-2-3-D o D-1-2-3-E.

La tabla mostrada a continuación ilustra también cinco combinaciones de ejemplo de etiquetas en 5' (por ejemplo, etiquetas experimentales o adaptadores de secuenciación), secuencias flanqueantes en 5', secuencias diana específicas, secuencias flanqueantes en 3' y etiquetas en 3'. Las etiquetas en 5' y las etiquetas en 3' se añaden 55 durante la amplificación. Las secuencias flanqueantes en 5' y las secuencias flanqueantes en 3' son secuencias que forman parte de la NPPF original (y, así, parte de la secuencia flanqueante en sí).

	Etiqueta en 5'	Secuencia flanqueante en 5'	Secuencia específica de diana	Secuencia flanqueante en 3'	Etiqueta en 3'
Ej. 1	Ninguna	Adaptador de secuenciador		Adaptador de secuenciador	Ninguna
Ej. 2	Adaptador de secuenciación	Identificador específico de secuencia		Identificador específico de secuencia	Adaptador de secuenciación
Ej. 3	Etiqueta experimental (secuencia corta o bases modificadas, identificador para discernir independientemente una/varias reacciones: (p. ej.) por paciente, muestra, tipo de célula, punto de curso temporal, tratamiento)	Etiqueta experimental (secuencia corta o bases modificadas, identificador para discernir independientemente una/varias reacciones: (p. ej.) por paciente, muestra, tipo de célula, punto de curso temporal, tratamiento)		Etiqueta experimental (secuencia corta o bases modificadas, identificador para discernir independientemente una/varias reacciones: (p. ej.) por paciente, muestra, tipo de célula, punto de curso temporal, tratamiento)	Etiqueta experimental (secuencia corta o bases modificadas, identificador para discernir independientemente una/varias reacciones: (p. ej.) por paciente, muestra, tipo de célula, punto de curso temporal, tratamiento)
Ej. 4	Biotina u otra etiqueta/secuencia de captura de detección (por ejemplo, hapteno)	Biotina u otra etiqueta/secuencia de captura de detección (por ejemplo, hapteno)		Biotina u otra etiqueta/secuencia de captura de detección (por ejemplo, hapteno)	Biotina u otra etiqueta/secuencia de captura de detección (por ejemplo, hapteno)
Ej. 5	Sitio para escisión (base enzimática/modificada)	Sitio para escisión (base enzimática/modificada)		Sitio para escisión (base enzimática/modificada)	Sitio para escisión (base enzimática/modificada)
		Secuencia "tampón" (por ejemplo, espaciador o universal)		Secuencia "tampón" (por ejemplo, espaciador o universal)	

En ejemplos específicos, cada secuencia flanqueante no se une específicamente a cualquier otra secuencia de NPPF (por ejemplo, secuencia (102) u otra secuencia flanqueante) o a cualquier componente de la muestra. En algunos ejemplos, si existen dos secuencias flanqueantes, la secuencia de cada secuencia flanqueante (104, 106) es diferente. Idealmente, si existen dos secuencias flanqueantes diferentes (por ejemplo, dos secuencias flanqueantes diferentes en la misma NPPF y/o a secuencias flanqueantes de otras NPPF en un conjunto de NPPF), cada secuencia flanqueante (104, 106) tiene una temperatura de fusión (T_f) similar, por ejemplo una T_f +/- aproximadamente 10°C o +/- 5°C con respecto a la otra, por ejemplo +/- 4°C , 3°C , 2°C o 1°C .

10

En ejemplos particulares, la secuencia flanqueante (104, 106) tiene al menos 12 nucleótidos de longitud, por ejemplo al menos 15, al menos 20, al menos 25, al menos 30, al menos 40 o al menos 50 nucleótidos de longitud, tal como 12-50 o 12-30 nucleótidos, por ejemplo, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 ó 30 nucleótidos de longitud, donde deben someterse a ensayo los nucleótidos contiguos no presentes en una molécula de ácido nucleico presente en la muestra. Las secuencias flanqueantes están protegidas de la degradación por la nucleasa mediante hibridación con moléculas en las secuencias flanqueantes que tienen una secuencia complementaria a las secuencias flanqueantes (CFS).

Los factores que influyen en la especificidad de hibridación NPPF-diana y NPPF-CFS incluyen la longitud de la NPPF y la temperatura de fusión de la CFS, la autocomplementariedad y la presencia de secuencia repetitiva o no única. Véase, por ejemplo, Sambrook y col., Molecular Cloning: A Laboratory Manual, 3ª ed., Cold Spring Harbor Press, 2001; Ausubel y col., Current Protocols in Molecular Biology, Greene Publishing Associates, 1992 (y

Supplements a 2000); Ausubel y col., Short Protocols in Molecular Biology: A Compendium de Methods from Current Protocols in Molecular Biology, 4ª ed., Wiley & Sons, 1999. Las condiciones que producen grados de hibridación (astringencia) determinados variarán dependiendo de la naturaleza del procedimiento de hibridación y de la composición y longitud de las secuencias de ácidos nucleicos de hibridación. En general, la temperatura de hibridación y la fuerza iónica (por ejemplo, la concentración de Na⁺) del tampón de hibridación determinarán la astringencia de hibridación. En algunos ejemplos, las NPPF usadas en los procedimientos divulgados tienen una T_f de al menos aproximadamente 37°C, al menos aproximadamente 42°C, al menos aproximadamente 45°C, al menos aproximadamente 50°C, al menos aproximadamente 55°C, al menos aproximadamente 60°C, al menos aproximadamente 65°C, al menos aproximadamente 70°C, al menos aproximadamente 75°C, al menos aproximadamente 80°C, tal como aproximadamente 42°C-80°C (por ejemplo, aproximadamente 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79 ó 80°C). En un ejemplo no limitativo, las NPPF usadas en los procedimientos divulgados tienen una T_f de aproximadamente 42°C. Los procedimientos de cálculo de la T_f de una sonda son conocidos para los expertos en la materia (véase por ejemplo, Sambrook y col., Molecular Cloning: A Laboratory Manual, 3ª ed., Cold Spring Harbor Press, 2001, Capítulo 10). En algunos ejemplos, las NPPF para una reacción en concreto se seleccionan de manera que cada una tenga la misma T_f o una similar con el fin de facilitar la detección o secuenciación simultánea de múltiples moléculas de ácidos nucleicos diana en una muestra, tal como T_f +/- aproximadamente 10°C entre sí, tal como +/- 10°C, 9°C, 8°C, 7°C, 6°C, 5°C, 4°C, 3°C, 2°C ó 1°C entre sí.

20 A. Secuencias flanqueantes

Una o las dos secuencias flanqueantes de la NPP (por ejemplo, (104 ó 106) de la FIG. 1) incluyen una secuencia que proporciona un punto de amplificación universal. Dicha secuencia es complementaria con al menos una parte de un cebador de amplificación. Esto permite que el cebador se hibride con la NPPF, y amplificar la NPPF. Dado que las secuencias flanqueantes pueden ser idénticas entre NPPF específicas para diferentes moléculas de ácidos nucleicos diana, esto permite usar el mismo cebador para amplificar cualquier número de NPPF diferentes. Por ejemplo, una NPPF puede incluir una secuencia flanqueante en 5' y una secuencia flanqueante en 3', donde las secuencias flanqueantes en 5' y en 3' son diferentes entre sí, pero son iguales para una pluralidad de NPPF para dianas diferentes. Así puede usarse un cebador de amplificación que incluye una secuencia complementaria a la secuencia flanqueante en 5' y un cebador de amplificación que incluye una secuencia complementaria a la secuencia flanqueante en 3', en una única reacción para amplificar múltiples NPPF, incluso si las NPPF son específicas para diferentes secuencias diana.

En algunos ejemplos, la secuencia flanqueante no incluye una secuencia de etiquetas experimentales y/o una secuencia de adaptadores de secuenciación. En algunos ejemplos, una secuencia flanqueante incluye o consiste en una secuencia de etiquetas experimentales y/o secuencia de adaptadores de secuenciación. En otros ejemplos, los cebadores usados para amplificar las NPPF incluyen una secuencia de etiquetas experimentales y/o una secuencia de adaptadores de secuenciación, permitiendo así la incorporación de la etiqueta experimental y/o el adaptador de secuenciación en el amplicón de NPPF durante amplificación de la NPPF.

En un ejemplo, una secuencia flanqueante se diseña de manera que la secuencia forma un bucle en sí mismo. Así, una región de una secuencia flanqueante es complementaria a una segunda región de la misma secuencia flanqueante, de manera que las regiones primera y segunda se hibridan entre sí, formando un bucle o una horquilla. Así se eliminaría la necesidad de CFS, ya que la segunda región protegería a la primera región durante la etapa de la nucleasa.

B. Cebadores que se unen a las secuencias flanqueantes

Los cebadores de amplificación que se unen o se hibridan específicamente con las secuencias flanqueantes pueden usarse para iniciar la amplificación, por ejemplo amplificación por PCR. Además, los cebadores de amplificación pueden usarse para introducir etiquetas de ácidos nucleicos (por ejemplo, etiquetas experimentales o adaptadores de secuenciación) y/o etiquetas detectables en las NPPF. Por ejemplo, además del cebador de amplificación que tiene una región complementaria a la secuencia flanqueante, también puede incluir una segunda región que tiene una secuencia de ácidos nucleicos que procede además de una etiqueta experimental, un adaptador de secuenciación, una etiqueta detectable o combinaciones de los mismos, con el amplicón de NPPF resultante. Puede introducirse una etiqueta experimental o un adaptador de secuenciación en la el extremo 5' y/o 3' de la NPPF. En algunos ejemplos, se añaden dos o más etiquetas experimentales y/o adaptadores de secuenciación a un único extremo o a los dos extremos del amplicón de NPPF, por ejemplo usando un único cebador que tiene una secuencia de ácidos nucleicos que procede además de dos o más etiquetas experimentales y/o adaptadores de secuenciación.

Las etiquetas experimentales pueden usarse, por ejemplo, para diferenciar una muestra o secuencia de otra, o para permitir la captura de un amplicón de NPPF por un sustrato. Las etiquetas de secuencia permiten la captura del amplicón de NPPF resultante por una plataforma de secuenciación en particular.

5 Una etiqueta detectable puede introducirse en cualquier punto de la NPPF, lo que incluye el extremo 5' y/o 3'. En un ejemplo, la etiqueta se introduce en un amplicón de NPPF por hibridación de una sonda complementaria etiquetada con el amplicón de NPPF. En un ejemplo, la etiqueta se introduce en un amplicón de NPPF mediante el uso de un cebador etiquetado durante la amplificación de la NPPF, generando así un amplicón de NPPF etiquetado. Las etiquetas detectables permiten la detección de los amplicones de NPPF.

10

En algunos ejemplos, dichos cebadores tienen al menos 12 nucleótidos de longitud, por ejemplo al menos 15, al menos 20, al menos 30, al menos 40 o al menos 50 nucleótidos (por ejemplo, 25 nucleótidos). En algunos ejemplos los cebadores incluyen una etiqueta detectable (y dichos cebadores pueden referirse como sondas), por ejemplo biotina, que se incorporan en los amplicones de NPPF.

15

C. Adición de etiquetas experimentales

Las etiquetas experimentales pueden formar parte de la NPPF cuando se generan (por ejemplo, formar parte de la secuencia flanqueante). En otro ejemplo, la etiqueta experimental se añade después, por ejemplo durante la amplificación de la NPPF, produciendo un amplicón de NPPF que contiene una etiqueta experimental. La presencia de las secuencias flanqueantes universales en la NPPF permiten el uso de cebadores universales, que pueden introducir otras secuencias en las NPPF, por ejemplo durante la amplificación.

Las etiquetas experimentales, por ejemplo una que diferencia una muestra de otra, pueden usarse para identificar la secuencia diana en particular asociada con la NPPF, o permitir la captura de un amplicón de NPPF por un sustrato (donde la etiqueta experimental es complementaria a una sonda de captura en el sustrato, lo que permite la hibridación entre los dos). En un ejemplo, la etiqueta experimental contiene los primeros tres, cinco, diez, veinte o treinta nucleótidos del extremo 5' y/o 3' de la NPPF o amplicón de NPPF.

En un ejemplo una etiqueta experimental se usa para diferenciar una muestra de otra. Por ejemplo, dicha secuencia puede actuar como un código de barras, para permitir que se correlacione una secuencia detectada en particular con una muestra, paciente o experimento en particular (por ejemplo, un pocillo de reacción, un día o un conjunto de condiciones de reacción en particular). Por ejemplo, esto permite asociar una NPPF detectada o secuenciada en particular con un paciente o una muestra o un experimento en concreto. El uso de dichas etiquetas proporciona una forma de reducir el coste por muestra y aumentar el rendimiento de la muestra, ya que pueden etiquetarse múltiples amplicones de NPPF y después combinarse (por ejemplo, de diferentes experimentos o pacientes), por ejemplo en una única ejecución de secuenciación o matriz de detección. Se obtiene así la capacidad de combinar diferentes muestras experimentales o de paciente en una única ejecución, dentro del primer canal de instrumentos. Por ejemplo, dichas etiquetas permiten realizar centenares o miles de experimentos diferentes para su secuenciación en una única ejecución, dentro de un único canal. Por ejemplo, en un grupo de 100 muestras por canal, pueden someterse a ensayo 8.000 muestras en una única ejecución de un secuenciador de 8 canales. Además, si el procedimiento incluye la etapa de purificación en gel de la reacción de amplificación completa (u otro procedimiento de purificación o de limpieza que no requiera una separación real) sólo se necesita utilizar un gel (o reacción o proceso de limpieza o de purificación) para la realización de la detección o secuenciación. A continuación los amplicones de NPPF secuenciados se clasifican, por ejemplo, por medio de las etiquetas experimentales.

En un ejemplo la etiqueta experimental se usa para identificar la secuencia diana en particular asociada con la NPPF. En este caso, el uso de una etiqueta experimental para correspondencia con una secuencia diana en particular puede abreviar el tiempo o la cantidad de secuenciación necesarios, ya que puede ser suficiente la secuenciación del extremo de la NPPF en lugar de toda la NPPF. Por ejemplo, si dicha etiqueta experimental está presente en el extremo 3' del amplicón de NPPF, no es preciso secuenciar la secuencia completa del amplicón de NPPF en sí para identificar la secuencia diana que se hibrida con la NPPF. En su lugar, sólo es preciso secuenciar el extremo 3' del amplicón de NPPF que contiene la etiqueta experimental. Este hecho puede reducir de manera significativa el tiempo y los recursos de secuenciación, ya que se necesita secuenciar menos material.

55

En un ejemplo la etiqueta experimental se usa para permitir capturar las NPPF, por ejemplo, para concentrar NPPF o amplicones de NPPF a partir de una muestra. Por ejemplo, la etiqueta experimental puede tener una secuencia que es complementaria a la secuencia de al menos una parte de una sonda de captura en una superficie de sustrato, permitiendo así la hibridación de la NPPF con la sonda de captura. Por ejemplo, después de la amplificación, los

amplicones de NPPF que contienen una etiqueta experimental (por ejemplo, una población de amplicones de NPPF que contienen la misma etiqueta experimental) pueden aislarse a partir de otros materiales incubando la muestra con un sustrato (por ejemplo, perlas magnéticas) que contiene una pluralidad de sondas de captura con secuencias complementarias a la etiqueta experimental. Después de su captura, los amplicones de NPPF pueden ser
 5 detectados o secuenciados, o pueden ser liberados del sustrato para su análisis posterior. En un ejemplo, el sustrato son perlas magnéticas, la reacción de PCR que contiene amplicones de NPPF es incubada con las perlas. A continuación se mantienen las perlas en un campo magnético mientras se elimina la solución de la muestra (que contiene moléculas de ácidos nucleicos no deseadas y otros materiales). Las NPPF capturadas pueden ser sometidas a elución en un volumen menor invirtiendo la hibridación, por ejemplo, por adición de bases y
 10 calentamiento. Se observará que pueden usarse procedimientos similares con otras NPPF y otros sustratos (por ejemplo, usando un sustrato sólido y un flujo a través del dispositivo), para hacer que las NPPF capturadas sean sometidas a elución en un volumen menor. Si se añade un haptano durante la amplificación, puede usarse para la captura. Una ventaja de dicho procedimiento es que las NPPF o los amplicones de NPPF pueden aislarse de una muestra grande, por ejemplo 1 ml de plasma, y someterse a elución en un volumen menor usado para ensayos, por
 15 ejemplo, 20 µl.

También pueden usarse etiquetas experimentales para amplificación, tal como amplificación anidada o en amplificación en dos fases.

20 En ejemplos concretos, la etiqueta experimental tiene al menos 3 nucleótidos de longitud, por ejemplo al menos 5, al menos 10, al menos 15, al menos 20, al menos 25, al menos 30, al menos 40 o al menos 50 nucleótidos de longitud, tal como 3-50, 3-20, 12-50 o 12-30 nucleótidos, por ejemplo, 3, 5, 10, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 ó 30 nucleótidos de longitud.

25 **D. Adición de adaptadores de secuenciación**

Los adaptadores de secuenciación pueden formar parte de la NPPF cuando se generan (por ejemplo, formar parte de la secuencia flanqueante). En otro ejemplo, el adaptador de secuenciación se añade posteriormente, por ejemplo durante amplificación de la NPPF, para producir un amplicón de NPPF que contiene un adaptador de secuenciación.

30 La presencia de las secuencias flanqueantes universales en la NPPF permite el uso de cebadores universales, que pueden introducir otras secuencias en las NPPF, por ejemplo durante la amplificación.

Puede usarse un adaptador de secuenciación para añadir una secuencia a un amplicón de NPPF necesario para una plataforma de secuenciación determinada. Por ejemplo, algunas plataformas de secuenciación (por ejemplo, las
 35 plataformas 454 y Illumina) requieren que se secuencie la molécula de ácido nucleico para incluir una determinada secuencia en su extremo 5' y/o 3', por ejemplo para capturar la molécula que se va a secuenciar. Por ejemplo, el adaptador de secuenciación apropiado es reconocido por una secuencia complementaria en el chip o perlas de secuenciación, y la NPPF capturada por la presencia del adaptador de secuenciación.

40 En un ejemplo, se añade un poli-A (o poli-T), por ejemplo un poli-A o poli-de T al menos 10 nucleótidos de longitud, a la NPPF durante la amplificación por PCR. En un ejemplo específico, el poli-A (o poli-T) se añade al extremo 3' de la NPPF. En algunos ejemplos, esta secuencia añadida se somete a poliadenilación en su extremo 3' usando una desoxinucleotidiltransferasa terminal (Tdt).

45 En ejemplos concretos, la etiqueta de secuenciación añadida tiene al menos 12 nucleótidos (nt) de longitud, por ejemplo al menos 15, al menos 20, al menos 25, al menos 30, al menos 40 o al menos 50 nt de longitud, tal como 12-50 ó 12-30 nt, por ejemplo, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 ó 30 nt de longitud.

E. Etiquetas detectables

50

En algunos ejemplos, las NPPF, cebadores de PCR o ambos divulgados incluyen una o más etiquetas detectables. Las etiquetas detectables son bien conocidas en la técnica. Una etiqueta detectable es una molécula o material que puede usarse para producir una señal detectable que indica la presencia o concentración de una NPPF o amplicón de NPPF (por ejemplo, la sonda unida o hibridada) en una muestra. Así, una NPPF etiquetada proporciona un
 55 indicador de la presencia o concentración de una secuencia de ácidos nucleicos diana (por ejemplo, un ADN diana o un ARN diana) en una muestra. La descripción no se limita al uso de etiquetas en particular, aunque se proporcionan ejemplos.

En algunos ejemplos, la etiqueta se incorpora en la NPPF durante la síntesis de la NPPF. En algunos ejemplos, la

etiqueta se incorpora en la NPPF durante la amplificación, por ejemplo usando cebadores etiquetados (generando así amplicones de NPPF etiquetados). En otros ejemplos más, la NPPF se etiqueta usando una sonda etiquetada que es complementaria a, y así se hibrida con, una parte de la NPPF (por ejemplo, un amplicón de NPPF), tal como una región flanqueante de la NPPF.

5

En algunos ejemplos, cada una de las NPPF incluidas en una pluralidad de NPPF usadas en los procedimientos divulgados es etiquetada con la misma etiqueta detectable. En otros ejemplos al menos una NPPF se etiqueta con una etiqueta detectable diferente a la de al menos otra NPPF en la pluralidad de NPPF. Por ejemplo, al menos una NPPF incluida en la pluralidad de NPPF puede etiquetarse con un fluoróforo (por ejemplo, Cy-3™) y al menos una

10

NPPF incluida en la pluralidad de NPPF puede etiquetarse con un fluoróforo diferente (por ejemplo, Cy-5™). En algunos ejemplos, la pluralidad de NPPF puede incluir al menos 2, 3, 4, 5, 6 o más etiquetas detectables diferentes. Análogamente, los cebadores de amplificación usados en los procedimientos proporcionados en la presente memoria descriptiva pueden etiquetarse con la misma etiqueta detectable o con etiquetas detectables diferentes.

