

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 614 810**

51 Int. Cl.:

C12Q 1/68 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **13.08.2009 PCT/US2009/053790**

87 Fecha y número de publicación internacional: **18.02.2010 WO2010019826**

96 Fecha de presentación y número de la solicitud europea: **13.08.2009 E 09791508 (6)**

97 Fecha y número de publicación de la concesión europea: **30.11.2016 EP 2331704**

54 Título: **Nanoindicadores estables**

30 Prioridad:

14.08.2008 US 88988 P

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

02.06.2017

73 Titular/es:

**NANOSTRING TECHNOLOGIES, INC (100.0%)
530 Fairview Ave N, Ste 2000
Seattle, WA 98109, US**

72 Inventor/es:

**WEBSTER, PHILIPPA, J. y
DAHL, TIM**

74 Agente/Representante:

SUGRAÑES MOLINÉ, Pedro

ES 2 614 810 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

Nanoindicadores estables

5 **Antecedentes de la invención**

Aunque todas las células del cuerpo humano contienen el mismo material genético, no se activan los mismos genes en todas esas células. Las alteraciones en los patrones de expresión génica pueden tener profundos efectos sobre las funciones biológicas. Estas variaciones en la expresión génica se encuentran en el núcleo de los procesos fisiológicos y patológicos alterados. Por lo tanto, la identificación y la cuantificación de la expresión de los genes en las células pueden ayudar al descubrimiento de nuevas dianas terapéuticas y de diagnóstico.

Hasta la fecha hay varias técnicas disponibles que permiten la detección del nivel de expresión de múltiples genes a la vez en una muestra compleja. La mayoría de estas tecnologías emplean micromatrices de ADN, dispositivos que consisten de miles de secuencias de ADN inmovilizadas presentes en una superficie miniaturizada que han hecho de este proceso más eficaz. Desafortunadamente, a pesar de la miniaturización de los formatos matriciales, este método todavía requiere cantidades importantes de la muestra biológica. Sin embargo, en varios casos, tales como las biopsias de tejidos o muestras enfermos de un tipo de células discreto, la muestra biológica es limitada. Además, la cinética de hibridación en la superficie de una micromatriz es menos eficaz que la hibridación en pequeñas cantidades de solución acuosa. Por otra parte, si bien existen métodos para estimar la cantidad de ácido nucleico presente en una muestra basándose en el resultado de la hibridación de micromatrices, la tecnología de las micromatrices hasta el momento no permite la detección de moléculas diana a nivel individual, ni existen métodos basados en micromatrices para cuantificar directamente la cantidad de molécula diana de una muestra dada.

Fortina P *et al.* "Digital mRNA Profiling" *Nature Biotechnology*, vol.25, n.º 3, 1 de marzo de 2008, introduce y resume los hallazgos de Geiss *et al.* *Nature Biotechnology*, vol. 26, n.º 3, marzo de 2008. Geiss *et al.* describen el sistema de expresión génica NanoString nCounter®, que consiste en ADN de M13 monocatenario linealizado, denominado estructura principal. Los documentos WO-A2-2007/139766 y WO-A2-2007/076128 proporcionan datos experimentales que se refieren a nanoindicadores que tienen una estructura principal basada en ADN de M13 linealizado.

Existe la necesidad de una detección precisa y sensible, identificación y cuantificación de moléculas diana en mezclas complejas.

35 **Sumario de la invención**

La presente divulgación se refiere, en general, al campo de la detección, la identificación y la cuantificación de moléculas diana en una muestra, e incluye nanoindicadores estables y las poblaciones de estos nanoindicadores, basándose, en parte, en ciertas características de diseño de las secuencias polinucleotídicas de las estructuras principales de los nanoindicadores y las secuencias polinucleotídicas complementarias unidas a las mismas.

La invención proporciona una población de sondas de nanoindicador marcadas de forma única, en la que cada sonda de nanoindicador comprende: i) una región específica de una diana única; e ii) una región que comprende un nanoindicador diseñado, único, en la que el nanoindicador comprende una estructura principal de ácido nucleico monocatenario, diseñada de forma única, comprendiendo la estructura principal al menos tres regiones de unión al marcador unidas covalentemente entre sí en una combinación lineal única, en la que cada región de unión al marcador de cada estructura principal se selecciona individualmente y es diferente de las otras regiones de unión al marcador de esa misma estructura principal, en la que las regiones de unión al marcador se seleccionan de una población de secuencias polinucleotídicas de diseño racional, en la que cada secuencia polinucleotídica se hibrida a una secuencia polinucleotídica complementaria que tiene unidas a la misma una o más moléculas detectables; en la que cada secuencia polinucleotídica complementaria tiene designada una molécula detectable específica; en la que cada región de unión al marcador tiene un contenido de G/C del aproximadamente 50 %, y en la que cada una de dichas secuencias polinucleotídicas complementarias tiene una proporción de G/C de aproximadamente 1/1 u, opcionalmente, de aproximadamente 3/2; y en la que cada nanoindicador tiene una señal detectable que lo distingue de otros nanoindicadores de dicha población. En algunas realizaciones, cada una de las sondas de nanoindicador comprende además una región constante, comprendiendo la región constante una pluralidad de secuencias de nucleótidos repetidas. La población de sondas de nanoindicador puede comprender dos o más sondas de nanoindicador.

En algunas realizaciones, la invención proporciona métodos de determinación de la presencia de al menos una molécula diana en una muestra, que comprenden: (1) formar al menos un complejo molecular que comprende: (a) al menos una molécula diana, y (b) al menos una sonda que comprende una región específica de una diana única y una región que comprende un nanoindicador diseñado, único, en la que dicho nanoindicador comprende una estructura principal de ácido nucleico monocatenario, diseñada de forma única, comprendiendo la estructura principal al menos tres regiones de unión al marcador unidas covalentemente entre sí en una combinación lineal única, en la que cada región seleccionada de unión al marcador se selecciona individualmente y es diferente del

resto de regiones seleccionadas de unión al marcador de la estructura principal, en la que cada región de unión al marcador se hibrida a una secuencia polinucleotídica complementaria que tiene unidas a la misma una o más moléculas detectables, en la que las regiones de unión al marcador se seleccionan de una población de secuencias polinucleotídicas de diseño racional, en la que cada secuencia polinucleotídica complementaria tiene designada una molécula detectable específica; en la que cada región de unión al marcador tiene un contenido de G/C del aproximadamente 50 %, y la secuencia polinucleotídica complementaria tiene una proporción de G/C de al menos 1/1, y (2) contar individualmente la presencia de uno o más complejos moleculares o al menos parte del al menos un complejo molecular para determinar la presencia de dicha al menos una molécula diana en la muestra. En algunas realizaciones, el porcentaje de recuentos moleculares válidos del complejo molecular es superior al aproximadamente 12,5 %. En algunas realizaciones, el número de recuentos es al menos dos veces superior a los recuentos obtenidos cuando se usa una sonda de nanoindicador que comprende ADN de M13. Un nanoindicador que comprende ADN de M13 puede comprender una estructura principal monocatenaria que comprenda una pluralidad de regiones de ADN de M13 unidas covalentemente entre sí, en la que cada región se hibride a una secuencia polinucleotídica complementaria que tenga unidas a la misma una o más moléculas detectables.

En algunas realizaciones de los métodos y la composición de la invención, los números de recuentos por encima del fondo de cada uno de los complejos moleculares tras la normalización de la muestra es al menos dos veces superior si se compara con una sonda de nanoindicador que comprende ADN de M13. En algunas realizaciones, la Tf de las secuencias polinucleotídicas complementarias cuando están hibridadas a sus regiones de unión al marcador es de aproximadamente 80 °C o superior. En algunas realizaciones, la temperatura de fusión (Tf) de las secuencias polinucleotídicas complementarias cuando están hibridadas a sus regiones de unión al marcador es superior a la Tf de las secuencias polinucleotídicas complementarias al ADN de M13 cuando están hibridadas a la sonda de nanoindicador que comprende ADN de M13.

En algunas realizaciones, los métodos de la invención comprenden además determinar la presencia de una pluralidad de moléculas diana mediante un método que comprende formar una pluralidad complejos moleculares, comprendiendo cada complejo (a) al menos una molécula diana y (b) al menos una sonda que comprende una región específica de una diana única y una región que comprende un nanoindicador diseñado, único, en la que cada nanoindicador comprende una estructura principal de ácido nucleico monocatenario, diseñada de forma única, comprendiendo la estructura principal al menos tres regiones de unión al marcador unidas covalentemente entre sí en una combinación lineal única, en la que cada región seleccionada de unión al marcador se selecciona individualmente y es diferente del resto de regiones seleccionadas de unión al marcador de la estructura principal, en la que las regiones de unión al marcador se seleccionan de una población de secuencias polinucleotídicas de diseño racional, en la que cada región de unión al marcador se hibrida a una secuencia polinucleotídica complementaria que tiene unidas a la misma una o más moléculas detectables, y en la que cada sonda comprende una región de nanoindicador diferente y en la que cada región de unión al marcador tiene un contenido de G/C del aproximadamente 50 %, y la secuencia polinucleotídica complementaria tiene una proporción de G/C de al menos 1/1. En algunas realizaciones, se determina la presencia de al menos 5, 10, 20, 30, 50, 100, 200, 300 o 500 moléculas diana diferentes. En algunas realizaciones, la molécula diana es un ácido nucleico. En algunas realizaciones, el ácido nucleico comprende al menos una mutación heredable, al menos una mutación somática, al menos un polimorfismo de un solo nucleótido (SNP), al menos una mutación puntual, al menos una mutación de eliminación, al menos una mutación de inserción, al menos una translocalización cromosómica o combinaciones de las mismas. En algunas realizaciones, la molécula diana es un indicador de diagnóstico.

En algunas realizaciones, la invención proporciona, una sonda de nanoindicador marcada de forma única que comprende: i) una región específica de una diana única; y ii) una región de una estructura principal de ácido nucleico monocatenario, diseñada de forma única, que comprende al menos tres regiones diseñadas de unión al marcador unidas covalentemente entre sí en una combinación lineal, en la que cada región de unión al marcador comprende de aproximadamente 800 a 1.300 bases de nucleótido y tiene un contenido de G/C del aproximadamente 50 %, en la que cada región seleccionada de unión al marcador se selecciona individualmente y es diferente del resto de regiones seleccionadas de unión al marcador de esa misma sonda, en la que las regiones de unión al marcador se seleccionan de una población de secuencias polinucleotídicas de diseño racional, y en la que cada región de unión al marcador se hibrida a una secuencia polinucleotídica complementaria que tiene unidas a la misma una o más moléculas detectables, en la que la secuencia polinucleotídica complementaria tiene una proporción de G/C de al menos 1/1. En algunas realizaciones, la secuencia polinucleotídica complementaria tiene una proporción de G/C de aproximadamente 3/2. En algunas realizaciones, la sonda de nanoindicador comprende además una región constante, comprendiendo la región constante una pluralidad de secuencias de nucleótidos repetidas.

En algunas realizaciones de los métodos y las composiciones de la invención, la región de unión al marcador comprende de aproximadamente 800 a 1.300 bases de nucleótido y tiene un contenido de G/C del aproximadamente 50 % y la secuencia polinucleotídica complementaria tiene una proporción de G/C de aproximadamente 1/1. En algunas realizaciones, la secuencia polinucleotídica complementaria tiene una proporción de G/C de aproximadamente 3/2. Los ejemplos de moldes que se pueden usar para generar las regiones de unión al marcador descritas en el presente documento incluyen, pero sin limitación, secuencias seleccionadas del grupo que consiste en SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23 y 24.

En algunas realizaciones de los métodos y las composiciones de la invención, las regiones de unión al marcador comprenden un contenido de adenina similar. En algunas realizaciones, las bases de adenina están separadas al menos una media de cada 8 a 16 bases de nucleótido. En algunas realizaciones, las regiones de unión al marcador comprenden un patrón repetido de forma regular de bases de adenina. En algunas realizaciones, las bases de adenina están separadas aproximadamente cada 8 a 16 bases de nucleótido. En algunas realizaciones, las regiones de unión al marcador comprenden un contenido de timina del aproximadamente 35-45 %.

En algunas realizaciones de los métodos y las composiciones de la invención, la secuencia polinucleotídica complementaria comprende una secuencia polinucleotídica de ARN. La secuencia polinucleotídica de ARN puede comprender al menos una base de uracilo modificada con aminoalilo. En algunas realizaciones, la molécula detectable de la secuencia polinucleotídica complementaria está unida a la base de uracilo modificada con aminoalilo. En algunas realizaciones, la secuencia polinucleotídica de ARN comprende una pluralidad de bases de uracilo modificadas con aminoalilo que están separadas a aproximadamente una media de cada 8 a 16 bases en dicha secuencia polinucleotídica de ARN. En algunas realizaciones, la molécula detectable está unida a cada una de las bases de uracilo modificadas con aminoalilo.

En algunas realizaciones de las composiciones y los métodos de la invención, las moléculas detectables son colorantes fluorescentes.

En algunas realizaciones, la invención proporciona métodos de preparación de al menos un nanoindicador marcado de forma única que comprende una estructura principal de ácido nucleico monocatenario diseñada de forma única, incluyendo el método las etapas de: i) combinar al menos tres regiones de unión al marcador, comprendiendo cada una de aproximadamente 800 a 1.300 bases de nucleótido y un contenido de G/C del aproximadamente 50 %, estando cada región de unión al marcador seleccionada individualmente y siendo diferente del resto de regiones seleccionadas de unión al marcador del mismo nanoindicador, ii) unir covalentemente las al menos tres regiones de unión al marcador entre sí en una combinación lineal; y iii) hibridar una secuencia polinucleotídica complementaria a dicha región de unión al marcador, en la que dicha secuencia polinucleotídica complementaria tiene unidas a la misma una o más moléculas detectables, y en la que la secuencia polinucleotídica complementaria tiene una proporción de G/C de al menos 1/1, y en la que las regiones de unión al marcador se seleccionan de una población de secuencias polinucleotídicas de diseño racional. En algunas realizaciones, los métodos comprenden además preparar una sonda de nanoindicador marcada mediante la unión del nanoindicador marcado a una región específica diana.

En algunas realizaciones, la invención proporciona kits de preparación de al menos un nanoindicador marcado de forma única que comprende una estructura principal de ácido nucleico monocatenario, diseñada de forma única, kit que incluye:

a) al menos tres regiones de unión al marcador, comprendiendo cada una de aproximadamente 800 a 1.300 bases de nucleótido, un contenido de G/C del aproximadamente 50 %, estando cada región de unión al marcador del nanoindicador seleccionada individualmente y siendo diferente de las otras regiones de unión al marcador del mismo nanoindicador; y b) al menos tres secuencias polinucleotídicas complementarias que tienen unida a las mismas una molécula detectable, en la que las secuencias polinucleotídicas complementarias tienen una proporción de G/C de al menos 1/1, y en la que las regiones de unión al marcador se seleccionan de una población de secuencias polinucleotídicas de diseño racional. En algunas realizaciones, los kits comprenden además al menos tres sondas específicas de la diana.

En algunas realizaciones, la invención proporciona kits que comprenden una población de nanoindicadores según lo descrito en el presente documento e instrucciones para su uso.

Breve descripción de las figuras

Las características novedosas de la invención se exponen con particularidad en las reivindicaciones adjuntas. Se obtendrá una mejor comprensión de las características y ventajas de la presente invención por referencia a la siguiente descripción detallada y a las figuras adjuntas, de las cuales:

La **Figura 1** es una ilustración esquemática de un nanoindicador doble con un código de nanoindicador en la posición 6, usando una sonda de captura y un único componente nanoindicador en la posición 6. La flecha es ilustrativa de un marcador de afinidad, que está opcionalmente incluido y que se puede usar para purificar el nanoindicador o inmovilizar el nanoindicador (o el complejo de nanoindicador-molécula diana) con el fin de formar imágenes.

Figura 2A-2E. La Figura 2A es una ilustración esquemática del experimento mostrado en las Figuras 2B y 2C. En este caso, el diamante representa la biotina que se usó para unir el complejo por un extremo a la superficie antes del estiramiento. Las Figuras 2B y 2C muestran imágenes de experimentos en los que se hibridaron la sonda de captura S2-A, el nanoindicador marcado S2-B y el ADN diana S2 (Figura 2B) o ARN diana S2 (Figura 2C). La Figura 2E muestra un primer plano de los complejos de nanoindicador de la Figura 2B, conteniendo cada uno sonda de captura S2-A, nanoindicador marcado S2-B y ADN diana S2. La Figura 2D muestra una imagen de

un experimento de control negativo, en el que se hibridaron la sonda de captura S2-A, el nanoindicador marcado S2-B y ningún ADN diana S2.

Figura 3A-3B. Muestra un nanoindicador marcado (a base de ácido nucleico) con un marcador de afinidad, A1, en un extremo. En la Figura 3, el nanoindicador marcado se inmoviliza a través de la unión de A1 a una pareja de afinidad inmovilizada. El otro extremo del nanoindicador está en solución (Figura 3A), pero se puede inmovilizar mediante hibridación a un oligonucleótido complementario que contenga otro marcador de afinidad (A2) usado para inmovilizar el nanoindicador (Figura 3B). A1 y A2 pueden ser los mismos, por ejemplo, biotina, que para la inmovilización en una superficie recubierta de avidina o estreptavidina. Tras la inmovilización de A1, el nanoindicador se puede estirar o "alargar" como se representa en la Figura 3, por ejemplo, por electroestiramiento, para la separación de las regiones de unión al marcador de una manera que permita la detección del código de nanoindicador. Opcionalmente, mientras que el nanoindicador está en un estado alargado, se introduce A2 y se une el extremo del nanoindicador que es complementario a A2 en sentido descendente hacia la superficie.

Figura 4A-4C. La Figura 4A ilustra la inmovilización de un extremo de una molécula de ADN en un dispositivo microfluídico; la Figura 4B ilustra la extensión del ADN en un campo eléctrico; y la Figura 4C ilustra la inmovilización selectiva de un segundo extremo de la molécula de ADN extendido por los marcadores de afinidad introducidos en el dispositivo tras el alargamiento.

La **Figura 5** muestra la detección de IL-8 en la misma muestra usando ambos sistemas nanoindicadores DV1 y M13. Los datos se recogieron de experimentos realizados de acuerdo con el Ejemplo 4.

La **Figura 6** muestra una comparación de los sistemas DV1 y M13 en la medición de la expresión de 40 genes en dos muestras. Los datos se recogieron de dos experimentos representativos realizados de acuerdo con el Ejemplo 4.

La **Figura 7** muestra la detección de la expresión de GusB. En este experimento, se usaron bibliotecas de M13 y DV1 que contenían 148 sondas idénticas para medir los niveles de expresión génica en 26 muestras de ARN de ratón. Se observó un aumento medio de 6 veces del número absoluto de recuentos medidos por los indicadores DV1 en comparación con los indicadores M13. Los datos se recogieron de experimentos representativos realizados de acuerdo con el Ejemplo 5.

La **Figura 8** muestra un gráfico de "indicadores válidos", que se refiere a cadenas de manchas que pueden ser interpretadas por el software de análisis de imágenes como indicador "contable". Los resultados se describen como un porcentaje del número total de uniones en un campo de visión. Un porcentaje significativamente superior de los indicadores DV1 son contables en este experimento, conduciendo a un aumento del triple en los datos (del 12,5 % al 38 %).

La **Figura 9** muestra la secuencia polinucleotídica de un vector de plásmido que se puede utilizar para clonar, propagar y generar las estructuras principales nanoindicadoras monocatenarias de la presente invención.

La **Figura 10** muestra las secuencias polinucleotídicas de dos moldes que se utilizaron en los experimentos de optimización de colorante para generar estructuras principales nanoindicadoras monocatenarias (véase el Ejemplo 6). En estas secuencias molde, la base repetida de manera regular es la timina, que, tras la transcripción, produce una estructura principal nanoindicadora monocatenaria complementaria que tiene la adenina como la base repetida de forma regular. La Figura 10A muestra la secuencia polinucleotídica de un molde que tiene una base repetida de forma regular en aproximadamente cada 8 nucleótidos. La Figura 10B muestra la secuencia polinucleotídica de un molde que tiene una base repetida de forma regular en aproximadamente cada 10 nucleótidos.

La **Figura 11** muestra las secuencias polinucleotídicas de dos moldes que se utilizaron en los experimentos de optimización de colorante para generar estructuras principales nanoindicadoras monocatenarias (véase el Ejemplo 6). La Figura 11A muestra la secuencia polinucleotídica de un molde que tiene la base repetida de forma regular en aproximadamente cada 12 nucleótidos. La Figura 11B muestra la secuencia polinucleotídica de un molde que tiene una base repetida de forma regular en aproximadamente cada 14 nucleótidos.

La **Figura 12** muestra las secuencias polinucleotídicas de dos moldes que se utilizaron en los experimentos de optimización de colorante para generar estructuras principales nanoindicadoras monocatenarias (véase el Ejemplo 6). La Figura 12A muestra la secuencia polinucleotídica de un molde que tiene una base repetida de forma regular en aproximadamente cada 16 nucleótidos. La Figura 12B muestra la secuencia polinucleotídica de un molde que tiene una base repetida de forma regular en aproximadamente cada 24 nucleótidos.

Descripción detallada de la invención

La presente divulgación proporciona composiciones y métodos de detección y cuantificación de moléculas diana individuales en muestras biomoleculares, incluyendo nanoindicadores estables que son capaces de unirse a moléculas diana individuales y que proporcionan una mejor detección de las moléculas diana. A través de los códigos de marcaje de los nanoindicadores, la unión de las sondas de nanoindicador a moléculas diana da como resultado la identificación de las moléculas diana. También se proporcionan métodos de fabricación y uso de dichos nanoindicadores. Los nanoindicadores se pueden usar en una amplia variedad de aplicaciones tales como aplicaciones de diagnóstico, pronóstico, control de calidad y detección.

La divulgación también se refiere a la selección de una biblioteca o una población de nanoindicadores diseñados (por ejemplo, secuencias sintéticas). Más concretamente, ciertas características de secuencia optimizadas proporcionan una mejor estabilidad molecular de un nanoindicador, así como una mejor detección cuando el

nanoindicador se combina con una secuencia específica de la diana. Por ejemplo, los métodos y las composiciones de la invención proporcionan nanoindicadores que comprenden estructuras principales sintéticas únicas que no producen una estructura secundaria y producen brillo uniforme.

5 La presente divulgación proporciona una población de nanoindicadores diseñados (por ejemplo, sintéticos), en la que dicho nanoindicador comprende una pluralidad de diferentes moléculas detectables, y en la que la pluralidad de diferentes moléculas detectables de cada nanoindicador tiene una señal detectable que lo distingue de otros nanoindicadores de dicha población. Así pues, la divulgación proporciona una población de nanoindicadores diseñados con una mejor estabilidad molecular, en la que cada nanoindicador de la población tiene una señal detectable que lo distingue de otros nanoindicadores de dicha población.

10 Los nanoindicadores pueden comprender estructuras principales de diseño único (por ejemplo, sintéticas) hibridadas a una secuencia polinucleotídica complementaria única que tenga unido a la misma un marcador detectable. Cada nanoindicador generará una señal única que no cambiará en el transcurso de un ensayo de detección. Se trata de la señal o del código de la señal única y distinguible asociada con el nanoindicador que permanecerá invariable en el transcurso de un ensayo de detección. En algunas realizaciones, la secuencia polinucleotídica complementaria que tiene unido a la misma el marcador detectable puede ser un segmento de ARN acoplado a un colorante, transcrito *in vitro*, que tenga un brillo uniforme. El brillo uniforme como se describe en el presente documento se refiere a la fuerza, al tamaño y/o a la intensidad de la señal producida por los segmentos de ARN acoplados al colorante. Es decir, en algunas realizaciones, la fuerza, el tamaño y/o la intensidad de la señal del colorante serán similares entre los segmentos de ARN acoplados al colorante que tengan el mismo colorante unido a los mismos. Por ejemplo, los segmentos de ARN acoplados al colorante que tienen un colorante de color verde unido a los mismos tendrán una señal similar o una señal con la misma intensidad y/o el mismo brillo. Esto es útil, entre otras cosas, porque permite que un software/algorithmo de análisis de imágenes o un usuario defina a qué mancha de un determinado color se parecería en términos de brillo y tamaño. Además, varios de los colorantes pueden migrar a los canales de los otros colorantes, si la fuerza de la señal es constante, entonces la migración a través de la señal también será constante, pudiéndose así ignorar la migración. Por otra parte, los ruidos de las imágenes que no entran dentro de los parámetros de las manchas pueden pasarse por alto. Como resultado de ello, por ejemplo, se pueden definir estrechamente los parámetros asociados con las señales de los nanoindicadores, permitiendo que el software/algorithmo o usuario haga caso omiso de un porcentaje más alto del ruido. Estos proporcionan lecturas más robustas y fiables.

15 En algunas realizaciones, los nanoindicadores comprenden una estructura principal con una disposición de regiones de unión al marcador basadas en nucleótidos, en la que cada región de unión al marcador tiene una secuencia específica designada para un marcador específico. En este sistema, la secuencia única de la estructura principal dicta el código de color del nanoindicador. Cada estructura principal se hibridará solo a la secuencia polinucleotídica complementaria a su secuencia, cada una de las cuales tiene un marcador designado específico (por ejemplo, un color). Así pues, cada estructura principal generará solo el código designado, por ejemplo, incluso si la secuencia polinucleotídica se separa de la estructura principal durante el proceso de detección. Si la secuencia polinucleotídica del nanoindicador se separa durante la síntesis o durante un procedimiento de hibridación, solo puede ser reemplazada por la secuencia polinucleotídica del mismo color, eliminando la posibilidad de marcadores compartidos o intercambiados. La secuencia polinucleotídica complementaria que tiene unido a la misma el marcador detectable puede ser un segmento de ARN acoplado a un colorante, transcrito *in vitro*. En algunas realizaciones, en la construcción de una estructura principal del nanoindicador dada de una biblioteca de moldes de región de unión al marcador basados en polinucleótidos individualmente únicos, se asigna a cada región de unión al marcador un marcador detectable (por ejemplo, una molécula detectable), y cada región de unión al marcador de una estructura principal dada se selecciona para que sea diferente de las otras regiones de unión al marcador de la misma estructura principal.

20 En algunas realizaciones, las secuencias se diseñan para tener una distribución uniforme de una base, lo que permite una distribución uniforme de los marcadores acoplados (por ejemplo, fluorescentes) cuando se introduce esta base como un nucleótido modificado en un ARN o ADN polimerizado *in vitro*. En algunas realizaciones, las secuencias se diseñan sin repeticiones directas o inversas significativas para fabricar los nanoindicadores lo más desestructurados y únicos como sea posible. En algunas realizaciones, las secuencias se diseñan sin ninguna repetición directa o inversa de 9 nucleótidos o más a lo largo de una secuencia que 1.100 pares de bases de longitud. En algunas realizaciones, las secuencias se diseñan sin ninguna repetición directa o inversa de 7 nucleótidos o más a lo largo de cualquier región de 100 pares de bases. En algunas realizaciones, las regiones de unión al marcador y/o la secuencia polinucleotídica complementaria comprenden un contenido y una proporción particulares de G/C.

25 Los ejemplos de moldes de polinucleótidos que se pueden utilizar para generar estas regiones de unión al marcador diseñadas se exponen en las secuencias polinucleotídicas de SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23 y 24.

30 Por consiguiente, ciertos aspectos de la invención reivindicada proporcionan una población de nanoindicadores o sondas de nanoindicador únicos, cada uno compuesto de una estructura principal a base de polinucleótidos única,

en la que cada nanoindicador de la población no solo es distinto de los otros nanoindicadores de la población, sino que también tiene una mejor estabilidad molecular y una señal detectable que lo distingue de otros nanoindicadores de dicha población. En algunas realizaciones, cada sonda de nanoindicador comprende una pluralidad de regiones de unión al marcador diseñadas, individuales. En algunas realizaciones, cada una de las regiones de unión al marcador es diferente de las otras regiones de unión al marcador de ese mismo nanoindicador. Por lo tanto, la invención reivindicada proporciona una población más estable de nanoindicadores únicos que tiene propiedades de detección mejoradas. En el presente documento, se describe un nanoindicador ilustrativo que tiene dichas características, denominado nanoindicador DV1.

10 **Nanoindicadores**

Una sonda de nanoindicador totalmente montada y marcada comprende dos partes principales, una secuencia específica de la diana que es capaz de unirse a una molécula diana, y un nanoindicador marcado que proporciona un "código" de señales (el "código de nanoindicador") asociado con la secuencia específica de la diana.

Tras la unión de la sonda de nanoindicador a la molécula diana, el código del nanoindicador identifica la molécula diana a la que se une el nanoindicador.

Los nanoindicadores son estructuras modulares. En algunas realizaciones, el nanoindicador comprende una pluralidad de diferentes moléculas detectables. En algunas realizaciones, un nanoindicador marcado es una entidad molecular que contiene ciertos elementos básicos: (i) una pluralidad de regiones únicas de unión al marcador unidas en una determinada combinación lineal única; e (ii) secuencias polinucleotídicas complementarias unidas a las regiones de unión al marcador de la estructura principal. En algunas realizaciones, el nanoindicador marcado comprende 2, 3, 4, 5, 6, 7, 8, 9, 10 o más regiones únicas de unión al marcador, unidas en una determinada combinación lineal única, y secuencias polinucleotídicas complementarias unidas a las regiones de unión al marcador de la estructura principal. En algunas realizaciones, el nanoindicador marcado comprende 3 o más regiones únicas de unión al marcador unidas en una determinada combinación lineal única y secuencias polinucleotídicas complementarias unidas a las regiones de unión al marcador de la estructura principal. En algunas realizaciones, el nanoindicador marcado comprende 6 o más regiones únicas de unión al marcador unidas en una determinada combinación lineal única, y secuencias polinucleotídicas complementarias unidas a las regiones de unión al marcador de la estructura principal. Una sonda de nanoindicador comprende además una secuencia específica de la diana, también unida a la estructura principal.

La expresión "región de unión al marcador" incluye una región de secuencia polinucleotídica definida dentro de una estructura principal dada que puede servir como un punto de unión individual para una molécula detectable. En algunas realizaciones, las regiones de unión al marcador comprenden secuencias diseñadas. Los ejemplos específicos de secuencias polinucleotídicas molde definidas que se pueden utilizar para generar regiones de unión al marcador monocatenarias incluyen las secuencias molde expuestas en SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23 y 24, además de las variantes adecuadas de las mismas (por ejemplo, secuencias que tienen un 60 %, 70 %, 80 %, 90 %, 95 %, 98 % o más, incluyendo todos los números enteros intermedios, de identidad de secuencia). La transcripción de las secuencias molde de SEQ ID NO: 1 a 24 genera regiones de unión al marcador monocatenarias que tienen una secuencia polinucleotídica que es complementaria a las secuencias de SEQ ID NO: 1-24.

En algunas realizaciones, el nanoindicador marcador también comprende una estructura principal que contiene una región constante. La expresión "región constante" incluye secuencias repetidas en tándem de aproximadamente 10 a aproximadamente 25 nucleótidos que están unidas covalentemente a un nanoindicador. La región constante puede estar unida bien a la región 5' o a la región 3' de un nanoindicador, y se puede utilizar para la captura y la inmovilización de un nanoindicador para la formación de imágenes o la detección, tal como mediante la unión a un sustrato sólido de una secuencia que sea complementaria a la región constante. En ciertos aspectos, la región constante contiene 2, 3, 4, 5, 6, 7, 8, 9, 10, o más secuencias repetidas en tándem, en la que cada una de las secuencias de repetición comprende aproximadamente 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 o más nucleótidos, incluyendo aproximadamente 12-18, 13-17 o aproximadamente 14-16 nucleótidos.

Los nanoindicadores descritos en el presente documento comprenden secuencias diseñadas sintéticas. En algunas realizaciones, los nanoindicadores descritos en el presente documento proporcionan una mayor uniformidad en el brillo del marcador. Un brillo uniforme, como se describe en el presente documento, se refiere a la fuerza, al tamaño y/o a la intensidad de la señal producida por los segmentos marcados. Es decir, en algunas realizaciones, la fuerza, el tamaño y/o la intensidad de la señal de los segmentos marcados serán similares entre los segmentos que tengan el mismo marcador unido a los mismos. Como se ha descrito anteriormente, esta mayor uniformidad conduce a datos más robustos. En algunas realizaciones, un nucleótido tiene al menos una media de 8, 9, 10, 12, 15, 16, 20, 30 o 50 bases de separación. En algunas realizaciones, un nucleótido tiene al menos una media de 8 a 16 bases de separación. En algunas realizaciones, un nucleótido tiene al menos una media de 8 bases de separación. En algunas realizaciones, las secuencias contienen un patrón espaciado de forma bastante regular de un resto de nucleótido (por ejemplo, adenina) en la estructura principal. Esto permite que un nucleótido complementario

espaciado de forma regular de la secuencia polinucleotídica complementaria tenga unido al mismo una molécula detectable. Por ejemplo, en algunas realizaciones, cuando las secuencias nanoindicadoras contienen un patrón espaciado de forma bastante regular de restos de adenina en la estructura principal, cuyo complemento es un patrón espaciado de forma regular de restos de uridina (U) en los segmentos de ARN complementarios, la transcripción *in vitro* de la segmentos se pueden realizar usando una base de uridina modificada con aminoalilo, lo que permite el acoplamiento de amina covalente de moléculas de colorante a intervalos regulares a lo largo del segmento. En algunas realizaciones, las secuencias contienen aproximadamente el mismo número o porcentaje de un nucleótido (por ejemplo, adenina) que tiene al menos una media de 8, 9, 10, 12, 15, 16, 20, 30 o 50 bases de separación en las secuencias. Esto permite que un número o porcentaje similar de la secuencia polinucleotídica complementaria tenga unida a la misma una molécula detectable. Por lo tanto, en algunas realizaciones, las secuencias contienen un nucleótido que no está separado de forma regular, pero que tiene al menos una media de 8, 9, 10, 12, 15, 16, 20, 30 o 50 bases de separación, el número de nucleótidos de las secuencias puede variar en función del brillo deseado del nanoindicador. En algunas realizaciones, el 20 %, 30 %, 50 %, 60 %, 70 %, 80 %, 90 % o 100 % del nucleótido complementario está acoplado a una molécula detectable. Por ejemplo, en algunas realizaciones, cuando las secuencias nanoindicadoras contienen un porcentaje similar de adenina en la estructura principal, y la transcripción *in vitro* de los segmentos complementarios se realiza usando una base de uridina modificada con aminoalilo, el 20 %, 30 %, 50 %, 60 %, 70 %, 80 %, 90 % o 100 % de la base de uridina modificada con aminoalilo se puede acoplar a una molécula detectable. Como alternativa, la proporción de las bases de uridina modificada con aminoalilo con respecto a las bases de uridina se puede cambiar durante el proceso de transcripción *in vitro* para alcanzar el brillo deseado. Por ejemplo, el proceso de transcripción *in vitro* puede tener lugar en presencia de una mezcla con una proporción de 1/1 de las bases de uridina con respecto a las bases de uridina modificadas con aminoalilo, cuando algunas o todas las bases de uridina modificadas con aminoalilo se pueden acoplar a una molécula detectable. Así pues, el experto habitual en la materia entenderá que hay varios métodos en los que se puede lograr un brillo uniforme entre los nanoindicadores.

En algunas realizaciones, los nanoindicadores descritos en el presente documento tienen una temperatura de fusión (Tf) bastante uniforme. En algunas realizaciones, la Tf de las secuencias polinucleotídicas complementarias cuando se hibridan a sus regiones de unión al marcador de los nanoindicadores descritos en el presente documento son superiores a la Tf de secuencias polinucleotídicas complementarias a un molde de ADN de M13 cuando se hibridan a una sonda de nanoindicador que comprende el ADN de M13. Sin pretender quedar vinculados a teoría alguna, la Tf de los nanoindicadores descritos en el presente documento proporciona enlaces más fuertes entre la estructura principal y la secuencia polinucleotídica complementaria del nanoindicador que tiene unida a la misma una molécula detectable, por lo tanto, evitándose la disociación durante los procedimientos de síntesis y de hibridación. Además, la Tf uniforme entre una población de nanoindicadores permite la optimización de los procedimientos de síntesis y de hibridación, ya que las condiciones óptimas son las mismas para todas las puntos y posiciones. En algunas realizaciones, las secuencias de los nanoindicadores tienen un 50 % de guanina/citosina (G/C), con no más de tres G en una fila. Por lo tanto, en algunas realizaciones, la invención proporciona una población de nanoindicadores en la que la Tf entre los nanoindicadores de la población es bastante uniforme. En algunas realizaciones, la invención proporciona una población de nanoindicadores en la que la Tf de las secuencias polinucleotídicas complementarias cuando se hibridan a sus regiones de unión al marcador es de aproximadamente 80 °C, 85 °C, 90 °C, 100 °C o superior. En algunas realizaciones, la invención proporciona una población de nanoindicadores en la que la Tf de las secuencias polinucleotídicas complementarias cuando se hibridan a sus regiones de unión al marcador es de aproximadamente 80 °C o superior.