15

Una etiqueta asociada con una o más moléculas de ácidos nucleicos (por ejemplo, una NPPF o cebador de amplificación) puede detectarse directa o indirectamente. Una etiqueta puede detectarse por cualquier mecanismo conocido o aún por descubrir que incluye absorción, emisión y/o difusión de un fotón (que incluye fotones de radiofrecuencia, frecuencia de microondas, frecuencia infrarroja, frecuencia visible y frecuencia ultravioleta). Las etiquetas detectables incluyen moléculas y materiales coloreados, fluorescentes, electroluminiscentes,

20

fosforescentes y luminiscentes, catalizadores (por ejemplo, enzimas) que convierten una sustancia en otra sustancia para proporcionar una diferencia detectable (por ejemplo, convirtiendo una sustancia incolora en una coloreada o a la inversa, o produciendo un precipitado o aumentando la turbidez de la muestra), haptenos y moléculas o materiales magnéticos y paramagnéticos. Las etiquetas detectables adicionales incluyen etiquetas (de difusión luminosa) Raman (por ejemplo, bioetiquetas Nanoplex®, Oxonica, Bucks, RU). Otras etiquetas detectables de ejemplo incluyen

25

la digoxina, el uso de pares de transferencia de energía e inactivación energética (por ejemplo, FRET), IR y etiquetas de absorbancia/colorimétricas.

En ejemplos no limitativos, las NPPF o cebadores se etiquetan con dNTP unidos de forma covalente a moléculas de hapteno (por ejemplo, un compuesto nitroaromático (por ejemplo, dinitrofenilo (DNP)), biotina, fluoresceína, digoxigenina, etc.). Los procedimientos para conjugar haptenos y otras etiquetas con dNTP (por ejemplo, para facilitar la incorporación en sondas etiquetadas) son bien conocidos en la técnica. Como ejemplos de procedimientos, pueden verse las patentes de EE.UU. n° 5.258.507, 4.772.691, 5.328.824 y 4.711.955. Una etiqueta puede unirse de forma directa o indirectamente a un dNTP en cualquier posición en el dNTP, tal como un fosfato (por ejemplo, fosfato α , β o γ) o un azúcar. En algunos ejemplos, la detección de moléculas de ácidos nucleicos etiquetadas puede realizarse mediante la puesta en contacto de la NPP etiquetada con hapteno con un anticuerpo anti-hapteno primario. En un ejemplo, el anticuerpo anti-hapteno primario (por ejemplo, un anticuerpo anti-hapteno de ratón) se etiqueta directamente con una enzima. En otro ejemplo, se usa un anti-anticuerpo secundario (por ejemplo, un anticuerpo IgG antirratón de cabra) conjugado con una enzima para amplificación de la señal. En otros ejemplos, el hapteno es biotina y se detecta mediante la puesta en contacto de la NPPF etiquetada con hapteno con

30

avidina o estreptavidina conjugada con una enzima, tal como peroxidasa de rábano picante (HRP) o fosfatasa alcalina (AP).

35

Los ejemplos adicionales de etiquetas detectables incluyen moléculas fluorescentes (o fluorocromos). Los expertos en la materia conocen numerosos fluorocromos, y pueden seleccionarse, por ejemplo, en Life Technologies (antes

40

Invitrogen); por ejemplo, véase, *The Handbook - A Guide to Fluorescent Probes and Labeling Technologies*. Se proporcionan ejemplos de fluoróforos que pueden unirse (por ejemplo, conjugarse químicamente) con una molécula de ácido nucleico (por ejemplo, una NPPF) en la patente de EE.UU. n° 5.866.366 para Nazarenko y col., tal como 4-

45

ácido acetamido-4'-isotiocianatoestilben-2,2'-disulfónico, acridina y derivados tales como acridina e isotiocianato de acridina, ácido 5-(2'-aminoetil)aminonaftalen-1-sulfónico (EDANS), 4-amino-N-[3-vinilsulfonil]fenil]naftalimida-3,5 disulfonato (Lucifer Yellow VS), N-(4-anilino-1-naftil)maleimida, antranilamida, Amarillo Brillante, cumarina y derivados tales como cumarina, 7-amino-4- metilcumarina (AMC, Cumarina 120), 7-amino-4-trifluorometilcumarina (Cumarina 151); cianosina; 4',6-diaminidino-2-fenilindol (DAPI); 5',5"-dibromopirogalol-sulfonftaleína (rojo de bromopirogalol); 7-dietilamino-3-(4'-isotiocianatofenil)-4-metilcumarina; pentaacetato de dietilenetriamina; ácido 4,4'-diisotiocianatodihidro-estilbeno-2,2'-disulfónico; ácido 4,4'-diisotiocianatoestilben-2,2'-disulfónico; cloruro de 5-[dimetilamino]naftaleno-1-sulfonilo (DNS, cloruro de dansilo); ácido 4-(4'-dimetilaminofenilazo)benzoico (DABCIL); 4-dimetilaminofenilazofenil-4'-isotiocianato (DABITC); eosina y derivados tales como eosina y eosina isotiocianato; eritrosina y derivados tales como eritrosina B e isotiocianato de eritrosina; etidio; fluoresceína y derivados tales como 5-carboxifluoresceína (FAM), 5-(4,6-diclorotriazin-2-il)aminofluoresceína (DTAF), 2'7'-dimetoxi-4'5'-dicloro-6-carboxifluoresceína (JOE), fluoresceína, isotiocianato de fluoresceína (FITC) y QFITC (XRITC); 2',7'-

55

- difluorofluoresceína (OREGON GREEN®); fluorescamina; IR144; IR1446; isotiocianato de verde de malaquita; 4-metilumbelliferona; orto-cresoltaleína; nitrotirosina; pararosanilina; rojo de fenol; B-ficoeritrina; o-ftaldialdehído; pireno y derivados tales como pireno, butirato de pireno y butirato de succinimidil 1-pireno; rojo reactivo 4 (Rojo Brillante Cibacron Red 3B-A); rodamina y derivados tales como 6-carboxi-X-rodamina (ROX), 6-carboxirrodamina (R6G), cloruro de lisamina rodamina B sulfonilo, rodamina (Rhod), rodamina B, rodamina 123, rodamina X, isotiocianato de verde de rodamina, sulforrodamina B, sulforrodamina 101 y cloruro de sulfonilo derivado de sulforrodamina 101 (Texas Red); N,N,N',N'-tetrametil-6-carboxirrodamina (TAMRA); tetrametil-rodamina; isocianato de tetrametil-rodamina (TRITC); riboflavina; ácido rosólico y derivados de quelatos de terbio.
- 10 Otros fluoróforos adecuados incluyen quelatos de europio reactivos con tiol que emiten a aproximadamente 617 nm (Heyduk y Heyduk, *Analyt. Biochem.* 248:216-27, 1997; *J. Biol. Chem.* 274:3315-22, 1999), así como GFP, Lissamine™, dietilaminocumarina, fluoresceína clorotriacínilo, naftofluoresceína, 4,7-diclororrodamina y xanteno (tal como se describe en la patente de EE.UU. nº 5.800.996 para Lee y col.) y derivados de los mismos. Pueden usarse también otros fluoróforos conocidos para los expertos en la materia, por ejemplo los disponibles en Life Technologies (Invitrogen; Molecular Probes (Eugene, OR)) y que incluyen la serie de colorantes ALEXA FLUOR® (por ejemplo, tal como se describe en las patentes de EE.UU. nº 5.696.157, 6.130.101 y 6.716.979), la serie de colorante BODIPY (colorantes de difluoruro de dipirrometeno-boro, por ejemplo tal como se describe en las patentes de EE.UU. nº 4.774.339, 5.187.288, 5.248.782, 5.274.113, 5.338.854, 5.451.663 y 5.433.896), Azul Cascada (un derivado de reactivo de amina del pireno sulfonatado descrito en la patente de EE.UU. nº 5.132.432) y Azul de Marina (patente de EE.UU. nº 5.830.912).

Además de los fluorocromos descritos anteriormente, una etiqueta fluorescente puede ser una nanopartícula fluorescente, tal como un nanocrystal de semiconductores, por ejemplo, a QUANTUM DOT™ (obtenido, por ejemplo, en Life Technologies (QuantumDot Corp, Invitrogen Nanocrystal Technologies, Eugene, OR); véanse también, las patentes de EE.UU. nº 6.815.064; 6.682.596; y 6.649.138). Los nanocristales de semiconductores son partículas microscópicas que tienen propiedades ópticas y/o eléctricas dependientes del tamaño. Cuando los nanocristales de semiconductores se iluminan con una fuente de energía primaria, tiene lugar una emisión de energía secundaria de una frecuencia que corresponde a la banda prohibida del material semiconductor usado en el nanocrystal de semiconductores. Esta emisión puede detectarse como luz coloreada de una longitud de onda o fluorescencia específica. Los nanocristales de semiconductores con diferentes características espectrales se describen, por ejemplo, en la patente de EE.UU. nº 6.602.671. Los nanocristales de semiconductores pueden acoplarse a diversas moléculas biológicas (que incluyen dNTP y/o ácidos nucleicos) o sustratos por técnicas descritas, por ejemplo, en Bruchez y col., *Science* 281:2013-2016, 1998; Chan y col., *Science* 281:2016-2018, 1998; y en la patente de EE.UU. nº 6.274.323.

35 La formación de nanocristales de semiconductores de diversas composiciones se divulga, por ejemplo, en las patentes de EE.UU. nº 6.927.069; 6.914.256; 6.855.202; 6.709.929; 6.689.338; 6.500.622; 6.306.736; 6.225.198; 6.207.392; 6.114.038; 6.048.616; 5.990.479; 5.690.807; 5.571.018; 5.505.928; 5.262.357 y en la publicación de patente de EE.UU. nº 2003/0165951 así como en la publicación PCT nº 99/26.299 (publicada el 27 de mayo de 1999). Pueden producirse poblaciones separadas de nanocristales de semiconductores que son identificables basándose en sus diferentes características espectrales. Por ejemplo, pueden producirse nanocristales de semiconductores que emiten luz de diferentes colores basándose en su composición, tamaño o tamaño y composición. Por ejemplo, en la presente memoria descriptiva se divulgan puntos cuánticos que emiten luz a diferentes longitudes de onda basándose en el tamaño (longitudes de onda de emisión de 565 nm, 655 nm, 705 nm o 800 nm), que son adecuados como etiquetas fluorescentes en las sondas, disponibles en Life Technologies (Carlsbad, CA).

Las etiquetas adicionales incluyen, por ejemplo, radioisótopos (por ejemplo, ³H), quelatos metálicos tales como quelatos DOTA y DPTA de iones metálicos radioactivos o paramagnéticos como Gd³⁺, y liposomas.

50 Las etiquetas detectables que pueden usarse con moléculas de ácidos nucleicos (por ejemplo, una NPPF o cebador de amplificación) incluyen también enzimas, por ejemplo HRP, AP, fosfatasa ácida, glucosa oxidasa, β-galactosidasa, β-glucuronidasa o β-lactamasa. Cuando la etiqueta detectable incluye una enzima, puede usarse un compuesto cromógeno, fluorógeno o luminógeno en combinación con la enzima para generar una señal detectable (muchos de dichos compuestos están disponibles comercialmente, por ejemplo, en Life Technologies, Carlsbad, CA). Los ejemplos concretos de compuestos cromógenos incluyen diaminobencidina (DAB), 4-nitrofenilfosfato (pNPP), rojo rápido, azul rápido, fosfato de bromocloroindolilo (BCIP), nitro azul tetrazolio (NBT), BCIP/NBT, AP Orange, azul AP, tetrametilbencidina (TMB), sulfonato de 2,2'-acino-di-[3-etilbenzotiazolina] (ABTS) o-dianisidina, 4-cloronaftol (4-CN), nitrofenil-β-D-galactopiranosido (ONPG), o-fenilendiamina (OPD), 5-bromo-4-cloro-3-indolil-β-

galactopiranosido (X-Gal), metilumbelliferil- β -D-galactopiranosido (MU-Gal), p-nitrofenil- α -D-galactopiranosido (PNP), 5-bromo-4-cloro-3-indolil- β -D-glucuronido (X-Gluc), 3-amino-9-etil-carbazol (AEC), fucsina, yodonitrotetrazolio (INT), azul de tetrazolio y violeta de tetrazolio.

5 Alternativamente, puede usarse una enzima en un esquema de detección metalográfica. Los procedimientos de detección metalográfica incluyen el uso de una enzima, tal como fosfatasa alcalina, en combinación con un ion metálico hidrosoluble y un sustrato activo redox de la enzima. El sustrato se convierte en agente activo redox mediante la enzima, y el agente activo redox reduce el ion metálico, para formar un precipitado detectable. (véase, por ejemplo, publicación de solicitud de patente de EE.UU. nº 2005/0.100.976, publicación PCT nº 2005/003.777 y
10 publicación de solicitud de patente de EE.UU. nº 2004/0.265.922). Los procedimientos de detección metalográficos incluyen también el uso de una enzima oxido-reductasa (por ejemplo, peroxidasa de rábano picante) junto con un ion metálico soluble en agua, un agente oxidante y un agente reductor, de nuevo para formar un precipitado detectable. (véase, por ejemplo, patente de EE.UU. nº 6.670.113).

15 En algunas realizaciones, la etiqueta detectable está unida a o incorporada en la NPPF o cebador en el extremo 5' o el extremo 3' (por ejemplo, la NPPF o cebador es una sonda etiquetada en el extremo). En otros ejemplos, la etiqueta detectable se incorpora en la NPPF o cebador en una posición interna, tal como 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 o más bases desde el extremo 5' de la NPPF o cebador, o 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 o más bases desde el extremo 3' de la NPPF o cebador.

20 En un ejemplo, una de las regiones flanqueantes de la NPPF contiene un aceptor o emisor (por ejemplo, un fluoróforo aceptor), mientras que el cebador de amplificación complementaria a la región flanqueante contiene el inverso (por ejemplo, un fluoróforo donador). Así el dúplex cebador-NPPF emite la señal detectable, pero los cebadores monocatenarios, o las NPPF monocatenarias, no lo hacen. El aspecto de la señal es una medida de la
25 cantidad de NPPF en la muestra, y puede medirse sin separación de los cebadores etiquetados en exceso de los aductos amplificados. Los ejemplos de pares aceptor-donador FRET son conocidos en la técnica y pueden incluir FAM como fluoróforo donador para su uso con JOE, TAMRA y ROX, 3-(ϵ -carboxipentil)-3'-etil-5,5'-dimetiloxacarbocianina (CYA) puede servir como fluoróforo donador para derivados de la rodamina (por ejemplo, R6G, TAMRA, y ROX) que pueden usarse como fluoróforos aceptores. Grant y col. (Biosens Bioelectron. 16:231-7,
30 2001) proporcionan ejemplos concretos de pares FRET que pueden usarse en los procedimientos divulgados en la presente memoria descriptiva.

V. Muestras

35 Una muestra es cualquier espécimen que comprende una o más dianas, tal como una muestra biológica o un espécimen biológico. La muestra puede ser recogida u obtenida usando procedimientos bien conocidos para los expertos en la materia. Las muestras de uso en los procedimientos divulgados pueden incluir cualquier espécimen que incluye ácido nucleico (por ejemplo, ADN genómico, cDNA, ADN o ARN vírico, rRNA, tRNA, mRNA, miRNA, oligonucleótidos, fragmentos de ácidos nucleicos, ácidos nucleicos modificados, ácidos nucleicos sintéticos, o
40 similares). En un ejemplo, la muestra incluye ARN inestable. En algunos ejemplos, la molécula de ácido nucleico para detección o secuenciación está reticulada en la muestra (por ejemplo, un ADN, mRNA, miRNA o vRNA reticulado) o es soluble en la muestra. En algunos ejemplos, la muestra es una muestra fija, tal como una muestra que incluye un agente que provoca la reticulación de la molécula diana. En algunos ejemplos, el ácido nucleico diana en la muestra no es extraído, solubilizado, o ambos, antes de detectar o secuenciar la molécula de ácido nucleico
45 diana.

En algunos ejemplos, los procedimientos divulgados incluyen la obtención de la muestra antes del análisis de la muestra. En algunos ejemplos, los procedimientos divulgados incluyen la selección de un sujeto que tiene un tumor, y en algunos ejemplos la selección adicional de uno o más ADN o ARN diana para detectar basándose en el tumor
50 del sujeto, por ejemplo, para determinar un diagnóstico o pronóstico para el sujeto o para selección de una o más terapias. En algunos ejemplos, las moléculas de ácidos nucleicos en una muestra para su análisis primero se aíslan, se extraen, se concentran, o combinaciones de lo anterior, a partir de la muestra.

En algunos ejemplos, el ARN en la muestra se somete a transcripción inversa antes de realizar los procedimientos proporcionados en la presente memoria descriptiva. Sin embargo, los procedimientos divulgados no requieren transcripción inversa, ya que la secuencia de ARN diana se convierte efectivamente en una secuencia de sonda complementaria a través de hibridación y actividad de la nucleasa. A veces es deseable secuenciar moléculas de ARN en lugar de secuencias de genes que codifican el ARN, dado que las moléculas de ARN no son necesariamente colineales con su plantilla de ADN. Y algunos organismos son ARN, por ejemplo, virus de ARN.

En algunos ejemplos, la muestra se somete a lisis. El tampón de lisis está diseñado para inactivar enzimas y prevenir la degradación de ARN, pero después de una dilución limitada en un tampón de dilución de hibridación permite actividad de nucleasa y facilita la hibridación con especificidad de astringencia. Puede añadirse un tampón de dilución para neutralizar la actividad inhibidora de la lisis y otros tampones, tal como la actividad inhibidora para otras enzimas (por ejemplo, polimerasa). Alternativamente, la composición del tampón de lisis y otros tampones puede modificarse a una composición que sea tolerada, por ejemplo, por una polimerasa.

En algunos ejemplos, los procedimientos incluyen el análisis de una pluralidad de muestras de forma simultánea o contemporánea. Por ejemplo, los procedimientos pueden analizar al menos dos muestras diferentes (por ejemplo, de pacientes diferentes) de forma simultánea o contemporánea. En un ejemplo, los procedimientos pueden detectar o secuenciar al menos dos moléculas de ácidos nucleicos diana diferentes (por ejemplo, 2, 3, 4, 5, 6, 7, 8, 9 ó 10 dianas diferentes) en al menos dos muestras diferentes (por ejemplo, al menos 5, al menos 10, al menos 100, al menos 500, al menos 1.000 o al menos 10.000 muestras diferentes) de forma simultánea o contemporánea.

Las muestras de ejemplo incluyen, pero no se limitan a, células, lisados celulares, frotis de sangre, preparaciones de citocentrífuga, frotis citológicos, fluidos corporales (por ejemplo, sangre y fracciones de la misma como, por ejemplo suero y plasma, saliva, esputo, orina, líquido raquídeo, fluidos gástricos, sudor, semen, etc.), frotis citológicos, células bucales, extractos de tejidos, células u órganos, biopsias de tejido (por ejemplo, biopsias de tumor), aspirados con aguja fina, biopsias de punciones, células tumorales circulantes, tejido fresco, tejido congelado, tejido fijo, tejido fijado e integrado con cera (por ejemplo, parafina), médula ósea y/o cortes de tejido (por ejemplo, cortes de tejido con criostato y/o cortes de tejido integrado con parafina). La muestra biológica puede ser también una muestra de investigación de laboratorio, tal como una muestra de cultivo celular o sobrenadante.

Los procedimientos de obtención de una muestra a partir de un sujeto son conocidos en la técnica. Por ejemplo, los procedimientos de obtención de muestras de tejidos o células son rutinarios. Pueden obtenerse muestras de ejemplo de células o tejidos normales, o de células o tejidos neoplásicos. La neoplasia es un estado biológico en el que una o más células han experimentado una anaplasia característica con pérdida de diferenciación, aumento en la tasa de crecimiento e invasión del tejido circundante, y dichas células pueden ser susceptibles de metástasis. En ejemplos concretos, una muestra biológica incluye una muestra tumoral, tal como una muestra que contiene células neoplásicas.

Las células o tejidos neoplásicos de ejemplo pueden estar incluidos en o ser aislados de tumores sólidos, que incluyen cáncer de pulmón (por ejemplo, cáncer de pulmón no microcítico, tal como carcinoma de pulmón de células escamosas), carcinomas de mama (por ejemplo, carcinomas lobulares y de los conductos), cáncer adrenocortical, ameloblastoma, cáncer ampular, cáncer de vejiga, cáncer de huesos, cáncer de cuello de útero, colangioma, cáncer colorrectal, cáncer de endometrio, cáncer de esófago, cáncer gástrico, glioma, tumor de células granulares, cáncer de cabeza y cuello, cáncer hepatocelular, lunar hidatiforme, linfoma, melanoma, mesotelioma, mieloma, neuroblastoma, cáncer oral, osteocondroma, osteosarcoma, cáncer de ovarios, cáncer de páncreas, pilomatricoma, nevo de Spitz, cáncer de células escamosas, cáncer teratoide y cáncer del tiroides. Las células neoplásicas de ejemplo también pueden estar incluidas en o ser aisladas de cánceres hematológicos que incluyen leucemias, lo que incluye leucemias agudas (por ejemplo, leucemia linfocítica aguda, leucemia mielocítica aguda, leucemia mielógena aguda y leucemia mieloblástica, promielocítica, mielomonocítica, monocítica y eritroleucemia), leucemias crónicas (por ejemplo, leucemia mielocítica (granulocítica) crónica, leucemia mielógena crónica y leucemia linfocítica crónica), policitemia vera, linfoma, enfermedad de Hodgkin, linfoma no hodgkiniano (formas indolente y de grado alto), mieloma múltiple, macroglobulinemia de Waldenström, enfermedad de la cadena pesada, síndrome mielodisplásico y mielodisplasia.

Por ejemplo, una muestra de un tumor que contiene material celular puede obtenerse por extirpación quirúrgica de parte o la totalidad del tumor, recogida de un aspirado con aguja fina del tumor, así como otros procedimientos conocidos en la técnica. En algunos ejemplos, se aplica una muestra de tejido o célula a un sustrato y se analiza para determinar la presencia de uno o más ADN o ARN diana. Un soporte sólido útil en un procedimiento divulgado sólo necesita llevar la muestra biológica y, opcionalmente, permitir la cómoda detección de los componentes (por ejemplo, proteínas y/o secuencias de ácidos nucleicos) en la muestra. Los soportes de ejemplo incluyen portaobjetos de microscopio (por ejemplo, portaobjetos de vidrio o portaobjetos de plástico), cubreobjetos (por ejemplo, cubreobjetos de vidrio o cubreobjetos de plástico), placas para cultivo de tejidos, placas de pocillos múltiples, membranas (por ejemplo, nitrocelulosa o polifluoruro de vinilideno (PVDF)) o chips BIACORE™.

Los procedimientos divulgados son sensibles y específicos y permiten la detección de moléculas de ácidos nucleicos diana en una muestra que contiene un número limitado de células. Las muestras que incluyen números bajos de células, por ejemplo menos de 250.000 células (por ejemplo, menos de 100.000, menos de 50.000, menos de 10.000, menos de 1.000, menos de 500, menos de 200, menos de 100 células o menos de 10 células), incluyen pero no se limitan a, muestras FFPE, aspirados con aguja fina (por ejemplo, de pulmón, próstata, linfa, mama o hígado), biopsias de punción, biopsias con aguja, pequeñas poblaciones de células clasificadas (por ejemplo, FACS) o células tumorales circulantes, aspirados pulmonares, pequeños números de células capturadas con láser o macrodisecionadas o células tumorales circulantes, exosomas y otras partículas subcelulares, o líquidos corporales (por ejemplo, plasma, suero, líquido raquídeo, saliva y aspirados de mama). Por ejemplo, un ADN diana o ARN diana puede detectarse en apenas 1.000 células (por ejemplo, una muestra que incluye 1.000 o más células, tal como 1.000, 2.000, 3.000, 4.000, 5.000, 6.000, 7.000, 8.000, 9.000, 10.000, 15.000, 20.000, 50.000 o más células). En algunos ejemplos, la expresión de un ADN diana o ARN diana puede detectarse en aproximadamente 1.000 a 100.000 células, por ejemplo aproximadamente de 1.000 a 50.000, de 1.000 a 15.000, de 1.000 a 10.000, de 1.000 a 5.000, de 3.000 a 50.000, de 6.000 a 30.000 o de 10.000 a 50.000 células). En algunos ejemplos, la expresión de un ADN diana o ARN diana puede detectarse en aproximadamente 100 a 250.000 células, por ejemplo de aproximadamente 100 a 100.000, 100 a 50.000, 100 a 10.000, 100 a 5.000, 100 a 500, 100 a 200 ó 100 a 150 células. En otros ejemplos, la expresión de un ADN diana o ARN diana puede detectarse en aproximadamente 1 a 1.000 células (por ejemplo, de aproximadamente 1 a 500 células, de aproximadamente 1 a 250 células, de aproximadamente 1 a 100 células, de aproximadamente 1 a 50 células, de aproximadamente 1 a 25 células o aproximadamente 1 célula).

Las muestras pueden tratarse en diversas formas conocidas para los expertos en la materia en la técnica antes de (o al mismo tiempo que) la puesta en contacto de la muestra con un reactivo específico de la diana (por ejemplo, una NPPF). Un tratamiento relativamente sencillo es la suspensión de la muestra en un tampón, por ejemplo, un tampón de lisis, que conserva todos los componentes de la muestra en una sola solución. Muchos procedimientos tradicionales para detectar dianas requieren un procesamiento de muestra más complejo (por ejemplo, que implica varias etapas y/o varios tipos de instrumentos especializados) para hacer la diana accesible a uno o varios reactivos específicos de la diana. Por ejemplo, algunos procedimientos de detección requieren el aislamiento parcial o completo (por ejemplo, extracción) de una diana (por ejemplo, ADN o mRNA) a partir de la muestra. Una diana (como ADN o ARN) se ha aislado o extraído cuando se purifica desde otros componentes biológicos no diana de una muestra. Purificación se refiere a separación de la diana de uno o más componentes extraños presentes también en una muestra. Por ejemplo, antes de una detección basada en PCR de mRNA con cebadores específicos de diana en pares, a menudo se separa mRNA (que incluye el mRNA diana) soluble o total de las proteínas celulares y otros ácidos nucleicos en la muestra. Los componentes que son aislados, extraídos o purificados a partir de un espécimen o muestra mezclado normalmente están enriquecidos en al menos el 50%, al menos el 60%, al menos el 75%, al menos el 90% o al menos el 98% o incluso al menos el 99% en comparación con la muestra no purificada o no extraída.

El aislamiento de componentes biológicos de una muestra consume tiempo y conlleva el riesgo de pérdida del componente que se aísla, por ejemplo, por degradación y/o eficacia deficiente o incompletitud del proceso o los procesos usados para aislamiento. Por otra parte, con algunas muestras, como los tejidos fijos, las dianas (por ejemplo, ADN o ARN (por ejemplo, mRNA o miRNA)) son claramente difíciles de aislar con alta fidelidad (por ejemplo, en comparación con tejidos frescos o congelados) dado que, según se cree, al menos una proporción de las dianas está reticulada con otros componentes en la muestra fija y, por tanto, no pueden aislarse o solubilizarse con facilidad y pueden perderse al separar fracciones solubles e insolubles. En consecuencia, en algunos ejemplos, los procedimientos de detección de un ácido nucleico diana no requieren ni implican purificación, extracción o aislamiento de una diana a partir de una muestra antes de la puesta en contacto de la muestra con una o más NPPF, y/o implican sólo la puesta en suspensión de la muestra en una solución, por ejemplo, tampón de lisis, que conserva todos los componentes de la muestra antes de la puesta en contacto de la muestra con un reactivo específico de la diana.

En algunos ejemplos, las células en la muestra son lisadas o permeabilizadas en una solución acuosa (por ejemplo, usando un tampón de lisis). La solución acuosa o tampón de lisis incluye detergente (por ejemplo, dodecilsulfato de sodio) y uno o más agentes caotrópicos (por ejemplo, formamida, HCl de guanidinio, isotiocianato de guanidinio o urea). La solución puede contener también un tampón (por ejemplo, SSC). En algunos ejemplos, el tampón de lisis incluye aproximadamente del 15% al 25% de formamida (v/v), aproximadamente del 0,01% al 0,1% de SDS y aproximadamente 0,5-6X SSC (por ejemplo, aproximadamente 3X SSC). El tampón puede incluir opcionalmente tRNA (por ejemplo, de aproximadamente 0,001 a aproximadamente 2,0 mg/ml) o una ribonucleasa; ADNasa; proteinasa K; enzimas (por ejemplo, colagenasa o lipasa) que degradan proteínas, matriz, hidratos de carbono,

lípidos o una especie de oligonucleótidos, o combinaciones de los mismos. El tampón de lisis puede incluir también un indicador de pH, tal como rojo de fenol. En un ejemplo concreto, el tampón de lisis incluye formamida al 20%, 3X SSC (79,5%), SDS al 0,05%, 1 µg/ml de tRNA y 1 mg/ml de rojo de fenol. Las células se incuban en la solución acuosa (opcionalmente con aceite superpuesto para evitar la evaporación o para servir como un sumidero de la parafina) durante un periodo de tiempo suficiente (por ejemplo, entre aproximadamente 1 minuto y aproximadamente 60 minutos, por ejemplo entre aproximadamente 5 minutos y aproximadamente 20 minutos, o entre aproximadamente 10 minutos) y a una temperatura suficiente (por ejemplo, entre aproximadamente 22°C y aproximadamente 110°C, por ejemplo, entre aproximadamente 80°C y aproximadamente 105°C, entre aproximadamente 37°C y aproximadamente 105°C o entre aproximadamente 90°C y aproximadamente 100°C) para lisar o permeabilizar la célula. En algunos ejemplos, la lisis se realiza a aproximadamente 95°C. En algunos ejemplos, la etapa de lisis incluye la incubación de la muestra a aproximadamente 95°C durante aproximadamente 5-15 minutos para desnaturalizar el ARN en la muestra, pero no ADN genómico. En otros ejemplos, la etapa de lisis incluye la incubación de la muestra a aproximadamente 105°C durante aproximadamente 5-15 minutos para desnaturalizar el ARN y el ADN genómico en la muestra. En un ejemplo se incluye proteinasa K con el tampón de lisis.

En algunos ejemplos, la lisis celular en crudo se usa directamente sin purificación adicional. Las células pueden someterse a lisis en presencia o ausencia de una o más de las NPPF divulgadas. Si las células se someten a lisis en ausencia de sonda, la una o más sondas pueden añadirse posteriormente al lisado en crudo. En otros ejemplos, los ácidos nucleicos (por ejemplo, ADN y/o ARN) se aíslan a partir del lisado celular antes de la puesta en contacto del lisado con una o más NPPF.

En otros ejemplos, las muestras de tejido se preparan por fijado e integración del tejido en un medio o incluyen una suspensión de células como monocapa en un soporte sólido (por ejemplo, un portaobjetos de vidrio), por ejemplo mediante frotis o centrifugado de células en el soporte sólido. En ejemplos adicionales, puede usarse tejido o cortes de tejido frescos congelados (por ejemplo, no fijados) en los procedimientos divulgados en la presente memoria descriptiva. En ejemplos concretos, se usan cortes de tejido FFPE en los procedimientos divulgados.

En algunos ejemplos se usa un medio de integración. Un medio de integración es un material inerte en el que se integran los tejidos y/o células para ayudar a conservarlos para futuro análisis. La integración permite también cortar las muestras de tejido en secciones finas. Los medios de integración incluyen parafina, celoidina, compuesto OCT™, agar, plástico o acrílico. Muchos medios de integración son hidrófobos; por tanto, puede ser necesario eliminar el material inerte antes del análisis, que usa principalmente reactivos hidrófilos. El término desparafinación o desencerado se usa ampliamente en la presente memoria descriptiva para referirse a la eliminación parcial o completa de cualquier tipo de medio de integración de una muestra biológica. Por ejemplo, los cortes de tejido integrados en parafina son desencerados mediante el paso a través de disolventes orgánicos, por ejemplo, tolueno, xileno, limoneno u otros disolventes adecuados. En otros ejemplos, los cortes de tejido integrados en parafina se usan directamente (por ejemplo, sin etapa de desencerado).

Los tejidos pueden fijarse mediante cualquier procedimiento adecuado, que incluye perfusión o inmersión en un fijador. Los fijadores pueden clasificarse como agentes de reticulación (por ejemplo, aldehídos, por ejemplo, formaldehído, paraformaldehído y glutaraldehído, así como agentes de reticulación sin aldehído), agentes oxidantes (por ejemplo, iones y complejos metálicos, como tetróxido de osmio y ácido crómico), agentes de desnaturalización de proteínas (por ejemplo, acético ácido, metanol y etanol), fijadores de mecanismo desconocido (por ejemplo, cloruro mercúrico, acetona y ácido pícrico), reactivos de combinación (por ejemplo, fijador de Carnoy, metacarn, líquido de Bouin, fijador B5, líquido de Rossman y líquido de Gendre), microondas y fijadores misceláneos (por ejemplo, fijación por exclusión de volumen y fijación por vapor). También pueden incluirse aditivos en el fijador, tal como tampones, detergentes, ácido tánico, fenol, sales metálicas (por ejemplo, cloruro de cinc, sulfato de cinc y sales de litio) y lantano. El fijador usado más comúnmente en la preparación de muestras de tejidos o células es el formaldehído, generalmente en forma de una solución de formalina (formaldehído al 4% en una solución tampón, referido como formalina tamponada al 10%). En un ejemplo, el fijador es formalina tamponada neutra al 10%, y así en algunos ejemplos la muestra es formalina fijada.

En algunos ejemplos, la muestra es una muestra ambiental (por ejemplo, una muestra de suelo, aire o agua, o una muestra obtenida de una superficie (por ejemplo, por aspiración)), o una muestra alimentaria (por ejemplo, una verdura, fruta, producto lácteo o carne que contiene una muestra), por ejemplo para detectar los patógenos que pueden estar presentes.

VI. Ácidos nucleicos diana

Una molécula de ácido nucleico diana es una molécula de ácido nucleico que es susceptible de detección, o de interés, o útil para su detección con los procedimientos divulgados. Las dianas incluyen moléculas de ácidos nucleicos monocatenarias, bicatenarias u otras con múltiples cadenas (tal como ADN (por ejemplo, genómico, 5 mitocondrial o sintético), ARN (por ejemplo, mRNA, miRNA, tRNA, siRNA, ARN largo no codificante (nc), ARN antisentido de ocurrencia biológica, ARN de interacción Piwi (piRNA) o ARN nucleolares pequeños (snoRNA)), ya sean eucariotas, procariotas, virus, hongos, bacterias u otros organismos biológicos. Los ADN genómicos dianas pueden incluir una o varias partes del genoma, tales como regiones codificantes (por ejemplo, genes o exones), regiones no codificantes (tengan o no una función biológica desconocida, por ejemplo, estimuladores, promotores, 10 regiones reguladoras, telómeros o ADN "sin sentido"). En algunas realizaciones, una diana puede contener o ser el resultado de una mutación (por ejemplo, mutación somática o de línea germinal) que puede ser de ocurrencia natural o inducida por otros medios (por ejemplo, mutación inducida químicamente o por radiación). Dichas mutaciones pueden incluir (o ser resultado de) reorganizaciones genómicas (por ejemplo, translocaciones, inserciones, deleciones o inversiones), variaciones de nucleótido individuales y/o amplificaciones genómicas. En algunas 15 realizaciones, una diana puede contener una o más unidades de monómeros modificadas o sintéticas (por ejemplo, ácido nucleico peptídico (PNA), ácido nucleico bloqueado (LNA), ácido nucleico metilado, aminoácido modificado después de traducción, ácido nucleico reticulado o aminoácido reticulado).

La parte de una molécula de ácido nucleico diana a la que puede unirse específicamente una NPPF puede referirse 20 como "diana", de nuevo, en virtud del contexto, pero más específicamente puede referirse como parte diana, región complementaria (CR), sitio diana, región diana protegida o sitio protegido, o similar. Una NPPF unida específicamente a su región complementaria forma un complejo, que puede permanecer integrado con la diana como un conjunto y/o muestra, o separarse (o estar separado) de la diana como un conjunto y/o muestra. En algunas realizaciones, un complejo NPPF/CR se separa (o se desasocia) de la diana como un conjunto y/o muestra, 25 por ejemplo, por la acción de una nucleasa, tal como la nucleasa S1.

Todos los tipos de moléculas de ácidos nucleicos diana pueden analizarse usando los procedimientos divulgados. En un ejemplo, la diana es una molécula de ácido ribonucleico (ARN), tal como un ARN mensajero (mRNA), un ARN ribosómico (rRNA), un ARN de transferencia (tRNA), un micro-ARN (miRNA), un siRNA, un ARN antisentido o un 30 ARN vírico (vRNA). En otro ejemplo, la diana es una molécula de ácido desoxirribonucleico (ADN), tal como ADN genómico (gDNA), ADN mitocondrial (mtDNA), ADN de cloroplastos (cpDNA), ADN vírico (vDNA), cDNA o un ADN transfectado. En un ejemplo específico, la diana es un nucleótido antisentido. En algunos ejemplos, puede analizarse el transcriptoma completo de una célula o un tejido usando los procedimientos divulgados. En un ejemplo, la molécula de ácido nucleico diana es una molécula de ácido nucleico rara, por ejemplo sólo que aparece menos de 35 aproximadamente 100.000 veces, menos de aproximadamente 10.000 veces, menos de aproximadamente 5.000 veces, menos de aproximadamente 100 veces, menos de 10 veces o sólo una vez en la muestra, tal como una molécula de ácido nucleico que aparece sólo de 1 a 10.000, de 1 a 5.000, de 1 a 100 o de 1 a 10 veces en la muestra).

40 Es posible detectar o secuenciar una pluralidad de dianas en la misma muestra o ensayo, o incluso en múltiples muestras o ensayos, por ejemplo de forma simultánea o contemporánea. Análogamente, puede detectarse o secuenciarse una sola diana en una pluralidad de muestras, por ejemplo de forma simultánea o contemporánea. En un ejemplo la molécula de ácido nucleico diana es un miRNA y un mRNA. Así, en dicho ejemplo, el procedimiento incluiría el uso de al menos una NPPF específica para el miRNA y al menos una NPPF específica para el mRNA. En 45 un ejemplo las moléculas de ácidos nucleicos diana son dos moléculas de ADN diferentes. Así, en dicho ejemplo, el procedimiento incluiría el uso de al menos una NPPF específica para la primera diana ADN y al menos una NPPF específica para la segunda diana ADN. En un ejemplo las moléculas de ácidos nucleicos diana son dos moléculas de ARN diferentes. Así, en dicho ejemplo, el procedimiento incluiría el uso de al menos una NPPF específica para la primera ARN diana y al menos una NPPF específica para la segunda ARN diana.

50 En algunos ejemplos, los procedimientos divulgados permiten la detección o secuenciación de polimorfismos de ADN o ARN de un solo nucleótido (SNP) o variantes (sNPV), uniones de splice, ADN metilado, fusiones de genes u otras mutaciones, ADN o ARN unidos a proteínas y también cDNA, así como niveles de expresión (por ejemplo, expresión de ADN o ARN, tal como expresión de cDNA, expresión de mRNA, expresión de miRNA, expresión de 55 rRNA, expresión de siRNA o expresión de tRNA). Con los procedimientos divulgados puede cuantificarse e identificarse cualquier molécula de ácido nucleico con la cual pueda diseñarse una sonda de protección para hibridación, aun cuando las moléculas de ácidos nucleicos diana en sí no necesiten secuenciarse e incluso en algunos ejemplos puedan destruirse.

En un ejemplo, se detecta metilación de ADN usando una NPPF que incluye una falta de correspondencia de bases en el sitio en que se ha producido o no metilación, de manera que tras el tratamiento de la muestra diana, las bases metiladas se convierten en una base diferente, complementaria a la base en la NPPF.

- 5 Un experto en la materia observará que la diana puede incluir bases naturales o no naturales, o combinaciones de las mismas.

En ejemplo específico no limitativos, se selecciona un ácido nucleico diana (por ejemplo, un ADN diana o ARN diana) asociado con un neoplasma (por ejemplo, un cáncer). Se han identificado numerosas anomalías cromosómicas (que incluyen translocaciones y otras reordenaciones, reduplicación o delección) o mutaciones en células neoplásicas, especialmente en células cancerosas, como las leucemias de linfocitos B y linfocitos T, linfomas, cáncer de mama, cáncer de colon, cánceres neurológicos y similares.