En algunas realizaciones, los nanoindicadores descritos en el presente documento tienen estructuras mínimas o no secundarias, tales como cualquier interacción de emparejamiento de bases intramolecular (por ejemplo, horquillas). Sin pretender quedar ligados a teoría alguna, la estructura secundaria mínima de los nanoindicadores proporciona una mejor hibridación entre la estructura principal y la secuencia polinucleotídica del nanoindicador que tiene unida a la misma una molécula detectable. Además, la estructura secundaria mínima de los nanoindicadores proporciona una mejor detección de las moléculas detectables en los nanoindicadores. En algunas realizaciones, los nanoindicadores descritos en el presente documento no tienen emparejamiento intramolecular significativo en condiciones de hibridación de 75 °C, 1 x SSPE. Las estructuras secundarias pueden predecirse mediante los programas conocidos en la técnica tales como MFOLD. En algunas realizaciones, los nanoindicadores descritos en el presente documento contienen menos del 1 % de repeticiones invertidas en cada cadena, en los que las repeticiones invertidas son de 9 bases o más. En algunas realizaciones, los nanoindicadores descritos en el presente documento no contienen repeticiones invertidas en cada cadena. En algunas realizaciones, los nanoindicadores no contienen ninguna repetición invertida de 9 nucleótidos o más a lo largo de una secuencia que es de 1.100 pares de bases de longitud. En algunas realizaciones, los nanoindicadores no contienen ninguna repetición invertida de 7 nucleótidos o más a lo largo de ninguna región de 100 pares de bases. En algunas realizaciones, los nanoindicadores descritos en el presente documento contienen menos de 1 % de repeticiones invertidas en cada cadena, en los que las repeticiones invertidas son de 9 nucleótidos o más a lo largo de una secuencia que 1.100 pares de bases de longitud. En algunas realizaciones, los nanoindicadores descritos en el presente documento contienen menos del 1 % de repeticiones invertidas en cada cadena, en los que las repeticiones invertidas son de 7 nucleótidos o más a lo largo de cualquier región de 100 pares de bases. En algunas realizaciones, los nanoindicadores descritos en el presente documento tienen un contenido específico de cadenas sesgadas tal como una cadena es rica en CT y la otra es rica en GA.

La invención proporciona nanoindicadores únicos. En algunas realizaciones, los nanoindicadores descritos en el presente documento contienen menos del 1 % de repeticiones directas. En algunas realizaciones, los nanoindicadores descritos en el presente documento no contienen repeticiones directas. En algunas realizaciones, los nanoindicadores no contienen ninguna repetición directa de 9 nucleótidos o más a lo largo de una secuencia de 1.100 pares de bases de longitud. En algunas realizaciones, los nanoindicadores marcados no contienen ninguna repetición directa de 7 nucleótidos o más a lo largo de ninguna región de 100 pares de bases. En algunas realizaciones, los nanoindicadores descritos en el presente documento contienen menos del 1 % de repeticiones directas en cada cadena, en los que las repeticiones directas son de 9 nucleótidos o más a lo largo de una secuencia que 1.100 pares de bases de longitud. En algunas realizaciones, los nanoindicadores descritos en el presente documento contienen menos del 1 % de repeticiones directas en cada cadena, en los que las repeticiones directas son de 7 nucleótidos o más a lo largo de cualquier región de 100 pares de bases. En algunas realizaciones, los nanoindicadores descritos en el presente documento contienen menos del 85, 80, 70, 60, 50, 40, 30, 20, 10 o 5 % de homología con cualquier otra secuencia usada en las estructuras principales o con cualquier secuencia descrita en la base de datos pública REFSEQ. En algunas realizaciones, los nanoindicadores descritos en el presente documento contienen menos del 85 % de homología con cualquier otra secuencia usada en las estructuras principales o con cualquier secuencia descrita en la base de datos pública REFSEQ. En algunas realizaciones, los nanoindicadores descritos en el presente documento contienen menos de 20, 16, 15, 10, 9, 7, 5, 3, 2 bases contiguas de homología con cualquier otra secuencia usada en las estructuras principales o con cualquier secuencia descrita en la base de datos pública REFSEQ. En algunas realizaciones, los nanoindicadores descritos en el presente documento no tienen más de 15 bases contiguas de homología y no más del 85 % de identidad a lo largo de toda la longitud del nanoindicador con cualquier otra secuencia usada en las estructuras principales o con cualquier secuencia descrita en la base de datos pública REFSEQ.

En algunas realizaciones, las características de secuencia de las sondas de nanoindicador descritas en el presente documento proporcionan una mejor detección de una molécula diana. Por ejemplo, la unión de las sondas de nanoindicador a moléculas diana que da lugar a la identificación de las moléculas diana se puede realizar mediante la detección de manera individual de la presencia del nanoindicador. Esto se puede realizar mediante el recuento de manera individual de la presencia de una o más de las moléculas nanoindicadoras de una muestra. En algunas realizaciones en las que se usan dichos métodos de recuento, las sondas de nanoindicador descritas en el presente documento permiten un aumento en el número de recuentos. En algunas realizaciones, el número de recuentos moleculares por encima del fondo de dicho complejo molecular tras la normalización de la muestra es superior a 300, 400, 450, 500, 600, 700, 800, 900 o 1.000 recuentos moleculares. En algunas realizaciones, el número de recuentos moleculares por encima del fondo de dicho complejo molecular tras la normalización de la muestra es superior a 400 recuentos moleculares. En algunas realizaciones, el porcentaje de recuentos moleculares válidos de las sondas de nanoindicador descritas en el presente documento es superior al aproximadamente 10, 11, 12, 12,5, 14, 15, 20, 25, 30, 35, 40, 50, 60, 70, 80 o 90 %. En algunas realizaciones, el porcentaje de recuentos moleculares válidos de las sondas de nanoindicador descritas en el presente documento es superior al aproximadamente 10 %. En algunas realizaciones, el porcentaje de recuentos moleculares válidos de las sondas de nanoindicador descritas en el presente documento es superior al aproximadamente 12,5 %. En algunas realizaciones, el número de recuentos moleculares por encima del fondo del nanoindicador descrito en el presente documento tras la normalización de dicha muestra es al menos 2, 3, 5, 6, 10, 20, 30, 40, 50, 60, 70, 80, 90 o 100 veces superior al de una sonda de nanoindicador comparable que comprende ADN de M13. Un nanoindicador comparable que comprende ADN de M13 es una nanoindicador que comprende la misma región específica de la diana unida a la estructura principal del ADN de M13. Los ejemplos de sondas de nanoindicador comparables que comprenden ADN de M13 se describen en el apartado de ejemplos. En algunas realizaciones, el número de recuentos moleculares por encima del fondo del nanoindicador descrito en el presente documento tras la normalización de dicha muestra es al menos 2 veces superior al de un nanoindicador comparable que comprende ADN de M13. En algunas realizaciones, el número de recuentos moleculares por encima del fondo del nanoindicador descrito en el presente documento tras la normalización de dicha muestra es al menos 6 veces superior al de un nanoindicador comparable que comprende ADN de M13. En algunas realizaciones, el número de recuentos moleculares por encima del fondo del nanoindicador descrito en el presente documento tras la normalización de dicha muestra es al menos 20 veces superior al de un nanoindicador comparable que comprende ADN de M13. En algunas realizaciones, el número de recuentos moleculares por encima del fondo del nanoindicador descrito en el presente documento tras la normalización de dicha muestra es al menos 100 veces superior al de un nanoindicador comparable que comprende ADN de M13.

Los elementos de un nanoindicador se pueden encontrar en una sola entidad molecular (un nanoindicador "simple") o en dos entidades moleculares distintas (un nanoindicador "doble"). Cada entidad molecular puede estar compuesta de una molécula o más de una molécula unidas entre sí por medios covalentes o no covalentes. En algunas realizaciones, cada componente de un nanoindicador doble tiene una secuencia específica de la diana que se une a un sitio diferente en la misma molécula diana. Esto permite componentes más pequeños del nanoindicador con cinéticas más eficientes de unión del nanoindicador a la molécula diana y mejores proporciones de señal:ruido como resultado de la mayor especificidad de unión. Cuando se usa un sistema de nanoindicador doble, una de las sondas de nanoindicador puede estar sin marcar. En algunas realizaciones, la sonda de nanoindicador sin marcar puede comprender una región de captura. En algunas realizaciones, la sonda de nanoindicador sin marcar puede comprender una región específica de la diana y una estructura principal que puede ser monocatenaria. En algunas realizaciones, la sonda de nanoindicador sin marcar puede comprender una región específica de la diana y una

estructura principal que puede ser bicatenaria.

Las secuencias polinucleotídicas complementarias unidas a una estructura principal del nanoindicador sirven para unir moléculas detectables, o monómeros de marcaje, a la estructura principal del nanoindicador. Las secuencias polinucleotídicas complementarias se pueden marcar directamente, por ejemplo, mediante la incorporación covalente de una o más moléculas detectables a la secuencia polinucleotídica complementaria. Como alternativa, las secuencias polinucleotídicas complementarias pueden marcarse indirectamente, tal como mediante la incorporación de biotina u otra molécula capaz de tener una interacción con un ligando específico en la secuencia polinucleotídica complementaria. En dichos casos, el ligando (por ejemplo, la estreptavidina en el caso de la incorporación de biotina a la secuencia polinucleotídica complementaria) se puede unir covalentemente a la molécula detectable. Cuando las moléculas detectables unidas a una región de unión al marcador no se incorporan directamente a la secuencia polinucleotídica complementaria, esta secuencia sirve como un puente entre la molécula detectable y la región de unión al marcador, y se puede denominar molécula puente, por ejemplo, ácido nucleico de puente.

El nanoindicador a base de ácido nucleico y los complejos de nanoindicador-diana de la presente invención comprenden ácidos nucleicos, que se pueden purificar por afinidad o inmovilizar usando un ácido nucleico, tal como un oligonucleótido que sea complementario a la región constante o al nanoindicador o al ácido nucleico diana. Como se ha indicado anteriormente, en algunas realizaciones, los nanoindicadores comprenden al menos una región constante, que puede servir como un marcador de afinidad para la purificación y/o para la inmovilización (por ejemplo, a una superficie sólida). La región constante normalmente comprende dos o más regiones repetidas en tándem de nucleótidos de repetición, tales como una serie de repeticiones de 15 bases. En dichas realizaciones ilustrativas, el nanoindicador, ya esté formando un complejo con una molécula diana o de otra manera, se puede purificar o inmovilizar mediante un reactivo de afinidad recubierto con un oligonucleótido de 15 bases que sea el complemento inverso de la unidad de repetición.

Los nanoindicadores o complejos de nanoindicador-molécula diana se pueden purificar en dos o más etapas de selección por afinidad. Por ejemplo, en un nanoindicador doble, una sonda puede comprender un primer marcador de afinidad y la otra sonda puede comprender un segundo marcador de afinidad (diferente). Las sondas se mezclan con moléculas diana, y los complejos que comprenden las dos sondas del nanoindicador doble se separan de los materiales no unidos (por ejemplo, la diana o las sondas individuales del nanoindicador) mediante purificación por afinidad contra uno o ambos marcadores de afinidad individuales. En la primera etapa, la mezcla se puede unir a un reactivo de afinidad para el primer marcador de afinidad, de modo que solo se purifican las sondas que comprenden el primer marcador de afinidad y los complejos deseados. Los materiales unidos se liberan del primer reactivo de afinidad y, opcionalmente, se unen a un reactivo de afinidad para el segundo marcador de afinidad, lo que permite la separación de los complejos de las sondas que comprenden el primer marcador de afinidad. En este momento, solo se unirían complejos completos. Los complejos se liberan finalmente del reactivo de afinidad para el segundo marcador de afinidad, y luego preferentemente se estiran y se fotografían. El reactivo de afinidad puede ser cualquier superficie sólida recubierta con una pareja de unión para el marcador de afinidad, tal como una columna, una perla (por ejemplo, de látex o perlas magnéticas) o un portaobjetos recubierto con la pareja de unión. La inmovilización y el estiramiento de nanoindicadores usando reactivos de afinidad se describe de manera completa en la solicitud provisional de EE.UU. n.º 60/753.816 por Sean M. Ferree y Dwayne L. Dunaway, titulada "Composiciones que comprenden macromoléculas inmovilizadas, orientadas, y métodos para su preparación", presentada el 23 de diciembre de 2005.

La secuencia de señales proporcionadas por los monómeros marcadores asociados a las diversas regiones de unión al marcador de la estructura principal de un nanoindicador dado permite la identificación única del nanoindicador. Por ejemplo, cuando se usan marcadores fluorescentes, un nanoindicador que tiene una identidad única o una firma espectral única se asocia con una secuencia específica de la diana que reconoce una molécula diana específica o una parte de la misma. Cuando un nanoindicador se expone a una mezcla que contiene la molécula diana en condiciones que permiten la unión de la/s secuencia/s específica/s de la diana del nanoindicador a la molécula diana, la/s secuencia/s específica/s de la diana se une/n preferentemente a la molécula diana. La detección de la señal del nanoindicador, tal como el código espectral de un nanoindicador marcado con fluorescencia, asociada con el nanoindicador permite la detección de la presencia de la molécula diana en la mezcla (análisis cualitativo). El recuento de todos los monómeros marcadores asociados con un código o con una firma espectral dado permite el recuento de todas las moléculas en la mezcla asociada con la secuencia específica de la diana acoplada al nanoindicador (análisis cuantitativo). Los nanoindicadores son, por tanto, útiles para el diagnóstico o el pronóstico de diferentes estados biológicos (por ejemplo, enfermedad frente a salud) mediante el análisis cuantitativo de los marcadores biológicos conocidos. Por otra parte, la exquisita sensibilidad de detección de moléculas individuales y la cuantificación proporcionada por los nanoindicadores de la invención permite la identificación de nuevos marcadores de diagnóstico y de pronóstico, incluyendo aquellos cuyas fluctuaciones entre los diferentes estados biológicos son demasiado leves para detectar una correlación con un determinado estado biológico usando métodos moleculares tradicionales. La sensibilidad de detección molecular basada en nanoindicadores permite el análisis farmacocinético detallado de agentes terapéuticos y de diagnóstico en muestras biológicas pequeñas.

Muchos nanoindicadores, denominados nanoindicadores simples, se componen de una entidad molecular. Sin embargo, para aumentar la especificidad de un nanoindicador y/o para mejorar la cinética de su unión a una molécula diana, un nanoindicador puede ser un nanoindicador doble, compuesto de dos entidades moleculares, conteniendo cada una de ellas una secuencia específica de la diana diferente que se une a una región diferente de la misma molécula diana. En un nanoindicador doble, al menos una de las dos entidades moleculares está marcada. La otra entidad molecular no tiene que estar necesariamente marcada. Dichos componentes no marcados de nanoindicadores dobles se pueden usar como sondas de captura (véanse las Figuras 1 y 2) y, opcionalmente, tienen marcadores de afinidad unidos, tal como biotina, que son útiles para inmovilizar y/o estirar el complejo que contiene el nanoindicador doble y la molécula diana para permitir la visualización y/o de formación de imágenes del complejo. Por ejemplo, en algunas realizaciones, se puede usar un nanoindicador doble con un código de nanoindicador en la posición 6, usando un nanoindicador codificado en la posición 6 y una sonda de captura. En algunas realizaciones, se puede usar un nanoindicador doble con un código de nanoindicador en la posición 7, usando un componente nanoindicador en la posición 8 y un componente nanoindicador en una sola posición. En algunas realizaciones, se puede usar un nanoindicador doble con un código de nanoindicador en la posición 6, usando una sonda de captura con un marcador de afinidad y un componente de nanoindicador en la posición 6. En alguna realización, se incluye opcionalmente un marcador de afinidad, y se puede usar para purificar el nanoindicador o inmovilizar el nanoindicador (o complejo de nanoindicador-molécula diana) con el fin de la formación de imágenes.

Debido a sus estructuras modulares, los nanoindicadores se pueden montar y marcar de una variedad de maneras diferentes. Por ejemplo, se puede unir una estructura principal de nanoindicador a una secuencia específica de la diana (por ejemplo, por hibridación y, opcionalmente, ligadura), y unirse la estructura que comprende la estructura principal y la secuencia específica de la diana a una o más secuencias polinucleotídicas complementarias que tengan unida a las mismas, ya sea directa o indirectamente, una molécula detectable. Como alternativa, se puede unir primero la estructura principal del nanoindicador a una o más secuencias polinucleotídicas complementarias, y unirse después la estructura principal a una secuencia específica diana. Por lo tanto, a menos que se indique lo contrario, una descripción o una enumeración de las etapas en el montaje del nanoindicador no implican que haya que seguirse una ruta específica de montaje.

La síntesis de los nanoindicadores se puede realizar mediante cualquier método adecuado conocido en la técnica. Se pueden cultivar plásmidos bicatenarios que porten secuencias de nanoindicador (por ejemplo, secuencias de DV1) en algunas cepas bacterianas en condiciones de temperatura baja (no superior a 34 °C). Las estructuras principales monocatenarias lineales de los indicadores se pueden fabricar a partir de ADN bicatenario de plásmido usando un protocolo de cuatro etapas, que incluye la linealización con una enzima de restricción, la desfosforilación con una fosfatasa termolábil, la digestión con una segunda enzima de restricción para separar el vector de clonación de la secuencia de la estructura principal, y una digestión con exonucleasa lambda específica de la cadena que deje solo una de las cadenas del fragmento de estructura principal intacta. La Figura 9 muestra un ejemplo de un vector que se puede usar para la síntesis de nanoindicadores.

El montaje y el uso de nanoindicadores se ilustran en el presente documento, en gran parte, a modo de descripción de una variedad de nanoindicadores a base de ácidos nucleicos. A continuación, se presentan realizaciones ilustrativas de nanoindicadores montados parcial y totalmente.

En su forma más simple, la divulgación proporciona una estructura principal de ácido nucleico que tiene una pluralidad (por ejemplo, 3) de las regiones de unión al marcador que son capaces de ser marcadas y resueltas, estando cada una de ellas hecha de una secuencia de nucleótidos de diseño racional. Estas secuencias abarcan una o más, o todas, las características descritas en el presente documento que hacen que el nanoindicador sea más estable. Los ejemplos de moldes de polinucleótidos que se pueden utilizar para generar estas regiones de unión al marcador diseñadas se exponen en las secuencias polinucleotídicas de SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23 y 24. Las regiones de unión al marcador de cada estructura principal individual están dispuestas en un orden lineal única o combinación, haciendo a cada estructura principal individual única en comparación con el resto de estructuras principales de una población. Dado que cada uno de los nanoindicadores individuales de una población está formado por una estructura principal única, cada nanoindicador es igualmente único en comparación con el resto de nanoindicadores de una población. En algunas realizaciones, cada región de unión al marcador es única en comparación con la otra región de unión al marcador de la estructura principal.

En algunas realizaciones, las secuencias de nucleótidos de las regiones de unión al marcador individuales dentro de cada nanoindicador son diferentes de las secuencias de nucleótidos de las otras regiones de unión al marcador de ese nanoindicador, evitando los reordenamientos, tales como la recombinación, la compartición o el cambio de las secuencias polinucleotídicas del marcador, y mejorando así la estabilidad molecular. El número de regiones de unión al marcador que se van a formar en una estructura principal se basa en la longitud y en la naturaleza de la estructura principal, los medios de marcaje del nanoindicador, así como en el tipo de monómeros marcadores que proporcionan una señal que se va a unir a las regiones de unión al marcador de la estructura principal. En algunas realizaciones, la secuencia de nucleótidos complementaria de cada región de unión al marcador recibe una molécula detectable específica.

En el presente documento, se describen nanoindicadores marcados, en los que una o más regiones de unión al marcador están unidas a una molécula detectable correspondiente, proporcionando cada molécula detectable una señal. Por ejemplo, en algunas realizaciones, un nanoindicador marcado de acuerdo con la invención se obtiene cuando al menos tres moléculas detectables se unen a tres regiones de unión al marcador correspondientes de la estructura principal, de manera que estas regiones de unión al marcador marcadas, o manchas, se distinguen basándose en su disposición lineal única. Una "mancha", en el contexto de la detección de los nanoindicadores, es la señal agregada detectada procedente de los monómeros marcadores unidos a un solo sitio de unión al marcador de un nanoindicador, y que, dependiendo del tamaño de la región de unión al marcador y de la naturaleza (por ejemplo, longitud de onda de emisión primaria) del monómero marcador, puede aparecer como una única fuente puntual de luz cuando al observar con un microscopio. Las manchas de un nanoindicador pueden estar superpuestas o no. El código del nanoindicador que identifica esa molécula diana puede comprender cualquier permutación de la longitud de una mancha, su posición relativa con respecto a otras manchas y/o la naturaleza (por ejemplo, la/s longitud/es de onda de emisión primaria/s) de su señal. En general, para cada sonda o par de sondas, las regiones de unión al marcador adyacentes no están superpuestas, y/o las manchas de las regiones de unión al marcador adyacentes se distinguen espacial y/o espectralmente, al menos en las condiciones de detección (por ejemplo, cuando el nanoindicador se inmoviliza, se estira y se observa bajo un microscopio, como se describe en la solicitud de EE.UU. n.º de serie 61/029.220).

En ocasiones, se hace referencia al tamaño de una mancha como un cierto número de bases o nucleótidos. Como un experto en la materia entenderá fácilmente, esto se refiere al número de bases o nucleótidos de la región de unión al marcador correspondiente.

El orden y la naturaleza (por ejemplo, la/s longitud/es de onda de emisión primaria/s, opcionalmente también la longitud) de las manchas de un nanoindicador sirven como un código del nanoindicador que identifica la molécula diana capaz de ser unida por el nanoindicador a lo largo de la/s secuencia/s específica/s diana del nanoindicador. Cuando el nanoindicador se une a una molécula diana, el código del nanoindicador también identifica la molécula diana. Opcionalmente, la longitud de una mancha puede ser un componente del código del nanoindicador.

Las moléculas detectables que proporcionan una señal asociada a diferentes regiones de unión al marcador de la estructura principal pueden proporcionar señales que son indistinguibles en las condiciones de las detecciones (señales "similares"), o pueden proporcionar señales que se pueden distinguir, al menos en las condiciones de la detección (por ejemplo, cuando el nanoindicador se inmoviliza, se estira y se observa con un microscopio).

La divulgación también proporciona un nanoindicador en el que dos o más moléculas detectables se unen a una región de unión al marcador. La señal proporcionada por las moléculas detectables asociadas con dicha región de unión al marcador produce una señal agregada que se detecta. La señal agregada producida puede estar compuesta de señales similares o compuesta de al menos dos señales distinguibles (por ejemplo, señales distinguibles espectralmente).

En una realización, la divulgación proporciona un nanoindicador en el que al menos tres moléculas detectables que proporcionan señales similares se unen a tres regiones de unión al marcador correspondientes de la estructura principal y dichas tres moléculas detectables se distinguen espacialmente. En otra realización, la divulgación proporciona un nanoindicador en el que al menos tres moléculas detectables que proporcionan tres señales distinguibles se unen a tres regiones de unión al marcador vecinas, por ejemplo, tres regiones de unión al marcador adyacentes, mediante lo que dichos al menos tres monómeros marcadores se distinguen espectralmente.

La divulgación también proporciona un nanoindicador en el que las manchas que proporcionan señales similares o distintas están separadas por una región espaciadora, de modo que la interposición de la región espaciadora permite la generación de manchas oscuras, que expanden la posible combinación de señales detectables de forma única. La expresión "mancha negra" se refiere a la falta de señal procedente de un sitio de unión al marcador de un nanoindicador. Las manchas oscuras se pueden incorporar al código del nanoindicador para añadir más permutaciones de codificación y generar una mayor diversidad de los nanoindicadores de una población de nanoindicadores. En una realización, las regiones espaciadoras tienen una longitud determinada por la resolución de un instrumento empleado en la detección del nanoindicador.

La divulgación proporciona un nanoindicador con una o más "manchas dobles". Cada mancha doble contiene dos o más (por ejemplo, tres, cuatro o cinco) manchas adyacentes que proporcionan señales similares sin estar separadas por una región espaciadora. Las manchas dobles se pueden identificar por sus tamaños.

Una molécula detectable que proporciona una señal de acuerdo con la invención se puede unir covalente o no covalentemente (por ejemplo, mediante hibridación) a una secuencia polinucleotídica complementaria que esté unida a la región de unión al marcador. Los monómeros marcadores también se pueden unir indirectamente a la secuencia polinucleotídica complementaria, tal como mediante unión covalente a una molécula de ligando (por ejemplo, estreptavidina) que se une por medio de su interacción con una molécula incorporada a la secuencia polinucleotídica complementaria (por ejemplo, biotina incorporada a la secuencia polinucleotídica complementaria) que, a su vez, se une mediante hibridación a la estructura principal.

El nanoindicador puede estar asociado a una señal detectable de forma única, tal como un código espectral, determinada por la secuencia de señales proporcionada por los monómeros marcadores unidos (por ejemplo, indirectamente) a las regiones de unión al marcador de la estructura principal del nanoindicador, por lo que la detección de la señal permite la identificación del nanoindicador.

El nanoindicador puede comprender además un marcador de afinidad unido a la estructura principal del nanoindicador, de manera que la unión del marcador de afinidad a un soporte permite a la estructura principal el estiramiento y la resolución de las señales proporcionadas por los monómeros marcadores correspondientes a diferentes regiones de unión al marcador de la estructura principal. El estiramiento del nanoindicador puede implicar cualquier medio de estiramiento conocido en la técnica, incluyendo, pero sin limitación, medios que incluyen medios físicos, hidrodinámicos o eléctricos. El marcador de afinidad puede comprender una región constante.

Un nanoindicador de acuerdo con la invención puede incluir además una secuencia específica de la diana acoplada a la estructura principal. La secuencia específica de la diana se selecciona para permitir que el nanoindicador reconozca, se una o se adhiera a una molécula diana. Los nanoindicadores de la invención son adecuados para la identificación de moléculas diana de todos los tipos. Por ejemplo, las secuencias específicas de la diana apropiadas se pueden acoplar a la estructura principal del nanoindicador para permitir la detección de una molécula diana. Preferentemente, la molécula diana es ADN (incluyendo ADNc), ARN (incluyendo ARNm y ARNc), un péptido, un polipéptido o una proteína.

Una realización de la divulgación proporciona una mayor flexibilidad en la detección de la molécula diana con monómeros marcadores. En dicha realización, un nanoindicador doble que comprende dos entidades moleculares diferentes, cada una con una región específica de la diana separada, al menos una de las cuales está marcada, se une a la misma molécula diana. Por lo tanto, las secuencias específicas de la diana de los dos componentes del nanoindicador doble se unen a diferentes partes de una molécula diana seleccionada, por lo que la detección del código espectral asociado con el nanoindicador doble proporciona la detección de la molécula diana seleccionada en una muestra biomolecular en contacto con dicho nanoindicador doble.

La divulgación también proporciona un método de detección de la presencia de una molécula diana específica en una muestra biomolecular, que comprende: (i) poner en contacto dicha muestra con un nanoindicador como se describe en el presente documento (por ejemplo, un nanoindicador simple o doble) en condiciones que permitan la unión de las secuencias específicas de la diana del nanoindicador doble a la molécula diana y (ii) detectar el código espectral asociado con el nanoindicador doble. Dependiendo de la arquitectura del nanoindicador, el nanoindicador doble se puede marcar antes o después de la unión a la molécula diana.

La singularidad de cada sonda de nanoindicador de una población de la sonda permite el análisis multiplexado de una pluralidad de moléculas diana. Por ejemplo, en algunas realizaciones, cada sonda de nanoindicador contiene seis regiones de unión al marcador, donde cada región de unión al marcador de cada estructura principal es diferente de las otras regiones de unión al marcador de la misma estructura principal. Si las regiones de unión al marcador van a ser marcadas con uno de los cuatro colores, y hay 24 posibles secuencias únicas para las regiones de unión al marcador y a cada región de unión al marcador se le asigna un color específico, cada región de unión al marcador de cada estructura principal consistirá en una de cuatro secuencias. Habrá 4.096 nanoindicadores posibles en este ejemplo. El número de posibles nanoindicadores se puede aumentar, por ejemplo, aumentando el número de colores, aumentando el número de secuencias únicas para las regiones de unión al marcador y/o aumentando el número de regiones de unión al marcador por estructura principal. Del mismo modo, el número de posibles nanoindicadores se puede disminuir reduciendo el número de colores, reduciendo el número de secuencias únicas para las regiones de unión al marcador y/o reduciendo el número de regiones de unión al marcador por estructura principal.

En ciertas realizaciones, los métodos de detección se realizan en ensayos multiplexados, mediante los que se detecta una pluralidad de moléculas diana en el mismo ensayo (una única mezcla de reacción). En una realización preferida, el ensayo es un ensayo de hibridación en el que la pluralidad de moléculas diana se detecta simultáneamente. En ciertas realizaciones, la pluralidad de moléculas diana detectada en el mismo ensayo es al menos 2, al menos 5 moléculas diana diferentes, al menos 10 moléculas diana diferentes, al menos 20 moléculas diana diferentes, al menos 50 moléculas diana diferentes, al menos 75 moléculas diana diferentes, al menos 100 moléculas diana diferentes, al menos 200 moléculas diana diferentes, al menos 500 moléculas diana diferentes o al menos 750 moléculas diana diferentes o al menos 1.000 moléculas diana diferentes. En otras realizaciones, la pluralidad de moléculas diana detectada en el mismo ensayo es de hasta 50 moléculas diana diferentes, de hasta 100 moléculas diana diferentes, de hasta 150 moléculas diana diferentes, de hasta 200 moléculas diana diferentes, de hasta 300 moléculas diana diferentes, de hasta 500 moléculas diana diferentes, de hasta 750 moléculas diana diferentes, de hasta 1.000 moléculas diana diferentes, de hasta 2.000 moléculas diana diferentes o de hasta 5.000 moléculas diana diferentes. En otras realizaciones más, la pluralidad de moléculas diana detectada es cualquier intervalo entre los números anteriores de diferentes moléculas diana, tales como, pero sin limitación, de 20 a 50 moléculas diana diferentes, de 50 a 200 moléculas diana diferentes, de 100 a 1.000 moléculas diana diferentes, de 500 a 5.000 moléculas diana diferentes, etcétera.

En ciertas realizaciones, la invención se dirige a la detección de diferentes formas de corte y empalme del mismo ARN. Las diferentes formas de corte y empalme se pueden detectar usando una pluralidad de sondas de nanoindicador, cada una con una secuencia específica de la diana diferente complementaria a un exón diferente del mismo gen.

5 Además de las capacidades analíticas cualitativas proporcionadas por los nanoindicadores de la invención y las técnicas analíticas basadas en las mismas, los nanoindicadores de la invención son especialmente adecuados para la realización de análisis cuantitativos. Al proporcionar una unión de uno a uno entre los nanoindicadores (ya sean nanoindicadores simples o dobles) de la invención y sus moléculas diana en una muestra biomolecular, se puede
10 identificar y contar la totalidad o una parte representativa de las moléculas diana presentes en la muestra. Este recuento individual de las diversas especies moleculares proporciona un método preciso y directo para determinar la concentración absoluta o relativa de la molécula diana en la muestra biomolecular. Por otra parte, la capacidad para dirigirse a cada molécula de la mezcla individualmente potencia los beneficios de la miniaturización incluyendo una elevada sensibilidad, requerimientos de cantidades mínimas de muestra, elevadas velocidades de reacción que son permitidas por la cinética en fase de solución en un pequeño volumen, y por último, costes de reactivos muy bajos.

Como se apreciará a partir de la descripción y de los ejemplos proporcionados más adelante, la presente invención proporciona numerosas ventajas. Por ejemplo, la modularidad del complejo al formar los nanoindicadores de acuerdo con la invención permite la creación sistemática de genotecas de nanoindicadores únicos que tienen un
20 grado de diversidad muy elevado (por ejemplo, millones de nanoindicadores reconocibles de manera única). Dicha modularidad permite la flexibilidad en la adaptación de las poblaciones de nanoindicadores a las aplicaciones específicas que, a su vez, proporcionan eficacias de fabricación significativas. Otra ventaja que se apreciará a través de la siguiente descripción proviene de la flexibilidad en el ensamblaje de los nanoindicadores de la invención. Es decir, debido a su estructura modular, los nanoindicadores de la invención se pueden ensamblar antes del envío a
25 un punto de uso o ensamblarse en el punto de uso.

Nanoindicadores dobles

Las Figuras 1 y 2 ilustran ciertas realizaciones que implican nanoindicadores dobles. En algunas realizaciones, cada uno de los dos componentes del nanoindicador se marca, de manera que el código espectral del nanoindicador solo se forma cuando los dos componentes del nanoindicador se juntan tras la unión del nanoindicador doble a su molécula diana. Sin embargo, en un nanoindicador doble, no es necesario que ambos componentes estén marcados. Por ejemplo, como se representa en las Figuras 1 y 2, un componente de un nanoindicador doble se marca con el código de nanoindicador, y el otro componente se une a un marcador de afinidad (flecha) que es útil para inmovilizar el nanoindicador para el estiramiento y la visualización.

Definiciones

A menos que se defina lo contrario, todos los términos técnicos y científicos usados en el presente documento tienen el mismo significado comúnmente entendido por los expertos habituales en la materia a la que pertenece la invención. Para los fines de la presente invención, a continuación, se definen los siguientes términos.

Los artículos "un" y "una" se usan en el presente documento para referirse a uno o a más de una (por ejemplo, a al menos uno) del objeto gramatical del artículo. A modo de ejemplo, "un elemento" significa un elemento o más de un elemento.

Los términos "complementarios" con respecto a polinucleótidos se refieren a polinucleótidos relacionados por las reglas de apareamiento de bases. Por ejemplo, la secuencia "A-G-T" es complementaria a la secuencia "T-C-A". La complementariedad puede ser parcial, en la que solo algunas de las bases de los ácidos nucleicos se emparejan de acuerdo con las reglas de apareamiento de bases; o puede haber complementariedad completa o total entre los ácidos nucleicos. El grado de complementariedad entre las cadenas de ácido nucleico tiene efectos significativos sobre la eficiencia y la fuerza de hibridación entre las cadenas de ácido nucleico.

Los términos "polinucleótidos", "ácidos nucleicos", "nucleótidos" y "oligonucleótidos" se usan indistintamente. Se refieren a una forma polimérica de nucleótidos de cualquier longitud, ya sean desoxirribonucleótidos o ribonucleótidos, o análogos de los mismos. Los siguientes son ejemplos de polinucleótidos no limitantes: regiones codificantes o no codificantes de un gen o fragmento de gen, locus definidos a partir del análisis de enlaces, exones, intrones, ARN mensajero (ARNm), ARN de transferencia, ARN ribosomal, ribozimas, ADNc, polinucleótidos recombinantes, polinucleótidos ramificados, plásmidos, vectores, ADN aislado de cualquier secuencia, ARN aislado de cualquier secuencia, sondas de ácido nucleico y cebadores. Un polinucleótido puede comprender nucleótidos modificados, tales como nucleótidos metilados y análogos de nucleótidos. Si están presentes, se pueden conferir modificaciones en la estructura de nucleótidos antes o después de la formación del polímero. La secuencia de nucleótidos se puede interrumpir con componentes no nucleótidos. Un polinucleótido puede modificarse además tras la polimerización, tal como mediante la conjugación con un componente de marcaje.

65

Las expresiones "identidad de secuencia" o, por ejemplo, que comprende una "secuencia idéntica en un 50 % a", como se usan en el presente documento, se refieren a la medida en que las secuencias son idénticas de nucleótido a nucleótido en una ventana de comparación. Por lo tanto, se puede calcular un "porcentaje de identidad de secuencia" mediante la comparación de dos secuencias óptimamente alineadas en una ventana de comparación, determinando el número de posiciones en las que se produce la base de ácido nucleico idéntica (por ejemplo, A, T, C, G, I) en ambas secuencias para producir el número de posiciones coincidentes, dividiendo el número de posiciones coincidentes entre el número total de posiciones de la ventana de comparación (por ejemplo, el tamaño de la ventana), y multiplicando el resultado por 100 para dar el porcentaje de identidad de secuencia.

Las expresiones y los términos usados para describir las relaciones de secuencia entre dos o más polinucleótidos incluyen "homología", "ventana de comparación", "identidad de secuencia", "porcentaje de identidad de secuencia" e "identidad sustancial". Debido a que dos polinucleótidos pueden comprender cada uno (1) una secuencia (por ejemplo, solo una parte de la secuencia polinucleotídica completa) que es similar entre los dos polinucleótidos; y (2) una secuencia que es divergente entre los dos polinucleótidos, las comparaciones de secuencias entre dos (o más) polinucleótidos normalmente se realizan comparando secuencias de los dos polinucleótidos en una "ventana de comparación" para identificar y comparar regiones locales de similitud de secuencia. Una "ventana de comparación" se refiere a un segmento conceptual de al menos 6 posiciones contiguas, en general, de aproximadamente 50 a aproximadamente 100, más habitualmente de aproximadamente 100 a aproximadamente 150 en la que una secuencia se compara con una secuencia de referencia del mismo número de posiciones contiguas tras alinear las dos secuencias de manera óptima. La ventana de comparación puede comprender adiciones o deleciones (es decir, huecos) del aproximadamente 20 % o menos en comparación con la secuencia de referencia (que no comprende adiciones ni deleciones) para la alineación óptima de las dos secuencias. La alineación óptima de secuencias para alinear una ventana de comparación se puede realizar mediante implementaciones informatizadas de algoritmos (GAP, BESTFIT, FASTA y TFASTA en el Wisconsin Genetics Software Package Release 7.0, Genetics Computer Group, 575 Science Drive Madison, WI, EE.UU.) o mediante inspección y la mejor alineación (es decir, dando como resultado el porcentaje más alto de homología en la ventana de comparación) generada mediante cualquiera de los diversos métodos seleccionados. También cabe hacer referencia a la familia de programas BLAST como, por ejemplo, los desvelados por Altschul *et al.*, 1997, *Nucl. Acids Res.* 25: 3389. Se puede encontrar una descripción detallada del análisis de las secuencias en la Unidad 19.3 de Ausubel *et al.*, "Current Protocols in Molecular Biology", John Wiley & Sons Inc, 1994-1998, Capítulo 15.