En algunos ejemplos, una molécula de ácido nucleico diana incluye GAPDH (por ejemplo, nº acceso GenBank 15 NM_002046), PPIA (por ejemplo, nº acceso GenBank NM_021130), RPLP0 (por ejemplo, nº acceso GenBank NM_001002 o NM_053275), RPL19 (por ejemplo, nº acceso GenBank NM_000981), ZEB1 (por ejemplo, nº acceso GenBank NM_030751), Zeb2 (por ejemplo, nº acceso GenBank NM_001171653 o NM_014795), CDH1 (por ejemplo, nº acceso GenBank NM_004360), CDH2 (por ejemplo, nº acceso GenBank NM_007664), VIM (por ejemplo, nº acceso GenBank NM_003380), ACTA2 (por ejemplo, nº acceso GenBank NM_001141945 o NM_001613), CTNNA1 (por ejemplo, nº acceso GenBank NM_001904, NM_001098209, o NM_001098210), KRT8 (por ejemplo, nº acceso GenBank NM_002273), SNAI1 (por ejemplo, nº acceso GenBank NM_005985), SNAI2 (por ejemplo, nº acceso GenBank NM_003068), TWIST1 (por ejemplo, nº acceso GenBank NM_000474), CD44 (por ejemplo, nº acceso GenBank NM_000610, NM_001001389, NM_00100390, NM_001202555, NM_001001391, NM_001202556, NM_001001392, NM_001202557), CD24 (por ejemplo, nº acceso GenBank NM_013230), FN1 (por ejemplo, nº 25 acceso GenBank. NM_212474, NM_212476, NM_212478, NM_002026, NM_212482, NM_054034), IL6 (por ejemplo, nº acceso GenBank. NM_000600), MYC (por ejemplo, nº acceso GenBank NM_002467), VEGFA (por ejemplo, nº acceso GenBank NM_001025366, NM_001171623, NM_003376, NM_001171624, NM_001204384, NM_001204385, NM_001025367, NM_001171625, NM_001025368, NM_001171626, NM_001033756, NM_001171627, NM_001025370, NM_001171628, NM_001171622, NM_001171630), HIF1A (por ejemplo, nº 30 acceso GenBank NM_001530, NM_181054), EPAS1 (por ejemplo, nº acceso GenBank NM_001430), ESR2 (por ejemplo, nº acceso GenBank NM_001040276, NM_001040275, NM_001214902, NM_001437, NM_001214903), PRKCE (por ejemplo, nº acceso GenBank NM_005400), EZH2 (por ejemplo, nº acceso GenBank NM_001203248, NM_152998, NM_001203247, NM_004456, NM_001203249), DAB2IP (por ejemplo, nº acceso GenBank NM_032552, NM_138709), B2M (por ejemplo, nº acceso GenBank NM_004048) y SDHA (por ejemplo, nº acceso 35 GenBank NM_004168).

En algunos ejemplos, un miRNA diana incluye hsa-miR-205 (MIR205, por ejemplo, nº acceso GenBank NR_029622), hsa-miR-324 (MIR324, por ejemplo, nº acceso GenBank NR_029896), hsa-miR-301a (MIR301A, por ejemplo, nº acceso GenBank NR_029842), hsa-miR-106b (MIR106B, por ejemplo, nº acceso GenBank NR_029831), 40 hsa-miR-877 (MIR877, por ejemplo, nº acceso GenBank NR_030615), hsa-miR-339 (MIR339, por ejemplo, nº acceso GenBank NR_029898), hsa-miR-10b (MIR10B, por ejemplo, nº acceso GenBank NR_029609), hsa-miR-185 (MIR185, por ejemplo, nº acceso GenBank NR_029706), hsa-miR-27b (MIR27B, por ejemplo, nº acceso GenBank NR_029665), hsa-miR-492 (MIR492, por ejemplo, nº acceso GenBank NR_030171), hsa-miR-146a (MIR146A, por ejemplo, nº acceso GenBank NR_029701), hsa-miR-200a (MIR200A, por ejemplo, nº acceso GenBank NR_029834), 45 hsa-miR-30c (por ejemplo, nº acceso GenBank NR_029833, NR_029598), hsa-miR-29c (MIR29C, por ejemplo, nº acceso GenBank NR_029832), hsa-miR-191 (MIR191, por ejemplo, nº acceso GenBank NR_029690), o hsa-miR-655 (MIR655, por ejemplo, nº acceso GenBank NR_030391).

En un ejemplo la diana es un ácido nucleico patógeno, tal como ARN o ADN vírico. Los patógenos de ejemplo 50 incluyen, pero no se limitan a, virus, bacterias, hongos y protozoos. En un ejemplo, la diana es un ARN vírico. Los virus incluyen virus de ARN de cadena positiva y virus de ARN de cadena negativa. Los virus de ARN de cadena positiva de ejemplo incluyen, pero no se limitan a: Picornavirus (por ejemplo, Aftoviridae [por ejemplo, virus de la fiebre aftosa (FMDV)], Cardioviridae; Enteroviridae (por ejemplo, virus Coxsackie, Echovirus, Enterovirus y Poliovirus); Rhinoviridae (Rhinovirus)); Hepataviridae (virus de la hepatitis A); Togavirus (cuyos ejemplos incluyen 55 virus de la rubeola; alfavirus (por ejemplo, virus de la encefalitis equina occidental, virus de la encefalitis equina oriental y virus de la encefalitis equina venezolana)); Flavivirus (cuyos ejemplos incluyen virus del dengue, virus del Nilo occidental y virus de la encefalitis japonesa); y Coronavirus (cuyos ejemplos incluyen coronavirus del SARS, como la cepa Urbani). Los virus de ARN de cadena negativa de ejemplo incluyen, pero no se limitan a: Ortomixovirus (por ejemplo, virus de la gripe), Rabdovirus (por ejemplo, virus de la rabia) y Paramixovirus (cuyos ejemplos incluyen

virus del sarampión, virus sincitial respiratorio y virus de la parainfluenza). En un ejemplo la diana es ADN vírico de un virus de ADN, tal como Herpesvirus (por ejemplo, virus de la varicela-zóster, por ejemplo la cepa Oka; citomegalovirus; y virus del herpes simple (HSV) tipos 1 y 2), Adenovirus (por ejemplo, Adenovirus tipo 1 y Adenovirus tipo 41), Poxvirus (por ejemplo, virus de vacuna) y Parvovirus (por ejemplo, Parvovirus B19). En otro ejemplo, la diana es un ácido nucleico retrovírico, tal como virus de inmunodeficiencia humana tipo 1 (HIV-1), por ejemplo subtipo C, HIV-2; virus de la anemia infecciosa equina; virus de inmunodeficiencia felina (FIV); virus de la leucemia felina (FeLV); virus de inmunodeficiencia de los simios (SIV); y virus del sarcoma aviar. En un ejemplo, el ácido nucleico diana es un ácido nucleico bacteriano. En un ejemplo el ácido nucleico bacteriano procede de una bacteria gramnegativa, tal como *Escherichia coli* (K-12 y O157:H7), *Shigella dysenteriae* y *Vibrio cholerae*. En otro ejemplo el ácido nucleico bacteriano procede de una bacteria grampositiva, tal como *Bacillus anthracis*, *Staphylococcus aureus*, neumococo, gonococo y meningitis estreptocócica. En un ejemplo, el ácido nucleico diana es un ácido nucleico de protozoos, nematodos u hongos. Los protozoos de ejemplo incluyen, pero no se limitan a, *Plasmodium*, *Leishmania*, *Acanthamoeba*, *Giardia*, *Entamoeba*, *Cryptosporidium*, *Isospora*, *Balantidium*, *Trichomonas*, *Trypanosoma*, *Naegleria* y *Toxoplasma*. Los hongos de ejemplo incluyen, pero no se limitan a, *Coccidioides immitis* y *Blastomyces dermatitidis*.

Un experto en la materia puede identificar ADN o ARN diana adicionales y/o miRNA diana adicionales que pueden detectarse usando los procedimientos divulgados en la presente memoria descriptiva.

20 VII. Resultado del ensayo

En algunas realizaciones, los procedimientos divulgados incluyen la determinación de la presencia o de una cantidad de una o más moléculas de ácidos nucleicos diana en una muestra. En otras realizaciones o adicionales, los procedimientos divulgados incluyen la determinación de la secuencia de una o más moléculas de ácidos nucleicos diana en una muestra, que puede incluir la cuantificación de secuencias detectadas. Los resultados de los procedimientos pueden proporcionarse a un usuario (por ejemplo, un científico, un clínico u otro profesional sanitario, o un paciente) en una forma perceptible que proporciona información acerca de los resultados de la prueba. En algunos ejemplos, los resultados pueden darse en papel (por ejemplo, en un formato escrito o impreso), una representación en una pantalla, una forma gráfica (por ejemplo, un gráfico, una representación u otro diagrama) o una forma audible. En un ejemplo, el resultado es una tabla o gráfico que incluye un indicador cualitativo o cuantitativo de la presencia o la cantidad (por ejemplo, una cantidad normalizada) de una molécula de ácido nucleico diana detectada (o no detectada) en la muestra. En otros ejemplos la salida es un mapa o imagen de señal presente en un sustrato (por ejemplo, una imagen digital de fluorescencia de una matriz). En otros ejemplos de las realizaciones, la salida es la secuencia de una o más moléculas de ácidos nucleicos diana en una muestra, de manera que dicho informe indica la presencia de una mutación en particular en la molécula diana.

En algunos ejemplos, la salida es un valor numérico, tal como una cantidad de una molécula de ácido nucleico diana en una muestra. En ejemplos adicionales, la salida es una representación gráfica, por ejemplo, un gráfico que indica el valor (por ejemplo, cantidad o cantidad relativa) de una molécula de ácido nucleico diana en la muestra en una curva estándar. En ejemplos adicionales, la salida es una representación gráfica, por ejemplo, un gráfico que indica la secuencia de una molécula de ácido nucleico diana en la muestra (por ejemplo, que podría indicar dónde está presente una mutación). En algunos ejemplos, la salida se comunica al usuario, por ejemplo proporcionando una salida por medio de medios físicos, audibles o electrónicos (por ejemplo, por correo postal, teléfono, transmisión por fax, correo electrónico o comunicación en un registro médico electrónico).

La salida puede proporcionar información cualitativa (por ejemplo, una cantidad de una molécula de ácido nucleico diana en particular o una cantidad de una molécula de ácido nucleico diana en particular con respecto a una muestra o valor de control) o puede proporcionar información cualitativa (por ejemplo, determinación de presencia o ausencia de una molécula de ácido nucleico diana en particular). En ejemplos adicionales, la salida puede proporcionar información cualitativa relacionada con la cantidad de una molécula de ácido nucleico diana en la muestra, por ejemplo, la identificación de un aumento o disminución con respecto a un control o ausencia de cambio con respecto a un control.

Tal como se expone en la presente memoria descriptiva los amplicones de NPPF pueden incluir una o más etiquetas experimentales, que pueden usarse por ejemplo para identificar a un paciente, muestra, experimento o secuencia diana en particular. El uso de dichas etiquetas permite "clasificar" o incluso contar el amplicón de NPPF detectado o secuenciado, y así permite el análisis de múltiples muestras diferentes (por ejemplo, de distintos pacientes), múltiples dianas diferentes (por ejemplo, al menos dos ácidos nucleicos diana diferentes), o combinaciones de los mismos, en una única reacción. En un ejemplo, puede usarse software Illumina y Bowtie para dicho análisis.

En un ejemplo, las NPPF incluyen una etiqueta experimental única para cada molécula de ácido nucleico diana diferente. El uso de dicha etiqueta permite simplemente secuenciar o detectar esta etiqueta, sin secuenciación de toda la NPPF, para identificar la NPPF como correspondiente a un ácido nucleico diana en particular. Además, si se van a analizar múltiples ácidos nucleicos diana, el uso de una etiqueta experimental única para cada diana simplifica el análisis, ya que cada etiqueta experimental detectada o secuenciada puede clasificarse, y si se desea, contarse. Esto permite la cuantificación del ácido nucleico diana que había en la muestra, ya que los amplicones de NPPF están en proporción estequiométrica con la diana en la muestra. Por ejemplo si se detectan o secuencian múltiples ácidos nucleicos diana en una muestra, los procedimientos permiten la generación de una tabla o gráfico que muestra cada secuencia diana y el número de copias detectadas o secuenciadas, simplemente por detección o secuenciación y después clasificación de la etiqueta experimental.

En otro ejemplo, las NPPF incluyen una etiqueta experimental única para cada muestra diferente (por ejemplo, una etiqueta única para cada muestra de paciente). El uso de dicha etiqueta permite asociar un amplicón de NPPF detectado en particular con una muestra determinada. Así, si se analizan múltiples muestras en la misma reacción (por ejemplo, el mismo pocillo o la misma reacción de secuenciación), el uso de una etiqueta experimental única para cada muestra simplifica el análisis, ya que cada NPPF detectada o secuenciada puede asociarse con una muestra en particular. Por ejemplo si se detecta o secuencia un ácido nucleico diana en muestras, los procedimientos permiten la generación de una tabla o gráfico que muestra el resultado del análisis para cada muestra.

Un experto en la materia observará que cada amplicón de NPPF puede incluir una pluralidad de etiquetas experimentales (por ejemplo, al menos 2, 3, 4, 5, 6, 7, 8, 9 ó 10 etiquetas experimentales), por ejemplo una etiqueta que representa la secuencia diana, y otra que representa la muestra. Una vez detectada o secuenciada cada etiqueta, puede usarse un puede usarse apropiado para clasificar los datos en cualquier formato deseado, tal como un gráfico o tabla. Por ejemplo, esto permite el análisis de múltiples secuencias diana en múltiples muestras de forma simultánea o contemporánea.

En algunos ejemplos, el amplicón de NPPF detectado o secuenciado se compara con una base de datos de referencia de secuencias conocidas para cada secuencia de ácidos nucleicos diana. En algunos ejemplos, dicha comparación permite la detección de mutaciones, tales como SNP. En algunos ejemplos, dicha comparación permite la comparación de la abundancia de la NPPF de referencia con la abundancia de una sonda de NPPF en una región que según se sabe contiene SNP.

La descripción se ilustra además por medio de los siguientes ejemplos no limitativos.

EJEMPLO 1

Secuenciación simultánea de una pluralidad de NPPF

Este ejemplo describe procedimientos usados para generar y secuenciar NPPF.

Se generaron siete NPPF diferentes. Cada NPPF incluía una región que era específica para una molécula de ácido nucleico diana en particular de 25 nucleótidos de longitud con una Tf media de 62°C, así como secuencias flanqueantes en los dos extremos. Aunque las secuencias flanqueantes en 5' y 3' eran diferentes, eran las mismas para las siete NPPF diferentes. La secuencia flanqueante en 5' tenía 25 nucleótidos con una Tf de 61°C y la secuencia flanqueante en 3' tenía 25 nucleótidos con una Tf de 63°C.

Las siete NPPF diferentes se agruparon en relaciones conocidas (1:1.5:2:4:5) y se amplificaron por PCR del modo siguiente. Se incubaron las NPPF con cebadores de PCR. Un cebador incluía una secuencia que era complementaria a la secuencia flanqueante en 5' y el segundo cebador incluía una secuencia que era complementaria a la secuencia flanqueante en 3'. El segundo cebador también incluía una secuencia para permitir la incorporación de una etiqueta experimental de seis nucleótidos en el amplicón resultante, de manera que cada NPPF amplificada usando este cebador tenía la misma etiqueta experimental de seis nucleótidos. Se llevaron a cabo varias de estas reacciones, cada una con una etiqueta diferente. El primer cebador tenía 49 bases de longitud. Veinte de estas bases eran idénticas a la secuencia flanqueante en 5'. Estas 20 bases tenían una Tf de 54°C y la Tf global de todo el cebador era 70°C. El segundo cebador complementario a la secuencia flanqueante en 3' tenía 57 nucleótidos en total con una Tf de aproximadamente 70°C. Los primeros 19 nucleótidos del segundo cebador eran exactamente complementarios a la región flanqueante en 3' y tenían una Tf de 54°C.

- Se ejecutaron ocho reacciones de PCR separadas, de manera que pudieron identificarse variaciones. Los amplicones resultantes se limpiaron usando purificación en gel o purificación basada en columna estándar (columnas de centrifugación Qiagen QIAQuick). A continuación se secuenciaron los amplicones que contenían la
- 5 NPPF y una etiqueta experimental usando la plataforma Illumina. Se clasificó cada amplicón secuenciado basándose en la secuencia de etiquetas experimentales, con cada etiqueta representativa de una reserva en replicado de las siete NPPF. En cada grupo de etiquetas experimentales, se contó el número de amplicones identificados para cada una de las siete etiquetas.
- 10 Se secuenciaron 128 millones de, y de ellos, 110 millones (87%) produjeron una etiqueta experimental perfectamente secuenciada. Los amplicones se compararon con las secuencias esperadas usando Bowtie, con el resultado de aproximadamente el 80% secuencias de correspondencias perfectas. Este es un buen porcentaje de secuencias de correspondencias perfectas para el sistema Illumina, basándose en sus especificaciones publicadas de errores y calidad. La FIG. 5 muestra el número de amplicones detectados para cada uno de los siete NPPF
- 15 especiales correspondientes a la proporción original de NPPF agrupados antes de PCR. Las sondas se midieron en ocho experimentos separados, cada uno de los cuales tenía una etiqueta experimental diferente añadida durante la amplificación. Fueron todos agrupados en un único canal del secuenciador, y secuenciados. Las barras de error indican la reproducibilidad (1 DT) de los ocho experimentos
- 20 La FIG. 6, una nueva representación de los datos mostrados en la FIG. 5, muestra los ocho resultados experimentales individuales para cada sonda, la media (sin barras de error, misma media que la representada en la FIG. 5 con barras de error) y lo esperado (basándose en la cantidad de NPPF añadida a la muestra). Las relaciones observadas para cada una de las siete NPPF tuvieron correspondencia con lo esperado (basándose en la cantidad original de NPPF añadida a la reacción de PCR).

25

EJEMPLO 2

Detección simultánea de una pluralidad de NPPF

- 30 Este ejemplo describe procedimientos usados para generar y detectar NPPF usando una matriz y cuantificar el grado de amplificación alcanzado.

Se generaron tres NPPF diferentes (uno que contenía una secuencia complementaria al gen BAX humano, uno que contenía una secuencia complementaria de gen de fusión EML4-ALK y uno que contenía una secuencia complementaria EML4). Así, cada NPPF incluía una región que era específica para una molécula de ácido nucleico

35 diana en particular, así como secuencias flanqueantes en los dos extremos. El extremo 5' de cada NPPF tenía una etiqueta de biotina. Las NPPF tenían 25 nucleótidos, que tenían una Tf de 63°C a 65°C. Aunque las secuencias flanqueantes en 5' y 3' eran diferentes, coincidían para cada una de las siete NPPF diferentes. La secuencia flanqueante en 5' tenía 25 nucleótidos con una Tf de 61°C y la secuencia flanqueante en 3' tenía 25 nucleótidos con

40 una Tf de 63°C.

En un experimento, las NPPF no amplificadas se hibridaban con una matriz después de qNPA. La matriz incluía una sonda de anclaje unida a conectores bifuncionales. La mitad del conector bifuncional es complementaria al anclaje, y la otra mitad es complementaria a la parte específica del gen de la NPPF. Así, el conector forma un puente entre el

45 anclaje y la NPPF. Las tres NPPF diferentes se agruparon en relaciones conocidas, y se hibridaron con ARN sintéticos que contenían las secuencias diana de interés, así como con las CFS complementarias a las regiones flanqueantes en las NPPF. Después de la digestión mediada por S1 de ARN, NPPF y CFS no hibridados, se dividió la reacción. Se incubó una fracción con la matriz en condiciones de permitir que las NPPF se unieran a su conector bifuncional apropiado. La unión de la NPPF a la matriz fue detectada por la etiqueta de biotina presente en la NPPF

50 usando estreptavidina ficoeritrina fluorescente.

En otro experimento, se amplificó por PCR otra fracción de la reacción agrupada antes de la hibridación con una matriz, y se diluyó el producto 1:10 ó 1:100 antes de la hibridación con la matriz. Para amplificación por PCR se incubó la reacción que contenían las NPPF con cebadores de PCR. Un cebador incluía una secuencia que era

55 idéntica a la secuencia flanqueante en 5' (e incluía una etiqueta de biotina) y el segundo cebador incluía una secuencia que era complementaria a la secuencia flanqueante en 3'. El primer cebador complementario a la secuencia flanqueante en 5' tenía 22 nucleótidos y tenía una Tf de 59°C, y el segundo cebador complementario a la secuencia flanqueante en 3' tenía 22 nucleótidos y tenía una Tf de 56°C. La ventaja de usar NPPF que tienen las mismas secuencias flanqueantes (pero diferentes regiones específicas de diana) reside en que las secuencias

flanqueantes permiten el uso de cebadores de PCR universales, de manera que sólo se necesita una secuencia de cebador en 5' y una secuencia de cebador en 3' para amplificar varias secuencias de NPPF diferentes. Los amplicones de NPPF se diluyeron 1:10 o 1:100 y después se hibridaron con la matriz y se detectaron tal como se describe anteriormente.

5

Tal como se muestra en la FIG. 7, el uso de amplificación por PCR, antes de captura de hibridación, aumenta la sensibilidad al menos 150 veces (teniendo en cuenta la dilución de los amplicones después de la etapa de PCR).

EJEMPLO 3

10

Secuenciación simultánea de una pluralidad de NPPF diseñadas para medir mRNA o miRNA

Este ejemplo describe procedimientos usados para generar y secuenciar NPPF. Se generaron dos conjuntos de NPPF. En el primer conjunto se generaron cuarenta y seis NPPF diferentes. Cada NPPF incluía una región que era específica para una molécula de ácido nucleico diana en particular de 25 nucleótidos de longitud con una Tf media de 56°C, así como secuencias flanqueantes en los dos extremos. Para el segundo conjunto, se generaron trece NPPF diferentes. Cada NPPF incluía una región que era específica para una molécula de ácido nucleico diana en particular de miRNA de 18-25 nucleótidos de longitud con una Tf media de 51°C, así como secuencias flanqueantes en los dos extremos.

20

Para todas las NPPF, con independencia de la diana, aunque las secuencias flanqueantes en 5' y 3' diferían, eran iguales para cada una de las NPPF diferentes. La secuencia flanqueante en 5' (5'-AGTTCAGACGTGTGCTCTTCCGATC 3'; SEQ ID NO: 17) tenía 25 nucleótidos con una Tf de 61°C y la secuencia flanqueante en 3' (5' GATCGTCCGACTGTAGAACTCTGAA 3'; SEQ ID NO: 18) tenía 25 nucleótidos con una Tf de 63°C.

25

Se realizó qNPA en lisados de líneas celulares a diferentes concentraciones, usando estas NPPF como sondas. La FIG. 10 muestra las reacciones de qNPA, las muestras usadas como material de entrada y las etiquetas experimentales añadidas antes de la secuenciación. Las reacciones se realizaron por triplicado para cada concentración celular. Algunas etiquetas experimentales no fueron reconocidas por el software secuenciador y así las reacciones marcadas con dichas etiquetas experimentales no se consideraron en este análisis. Las NPPF diferentes se agruparon y se hibridaron con el ARN de un lisado celular, así como las CFS complementarias a las regiones flanqueantes en las NPPF. La hibridación se realizó a 50°C para las cuarenta y seis NPPF del conjunto 1, pero se realizaron a 37°C para las trece NPPF del conjunto 2. La diferencia en la temperatura tiene en cuenta la menor longitud de las NPPF del miRNA y sus Tf menores correspondientes.

35

Después de la digestión mediada por S1 de ARN, NPPF y CFS no hibridados, se neutralizó la reacción por adición de Tris 1 M pH 9,0 y se inactivó la nucleasa S1 por calentamiento a 95°C durante 20 minutos. A continuación se incubó cada reacción resultante, que contenía las NPPF como representativas de las transcripciones originales en la muestra, con cebadores de PCR. Un cebador incluía una secuencia que era complementaria a la secuencia flanqueante en 5' y el segundo cebador incluía una secuencia que era complementaria a la secuencia flanqueante en 3'. El segundo cebador también incluía una secuencia para permitir la incorporación de una etiqueta experimental de seis nucleótidos en el amplicón resultante, de manera que cada NPPF amplificada usando este cebador tenía la misma etiqueta experimental de seis nucleótidos.

45

El primer cebador (5'-AATGATACGGCGACCACCGACAGGTTTCAGAGTTCTACAGTCCGACGATC 3'; SEQ ID NO: 19) tenía 49 bases de longitud. Veinte de estas bases eran idénticas a la secuencia flanqueante en 5'. Estas 20 bases tenían una Tf de 54°C y la Tf global de todo el cebador era de 70°C. El segundo cebador, (5'-CAAGCAGAAGACGGCATAACGAGATnnnnnnGTGACTGGAGTTTCAGACGTGTGCTCTT 3'; SEQ ID NO: 20) complementario a la secuencia flanqueante en 3' tenía 57 nucleótidos en total con una Tf de aproximadamente 70°C. Los 19 primeros nucleótidos del segundo cebador eran exactamente complementarios a la región flanqueante en 3' y tenían una Tf de 54°C. Las seis bases marcadas con "nnnnnn" anteriores eran una de las siguientes 24 secuencias de la Tabla 2. La secuencia resultante se muestra en la columna de la derecha, con su SEQ ID NO: entre paréntesis.

55

Tabla 2: Secuencia de cebadores y códigos de barras

Secuencia de códigos de barras (nnnnnn en SEQ ID NO: 20)	Secuencia de cebadores resultante con código de barras (SEQ ID NO:)
ATCACG	CAAGCAGAAGACGGCATAACGAGATTCACGGTGACTGG AGTTCAGACGTGTGCTCTT (21)
CGATGT	CAAGCAGAAGACGGCATAACGAGATCGATGTGTGACTG GAGTTCAGACGTGTGCTCTT (22)
TTAGGC	CAAGCAGAAGACGGCATAACGAGATTTAGGCCTGACTG GAGTTCAGACGTGTGCTCTT (23)
TGACCA	CAAGCAGAAGACGGCATAACGAGATTGACCAGTGACTG GAGTTCAGACGTGTGCTCTT (24)
ACAGTG	CAAGCAGAAGACGGCATAACGAGATACAGTGGTGACTG GAGTTCAGACGTGTGCTCTT (25)
GCCAAT	CAAGCAGAAGACGGCATAACGAGATGCCAATGTGACTG GAGTTCAGACGTGTGCTCTT (26)
CAGATC	CAAGCAGAAGACGGCATAACGAGATCAGATCGTGACTG GAGTTCAGACGTGTGCTCTT (27)
ACTTGA	CAAGCAGAAGACGGCATAACGAGATACTTGAGTGACTG GAGTTCAGACGTGTGCTCTT (28)
GATCAG	CAAGCAGAAGACGGCATAACGAGATGATCAGGTGACTG GAGTTCAGACGTGTGCTCTT (29)
TAGCTT	CAAGCAGAAGACGGCATAACGAGATTAGCTTGTGACTG GAGTTCAGACGTGTGCTCTT (30)
GGCTAC	CAAGCAGAAGACGGCATAACGAGATGGCTACGTGACTG GAGTTCAGACGTGTGCTCTT (31)
CTTGTA	CAAGCAGAAGACGGCATAACGAGATCTTGTAGTGACTG GAGTTCAGACGTGTGCTCTT (32)
AGTCAA	CAAGCAGAAGACGGCATAACGAGATAGTCAAGTGACTG GAGTTCAGACGTGTGCTCTT (33)
AGTTCC	CAAGCAGAAGACGGCATAACGAGATAGTTCCGTGACTG GAGTTCAGACGTGTGCTCTT (34)
ATGTCA	CAAGCAGAAGACGGCATAACGAGATATGTCAAGTGACTG GAGTTCAGACGTGTGCTCTT (35)
CCGTCC	CAAGCAGAAGACGGCATAACGAGATCCGTCCGTGACTG GAGTTCAGACGTGTGCTCTT (36)
GTAGAG	CAAGCAGAAGACGGCATAACGAGATGTAGAGGTGACTG GGAGTTCAGACGTGTGCTCTT (37)
GTCCGC	CAAGCAGAAGACGGCATAACGAGATGTCCGCCTGACTG GAGTTCAGACGTGTGCTCTT (38)
GTGAAA	CAAGCAGAAGACGGCATAACGAGATGTGAAAAGTGACTG GGAGTTCAGACGTGTGCTCTT (39)
GTGGCC	CAAGCAGAAGACGGCATAACGAGATGTGGCCGTGACTG GAGTTCAGACGTGTGCTCTT (40)
GTTTCG	CAAGCAGAAGACGGCATAACGAGATGTTTCGGTGACTG GAGTTCAGACGTGTGCTCTT (41)
CGTACG	CAAGCAGAAGACGGCATAACGAGATCGTACGGTGACTG GAGTTCAGACGTGTGCTCTT (42)
GAGTGG	CAAGCAGAAGACGGCATAACGAGATGAGTGGGTGACTG GGAGTTCAGACGTGTGCTCTT (43)
GGTAGC	CAAGCAGAAGACGGCATAACGAGATGGTAGCGTGACTG GAGTTCAGACGTGTGCTCTT (44)

Cada reacción por triplicado se amplificó en una reacción PCR separada, y tenía una etiqueta experimental separada, de manera que pudiera identificarse la variancia (véase FIG. 10). Se limpiaron los amplicones resultantes usando purificación en gel o purificación basada en columna estándar (columnas de centrifugación Qiagen QIAQuick). A continuación se secuenciaron los amplicones que contenían la NPPF y una etiqueta experimental usando la plataforma Illumina. Aunque la etiqueta experimental puede localizarse en varios lugares, en este ejemplo,

estaba situada en el extremo 3' del amplicón, inmediatamente en la dirección 3' de una región complementaria a un cebador de secuenciación de lectura de índice. Así la secuenciación Illumina se realizó en dos etapas, una lectura inicial de la secuencia seguida por una segunda lectura de la etiqueta experimental usando un segundo cebador de secuenciación. El uso de esta manera de dos cebadores de secuenciación es un procedimiento estándar para
5 multiplexación de muestras en la plataforma Illumina.

A continuación se clasificó cada amplicón secuenciado basándose primero en la etiqueta experimental (código de barras), y después en cada grupo de etiquetas experimentales, y se contó el número de amplicones identificados para cada una de las diferentes etiquetas. Los amplicones se compararon con las secuencias esperadas usando
10 Bowtie.

La FIG. 11 muestra los resultados de las reacciones de qNPA en triplicado en células THP1, usando las 46 NPPF de mRNA. Se observó una reproducibilidad excelente entre replicados y los valores CV son bajos. el gráfico representa el número de amplicones detectados para cada una de las cuarenta y seis NPPF especiales correspondientes a la
15 relación original de NPPF agrupadas antes de PCR. Las barras de error representan 1 desviación típica con respecto a la media. Las sondas se midieron en tres experimentos separados, cada uno de los cuales tenía una etiqueta experimental diferente añadida durante la amplificación. Todos ellos se agruparon en un único canal del secuenciador y se secuenciaron. Las barras de error indican la reproducibilidad (1 DT) de los tres experimentos.

20 Las FIG. 12A y 12B muestran los recuentos representados obtenidos para 12 de las 46 NPPF de mRNA de las reacciones que se realizan en una valoración celular THP 1 de cuatro puntos. Los datos mostrados representan los valores de NPPF de mínima (A) y máxima (B) abundancia, y muestran el gran intervalo de detección que puede obtenerse usando secuenciación. También muestran la linealidad de la reacción qNPS para sondas de alta y baja abundancia (que representan alta y baja expresión del ARN correspondiente en la muestra).

25 La FIG. 13 representa los resultados para cinco de las trece NPPF de miRNA de reacciones realizadas en una valoración celular HepG2 de tres puntos (de 5.000 células a 50.000 células). Estas cinco se eligieron porque tenían niveles semejantes y podían verse claramente en el mismo gráfico. La representación gráfica muestra que los miRNA son detectables en lisados celulares que usan los procedimientos divulgados, y muestran buena linealidad
30 en los tamaños de muestra sometidos a ensayo.

EJEMPLO 4

35 **Detección de una pluralidad de NPPF diseñadas para medir mRNA usando secuenciación y captura de las NPPF en una matriz**

Este ejemplo describe procedimientos usados para generar y secuenciar NPPF.

Se generaron nueve NPPF diferentes. Cada NPPF incluía una región que era específica para una molécula de ácido nucleico diana en particular de 25 nucleótidos de longitud con una Tf media de 57°C, así como secuencias
40 flanqueantes en los dos extremos. Aunque las secuencias flanqueantes en 5' y 3' diferían, eran iguales para cada una de las NPPF diferentes. La secuencia flanqueante en 5' (5'-AGTTCAGACGTGTGCTCTTCCGATC-3'; SEQ ID NO: 17) tenía 25 nucleótidos con una Tf de 61°C y la secuencia flanqueante en 3' (5'-GATCGTCGGACTGTAGAACTCTGAA-3'; SEQ ID NO: 18) tenía 25 nucleótidos con una Tf de 63°C. Se incluyó una
45 etiqueta de biotina en la 5' secuencia flanqueante de cada NPPF.

Se realizó qNPA en muestras formadas por diluciones de ARN sintéticos (ARN transcritos *in vitro*) en tampón de lisis de qNPA. Las reacciones se realizaron por triplicado para cada concentración de muestra. Se agruparon las NPPF diferentes a 166 pM cada una, y se hibridaron con las muestras descritas anteriormente, así como las CFS
50 complementarias a las regiones flanqueantes en las NPPF. Las CFS se incluyeron en la reacción en una razón molar de 10 veces (1,6 nM cada CFS). La hibridación se realizó a 50°C durante al menos 16 horas en un volumen de reacción total de 30 µl. Después de la hibridación, se añadieron 20 µl de tampón de reacción S1 a la reacción. Este tampón estaba formado por: NaOAc 100 mM pH 5,0, KCl 250 mM, ZnSO4 22,5 nM y 25U de nucleasa S1. Se dejó avanzar la reacción S1 durante 90 minutos a 50°C. Después de digestión mediada por S1 de ARN, NPPF y
55 CFS no hibridados, se neutralizó la reacción mediante adición de 1,5 µl de Tris 1 M pH 9,0 y se inactivó la nucleasa S1 por calentamiento a 95°C durante 20 minutos. Cada reacción resultante contenía NPPF como representantes de las transcripciones originales en la muestra. En este punto, la reacción se dividió en dos partes.

Una parte de las NPPF no amplificadas se hibridó con una matriz después de qNPA. La matriz incluía una sonda de

anclaje unida a conectores bifuncionales. La mitad del conector bifuncional es complementaria al anclaje, y la otra es complementaria a la parte de la NPPF específica del gen. El conector forma así un puente entre el anclaje y la NPPF. Las NPPF de la reacción anterior se complementaron con un tampón de sustitución de sal para ajustar la reacción a las condiciones usadas para hibridación de matriz (la tampón de sustitución salina es: NaCl 3,225 M; 5 EDTA 67,5 mM pH 8,0; 3x SSC; HEPES 500 mM pH 7,5) y se incubaron con la matriz durante 16 horas a 50°C. La unión de la NPPF a la matriz fue detectada por la etiqueta de biotina presente en la NPPF usando estreptavidina-ficoeritrina fluorescente.

La otra parte de la reacción se preparó para secuenciación. Primero se incubó la reacción con cebadores de PCR. 10 Un cebador incluía una secuencia que era complementaria a la secuencia flanqueante en 5' y el segundo cebador incluía una secuencia que era complementaria a la secuencia flanqueante en 3'. El segundo cebador también incluía una secuencia para permitir para la incorporación de una etiqueta experimental de seis nucleótidos en el amplicón resultante, de manera que cada NPPF amplificada usando este cebador tenía la misma etiqueta experimental de seis nucleótidos.

15 El primer cebador tenía 49 bases de longitud. Veinte de estas bases eran idénticas a la secuencia flanqueante en 5'. Estas 20 bases tenían una Tf de 54°C y la Tf global de todo el cebador era de 70°C. El segundo cebador, complementario a la secuencia flanqueante en 3', tenía 57 nucleótidos en total con una Tf de aproximadamente 70°C. Los 19 primeros nucleótidos del segundo cebador eran exactamente complementarios a la región flanqueante 20 en 3' y tenían una Tf de 54°C.

Se amplificó cada reacción por triplicado en una reacción de PCR separada, con una etiqueta separada, de manera que pudiera identificarse la variancia. Se limpiaron los amplicones resultantes usando purificación en gel o purificación basada en columna estándar (columnas de centrifugación Qiagen QIAQuick). A continuación se 25 secuenciaron los amplicones que contenían la NPPF y una etiqueta experimental usando una plataforma Illumina, usando la segunda técnica de índice de lectura para secuenciar la etiqueta experimental, tal como se describe en el Ejemplo 3.

Las reacciones de PCR se configuraron también para determinar el impacto del número de ciclos en los resultados 30 de la secuenciación. Brevemente, cada reacción por triplicado se amplificó en tres reacciones de PCR separadas, cada reacción, con una etiqueta separada, de manera que pudiera identificarse la variancia. Estas tres reacciones de PCR se sometieron a 10, 12 ó 15 ciclos de PCR y los amplicones resultantes se limpiaron usando purificación en gel o purificación basada en columna estándar (columnas de centrifugación Qiagen QIAQuick). A continuación se secuenciaron los amplicones que contenían la NPPF y una etiqueta experimental usando una plataforma Illumina, 35 usando la segunda técnica de índice de lectura para secuenciar la etiqueta experimental, tal como se describe en el Ejemplo 3.

Se clasificó cada amplicón secuenciado basándose primero en la etiqueta experimental (código de barras), y después en cada grupo de etiquetas experimentales, y se contó el número de amplicones identificados para cada 40 una de las diferentes etiquetas. Los amplicones se compararon con las secuencias esperadas usando Bowtie.

La FIG. 14 muestra que números de ciclo de PCR bajos (10, 12 y 15) no influyen de forma indebida en los resultados de secuenciación. El gráfico de barras muestra los recuentos generados para cada NPPF después de la 45 secuenciación. Se indica el número de ciclos y la cantidad de material de entrada en la muestra original. Se normalizaron los datos para permitir la comparación de los diferentes ciclos y niveles de entrada. Aunque está claro que cualquiera de estos ciclos podría usarse con los procedimientos divulgados, el aumento en el material en las muestras después de 15 ciclos de PCR hizo más fácil la posterior limpieza de la biblioteca de secuenciación. Más de 15 ciclos producen productos espurios más grandes y más pequeños que el tamaño de amplicón deseado. Así, en algunos ejemplos, el procedimiento divulgado usa de 10 a 15 PCR ciclos, por ejemplo, 10, 11, 12, 13, 14 ó 15 ciclos.

50 Las FIG. 15A y 15B muestran los resultados a partir de las mismas reacciones de qNPA por triplicado después de la división. Las NPPF se detectaron por hibridación en una matriz (FIG. 15A) o contando las NPPF secuenciadas (FIG. 15B). Las barras mostradas son promedios de los triplicados, y las barras de error representan una desviación típica con respecto a la media.

55 En vistas de las muchas realizaciones posibles a las que pueden aplicarse los principios de la invención divulgada, debe reconocerse que las realizaciones ilustradas son sólo ejemplos de la descripción y no deben tomarse como limitativos del alcance de la descripción. Al contrario, el alcance de la invención se define mediante las reivindicaciones siguientes.

LISTADO DE SECUENCIAS

<110> HTG Molecular Diagnostics, Inc.
 5 <120> MEJORAS EN UN ENSAYO CUANTITATIVO DE PROTECCIÓN CONTRA LA NUCLEASA (QNPA) Y SECUENCIACIÓN (QNPS)
 <130> 8736-87683-07
 10 <150> 61/482.486
 <151> 2011-05-04
 <150> 61/537.492
 15 <151> 2011-09-21
 <150> 61/576.143
 <151> 2011-12-15
 20 <160> 44
 <170> PatentIn versión 3.5
 <210> 1
 25 <211> 16
 <212> ADN
 <213> Secuencia artificial
 <220>
 30 <223> secuencia de anclaje de ejemplo
 <400> 1
 tgattcagac cggccg 16
 35 <210> 2
 <211> 16
 <212> ADN
 <213> Secuencia artificial
 40 <220>
 <223> secuencia de anclaje de ejemplo
 <400> 2
 cccggggcgt ctaac 16
 45 <210> 3
 <211> 16
 <212> ADN
 <213> Secuencia artificial
 50 <220>
 <223> secuencia de anclaje de ejemplo
 <400> 3
 55 ggaccata tgcgct 16
 <210> 4
 <211> 16
 <212> ADN

<213> Secuencia artificial

<220>
<223> secuencia de anclaje de ejemplo

5 <400> 4
tgagggtccc gccata 16

<210> 5
10 <211> 16
<212> ADN
<213> Secuencia artificial

<220>
15 <223> secuencia de anclaje de ejemplo

<400> 5
aaccctgtgac gtgtgc 16

20 <210> 6
<211> 16
<212> ADN
<213> Secuencia artificial

25 <220>
<223> secuencia de anclaje de ejemplo

<400> 6
agcatcgccg gtcctg 16

30 <210> 7
<211> 16
<212> ADN
<213> Secuencia artificial

35 <220>
<223> secuencia de anclaje de ejemplo

<400> 7
40 cctgcaaggc tgacgt 16

<210> 8
<211> 16
<212> ADN
45 <213> Secuencia artificial

<220>
<223> secuencia de anclaje de ejemplo

50 <400> 8
cagttgtcga ccccgg 16

<210> 9
<211> 16
55 <212> ADN
<213> Secuencia artificial

<220>
<223> secuencia de anclaje de ejemplo

<400> 9
 cggcgcgtcc aattcg 16

5 <210> 10
 <211> 16
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> secuencia de anclaje de ejemplo

<400> 10
 atcgatctga gggccc 16

15 <210> 11
 <211> 16
 <212> ADN
 <213> Secuencia artificial

20 <220>
 <223> secuencia de anclaje de ejemplo

<400> 11
 25 gtacatgcgg cctgca 16

<210> 12
 <211> 16
 <212> ADN
 30 <213> Secuencia artificial

<220>
 <223> secuencia de anclaje de ejemplo

35 <400> 12
 tagccgctcg ctagag 16

<210> 13
 <211> 16
 40 <212> ADN
 <213> Secuencia artificial

<220>
 <223> secuencia de anclaje de ejemplo

45 <400> 13
 cctagtgatg accggc 16

<210> 14
 50 <211> 16
 <212> ADN
 <213> Secuencia artificial

<220>
 55 <223> secuencia de anclaje de ejemplo

<400> 14
 gtctgagggc aacctc 16

<210> 15
 <211> 16
 <212> ADN
 <213> Secuencia artificial
 5
 <220>
 <223> secuencia de anclaje de ejemplo

 <400> 15
 10 ctagctggct acgcag 16

 <210> 16
 <211> 16
 <212> ADN
 15 <213> Secuencia artificial

 <220>
 <223> secuencia de anclaje de ejemplo

 20 <400> 16
 gccatccgct tggagc 16

 <210> 17
 <211> 25
 25 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> secuencia flanqueante de ejemplo del extremo 5'
 30
 <400> 17
 agttcagacg tgtgctcttc cgatc 25