Estructura principal

La estructura principal del nanoindicador de la presente invención es una molécula de ácido nucleico que contiene una pluralidad de regiones de unión al marcador (por ejemplo, al menos tres regiones de unión al marcador) dispuestas en una combinación lineal a la que los monómeros marcadores se pueden unir directa o indirectamente. En una realización, la estructura principal del nanoindicador es una estructura principal de ácido nucleico en la que las regiones de unión al marcador son regiones monocatenarias a la que otros ácidos nucleicos, tales como oligonucleótido complementario, secuencias de ARN complementarias o secuencias de ADN complementarias, pueden unirse por hibridación. En realizaciones específicas, la estructura principal del nanoindicador es una molécula de ácido nucleico monocatenario.

En algunas realizaciones, la estructura principal de la presente invención es ADN. Las estructuras basadas en ADN ofrecen numerosas ventajas en el contexto de la presente invención debido, al menos en parte, al gran universo de técnicas y metodologías existentes que permiten la manipulación de las construcciones de ADN. Como se ha indicado anteriormente, la estructura principal puede ser monocatenaria.

Cada estructura principal se compone de una sola disposición de regiones de unión al marcador. Las regiones de unión al marcador de una estructura principal de nanoindicador variarán en tamaño dependiendo del método de marcaje. En diversas realizaciones, una región de unión al marcador puede tener una longitud de cualquier valor de 10 nm a 10.000 nm, pero es más preferentemente de 50 nm a 5.000 nm, y es más preferentemente de 100 nm a 1.000 nm. En diversas realizaciones, la región de unión al marcador es de aproximadamente 100 nm a aproximadamente 500 nm, de aproximadamente 150 nm a aproximadamente 450 nm, de aproximadamente 200 nm a aproximadamente 400 nm, o de 250 a aproximadamente 350 nm. En una realización preferida, la región de unión al marcador se corresponde estrechamente con el tamaño de una mancha limitada por la difracción, por ejemplo, la mancha más pequeña que se puede detectar con la óptica convencional, que es de aproximadamente 300 nm.

En algunas realizaciones, las regiones de unión al marcador están formadas por secuencias polinucleotídicas. Para las estructuras principales de ácido nucleico, 1 nm corresponde a aproximadamente 3 nucleótidos; por lo tanto, una región de unión al marcador de aproximadamente 300 nm corresponde a una secuencia de nucleótidos de aproximadamente 900 bases. En ciertas realizaciones, la región de unión al marcador es de aproximadamente 300 nucleótidos a aproximadamente 1,5 kb, de aproximadamente 450 nucleótidos a aproximadamente 1,35 kb, de aproximadamente 0,6 kb a aproximadamente 1,2 kb, o de 0,75 kb a aproximadamente 1,05 kb. En ciertos aspectos, la región de unión al marcador es de aproximadamente 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1.000, 1.050, 1.100, 1.150, 1.200, 1.250, 1.300, 1.350, 1.400, 1.450, 1.500, 1.600, 1.700, 1.800, 1.900, 2.000, 2.500, 3.000, 3.500, 4.000, 4.500, 5.000 o más nucleótidos, incluyendo todos los números enteros

intermedios. En ciertas realizaciones preferidas, la región de unión al marcador es una secuencia de nucleótidos de aproximadamente 1.100 nucleótidos de longitud.

Una estructura principal puede tener 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24-100 regiones de unión al marcador o más, incluyendo todos los números enteros intermedios. En ciertas realizaciones, una estructura principal tiene 6 regiones de unión al marcador. En alguna realización, las estructuras principales individuales tienen un conjunto de regiones de unión al marcador distintas, cada una de las cuales tiene una secuencia polinucleotídica que es diferente de las secuencias de las otras regiones de unión al marcador de la misma estructura principal.

En algunas realizaciones, una estructura principal tiene 6 regiones únicas de unión al marcador. En algunas realizaciones, cada estructura principal comprende una pluralidad de diferentes moléculas detectables en la que dicha pluralidad de diferentes moléculas detectables de cada estructura principal tiene una señal detectable que la distingue de otras estructuras principales de la población.

Cada estructura principal de una población de estructuras principales se distingue del resto por tener una combinación lineal única, por el orden o por la disposición de las regiones de unión al marcador. La secuencia polinucleotídica de cada región de unión al marcador de una estructura principal dada es diferente de la secuencia polinucleotídica de las otras regiones de unión al marcador de la misma estructura principal. En una disposición lineal de múltiples regiones de unión al marcador de la misma estructura principal, la repetición de más de una de las mismas regiones de unión al marcador puede hacer que los plásmidos a partir de los cuales se generan las cadenas principales sean inestables durante la replicación en *E. coli* y, por lo tanto, difíciles de propagar. Además, la introducción de segmentos de ARN de complemento inverso a dicha estructura principal puede causar la formación de indicadores "nudosos" en los que un segmento de ARN une y reúne dos regiones de unión al marcador idénticas, alterando la secuencia lineal del indicador e interfiriendo en la precisión y la legibilidad del código del indicador, entre otros problemas. En realizaciones en las que la inestabilidad del plásmido y la legibilidad del indicador potencialmente baja son un problema, estos problemas se evitan mediante la construcción de cada estructura principal individual para que tenga un conjunto de regiones de unión al marcador distintas, teniendo cada una de ellas una secuencia polinucleotídica que es diferente de las secuencias de las otras regiones de unión al marcador de la misma estructura principal.

Además, estas regiones de unión al marcador se seleccionan de una población de secuencias de nucleótidos de diseño racional (por ejemplo, secuencias sintéticas) únicas, teniendo cada una de ellas una secuencia polinucleotídica única representada por ciertas características (por ejemplo, contenido de G/C, proporción de G/C, repeticiones de adenina), como se detalla en el presente documento, designándose a cada una de ellas una molécula detectable dada. Estas secuencias de diseño racional no solo mejoran la estabilidad mediante la prevención de la formación de la estructura secundaria (por ejemplo, manteniendo la secuencia lineal del indicador para mantener la precisión y la legibilidad del código indicador), sino que optimizan la concentración y la separación de las moléculas detectables unidas a cada región de unión al marcador, lo que mejora la legibilidad de una estructura principal marcada entre una población de estructuras principales marcadas de forma única.

Regiones de unión al marcador

La presente invención proporciona estructuras principales de nanoindicador que son moléculas de ácido nucleico sintéticas (ADN, ARN o híbridos de ADN/ARN), de diseño racional para que tengan características que optimicen el marcaje y la detección del nanoindicador, y que eviten la formación de una estructura secundaria. En algunas realizaciones de la invención, una estructura principal de nanoindicador es una secuencia polinucleotídica diseñada que comprende una o más secuencias de 50 a 50.000 bases de longitud.

En algunas realizaciones, la estructura principal del nanoindicador está diseñada para minimizar las estructuras secundarias predecibles. En algunas realizaciones, la estructura principal del nanoindicador está desprovista de cualquier estructura secundaria. Las estructuras secundarias putativas (por ejemplo, horquillas, plegamientos o apareamiento de bases internas) se pueden predecir mediante métodos conocidos en la técnica tales como MFOLD. Sin pretender quedar limitados a teoría alguna, la estructura secundaria predecible de la estructura del nanoindicador puede minimizarse evitando las repeticiones invertidas y distorsionando el contenido específico de la estructura principal de manera que la estructura principal sea rica en CT o en GA. En algunas realizaciones, la estructura principal del nanoindicador no tiene ningún apareamiento intramolecular significativo en condiciones de hibridación de 75 °C, 1 x SSPE. En algunas realizaciones, la estructura principal del nanoindicador tiene menos del aproximadamente 20, 15, 10, 5, 4, 3, 2 o 1 % de repeticiones invertidas, en la que las repeticiones invertidas son de 9 nucleótidos o más. En algunas realizaciones, la estructura principal tiene menos del aproximadamente 1 % de repeticiones invertidas, en la que las repeticiones invertidas son de 9 nucleótidos o más. En algunas realizaciones, los nanoindicadores no contienen ninguna repetición invertida de 9 nucleótidos o más a lo largo de una secuencia de 1.100 pares de bases de longitud. En algunas realizaciones, los nanoindicadores no contienen ninguna repetición invertida de 7 nucleótidos o más a lo largo de cualquier región de 100 pares de bases. En algunas realizaciones, la estructura principal del nanoindicador está diseñada para sesgar el contenido específico de la cadena de la estructura principal de manera que la estructura principal sea rica en CT o en GA. Una región CT o GA es una región

compuesta de cualquier combinación de los nucleótidos C y T, o G y A. Por ejemplo, la región CT más pequeña es una región de dos nucleótidos en la que los dos nucleótidos son C y T. Las regiones pueden ser de 2, 3, 4, 5, 6, 7, 8 o más nucleótidos de longitud. En algunas realizaciones, aproximadamente el 30, 35, 40, 45, 50, 55, 60, 65 o 70 % de la estructura principal del nanoindicador se compone de regiones CT o GA. En algunas realizaciones, del aproximadamente 50 al 65 % de la estructura principal del nanoindicador se compone de regiones CT o GA. En algunas realizaciones, aproximadamente el 60 % de la estructura principal del nanoindicador se compone de regiones CT o GA. En algunas realizaciones, la estructura principal del nanoindicador está diseñada para sesgar el contenido específico de la cadena de la estructura principal de manera que la estructura principal sea rica en CT. En algunas realizaciones, aproximadamente el 60 % de la estructura principal del nanoindicador se compone de regiones CT.

En algunas realizaciones, la estructura principal del nanoindicador está diseñada para maximizar la singularidad de toda la secuencia de una población de nanoindicadores evitando repeticiones directas y mediante el rastreo de toda la secuencia que tenga cualquier homología significativa con alguna otra secuencia usada en otras estructuras principales de una población de nanoindicadores o a cualquier secuencia descrita en la base de datos pública REFSEQ. En algunas realizaciones, la estructura principal del nanoindicador tiene menos del aproximadamente 20, 15, 10, 5, 4, 3, 2 o 1 % de repeticiones directas. En algunas realizaciones, la estructura principal del nanoindicador tiene menos del aproximadamente el 1 % de repeticiones directas, en la que las repeticiones directas son de 9 nucleótidos o más en una secuencia de 1.100 pares de bases de longitud o de 7 nucleótidos o más a lo largo de cualquier región de 100 pares de bases. En algunas realizaciones, los nanoindicadores descritos en el presente documento contienen menos del 85, 80, 70, 60, 50, 40, 30, 20, 10 o 5 % de homología con cualquier otra secuencia usada en otras estructuras principales de una población de nanoindicadores o con cualquier secuencia descrita en la base de datos pública REFSEQ. En algunas realizaciones, los nanoindicadores descritos en el presente documento contienen menos del 60 % de homología con cualquier otra secuencia usada en otras estructuras principales de una población de nanoindicadores o con cualquier secuencia descrita en la base de datos pública REFSEQ. En algunas realizaciones, los nanoindicadores descritos en el presente documento contienen menos del 50 % de homología con cualquier otra secuencia usada en otras estructuras principales de una población de nanoindicadores o con cualquier secuencia descrita en la base de datos pública REFSEQ. En algunas realizaciones, la estructura principal del nanoindicador está diseñada para contener menos del 85 % de homología con cualquier otra secuencia usada en las estructuras principales o con cualquier secuencia descrita en la base de datos pública REFSEQ. En algunas realizaciones, la estructura principal del nanoindicador está diseñada para contener menos de 20, 16, 15, 10, 9, 7, 5, 3, 2 bases contiguas de homología con cualquier otra secuencia usada en las estructuras principales o con cualquier secuencia descrita en la base de datos pública REFSEQ. En algunas realizaciones, la estructura principal de los nanoindicadores está diseñada para tener no más de 15 bases contiguas de homología y no más del 85 % de identidad en toda la longitud del nanoindicador con cualquier otra secuencia usada en las estructuras principales o con cualquier secuencia descrita en la base de datos pública REFSEQ.

Las regiones de unión al marcador de ácido nucleico de la invención preferentemente no tienen repeticiones directas o invertidas que sean mayores de 12 bases de longitud. En otras realizaciones, las regiones de unión al marcador de ácido nucleico no tienen repeticiones directas o invertidas que sean mayores de aproximadamente 7, 8, 9, 10, 11, 12, 13, 14, 15, 16 o más bases de nucleótidos de longitud. En algunas realizaciones, las regiones de unión al marcador de ácido nucleico no contienen ninguna repetición invertida de 9 nucleótidos o más a lo largo de una secuencia que 1.100 pares de bases de longitud. En algunas realizaciones, las regiones de unión al marcador de ácido nucleico no contienen ninguna repetición invertida de 7 nucleótidos o más a lo largo de cualquier región de 100 pares de bases.

Las secuencias también se pueden explorar para evitar los sitios de reconocimiento de las enzimas de restricción que cortan seis bases comunes. Las secuencias seleccionadas se pueden someter además al análisis de la estructura secundaria predicha, y aquellas con la mínima estructura secundaria se pueden seleccionar para una evaluación adicional. Se puede usar cualquier programa conocido en la técnica para predecir la estructura secundaria, tal como el programa MFOLD (Zuker, 2003, *Nucleic Acids Res.* 31 (13):3406-15; Mathews *et al.*, 1999, *J. Mol. Biol.* 288:911-940).

La estructura principal del nanoindicador, que en algunas realizaciones es una molécula de ácido nucleico monocatenario, se diseña para que tenga una o más regiones de unión al marcador, que comprenden un patrón regular de una determinada base (la "base repetida de forma regular"). En dichas regiones, la base repetida de forma regular normalmente aparece con una periodicidad de aproximadamente cada n restos, donde n es cualquier número, y preferentemente de 4 a 25. Preferentemente, no más del 25 % de las bases repetidas de forma regular en una región de unión al marcador aparece en un intervalo distinto de dichos intervalos regulares. Por ejemplo, si en una región de unión al marcador de 100 nucleótidos existen 12 bases de timina, y la timina es la base repetida de forma regular, no más del 25 % de estas, es decir, 3 bases de timina, aparecen fuera del patrón regular de timinas. De igual forma, si en una región de unión al marcador de 100 nucleótidos existen 12 bases de adenina, y la adenina es la base repetida de forma regular, no más del 25 % de estas, es decir, 3 bases de adenina, aparecen fuera del patrón regular de adeninas. En realizaciones específicas, 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 % o no más del 1 % de dichas bases aparecen en otros intervalos distintos de dichos intervalos regulares

de dicha región. En algunas realizaciones, las bases no están separadas de forma regular, pero las bases son están, al menos, una media de 8, 9, 10, 12, 15, 16, 18, 20, 30 o 50 nucleótidos de separación. En algunas realizaciones, las bases no están separadas de forma regular, pero las bases están, al menos, una media de 8 nucleótidos de separación.

En ciertas realizaciones, la base repetida de forma regular en una región de unión al marcador dada está separada al menos aproximadamente cada 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 o más bases de nucleótidos. En ciertos aspectos, la base repetida de forma regular está espaciada aproximadamente cada 5 a 15 bases de nucleótidos, aproximadamente cada 6 a 12 bases de nucleótidos, aproximadamente cada 7 a 10 bases de nucleótidos o aproximadamente cada 8 a 10 bases de nucleótidos. En ciertos aspectos, la base repetida de forma regular está separada aproximadamente cada 8 bases de nucleótidos. En ciertas realizaciones específicas, una base adenina se repite al menos cada 8 bases de nucleótidos. En ciertas realizaciones específicas, una base de adenina se repite aproximadamente cada 15 a 20 bases de nucleótidos. En ciertas realizaciones específicas, una base de adenina se repite aproximadamente cada 16 bases de nucleótidos.

La base repetida de forma regular puede estar presente en una la secuencia polinucleotídica de una región de unión al marcador dada en un contenido de nucleótidos del aproximadamente 5 %, 6 %, 7 %, 8 %, 9 %, 10 %, 11 %, 12 %, 13 %, 14 %, 15 %, 16 %, 17 %, 18 %, 19 %, 20 %. La presencia de una base repetida de forma regular en una media de aproximadamente cada 8 a 16 nucleótidos en una región de unión al marcador representa un contenido de nucleótidos del aproximadamente 12,5 %. En ciertos aspectos, el contenido de nucleótidos de la base repetida de forma regular puede ser del aproximadamente 12,5 %. En ciertos aspectos, la región de unión al marcador comprende un número seleccionado o porcentaje seleccionado de bases repetidas de forma regular por región, que se distribuyen aleatoriamente a excepción de que deben estar separadas al menos una cierta distancia mínima, tal como al menos aproximadamente 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 o más bases de separación.

La base repetida de forma regular puede ser cualquier base de nucleótido de adenina, timina, guanina, citosina o uracilo. En ciertas realizaciones, la base repetida de forma regular en la región de unión al marcador monocatenaria es adenina.

En ciertos aspectos, la secuencia polinucleotídica de las regiones de unión al marcador también está diseñada para que tenga un cierto contenido de guanina/citosina (G/C). Por ejemplo, ciertas regiones de unión al marcador están diseñadas para tener un contenido total de G/C del aproximadamente 45 %, 46 %, 47 %, 48 %, 49 %, 50 %, 51 %, 52 %, 54 %, 55 %, 56 %, 57 %, 58 %, 59 %, 60 %, más preferentemente aproximadamente un 50 %, ya que esto proporciona una Tf bastante uniforme entre las numerosas secuencias polinucleotídicas, únicas, de las regiones de unión al marcador. En ciertos aspectos, las secuencias polinucleotídicas tienen no más de tres guanines en una fila, para reducir al mínimo los problemas de síntesis. El contenido de GC total también es preferentemente uniforme en tramos relativamente cortos para hacer Tf locales similares. Sin pretender limitarse a ninguna teoría, cuando el nanoindicador es un ácido nucleico bicatenario, la Tf de los nanoindicadores descritos en el presente documento proporciona enlaces más fuertes entre la estructura principal del nanoindicador y la secuencia polinucleotídica complementaria que tiene unida a la misma una molécula detectable, evitándose así la disociación durante los procedimientos de síntesis y de hibridación. Además, la Tf uniforme entre la población de nanoindicadores permite la buena optimización de los procedimientos de síntesis y de hibridación, ya que las condiciones óptimas son las mismas para todas las manchas y posiciones. En algunas realizaciones, las secuencias de los nanoindicadores tienen aproximadamente el 50 % de guanina/citosina (G/C), con no más de tres G en una fila.

En ciertos aspectos, la secuencia polinucleotídica de las regiones de unión al marcador también está diseñada para que tenga una cierta proporción de G/C en una cadena dada. Por ejemplo, en ciertas realizaciones, la secuencia polinucleotídica de las regiones de unión al marcador están diseñadas de modo que las secuencias polinucleotídicas complementarias tengan una proporción de G/C de aproximadamente 1/1, 2/1, 3/2, aproximadamente 5/3 o aproximadamente 7/5. En ciertos aspectos, la proporción de G/C es de al menos 1/1. En ciertos aspectos, la proporción de G/C es de aproximadamente 1/1, 2/1, 3/2, 5/3 o aproximadamente 7/5 en una cadena dada. En ciertos aspectos, la proporción de G/C es de al menos 1/1 en la secuencia polinucleotídica complementaria. En ciertos aspectos, la proporción de G/C es de aproximadamente 3/2 en la secuencia polinucleotídica complementaria.

En ciertas realizaciones específicas, en las que la adenina es la base repetida de forma regular, espaciada al menos aproximadamente cada 8 nucleótidos (aproximadamente un contenido de adenina del 12,5 %), y el contenido de G/C es del aproximadamente 50 %, el contenido de timina puede ser del aproximadamente 37-38 % o del aproximadamente 37,5 %. En ciertas realizaciones específicas, en las que la adenina es la base repetida de forma regular, y el contenido de G/C es del aproximadamente 50 %, en algunas realizaciones, el contenido de timina puede ser del aproximadamente 35-45 %. En una región de unión al marcador ilustrativa, en la que el nucleótido repetido de forma regular es adenina y un contenido de GC del aproximadamente 50 %, el exceso de adeninas se puede utilizar para compensar la pérdida en la abundancia de T. Para generar la secuencia seleccionada, se pueden crear secuencias aleatorias con patrones fijos de adeninas que varían de cada 4 bases a cada 25 bases, y explorarse para reducir al mínimo la presencia de repeticiones invertidas y directas.

En algunas realizaciones, la presente divulgación proporciona poblaciones de nanoindicadores o de unidades marcadoras de nanoindicadores, por ejemplo, bibliotecas de nanoindicadores o de unidades marcadoras de nanoindicadores que contienen al menos 3, al menos 5, al menos 10, al menos 15, al menos 20, al menos 25, al menos 30, al menos 35, al menos 40, al menos 45, al menos 50, al menos 60, al menos 70, al menos 80, al menos 90, al menos 100, al menos 200, al menos 300, al menos 400, al menos 500, al menos 750 o al menos 1.000 nanoindicadores únicos, unidades marcadoras de nanoindicadores únicas o nanoindicadores marcados únicos. Como se usa en el presente documento, el término "único" cuando se usa en referencia a un nanoindicador o unidades marcadoras de nanoindicadores de una población pretende significar un nanoindicador que tiene un código que lo distingue de otros nanoindicadores de la misma población, por ejemplo, cada nanoindicador tiene una señal detectable que lo distingue de otros nanoindicadores de dicha población. Por lo general, un nanoindicador se vuelve "único" por la combinación lineal particular de las regiones de unión al marcador, teniendo cada una unido a la misma un monómero marcador seleccionado.

En realizaciones específicas, las poblaciones de nanoindicadores comprenden al menos 1.000, 5.000, al menos 10.000, al menos 20.000 o al menos 50.000 nanoindicadores únicos o unidades marcadoras de nanoindicadores únicas.

Los nanoindicadores de una población de nanoindicadores pueden ser nanoindicadores simples, nanoindicadores dobles o una combinación de los mismos. Los nanoindicadores pueden estar marcados o no marcados. La población de nanoindicadores descrita en el presente documento puede comprender dos o más nanoindicadores.

En algunas realizaciones, la invención proporciona una población de estructuras principales de nanoindicador en la que una base (por ejemplo, adenina) está espaciada de forma regular, los nanoindicadores no tienen ninguna formación significativa de estructuras secundarias y la Tf entre los nanoindicadores de la población es bastante uniforme. En algunas realizaciones, la invención proporciona una población de nanoindicadores en la que la Tf de las secuencias polinucleotídicas complementarias cuando se hibridan a sus regiones de unión al marcador es de aproximadamente 80 °C o superior.

En algunas realizaciones, la divulgación proporciona una población de estructuras principales de nanoindicador en la que cada estructura principal de nanoindicador comprende una región que comprende una pluralidad de regiones de unión al marcador unidas covalentemente entre sí en una combinación lineal, en la que cada región de unión al marcador comprende aproximadamente de 800 a 1.300 bases de nucleótidos y tiene un contenido de G/C del aproximadamente 50 %, en la que cada región de unión al marcador seleccionada es diferente de las otras regiones de unión al marcador seleccionadas, y en la que cada región de unión al marcador se hibrida con una secuencia polinucleotídica complementaria que tiene unidas a la misma una o más moléculas detectables, en la que la secuencia polinucleotídica complementaria tiene una proporción de G/C de al menos 1/1. En algunas realizaciones, la secuencia polinucleotídica complementaria tiene una proporción de G/C de aproximadamente 3/2. En algunas realizaciones, la estructura principal del nanoindicador está diseñada para sesgar el contenido específico de la cadena de la estructura principal de manera que la estructura principal sea rica en CT. En algunas realizaciones aproximadamente el 60 % de la estructura principal del nanoindicador se compone de regiones de CT.

En algunas realizaciones, la divulgación proporciona una población de estructuras principales de nanoindicador en la que cada estructura principal de nanoindicador comprende una estructura principal de ácido nucleico monocatenario, comprendiendo dicha estructura principal una pluralidad de regiones de unión al marcador unidas covalentemente entre sí en una combinación lineal única, en la que dicha estructura principal del nanoindicador comprende una base (por ejemplo, adenina) espaciada de forma regular, las estructuras principales de los nanoindicadores no tienen ninguna formación significativa de estructuras secundarias y contienen menos del 85 % de homología con cualquier otra secuencia usada en otras estructuras principales de la población o con cualquier secuencia descrita en la base de datos pública REFSEQ.

En algunas realizaciones, la divulgación proporciona una población de estructuras principales de nanoindicador en la que cada estructura principal de nanoindicador comprende una estructura principal de ácido nucleico monocatenario, comprendiendo dicha estructura principal una pluralidad de regiones de unión al marcador unidas covalentemente entre sí en una combinación lineal única, en la que dicha estructura principal del nanoindicador tiene un contenido de G/C del aproximadamente 50 %, las estructuras principales de nanoindicadores no tienen ninguna formación significativa de estructuras secundarias y contienen menos del 85 % de homología con cualquier otra secuencia usada en otras estructuras principales de la población o con cualquier secuencia descrita en la base de datos pública REFSEQ.

En algunas realizaciones, la divulgación proporciona una población de estructuras principales de nanoindicador en la que cada estructura principal de nanoindicador comprende una estructura principal de ácido nucleico monocatenario, comprendiendo dicha estructura principal una pluralidad de regiones de unión al marcador unidas covalentemente entre sí en una combinación lineal única, en la que dicha estructura principal de nanoindicador tiene un contenido de G/C del aproximadamente 50 %, las estructuras principales de los nanoindicadores no tienen ninguna formación significativa de estructuras secundarias y contienen menos del 1 % de secuencias de repetición directa.

En algunas realizaciones, la divulgación proporciona una población de estructuras principales de nanoindicador en la que cada estructura principal de nanoindicador comprende una estructura principal de ácido nucleico monocatenario, comprendiendo dicha estructura principal una pluralidad de regiones de unión al marcador unidas covalentemente entre sí en una combinación lineal única, en la que las regiones de unión al marcador se seleccionan de una población de secuencias polinucleotídicas diseñadas, únicas, en la que cada secuencia polinucleotídica se hibrida a una secuencia polinucleotídica complementaria única que tiene unidas a la misma una o más moléculas detectables, en la que cada nanoindicador tiene una señal detectable que lo distingue de otros nanoindicadores de dicha población, y en la que cada nanoindicador es más estable que un nanoindicador comparable que comprende ADN de M13.

En algunas realizaciones, la divulgación proporciona una población de estructuras principales de nanoindicador en la que cada estructura principal del nanoindicador comprende una estructura principal de ácido nucleico monocatenario, comprendiendo dicha estructura principal una pluralidad de regiones de unión al marcador unidas covalentemente entre sí en una combinación lineal única, en la que las regiones de unión al marcador se seleccionan de una población de secuencias polinucleotídicas diseñadas, únicas, en la que cada secuencia polinucleotídica se hibrida a una secuencia polinucleotídica complementaria única que tiene unidas a la misma una o más moléculas detectables, en la que cada nanoindicador tiene una señal detectable que lo distingue de otros nanoindicadores de dicha población, en la que la Tf de las secuencias polinucleotídicas complementarias cuando se hibridan a sus regiones de unión al marcador en los nanoindicadores descritos en el presente documento son más altas que la Tf de una secuencia polinucleotídica complementaria a un molde de ADN de M13 cuando se hibrida a una sonda de nanoindicador que comprende ADN de M13.

En algunas realizaciones, la divulgación proporciona una población de estructuras principales de nanoindicador en la que las características de la secuencia de las sondas de nanoindicador descritas en el presente documento proporcionan una mejor detección de una molécula diana. Por ejemplo, cuando se realiza la unión de las sondas de nanoindicador con las moléculas diana contando individualmente la presencia de una o más moléculas del nanoindicador en una muestra, las sondas de nanoindicador descritas en el presente documento permiten un aumento en el número de recuentos. En algunas realizaciones, el número de recuentos moleculares por encima del fondo de dicho complejo molecular tras la normalización de la muestra es superior a 300, 400, 450, 500, 600, 700, 800, 900 o 1.000 recuentos moleculares. En algunas realizaciones, las sondas de nanoindicador descritas en el presente documento permiten un aumento del número de recuentos. En algunas realizaciones, el número de recuentos moleculares por encima del fondo de dicho complejo molecular tras la normalización de la muestra es superior a 400 recuentos moleculares. En algunas realizaciones, el porcentaje de los recuentos moleculares válidos de las sondas de nanoindicador descritas en el presente documento es superior al aproximadamente 10, 11, 12, 12,5, 14, 15, 20, 25, 30, 35, 40, 50, 60, 70, 80 o 90 %. En algunas realizaciones, el porcentaje de recuentos válidos moleculares de las sondas de nanoindicador descritas en el presente documento es superior al aproximadamente 12,5 %. En algunas realizaciones, el número de recuentos moleculares por encima del fondo de las sondas de nanoindicador descritas en el presente documento tras la normalización de dicha muestra es al menos 2, 5, 6, 10, 15, 20, 30, 40, 50, 60, 70, 80 o 90 veces superior al de un nanoindicador comparable que comprende ADN de M13.

En algunas realizaciones, el número de recuentos moleculares por encima del fondo de las sondas de nanoindicador descritas en el presente documento tras la normalización de dicha muestra es al menos 6 veces superior al de un nanoindicador comparable que comprende ADN de M13. En algunas realizaciones, el número de recuentos moleculares por encima del fondo de las sondas de nanoindicador descritas en el presente documento tras la normalización de dicha muestra es al menos 2 veces superior al de un nanoindicador comparable que comprende ADN de M13.

Una sonda de nanoindicador de M13 comparable es un nanoindicador que comprende la misma región específica de la diana que las sondas de nanoindicador descritas en el presente documento unidas a una estructura principal monocatenaria que comprende una pluralidad de regiones de ADN de M13 unidas covalentemente entre sí, en la que cada región se hibrida con una secuencia polinucleotídica complementaria que tiene unidas a la misma una o más moléculas detectables.

El tamaño de una población de sondas de nanoindicador y la naturaleza de las secuencias específicas de la diana de los nanoindicadores que contiene dependerán del uso previsto del nanoindicador. Se pueden crear poblaciones de sondas de nanoindicador en las que las secuencias específicas de la diana correspondan a los marcadores de un determinado tipo de células, incluyendo un tipo de célula enferma. En una realización específica, la célula o el tejido es una célula o un tejido de mamífero, y más preferentemente es una célula o un tejido humano.

En algunas realizaciones, una secuencia apropiada se divide en regiones de unión al marcador que varían de 50 bases a 2 kilobases de longitud (pudiendo ser más largas). Las regiones de unión al marcador también pueden variar de aproximadamente 200 bases a aproximadamente 1.800 bases, de aproximadamente 400 bases a aproximadamente 1.600 bases, de aproximadamente 600 bases a aproximadamente 1.500 bases, de aproximadamente 800 bases a aproximadamente 1.300 bases, de aproximadamente 900 bases a aproximadamente 1.200 bases, de aproximadamente 1.000 bases a aproximadamente 1.150 bases, de aproximadamente 1.050 bases a aproximadamente 1.150 bases, incluyendo combinaciones de estos intervalos ilustrativos y de todos los números enteros intermedios (por ejemplo, 850, 950, etc.). En ciertas realizaciones, las regiones de unión al marcador son de

aproximadamente 1.100 bases.

5 En algunas realizaciones, cada región de unión al marcador es una secuencia única, pero contiene un número y espaciado de adeninas constante y mínimo, o de otra base seleccionada, en relación con las otras regiones de unión al marcador de una secuencia indicadora dada. Estas regiones de unión al marcador pueden intercalarse con otras regiones cuya secuencia no sea relevante. Las regiones de unión al marcador de una estructura principal de nanoindicador pueden ser de diferentes longitudes y/o tener diferentes bases repetidas de forma regular.

10 Los ejemplos específicos de secuencias polinucleotídicas que se pueden utilizar como moldes para sintetizar regiones de unión al marcador monocatenarias incluyen las secuencias polinucleotídicas establecidas en SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23 y 24. Para estas secuencias, la base repetida de forma regular citada es timina, que produce adenina como la base repetida de forma regular tras la síntesis de las regiones de unión al marcador monocatenarias, normalmente en forma de una estructura principal de nanoindicador. También se contemplan las variantes de estas secuencias polinucleotídicas, incluyendo secuencias polinucleotídicas que tienen al menos un 60 %, 70 %, 80 %, 90 %, 95 %, 98 % o más de identidad de secuencia con las secuencias polinucleotídicas establecidas en SEQ ID NO: 1-24, variantes que conservan las características deseadas o adecuadas de una región de unión al marcador, según lo descrito en mayor detalle en el presente documento (por ejemplo, la base repetida de forma regular, el contenido de G/C, la proporción de G/C de una cadena dada, Tf, falta de repeticiones invertidas). Como el experto habitual en la materia apreciará, se pueden diseñar otras secuencias para que posean las características deseadas o adecuadas de una región de unión al marcador descrita en el presente documento.

25 Se puede añadir una secuencia de inicio optimizada para la transcripción por la ARN polimerasa T7, T3 o SP6 (comenzando en la posición +1 de la transcripción) al extremo 5' de cada región de unión al marcador. Los sitios de restricción se añaden opcionalmente a los límites de cada región de unión al marcador para permitir la adición específica o la eliminación específica de regiones de unión al marcador individuales en la secuencia usando técnicas de clonación convencionales.

30 El número de regiones de unión al marcador de una estructura principal de nanoindicador varía preferentemente de 1 a 50. En otras realizaciones más, el número de regiones de unión al marcador sintéticas de una estructura principal de nanoindicador varía de 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10 regiones de unión al marcador de diseño racional a 15, 20, 30, 40 o 50 regiones de unión al marcador, o cualquier intervalo intermedio. En ciertas realizaciones, el número de regiones de unión al marcador de una estructura principal es de seis.

35 En algunas realizaciones, la base repetida de forma regular en la región de unión al marcador de una estructura principal de nanoindicador o su base complementaria repetida de forma regular en una secuencia (o segmento) polinucleotídica complementaria hibridada se pueden usar para unir monómeros marcadores, preferentemente monómeros marcadores de emisión de luz, al nanoindicador en un patrón regular, espaciado uniformemente, para una mejor distribución de la señal del nanoindicador. Preferentemente, cuando una región de unión al marcador se marca, al menos el 20 %, al menos el 30 %, al menos el 40 %, al menos el 50 %, al menos el 60 %, al menos el 70 %, al menos el 80 %, al menos el 90 %, al menos el 95 % o al menos el 98 % de las apariciones de la base repetida de forma regular está unido a al menos un monómero marcador emisor de luz, ya sea por unión covalente de un monómero marcador a una base o por hibridación con un ácido nucleico en el que los complementos de la base repetida de forma regular se marcan de esta manera.

45 Este porcentaje de apariciones se puede medir mediante cualquier medio conocido en la técnica. En un método, se purifica la cantidad de ácido nucleico producida en una reacción de marcaje (por ejemplo, el ARN se puede purificar usando un kit Qiagen RNeasy) y se somete a espectrofotometría de UV. Se mide la absorbancia ("A") a las longitudes de onda apropiadas para cada uno de los ácidos nucleicos (260 nm) y el monómero marcador cuya aparición se va a medir (por ejemplo, 495 nm para Alexa Fluor 488; 590 nm para Alexa Fluor 594; 650 de Alexa Fluor 647; y 550 nm para Cy3). Se corrige la absorbancia del ácido nucleico ajustando el valor de la absorbancia a 260 nm ("A260") para eliminar la contribución del "ruido" del monómero marcador restando la absorbancia a la longitud de onda máxima para el monómero marcador (A_{MM}) menos el factor de corrección para ese monómero marcador. Cuando el ácido nucleico es ARN, el número de monómeros marcadores por cada mil nucleótidos se calcula de acuerdo con la fórmula:

$$\frac{\text{n.º de monómeros marcadores}}{1.000 \text{ nucleótidos}} = \frac{A_{260}}{A_{MM}} \times \frac{9010}{CE_{MM}} \times 1000$$

60 en la que CE_{MM} es el coeficiente de extinción para el monómero marcador. A partir de esta fórmula, se puede calcular el porcentaje de apariciones de la base repetida de forma regular que se une a un monómero marcador de emisión de luz.

En algunas realizaciones, la base repetida de forma regular preferida de una región de unión al marcador es la adenina, de manera que la región se puede marcar mediante la hibridación a una o más secuencias polinucleotídicas complementarias (por ejemplo, segmentos de ARN) en las que la base repetida de forma regular sea el uracilo. En algunas realizaciones, la adenina no se repite de forma regular, pero está espaciada al menos una

5 media de 8 bases. Esto permite el uso de UTP modificados con aminoalilo, que se pueden obtener fácilmente a nivel comercial, como sitios de unión al monómero marcador, en una secuencia por lo demás aleatoria. Preferentemente, además de la periodicidad regular de las regiones de unión al marcador, las regiones (y el ácido nucleico que las comprende) contienen una estructura secundaria mínima y Tf bastante uniforme.

10 Los ácidos nucleicos se pueden sintetizar químicamente usando nucleótidos de origen natural o nucleótidos modificados de diversas formas diseñados para aumentar la estabilidad biológica de las moléculas o para aumentar la estabilidad física del dúplex formado entre la región de unión al marcador y las secuencias, o segmentos, polinucleotídicas complementarias hibridadas, por ejemplo, se pueden usar derivados de fosforotioato y nucleótidos sustituidos con acridina. Los ejemplos de los nucleótidos modificados que se pueden usar para generar el ácido

15 nucleico sintético incluyen 5-fluorouracilo, 5-bromouracilo, 5-clorouracilo, 5-yodouracilo, hipoxantina, xantina, 4-acetilcitosina, 5-(carboxihidroxi)metiluracilo, 5-carboximetilaminometil-2-tiouridina, 5-carboximetilaminometiluracilo, dihidouracilo, beta-D-galactosilqueosina, inosina, N6-isopenteniladenina, 1-metilguanina, 1-metilinosina, 2,2-dimetilguanina, 2-metiladenina, 2-metilguanina, 3-metilcitosina, 5-metilcitosina, N6-adenina, 7-metilguanina, 5-metilaminometiluracilo, 5-metoxiaminometil-2-tiouracilo, beta-D-manosilqueosina, 5'-metoxicarboximetiluracilo, 5-

20 metoxiuracilo, 2-metil-N6-isopenteniladenina, ácido uracil-5-oxiacético (v), wibutoxosina, pseudouracilo, queosina, 2-tiocitosina, 5-metil-2-tiouracilo, 2-tiouracilo, 4-tiouracilo, 5-metiluracilo, éster metílico de ácido uracil-5-oxiacético, ácido uracil-5-oxiacético (v), 5-metil-2-tiouracilo, 3-(3-amino-3-N-2-carboxipropil)uracilo, (acp3)w y 2,6-diaminopurina.