 <210> 18
 35 <211> 25
 <212> ADN
 <213> Secuencia artificial

 <220>
 40 <223> secuencia flanqueante de ejemplo del extremo 3'

 <400> 18
 gatcgtcgga ctgtagaact ctgaa 25

 45 <210> 19
 <211> 49
 <212> ADN
 <213> Secuencia artificial

 50 <220>
 <223> cebador de PCR

 <400> 19
 55 aatgatacgg cgaccaccga cagggtcaga gttctacagt cgcacgatc 49

 <210> 20
 <211> 57
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> cebador de PCR

5 <220>
 <221> misc_feature
 <222> (25).(30)
 10 <223> nnnnnn en posiciones 25 a 30 se muestra en posiciones 25 a 30 SEQ ID NO: 21-44

<400> 20
 caagcagaag acggcatacg agatnnnnn gtgactggag ttcagacgtg tgctctt 57

15 <210> 21
 <211> 56
 <212> ADN
 <213> Secuencia artificial

20 <220>
 <223> cebador de ejemplo con una secuencia de código de barras presente en los nucleótidos 25-30

<400> 21
 caagcagaag acggcatacg agattcacgg tgactggagt tcagacgtg gctctt 56

25 <210> 22
 <211> 57
 <212> ADN
 <213> Secuencia artificial

30 <220>
 <223> cebador de ejemplo con una secuencia de código de barras presente en los nucleótidos 25-30

<400> 22
 35 caagcagaag acggcatacg agatcgatgt gtgactggag ttcagacgtg tgctctt 57

<210> 23
 <211> 57
 <212> ADN
 40 <213> Secuencia artificial

<220>
 <223> cebador de ejemplo con una secuencia de código de barras presente en los nucleótidos 25-30

45 <400> 23
 caagcagaag acggcatacg agatttaggc gtgactggag ttcagacgtg tgctctt 57

<210> 24
 <211> 57
 50 <212> ADN
 <213> Secuencia artificial

<220>
 <223> cebador de ejemplo con una secuencia de código de barras presente en los nucleótidos 25-30

55 <400> 24
 caagcagaag acggcatacg agattgacca gtgactggag ttcagacgtg tgctctt 57

<210> 25

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

5 <220>
 <223> cebador de ejemplo con una secuencia de código de barras presente en los nucleótidos 25-30

<400> 25
 caagcagaag acggcatacg agatacagtg gtgactggag ttcagacgtg tgctctt 57

10 <210> 26
 <211> 57
 <212> ADN
 <213> Secuencia artificial

15 <220>
 <223> cebador de ejemplo con una secuencia de código de barras presente en los nucleótidos 25-30

<400> 26
 20 caagcagaag acggcatacg agatgccaat gtgactggag ttcagacgtg tgctctt 57

<210> 27
 <211> 57
 <212> ADN

25 <213> Secuencia artificial

<220>
 <223> cebador de ejemplo con una secuencia de código de barras presente en los nucleótidos 25-30

30 <400> 27
 caagcagaag acggcatacg agatcagatc gtgactggag ttcagacgtg tgctctt 57

<210> 28
 <211> 57

35 <212> ADN
 <213> Secuencia artificial

<220>
 <223> cebador de ejemplo con una secuencia de código de barras presente en los nucleótidos 25-30

40 <400> 28
 caagcagaag acggcatacg agatactga gtgactggag ttcagacgtg tgctctt 57

<210> 29

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

<220>
 50 <223> cebador de ejemplo con una secuencia de código de barras presente en los nucleótidos 25-30

<400> 29
 caagcagaag acggcatacg agatgatcag gtgactggag ttcagacgtg tgctctt 57

55 <210> 30
 <211> 57
 <212> ADN
 <213> Secuencia artificial

ES 2 597 032 T3

<220>
<223> cebador de ejemplo con una secuencia de código de barras presente en los nucleótidos 25-30

<400> 30
5 caagcagaag acggcatacg agattagctt gtgactggag ttcagacgtg tgctctt 57

<210> 31
<211> 57
<212> ADN
10 <213> Secuencia artificial

<220>
<223> cebador de ejemplo con una secuencia de código de barras presente en los nucleótidos 25-30

15 <400> 31
caagcagaag acggcatacg agatggctac gtgactggag ttcagacgtg tgctctt 57

<210> 32
<211> 57
20 <212> ADN
<213> Secuencia artificial

<220>
<223> cebador de ejemplo con una secuencia de código de barras presente en los nucleótidos 25-30

25 <400> 32
caagcagaag acggcatacg agatcttga gtgactggag ttcagacgtg tgctctt 57

<210> 33
30 <211> 57
<212> ADN
<213> Secuencia artificial

<220>
35 <223> cebador de ejemplo con una secuencia de código de barras presente en los nucleótidos 25-30

<400> 33
caagcagaag acggcatacg agatagtcaa gtgactggag ttcagacgtg tgctctt 57

40 <210> 34
<211> 59
<212> ADN
<213> Secuencia artificial

45 <220>
<223> cebador de ejemplo con una secuencia de código de barras presente en los nucleótidos 25-30

<400> 34
50 agcaagcaga agacggcata cgagatagtt cctgactgg agttcagacg tggctctt 59

<210> 35
<211> 57
<212> ADN
<213> Secuencia artificial

55 <220>
<223> cebador de ejemplo con una secuencia de código de barras presente en los nucleótidos 25-30

<400> 35

caagcagaag acggcatacg agatagtca gtgactggag ttcagacgtg tgctctt 57

<210> 36
 <211> 57
 5 <212> ADN
 <213> Secuencia artificial

<220>
 <223> cebador de ejemplo con una secuencia de código de barras presente en los nucleótidos 25-30

10 <400> 36
 caagcagaag acggcatacg agatccgtcc gtgactggag ttcagacgtg tgctctt 57

<210> 37
 15 <211> 57
 <212> ADN
 <213> Secuencia artificial

<220>
 20 <223> cebador de ejemplo con una secuencia de código de barras presente en los nucleótidos 25-30

<400> 37
 caagcagaag acggcatacg agatgtagag gtgactggag ttcagacgtg tgctctt 57

25 <210> 38
 <211> 57
 <212> ADN
 <213> Secuencia artificial

30 <220>
 <223> cebador de ejemplo con una secuencia de código de barras presente en los nucleótidos 25-30

<400> 38
 caagcagaag acggcatacg agatgtccgc gtgactggag ttcagacgtg tgctctt 57

35 <210> 39
 <211> 57
 <212> ADN
 <213> Secuencia artificial

40 <220>
 <223> cebador de ejemplo con una secuencia de código de barras presente en los nucleótidos 25-30

<400> 39
 45 caagcagaag acggcatacg agatgtgaaa gtgactggag ttcagacgtg tgctctt 57

<210> 40
 <211> 57
 <212> ADN
 50 <213> Secuencia artificial

<220>
 <223> cebador de ejemplo con una secuencia de código de barras presente en los nucleótidos 25-30

55 <400> 40
 caagcagaag acggcatacg agatgtggcc gtgactggag ttcagacgtg tgctctt 57

<210> 41
 <211> 57

ES 2 597 032 T3

<212> ADN
<213> Secuencia artificial

<220>

5 <223> cebador de ejemplo con una secuencia de código de barras presente en los nucleótidos 25-30

<400> 41

caagcagaag acggcatacg agatgttctg gtgactggag ttcagacgtg tgctctt 57

10 <210> 42

<211> 57

<212> ADN

<213> Secuencia artificial

15 <220>

<223> cebador de ejemplo con una secuencia de código de barras presente en los nucleótidos 25-30

<400> 42

caagcagaag acggcatacg agatcgtacg gtgactggag ttcagacgtg tgctctt 57

20

<210> 43

<211> 57

<212> ADN

<213> Secuencia artificial

25

<220>

<223> cebador de ejemplo con una secuencia de código de barras presente en los nucleótidos 25-30

<400> 43

30 caagcagaag acggcatacg agatgagtgg gtgactggag ttcagacgtg tgctctt 57

<210> 44

<211> 57

<212> ADN

35 <213> Secuencia artificial

<220>

<223> cebador de ejemplo con una secuencia de código de barras presente en los nucleótidos 25-30

40 <400> 44

caagcagaag acggcatacg agatgtagc gtgactggag ttcagacgtg tgctctt 57

REIVINDICACIONES

1. Un procedimiento de determinación de una secuencia de al menos una molécula de ácido nucleico diana en una muestra, que comprende:
- 5 la puesta en contacto de la muestra con al menos una sonda de protección contra la nucleasa flanqueada (NPPF) en condiciones suficientes para que la NPPF se una específicamente a la molécula de ácido nucleico diana, donde la NPPF comprende:
- 10 un extremo 5' y un extremo 3',
una secuencia complementaria a una región de la molécula de ácido nucleico diana, que permite una unión específica entre la NPPF y la molécula de ácido nucleico diana,
una secuencia flanqueante situada en 5', 3' o en ambos, con respecto a la secuencia complementaria a la molécula de ácido nucleico diana, donde la secuencia flanqueante comprende al menos 12 nucleótidos contiguos no
15 presentes en una molécula de ácido nucleico presente en la muestra, que proporciona una secuencia de amplificación universal, y donde la secuencia flanqueante es complementaria a al menos una parte de un cebador de amplificación;
- la puesta en contacto de la muestra con una molécula de ácido nucleico que comprende una secuencia
20 complementaria a la secuencia flanqueante (CFS) en condiciones suficientes para que la secuencia flanqueante se una específicamente a la CFS;
la puesta en contacto de la muestra con una nucleasa específica para moléculas de ácidos nucleicos monocatenarias en condiciones suficientes para eliminar las moléculas de ácidos nucleicos no unidas, generando así una muestra digerida que comprende NPPF hibridadas con la molécula de ácido nucleico diana y con la o las
25 CFS;
la amplificación de NPPF en la muestra digerida con el cebador de amplificación, generando así amplicones de NPPF; y la secuenciación de al menos una parte de los amplicones de NPPF, determinando así la secuencia de la al menos una molécula de ácido nucleico diana en la muestra.
- 30 2. Un procedimiento de detección de al menos una molécula de ácido nucleico diana en una muestra, que comprende:
- la puesta en contacto de la muestra con al menos una sonda de protección contra la nucleasa flanqueada (NPPF) en condiciones suficientes para que la NPPF se una específicamente a la molécula de ácido nucleico diana,
35 donde la NPPF comprende:
- un extremo 5' y un extremo 3',
una secuencia complementaria a una región de la molécula de ácido nucleico diana, que permite una unión específica entre la NPPF y la molécula de ácido nucleico diana,
40 una secuencia flanqueante situada en 5', 3' o en ambos, con respecto a la secuencia complementaria a la molécula de ácido nucleico diana, donde la secuencia flanqueante comprende al menos 12 nucleótidos contiguos no presentes en una molécula de ácido nucleico presente en la muestra que proporciona una secuencia de amplificación universal, y donde la secuencia flanqueante es complementaria a al menos una parte de un cebador de amplificación;
- 45 la puesta en contacto de la muestra con una molécula de ácido nucleico que comprende una secuencia complementaria a la secuencia flanqueante (CFS) en condiciones suficientes para que la secuencia flanqueante se una específicamente a la CFS;
la puesta en contacto de la muestra con una nucleasa específica para moléculas de ácidos nucleicos
50 monocatenarias en condiciones suficientes para eliminar moléculas de ácidos nucleicos no unidas, generando así una muestra digerida que comprende NPPF hibridadas con la molécula de ácido nucleico diana y CFS;
la amplificación de NPPF en la muestra digerida con cebadores de amplificación, generando así amplicones de NPPF; y la detección de amplicones de NPPF, detectando así la al menos una molécula de ácido nucleico diana en la muestra.
- 55 3. El procedimiento de acuerdo la reivindicación 1 ó 2, donde la NPPF comprende una molécula de ADN.
4. El procedimiento de acuerdo con cualquiera de las reivindicaciones 1 a 3, donde la NPPF comprende 35-150 nucleótidos, la secuencia complementaria a una región de la molécula de ácido nucleico diana tiene 10-60

nucleótidos de longitud, la secuencia flanqueante tiene 12-50 nucleótidos de longitud, o combinaciones de los mismos.

5. El procedimiento de acuerdo con cualquiera de las reivindicaciones 1 a 4, donde:
- 5
- i) la NPPF comprende una secuencia flanqueante en el extremo 5' y el extremo 3', donde la secuencia flanqueante en el extremo 5' difiere de la secuencia flanqueante en el extremo 3'; o
 - ii) el al menos un cebador de amplificación comprende además una secuencia que permite la fijación de una etiqueta experimental o un adaptador de secuenciación al amplicón de NPPF durante la etapa de amplificación.
- 10
6. El procedimiento de acuerdo con cualquiera de las reivindicaciones 1 a 5, donde la secuencia flanqueante comprende además una etiqueta experimental, un adaptador de secuenciación, o ambos, donde opcionalmente la etiqueta experimental comprende una secuencia de ácidos nucleicos que permite la identificación de una muestra, sujeto, tratamiento o secuencia de ácidos nucleicos diana.
- 15
7. El procedimiento de acuerdo la reivindicación 5 ó 6, donde el adaptador de secuenciación comprende una secuencia de ácidos nucleicos que permite la captura en una plataforma de secuenciación.
8. El procedimiento de acuerdo con cualquiera de las reivindicaciones 5 a 7, donde la etiqueta experimental o adaptador de secuenciación está presente en el extremo 5' o extremo 3' del amplicón de NPPF.
- 20
9. El procedimiento de acuerdo con cualquiera de las reivindicaciones 1 a 8, donde una o más moléculas de ácidos nucleicos diana son fijas, reticuladas o insolubles.
- 25
10. El procedimiento de acuerdo con cualquiera de las reivindicaciones 1 a 9, donde la NPPF es un ADN y la nucleasa comprende una exonucleasa, una endonucleasa, por ejemplo nucleasa S1, o una combinación de las mismas.
- 30
11. El procedimiento de acuerdo con cualquiera de las reivindicaciones 1 a 10, donde:
- i) el procedimiento secuencia o detecta al menos una molécula de ácido nucleico diana en una pluralidad de muestras simultáneamente; o
 - ii) el procedimiento secuencia o detecta al menos dos moléculas de ácidos nucleicos diana en la misma muestra, y donde la muestra se pone en contacto con al menos dos NPPF diferentes, con cada NPPF específica de una
 - 35 molécula de ácido nucleico diana diferente; o
 - iii) el procedimiento se realiza en una pluralidad de muestras y se detectan al menos dos moléculas de ácidos nucleicos diana en cada una de la pluralidad de muestras; o
 - iv) al menos una NPPF es específica para una molécula de ácido nucleico diana de miRNA y al menos una NPPF es específica para una molécula de ácido nucleico diana de mRNA.
- 40
12. El procedimiento según una cualquiera de las reivindicaciones 1 a 11, que comprende además la lisis de la muestra.
13. El procedimiento de acuerdo con cualquiera de las reivindicaciones 1 ó 3 a 12, que comprende además: la comparación de una secuencia de NPPF obtenida con una base de datos de secuencia de referencia; y la determinación de un número de cada secuencia de NPPF identificada.
- 45
14. El procedimiento según una cualquiera de las reivindicaciones 2 a 13, donde la detección de los amplicones de NPPF comprende:
- 50
- i) la puesta en contacto de los amplicones de NPPF con una superficie que comprende múltiples regiones espacialmente discretas, comprendiendo cada región al menos un anclaje en asociación con un conector bifuncional, donde el conector bifuncional comprende una primera parte que se une específicamente al anclaje y una segunda parte que se une específicamente a al menos una parte de uno de los amplicones de NPPF, en condiciones
 - 55 suficientes para que los amplicones de NPPF se unan específicamente a la segunda parte del conector bifuncional; o
 - ii) la puesta en contacto de los amplicones de NPPF con una superficie que comprende múltiples regiones espacialmente discretas, comprendiendo cada región al menos un anclaje de ácido nucleico que tiene una región complementaria a al menos una parte de uno de los amplicones de NPPF, en condiciones suficientes para que los

- amplicones de NPPF se unan específicamente al anclaje de ácido nucleico; o
- iii) la puesta en contacto de los amplicones de NPPF con una población de superficies, donde la población de superficies comprende subpoblaciones de superficies, y donde cada subpoblación de superficies comprende al menos un anclaje en asociación con un conector bifuncional que comprende una primera parte que se une específicamente a al menos una parte de uno de los amplicones de NPPF, en condiciones suficientes para que los amplicones de NPPF se unan específicamente a la segunda parte del conector bifuncional; o
- iv) la puesta en contacto de los amplicones de NPPF con una población de superficies, donde la población de superficies comprende subpoblaciones de superficies, y donde cada subpoblación de superficies comprende al menos un anclaje de ácido nucleico que tiene una región complementaria a al menos una parte de uno de los amplicones de NPPF, en condiciones suficientes para que los amplicones de NPPF se unan específicamente al anclaje de ácido nucleico, opcionalmente la población de superficies de acuerdo con i)-iv) comprende una población de perlas.
15. El procedimiento de acuerdo con cualquiera de las reivindicaciones 2 a 14, donde la segunda parte del conector bifuncional es complementaria a la región de NPPF complementaria a la región de la molécula de ácido nucleico diana, permitiendo así una unión específica entre el amplicón de NPPF y el conector bifuncional.
16. El procedimiento de acuerdo la reivindicación 15, donde el amplicón de NPPF comprende una etiqueta detectable.
17. El procedimiento según una cualquiera de las reivindicaciones 1 a 16, donde la al menos una NPPF comprende al menos 10 NPPF.