Como alternativa, el ácido nucleico sintético se puede producir biológicamente usando un vector en el que se haya subclonado un ácido nucleico. Como ejemplo, se puede crear una estructura principal de ADN monocatenario lineal a partir de un ADN bicatenario de plásmido usando un protocolo de cuatro etapas que incluya (i) la linealización del ADNbc con una enzima de restricción; (ii) la desfosforilación con una fosfatasa termolábil; (iii) la digestión con una segunda enzima de restricción para separar el vector de clonación de la secuencia de la estructura principal; y (iv) la digestión con una exonucleasa lambda específica de la cadena, dejando solo una de las cadenas del fragmento de

25 estructura principal intacta.

En diversas realizaciones, las moléculas de ácido nucleico se pueden modificar en la fracción de la base, la fracción de azúcar o la estructura principal de fosfato para mejorar, por ejemplo, la estabilidad, hibridación o solubilidad de la molécula. Por ejemplo, la estructura principal de desoxirribosa fosfato de los ácidos nucleicos se puede modificar para generar ácidos peptidonucleicos (véase Hyrup *et al.*, 1996, *Bioorganic & Medicinal Chemistry* 40:5-23). Como se usan en el presente documento, los términos "ácidos peptidonucleicos" o "PNA" hacen referencia a un mimético de ácido nucleico, por ejemplo, miméticos de ADN, en los que la estructura principal de desoxirribosa fosfato es

35 reemplazada por una estructura principal de pseudopéptido y solo se conservan las cuatro nucleobases naturales. Se ha demostrado que la estructura principal neutra de los PNA permite la hibridación específica al ADN y ARN en condiciones de baja fuerza iónica. La síntesis de oligómeros de PNA se puede realizar usando protocolos de síntesis peptídica en fase sólida convencionales como describe Hyrup *et al.*, 1996, en *Bioorganic & Medicinal Chemistry* 4(1): 5-23; Perry-O'Keefe *et al.*, 1996, *Proc. Natl. Acad. Sci. EE.UU.* 93:14670-675.

En una realización ilustrativa, se puede construir la secuencia sintética novedosa seleccionada sintéticamente en forma de un ADN bicatenario por medio de una empresa comercial de síntesis de genes y se puede clonar de una manera orientada en un "fagémido", un vector plasmídico que contenga una región intergénica (IG) de un fago M13 o f1 que contiene las secuencias que actúan en cis necesarias para la replicación del ADN y la encapsidación del fago, tal como pUCI 19. La orientación apropiada de la inserción clonada con respecto al origen de replicación del fago permite la generación de una estructura principal de ADN monocatenario que es el complemento inverso de las moléculas de ARN generadas mediante la transcripción *in vitro* para cada región de unión al marcador.

45

50

Para generar la cadena principal de ADN monocatenario del nuevo indicador, el fagémido se transforma en una cepa de *E. coli* que contiene un episoma F'. La posterior infección de la bacteria transformada con un fago auxiliar tal como el mutante K07 de M13 produce la secreción del fagémido que porta la nueva secuencia indicadora en forma de un fago empaquetado monocatenario a partir del cual se prepara el ADN monocatenario circular usando un protocolo normalizado. Este ADN se linealiza y la parte del vector se separa mediante la hibridación de oligonucleótidos complementarios, cortos, a cualquier extremo de la nueva secuencia indicadora para generar sitios de restricción bicatenarios, seguido del tratamiento con las enzimas de restricción apropiadas. La Figura 9 muestra un ejemplo de vector plasmídico que se puede usar para clonar, propagar y generar los nanoindicadores descritos en el presente documento.

55

60

Secuencias polinucleotídicas complementarias

En algunas realizaciones, las moléculas detectables o los monómeros marcadores que proporcionan o emiten señales que constituyen todo o parte del código del nanoindicador están unidos a la región o a las regiones de unión al marcador de la estructura principal del nanoindicador a través de una estructura denominada en el presente

65

documento "secuencia polinucleotídica complementaria". Las moléculas detectables pueden estar unidas directamente (por ejemplo, covalente o no covalentemente) a una secuencia polinucleotídica complementaria, o unidas indirectamente a dicha secuencia (por ejemplo, a través de un componente intermedio, tal como un ligando).

5 En algunas realizaciones, las secuencias polinucleotídicas complementarias pueden tener de 25 nucleótidos a varias kilobases (por ejemplo, 5 kb) de longitud, y tienen preferentemente de 50 nucleótidos a 2 kb de longitud. En realizaciones específicas, las secuencias polinucleotídicas complementarias tienen aproximadamente de 25 a 250, de 50 a 200, de 50 a 150 o de 50 a 100 nucleótidos de longitud. En otras realizaciones, las secuencias polinucleotídicas complementarias tienen aproximadamente de 500 a 2.000, de 500 a 1.500, de 500 a 1.000, de 750 a 1.250 o de 750 a 1.000 nucleótidos de longitud. En ciertos aspectos, las secuencias polinucleotídicas complementarias tienen aproximadamente 800, 850, 900, 950, 1.100, 1.150 o 1.100 nucleótidos de longitud, incluyendo todos los números enteros intermedios. Las secuencias polinucleotídicas complementarias pueden ser polinucleótidos de ARN complementarios o polinucleótidos de ADN complementarios. En una realización preferida, las secuencias polinucleotídicas complementarias son polinucleótidos de ARN complementarios.

15 Una molécula detectable o un monómero marcador se pueden unir covalentemente a una secuencia polinucleotídica complementaria antes o después de que la secuencia se una a la región de unión al marcador de una estructura principal de nanoindicador. Por ejemplo, en la unión de una molécula detectable a una secuencia polinucleotídica complementaria, el marcador se puede unir covalentemente mediante la incorporación de un nucleótido que contenga una molécula detectable al ácido nucleico durante la síntesis, pero antes de que se hibride a la región de unión al marcador de la estructura principal. Como alternativa, durante la síntesis de una secuencia polinucleotídica complementaria, se puede incluir un nucleótido que contenga un grupo aceptor de moléculas detectables, y se puede añadir la molécula detectable a la secuencia polinucleotídica complementaria tras su síntesis, ya sea antes o después de su hibridación a la región de unión al marcador de la estructura principal. Como alternativa, el monómero marcador se puede unir indirectamente a la secuencia polinucleotídica complementaria, por ejemplo, mediante la incorporación de un nucleótido que contenga una molécula de unión al ligando (por ejemplo, biotina) a la secuencia polinucleotídica complementaria durante la síntesis, y mediante la adición de un ligando (por ejemplo, estreptavidina) que esté unido covalentemente a la molécula detectable.

30 En algunas realizaciones, una secuencia polinucleotídica complementaria puede variar de 20 nucleótidos a más de 5 kb de longitud, dependiendo del método de ensamblaje del nanoindicador. Por ejemplo, cuando una secuencia polinucleotídica complementaria tiene incorporado covalentemente uno o más monómeros marcadores que emiten señales que son parte del código del nanoindicador en el contexto del nanoindicador marcado, la secuencia polinucleotídica complementaria tiene preferentemente de aproximadamente 100 a aproximadamente 10.000 bases, más preferentemente de 200 a aproximadamente 2.000 bases, y aún más preferentemente de 700 a aproximadamente 1.200 nucleótidos de longitud, o aproximadamente 1.100 pares de bases de longitud, y, en general, en el presente documento, se denomina segmento, siendo un segmento oscuro la secuencia polinucleotídica complementaria antes de la incorporación del monómero marcador (pero, en una realización preferida, que contiene sitios aceptores de monómeros marcadores, tales como nucleótidos de aminoalilo), y siendo un segmento coloreado aquel que contiene el monómero marcador o los monómeros marcadores deseados. La T_f de un segmento cuando se hibrida con su región de unión al marcador es preferentemente >80 °C, más preferentemente >90 °C, en Na^+ 825 mM (5 x SSC).

45 Al igual que con las regiones de unión al marcador, una secuencia polinucleotídica complementaria puede comprender un patrón de bases repetidas de forma regular o puede comprender un porcentaje similar de bases, tales como una base de repetición de guanina, adenina, timina, citosina o uracilo. Cuando la base de repetición de la región de unión al marcador monocatenaria es adenina, la base de repetición de la secuencia polinucleotídica complementaria puede ser timina si la secuencia polinucleotídica complementaria es un polinucleótido de ADN complementario, o uracilo si la secuencia polinucleotídica complementaria es un polinucleótido de ARN complementario.

50 La base repetida de forma regular de una secuencia polinucleotídica complementaria dada puede estar separada aproximadamente cada 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 o más bases de nucleótidos. En ciertos aspectos, la base repetida de forma regular está espaciada aproximadamente cada 5 a 15 bases de nucleótidos, aproximadamente cada 6 a 12 bases de nucleótidos, aproximadamente cada 7 a 10 bases de nucleótidos o aproximadamente cada 8 a 10 bases de nucleótidos. En ciertos aspectos, la base repetida de forma regular está espaciada aproximadamente cada 8 bases de nucleótidos. En ciertos aspectos, la base repetida de forma regular está espaciada aproximadamente cada 15 a 20 bases de nucleótidos. En ciertos aspectos, la base repetida de forma regular está espaciada aproximadamente cada 16 bases de nucleótidos.

60 En ciertos aspectos, la base repetida de forma regular es modificada con aminoalilo. Esta modificación permite la unión (por ejemplo, directa o indirecta, como anteriormente) de una molécula detectable o un monómero marcador a la base repetida de forma regular. La base modificada con aminoalilo puede ser una base de guanina, adenina, timina, citosina o uracilo. En ciertas realizaciones, la base repetida de forma regular es una base de uracilo modificada con aminoalilo.

65

En ciertos aspectos, la molécula detectable o el monómero marcador se une a la base repetida de forma regular, tal como mediante la modificación con aminoalilo de esa base. En ciertos aspectos, la molécula detectable se une a todas o a casi todas las bases repetidas de forma regular de una secuencia polinucleotídica complementaria dada. En ciertas realizaciones, la molécula detectable se une a aproximadamente el 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 % o 100 % de las bases repetidas de forma regular de una secuencia polinucleotídica complementaria dada.

En cierto aspecto, la molécula detectable o el monómero marcador se une a una base que no se repite de forma regular, pero que está al menos a 8 bases de distancia, tal como mediante la modificación con aminoalilo de esa base. El brillo de la secuencia polinucleotídica complementaria puede variar de acuerdo con el número de bases de la secuencia o el número de bases marcadas de la secuencia.

Como ejemplo específico, cuando la adenina se repite de forma regular en la región de unión al marcador monocatenaria y está presente aproximadamente cada 8 bases de nucleótidos, también puede haber un uracilo modificado con aminoalilo en el polinucleótido de ARN complementario aproximadamente cada 8 bases de nucleótidos. Además, los uracilos modificados con aminoalilo pueden incorporarse aleatoriamente en las secuencias polinucleotídicas complementarias, tal como mediante el ajuste de la proporción de bases de uracilo con respecto a las bases de uracilo modificadas en la reacción de síntesis. En estos casos, la molécula detectable, tal como un colorante fluorescente, se puede incorporar en todos o casi todos los uracilos modificados con aminoalilo (por ejemplo, aproximadamente cada 8 bases) o se puede incorporar en las bases de uracilo modificadas con aminoalilo aleatoriamente.

Al modificar el número de bases repetidas de forma regular (por ejemplo, cada 5, 6, 7, 8, 9, 10 o más bases), y la incorporación de la molécula detectable en cada base repetida de forma regular, cada dos bases repetidas de forma regular, cada tres bases repetidas de forma regular, o aleatoriamente, etc., o sus combinaciones, la molécula detectable se pueden incorporar aproximadamente cada 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 o más bases de nucleótidos, dependiendo del nivel deseado de la señal detectable.

En ciertas realizaciones, la secuencia polinucleotídica complementaria es un polinucleótido de ARN que comprende una secuencia que es complementaria a las regiones de unión al marcador monocatenarias generadas por las secuencias polinucleotídicas de molde expuestas en las SEQ ID NO: 1-24. Un polinucleótido de ARN complementario puede ser complementario a la secuencia de longitud completa de cualquiera de las regiones de unión al marcador generadas a partir de las secuencias molde de SEQ ID NO: 1-24, o puede ser complementaria a una región dentro de esa secuencia polinucleotídica, tal como una región que comprende 100, 200, 300, 400, 500, 600, 700, 800, 900, 1.000 o más bases de nucleótidos, incluyendo todos los números enteros intermedios.

En las realizaciones en las que la secuencia polinucleotídica complementaria es un polinucleótido de ARN, los nucleótidos de ARN complementarios (por ejemplo, segmentos) para cada región de unión al marcador se pueden crear usando métodos conocidos en la técnica, tales como los protocolos de transcripción *in vitro* (IVT) convencionales. Para sintetizar los nucleótidos de ARN complementarios, se pueden diseñar cebadores de la reacción en cadena de la polimerasa ("PCR") para generar un molde bicatenario a partir de un promotor de la ARN polimerasa (T7, T3 o SP6) directamente aguas arriba (5') del sitio de inicio de la transcripción y el final tras el sitio de la enzima de restricción de 3'. En ciertos aspectos, los moldes para estas reacciones IVT pueden incluir los productos de síntesis de ADN de los moldes que se describen en SEQ ID NO: 1-24, o las secuencias variantes de los mismos, que se han clonado aguas abajo de los promotores de la ARN polimerasa T7, T3 o SP6. Usando dichos moldes, se puede realizar la transcripción *in vitro* de moléculas de ARN en presencia de una base repetida de forma regular modificada con aminoalilo en el ARN (por ejemplo, UTP) y otras bases no modificadas (por ejemplo, ATP, CTP y GTP). Esto conduce a un producto de ARN en el que se modifica cada base repetida de forma regular (por ejemplo, U) para permitir el acoplamiento covalente de un monómero marcador en esa posición de la molécula de ARN.

Las transcripciones modificadas con aminoalilo resultantes (segmentos de aa) se pueden acoplar entonces a diversos fluoróforos (por ejemplo, Alexa 488 (azul), Alexa 546 (verde), Alexa 594 (amarillo) y Alexa 647 (rojo) - pudiéndose usar también otros colorantes). Los fluoróforos Alexa se pueden adquirir en forma de succinimidil-ésteres o TFP-ésteres (Invitrogen™), y se pueden acoplar a las transcripciones mediante la formación de enlaces de amida usando protocolos convencionales. El nivel de incorporación de colorante en un segmento se correlaciona con el número de posibles sitios de unión (fracciones de aminoalilo (aa)). Para fabricar los segmentos de color más brillantes posibles, se usa el 100 % de aaUTP. Para cambiar el brillo, se puede usar una mezcla de aaUTP y UTP sin modificar en cualquier proporción en las reacciones de IVT para modificar el número de sitios de aa presentes en un segmento de aa. El brillo deseado varía con la óptica del sistema de detección.

Moléculas detectables o monómeros marcadores

Los nanoindicadores de la presente invención pueden marcarse con cualquiera de una variedad de monómeros marcadores, tales como un radioisótopo, fluorocromo, colorante, enzima, nanopartícula, marcador quimioluminiscente, biotina u otro monómero conocido en la técnica que se pueda detectar directamente (por

ejemplo, mediante la emisión de luz) o indirectamente (por ejemplo, mediante la unión de un anticuerpo marcado con fluorescencia). En general, una o más de las regiones de unión al marcador del nanoindicador se marca con uno o más monómeros marcadores, y las señales proporcionadas por los monómeros marcadores unidos a las regiones de unión al marcador de un nanoindicador constituyen un código detectable que identifica la diana a la que se une la región específica de la diana del nanoindicador. En ciertas realizaciones, la falta de una señal dada de la región de unión al marcador (por ejemplo, una mancha oscura) también puede constituir parte del código del nanoindicador.

En ciertas realizaciones, por ejemplo, cuando una estructura principal del nanoindicador se construye a partir de una biblioteca (por ejemplo, los moldes de las regiones de unión al marcador de ejemplo de SEQ ID NO: 1-24), se puede asignar a cada región de unión al marcador única una molécula detectable dada. Meramente a modo de ilustración específica, en una biblioteca de las 24 regiones de unión al marcador únicas, por ejemplo, productos de los moldes de SEQ ID NO: 1-24, se puede asignar a los productos de los moldes de SEQ ID NO: 1, 5, 9, 13, 17 y 21 una primera molécula detectable, tal como un fluoróforo azul, se puede asignar a los productos de los moldes de SEQ ID NO: 2, 6, 10, 14, 18 y 22 una segunda molécula detectable, tal como un fluoróforo verde, se puede asignar a los productos de los moldes de SEQ ID NO: 3, 7, 11, 19 y 23 una tercera molécula detectable, tal como un fluoróforo amarillo, y se puede asignar a los productos de los moldes de SEQ ID NO: 4, 8, 12, 20 y 24 una cuarta molécula detectable, tal como un fluoróforo rojo. El experto en la materia apreciará que este ejemplo es meramente ilustrativo, y que se pueden utilizar numerosas combinaciones o alternativas para realizar el objetivo de crear una población de nanoindicadores detectables de forma única, incluyendo el uso de bibliotecas de diferentes tamaños y con secuencias distintas a las expuestas en SEQ ID NO: 1-24.

Otro ejemplo de monómeros marcadores que se pueden utilizar mediante la invención son los fluoróforos. Se pueden usar varios fluoróforos como monómeros marcadores para marcar nucleótidos incluyendo, por ejemplo, fluoresceína, tetrametilrodamina y Rojo Texas. Se conocen varios fluoróforos diferentes, y se siguen produciendo más, que abarcan el espectro completo. Además, se han producido diferentes formulaciones del mismo fluoróforo para diferentes aplicaciones. Por ejemplo, se puede usar fluoresceína, en su forma de isotiocianato (FITC), en forma de isómeros mezclados o isómeros individuales de succinimidiléster de carboxifluoresceína (FAM), o como formas de diclorotriazina isoméricas de fluoresceína (DTAF). Estos monómeros son químicamente distintos, pero todos emiten luz con un pico entre 515-520 nm, generando de ese modo una señal similar. Además de las modificaciones químicas de la fluoresceína, se han sintetizado fluoróforos completamente diferentes que tienen los mismos picos de emisión o muy similares a los de la fluoresceína. Por ejemplo, el colorante Verde Oregón tiene espectros de excitación y emisión casi superponibles en comparación con la fluoresceína. Otros fluoróforos tales como Verde Rhodol y Verde Rodamina tienen picos de emisión solo ligeramente desplazados y, por tanto, también sirven funcionalmente como sustituyentes de la fluoresceína. Además, se han desarrollado diferentes formulaciones o colorantes relacionados en torno a otros fluoróforos que emiten luz en otras partes del espectro.

También se dispone de monómeros marcadores no radiactivos y no fluorescentes. Por ejemplo, se puede unir la biotina directamente a los nucleótidos y detectarla mediante la unión específica y de alta afinidad a la avidina o a la estreptavidina que ha sido acoplada químicamente a una enzima que cataliza una reacción colorimétrica (tal como fosfatasa, luciferasa o peroxidasa). Los nucleótidos marcados con digoxigenina también se pueden usar de manera similar para la detección no isotópica de ácidos nucleicos. Los nucleótidos biotinilados y marcados con digoxigenina se encuentran disponibles en el mercado.

También se pueden usar partículas muy pequeñas, denominadas nanopartículas, como monómeros marcadores para marcar ácidos nucleicos. Estas partículas varían de 1 a 1.000 nm de tamaño, e incluyen diversas estructuras químicas tales como partículas de oro y plata y puntos cuánticos.

Cuando se irradia con luz blanca que incide en un ángulo, las nanopartículas de plata y oro que varían de 40 a 120 nm dispersarán la luz monocromática con una alta intensidad. La longitud de onda de la luz dispersada depende del tamaño de la partícula. Cuatro a cinco partículas diferentes muy próximas dispersarán cada una luz monocromática que, cuando se superponga, dará un único color específico. Las partículas están siendo fabricadas por compañías tales como Genicon Sciences. Se pueden unir partículas de plata y oro derivatizadas a una amplia matriz de moléculas que incluyen, proteínas, anticuerpos, moléculas pequeñas, ligandos de receptores y ácidos nucleicos. Por ejemplo, la superficie de la partícula se puede derivatizar químicamente para permitir la unión a un nucleótido.

Otro tipo de nanopartículas que se puede usar como monómero marcador son los puntos cuánticos. Los puntos cuánticos son cristales fluorescentes de 1 a 5 nm de diámetro que son excitables por un amplio intervalo de longitudes de onda de luz. Estos cristales emiten luz, tal como luz monocromática, con una longitud de onda dependiente de su composición química y tamaño. Los puntos cuánticos tales como CdSe, ZnSe, InP o InAs poseen propiedades ópticas únicas.

Se pueden crear muchas docenas de clases de partículas de acuerdo con el número de clases de tamaños de los cristales de los puntos cuánticos. Las clases de tamaños de los cristales se crean bien 1) mediante un estricto control de los parámetros de formación de los cristales para crear cada clase de partículas del tamaño deseado; o 2) mediante la creación de lotes de cristales según parámetros para la formación controlada con menos rigor, seguida

de la clasificación de acuerdo con el tamaño deseado y/o las longitudes de onda de emisión. El uso de puntos cuánticos para el marcaje de partículas, en el contexto de la presente invención, es nuevo, pero en la técnica de los semiconductores es antiguo. Dos ejemplos de referencias anteriores en las que los puntos cuánticos se embeben en capas epitaxiales de silicio intrínsecas de dispositivos semiconductores de emisión/detección de luz son las patentes de los Estados Unidos n.º 5.293.050 y 5.354.707 concedidas a Chapple Sokol, *et al.*

En realizaciones específicas, una o más de las regiones de unión al marcador del nanoindicador están marcadas con uno o más colorantes emisores de luz, conteniendo cada región de unión al marcador, directa o indirectamente, una o más moléculas detectables o monómeros marcadores. La luz emitida por los colorantes puede ser luz visible o luz invisible, tal como luz ultravioleta o infrarroja. En realizaciones ilustrativas, el colorante es un colorante de transferencia de energía por resonancia de fluorescencia (FRET); un colorante de xanteno, tal como fluoresceína y rodamina; un colorante que tiene un grupo amino en la posición alfa o beta (tal como un colorante de naftilamina, 1-dimetilaminonaftil-5-sulfonato, 1-anilino-8-naftalenosulfonato y 2-p-touidinil-6-naftalensulfonato); un colorante que tiene 3-fenil-7-isocianatocumarina; una acridina, tal como 9-isotiocianatoacridina y naranja de acridina; un pireno, un benzoxadiazol y un estilbeno; un colorante que tiene 3-(ϵ -carboxipentil)-3'-etil-5,5'-dimetiloxacarbocianina (CYA); 6-carboxifluoresceína (FAM); 5 y 6-carboxirrodamina-110 (R110); 6-carboxirrodamina-6G (R6G); *N,N,N',N'*-tetrametil-6-carboxirrodamina (TAMRA); 6-carboxi-X-rodamina (ROX); 6-carboxi-4',5'-dicloro-2',7'-dimetoxifluoresceína (JOE); ALEXA Fluor™; Cy2; Rojo Texas y Rojo Rodamina; 6-carboxi-2',4,7,7'-tetraclorofluoresceína (TET); 6-carboxi-2',4,4',5',7,7'-hexaclorofluoresceína (HEX); 5-carboxi-2',4',5',7'-tetraclorofluoresceína (ZOE); NAN; NED; Cy3; Cy3.5; Cy5; Cy5.5; Cy7; yd Cy7.5; Alexa Fluor 350; Alexa Fluor 488; Alexa Fluor 532; Alexa Fluor 546; Alexa Fluor 568; Alexa Fluor 594; o Alexa Fluor 647.

Los monómeros marcadores se pueden incorporar a un nanoindicador en diferentes fases de su ensamblaje, o en un componente (por ejemplo, un nucleótido de ARN complementario del nanoindicador antes de ensamblaje en el nanoindicador).

Un monómero marcador se puede unir directamente a un nucleótido usando métodos bien conocidos en la técnica. Los nucleótidos también se pueden modificar químicamente o derivatizar para unirlos a un monómero marcador. Por ejemplo, se puede unir un monómero fluorescente tal como una molécula de fluoresceína a dUTP (desoxiuridin-trifosfato) usando un grupo aminoalquilino de cuatro átomos. Cada monómero marcador se une a un nucleótido formando un complejo de monómero marcador:nucleótido.

Este complejo de monómero marcador/nucleótido se puede incorporar a los ácidos nucleicos (por ejemplo, un parche de ADN o un oligonucleótido de detección) de una variedad de maneras. Por ejemplo, se puede incorporar un complejo de monómero marcador:nucleótido solamente en una ubicación dentro del ácido nucleico o en dos o más ubicaciones dentro del ácido nucleico.

Se dispone de fluoróforos reactivos con amina y reactivos con tiol, y se usan para el marcaje de nucleótidos y biomoléculas. En general, los nucleótidos se marcan fluorescentemente durante la síntesis química, por ejemplo, la incorporación de aminas o tioles durante la síntesis de nucleótidos permite la adición de fluoróforos. Los nucleótidos marcados fluorescentemente se encuentran disponibles en el mercado. Por ejemplo, se dispone de uridina y desoxiuridina trifosfatos que están conjugados a diez fluoróforos diferentes que cubren el espectro.

Se puede unir un nucleótido a un monómero marcador primero y después se puede incorporar a un ácido nucleico. Como alternativa, se puede marcar un ácido nucleico existente uniendo un monómero marcador a un nucleótido dentro del ácido nucleico. Por ejemplo, se pueden incorporar nucleótidos UTP modificados con aminoalil-("AA") al producto de ARN durante la transcripción. En diversas realizaciones, el 20 % o más de los nucleótidos UTP de una reacción de transcripción para generar nucleótidos de ARN complementarios se modifican con AA. En diversas realizaciones, aproximadamente el 10 % o del aproximadamente 20 % al 100 %, del 20 % al 80 %, del 30 % a 80 %, del 40 % al 60 % o del 50 % al 75 % de los UTP de una reacción de transcripción están modificados con AA, en una realización preferida, aproximadamente el 40 % o aproximadamente el 50 % de los UTP de una reacción de transcripción están modificados con AA.

Además, por ejemplo, se pueden incorporar diferentes tipos de complejos de monómero marcador:nucleótido a un único ácido nucleico, donde un componente del código del nanoindicador comprende más de un tipo de señal.

También se pueden utilizar como monómeros marcadores colorantes fluorescentes que se puedan unir directamente a los nucleótidos. Por ejemplo, FAM, JOE, TAMRA y ROX son colorantes fluorescentes reactivos con aminas que han sido unidos a nucleótidos y se usan en la secuenciación automática de ADN. Estos nucleótidos marcados fluorescentemente, por ejemplo, ROX-ddATP, ROX-ddCTP, ROX-ddGTP y ROX-ddUTP, se encuentran disponibles en el mercado.

Otros tipos de monómeros marcadores que se pueden usar para marcar un nanoindicador son los puntos cuánticos. Debido a su muy pequeño tamaño, los puntos cuánticos se pueden acoplar a oligonucleótidos directamente sin afectar a la solubilidad o al uso del oligonucleótido. En una realización preferida, solo se acopla una molécula de oligonucleótido a cada nanopartícula. Para sintetizar un complejo de oligonucleótido-nanopartícula en una

proporción 1:1 por medio de la química por lotes convencional, tanto el oligonucleótido como el nanopartícula requieren un solo grupo reactivo de diferentes clases que pueden reaccionar entre sí. Por ejemplo, si un oligonucleótido tiene un grupo amino y una nanopartícula tiene un grupo aldehído, estos grupos pueden reaccionar para formar una base de Schiff. Se puede derivatizar un oligonucleótido para unir un solo grupo amino u otro grupo funcional usando la química bien conocida en la técnica. Sin embargo, cuando una nanopartícula es derivatizada, ésta se recubre con un reactivo químico que da como resultado el recubrimiento de la superficie completa de la nanopartícula con varios grupos funcionales.

Marcadores de afinidad

Se puede usar una variedad de marcadores de afinidad conocidos en la técnica para purificar y/o inmovilizar los nanoindicadores.

Cuando se usa un marcador de afinidad para inmovilizar un nanoindicador para la detección y la formación de imágenes, este puede denominarse en el presente documento "ancla" (véase la Figura 3). En algunas realizaciones, se une un ancla de biotina al nanoindicador, permitiendo la inmovilización del nanoindicador sobre un portaobjetos recubierto con estreptavidina.

En algunas realizaciones, un nanoindicador marcado contendrá un marcador de afinidad en cada extremo, A1 y A2. El nanoindicador marcado se puede inmovilizar a una superficie a través de la unión de A1 a una pareja de afinidad inmovilizada. En ausencia de una pareja de unión de afinidad para A2, el extremo A2 del nanoindicador permanece en solución, pero en presencia de una pareja de unión de afinidad (A2'), el extremo A2 del nanoindicador también está inmovilizado. En algunas realizaciones, un nanoindicador marcado contendrá un solo marcador de afinidad, A1. Otro marcador de afinidad, A2, se puede unir al nanoindicador mediante la unión directa del nanoindicador a una molécula que contenga A2 (por ejemplo, si el nanoindicador es o comprende un ácido nucleico, puede hibridarse directamente con otro ácido nucleico al que A2 está unido). Como alternativa, cualquier marcador de afinidad se puede unir al nanoindicador marcado a través de una molécula puente, tal como el ácido nucleico puente. La Figura 3 ilustra otra realización más en la que un nanoindicador marcado (a base de ácido nucleico) contiene un marcador de afinidad, A1, en un extremo. En la Figura 3, el nanoindicador marcado se inmoviliza a través de la unión de A1 a una pareja de afinidad inmovilizada. El otro extremo del nanoindicador está en solución (Figura 3A), pero se puede inmovilizar mediante hibridación con un oligonucleótido complementario que contenga otro marcador de afinidad (A2) usado para inmovilizar el nanoindicador (Figura 3B). Tras la inmovilización de A1, el nanoindicador se puede estirar o "alargar" como se representa en la Figura 3, por ejemplo, mediante electroestiramiento, para la separación de las regiones de unión al marcador de una manera que permita la detección del código del nanoindicador. Opcionalmente, mientras que el nanoindicador está en un estado alargado, se introduce A2 y se une el extremo del nanoindicador que es complementario a A2 en sentido descendente hacia la superficie.

Un marcador de afinidad se puede usar para la unión a perlas u otras matrices para una variedad de aplicaciones útiles, incluyendo, pero sin limitación, la purificación.

Más adelante, se proporcionan ejemplos no limitantes de marcadores de afinidad adecuados. Se ha de entender que la mayor parte de los marcadores de afinidad podrían servir para un doble fin: como anclas para la inmovilización de los nanoindicadores y como marcadores para la purificación de los nanoindicadores (bien total o solo parcialmente ensamblados) o sus componentes.

En ciertas realizaciones, el marcador de afinidad es un monómero de proteína. Los ejemplos de los monómeros de proteína incluyen, pero sin limitación, las regiones constantes de inmunoglobulinas (véase Petty, 1996, "Metal-chelate affinity chromatography", en *Current Protocols in Molecular Biology*, Vol. 2, Ed. Ausubel et al., Greene Publish. Assoc. & Wiley Interscience), glutatona S-transferasa (GST; Smith, 1993, *Methods Mol. Cell Bio.* 4:220-229), la proteína de unión a maltosa de *E. coli* (Guan et al., 1987, *Gene* 67:21-30), y diversos dominios de unión a celulosa (patentes de EE.UU. n.º 5.496.934; 5.202.247; 5.137.819; Tomme et al., 1994, *Protein Eng.* 7:117-123), etc.

Otros marcadores de afinidad son reconocidos por patrones de unión específicos y, de este modo, facilitan el aislamiento y la inmovilización mediante la unión por afinidad a la pareja de unión, que puede estar inmovilizada sobre un soporte sólido. Por ejemplo, el marcador de afinidad puede ser un epítipo, y la pareja de unión un anticuerpo. Los ejemplos de dichos epítipos incluyen, pero sin limitación, el epítipo FLAG, el epítipo myc en los aminoácidos 408-439, el epítipo de la hemaglutinina del virus de la gripe (HA) o la digoxigenina ("DIG"). En otras realizaciones, el marcador de afinidad es una secuencia de proteína o aminoácidos que es reconocida por otra proteína o aminoácido, por ejemplo, la avidina/estreptavidina y la biotina.

En ciertos aspectos de la invención, el marcador de afinidad es una secuencia de nucleótidos. Se pueden usar una gran variedad de secuencias de aproximadamente 8 a aproximadamente 30 bases, más preferentemente de aproximadamente 10 a aproximadamente 20 bases, para la purificación e inmovilización de los nanoindicadores, y la secuencia se puede repetir en tándem (por ejemplo, de 1 a 10 repeticiones en tándem). Preferentemente, dicha secuencia no está ampliamente representada (es decir, presente en menos del 5 % de los genes, más preferentemente, presente en menos del 3 % de los genes y, lo más preferentemente, presente en menos del 1 %

de los genes) en la muestra que se está analizando (por ejemplo, cuando el nanoindicador se usa para la detección de ARN celular humano, la secuencia preferentemente no está ampliamente representada en el genoma humano); tiene poca o ninguna estructura secundaria o autocomplementariedad bien internamente o con copias de sí misma cuando se multimeriza (es decir, todas las estructuras secundarias del marcador multimerizado tienen preferentemente una T_f inferior a 25 °C a NaCl 1 M); no tiene una identidad significativa ni complementariedad con las secuencias de la estructura principal o del segmento (es decir, la T_f de las secuencias complementarias es preferentemente inferior a 25 °C a NaCl 0,2 M); y tiene una T_f de aproximadamente 35-65 °C, más preferentemente de aproximadamente 40-50 °C, en Na⁺ 50 mM.

En ciertas realizaciones, se usan diferentes secuencias como marcadores de purificación e inmovilización. En este caso, por ejemplo, el marcador de purificación puede ser como se ha descrito anteriormente, pero el marcador de inmovilización puede estar en el intervalo de 10 a 100 bases, con una T_f de hasta 95 °C en Na⁺ 50 mM. Una realización alternativa tendría el marcador de purificación anidado dentro del marcador de inmovilización (por ejemplo, el marcador de afinidad comprendería una secuencia de 25 bases, de las cuales 15 bases se usan como marcador de purificación y las 25 bases completas se usan como marcador de inmovilización).

En ciertos casos, el marcador de afinidad se puede usar para el marcaje de un nanoindicador además de para purificar o inmovilizar el nanoindicador.

Como los expertos en la materia apreciarán, se pueden usar muchos métodos para obtener la región codificante de los marcadores de afinidad, incluyendo, pero sin limitación, la clonación de ADN, amplificación de ADN y métodos sintéticos. Algunos de los marcadores de afinidad y reactivos para su detección y aislamiento se encuentran disponibles en el mercado.

25 **Región constante**

Los nanoindicadores de la presente invención pueden comprender al menos una región constante. Una región constante puede comprender una secuencia polinucleotídica. En ciertos aspectos, la secuencia polinucleotídica de una región constante comprende una pluralidad de secuencias de nucleótidos repetidas individuales, tal como 2, 3, 4, 5, 6, 7, 8, 9, 10 o más copias de una secuencia de repetición. En ciertos aspectos, la región constante comprende cuatro copias de una secuencia de repetición.

En ciertas realizaciones, una secuencia de repetición individual comprende aproximadamente 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 o más bases de nucleótidos. En ciertas realizaciones, una secuencia de repetición individual comprende aproximadamente 15 bases de nucleótidos.

Un ejemplo de una secuencia de repetición de 15 bases es 5'-GGTOTGTGTGATGTT-3' (SEQ ID NO: 25). En ciertas realizaciones, la región constante contiene cuatro copias de una repetición de 15 bases, tal como la secuencia de repetición de SEQ ID NO: 25.

En ciertos aspectos, la región constante se puede ligar a una estructura principal de nanoindicador tras el aislamiento y la preparación de la secuencia polinucleotídica de la estructura principal, como se describe en el presente documento. En algunas realizaciones, la región constante representa una parte permanente de la estructura principal del polinucleótido, como si estuviera fusionada con la secuencia de "codificación" de la estructura principal (es decir, la región constante se clona en la estructura principal). En ciertas realizaciones, la región constante es adyacente a un sitio de enzima de restricción para la incorporación preparada de las regiones de unión al marcador seleccionadas.

50 **Regiones específicas de la diana**

La expresión "secuencia específica de la diana" se refiere a una entidad molecular que es capaz de unirse a una molécula diana. En el contexto de un nanoindicador, la secuencia específica de la diana se ancla a la estructura principal del nanoindicador.

La secuencia específica de la diana es, en general, una secuencia de aminoácidos (es decir, una secuencia polipeptídica o peptídica) o una secuencia de ácido nucleico.

En realizaciones específicas, cuando la secuencia específica de la diana es una secuencia de aminoácidos, la secuencia específica de la diana es un fragmento de un anticuerpo, tal como un fragmento Fab' de un anticuerpo, un anticuerpo Fv monocatenario.

La secuencia específica de la diana es preferentemente una secuencia de ácido nucleico, y muy preferentemente se encuentra en un oligonucleótido que está unido bien covalentemente (por ejemplo, mediante ligación) o no covalentemente (por ejemplo, mediante hibridación) a la estructura principal del nanoindicador. Una secuencia de ácido nucleico específica de la diana tiene preferentemente al menos 15 nucleótidos de longitud, y más preferentemente tiene al menos 20 nucleótidos de longitud. En realizaciones específicas, la secuencia específica de

la diana tiene aproximadamente de 10 a 500, de 20 a 400, de 30 a 300, de 40 a 200 o de 50 a 100 nucleótidos de longitud. En otras realizaciones, la secuencia específica de la diana tiene aproximadamente de 30 a 70, de 40 a 80, de 50 a 90, o de 60 a 100, de 30 a 120, de 40 a 140 o de 50 a 150 nucleótidos de longitud.

- 5 Una secuencia de nucleótidos específica de la diana tiene preferentemente una T_f de aproximadamente 65-90 °C para cada sonda en Na^+ 825 mM (5 x SSC), lo más preferentemente aproximadamente 78-83 °C.

10 En ciertas realizaciones preferidas, la secuencia específica de la diana de cada sonda de un nanoindicador doble tiene aproximadamente de 35 a 100 nucleótidos (para una secuencia diana total de aproximadamente 70 a 200 nucleótidos, cubierta por 2 sondas), lo más preferentemente de aproximadamente 40 a 50 nucleótidos para cada sonda (para un total de aproximadamente 80 a 100 nucleótidos).

Moléculas diana

15 La expresión "molécula diana" es la molécula detectada o medida mediante la unión de un nanoindicador marcado cuya/s secuencia/s específica/s de la diana se reconoce/n (son parejas de unión específicas). Preferentemente, una molécula diana puede ser, pero sin limitación, cualquiera de las siguientes: ácido nucleico, péptido, un polipéptido/una proteína (por ejemplo, una proteína bacteriana o viral o un anticuerpo), un lípido, un hidrato de carbono, una glicoproteína, un glicolípido, una molécula pequeña, un monómero orgánico o un fármaco. Los ácidos nucleicos que se pueden analizar mediante los métodos de la presente invención incluyen: ADN bicatenario, ADN monocatenario, horquillas de ADN monocatenario, híbridos de ADN/ARN, ARN (por ejemplo, ARNm o miARN) y horquillas de ARN. En general, una molécula diana es una molécula de origen natural o un ADNc de una molécula de origen natural o el complemento de dicho ADNc.

25 Una molécula diana puede ser parte de una muestra biomolecular que contenga otros componentes o puede ser el único o principal componente de la muestra. Una molécula diana puede ser un componente de una célula o de un tejido completo, un extracto celular o tisular, un producto lisado fraccionado del mismo o una molécula esencialmente purificada. La molécula diana puede estar unida en solución o en fase sólida, incluyendo, por ejemplo, una superficie sólida tal como una microplaca, una micromatriz o una perla. Asimismo, la molécula diana puede tener una estructura o secuencia conocida o desconocida.

35 En ciertas realizaciones específicas, esa molécula diana no es un cromosoma. En otras realizaciones específicas, la molécula diana tiene un tamaño no superior a 1.000 kb (o 1 mb), un tamaño no superior a 500 kb, un tamaño no superior a 250 kb, un tamaño no superior a 175 kb, un tamaño no superior a 100 kb, un tamaño no superior a 50 kb, un tamaño no superior a 20 kb o un tamaño no superior a 10 kb. En otras realizaciones específicas más, la molécula diana está aislada de su medio celular.

Muestras biomoleculares

40 Los sistemas de nanoindicadores de la invención se pueden usar para detectar la molécula diana en cualquier muestra biomolecular. Como apreciarán los expertos en la materia, la muestra puede comprender cualquier número de objetos, incluyendo, pero sin limitación: muestras biológicas tales como células (incluyendo tanto células primarias como líneas celulares cultivadas), productos lisados o extractos celulares (incluyendo, pero sin limitación, extractos de ARN; ARNm purificado), tejidos y extractos tisulares (incluyendo, pero sin limitación, extractos de ARN; ARNm purificado); fluidos corporales (incluyendo, pero sin limitación, sangre, orina, suero, linfa, bilis, fluido cefalorraquídeo, fluido intersticial, humero vítreo y acuoso, calostro, esputo, fluido amniótico, saliva, secreciones anales y vaginales, sudor y semen, un transudado, un producto exudado (por ejemplo, fluido obtenido de un absceso o cualquier otro lugar de infección o inflamación) o fluido obtenido de una articulación (por ejemplo, una articulación normal o una articulación afectada por una enfermedad tal como artritis reumatoide, osteoartritis, gota o artritis séptica) o casi de cualquier organismo, prefiriéndose las muestras de mamífero y, en particular, las muestras de ser humano; muestras medioambientales (incluyendo, pero sin limitación, muestras de aire, agrícolas, agua y suelo); muestras de agentes biológicos de conflictos bélicos; muestras de investigación incluyendo fluidos extracelulares, sobrenadantes extracelulares de cultivos celulares, cuerpos de inclusión de bacterias, compartimentos celulares, periplasma celular, compartimentos mitocondriales, etc.

55 Las muestras biomoleculares pueden derivar indirectamente de muestras biológicas. Por ejemplo, cuando la molécula diana de interés es el producto de una transcripción celular, por ejemplo, un ARN mensajero, la muestra biomolecular puede ser una muestra que contenga ADNc producido mediante la transcripción inversa del ARN mensajero. En otro ejemplo, la muestra biomolecular se genera sometiendo una muestra biológica a fraccionamiento, por ejemplo, fraccionamiento por tamaño o fraccionamiento de membrana.

60 Las muestras biomoleculares pueden ser bien nativas, por ejemplo, no sometidas a manipulación ni a tratamiento, o tratadas, que pueden incluir cualquier número de tratamientos, incluyendo la exposición a agentes candidato incluyendo fármacos, ingeniería genética (por ejemplo, la adición o la delección de un gen), etc.

65

Las muestras biomoleculares también pueden incluir muestras ambientales, tales como las que contienen bacterias u otros organismos, tales como diatomeas, dinoflagelados, algas, entre otros, tales como en ciertas muestras marinas o a base de tierra.

5 **Métodos**

En el presente documento, se describen métodos de detección y/o cuantificación de una o más moléculas diana. En particular, la invención reivindicada proporciona nanoindicadores estables que son capaces de unirse a moléculas diana individuales, y proporcionan una mejor detección de las moléculas diana. A través de los códigos de marcador de los nanoindicadores, la unión de las sondas de nanoindicador a las moléculas diana da lugar a la identificación de las moléculas diana. La detección y/o cuantificación se puede realizar usando un sistema de sondas de nanoindicador simples o un sistema de sondas de nanoindicador dobles como se describe en el presente documento.

En algunas realizaciones, la divulgación proporciona métodos de detección de la presencia de una molécula diana específica en una muestra biomolecular que comprenden poner en contacto la muestra con una sonda de nanoindicador como se describe en el presente documento en condiciones que permitan la unión de las secuencias específicas de la diana de las sondas con una molécula diana en la muestra, y la detección de la señal asociada con la sonda de nanoindicador marcada de forma única. En algunas realizaciones, la muestra biomolecular es una muestra biológica. En algunas realizaciones, la muestra biológica se selecciona entre una célula, un lisado celular, una muestra de tejido, un extracto de tejido o un fluido corporal. En algunas realizaciones, la muestra biomolecular es una muestra ambiental.

En algunas realizaciones, la divulgación proporciona métodos de detección de la presencia de una pluralidad de moléculas diana específicas en una muestra biomolecular que comprenden poner en contacto la muestra con una pluralidad de sondas de nanoindicador como se describe en el presente documento en condiciones que permitan la unión de las secuencias específicas de la diana de las sondas con una molécula diana en la muestra, y la detección de la señal asociada con las sondas de nanoindicador marcadas de forma única. En algunas realizaciones, la divulgación proporciona métodos de determinación de la presencia de al menos una molécula diana en una muestra, que comprenden formar al menos un complejo molecular que comprenda (a) al menos una molécula diana y (b) al menos una sonda de nanoindicador que comprenda una región específica de la diana única y una región que comprenda un nanoindicador diseñado, único, en la que dicho nanoindicador comprende una pluralidad de diferentes moléculas detectables, y detectar individualmente dicho al menos un complejo molecular o al menos parte de dicho al menos un complejo molecular para determinar la presencia de al menos una molécula diana en la muestra mediante un método que comprende contar individualmente la presencia de una o más moléculas de nanoindicador. En algunas realizaciones, los métodos proporcionan la determinación de la presencia de una pluralidad de moléculas diana mediante un método que comprende formar una pluralidad de complejos moleculares, comprendiendo cada complejo (a) al menos una molécula diana y (b) al menos una sonda de nanoindicador que comprende una región específica de la diana única y una región que comprende un nanoindicador diseñado, único, en la que cada nanoindicador comprende una pluralidad de diferentes moléculas detectables; y en la que cada sonda de nanoindicador comprende una región nanoindicadora diferente. En algunas realizaciones, el número de recuentos moleculares por encima del fondo de dicho complejo molecular tras la normalización de la muestra es superior a 300, 400, 450, 500, 600, 700, 800, 900 o 1.000 recuentos moleculares. En algunas realizaciones, el número de recuentos moleculares por encima del fondo del complejo molecular tras la normalización de la muestra es superior a 400 recuentos moleculares. En algunas realizaciones, el porcentaje de los recuentos moleculares válidos de las sondas de nanoindicador es superior al aproximadamente 10, 11, 12, 12,5, 14, 15, 20, 25, 30, 35, 40, 50, 60, 70, 80 o 90 %. En algunas realizaciones, el porcentaje de recuentos moleculares válidos de las sondas de nanoindicador es superior al aproximadamente 12,5 %. En algunas realizaciones, el número de recuentos moleculares por encima del fondo del nanoindicador descrito en el presente documento tras la normalización de dicha muestra es al menos 6 veces mayor que el de un nanoindicador comparable que comprende ADN de M13. En algunas realizaciones, el número de recuentos moleculares por encima del fondo del nanoindicador descrito en el presente documento tras la normalización de dicha muestra es al menos 2 veces mayor que un nanoindicador comparable que comprende ADN de M13.

En algunas realizaciones, la divulgación proporciona métodos de determinación de la presencia de al menos una molécula diana en una muestra, que comprenden (1) formar al menos un complejo molecular que comprenda (a) al menos una molécula diana y (b) al menos una sonda que comprenda una región específica de la diana única y una región que comprende un nanoindicador diseñado, único, en la que dicho nanoindicador comprende una estructura principal de ácido nucleico monocatenario, comprendiendo dicha estructura principal una pluralidad de regiones de unión al marcador unidas covalentemente entre sí en una combinación lineal única, en la que cada región de unión al marcador se hibrida a una secuencia polinucleotídica complementaria única que tiene unidas a la misma una o más moléculas detectables; y (2) detectar individualmente el al menos un complejo molecular o al menos parte de dicho al menos un complejo molecular para determinar la presencia de al menos una molécula diana en la muestra. En algunas realizaciones, el porcentaje de los recuentos moleculares válidos de las sondas de nanoindicador es superior al aproximadamente 10, 11, 12, 12,5, 14, 15, 20, 25, 30, 35, 40, 50, 60, 70, 80 o 90 %. En algunas realizaciones, el porcentaje de recuentos moleculares válidos de las sondas de nanoindicador es superior al

aproximadamente 12,5 %.

En algunas realizaciones, el número de recuentos moleculares por encima del fondo del nanoindicador descrito en el presente documento tras la normalización de dicha muestra es al menos 2 veces superior al de un nanoindicador que comprende ADN de M13. En algunas realizaciones, la T_f de las secuencias polinucleotídicas complementarias cuando se hibridan a sus regiones de unión al marcador es superior a la T_f de las secuencias polinucleotídicas complementarias a un molde de ADN de M13 cuando se hibrida con una sonda de nanoindicador que comprende el ADN de M13. La sonda de nanoindicador de M13 comprende la misma secuencia específica de la diana de la sonda de nanoindicador, y comprende una estructura principal monocatenaria que comprende una pluralidad de regiones de ADN de M13 unidas covalentemente entre sí, en la que cada región se hibrida con una secuencia polinucleotídica complementaria que tiene unidas a la misma una o más moléculas detectables. En algunas realizaciones, la divulgación proporciona métodos de determinación de la presencia de una pluralidad de moléculas diana mediante un método que comprende formar una pluralidad de complejos moleculares, comprendiendo cada complejo (a) al menos una molécula diana y (b) al menos una sonda de nanoindicador, donde cada sonda de nanoindicador tiene una temperatura de fusión superior a la de un nanoindicador comparable que comprende ADN de M13.

Los ejemplos de realizaciones específicas de la invención se muestran en las Figuras 1 y 2.

Además de detectar una señal global generada a partir de un nanoindicador marcado, la divulgación proporciona la determinación de la ubicación espacial de las señales que emanan de los monómeros marcadores (por ejemplo, las manchas) en un nanoindicador, representando cada mancha la señal combinada procedente de monómeros marcadores unidos a una región de unión al marcador determinada. Una mancha puede contener señales de la misma longitud de onda o de diferentes longitudes de onda. Por lo tanto, la naturaleza de las manchas de un nanoindicador y su ubicación constituye el código del nanoindicador.

En algunas realizaciones, el nanoindicador marcador se captura y se estira en una ubicación (véase la Figura 4). En algunas realizaciones, antes de estirar un nanoindicador, el nanoindicador se inmoviliza en una superficie sólida usando un marcador de afinidad, como se ha descrito anteriormente (véase la Figura 3). En ciertos aspectos, se inmoviliza un extremo de un nanoindicador, ya sea a través unión específica o inespecífica a una superficie sólida, se estira el nanoindicador, y luego se inmoviliza el otro extremo del indicador, también ya sea a través de unión específica o inespecífica a una superficie sólida (véase la Figura 3). Los métodos de captura, estiramiento e inmovilización de ácidos nucleicos son bien conocidos en la técnica. Los ejemplos de métodos que se pueden usar en los métodos descritos en el presente documento se describen en la patente de EE.UU. n.º 7.473.767 titulada "Métodos de detección y cuantificación de analitos en mezclas complejas"; la publicación de patente de EE.UU. n.º 2007/0166708 titulada "Métodos de detección y cuantificación de analitos en mezclas complejas"; la solicitud de EE.UU. n.º 11/645.270, titulada "Composiciones que comprenden macromoléculas inmovilizadas, orientadas, y métodos para su preparación"; la solicitud PCT n.º US06/049.274 titulada "Nanoindicadores y métodos de fabricación y uso de los mismos"; y la solicitud provisional de Estados Unidos n.º 60/088.988 titulada "Nanoindicador estable".

Por lo tanto, en algunas realizaciones, el nanoindicador está "congelado" en su estado estirado o extendido, para facilitar la resolución del código del nanoindicador mediante la detección y/o la generación de imágenes de las señales proporcionadas (por ejemplo, emitidas) por los monómeros marcadores unidos a un nanoindicador y sus ubicaciones de unas con respecto a otras. Estos aspectos de la invención se describen en la patente de EE.UU. n.º 7.473.767 titulada "Métodos de detección y cuantificación de analitos en mezclas complejas"; la publicación de patente de EE.UU. n.º 2007/0166708 titulada "Métodos de detección y cuantificación de analitos en mezclas complejas"; la solicitud de EE.UU. n.º 11/645.270, titulada "Composiciones que comprenden macromoléculas inmovilizadas, orientadas, y métodos para su preparación"; la solicitud PCT n.º US06/049.274 titulada "Nanoindicadores y métodos de fabricación y uso de los mismos"; y la solicitud provisional de Estados Unidos n.º 60/088.988 titulada "Nanoindicador estable".

En algunas realizaciones, el nanoindicador se puede inmovilizar al sustrato. En estos métodos, el sustrato para la inmovilización puede ser cualquier sustrato capaz de unirse selectivamente al nanoindicador evidente para los expertos en la materia. En algunas realizaciones, se puede inmovilizar una primera parte del nanoindicador sobre el sustrato por medio de un par de unión de avidina-biotina. En ciertas realizaciones, el nanoindicador puede comprender una fracción de biotina en su primera parte. Por ejemplo, un nanoindicador de polinucleótido puede comprender un resto de nucleótido biotinilado. En un sistema de sondas de nanoindicador dobles, una de las sondas de nanoindicador (por ejemplo, una sonda de nanoindicador de captura) puede comprender una fracción de biotina. Esta sonda podría estar marcada o no marcada. En realizaciones preferidas, cuando se usa un sistema de nanoindicadores dobles, la sonda de nanoindicador que comprende una fracción de biotina no está marcada. El sustrato que comprende avidina puede ser cualquier sustrato que comprende avidina conocido por los expertos en la materia. Los sustratos útiles que comprenden avidina se encuentran disponibles en el mercado, incluyendo TB0200 (Accelr8), SAD6, SAD20, SAD100, SAD500, SAD2000 (Xantec), SuperAvidin (Array-It), portaobjetos de estreptavidina (n.º de catálogo MPC 000, Xenopore) y STREPTAVIDINnslide (n.º de catálogo 439003, Greiner Bio-One).

En ciertas realizaciones, la primera parte del nanoindicador puede comprender una secuencia de nucleótidos que sea capaz de unirse selectivamente a una secuencia de nucleótidos sobre el sustrato (por ejemplo, biotina). En ciertos aspectos, la región constante (por ejemplo, múltiples copias de una secuencia de repetición de 15 bases) del nanoindicador es capaz de unirse selectivamente a una secuencia polinucleotídica complementaria que esté inmovilizada sobre el sustrato. En una realización específica, si el nanoindicador comprende cuatro copias de la secuencia de repetición de 15 bases ilustrativa de 5'-GGTCTGTGTGATGTT-3' (SEQ ID NO: 25), se puede utilizar un oligonucleótido que tiene la secuencia 5'-AACATCACACAGACC AACATCACACAGACC AACATCACACAGACC AACATCACACAGACC AGCCCTTG-3' (SEQ ID NO: 26) para inmovilizar el nanoindicador a un sustrato.

En algunas realizaciones en las que se usa un sistema de sondas de nanoindicador dobles, se usan una primera sonda de nanoindicador de captura no marcada que comprende biotina y una segunda sonda de nanoindicador marcada que comprende una región constante para la detección y/o cuantificación de una molécula diana (véanse las Figuras 1 y 2). Ambas sondas de nanoindicador se unen a la molécula diana a través de su secuencia específica de la diana formando un complejo que comprende la molécula diana y las sondas de nanoindicador. A continuación, se captura el complejo y se inmoviliza en un sustrato que contiene avidina tal como los descritos en el presente documento. El complejo se puede estirar y extender después como se describe en el presente documento. La región constante de la sonda de nanoindicador se puede unir luego a una secuencia polinucleotídica complementaria que esté inmovilizada sobre el sustrato. Los ejemplos de métodos que se pueden usar en los métodos descritos en el presente documento con respecto a los sistemas de nanoindicadores dobles se describen en la solicitud PCT n.º US06/049274 titulada "Nanoindicadores y métodos de fabricación y uso de los mismos"; y la solicitud provisional de EE.UU. n.º 60/088.988, titulada "Nanoindicador estable".

Los ejemplos de realizaciones específicas de la invención se muestran en las Figuras 2-4.

En ciertos aspectos, la divulgación proporciona métodos de preparación de una población adecuada de estructuras principales de ácido nucleico únicas, que comprenden (a) seleccionar un conjunto de regiones de unión al marcador a partir de una biblioteca de regiones de unión al marcador diseñadas, únicas; y (b) unir covalentemente el conjunto de regiones de unión al marcador entre sí en una combinación lineal que es diferente de las otras estructuras principales de la población, y repetir las etapas (a)-(b) hasta que se haya preparado una población adecuada de estructuras principales de ácido nucleico únicas. En algunas realizaciones, las regiones de unión al marcador comprenden de aproximadamente 800 a 1.300 bases de nucleótidos. En algunas realizaciones, las regiones de unión al marcador tienen un contenido de G/C del aproximadamente 50 %. En algunas realizaciones, las bases de adenina están separadas, como media, aproximadamente cada 8 a 16 bases en las regiones de unión al marcador. En algunas realizaciones, las regiones de unión al marcador están desprovistas de estructuras secundarias. En algunas realizaciones, cada región de unión al marcador seleccionada es diferente de las otras regiones de unión al marcador seleccionadas en el conjunto. En algunas realizaciones, la etapa de selección (a) comprende la selección de un conjunto de 3, 4, 5, 6, 7, 8 regiones de unión al marcador. En algunas realizaciones, la etapa de selección (a) comprende la selección de un conjunto de 6 regiones de unión al marcador. En algunas realizaciones, las regiones de unión al marcador se hibridan con un polinucleótido de ARN complementario que tiene unida al mismo una molécula detectable. En algunas realizaciones, el polinucleótido de ARN complementario tiene una proporción de G/C de al menos 1/1. En algunas realizaciones, el polinucleótido de ARN complementario tiene una proporción de G/C de aproximadamente 3/2.

En ciertos aspectos, la divulgación proporciona métodos de preparación de una población adecuada de estructuras principales de ácido nucleico únicas, que comprenden (a) seleccionar un conjunto de regiones de unión al marcador a partir de una biblioteca de regiones de unión al marcador diseñadas, únicas, en la que cada región de unión al marcador comprende de aproximadamente 800 a 1.300 bases de nucleótidos y tiene un contenido de G/C del aproximadamente 50 %, en la que cada región de unión al marcador comprende un patrón espaciado de forma regular de bases de adenina; y (b) unir covalentemente el conjunto de regiones de unión al marcador entre sí en una combinación lineal que es diferente de las otras estructuras principales de la población, y repetir las etapas (a)-(b) hasta que se haya preparado una población adecuada de estructuras principales de ácido nucleico únicas. En algunas realizaciones, las bases de adenina están separadas aproximadamente cada 8 a 16 bases en las regiones de unión al marcador. En algunas realizaciones, las regiones de unión al marcador están desprovistas de estructuras secundarias. En algunas realizaciones, cada región de unión al marcador seleccionada es diferente de las otras regiones de unión al marcador seleccionadas del conjunto. En algunas realizaciones, la etapa de selección (a) comprende la selección de un conjunto de 3, 4, 5, 6, 7, 8 regiones de unión al marcador. En algunas realizaciones, la etapa de selección (a) comprende la selección de un conjunto de 6 regiones de unión al marcador. En algunas realizaciones, las regiones de unión al marcador se hibridan con un polinucleótido de ARN complementario que tiene unida al mismo una molécula detectable. En algunas realizaciones, el polinucleótido de ARN complementario tiene una proporción de G/C de al menos 3/2.

Métodos de uso de nanoindicadores selectivamente inmovilizados, extendidos u orientados

En ciertas realizaciones, la divulgación proporciona nanoindicadores alargados, inmovilizados selectivamente que se pueden usar para crear códigos de barras macromoleculares para la separación y la detección secuencial de marcadores. Estos marcadores espaciados a lo largo de la molécula proporcionan un código único que se puede

leer cuando el nanoindicador es extendido e inmovilizado. La extensión e inmovilización selectivas pueden facilitar la descodificación del código de barras macromolecular.

Los nanoindicadores alargados, inmovilizados selectivamente se pueden usar en cualquier contexto en el que pueda ser útil la detección o la formación de imágenes de un nanoindicador. Se pueden usar con fines de diagnóstico, pronóstico, terapéuticos y de exploración. Por ejemplo, se pueden aplicar al análisis de muestras biomoleculares obtenidas o derivadas de un paciente para determinar si se encuentra presente un tipo de célula enferma en la muestra y/o para determinar la fase de la enfermedad. Se pueden usar para diagnosticar infecciones por patógenos, por ejemplo, infecciones por bacterias y virus intracelulares, determinando la presencia y/o cantidad de marcadores de bacteria o virus, respectivamente, en la muestra. Las composiciones y los métodos de la invención se pueden usar para cuantificar moléculas diana cuya abundancia es indicativa de un estado biológico o un estado patológico, por ejemplo, marcadores sanguíneos que están regulados positiva o negativamente como resultado de un estado patológico. Además, las composiciones y los métodos de la invención se pueden usar para proporcionar información de pronóstico que ayude a determinar el ciclo de tratamiento para un paciente.

Detección de nanoindicadores

Los nanoindicadores se detectan mediante cualquiera de los medios disponibles en la técnica que sea capaz de detectar las señales específicas sobre un nanoindicador dado. Cuando el nanoindicador está marcado fluorescentemente, se pueden investigar la consideración adecuada de fuentes de excitación apropiadas. Las posibles fuentes pueden incluir, pero sin limitación, lámparas, lámparas de xenón, láseres, diodos emisores de luz o alguna combinación de los mismos. La fuente de excitación apropiada se usa junto con un sistema de detección óptica apropiado, por ejemplo, un microscopio invertido de fluorescencia, un microscopio epi-fluorescente o un microscopio confocal. Preferentemente, se usa un microscopio que pueda permitir la detección con suficiente resolución espacial para determinar la secuencia de las manchas del nanoindicador. Por ejemplo, en una realización se puede obtener una imagen de un nanoindicador doble hibridado a una molécula diana. Si, por ejemplo, los nanoindicadores se marcan con tres colores diferentes, Alexa 488, Cy3 y Alexa 647 (marcados 1, 2 y 3, respectivamente). Cada uno de los colores 1, 2 y 3 se adquiere en diferentes canales y los registros primero y segundo, que se pueden ver como filas de manchas, se desplazan hacia arriba por varios píxeles para poder mostrar cada registro individualmente.

Los ejemplos de métodos de detección de nanoindicadores que se pueden usar en los métodos de la invención se describen en la patente de EE.UU. n.º 7.473.767 titulada "Métodos de detección y cuantificación de analitos en mezclas complejas"; la publicación de patente de EE.UU. n.º 2007/0166708 titulada "Métodos de detección y cuantificación de analitos en mezclas complejas"; la solicitud de EE.UU. n.º 11/645.270, titulada "Composiciones que comprenden macromoléculas inmovilizadas, orientadas, y métodos para su preparación"; la solicitud PCT n.º US06/049.274 titulada "Nanoindicadores y métodos de fabricación y uso de los mismos"; y la solicitud provisional de Estados Unidos n.º 60/088.988 titulada "Nanoindicador estable".

Selección del microscopio y de la lente del objetivo

La principal consideración referente a la lente del objetivo del microscopio es con la resolución óptica, que se determina mediante su apertura numérica (AN). En general, a mayor AN, mejor resolución óptica. La AN requerida es preferentemente de al menos 1,07 basándose en la relación de $\delta = 0,61\lambda/AN$ (δ = resolución óptica y λ = longitud de onda). La cantidad de luz que es recogida por el objetivo se determina mediante AN^4/Aum^2 (Aum = aumento del objetivo). Por lo tanto, para recoger tanta luz como sea posible, se deben seleccionar objetivos con AN elevada y pocos aumentos.

Selección de la cámara CCD y técnicas de captura de la imagen

Cuando se selecciona una cámara CCD, la primera consideración es el tamaño del píxel, que determina parcialmente la resolución final del sistema de formación de la imagen. Óptimamente, la resolución óptica no debe verse comprometida por la cámara CCD. Por ejemplo, si la resolución óptica es de 210-300 nm, lo que corresponde a 12,6-18 μm en una microplaca de CCD tras un aumento 60 veces, para resolver y mantener la resolución óptica debe ser de al menos dos píxeles para muestrear cada mancha. O el tamaño del píxel de la microplaca de CCD debe ser, a lo sumo, de 6,3-9 μm .

La segunda consideración es la sensibilidad de la detección que se puede determinar por medio de muchos factores que incluyen, pero sin limitación, el tamaño del píxel, la eficacia cuántica, el ruido de la lectura y el ruido base. Para lograr una alta sensibilidad, se selecciona una cámara cualitativa con un tamaño de píxel grande (que puede dar una zona de recepción grande), una alta eficacia cuántica y poco ruido. Una cámara ilustrativa con estos criterios es la cámara Orca-Ag de Hamamatsu Inc. El tamaño de la microplaca es de 1.344 x 1.024 píxeles; cuando se usa el objetivo de 60 aumentos, siendo el campo de visión de 144 x 110 μm^2 .

Aplicaciones de la tecnología de nanoindicadores

Las composiciones y los métodos de la invención se pueden usar con fines de diagnóstico, pronóstico, terapéuticos y de exploración. La presente invención proporciona la ventaja de que se pueden analizar muchas moléculas diana diferentes de una vez a partir de una única muestra biomolecular usando los métodos de la invención. Esto permite, por ejemplo, realizar varios ensayos de diagnóstico en una muestra.

Los métodos descritos en el presente documento diferencian entre secuencias de nucleótidos. La diferencia entre las secuencias de nucleótidos diana puede ser, por ejemplo, una única diferencia de base de ácido nucleico, una delección de ácido nucleico, una inserción de ácido nucleico o un reordenamiento. También se pueden detectar diferencias en las secuencias que implican más de una base. En algunas realizaciones, los conjuntos de sondas de oligonucleótido tienen esencialmente la misma longitud para hibridarse con secuencias de nucleótidos diana en condiciones de hibridación esencialmente similares. Como resultado de ello, el proceso de la presente invención es capaz de detectar enfermedades infecciosas, enfermedades genéticas y cáncer. También es útil en el control ambiental, en la medicina forense y en la ciencia de los alimentos. Los ejemplos de análisis genéticos que se pueden realizar sobre los ácidos nucleicos incluyen, por ejemplo, la detección de SNP, detección de STR, el análisis de la expresión del ARN, la metilación del promotor, la expresión de genes, la detección de virus, la subtipificación viral y la resistencia a fármacos.

En el ámbito del control del medio ambiente, la presente invención se puede usar para la detección, identificación y control de microorganismos patógenos e indígenas en los ecosistemas naturales y artificiales, y en microcosmos tales como en los sistemas municipales de depuración de aguas residuales y los depósitos de agua o en zonas contaminadas sometidas a biorremediación. También es posible detectar plásmidos que contienen genes que pueden metabolizar xenobióticos, controlar microorganismos diana específicos en estudios de dinámica de población, o bien detectar, identificar o controlar microorganismos modificados genéticamente en el medio ambiente y en plantas industriales.

La presente invención también se puede usar en una variedad de campos forenses, incluyendo la identificación humana para el personal militar y la investigación criminal, pruebas de paternidad y análisis de parentescos, tipificación de compatibilidad HLA, y análisis de sangre, esperma u órganos de trasplante para determinar contaminantes.

En la industria de alimentos y piensos, la presente invención tiene una amplia variedad de aplicaciones. Por ejemplo, se puede usar para la identificación y caracterización de organismos de producción tales como levadura para la producción de cerveza, vino, queso, yogur, pan, etc. Otro campo de uso es el relativo al control de la calidad y la certificación de productos y procesos (por ejemplo, animales de granja, pasteurización y procesamiento de carne) para contaminantes. Otros usos incluyen la caracterización de plantas, bulbos y semillas con fines reproductivos, la identificación de la presencia de patógenos específicos de plantas y la detección e identificación de infecciones veterinarias.

Métodos de diagnóstico/pronóstico

Los presentes métodos se pueden aplicar al análisis de muestras biomoleculares obtenidas o derivadas de un paciente para determinar si un tipo de célula enferma está presente en la muestra y/o la etapa de la enfermedad.

En algunas realizaciones, los métodos descritos en el presente documento se usan en el diagnóstico de una afección. Como se usa en el presente documento, el término "diagnosticar" o "diagnóstico" de una afección incluye la predicción o el diagnóstico de la afección, la determinación de la predisposición a la afección, el control del tratamiento de la afección, el diagnóstico de una respuesta terapéutica de la enfermedad y el pronóstico de la afección, la progresión de la afección y la respuesta al tratamiento en particular de la afección. Por ejemplo, se puede ensayar una muestra sanguínea de acuerdo con cualquiera de los métodos descritos en el presente documento para determinar la presencia y/o cantidad de marcadores de un tipo de célula cancerosa en la muestra, diagnosticando y determinando, de ese modo, la fase del cáncer.

Los cánceres que se pueden detectar mediante el proceso de la presente invención implican, en general, oncogenes, genes supresores de tumores o genes implicados en la amplificación, la replicación, la recombinación o la reparación del ADN. Los ejemplos de estos incluyen: gen BRCA1, gen p53, gen APC, la amplificación de Her2/Neu, Bcr/Ab1, gen K-ras, y los tipos de virus del papiloma humano 16 y 18. Diversos aspectos de la presente invención se pueden usar para identificar amplificaciones, delecciones grandes, así como mutaciones puntuales y pequeñas delecciones/inserciones de los genes anteriores en los siguientes cánceres humanos comunes: leucemia, cáncer de colon, cáncer de mama, cáncer de pulmón, cáncer de próstata, tumores cerebrales, tumores del sistema nervioso central, tumores de vejiga, melanomas, cáncer de hígado, osteosarcoma y otros cánceres óseos, testiculares y carcinomas de ovario, tumores de cabeza y cuello, y las neoplasias cervicales.

También se pueden detectar enfermedades genéticas mediante el proceso de la presente invención. Esto puede llevarse a cabo mediante la exploración prenatal o postnatal para determinar aberraciones cromosómicas y

genéticas o enfermedades genéticas. Los ejemplos de enfermedades genéticas detectables incluyen: deficiencia de 21 hidroxilasa, fibrosis quística, síndrome de X frágil, síndrome de Turner, distrofia muscular de Duchenne, síndrome de Down u otras trisomías, enfermedades del corazón, enfermedades de un solo gen, tipificación de HLA, fenilcetonuria, anemia de células falciformes, enfermedad de Tay-Sachs, talasemia, síndrome de Klinefelter, enfermedad de Huntington, enfermedades autoinmunes, lipidosis, defectos de obesidad, hemofilia, errores innatos del metabolismo y diabetes.

Como alternativa, los métodos descritos en el presente documento se pueden usar para diagnosticar infecciones por patógenos, por ejemplo, infecciones por bacterias y virus intracelulares, mediante la determinación de la presencia y/o la cantidad de marcadores de bacterias o virus, respectivamente, en la muestra.

Se puede detectar una amplia variedad de enfermedades infecciosas mediante el proceso de la presente invención. Por lo general, estas están causadas por agentes infecciosos de bacterias, virus, parásitos y hongos. Con el uso de la presente invención, también se puede determinar la resistencia de diversos agentes infecciosos a fármacos.

Los agentes infecciosos bacterianos que se pueden detectar mediante la presente invención incluyen *Escherichia coli*, *Salmonella*, *Shigella*, *Klebsiella*, *Pseudomonas*, *Listeria monocytogenes*, *Mycobacterium tuberculosis*, *Mycobacterium aviumintracellulare*, *Yersinia*, *Francisella*, *Pasteurella*, *Brucella*, *Clostridia*, *Bordetella pertussis*, *Bacteroides*, *Staphylococcus aureus*, *Streptococcus pneumoniae*, estreptomycin B-hemolítica, *Corynebacteria*, *Legionella*, *Mycoplasma*, *Ureaplasma*, *Chlamydia*, *Neisseria gonorrhoea*, *Neisseria meningitidis*, *Hemophilus influenzae*, *Enterococcus faecalis*, *Proteus vulgaris*, *Proteus mirabilis*, *Helicobacter pylori*, *Treponema palladium*, *Borrelia burgdorferi*, *Borrelia recurrentis*, *Rickettsial pathogens*, *Nocardia* y *Acitnomycetes*.

Los agentes infecciosos fúngicos que se pueden detectar mediante la presente invención incluyen *Cryptococcus neoformans*, *Blastomyces dermatitidis*, *Histoplasma capsulatum*, *Coccidioides immitis*, *Paracoccidioides brasiliensis*, *Candida albicans*, *Aspergillus fumigatus*, *Phycomycetes (Rhizopus)*, *Sporothrix schenckii*, *Chromomycosis* y *Maduromycosis*.

Los agentes infecciosos virales que se pueden detectar mediante la presente invención incluyen el virus de la inmunodeficiencia humana, virus linfocitotrófico de linfocitos T humanos, virus de la hepatitis (por ejemplo, virus de la hepatitis B y virus de la hepatitis C), virus de Epstein-Barr, citomegalovirus, virus del papiloma humano, ortomixovirus, paramixovirus, adenovirus, virus corona, virus rabd, virus de la polio, virus toga, virus bunya, arenavirus, virus de la rubéola y virus REO.

Los agentes parasitarios que se pueden detectar mediante la presente invención incluyen *Plasmodium falciparum*, *Plasmodium malariae*, *Plasmodium vivax*, *Plasmodium ovale*, *Onchocerca volvulus*, *Leishmania*, *Trypanosoma* sp., *Schistosoma* sp., *Entamoeba histolytica*, *Cryptosporidium*, *Giardia* sp., *Trichomonas* sp., *Balantidium coli*, *Wuchereria bancrofti*, *Toxoplasma* sp., *Enterobius vermicularis*, *Ascaris lumbricoides*, *Trichuris trichiura*, *Draunculus medinensis*, trematodos, *Diphyllobothrium latum*, *Taenia* sp., *Pneumocystis carinii* y *Necator americanus*.

La presente invención también es útil para la detección de resistencia a los fármacos por agentes infecciosos. Por ejemplo, con la presente invención, se pueden identificar *Enterococcus faecium* resistente a la vancomicina, *Staphylococcus aureus* resistente a la metilina, *Streptococcus pneumoniae* resistente a la penicilina, *Mycobacterium tuberculosis* resistente a múltiples fármacos y el virus de la inmunodeficiencia humana resistente a AZT.

Por lo tanto, las moléculas diana detectadas usando las composiciones y los métodos de la invención pueden ser bien marcadores de los pacientes (tales como un marcador de cáncer) o marcadores de infección con un agente foráneo, tales como marcadores bacterianos o virales.