FIG. 1

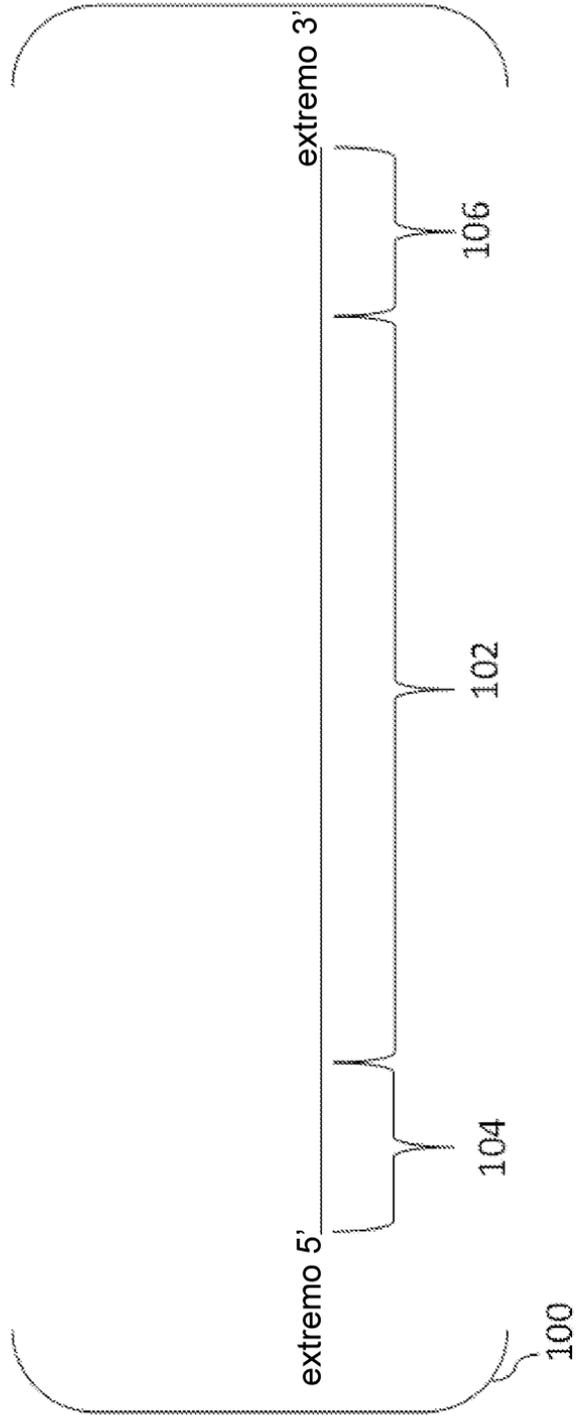


FIG. 2

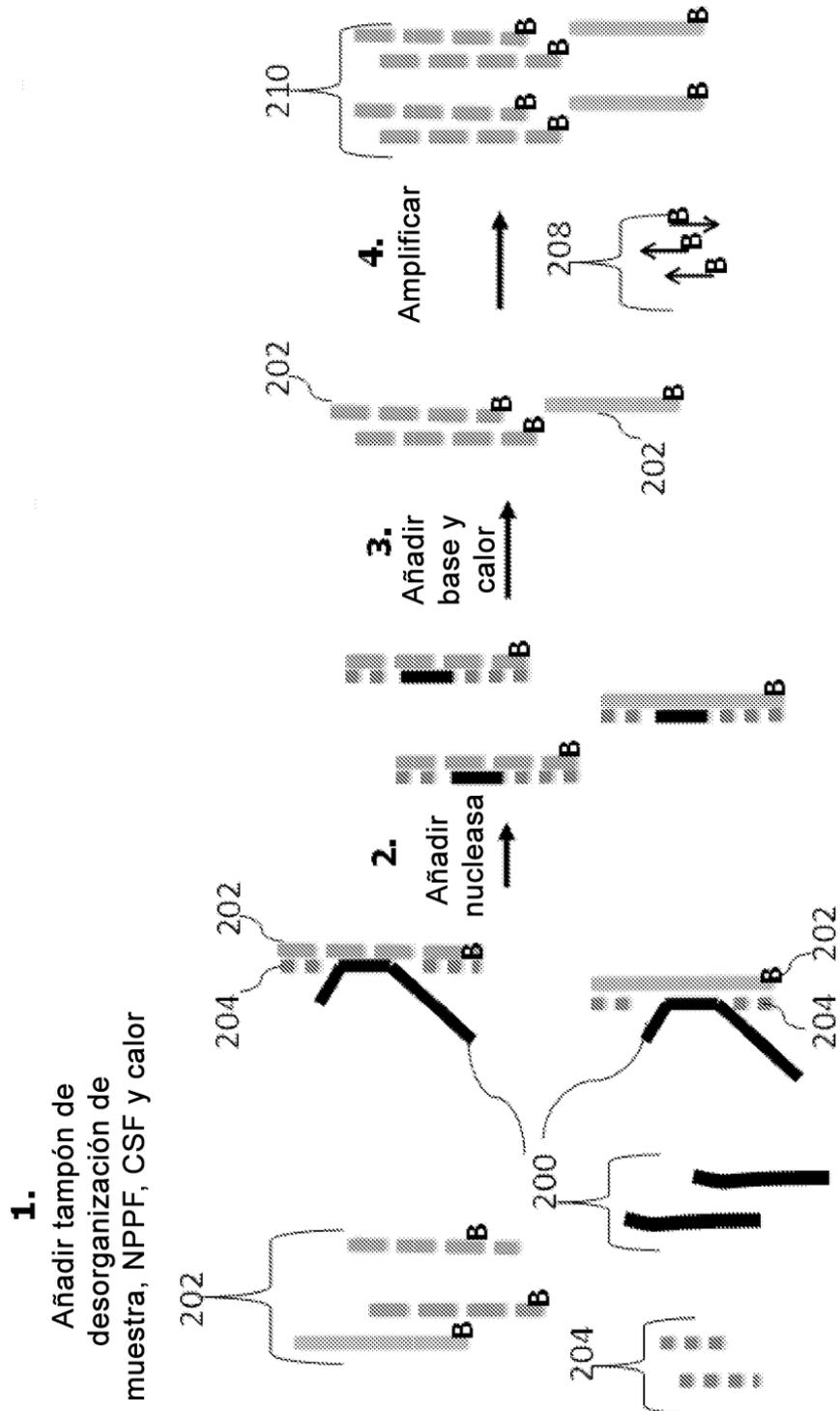


FIG. 3

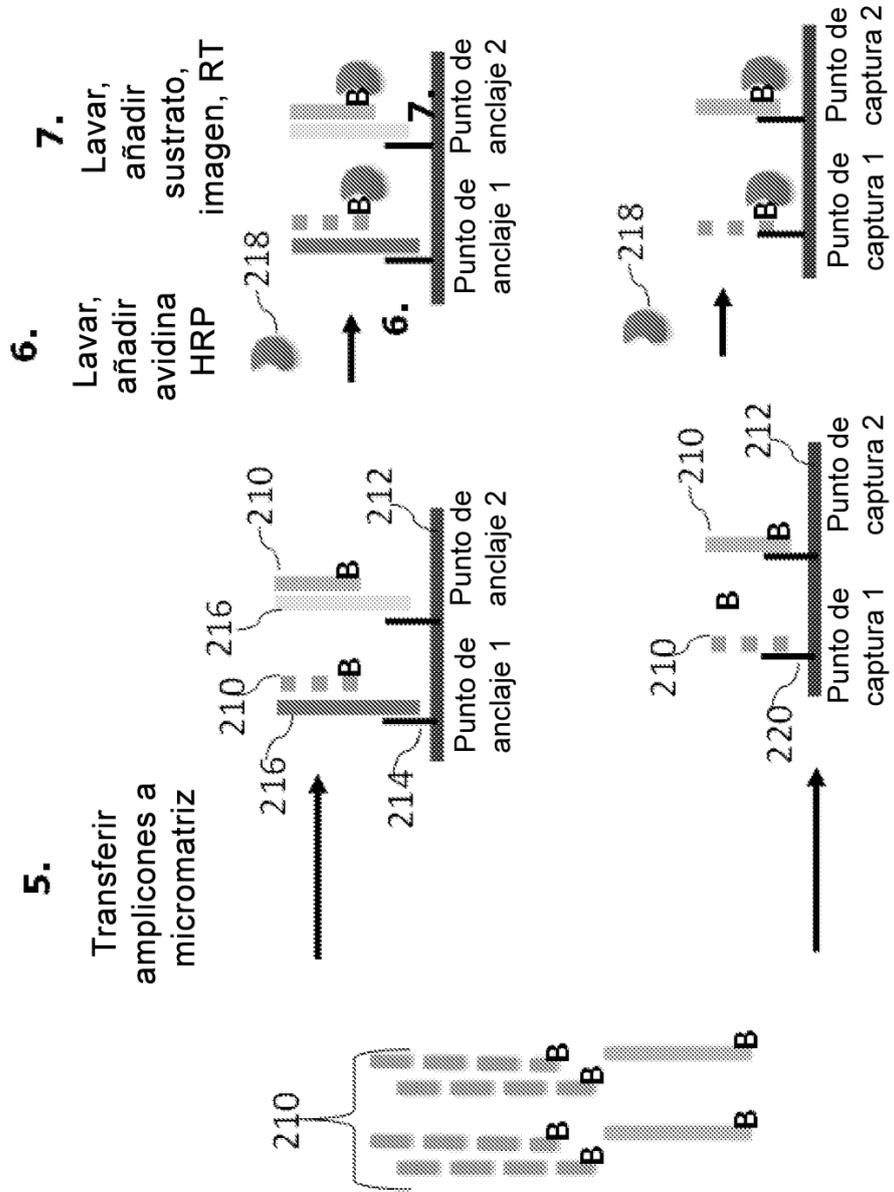


FIG. 4

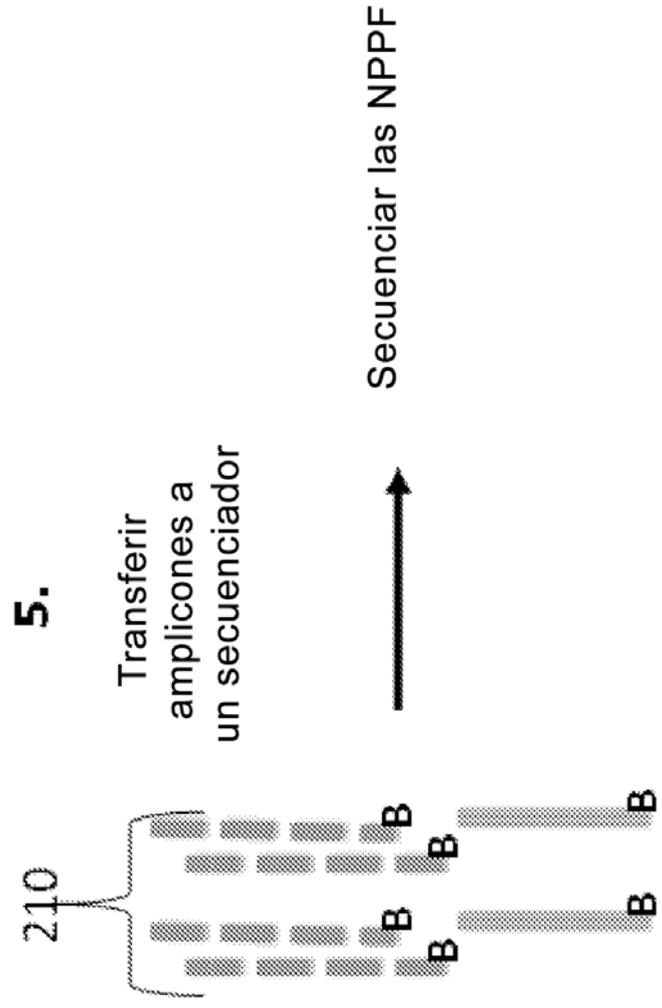


FIG. 5A

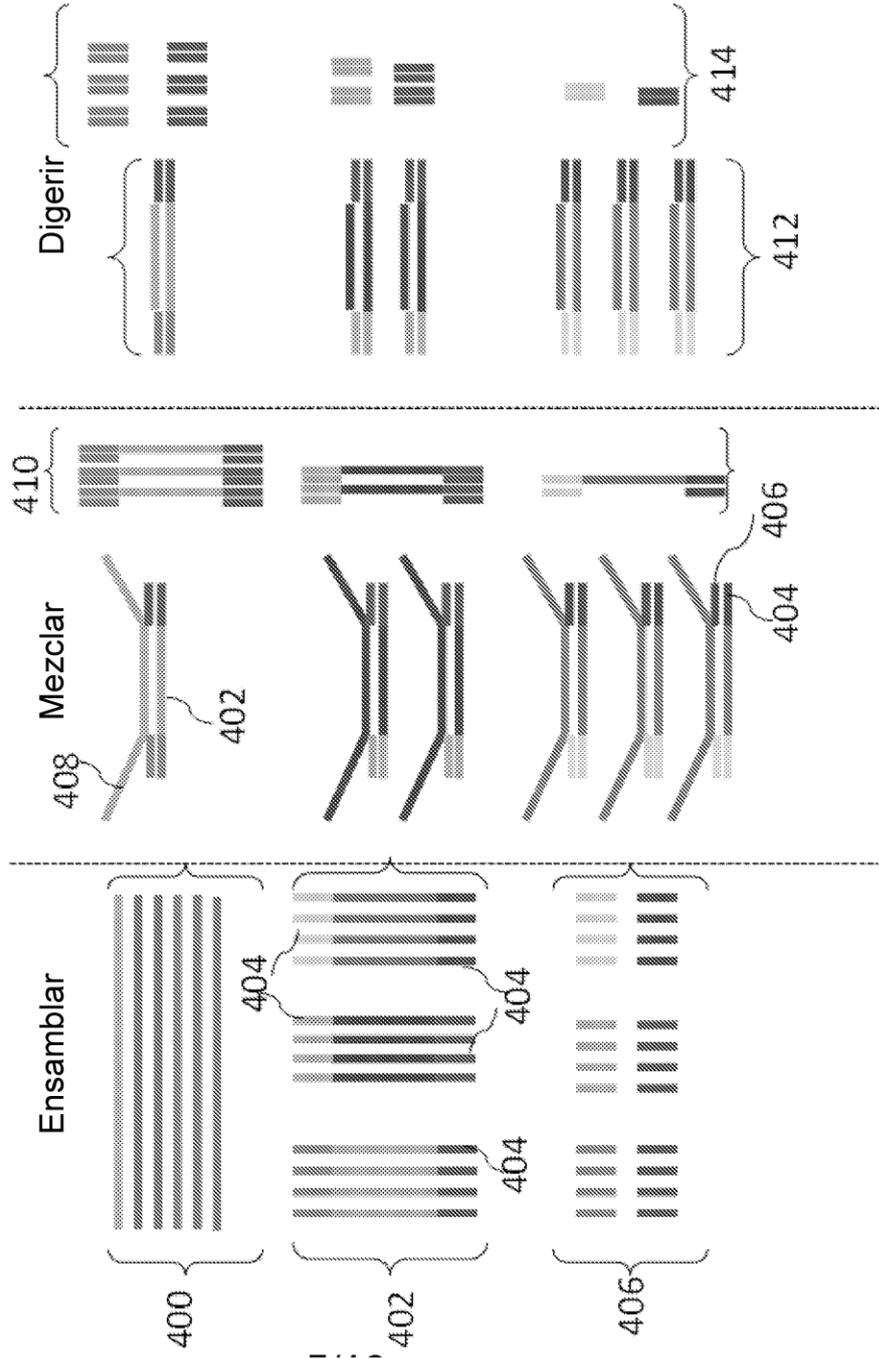


FIG. 5B

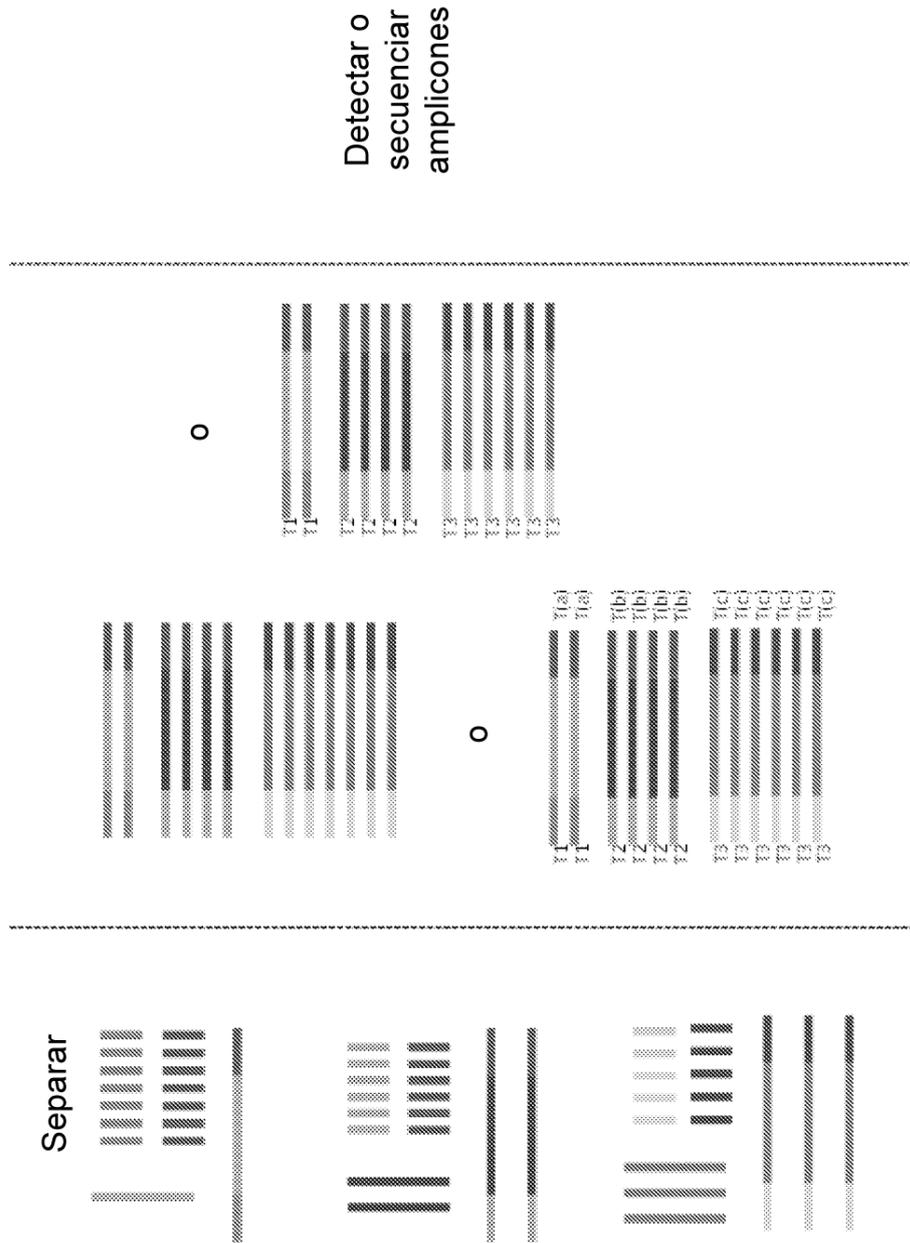
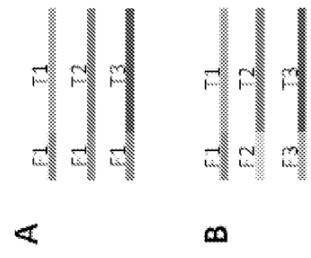


FIG. 6 Realizaciones de dos "alas"

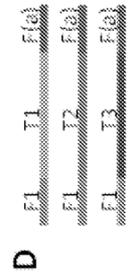
Realizaciones de un "ala"

(Las secuencias flanqueantes pueden estar en el extremo 3' o en el extremo 5')



Todas las secuencias flanqueantes en 5' y 3' son iguales

Todas las secuencias flanqueantes en un extremo (5' o 3') son iguales; Todas las secuencias flanqueantes en el otro extremo (3' o 5', respectivamente) difieren entre sí y con las del otro extremo



Todas las secuencias flanqueantes en cada extremo son iguales, pero la secuencia flanqueante en 5' difiere de la secuencia flanqueante en 3'

Todas las secuencias flanqueantes difieren entre sí con independencia de su posición

Las NPPF también pueden etiquetarse (de forma similar o diferente) en uno o los dos extremos 5' o 3' usando cebadores etiquetados en etapa(s) de amplificación posterior(es)