Debido a la naturaleza cuantitativa de los nanoindicadores, las composiciones y los métodos de la invención se pueden usar para cuantificar moléculas diana cuya abundancia es indicativa de un estado biológico o de una enfermedad, por ejemplo, marcadores sanguíneos que se regulan positiva o negativamente como resultado de un estado patológico.

Además, las composiciones y los métodos de la invención se pueden usar para proporcionar información de pronóstico que ayude en la determinación de un curso de tratamiento para un paciente. Por ejemplo, la cantidad de un determinado marcador para un tumor se puede cuantificar con precisión de incluso una pequeña muestra de un paciente. Para ciertas enfermedades como el cáncer de mama, la sobreexpresión de ciertos genes, tales como Her2-neu, indica que será necesario un curso más agresivo del tratamiento.

Análisis de muestras patológicas

El ARN extraído de muestras tisulares embebidas en parafina fijadas con formaldehído o paraformaldehído normalmente es de poca calidad (fragmentado) y tiene un bajo rendimiento. Esto hace sumamente difícil y a menudo completamente inviable el análisis de la expresión génica de genes de baja expresión en muestras histológicas o

tejidos de patología de archivo. La tecnología de nanoindicadores puede satisfacer esta necesidad no cubierta permitiendo el análisis de cantidades muy pequeñas de ARN total de baja calidad.

5 Para usar la tecnología de nanoindicadores en dicha aplicación, se puede extraer el ARN total de muestras de tejido embebidas en parafina fijadas con formaldehído o paraformaldehído (o similares) usando kits disponibles en el mercado tales como el kit de aislamiento de ácido nucleico total RecoverAll (Ambion), siguiendo los protocolos del fabricante. El ARN de dichas muestras se degrada con frecuencia en fragmentos pequeños (200 a 500 nucleótidos de longitud), y muchas muestras histológicas embebidas en parafina solo proporcionan decenas de nanogramos de ARN total. Se pueden usar pequeñas cantidades (5 a 100 ng) de este ARN total fragmentado directamente como material diana en una hibridación con nanoindicador siguiendo las condiciones de ensayo descritas en el presente documento.

Métodos de selección

15 Los métodos de la presente invención se pueden usar, entre otros, para determinar el efecto de una perturbación, incluyendo compuestos químicos, mutaciones, cambios de temperatura, hormonas de crecimiento, factores de crecimiento, enfermedades o cambios en las condiciones de cultivo, sobre diversas moléculas diana, identificando de ese modo las moléculas diana cuya presencia, ausencia o niveles son indicativos de determinados estados biológicos. En una realización preferida, la presente invención se usa para dilucidar y descubrir los componentes y las rutas de los estados patológicos. Por ejemplo, la comparación de las cantidades de moléculas diana presentes en un tejido enfermo con el tejido "normal" permite dilucidar las moléculas diana importantes implicadas en la enfermedad, identificando de ese modo las dianas para el descubrimiento/la exploración de nuevos fármacos candidato que se pueden usar para tratar la enfermedad.

Kits que comprenden nanoindicadores

25 La invención proporciona además kits que comprenden uno o más componentes de la invención. Los kits pueden comprender, por ejemplo, un sustrato y uno o más nanoindicadores extendidos u orientados, o ambos, inmovilizados selectivamente sobre el sustrato. Los kits se pueden usar para cualquier fin evidente para los expertos en la materia, incluyendo los descritos anteriormente.

35 En ciertas realizaciones, la divulgación proporciona kits para la preparación de al menos un nanoindicador marcado de forma única que comprende al menos tres regiones de unión al marcador, comprendiendo cada una aproximadamente de 800 a 1.300 bases de nucleótidos, un contenido de G/C del aproximadamente 50 %; y al menos tres secuencias polinucleotídicas complementarias que tienen unida a las mismas una molécula detectable, en las que las secuencias polinucleotídicas complementarias tienen una proporción de G/C de al menos 1/1. En algunas realizaciones, las secuencias polinucleotídicas complementarias tienen una proporción de G/C de aproximadamente 3/2. El kit puede comprender además al menos tres sondas específicas de la diana.

40 En ciertas realizaciones, los kits son útiles para la extensión y la inmovilización selectiva de nanoindicadores. Los kits pueden comprender un sustrato para la inmovilización y una o más parejas de unión para facilitar la extensión o la inmovilización de un nanoindicador. Las parejas de unión podrían, en ciertas realizaciones, comprender una fracción útil para la extensión del nanoindicador en una fuerza apropiada. En ciertas realizaciones, las parejas de unión podrían facilitar la inmovilización o la inmovilización selectiva del nanoindicador a la superficie. En realizaciones adicionales, el kit podría comprender un nanoindicador para la extensión y la inmovilización. En realizaciones adicionales, el kit podría comprender un dispositivo capaz de extender el nanoindicador.

Los kits pueden contener una población de nanoindicadores como se describe en el presente documento.

50 Los kits pueden contener nanoindicadores previamente marcados o nanoindicadores no marcados con uno o más componentes para el marcaje de nanoindicadores. Por otra parte, los nanoindicadores proporcionados en un kit pueden o no pueden tener secuencias específicas de la diana previamente unidas. En una realización, las secuencias diana se proporcionan en el kit sin unir a la estructura principal del nanoindicador.

55 El kit también puede incluir otros reactivos, por ejemplo, tampones para la realización de las reacciones de hibridación, enlazadores, endonucleasas de restricción y ADN ligasas I.

El kit también incluirá instrucciones para el uso de los componentes del kit, y/o para la fabricación y/o el uso de los nanoindicadores marcados.

60 La presente invención se puede entender además mediante los ejemplos no limitantes proporcionados a continuación.

Ejemplos

Ejemplo 1: Diseño y protocolo de fabricación de una biblioteca de estructuras principales de nanoindicador De Novo 1 (DV1)

Para construir una biblioteca de diversas estructuras principales de nanoindicador, se seleccionaron regiones de unión al marcador de una biblioteca de secuencias polinucleotídicas de diseño racional, únicas, y se clonaron en diversas combinaciones en un vector de plásmido que tenía la secuencia polinucleotídica expuesta en la SEQ ID NO: 27. Esta secuencia del vector no finaliza en la estructura principal del indicador final, pero se utiliza para la clonación y la propagación de las secuencias de la estructura principal.

En concreto, se construyó cada plásmido DV1 específico del código, único, de regiones de unión al marcador de seis a aproximadamente 1.100 pares de bases, que se habían seleccionado de las secuencias polinucleotídicas de SEQ ID NO: 1-24. Las regiones de unión al marcador definidas por estas secuencias polinucleotídicas se clonaron en diversas combinaciones en el vector de SEC ID NO: 27, cada región de unión al marcador correspondiente a una posición fija sobre una estructura principal dada, es decir, la posición 1, posición 2, hasta la posición 6. Las regiones de unión al marcador de cada posición de una estructura principal dada eran diferentes de las otras regiones de unión al marcador de la misma estructura principal. Se clonó el extremo 3' de cada secuencia de la posición 6 junto a 4 copias de una repetición de 15 bases común conocida como la repetición G-4 (SEQ ID NO: 26), que se usó en la purificación e inmovilización de los indicadores. Se diseñaron 972 estructuras principales de nanoindicador, teniendo cada una una combinación lineal única de regiones de unión al marcador, se diseñaron y se clonaron de acuerdo con técnicas de biología molecular rutinarias.

Cuando a cada región de unión al marcador se asigna una de las cuatro moléculas detectables, el uso de regiones de unión a seis marcadores en un sistema indicador de cuatro colores ofrece 4.096 nanoindicadores únicos posibles. En este ejemplo, se asignó a SEQ ID NO: 1, 5, 9, 13, 17 y 21 un fluoróforo azul, se asignó a SEQ ID NO: 2, 6, 10, 14, 18 y 22 un fluoróforo verde, se asignó a SEQ ID NOS: 3, 7, 11, 15, 19 y 23 un fluoróforo amarillo y se asignó a SEQ ID NO: 4, 8, 12, 16, 20 y 24 un fluoróforo rojo. Por lo tanto, en este ejemplo, incluso si una estructura principal dada comprende 6 regiones de unión al marcador a las que se asignó un fluoróforo azul (por ejemplo, SEQ ID NO: 1, 5, 9, 13, 17 y 21 en las posiciones 1-6), cada región de unión al marcador individual tendría una secuencia polinucleotídica diferente de las otras regiones de unión al marcador de la estructura principal.

Se amplificaron plásmidos individuales en bacterias, aisladas, y se convirtieron en estructuras principales monocatenarias, lineales. En concreto, se construyeron estructuras principales de ADN monocatenario lineales a partir de un ADN de plásmido bicatenario usando un protocolo de cuatro etapas: (i) se linealizó el ADNbc con una enzima de restricción; (ii) se desfosforiló el ADN linealizado con una fosfatasa termolábil; (iii) se sometió a digestión el ADN con una segunda enzima de restricción para separar el vector de clonación de la secuencia de la estructura principal; e (iv) se sometió a digestión la mezcla con una digestión de exonucleasa lambda específica de la cadena, dejando solo una de las cadenas del fragmento de estructura principal intacta.

Se ligaron las estructuras principales monocatenarias a secuencias de nucleótidos específicas de la diana, y se incubaron con secuencias de nucleótidos del ARN complementarias coloreadas con colorante para fabricar nanoindicadores marcados.

Ejemplo 2: Generación de polinucleótidos de ARN complementarios coloreados con colorante

Se utilizaron reacciones de transcripción *in vitro* (IVT) para generar polinucleótidos de ARN complementarios modificados con aminoalilo usando kits de T3, T7 o SP6 MEGAscript (Ambion®), siguiendo las instrucciones del fabricante con las siguientes características específicas y modificaciones.

Se linealizaron los plásmidos que contenían un promotor de ARN polimerasa y la secuencia polinucleotídica de interés mediante digestión de restricción, se precipitaron en etanol y se usaron como moldes. En este ejemplo, se usaron las secuencias expuestas en SEQ ID NO: 1 a 24 como moldes en la reacción de IVT, generando 24 polinucleótidos de ARN complementarios únicos.

Se sustituyó el amino-alil-UTP (aaUTP) (Fermentas) con el UTP suministrado en el kit, o se usó una mezcla de aaUTP y el UTP suministrado. El nivel de incorporación de colorante a un polinucleótido de ARN complementario se correlaciona con el número de posibles sitios de unión a aminoalilo (aa). Para crear los segmentos de color más brillante posible, se debe usar aaUTP al 100 %. Para cambiar el brillo, se pueden usar una mezcla de aaUTP y UTP no modificado en cualquier proporción en las reacciones de IVT con el fin de modificar el número de sitios aa presentes en un polinucleótido de ARN complementario. Por ejemplo, para incorporar colorantes al 50 % de la base repetida de forma regular, se utilizó una mezcla 1:1 de aaUTP y UTP. La concentración final de aaUTP + UTP juntos fue la misma que la de cada uno de los otros tres nucleótidos (es decir, ATP, CTP, GTP).

Se dejó proceder la reacción de IVT a 37 °C durante 22 horas. Tras la reacción, se purificaron las transcripciones de ARN modificados con aminoalilo usando un kit RNeasy (Qiagen®) siguiendo las instrucciones del fabricante.

La reacción de acoplamiento de colorante que se describe a continuación es para una reacción de 1 mg. Se realizó la precipitación en etanol en 1 mg de un polinucleótido de ARN modificado con aminoalilo (aa). Se volvieron a suspender los colorantes sólidos (TFP-éster Alexa 488, succinimidiléster Alexa 546, succinimidiléster Alexa 594 y succinimidiléster Alexa 647; Invitrogen) a 11 µg/µl en DMSO anhidro. Se volvieron a suspender los polinucleótidos de ARN modificado con aa en un volumen final de 90 µl de Na₂B₄O₇ 100 mM, pH 8,5, y se calentaron hasta 37 °C durante 20 min. Se mezclaron individualmente 110 µl de los colorantes resuspendidos con 90 µl de un polinucleótido de ARN modificado con aa. Se incubó la mezcla a oscuras a temperatura ambiente durante 30 min-2 h.

Se purificaron los polinucleótidos de ARN modificados con aa coloreados con colorante de la mezcla usando un kit RNeasy (Qiagen) siguiendo las instrucciones del fabricante. Se mezclaron cantidades iguales de cada uno de los 24 polinucleótidos de ARN coloreados con colorante para una concentración final de 40 nM de cada polinucleótido ("mezcla de segmentos de color"). Esta mezcla se puede combinar con una población de estructuras principales de nanoindicador para generar una población de nanoindicadores únicos, como se describe en el siguiente Ejemplo 3.

50 Ejemplo 3: Generación de una biblioteca de nanoindicadores DV1 marcados

Para producir una biblioteca de moléculas de nanoindicador DV1 marcadas, se ligó cada estructura principal de ADN específico del código del Ejemplo 1 por separado a una región de sonda, es decir, secuencia de nucleótidos específica de la diana, para un gen específico de interés. Se ligó la sonda a la estructura principal a través de un oligonucleótido, que servía como un puente entre la estructura principal y la sonda específica. En concreto, se añadió una mezcla maestra que contenía un oligonucleótido universal que servía como una ligadura de "puente" más tampón ligasa a pocillos individuales de placas de 96 pocillos que contenían sondas de oligonucleótidos específicas de la diana normalizadas (10 µM) (35-50 bases). Tras una corta incubación a 37 °C para hibridar la sonda de oligonucleótido a la parte complementaria del puente de oligonucleótido, se inició la ligación mediante la adición de 1,2 pmoles de estructura principal de nanoindicador monocatenaria individual por pocillo, tampón de ligación adicional y T4 ligasa. Se incubaron las placas a 37 °C en un termociclador de 96 pocillos durante 2 h. Se desalaron las reacciones de ligación mediante centrifugación a través de columnas Sephadex G-50 en un formato de 96 pocillos. Se combinaron las estructuras principales individuales ligadas, se precipitaron en etanol y se volvieron a suspender en Tris 10 mM, pH 8,0 a 20 nM.

Para producir nanoindicadores marcados con colorante, se mezcló lo siguiente: 2 ml de 20 x SSPE, pH 6,5, 28,05 ml de H₂O, 5 ml de estructuras principales de nanoindicador DV1 combinadas anteriores, 1,2 ml de etanol y 3,75 ml de mezcla de segmentos coloreados del Ejemplo 2. Se incubó la mezcla a 75 °C durante 2 horas. Para eliminar los segmentos en exceso, se purificó la mezcla sobre una columna de oligonucleótido acoplada al complemento inverso de la secuencia de repetición de G común a todas las secuencias de estructura principal (véase la SEQ ID NO: 26 para la secuencia de complemento inverso). Para producir sondas de "captura" no marcadas, se biotilaron las estructuras principales de nanoindicador cerca o en la región 3'.

40 Ejemplo 4: Comparación de los sistemas de nanoindicador DV1 y M13

Se creó una biblioteca de indicadores DV1 de 40 genes como se describe en los Ejemplos 1 a 3, y también se creó una biblioteca de M13 comparable, como se describe en la solicitud de EE.UU. n.º 12/100.990. Las bibliotecas de DV1 y M13 contenían secuencias de sonda idénticas (es decir, nucleotídicas específicas de la diana), y se usaron para ensayar las mismas muestras - 36 lisados celulares, ensayados por triplicado. Se siguieron los protocolos de hibridación, purificación y formación de imágenes convencionales, como se describe en Geiss *et al.*, *Nature Biotechnology* 26:317-324, 2008. En general, se hibridaron las bibliotecas de nanoindicadores con lisados celulares que contenían ARN, se eliminaron los indicadores en exceso por lavado, se unieron los indicadores a una superficie sobre la que se inmovilizaron y se estiraron, se fotografió la superficie y se analizaron las imágenes.

50 Reacciones de hibridación

Se llevó a cabo la detección de las transcripciones celulares en reacciones de hibridación de multiplexación, que utilizaron un sistema de nanoindicadores dobles que tenía tanto sondas de nanoindicador marcadas como sondas de nanoindicador no marcadas, es decir, sondas de "captura". Se hibridó cada muestra por triplicado con las siguientes concentraciones finales de los reactivos de hibridación: 200 pM de cada sonda de captura biotilada, sin marcar, (sonda de captura), 40 pM de cada sonda de indicador marcada, 5 x SSPE (pH 7,5), 5 x reactivo de Denhardt (Sigma), 100 ng/µl de ADN de esperma de salmón cortado (Sigma) y Tween-20 al 0,1 %. Cada 30 µl de reacción de hibridación también contenían 100 ng de ARN total a una concentración final de 3,3 ng/µl. Los reactivos se mezclaron y se incubaron a 65 °C en un bloque termociclador con una tapa calentada durante 20 horas.

60 Purificación tras la hibridación

Para eliminar los indicadores no hibridados, se purificaron las reacciones sobre perlas magnéticas (Invitrogen™) acopladas a oligonucleótidos complementarios a la secuencia de repetición 3' contenida en cada sonda de captura. Primero se diluyeron las reacciones hasta 1 x SSPE en Tween-20 al 0,1 %/TE y se dejó que se unieran a las perlas a 22,5 °C durante 30 minutos con rotación continua. Se lavaron las perlas tres veces en 150 µl de 0,1 x

SSPE/Tween-20 al 0,1 %, y se eluyeron los complejos hibridados en 100 µl de 0,1 x SSPE/Tween-20 al 0,1 % durante 15 minutos a 45 °C. Tras la elución, se purificaron las muestras por segunda vez para eliminar el exceso de sondas de captura mediante la unión a perlas magnéticas acopladas a oligonucleótidos complementarios a la secuencia de repetición 5' contenida en cada sonda indicadora. Se llevaron las eluciones de perlas anti-repetición 3' a una concentración final de 1 x SSPE mediante la adición de 50 µl de 3 x SSPE/Tween-20 al 0,1 % y se unieron durante 15 minutos a 22,5 °C con rotación. Se lavaron las perlas como antes y se eluyeron en 30 µl de 0,1 x SSPE/Tween-20 al 0,1 % a 45 °C. A continuación, se prepararon las muestras purificadas doblemente para la captura como se describe a continuación.

10 Captura, estiramiento y generación de imágenes de indicadores NanoString

Se añadió un microlitro de dilución a 1/5.000 de una solución de sólidos al 0,1 % de una formulación a petición del cliente de microesferas fluorescentes Tetraspeck (Invitrogen™) a cada muestra. Se cargaron las muestras en un dispositivo fluido NanoString fabricado mediante la laminación de acrílico fundido mecanizado con láser con un cubreobjetos recubierto con estreptavidina (Optichem®, Accelr8 Technology Corporation) usando una capa adhesiva de doble cara cortada con láser (Fralock) para generar canales de microfluidos profundos 30i.µm. Se impulsaron las muestras a lo largo del canal por presión hidrostática y se unieron específicamente por el extremo biotinilado 3' de la sonda de captura. Tras la captura, se lavó la superficie una vez con 90 µl de 1 x TAE y se preparó para el estiramiento mediante la adición de 40 µl de TAE a cada pocillo. Se estiraron las sondas de indicador y se alinearon mediante la aplicación de 160 V/cm durante 1 minuto a lo largo del canal fluido. Entonces, se inmovilizaron los indicadores estirados en la superficie mediante la adición de 60 µl de una solución 500 nM de un oligonucleótido biotinilado complementario a las repeticiones 5' presentes en el extremo 5' de todas las sondas indicadoras. Se mantuvo la corriente durante 5 minutos, durante todo el proceso de inmovilización. Tras la inmovilización, se retiró la solución TAE y se reemplazó por una formulación a petición del cliente de reactivo anti-fotoblanqueamiento SlowFade (Invitrogen) para la formación de imágenes.

Se fotografiaron los portaobjetos con una Nikon Eclipse TE2000E dotada de un enfoque perfecto, un objetivo de inmersión en aceite 1.4 NA Plan Apo VC 60X (Nikon), una fuente de luz de haluro metálico X-Cite 120 (Exfo Corporation), una fase H117 automatizada (Prior Scientific) y un SmartShutter (Sutter Instrument). Para cada campo de visión, se adquirieron 4 imágenes a diferentes longitudes de onda de excitación (480, 545, 580 y 622) con una cámara CCD Orca Ag (Hamamatsu) bajo el control de bien el software Metamorph (Universal Imaging Corporation) o un software a petición del cliente.

35 Procesamiento de imágenes

El procesamiento de imágenes se realizó en 4 imágenes (uno para cada longitud de onda) a nivel de campo de visión por campo de visión. El algoritmo personalizado trata cada campo de visión como un bloque fundamental en el que se realizan las siguientes etapas básicas: 1) identificación de las manchas; 2) registro de las imágenes; 3) agrupación espacial para producir sucesiones; y 4) clasificación de las sucesiones.

En la primera etapa del algoritmo, se identificaron las manchas. Se calculó el nivel de intensidad de fondo de cada canal, y se usó para el umbral de la imagen en la señal y en el fondo, donde las regiones de señal son el resultado de una longitud de onda específica de la luz observada en función de la dispersión de los puntos (PSF). La máscara de señal se segmentó usando un algoritmo de Watershed a petición del cliente. Entonces, se marcaron las regiones segmentadas, se parametrizaron y se filtraron para eliminar las manchas no PSF. El resto de manchas se archivó de forma centralizada para su uso en el registro y la validación de indicadores.

El registro de imágenes se realizó en cada campo de visión, basándose en las manchas archivadas que se correspondían con perlas fluorescentes (fiduciales) que se habían unido a la superficie de formación de imágenes (véase la captura, el estiramiento y la generación de imágenes de indicadores NanoString). Se referenciaron de forma cruzada las manchas archivadas para identificar los grupos de manchas entre canales que cumplían los requisitos fiduciales (umbrales y proporciones de intensidad entre canales). Los grupos que cumplían los requisitos se archivaron como fiduciales. La lista final de los fiduciales representaba las transformadas espaciales que se produjeron entre los canales durante la adquisición de imágenes. Los desplazamientos espaciales fueron tan grandes como de 5-6 píxeles. La transformada espacial se resolvió usando los centroides fiduciales observados y sus centroides coincidentes con la transformada previa (asumida) ($X_2 = X_1 * T$). Entonces, se aplicó la transformada inversa a todas las manchas identificadas para restablecer sus centroides originales.

Tras la identificación de las manchas y el registro de las imágenes, se reunieron las manchas en "sucesiones" mediante agrupación. En este punto, se filtró cada sucesión para eliminar cualquier mancha atribuida a una señal de sangre. Entonces, se clasificaron las sucesiones filtradas como indicadores o no indicadores. Para clasificarse como un indicador, la sucesión debe contener el número correcto de manchas, cumplir umbrales específicos de distancia de mancha a mancha (1,2-2,9 píxeles), y cumplir con los requisitos de linealidad y orientación aceptables. Luego se contaron las agrupaciones que se clasificaron como indicadores y se sumaron para cada gen sobre todos los campos de visión.

Normalización y análisis de los datos de NanoString

Para tener en cuenta las diferencias leves en la eficacia de hibridación y purificación, los datos se normalizaron con respecto a los recuentos medios para todos los picos de control de cada muestra. Para determinar si un gen era "detectado" por el sistema NanoString, se compararon las mediciones por triplicado obtenidas para cada gen experimental con mediciones por triplicado para los controles negativos. Para que un gen se clasificara como detectado, los recuentos medios de el gen experimental tenían que ser superiores a los recuentos medios para los 2 controles negativos, y el valor p de la prueba t de Student tenía que ser inferior a 0,05.

Se muestran los resultados de tres genes representativos de los 40 genes sondeados, comparando el sistema DV1 con el sistema M13. La Figura 5 muestra los resultados para IL-8 (los resultados para GAPDH y TSC22D3 no se muestran). El perfil global de la expresión de cada gen es el mismo para ambos sistemas (es decir, mayor o menor expresión relativa entre las muestras), pero el número de recuentos para cada ensayo es > 100 veces mayor usando DV1 (cabe señalar la diferencia de escala para la gráfica de M13 y de DV1). También se muestran los resultados de una comparación de la expresión de los 40 genes entre los sistemas en dos muestras representativas (véase la Figura 6). En la Figura 6, el valor de R^2 muestra una correlación del aproximadamente 90 % entre la expresión relativa de todos los genes de los dos sistemas, mientras que la pendiente de la línea indica un aumento superior a 100 veces en el número real de recuentos usando los indicadores DV1 en comparación con los indicadores M13.

Ejemplo 5: Comparación de sistemas de nanoindicadores DV1 y M13

Se creó una biblioteca de indicadores DV1 de 148 sondas como se ha descrito en los Ejemplos 1 a 3, y también se creó una biblioteca de M13 comparable, como se describe en la solicitud de EE.UU. n.º 12/100.990, incorporada en el presente documento por referencia. Las bibliotecas de DV1 y M13 contenían secuencias de sondas idénticas (es decir, nucleotídicas específicas de la diana), y se usaron para ensayar las mismas muestras - 26 muestras de ARN de ratón. Se siguieron los protocolos de hibridación, de purificación y de generación de imágenes convencionales que se describen en el Ejemplo 4.

Los resultados para la detección de la expresión de GusB se muestran en las Figuras 7, respectivamente. Los resultados para la detección de la expresión de Mod1, Acot4 y Atg7 no se muestran. La medición de los niveles relativos de expresión de un gen dado en las diversas muestras fue comparable entre los dos tipos de indicador. Sin embargo, se observó un aumento de 6 veces como media para el número absoluto de recuentos medidos por los indicadores DV1 en comparación con los indicadores M13. La Figura 8 proporciona un gráfico que muestra los "indicadores válidos", lo que se refiere a las sucesiones de manchas que pueden ser interpretadas por el software de análisis de imágenes como un indicador "contable", descrito como un porcentaje del número total de uniones en un campo de visión. Un porcentaje significativamente superior de los indicadores DV1 son contables en este experimento, lo que conduce a un aumento del triple en los datos (del 12,5 % al 38 %).

Se determinó una correlación de la expresión de cada gen en una muestra de acuerdo con lo medido por M13 y DV1. Los resultados muestran una correlación del 77 % entre los dos sistemas y un aumento medio de 6 veces con DV1 (datos no mostrados).

Ejemplo 6: Optimización de la incorporación de colorante en polinucleótidos de ARN complementarios

Para determinar la concentración óptima de colorante fluorescente en una molécula nanoindicadora dada, se incorporaron colorantes fluorescentes a las secuencias polinucleotídicas de ARN complementarias en diversas cantidades. Entonces se ensayaron las moléculas nanoindicadores que tenían diversas cantidades de incorporación de colorante para determinar la intensidad de la señal en un ensayo de detección de nanoindicadores convencional.

En concreto, para determinar el grado óptimo de separación de los nucleótidos entre cada molécula de colorante incorporada, se incorporaron moléculas de colorante fluorescente a los polinucleótidos de ARN complementarios cada 8, 10, 12, 16 o 24 bases de nucleótidos. Esto se realizó mediante la generación de nuevas secuencias que contenían un patrón regular de T separadas cada 8, 10, 12, 14, 16 y 24 bases, como se muestra en las Figuras 9-12. Estas secuencias eran de aproximadamente 1.100 bases de longitud y se reunieron en un solo plásmido de estructura principal de nanoindicador de 6.600 bases, que se convirtió en una estructura principal monocatenaria. Se transcribió cada molde para producir el ARN complementario inverso, y se acoplaron alícuotas de cada ARN por separado a cuatro colores: azul (Alexa 488), verde (Cy3 (Amersham)), amarillo (Alexa 594) y rojo (Alexa 647). Se generó una serie de indicadores marcados usando los diversos colores mediante la hibridación de segmentos de ARN marcados con la estructura principal. Se fotografiaron estos indicadores como se describe en el presente documento, y se determinó el brillo de la mancha resultante del acoplamiento del colorante en cada espacio en cada color (datos no presentados). Las mediciones de la intensidad de las manchas están en unidades arbitrarias.

Estos experimentos muestran que se puede manipular el brillo de las manchas mediante el diseño de la secuencia subyacente para que contengan bases repetidas con mayor o menor regularidad. En el presente documento, una separación de 8 nucleótidos da las manchas más brillantes. Las separaciones de menos de 8 no se analizaron, debido a las predicciones de interferencia esteérica de las moléculas de colorante vecinas a menores distancias. Sin

embargo, es probable que algunas manchas pudieran hacerse aún más brillantes mediante menores distancias. La menor distancia posible variará de un colorante a otro, pero 8 bases proporcionan una separación que debe dar cabida a todos los colorantes comercialmente adecuados. No es necesario que las bases estén separadas de forma regular.

5

Ejemplo 7: Consideraciones relativas a las cinéticas de hibridación de los nanoindicadores

ANTECEDENTES

10 Las hibridaciones en solución con un gran exceso de sonda frente a la diana siguen una cinética de pseudo primer orden. En este régimen, la velocidad de la reacción solo depende de la concentración de la sonda y no de la concentración de la diana. Para que una estrategia de una sola diana y dos sondas proporcione información exacta sobre la concentración de una diana en solución, ambas sondas deben estar presentes en exceso de la diana. El posible intervalo de concentraciones está preferentemente, por lo tanto, limitado en el extremo inferior por la concentración de la diana. Sin embargo, el intervalo de concentraciones útiles para la tecnología de nanoindicadores descrito en el presente documento está prácticamente limitado en el extremo inferior por la cantidad de tiempo necesario para realizar la hibridación.

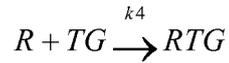
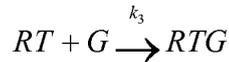
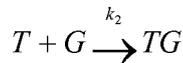
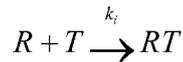
15

CINÉTICAS DE HIBRIDACIÓN

20

En realizaciones preferidas, se realizan análisis de detección y cuantificación de la diana en los que la diana (T) debe hibridarse tanto con una sonda de indicador (R) como con una sonda de captura (G) que se van a detectar (por ejemplo, por medio de selección de afinidad y detección de complejos que comprenden solo (R) y (G), que, a su vez, solo forman complejos en presencia de (T)). Suponiendo que estas reacciones son irreversibles, se producen cuatro reacciones elementales posibles.

25



Puesto que RT y TG son complejos intermedios de dos de las tres especies, estas cuatro reacciones se pueden simplificar en $R + T + G \rightarrow RTG$.

30

Sin embargo, para calcular cuantitativamente la velocidad de producción de RTG (el complejo de indicador-diana-sonda de captura), se deben considerar las cuatro reacciones. Las ecuaciones diferenciales que describen el sistema son:

$$\frac{dC_G}{dt} = -k_2 C_G C_T - k_3 C_G C_{RT}$$

$$\frac{dC_G}{dt} = -k_1 C_R C_T - k_4 C_R C_{TG}$$

$$\frac{dC_T}{dt} = -k_2 C_G C_T - k_3 C_R C_T$$

35

$$\frac{dC_{TG}}{dt} = k_2 C_G C_T - k_4 C_R C_{TG} \qquad \frac{dC_{RT}}{dt} = k_1 C_R C_T - k_3 C_G C_{RT}$$

$$\frac{dC_{RTG}}{dt} = k_4 C_R C_{TG} + k_3 C_G C_{RT}$$

5 en las que C_R , C_T , C_G , C_{RT} , C_{TG} y C_{RTG} son las concentraciones de las diversas especies, y 1(1-1 (4 son las constantes cinéticas para las cuatro reacciones elementales. Los valores para estas constantes cinéticas cuando las sondas y las dianas son moléculas monocatenarias complementarias (es decir, cuando no hay marcador de purificación sobre la sonda de captura y no hay indicador) se pueden calcular a partir de los datos disponibles en la literatura (Wetmur, *J. Annu. Rev. Biophys. Bioeng.* 1976.5:337-361).

$$k = k_N \frac{\sqrt{L} \alpha_{sal}}{N \alpha_{ref}}$$

10 En la ecuación anterior, k_N es la constante de la velocidad de nucleación, L es la longitud del ácido nucleico (en pares de bases), N es la complejidad del ácido nucleico (igual a L para secuencias no repetitivas), y α_{sal} y α_{ref} son las correcciones para la concentración de sal (Britten *et al.*, 1974, *Methods in Enzymology* 29E:363-406). En los sistemas de nanoindicadores descritos en el presente documento, las constantes cinéticas dependerán de los tamaños de los marcadores de la sonda de captura y la sonda de indicador unidas. Sin quedar ligados a ninguna teoría, la creencia de los inventores es que las constantes cinéticas tendrán la misma dependencia en la longitud que una reacción elemental tiene en las constantes de difusión de los reaccionantes.

$$k = k_N \frac{\sqrt{L} \alpha_{sal}}{N \alpha_{ref}} \frac{D_1 + D_2}{2D_{50}}$$

20 En la ecuación anterior, D_1 y D_2 son las constantes de difusión de las dos especies reaccionantes (véanse las reacciones anteriores) y D_{50} es la constante de difusión de una molécula de ADN monocatenaria de 50 unidades monoméricas. Suponiendo una diana monocatenaria de 100 bases, una sonda de captura monocatenaria de 100 bases y un indicador bicatenario de 7.200 bases, las constantes cinéticas relevantes son:

$$K_1 = 2,64 \times 10^5 \text{ l/mol/s}$$

$$K_2 = 6,55 \times 10^5 \text{ l/mol/s}$$

$$K_3 = 3,99 \times 10^5 \text{ l/mol/s}$$

$$K_4 = 1,91 \times 10^5 \text{ l/mol/s}$$

35 La resolución numérica del sistema de ecuaciones diferenciales con estas constantes cinéticas (suponiendo al menos un exceso de 10 veces de las sondas frente a la diana) produce la predicción de que el indicador 5 pM y la sonda de captura 5 pM conducirán a que la hibridación se complete hasta el 10 % en una reacción de una noche (16-18 horas). A concentraciones inferiores a 5 pM, es probable que sea poco práctico medir la cantidad de moléculas completamente hibridadas. De este modo, en una realización preferida, la concentración inferior de un componente nanoindicador (sonda de captura y/o sonda indicadora) es de 5 pM.

40 ENTRELAZAMIENTO DE LOS INDICADORES

45 A medida que aumentan las concentraciones de la sonda, la teoría pronostica que la cinética de hibridación se acelera sin límite - siendo el único límite la solubilidad de las sondas. Sin embargo, la sonda de indicador puede ser muy larga en comparación con la secuencia específica de la diana en los sistemas de nanoindicadores de la invención. Sin quedar ligados a ninguna teoría, los inventores creen que, mediante su unión a la sonda indicadora, las cinéticas de la secuencia específica de la diana se alteran con respecto a las cinéticas de hibridación en solución clásicas. Debido a que la sonda de indicador es una molécula polimérica, grande, puede tener interacciones de larga duración (entrelazamientos) con otros nanoindicadores cuando entran en contacto. A baja concentración, la probabilidad de que dos polímeros se entrelacen es baja, pero a medida que aumenta la concentración y/o el tamaño de un polímero en solución, estas interacciones se vuelven cada vez más comunes. En el caso extremo de moléculas muy grandes a muy alta concentración, los polímeros forman una red permanente, o gel, en solución. Para que se produzca hibridación en solución, un par de sonda (por ejemplo, sonda de nanoindicador)/diana se

debe difundir a través de la solución hasta que entren en contacto entre sí y se forme un núcleo de hibridación. Tradicionalmente, las reacciones de hibridación no están limitadas por la difusión debido a que la difusión translacional de las moléculas es más rápida que la nucleación de la hibridación (es decir, la sonda y la diana se difunden juntas e interaccionan muchas veces antes de que se produzca la nucleación). En una solución diluida, su gran tamaño ralentizará la difusión translacional de la sonda indicadora, pero puede no afectar significativamente a las cinéticas. A cierta concentración intermedia, las sondas de indicador ocupan casi todo el espacio en la solución, formando eficazmente un gel entrelazado permanentemente, y no se pueden difundir más en solución. Sin embargo, la sonda de captura y las dianas son moléculas más pequeñas que se cree que aún se difunden a través de las sondas de indicador entrelazadas, permitiendo que tenga lugar la hibridación (aunque posiblemente a una menor velocidad). Los inventores también creen que, a cierta concentración más alta, la sonda de indicador en solución también impedirá el movimiento de la sonda de captura y las dianas hasta el punto de que la reacción se vea limitada por la difusión. Esta concentración (que no se conoce cuantitativamente y depende de la estructura de la sonda indicadora, la estructura de la sonda de captura y el tamaño de la diana) es el límite superior del intervalo de concentración útil en el sistema de nanoindicadores, y puede ser determinada empíricamente por un experto en la materia guiado por los principios descritos en el presente documento.

DEPENDENCIA EN LA LONGITUD DE LA CINÉTICA

Dado que la concentración superior limitante para la hibridación depende tanto de la estructura del indicador como de la estructura de la sonda de captura (de las que hay muchas variaciones posibles), resulta útil en la práctica de la invención un marco teórico para pronosticar las permutaciones de los intervalos de concentraciones útiles. La teoría clásica predice que la cinética de hibridación solo depende del tamaño de la sonda más pequeña. La teoría podría pronosticar, por tanto, que el tamaño del indicador no desempeñará ningún papel en la cinética de hibridación siempre y cuando tanto la molécula diana como la sonda de captura sean significativamente más pequeñas. La teoría predice entonces que la velocidad de hibridación (para una longitud de la diana constante) depende de $1/L^{12}$, donde L es la longitud de la sonda de captura, debido a la inhibición estérica de la hibridación. Por consiguiente, la cinética de hibridación será más rápida con sondas de captura más pequeñas. A medida que aumenta la longitud de la sonda de captura, la velocidad de hibridación deberá reducirse como $1/L^{1/2}$. Si se supone una longitud de la sonda de captura constante, se pueden definir el intervalo de longitudes del indicador y las concentraciones que darán lugar a una cantidad medible de hibridaciones. Una vez que se ha definido el tamaño del indicador, se puede determinar el intervalo aproximado de tamaños de la sonda de captura. Se trata de un proceso iterativo, pero puede producir buenos puntos de partida a partir de los cuales se pueden reunir datos para generar directrices empíricas detalladas, dado que las teorías en las que están basados los fundamentos de los inventores fueron generadas a partir de datos de hibridación de sistemas que no emplean una sonda indicadora.