FIG. 7

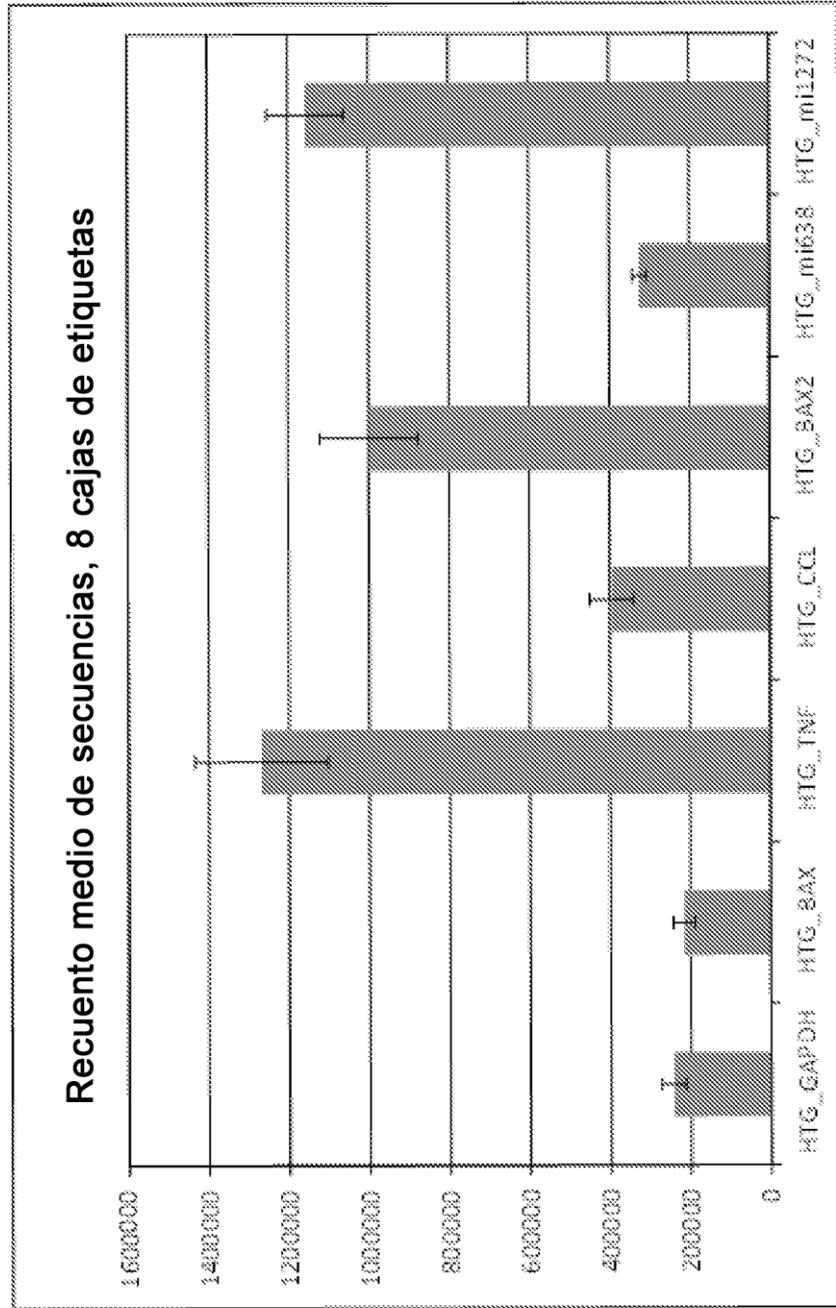


FIG. 8

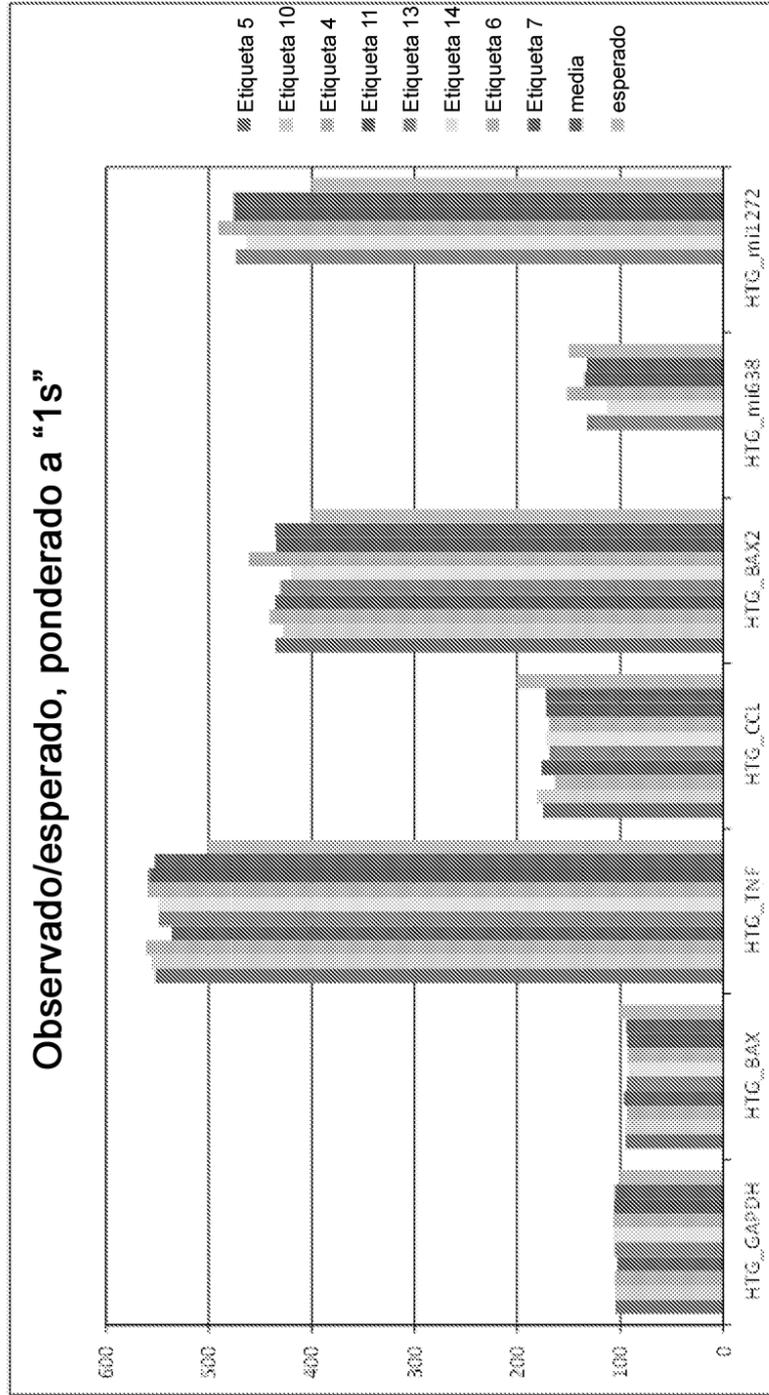
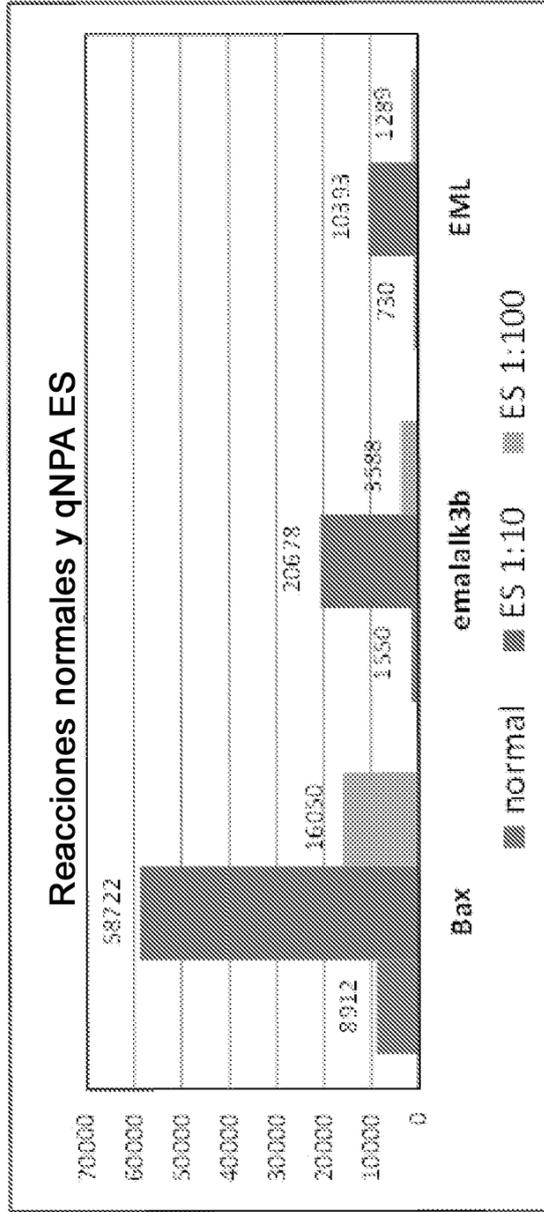


FIG. 9



Gene	normal	ES 1:10	ES 1:100	Incremento de número de veces
BAX	8912	58722	16050	180
EML4-ALK 3b	1550	20678	3588	232
EML4	730	10393	1289	176

	CV (normal)	CV (ES 1:10)	CV (ES 1:100)
Bax	10%	3%	5%
EML4-ALK 3b	31%	10%	17%
EML4	16%	9%	13%

FIG. 10

Etiqueta (conjunto1)	Etiqueta (conjunto2)	Etiqueta (conjunto3)	Lisado para reacción qNPA (mRNA)	Lisado para reacción qNPA (miRNA)
1	9	17	THP1, 50K	
2	10	18	THP1, 25K	HepG2, 50K
3	11		THP1, 10K	HepG2, 10K
4	12		THP1, 5K	HepG2, 5K
5	13		HepG2, 50K	THP1, 50K
6	14	22	HepG2, 25K	THP1, 10K
7	15		HepG2, 10K	THP1, 5K
8	16		HepG2, 5K	

FIG. 11

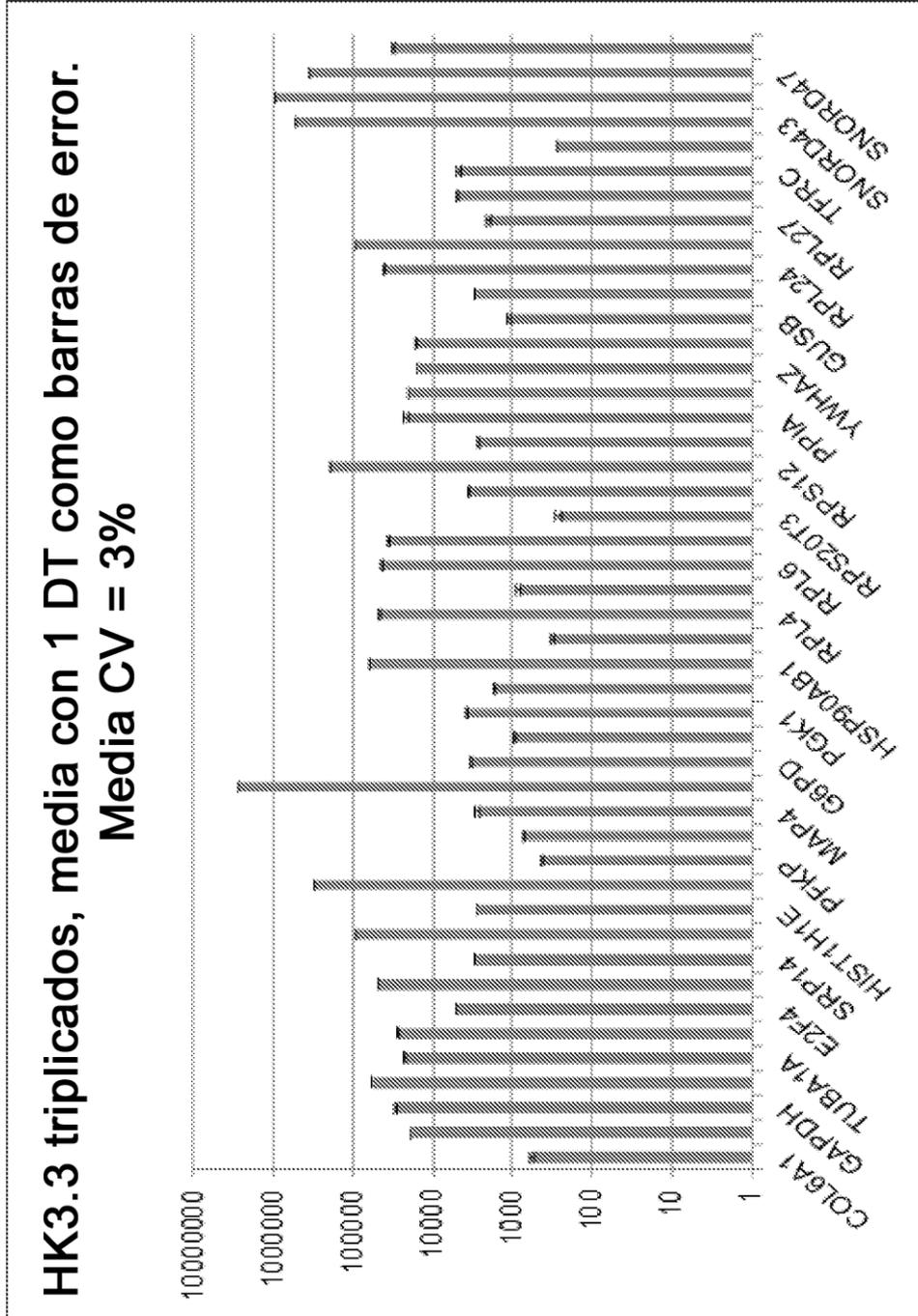


FIG. 12B

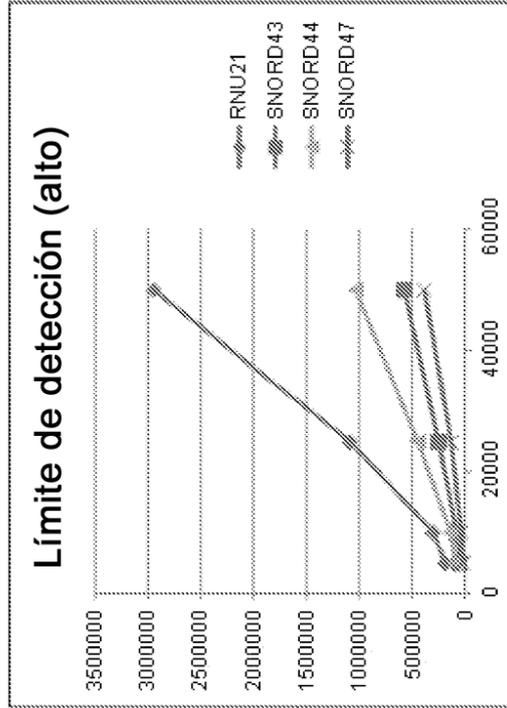


FIG. 12A

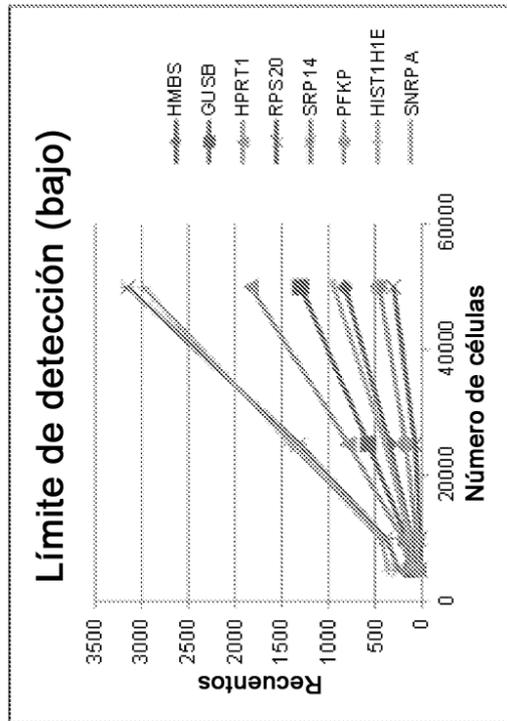


FIG. 13

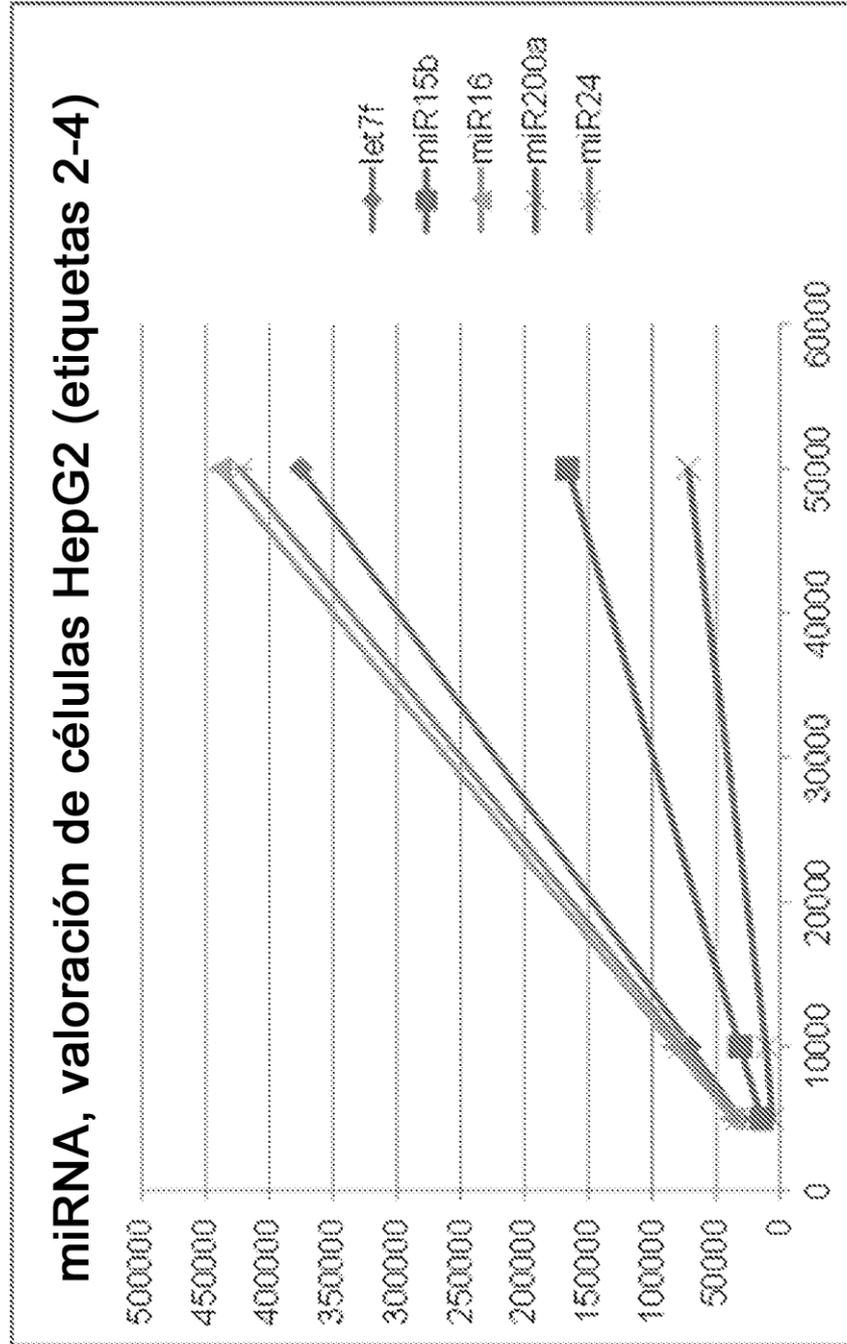


FIG. 14

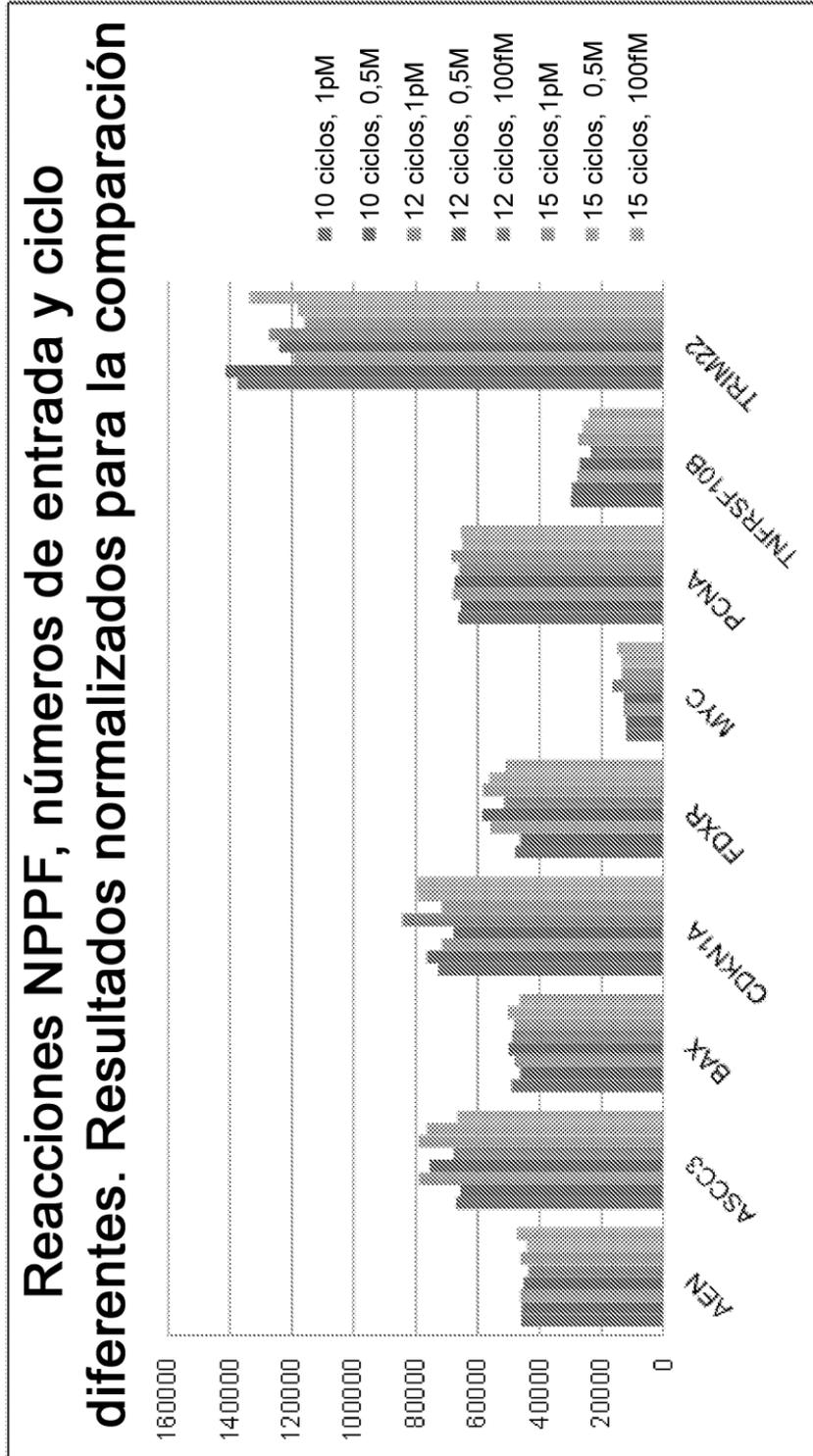


FIG. 15B

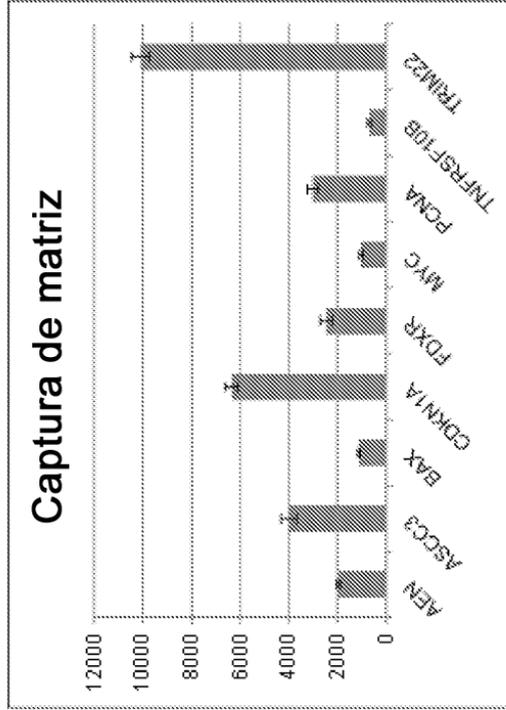


FIG. 15A