UMBRAL DE ENTRELAZAMIENTO

Una sonda de indicador es esencialmente un polímero en solución libre, que se comporta como una espiral aleatoria. El volumen ocupado por un solo indicador, V_p , se puede calcular a partir de las teorías físicas de polímeros de acuerdo con el modelo de la cadena articulada libremente (FJC, para un polímero flexible, tal como un ADN o ARN monocatenario) o el modelo la cadena vermiforme (WLC, para un polímero rígido tal como ADN bicatenario o un indicador). Para cualquier modelo:

$$V_p = \frac{4}{3} \pi R_g^3$$

en la que R_g es el radio de giro. Para la FJC:

$$R_g = b \left(\frac{N}{6} \right)^{0,6}$$

en la que b es la longitud del segmento y N es el número de segmentos de la cadena. Para la WLC:

$$R_g = \sqrt{\frac{1}{6} N b^2 - \frac{b^2}{4} + \frac{b^2}{4N} \left(1 + \frac{1}{2N} (e^{-2N} - 1) \right)}$$

La concentración umbral de entrelazamiento se define como la concentración a la que el volumen total de la solución está ocupado por los indicadores.

$$C^* = \frac{3}{4\pi R_g^3 N_A}$$

en la que N_A es el número de Avogadro. Por encima de esta concentración, se supone que la difusión translacional de los indicadores está fuertemente restringida. La concentración umbral de entrelazamiento varía con la estructura del indicador. A medida que aumenta la longitud del indicador, se reduce el umbral de entrelazamiento (como $1/L^{1.5}$). A partir de las ecuaciones anteriores, se puede calcular el umbral de entrelazamiento teórico para sondas de indicador con diferentes tamaños de mancha y diferentes longitudes. El resultado de dichos cálculos muestra que muestra que, para una sonda de indicador híbrida de ARN/ADN de 7.200 pb con 8 regiones de unión al marcador de aproximadamente 900 pb cada una, el umbral de entrelazamiento es de aproximadamente 70 nM.

Si tanto la diana como la sonda de captura son mucho más pequeñas que los indicadores, será más probable que estén libres para difundirse a través de la solución incluso a estas altas concentraciones de indicadores. Los datos iniciales indican que la cinética de hibridación no muestra ralentización apreciable hasta una concentración de 80 nM con una sonda de indicador de 7.200 pb, una diana de 100 bases y una sonda de captura de 100 bases.

EFFECTO DEL UMBRAL DE ENTRELAZAMIENTO SOBRE LA MULTIPLEXACIÓN

Suponiendo que la concentración máxima para los indicadores en una reacción de hibridación es C^* , la concentración de cada indicador (específica de una determinada diana) es igual a C^*/M , en la que M es la multiplexación de la reacción (número de diferentes dianas hacia las que se realiza la dirección simultáneamente). Por el contrario, el nivel de multiplexación posible para una estructura de indicador concreta se puede calcular a partir del límite inferior de concentración de la sonda (C_p de la cinética ~ 10 nM) y del umbral de entrelazamiento:

$$M = \frac{C^*}{C_p}$$

Si el número de códigos de nanoindicador disponible no depende del tamaño de la sonda indicadora, entonces, la multiplexación del nanoindicador depende principalmente del tamaño y de la concentración de la sonda de indicador (dado que es mucho mayor que la sonda de captura). Debido a que la sonda de captura contribuye de manera irrelevante al entrelazamiento durante la hibridación, los inventores creen que la concentración de la sonda de captura se puede aumentar bastante por encima de la concentración de la sonda indicadora. En siguiente Tabla 4, la concentración de sonda de captura total máxima ([G]) se ajusta a 1.000 nM para todas las concentraciones de indicador. Esta diferencia en la concentración de la sonda de captura y la sonda de indicador es un parámetro ajustable. Experimentos preliminares muestran que en una reacción de hibridación de multiplexación con un indicador de 7.200 pb y una sonda de captura de 100 b, 40pM de cada sonda de indicador y 200 pM de cada sonda de captura dan como resultado una hibridación casi completa en una reacción de una noche de duración.

INTERVALOS DE TAMAÑO Y DE CONCENTRACIÓN ÓPTIMOS

En la siguiente Tabla 4, hay un resumen de los intervalos de tamaño y de concentración útiles óptimos de la sonda de captura y la sonda de indicador a diferente multiplexación según las estimaciones de las teorías anteriores. Los inventores creen que las sondas de captura de hasta aproximadamente 200 bases serán prácticas para la mayoría de las aplicaciones.

Tabla 4: Intervalos de tamaño y de concentración óptimos de la sonda indicadora, la sonda de captura y la diana, así como multiplexación de las sondas en los sistemas de nanoindicadores de la invención

Longitud del indicador (pb)	Longitud de la sonda de captura (b)	[R] mínima (pM)	[G] mínima (pM)	[R] máxima (nM)	[G] máxima (nM)	Multiplexación máxima
2.000	100	5	5	603	1.000	114.417
2.000	50	4	4	603	1.000	161.811
2.000	200	7	7	603	1.000	80.905
3.000	100	6	6	292	1.000	45.182
3.000	50	5	5	292	1.000	63.897
3.000	200	9	9	292	1.000	31.948
4.000	100	7	7	178	1.000	23.912

Longitud del indicador (pb)	Longitud de la sonda de captura (b)	[R] mínima (pM)	[G] mínima (pM)	[R] máxima (nM)	[G] máxima (nM)	Multiplexación máxima
4.000	50	5	5	178	1.000	33.817
4.000	200	11	11	178	1.000	16.908
5.000	100	8	8	123	1.000	14.746
5.000	50	6	6	123	1.000	20.854
5.000	200	12	12	123	1.000	10.427
6.000	100	9	9	91	1.000	9.988
6.000	50	6	6	91	1.000	14.125
6.000	200	13	13	91	1.000	7.062
7.200	100	10	10	68	1.000	6.792
7.200	50	7	10	68	1.000	6.792
7.200	200	14	10	68	1.000	6.792
8.000	100	11	11	57	1.000	5.444
8.000	50	7	7	57	1.000	7.699
8.000	200	15	15	57	1.000	3.850
10.000	100	12	12	40	1.000	3.419
10.000	50	8	8	40	1.000	4.835
10.000	200	17	17	40	1.000	2.417

LISTADO DE SECUENCIAS

- 5 <110> NANOSTRING TECHNOLOGIES, INC.
- <120> NANOINDICADORES ESTABLES
- <130> 31353-701.601
- 10 <140>
- <141>
- <150> 61/088.988
- <151> 14-08-2008
- 15 <160> 33
- <170> PatentIn versión 3.5
- 20 <210> 1
- <211> 1140
- <212> ADN
- <213> Secuencia Artificial
- 25 <220>
- <221> Fuente
- <223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"
- 30 <400> 1

ES 2 614 810 T3

gaattcaaag ggctgggtga agcggtgaaa cggctcgagcg ctgaaggaat gagaaagtga 60
 ggcggtagac aaaagtaagc cagtggcaca gtgaggaaga tgagcgagct gaggacaatg 120
 acggagtccg aggaatcaga gcggtgagac aagtggagga tatcaaagat aagagcatag 180
 ggaaatgcaa caatggaac gtccaaggt ggaagcgtgg gagaatgaag aggtaagcaa 240
 atagaagacg tagggaacat gaaaccatgc agaagataag aaaatgccag aatacgacgg 300
 tgagagaaat caaccagtac aagcgtgaa cagctaccga ggtagcgaga tgaacaagat 360
 gcgaacctca ggaactcaag aagtagcgaa atcgaccggg tgggaaagt cgagaaataa 420
 gaacgtacca gggatacaga actagggacg taggaggggtg ggacgatacg gcgctgaaac 480
 ggggtggagg gtaacagggg gaaaagtaa gagactaagg aactgaaaca gctaacaggc 540
 taagggaaca tggagaaata aagacactgg agcgcagctg gaagatagag aaaatgagag 600
 cgtgaaacca tgaagggat caagaggtga cggagcatag aaagctgaaa caaatagggg 660
 agctgaagac cataagcggg ctgcaaaga taagagagtg acaagatacg cgccgtggag 720
 aagtgcagga cataaaaca tggccgcatc aggccgggtg agggcaatac aagagctaga 780
 agagtaccgc gataggaagg tggcaccagt aaggaaataa gcccatgagg acatacagca 840
 gtcgaaaaat aagcgagtca aacgctagc caactggcgg catgggacgg tgcgcgggtc 900
 gacagaggtg tacaagtgc aggatgaaag cataagaagg tgacgcaact agggccatac 960
 aaagagtgga ccaatccaaa cctgcgaaga taggaggata acaccggtag ggcaactaca 1020
 aggatcaaag gatgaaagaa taaacacta agggcgtcca acagtaccga agtcaggcgg 1080
 tcacaggctg aacagaactc aaccgaagtc tagaaggatc cactcgagtt tagatctttt 1140

5 <210> 2
 <211> 1141
 <212> ADN
 <213> Secuencia Artificial

10 <220>
 <221> Fuente
 <223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

15 <400> 2

ES 2 614 810 T3

gaattcaaag ggctgaacgc aggtacgaga gtaccgggaa tagcggagtg aagccctaaa 60
 gaaataacag gaatagcaca agtaggaaca taaacagatg cacaccaata aagagatacg 120
 gaaataaaga catagagaca tccgcaataa aagggatgaa acaataaccg ggtcgcgaag 180
 ctaacacca tcgccagctc gggcaactg atacctcga aggactagca agatggaaac 240
 atggacaagt aaaaggggta agaagtacca caaactcaac agctaagaca agtcgggaag 300
 taaaaaatc caagggtaac agaatggcga agtgagcgaa agtagagcaa tgacggcatg 360
 gagaatggg aaaaaatcac gagatagacg agtgggcagg taggcgggtg accggaatac 420
 gacaatgaga gactgccgca gtgcgaaagt ggcgggataa aaaagctgaa gggagtcaaa 480
 ccatcagcgg gatacagaag tagaacaatg cacagatgcg caggtaagcg ggtgcaagac 540
 tagaaaagct cggcaaatcg ccgaatagac aatcagctg gtgagcgggt acgaagactg 600
 gcggccataa cgccatgagg gcatagcaaa agtaccctaa taaggcagtc agagagtgc 660
 gggctacagc ggctaggccg actacagcag tgcgggaagt gaaaaagct cgaaggatgc 720
 aacctgagc caataaaggc gtcgagagct gccaatgac gaaagtgagg acatagacga 780
 atagaaaaat gagacgatag cgaggtgaca ccataccagg gtccggcagg tagccagata 840
 gggaaataga acgatcaaag gatagcgaaa cgtcagaaac tcgggagcta aaaggataac 900
 ggaatcagat gtacaacgca gatcagaacc tgggagcaag tggagccaat caacaagata 960
 agaaaaatga aggcattgaa acggtccggg cagtggaaac atccaagagt aggaaaatca 1020
 gaagcgtacg aaactacaag cgtcagggag tagaaaagta acacaatgaa gaaactaagc 1080
 acgtgaaaga gtagaggaac taggaaacat ctagaaggat ccactcgagt ttagatcttt 1140

t 1141

- 5 <210> 3
- <211> 1142
- <212> ADN
- <213> Secuencia Artificial
- 10 <220>
- <221> Fuente
- <223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"
- 15 <400> 3

ES 2 614 810 T3

gaattcaaag ggctgacaac cactgaaggc ataaacaagt accaaagtaa cagcatggag 60
 agctgagggc gtaacaaagg tcgaaaactc aaactactaca aaaatgggac aaatagcgcc 120
 catcggaagc ctgagcggat agcggactac ggcagtaaag agatagcggg gctgaaagcc 180
 tcaggacgtg aacaaatggg aaggataacc gggtaggagg gtgaaacaga tgagacggta 240
 agaaaagtaa gagaaggtgc gatatcgcg ggcacacgca acgtgcaaaa atgacggaat 300
 aaaagaatcg aggaggtcaa ggcgataagc gcgtgagagg atagaaagat cgagccatag 360
 cgggccatca gcggcgtgcg aggagtcgca cactcaaga gctaaccgga tcagcgagta 420
 cagcgggtag aagcgtcgcg ggatagagga agtcaaaga tcccgaactg ccagcgtagg 480
 aactgacc acatagcacc atcaaaagct gaacgagatg agacactacg caggatgaga 540
 acgtcgcaag catgaaccgg gtgcagagct aggacagctg cgccgatcag gccacgtagg 600
 aagatccaag cctggcacag agtcaagacg ctgaaaaat gaagaagtca gcaaagtagg 660
 gaggggtggg gcatgaacga gtgaggacat aggacgatcc caaagtgacg gaatgaccag 720
 gtgagaaagt agggcaaatc aagcagtcaa agcgtgaaga aaactagaag gcgtaaaaga 780
 gtggagaagc tccgacagat acaaaagtag aggcctgaga gggtcgggca aaatcccaga 840
 ctcggagaat cgcgacaatg caaacgtggc gcggtgggcg aggtgccgaa atcacgcgaa 900
 tggacggata tgtactga aaaagctcac aaaataagcg gatcgggaca tcggaacctg 960
 agacgaatac agacgataaa gcaataaccg actagacgag tgccagggta ggaacgtacc 1020
 aaaagtccag aagtccacgg gtggacagac tagagaacag tacagaaaat gccaccctaa 1080
 acgggtaagg gaatgcggga gtggacaaac tctagaagga tccactcgag tttagatctt 1140
 tt 1142

5 <210> 4
 <211> 1140
 <212> ADN
 <213> Secuencia Artificial

10 <220>
 <221> Fuente
 <223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

15 <400> 4

ES 2 614 810 T3

gaattcaaag ggctgaaaga atagacgggt aaacaagata gaggcattggg aggcctgcga 60
aagtggacaa atgacgggat gccagggatg caccactggc aggatggcaa cgtaagaggc 120
taccacgctc cgggaggtaa ggcgatcgaa agagttagaga agcatacccg cataggacaa 180
taaaagggtg agggaagtgg gagcatccga aaatagaacc ctacaggcaa gtagacaggt 240
aaacgcatga gaagacctgg gcgcggtcg gaacactgcg ccagtaaggc ccatgaaaaa 300
cggtgagaga tatcacagaa ctggaacaac tccggacata cagaaatata cgaataggc 360
aatagacgag taacaacgat aacgcgagta gcgagagata cgaagagtag caagactcgg 420
aggagtcaa gaaaataagc agatggcaaa gatagggag aataaacgac tggagccagt 480
gaggagctaa aagggtcagc agactggcgg ggtaaggaga taaaaactc agcgaatcaa 540
aaaatacagg caagtaccag ccatccacgg gtaacagctg gccaccgtcc ggcaaataaa 600
agaatccagc agctagagcc ggtaacccaa tcccgagagt cgaagagatg gcagaatgag 660
aacaatacaa gggataagca gggtagcggg atgaaacggt agaggacata cgaggataac 720
aacactaaag gactagagag aataaacgac gtgacacgac taaaaaagg tacgggcgat 780
caaacaagtg gaagactcca aaagataacg cggtgagcga atacggaggc gaaaaaaat 840
agacgggatg gcgggatgcc agagttagaag ccctaagaga atagaaagca taaaaccta 900
ggcaaaactg tacaatagac gggtagaaaag aatgaggcga tcaaagcata gcaaaatcag 960
gaactaagca gggtagcggg catggaaca tagccaaata gggaggacta gcagggtgga 1020
aggatgacga actagaagcg atcaacggat aaggagggtc acagatagca cgaataacgc 1080
gggtaagaaa gtacaagaat gagcaacgctc tagaaggatc cactcgagtt tagatctttt 1140

5 <210> 5
<211> 1106
<212> ADN
<213> Secuencia Artificial

10 <220>
<221> Fuente
<223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

15 <400> 5

ES 2 614 810 T3

tctagagaag gatacaggaa tcaagacatg cagggaaacta aagacatacc ggaataagag 60
 aggctaaagg gataaggaga gtacgaggag taaagacaga tcgacgggaa taacaagagt 120
 caagacgtga ccgacctgag cgcatcgaga ccgcgggaga ctcggcgggt gaacaaatgg 180
 aaggataaga acgggtagaa gaactacgag aatgaaggga tgtacaaggt agggcagata 240
 agaggggtgc gagatgaaaa ggctcggaaa actcaaccaa ctggagaact aggccgggtcg 300
 aaggaataga aacgtaagcg agataagagg atggcgcaga ctgggacaat cgggaggtcg 360
 agagatcacg ggaatagaca cgtccaagaa atgaacaagg taaggacatg cgggaataag 420
 aaactcaaac cctaggccag tcacagaagt aaccagtac gcaagtcacc gagtcggaac 480
 gatgacaaag taaaaaagg taaagaatga gagcactaaa aaagtaaaag agatcaggcg 540
 cgtgggcgag atgcagaggt aacgaaataa aaagatcgca gaatccaaag actacggagg 600
 atcaggcaaa taagaagata aaaaagatcc gaagataacg aggtaggcca atgagagaat 660
 accgagcgtg gagccatagc agagtaaaga gagtagggca gtaaaaagat ccagcggtag 720
 gagcatcgcc gcctcaagaa gatccagaga tggaaacgta ggagaataaa gcggtgaaag 780
 cgtacaaacg tagggaagtc gggagagtag agagatacag agagtaagag ccatagacac 840
 ctgagacgat cggcaactgg gcaacatcca gagatgggca cctaggcaca tcaccgggtg 900
 caaaggatca cagagtagaa cgctcaaaga agtcaccaag tgcacgggta agggacatgc 960
 gaaggtgaga gcgggtggaa aactcgacag ctcaacacat gagggaatgc cagagatgga 1020
 gcaaataaga gacatgccgg gcgatgagga actacaacag tagcccaatg caaggctaag 1080
 gaagtacaaa agtcgggaaa ggatcc 1106

5 <210> 6
 <211> 1107
 <212> ADN
 <213> Secuencia Artificial

10 <220>
 <221> Fuente
 <223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

15 <400> 6

ES 2 614 810 T3

tctagagaat gaaacagtgg aaaggtcggg aactgaggcg gagtagaac aagtcgag 60
 atacgagaga tgaaagagcg tcaaagaatg gaggactcag aagatagaaa catcgggaca 120
 ggtgagaaca ctgggaaaga atcgaaaaaa ctacaaaac tacaaaaact gcaaaaatga 180
 aaggagtggg cactggaca gccgcggaag ctcaagaact gtgtacaggt gcagcggta 240
 gaaacatcaa aaggtaacag catcgcggga tcggaaaatc cgggagtaa aaagtgcaga 300
 aatgacacgg gtacgaaacc tcagaaacat aaacagctac gggcatgaga gcataaagag 360
 ataaaaaat gcagagcgat accacacgta aggagatcgc aaggtgccac gatcagaaaa 420
 tagagacatc gacgggtggc gaactgagca acatccggga ataagaacct caagggcgtg 480
 caaccggtac acgaatgaga ggatgcggga gtcagcggag tgggaagggt aggaaagtag 540
 aaaaaataca gccaggtaca aaagtaaagc gggtaaccg gtgccaagg tgcgaaagta 600
 ggcccgtgag aaaaaccggc aatagaacag taaaagggtg agacgcataa cacgggtgca 660
 aagctaagag agatccacca gtggaagat gggaaagcat aacaagata gaagaataag 720
 gcagtaggag aatgaaaaca atacagaaaa gtagggcaga taaagcacta acgagctgac 780
 aggctggcag gatgaagaca tgacacggat caaggcatgg acgagtgcac aggtcaagcg 840
 ctccggcgat gagcgaatag cccgcgtaac acagtccgac aaggtgggag catagcccga 900
 tcacgagggt gcaggcgtga gaaaagatcc agggctaaa cactagaac atggcacagt 960
 acagggactg aggagctgac aagatgcaca cggtaaaagg ataagaaacg tcacgaggat 1020
 accaccatca cgaataaagc gggtcggaaa catgccagac tgggcaagtc accaggatga 1080
 agaaatgaag aagtagcaag aggatcc 1107

5 <210> 7
 <211> 1106
 <212> ADN
 <213> Secuencia Artificial

10 <220>
 <221> Fuente
 <223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

15 <400> 7

ES 2 614 810 T3

tctagagaaa aacataagca gcatggagcg ctacgggagg atggcacggt cgcgggaatc 60
 gaacgctgaa gcaatcaaag ggtagggagg tctgggcaata cagcgatcgg aggatgagac 120
 gggtcacaag aatgggacc atccgcaaat caagagataa ggaatacaa ggaataagaa 180
 ggatgaaaac gtgcggagat gaagacgatg acgacaataa tgtacaatga gcagatacac 240
 acggtgaacc ccgcggaacg ctaacaaaat accgaaataa aaaaatagcg agatacaggg 300
 ctaaagaagt gcaagggagt aaggccctgg cagggtaca agctcgcgga cgtcaccgcc 360
 gtgcgggcaa tacgaccgct aaagaagcta gagagaatca gcagggtaaa gaagtaggac 420
 cgactcgcaa gatcgagaaa taaaaaagtg gaccgggtga gcacggatga acaagtgaag 480
 aaatacagag gtacaagaat cagaaacctc gggacataga aacatccaag aagatgggca 540
 gaatcgaaca ggtagacgag tgagaagact aaaacaatga agaactaggc ccgcatacga 600
 gactgaaagc atcaagagga taaggaactc agaaaactag agaaactaag gcgatcagga 660
 ggatgagaaa atcgacgagt aagggaagtc cgggagtaca cggatgaacc gatacaagga 720
 tggcgacgtc cagcaatgga aaggctcagc cgataacacg gtaggagcaa tagaagaaat 780
 aaagcgtgc gcggaataag aaaggataac gagatcaaga caatggaacg agtacaaaag 840
 tcaaacgctc gaaagggtaa agcactgacg aacatagagc gataaacacc tgggaagatc 900
 cggactagc aaaatcacga caaatgcaaa gatcagccga tctaggcata ccagggtgaa 960
 aaaatcaaaa aagtaaccgg acctaaacca acgtgagaaa atgctgagcgt acagacctga 1020
 cgaaatcaac aatgaagaa gtaacagaga tccaaaactg aaaaggtaaa agcatggaca 1080
 cgtaagcaag atagacaagg ggatcc 1106

5 <210> 8
 <211> 1107
 <212> ADN
 <213> Secuencia Artificial

10 <220>
 <221> Fuente
 <223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

15 <400> 8

ES 2 614 810 T3

tctagagagg tagcacacgg tgaaaagcta agaacctcaa acgatcgac catgacgcga 60
atgagacaaa tgccgaaata cacacatacg aaacctcaag accctaacgg gctaaaagag 120
tcggagccat agaaaggtaa cagcgatacc gaaatggaac ggggcagca ataaagaagc 180
taaacgaagt aggaaaatag ggaaatggaa ccagtgggac atgtacagac gaatcggaaa 240
catgacgacc tggacgggta cggcaagtaa aaaaataaaa gagtagagac atgaaaacat 300
accgcggtga gagcgtagaa gcgtgaaaca gtaagcggct aagggagtcg gaggataaac 360
ggctgacgga gactgacaac cataaagcag gtgcaaaggg atggcaacgt caaggcctcg 420
cagaagtaag ggaatggcaa aactaagcaa agtagagcac taacaaaggt aggacgaagt 480
acgaaggta ggaggtggaa agctggcgca catgagagaa tacaagaact agcgaaggta 540
cagcactgga cggatcaaaa cggtcgcaac gtaagacggt agggcggtag aaaagctcaa 600
cgactgagag gcatgaggac gctaaaggga gtagcaagct gagaaaatgc cggaagtga 660
aaactggaaa gataagcaag tccccaagt cgagaacgta gaagagtacg ggagtacaga 720
acctggcgga agatcaggcg agtagcagag tcaaaaactg gcaagatcag gagtaacca 780
aggctcaggc aatgcaagac tgaacgggtc acaagcgtca ggacatagaa gaatgaaaga 840
gtaccagaat acacgggata gcagacgtaa ggacgatagg agcgtgagcc aatccgagga 900
tagaggaata aagcacataa gcgggtgaga ccatgagagg atcgaagaat acggaataa 960
ggcgggtaag aacgtgacgc aagtgaggcg ataagaacaa tagcgcgaat gacggagtac 1020
ccaatcaag ggatcgccgg gtgaaaaagt aagagaatcg ccagcgtcca aggaagtaga 1080
ggaatacaac aagtcacgga aggatcc 1107

5 <210> 9
<211> 1108
<212> ADN
<213> Secuencia Artificial

10 <220>
<221> Fuente
<223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

15 <400> 9

ES 2 614 810 T3

ggatccgatc gggagataaa gggagtaaac gcatcaaacg atggcgagat gaacagggatg 60
 acagggtccc gaacataccg ggatggagcg ctgggaacct ggaggaacta gaaacctccg 120
 aagctccacg ggtgaacagg tggacaggta aaagaatata acaagtataa ccgtaaaaa 180
 ccctcaggcg gtagggaact cgaagaaatg agcgggtaag gaaggtaaag ggatgaacgg 240
 gtggaaagat gggaaagtaa agagaatggg acaccataac cgcatacgcg gaatggcaag 300
 agtgacacaac taggagaagt cacaaagtgg aaaactcgca ccgtaaggac ctggacggga 360
 atagcgggga tcgaacggat atcgagtacg gaaagtccaa gagctgac actagaacac 420
 atgcacgact acgccagggt gaaagcgtaa acgccgtgca aagctgccc ggaatcagac 480
 agtggcgga taacgaaagt aaccagggt gacgcgctgc cagactaaag gagtaaggga 540
 gatacaggca ctaaaggga taagcggatg acacggatag caaacgtaag gagatgacag 600
 gcatagcaga actacgaaaa tgaacgaata gagaggctaa gagggtgag acgatgcaaa 660
 gaataagaga ataaaagacg tgcaggcatc gaaaactgca agagttaggc catcggcaac 720
 tcgaccgggt acaacgctaa cgaacgtaca gggctcggcg aatcgacaaa ctcgaaggat 780
 cagaccacat ggaacgctc ggacaataac ggaataagca actacgacgg tgaaggcgtc 840
 aacaagtgga cgagtaaagg cgtgaagacg tccagagctg gaacctgcg ggcataaaga 900
 gagagctcaa tgcgcagact gaaacaataa ccaggtagag gagatccacc ggtcaggcga 960
 taaagaaagt cgacaggtga cgaggtgacg aggtacgca ataaaccaat accggaatga 1020
 gaaggtagc gcgtagaaaa gatgaggaga tcagagcgat gaagagatag cgaagtaaag 1080
 aaaatacga agtccgggag aactcgag 1108

- 5 <210> 10
- <211> 1107
- <212> ADN
- <213> Secuencia Artificial
- 10 <220>
- <221> Fuente
- <223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"
- 15 <400> 10

ES 2 614 810 T3

ggatccgagc tacaagacta caagagtcgc gcgccgtaaa aaaagtggcg caggatggaa 60
 acgtgaggag gtagecgaagg tacagacctg gagagatcag gaactggcag aatgaggaca 120
 tcagccgatg aacgcctggg caaataacgc aatagaagag taggcagggt agagcgctaa 180
 cggaatccgc acgataagag cgtgggaggc taacagaatg gcccaatgac gggacgtaag 240
 cagatgaaaa cagtggcggg gctggaagaa tcagagcagt ccgaggggtgc agaggttaagg 300
 agatgggaga aatgcaaaga atgcggaagt aaacaaataa gaggatcggg gggctaaaac 360
 gggtcggcga gtgaaaaaat cgggaagtca gcgcataaag gcctgaaaaa gataagccga 420
 tgaagaaact gggcggatga cgcgggtgaa agatatccgg gaatggcacg gtggcgggat 480
 ggaaaagtgc gggaatcggg ccgtgaacga atagaacggg tacacgcgat aggcaagata 540
 accggatcga gagcactcaa gcgataccaa gctcgcaaga gtaggagcgt accgggagtg 600
 acaagataag agagtaggga gatccagaag tcacgacatc gggagctgcc gggatcagag 660
 ggtcagggaa tagcgaaaaa tgcaagaata gaaacatgca gacatgaacc gggtaagcca 720
 atgaaccact gacagaggtg agagcatagc gagctaggac ggtaaaagag gtccaacgcg 780
 gtccagagat ccgaaaatcg aaagagtaaa aagataggaa ggtgaccac tcgaggggat 840
 gacaacgtga aaaaaatggc gcgataacaa cagtacgaaa gctacgggaa ctgacagggg 900
 cgagctctag caaaactcgc ggagtggaac ggtagaggag tagggcaata gaaacatgaa 960
 gcaatgaaag ccgtagggaa actccaccag tcggagaata cgacaggtca cgcaataaaa 1020
 gccgtcaacg ggtgaccggc tacaagcgtg caacaaaata agaacagtgg gacagagtag 1080
 gccactagaa ccataaccaa actcgag 1107

5 <210> 11
 <211> 1107
 <212> ADN
 <213> Secuencia Artificial

10 <220>
 <221> Fuente
 <223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

<400> 11

ES 2 614 810 T3

ggatccgact acaaaaatgc acagatggcg aggtaaaaag gtagccaagc tgcgcgggtg 60
 agcggatggc gaagtgcaca ggtaagagag tacacagctc aagcagtagg gcaagtaaag 120
 aactcacaag ataagccgat caaggcgtcc aagcgtggcg aggtgggaga ctacaacacg 180
 atagagcagg taaaaaagta aaaagataac aaaaatccgg gaagtgaagg aatcaaacia 240
 tggagaagtg aggcaactgg agaactcaga gcatacggca gatggaacgg tccgggaaat 300
 caagcgagtg agagggatcg ggaatatggc agaatcaaac aagtgggaag atcaaccggg 360
 gaggagacta aacgcataac cggatagagc cgctacggca ctaagaccgg tagggagctg 420
 cgggagtgcg agacatggac gagtggacac atcgagaggt caaggagata gagaggactg 480
 agggactccc aggataaaaa ggagtaagag agtccggcag gtcccagata tccgacgtca 540
 aacgataaaa aatggggcgg atacagacgg gtcaagaaat cgcacggctc ggacgatgac 600
 ggaataggaa aatgaggcgc taaaaaacta gagaaatgag gaagtacgag cgtcgaggaa 660
 tgaggaaata caaggatagg caaatgagac gatggaggcg catcgacaag tcagaaaaat 720
 gccgcgatcg aggacgtcga cagagtagga cactgcggga aataaaggaa tcagaaggta 780
 gagagcgtca cgcggtggaa agctcagcgc aatcacggac tagaaggaat ccggacatcg 840
 agaaactacg agaagtgaaa aagtaagccg atccagacat gcccacaatg gcaccataac 900
 agagctcaaa actggcgagg agtgagccac agtaagacag tcccgaata aacacataca 960
 aggctagaaa aatgcgaaga gtgcgagaag gtcgggaagg tagagagaat acaaggctga 1020
 cggagtgagg aaataagaga gtggacagag tacgagagag tgcgaccaat aaacgggtag 1080
 gcgaatcaac ggataggaga actcgag 1107

5 <210> 12
 <211> 1107
 <212> ADN
 <213> Secuencia Artificial

10 <220>
 <221> Fuente
 <223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

<400> 12

ES 2 614 810 T3

ggatccgagc cagtcaagca atcgggagaa tggcaccgtg ggcggctaag caaagtacga 60
 aaaataacaa catgcgggaa tcggagcgtc cagaagtcaa ggaatggcaa ggggtgcgcaa 120
 atggaaggt aagagggatg aacccataga gaaggtgacg ggatgagaac gctaagaaaa 180
 tcccaaaata ggagagatag gaggggtggac gaaatgcaa gatgacgcaa tgacagaagt 240
 aacgggaagt gacagactgg cccacaatgc agaggttaagg ccctcaaccg aatcgggca 300
 tgaaacactg aaaacgtaga aaagtgcagc gactagcagg agtaaacga tgcccgaata 360
 aaaacctaaa ggagtgggag aatccacgag taagaaaatc aaggagtaac cgactcagaa 420
 agatgaccaa ggtcaaggac tcaaaagctc ggagagtaac ccagcgtcac aaaatgagag 480
 cctgcaaaaa tagaaacatg cggcaacctg aaacgctgcc agacagtggg aacgtccaca 540
 ggtggagggg tgggaaagta caacactgac cagatcggaa aatggcaaca ctacgaaggt 600
 ggatatctag cagaatgaaa agggtaaaga gactaaagca atccagaagt acaacgatag 660
 gagcgtacga gaatgcaaag atgaacgggt acgacaagtg aaaaactcaa gagatacaac 720
 catgcacgag atagacacgt acaaaactac aaaaatgaac acacgtagga gagctgaaca 780
 aagtaggccg catggagaag tacggcggct ggcagaatga aggcgtaaga gactaagcg 840
 gagatggaga cataagcaca tgggaacgtc aaaaaatcag agagtgcaaa cataaacaca 900
 tgagctcatg cgagagtgc acaaatcag aggtagaggc gtcgaggag tagcggagta 960
 cccaacatgg accactggaa aactcagggc ggtggacgga tgaggaaggt acacgggtag 1020
 aaacatagcg cgatgagagc aataagaaag gtgaaagcat gagagactac agaagatgga 1080
 caccgataaaa acgtaggcaa gctcgag 1107

5 <210> 13
 <211> 1125
 <212> ADN
 <213> Secuencia Artificial

10 <220>
 <221> Fuente
 <223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"
 <400> 13

ES 2 614 810 T3

ctcgaggacc cactaaggca atgaaagcct agaggcatga cgggataaac gggctcggga 60
 agctgccggg ctgacgagaa tggaggccta aagaaactgg agcgatcaga cagggtagca 120
 cgctcacgca ggtgacagc aatacgacgc tacaggaatg gaaagaatgc gagagctaaa 180
 acagtccaag ggtagagcgg tagagctcta gggaaactaag aagagatgga gaaatagaga 240
 cgatagaaac ctagcaaat ccgaagctcg ggagatccag cgagtgagga gatacacgaa 300
 tgcgagcgat ggcgaactgc caacggctga acacaatgag caaatggaga aataagcgaa 360
 catagggcga taagagaccg cggcaacgctc gggaggctca caagtagagg aataaacgag 420
 tgggaaaatc agaaaataaa gaggataaag gcgtcaggaa atgggaggaa tcgggaaatg 480
 aacgcgtaag aagatagaaa ggtgcccag gagtcaaccg aatgacagac gtagggaag 540
 atacaacaat caccaaatca aaaagtgcga gcagtcgagg aatcgggagg tagaaggaat 600
 gacaacgatg aagaacatca aaacgtgaca ccctaaggcc ctgagcggat aaaagagtga 660
 ggcaatcgag ggtgcccgc gtagacaagt gcgaacgtag accactaaga aagtcagcag 720
 aataaacaga gtaaaccact agcaagatca aagactaaa aactacgcac cgtagccaga 780
 ctcgggcaat gaacggatca aaagataagg gaaatgacag gatgcagaag aataacggga 840
 ctacaaggct aaggcagtc gaaagtaacg cactggcagg gtgaagacct gcaacagtga 900
 aagggtcggg caatagcaga agtgaccaca tacgccaatc gacaagtaaa gaggtaaagc 960
 aactaacgaa cgtacaagaa ataaccggct gaaaggaact agcgaaatga gggagtgaag 1020
 aaataagcag aactagggag ctgcgacggg tggcagagag tcgagagata caaaggtaaa 1080
 caaactcgca agggatgaaga gaccatggaa gcatgcaaaa gatct 1125

5 <210> 14
 <211> 1125
 <212> ADN
 <213> Secuencia Artificial

10 <220>
 <221> Fuente
 <223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

<400> 14

ES 2 614 810 T3

ctcggaggacg ataagcaaaa tagggccgtg agacaaatgg cgacatgaag caatacccaa 60
 gtgacaagct agagaggtaa cagcatagac aaccctaacg ggactagccc aatagaaga 120
 gtaaaccgggt agggaactga ggacctgaaa acctgcaaag actgggcagg gataggaaca 180
 ataagaagat gaacagatga gcgagctcgg gaaatgagaa aataaaaggc gtacgggagt 240
 gggacagtca cgagaataaa ggcgtcggca gatcaacgga tgcagggcat gggacagtac 300
 gggagagtac gcggctgaag agctgacacc ctgagggaag taggcaata aaagggtagc 360
 ccactagcga gcgtcaccgg gtggaaagct aggaacgtcg gaaactagga gagtcagcag 420
 ctccaagct gagaaactcg gcagatcacg gaccgcggga agatgaaaag ctgaggaggg 480
 tggacgggtc agaactggtg aaactagacg acgtgggcga atacgcacgt aaaggagtac 540
 gacacatcag ggcttgggaa gataccaaga tgccggaaga tcgaaaacta aagcagtgga 600
 acagtcaaca aatcaagggc gataagcgaa taaggaggtc agcaggtgga aaccgctaag 660
 accgtgggaa actcaggaaa tacgggcagt aaggcggctg gcagactagc gacatgagcg 720
 ggtcacagaa ggtaacgcaa taacaaaatc gcgcagtggc acagtaaagg cctcgggaag 780
 tgcggagatg caagagtacc aaaggctaag gcaactggca cgatacggga actagggcga 840
 ctgacagcat ccacgcagta agagaatggc gggatggagc gctaagacgg tgaaccaata 900
 agggcctaac aggatgacaa ggatgcggga atgagccact aaaggaagta gaggagtaag 960
 ccgggtgcga agctggaggg aataggaaaa tacgaggatg ggcaggtgag gagaaatcgg 1020
 acggatgaac ggctggcaag gtaaaagggt agcggaatag agagcagtcc gcaggtcaaa 1080
 cgggtgcacg ggctcagacg ggccatggaa gcatgcaaaa gatct 1125

5 <210> 15
 <211> 1123
 <212> ADN
 <213> Secuencia Artificial

10 <220>
 <221> Fuente
 <223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"
 <400> 15

ES 2 614 810 T3

ctcgaggaaa aatggcaagc taagaggaat caagaactgc ccacctaaga ccaatgaggc 60
gataaccgaa atcgggcaat agccgaataa agggaatgag acgggtgctc gactacacaa 120
gtgcgcaact aaaaaataa cgaagtggga agcgtacaa agatgggcag atcggcaggt 180
aaggacgtag ggaagtacag gagctccgcg actaggacca tccaacactg gcaggataga 240
gcggtgaaaa gctggagaaa cgtcgggacg tgggcaggtc agagggtgca gaagtgcgg 300
gcataagcac atacaacggt agcagagtcg aaaacataaa gagactggac gaatcagaga 360
ctagcggact agccacctgg gagagtaccg gggatacaag ggataagagg aatggcgag 420
tggacagatg gaagcatggg aggatcacag aactagggag atagcgagat accagcgtgg 480
agaagtaaaa aatgagggga ctaaggaat gaaaaagtaa gaaacccgcg gggtagacca 540
gtgcagcagt aggagaatac acgaatgcag ccagtcaaga gaaatgaaga actaagagag 600
tgcgagagta cagagctacg caagtgccca agtgagagaa tagcggcct caagcgatca 660
acgacataac aggagtcaag agaatcgcaa agtcacggga tgcgagcagt ggacaaatca 720
gcaaaatcaa agactcacia gatccgacia tagagggata aacaagatac aaacatccag 780
agactggcag gataaagaaa gtacaaagcg tgccaagcta aggaagtaga gaaatccaag 840
aatacagagg tgacgccgtg aagacatgctg cagggtgagca ggataggaga ctaaaggcgt 900
acgggaatgc gaaactagaa ggctgaaagg atcgaccac tgcgagcgtg gagggctacg 960
acaactaaag acataagcag atgaagccca tcaaggacat ggcgcatgg gaagatccaa 1020
gagatccaag ccctaggaaa gatgaacgcc tgaacagcta agagcgggtc caaggataag 1080
aacatgcaga aatggacgag ccatggaagc atgcaaaaga tct 1123

5 <210> 16
<211> 1125
<212> ADN
<213> Secuencia Artificial

10 <220>
<221> Fuente
<223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

<400> 16

ES 2 614 810 T3

ctcaggacc aaaagtcaag aaagtaccgg gctagaagag ctgaagccat aagcgagtag 60
 cagaatagaa agatcccaaa agtcccaggg atagacgagt aggaaggtga aaaaatggcg 120
 aggtagcgac atgcaaaggt aaaacgatga aaaactacga ggggtgaaga ataagcaggt 180
 gacgaagtaa acgggtcaaa ccgagctcag atcgaacgat aggaaacatg accgggtcac 240
 acgatcgaga ggtcccaaag gatagaagaa gtgaaaaggt gagcaaggtg gcgaaaataa 300
 aaagataaaa gaataagaca gtagcgggaa tacgacacta gaaggatcgg gacatgcagc 360
 agtaaaccba taggaggata acagggcatg gaagagtggg acggtaagac cctacacgaa 420
 tacaagcagt gccaggatgg cgcgagtgc aaaaagtaga agggtgaccg agtcgagaga 480
 tagagacgta aggaagtagg gaaggtggga cgatcgaaag atcgaagagt aggaggcgtc 540
 gaaaaatccg gaaaataggg aagatggacg gatcgggacg gtgaggagga atagcaaaag 600
 tcccgcgggt acgaaaggtc gggaaggtca agcaatcagg cgctcaacgg gactgaacaa 660
 ataaggacat acacaagtcg gcacgtcacg aactcaaaaag gtggagaagg tagcgaatc 720
 gaggagtgga gaaggtaaag aatgggaag gctaagaaa tggcagggtg gagaactggg 780
 acggtaaacg catgaaagaa tcaggagta gaagaacgtg aaggatagg agaactcaac 840
 agggtagcag aagtggaaag catggcaaga atggcagcat gaaaagatcc aggagtaagc 900
 gagctgaaga aatggagacg taacaacata gcgggagtag gcgcgtgaca agataacggc 960
 aatgcccagg tcgaggaatc cgcaagtga cactcaacg caatgaacgg atgaacacat 1020
 gcacgaagtc gacaagtaaa aaactggaa gccatgacaa cataacggga tggcaggata 1080
 agagagtga acgatgcacg agccatggaa gcatgcaaaa gatct 1125

5 <210> 17
 <211> 1108
 <212> ADN
 <213> Secuencia Artificial

10 <220>
 <221> Fuente
 <223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

<400> 17

ES 2 614 810 T3

ccatgggaca gcgatcagag ggtaaaacgg gatgaagcag tgaaggacc tcagcgaatg 60
 aaaaacgatg gccagatcca aagataaaaa actgaaagac tacggaata caagaataga 120
 agggtaaacg actgagaaag tacgaagcct agacgggtaa aaaaggcgg gaagggtaac 180
 gccatagaca aatgagaagg taaaggcatg gaaaaaatgg aggcacgac gaatccccg 240
 ctcaaaggat aacggactag cgcggtaaaa gggaatgcgg acgatcgaag aagctccgga 300
 ccgatccacg gaatagagac atacgacagt gcgcaaatg gaccgataaa agggtagacg 360
 aaataacagg atgaacagga ctcggagaaa taacaagtg gagaaagtac aaaagtcaac 420
 gaataggcca gtggcaaaag tgagcgagtg aacaggtaga ggagtggaaa agtacaagg 480
 atgcaaaaat gaaaggtag aaaactaagg cagtacaggc ataaacgact cacaactag 540
 aaaactacag cagatcgaag cataaaggaa ataggagaga ataaaaacgc ctaacaaact 600
 acaagaacta cgcggagtgc gagacgtcag gcaagtgagg acctgaaaca atgcaagaat 660
 ggcaagtgg acgcggtagc gggataagca aatacaccgg tagatatcat aggaaggta 720
 cgcaaatgga ggagtcaaga aactggcca gtgaagccct cggcagatac gcaaagtacg 780
 acaaataaga ggctcagaag atccagacga gtgcaggaat aagacaatca agagaatgaa 840
 cgcacggaa cactaggcag cagtgggacc ggtaaagca tagctagctc gggcgatgga 900
 ggcacgtaca aagggtacaa aagtaacggg aatacccacc gtaccaggat gaccagggat 960
 cgcgaagata gcggaatcga gccctcagga gctaggccag catagagacg tcacgagatg 1020
 aagggataaa ggaagtcaaa gaatgggaga actccggacg tacgaggcta cagaggtgaa 1080
 aagataaagc aggggtgagga aggcacg 1108

5 <210> 18
 <211> 1107
 <212> ADN
 <213> Secuencia Artificial

10 <220>
 <221> Fuente
 <223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

<400> 18

ccatgggaca agagatagaa agcagtgaaa gaataggacg gtcagaggat ggagggcctc 60

ES 2 614 810 T3

cgggcgtggc acgactagga cgatgcggaa ataacgacaa gtggaaacct agccagctaa 120
 ggaagctaag ggcaatggca aagtaaggaa gtacggaaat aggaccatag aagactggac 180
 cgatacagcg ctaggcgggg taaacgagtg aaagggtgga acaataagga cagtgcagca 240
 ggtaagacca ctaaaagact ccacgacgta cagagactcc gcgcctggaa cgatcgaagc 300
 gtaacgggca taaggaaaat agagaaggtc gaggaaataa agggaaatgg agaaactaag 360
 cggatagggg aataaacgaa cctcagggaa tcccaaagtc cgaccaatga cagactgcac 420
 gcatccgagc gtaaaaacat gagaccaata acgagatcgg caagtcgaga agtcgcagaa 480
 tcaaacaacc tgagaacctg cgggagtgaa cggataagac ggtaagagaa ataagagcat 540
 gagaaactgg gacgatagaa cgatagccaa gtaaaaggt aagagaatgg aacgaagtgc 600
 aaaaagtggc agaatgaaca gataggcaga tcagaagaat gaggaagtgc cagaaataag 660
 aggtggggag cgatgccggg atgcgcgaat gcgaaggtaa gagaatgcag gagtaagag 720
 gactgaaaag atcgggacat gaaacgatag gaaggtacgg cgatgagaca gtacaaaagt 780
 gaaaggtgga cagctgcgc ggatatcagg gatgccacga tgggcacatg cccaaaatga 840
 aaaaatcacc aaataccaaa atgaagccga tgcgggaatg cgctagctaa cgagcataaa 900
 acggtaggaa aatggaaagc taaccgcagt ggaagaataa ggagctacgc aaatacggc 960
 aataaggaag tagcggacta aggaggtaga cgaataggcg aataacgca gtcgagaat 1020
 ccaagactac aggactcagc agatgaaaaa actagaaaacg tgaaggccta agaaacataa 1080
 gacactgaaa gagtagcgga ggcctgc 1107

5 <210> 19
 <211> 1106
 <212> ADN
 <213> Secuencia Artificial

10 <220>
 <221> Fuente
 <223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

<400> 19

ccatgggaga gtaaggaaat gagaacagtg aagacatccc aagaaatgaa aaaagtggag 60
 aggtcgaacg gtagagcagt ggagaaggat acgccgatcg ccgggataac cgggctaacc 120

ES 2 614 810 T3

acaatgaaac acgtggccga atacggagga atcagaggag gtggcaggac tgaacgagtg 180
 ggccgggatag aaaaactaca gcgatacgcg cgataggaac ctacgagaac taagaggata 240
 aagacataag ggcctacgca cgagtaaaag agtaccgacg tcagacaata gaagggtaaa 300
 aagatgaacc gatgagcaaa atccaggcgt cgcaagggtca cgcagtcaag acataagaga 360
 atgccagaag tacaagcctg ggacggctga gagagatcgg gcagtcaaaa gggtcaggac 420
 atagcgggat agccgaatgc aaagatacga cggtgcaaga atcacggcat aggcaagtgc 480
 aaaacgtaac aacctggcc aaaatggaag actgaacgca tgacacgggt cacgcagtgc 540
 agacctgcaa cagtacagaa atggaaaact agaagagtaa gcaaatcgaa acctccaagg 600
 gtggaaggat ggacaggtga acaggtaaag agatcgcgga catgagaagg taaaaagcta 660
 aacaagtccg gaggtgaacg aataaggacg ggtaagggac ctggaccgga atggcaacat 720
 gcaaacataa gaggtcaac caatggaagg ctgaaaagat cgaaaaatgg gcgaatacaa 780
 aaggtaaagg gatagcggga tcagaagggt ggacgatgaa agaatgaaga aataccaagc 840
 gataacacga tccggaactg cgggagctga acaagtcacc gctagccaga gggtcagggg 900
 gatatcaaat gggcagatac ggagcgataa aaacatgaaa ggagtgggcc actggaagga 960
 tcagcacgta acggcctaaa ggactaagag cacatgagcg aagtgaagag agtgaagaaa 1020
 ctaagaaaga gtagagagat gcgagagata gcgaaaatcg acaacatcgc gggagtggaa 1080
 aactgacggg atgacgagaa gcatgc 1106

5 <210> 20
 <211> 1107
 <212> ADN
 <213> Secuencia Artificial

10 <220>
 <221> Fuente
 <223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

<400> 20

ccatgggaac agactaaaag aatcaacaga tagaggaatg aggaagtgca ggaagtaagg 60
 aaaactggaa gagtaacaca atgggagaat acaaaagtca accagatgga cagatagaga 120
 aatgacgaga tggaaacagt cacacgctaa gggaatggac gcgtgaccag atcgggaaga 180

ES 2 614 810 T3

tccgggcaat aggacagtag aaaggtgcag gaatgacaag atggccaaat cacagcatag 240
 agccaataag acggtaaaag gcgtagcacg ataggacggg tcacgagagt gagagaggtg 300
 ccaacaacta ggacaataag ccgataaaag cgtcgaaca tagagacgtc cagagaatgg 360
 acccaatcca ccggatagca agagtagcgg agataagaac gtcaggagat agcgaagtca 420
 cgaaatgaga gagtcacgcg gtagcacaat aaagacgtac aaaagtaca caaagtgaga 480
 acatcaaac gataagcagg ataaaaaggt aaaacgggat cgggacgctc gaaaactgac 540
 gaactagaaa agtaagaacc tagaaaaata gcggcataga aaactaaggg aatggcgaac 600
 ataagaggaa taggaaggtg gcgaagtgga gcaataaagg aggtgggacg gtcaagagct 660
 agagaaatcg caacgatacc ggaatcggga agtagaaca atcagcggcg taggacaagt 720
 ccgggaatcc acgagtcgaa aaaataagac actcagagag tgcggaggct aaacgggtgg 780
 aaagaactgg aaagatccag agcatcgag aataagacga gtgcgaggat caacggataa 840
 ccacctaagg gcgtcgggcg ataacgcag taccgaaatg ggctagccac gggatgggag 900
 aatggaaccg taggaccggt agagaaagt gaaactacg gagaatgcaa cagtaggaaa 960
 gctgacgggc taaggccggt aggacaagct ccgcgcgtga agatatcgcg gagtagagga 1020
 ataccgggcc atgacgggca ctcaacggct gaaagagctg aagcagggtc aggactggga 1080
 ggagtaggga cgatcacagg agcatgc 1107

5 <210> 21
 <211> 1107
 <212> ADN
 <213> Secuencia Artificial

10 <220>
 <221> Fuente
 <223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

<400> 21

gcatgcgaag aaaagtggaa cactagggcg gtgaacggga tgaaaaggat agagcgggtc 60
 gaaagtcaag aaataggaga actaacgaag tgggagagac taaagacgtc agaagatcaa 120
 aaggtcggag gcaatggaga agtcagcaa tcagaccgtc cggcgctgaa aaccaatgaa 180
 gacagatcca agagctgagc agtagcgaa gagtggcgag cgtaaagcag tagggaggtg 240

ES 2 614 810 T3

aaagaactac aggcctcaga caatccagac gtaagagggga tacagccgtc agggaggtaa 300
gagaatgaaa ggatgcgcc atgagccaac tgggagggct aaaaaactag aagagtgaaa 360
aactggcacc gtgagaaaat agcacaatac gaaaggtcga gaaatcggga aagtgcggaa 420
gtggaccact agcgagatca acagagtagg gacatcgaag aataaggcag aatgcgacag 480
tacgggggat cgaagactg cgagcgtgac gagatgaaga ggtaaaaaac tgaaaacatg 540
agggcgtgaa agggaataac caggtgggac acagtgcgg gcattggcggg cctcaagaga 600
tgcacgactg agcaggtgag aaggtaacac gatacagaaa tggaacggct caaacaata 660
agaaggtcgg cccgtgagac caggtgaagag agctggagga cccgcggagg tgagcgggtc 720
aaaaaatcga gagataagga gagtgaacgg gtgaagacag tagaaaaatg agagaaatcc 780
ggcaaatagc aggactggac gcgtacgaaa gagtggcaaa ataagcggat ggcgagatgg 840
cgggctcggc gagatagcaa gatgaagacg taagaacggt aaaaggctca agagatgaac 900
aaataaagag atacgcggct aaaagaccta aaaggataag aaactcaagg cagtgcgaa 960
acgtgcaagc agtaacagaa tgagaaagga tgaacaccgt ggcgggatag gagagtggag 1020
aaatggaaga atacaagaat aagaacggtc ggacgcataa acaggtgaag caataagaga 1080
ggtcagcaca cgtacgaagg aagatct 1107

5 <210> 22
<211> 1107
<212> ADN
<213> Secuencia Artificial

10 <220>
<221> Fuente
<223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

<400> 22

gcatgcgacg gatagaagca tggcagagga tcaggggaagt caaacgaata caaagatagc 60
caactaggac caatcagaga ctgagaagcg taagcgaagt acgacaatcg gaaagtcgag 120
ggatgcgaga gatacaaaag taagagggta gaaggcagtc gggagccatc cgggagctaa 180
caaaataaag aactggcaga ggctagcgtc cggggagata ggaaggatga caggccgtcc 240
gaaagtacgg aatcaaaaa gtgaggcact cagagagtga aagcgtgaaga acgggtaggg 300

ES 2 614 810 T3

agctaggcgg gtaggccaca tcacgggata aagagatgac aagcgtgaag gaatccacgg 360
 agtgcgcaga cgtccaaagg tccacaactc cgccgggtac aacggtgaag caaatcacgg 420
 gctcgaaaag gtggcgagat gggaccatga aagaatcgaa gggtaagacg atccagagat 480
 gagcccataa caaggaaca acctaacaga agtaccagaa tagagcaact gaaaaagtac 540
 agcactggca acgtccaagg ctgcggcct gcgcgctgc gccgaaataa ggacaataac 600
 gacctaggac cgaatacaaa agctaacaga ctgagagcat ccaacgctga gccagtcaga 660
 aggtagagaa gtaagcgaag taaagaaaat aaagaaatgc cgcgatcggg aggtgaagaa 720
 ctaccaggct acggagaatg aaagggtgaag agggtaagga catacaagaa taaagaagct 780
 aggacaataa acgcctaaa accgcggaga gaatagcccg atacaagcgt ccgggcatag 840
 aaacctggag aataggaca ggtgaaaggc tgggcacatg accgactgga gagcatggaa 900
 acatacacc agtcaaaaag tcgaggaata gccgggtggc acggaacaa aatcaagaaa 960
 ataccggagg gtgcagaagg gtacgaccgg tacaggacct aagaacgatg gcgagagtcg 1020
 ggccgtgagg acagtaagga cgtggaggaa gtagcagaat agcgggatag ccagctacga 1080
 agaaatgggc aagtgcggag aagatct 1107

- <210> 23
- <211> 1107
- <212> ADN
- <213> Secuencia Artificial
- <220>
- <221> Fuente
- <223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"
- <400> 23

gcatgcatc gaaaaatagg gaggtggccg gctgcgaaag tcgggcgggt gaaagcaact 60
 aaaaggatcg aagcatgaaa gagtaggaaa gtggaagaat gagaagataa cgaaataagg 120
 aaggtaaaac ggtggagaga tagaggacat acaaagtgg aaacactcaa gagctaaggg 180
 aactagaagc ataaccgaat ggctagcatc gggagagtag cggcatgaag cgatagggaa 240
 cgctgacaag aaataaccga atcgggaaat caaacatag aagactagac gggtaacaga 300
 atggaggcaa taggaaacgt aagacactaa cgggatacca cgagtgacaa gatcggagga 360

ES 2 614 810 T3

tggcaccatg aaaagataga gagctagaag ggtacgggaa tagaaaaatg caaagctagg 420
 gccgtcgacc ggtgagaagg gtaaaaaggg tgacaaaatg agagaataac cagatagggga 480
 cgtgaaaggc taggcacctg gagacaatga ggcagtacac gcgtaccaag atccagagaa 540
 tcagacggtg agacactgga caccataaga agatgaggag gtgagggaca tgaacaata 600
 acaacaataa ccacataagg gcctcgaaac gtggagagca tacagccggt gcaaaagtga 660
 gagggatgcg aaaaatgaac aggctgggac gatacccaga atgccaagat ggcgcctgc 720
 cgggatcacc gagtagccaa cctgacaaaa agtgcggaaa tacgggatg agaaggtgga 780
 gcactcagac gatgaagacg atacggacgt accaaaatgg aaaacaatgg gagcgtaac 840
 aaaatcagac ggtagaacgg taaagagatg cacaagatga agagcatcaa cacatgacag 900
 cccgcgggag taggcggagt aagcgggtaa cgaggtgagc acataaaaaag gtcgcaaagt 960
 acgaaggtaa gggagtggag agagtgaagg gcctggaggg agtcggaaca atgccgacgt 1020
 aagcaagta aacgctaagg caagataacg gaatgcgaaa ctggcgcct aagaacgtag 1080
 gaccataaca aggtagaaaa aagatct 1107

5 <210> 24
 <211> 1107
 <212> ADN
 <213> Secuencia Artificial

10 <220>
 <221> > Fuente
 <223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"
 <400> 24

gcatcgcgatg gaaagctgag agaaagtcag aagatcacag actcaggagg atcgggcagt 60
 agacacgtaa gaaggtagag gccaatgaca ggctgaagag gtgaaggcct aaaagaatac 120
 gcgggtcagg aacaatgcga agcctgggac gatcgggaga taaagggat agagaaatgg 180
 ccgaatagac cggtagcgcg cgctagcgat aaggaactag acgggtaaaa ccgtaggaag 240
 ctcagaaaca taacgaaagc tacagcaggat aagcaagcta agaacctgag caggataacg 300
 cagtaaggac actacgggaa ctaagccaga tgaccgggta cgaacctcaa acgaatagca 360
 aaggtgcggg acgtggccga agatagaaaa catacccaaa tgccggaatg ggagaatgcg 420

ES 2 614 810 T3

caagtaaaaa cataggaaca gtaaggcaat cgggaggtga aagggaatga gacaactaac 480
aaggtgggaa aaatccagca gtaaacagca tagggcaatg agaaggtaga acaataaaga 540
cgatgaggaa gtaaaaacgg gtgggcaaat gacacgatga aaaagtaacg gagtaaacc 600
actcagaaac tggaaaaagt cgaacatgg gaagaataca ccacatccag cgaataacgc 660
gactccaac caatagacga gtgaagagat ggaagccctg gcgaacatgg agacaatagg 720
aggtcaagg acgtggacac gtacagcggg aacggcctca gcaggtggga ggctacgacg 780
aatggaggag tgcagcaata ggcgggatag cagcctgcaa ggatcggcaa ctgagaaagt 840
gaagagaatg ccaaccctgc aagcgtaaaa agcgtgaagg cgtccagaag tacgcacact 900
ggacaaatag acgaataagg gcctcaaaag catcgcgagg ataagaaagt gcgggagtaa 960
caacctcggg aagtcggaaa gtagaaagg ggaccgataa cccgcgggca caataaacga 1020
actgcgagca tcagagaggt gcagaagata ccacgagtag aggcattgga acgtgaaaac 1080
aataacgcaa gtcaggacga gagatct 1107

5 <210> 25
<211> 15
<212> ADN
<213> Secuencia Artificial

10 <220>
<221> Fuente
<223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"

<400> 25
ggtctgtgtg atgtt 15

15 <210> 26
<211> 69
<212> ADN
<213> Secuencia Artificial

20 <220>
<221> Fuente
<223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético "

25 <400> 26
aacatcacac agaccaacat cacacagacc aacatcacac agaccaacat cacacagacc 60
agccctttg 69

30 <210> 27
<211> 3194
<212> ADN
<213> Secuencia Artificial

35 <220>
<221> Fuente
<223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

<400> 27

ES 2 614 810 T3

aagcttgcc acacagcttg agtattctat agtgtcacct aaatagcttg gcgtaatcat 60
 ggcatagct gtttctgtg tgaattggt atccgctcac aattccacac aacatacgag 120
 ccggaagcat aaagtgtaaa gcctggggtg cctaatgagt gagctaactc acattaattg 180
 cgttgcgctc actgcccgtt tccagtcgg gaaacctgtc gtgccagctg cattaatgaa 240
 tcggccaacg cgcggggaga ggcggtttc gtattggcg ctttccgct tctcgtca 300
 ctgactcgtc gcgctcggc gttcggctgc ggcgagcgg atcagctcac tcaaaggcgg 360
 taatacggtt atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaggcc 420
 agcaaaaggc caggaaccgt aaaaaggccg cgttgcctgc gttttccat aggtcctgcc 480
 cccctgacga gcatcacaaa aatcgacgct caagtcagag gtggcgaac ccgacaggac 540
 tataaagata ccaggcgtt cccctggaa gtcctcctgt gcgctctct gttccgacct 600
 tgccgctt acgatacctg tccgccttc tccctcggg aagcgtggcg ctttctcata 660
 gctcagctg taggtatctc agtccggtg aggtcgttcg ctccaagctg ggctgtgtc 720
 acgaaccccc cgttcagccc gaccgctgc cttatccgg taactatcgt cttgagtcca 780
 acccgtaag acacgactta tcgccactgg cagcagccac tgtaacagg attagcagag 840
 cgaggatgt aggcgggtct acagagtct tgaagtggg gcctaactac ggctacacta 900
 gaagaacagt atttggtatc tgcgctctc tgaagccagt tacctcggg aaaagagttg 960
 gtagctctg atccggcaaa caaaccaccg ctgtagcgg tggtttttt gttgcaagc 1020
 agcagattac gcgcagaaaa aaagatctc aagaagatcc ttgatcttt tctacggggt 1080
 ctgacgctca gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa 1140

ES 2 614 810 T3

ggatcttcac ctatgatcctt ttaaattaa aatgaagttt taaatcaatc taaagtatat 1200
 atgagtaaac ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga 1260
 tctgtctatt tcgttcaccc atagttgcct gactccccgt cgtgtagata actacgatac 1320
 gggagggcct accatctggc cccagtgctg caatgatacc gcgagacca cgctcaccgg 1380
 ctccagatt atcagcaata aaccagccag ccggaagggc cgagcgcaga agtggctctg 1440
 caactttatc cgctccatc cagtctatta atgttgccg ggaagctaga gtaagtagtt 1500
 cgccagtaa tagtttgcg aacgttggtg ccattgctac aggcacgtg gtgtcacgct 1560
 cgtcgtttgg tatggcttca ttcagctccg gttcccaacg atcaaggcga gttacatgat 1620
 cccccatgtt gtgcaaaaaa gcggtagct ccttcggctc tccgatcgtt gtcagaagta 1680
 agttggccgc agtgttatca ctcatggtta tggcagcact gcataattct ettactgtca 1740
 tgccatccgt aagatgcttt tctgtgactg gtgagtactc aaccaagtc tctgagaat 1800
 agtgtatgcg gcgaccgagt tgctcttgcc cggcgtcaat acgggataat accgcgccac 1860
 atagcagaac ttaaaaagtg ctcatcattg gaaaacgttc ttcggggcga aaactctcaa 1920
 ggatcttacc gctgttgaga tccagttcga tgtaaccac tcgtgcacc aactgatctt 1980
 cagcatcttt tactttcacc agcgtttctg ggtgagcaaa aacaggaagg caaaatgccg 2040
 caaaaaaggg aataagggcg acacggaat gttgaatact catactcttc cttttcaat 2100
 attatgaag catttatcag ggttattgtc tcatgagcgg atacatattt gaatgtattt 2160
 agaaaaataa acaaataggg gttccgcga catttccccg aaaagtcca cctgacgtct 2220
 aagaaacat tattatcatg acattaacct ataaaaatag gcgtatcacg aggccctttc 2280
 gtctcgcgcg tttcggatg gacggtgaaa acctctgaca catgcagctc ccggagacgg 2340
 tcacagcttg tctgtaagcg gatgccggga gcagacaagc ccgtcagggc gcgtcagcgg 2400
 gtgttggcgg gtgtcggggc tggcttaact atgcggcatc agagcagatt gtactgagag 2460
 tgcaccat at gcggtgtgaa ataccgcaca gatgcgtaag gagaaaatac cgcaltcagga 2520
 aattgtaagc gttaatattt tgttaaaatt cgcgttaaat tttgttaaa tcagctcatt 2580
 ttttaaccaa taggccgaaa tcggcaaaat ccttataaa tcaaaagaat agaccgagat 2640
 agggttgagt gttgtccag ttggaacaa gagtccacta ttaagaacg tggactccaa 2700

ES 2 614 810 T3

cgtaaagg cgaaaaaccg tctatcagg cgatggcca ctacgtgaac catcaccta 2760
 atcaagttt ttgggtcga ggtgccgtaa agcactaat cggaacccta aaggagccc 2820
 ccgatttaga gcttgacggg gaaagccggc gaacgtggcg agaaaggaag ggaagaaagc 2880
 gaaaggagcg ggcgctaggg cgctggcaag tgtagcggtc acgctgcgag taaccaccac 2940
 accgcccgcg ctaaatgcgc cgctacaggc cgcgtccatt cgccattcag gctgcgcaac 3000
 tgttgggaag ggcgatcggg gcgggcctct tcgctattac gccagctggc gaaaggggga 3060
 tgtgctgcaa ggcgattaag ttgggtaacg ccagggtttt cccagtcacg acgttgtaa 3120
 acgacggcca gtgaattgta atacgactca ctataggcg aattagtga tgagaagacc 3180
 tgcagaaaga attc 3194

5 <210> 28
 <211> 1106
 <212> ADN
 <213> Secuencia Artificial

10 <220>
 <221> Fuente
 <223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

<400> 28

tctagagaga aaggctacga agatccggac atcaaaagat caaaacatgc ggagatcgac 60
 acatccaaaa atgaacagat aacaggatca gagggtaag aagtgaacc atcgaggagt 120
 cgcagactac gaagatggac aagtgagca cagctcgat ggagaactaa aggagtcggg 180
 aggtacggca gtgaagaaat gtgcacatga aaagctgcc ggtgcgcaa ctggaggact 240
 agaagaatgg cacagtgacc cggtgagagc ctgaggaggt gcgagaatag ggcgataggg 300
 aggtaacccg atggcggact aaagagatac ggacatggac ggataggaga gtcgcaagat 360
 ggaaggatca caaagtgagc cggccaagc atacgcggg aaacggatca gaacgtaaag 420
 cggtggaaaa ataccgaggt aggaagatag caaaatacc aggtcgagcg atagcccact 480
 cggcgagtcc aaaactgaaa aaatggagag atgagacact agcggcatcg gacagtagag 540
 cagtgaacgc atcagagagt agaaaagtga agaactgca cggtgaaacc ataacagagt 600
 cagagggtag acccatcga ccctaccgca atacaaggct aagacggta gaggatggac 660

ES 2 614 810 T3

ggctcgcgaa ctcacgccgt gccaaagtga agaagtacca aaataggagc gtaacggcgt 720
agagaggtag acgagtgaga cgggtggcag ctccaaggt ccgagcatgg acaggtggca 780
aaatccagcg gtcagaaact acgagactgg ccggatgaag aaataaggcg ctagagagat 840
acgggagtga caagatggac cactagcga gtcggagggt gaagacgtga acaactagac 900
cagtacggca ctacagaaat agaccgatga aaacctgaaa agataacaga gtgaaggcat 960
cggagggtcg agaagtaaga acgtggagcc atacacgaat aaagcagta aaggggtggac 1020
aaataggaa atcgggaggt cccacggtga aaacgtaggg aggtgaacca gtgggagagt 1080
caaccaatac gaagatgcga ggatcc 1106

5 <210> 29
 <211> 1106
 <212> ADN
 <213> Secuencia Artificial

10 <220>
 <221> Fuente
 <223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"
 <400> 29

ccatgggaaa atgagagccg ctgacaagg ataaagaaaa gtcgccaac atccgccacg 60
ctaagcggcc ctaggagccg gtaacaagc atgaagaac atgcgaggca atgacacaaa 120
ctccaggcgg gtgaaaaga ctgacggag atgccacgca atagcgaac ctggaagaga 180
gtggaagcaa ataggaaaca gtagagcaaa atgccgagg atagaaaaca gtggcgcgca 240
ataagcagga ataaagcaca ataaaaagaa gtaaaccgca ctaaaacagc atccagaccg 300
ctgaaaagaa atggcggcga atgaaaaaaaa gtcacaaccg gtgcaacaac ctgaaaaagg 360
gtagagagcc ctcccgggca ataaaagacg gtaagaggaa atcccgaaaa gtgcaaaaag 420
gtcaaaacgg gtgcaacaaa atgccaaggc gtgcaagaaa ctgaggacgg ataaaagaca 480
atgacgacgg gtacgacacg ataaggaaag ataaggcga gtaaaagccg gtgcggcaaa 540
atacaacagg ctgccagaa gtaccgagcg ctaggcggga atgagcgagg gtaaaagaga 600
atacacaacc ctgcaagaaa gtccgggaag ctaggaccgg atacagagga ataaggaagg 660
gtacacagca atacgagcag atgaacgagc ataccagaga gatatcaggc ctaaaaaacg 720

ES 2 614 810 T3

gtaacaagaa ctaggaggag gtgccaaggc ctggaccaa atgggacgaa gtaggagggg 780
 atacacaacg atgggacgag atagagccgg gtaggcaaaa gtagcacggg ataccacaga 840
 atacagcgac gtaccggaaa gtaggagaga ctggaacccc gtaaggcaag ataagcgagg 900
 gctagcagaa ctaaacaga ctgaacacgg atcaaggcga gtaggcacga gtacacgccg 960
 atacggacaa gtaggccgca gtacggagaa gtaaacaggg atgagagaca gtgaacaagg 1020
 ctgagacgaa atcacagcaa atacaacaaa gtagcggcaa ataaacggga gtggaacggc 1080
 atagcaggaa gtgacggaca gcatgc 1106

5 <210> 30
 <211> 1169
 <212> ADN
 <213> Secuencia Artificial
 10 <220>
 <221> Fuente
 <223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"
 <400> 30

gaattcgagt accacctgat gagcagcaag gtaacggcga agatgaaaag aaaaatggag 60
 cggcactac aaaaaggat ggcaccagac atagaagagg agctcagcaa acggctggac 120
 agcgaagtcc caggaagcct accaaggaca gatatcaaag aaataagaaa caaactgaga 180
 aagcgcgtcg gaagcgacgt acaagcggca atggcagaga gaggtaaaaca aacagtgaaa 240
 ccaacaataa acgcccaggt agggcgaaag gtaaacaaaa ggatcgaaag acgagtaagg 300
 cgggagatgg aaacaacggt gcgaaaacca gtgaaaaagg gagtccaaag caccatgaag 360
 accacaatac gacacacct gggaggacac gtggaggagc cggtacgaag cagggtacaa 420
 gacaacatgc aaacagaggt aaaggaaccg ataaagaacg acatggaagg aaacgtagga 480
 acccgagtag aggaacggat agcgaacgac gtgcaacgcc acctagaacg agaagtcgca 540
 agggaaataa acggagaggt cccacggccg atgagcagaa aagtggaaaa ggacctgaag 600
 agaaaaatga ggaccacaat aaaggacgga ataaaaaaca cactcccaga ggaaatacaa 660
 cgagaaatag ggaacaggat gcaaagaagc atcagaaggg cagtccaaa gggaaatcga 720
 cgcaagctaa gacagccct agaccgcaa gtaaacgaac aactcggagg gcgaatcagg 780

ES 2 614 810 T3

cccaccatgc agagcaaat aagagaccag atgcgcgcca acctgagacc gaacctaaag 840
 gaaaggatag agggagaggt ccaagaaacg atggcgggca ccatgagaaa ggaagtcaa 900
 gtgcacctag caacacagat accacgagac atcgaagcga agatcggaa aaagctggag 960
 agggcaatga acagaaagct acgaaagaa gtcaaggaca caatgagga gcaagtcgag 1020
 agcgaggtgg cccgccaat aggagcaaca atgcgagggc aagtgggcac aacgatgaaa 1080
 gaccaactaa cacagccact tctagagtgc acggatcct cgaggtacc aacctggaa 1140
 gctagcaagc atgcaaactg cagaagctt 1169

5 <210> 31
 <211> 1106
 <212> ADN
 <213> Secuencia Artificial

10 <220>
 <221> Fuente
 <223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"
 <400> 31

ctcgaggaaa gacagagggt agcagggcac agatcaaggc ggcgaaatgg cccggagcac 60
 gtaagcacca cacggtcacc aagagagaat cagaagaggc aagtaaaaa aagcgcgtac 120
 ggcacccaag gtggacgcac aacagtcacg ggaagaagt cgaaggcggg cgctcagaag 180
 agcgaagtac acaaccagac ggtaccgga cacgatgcgc caacaagaat cagcggaaaag 240
 acataccggc ggacgggtcg aaacacgaaa atgaaaccgc agaagtaaag agaaaaggat 300
 cgcaggaaga ggatcaagac agaaaaatcg acaaagcaaa gtggaccgaa gaagatcgaa 360
 aagaagcact caggacagct gaatcgaca gacacggtga caagacgaaa atggacgaca 420
 accgataagc gaaaaaagat accggcgagc gagtcgagaa cacgaagtca aggcgcaaga 480
 atagagcaaa cgaagtaaac aggacagaat aggaaggcca aagtgggaga gaaaggctcg 540
 gaaccgacga ctgcaccgaa caacgtggga gcccacggct acaaacgcc agatcgaagc 600
 ccaggagtgc aagaacaaaa gtaacggcgc agaaatagag agaggaggct aaagcggaac 660
 cagtgaagga aaagaaataa cgggcccaag gtgaagaaaa agaagtggca ggaacggagt 720
 acagcccagc aactgggcaa aagcggctga cgagaacaag atagagaacg aggaatgacg 780

ES 2 614 810 T3

agacacagat agaagacaaa gactgctggac aggcaagtaa gagaaggaaa ataaagcggag 840
 agaagtaggc cgcccgaat gcaaaagaaa aagtgcacaca gagaaaatga ccggcgggag 900
 gtaacacgcc caggatgacg gagcaccaat agaaaaaac acgtaggaga gggaaacatgc 960
 ggacaagaca atcccaacc gaagataaac accgagaggt agacagcgca gactggagga 1020
 acagccgtgg aaagggaaag gtacgcaaac ggacctaaa aaaacaaagt cagggaaaaa 1080
 ggctaacgga acaaagctgg ccatgg 1106

5 <210> 32
 <211> 1241
 <212> ADN
 <213> Secuencia Artificial

10 <220>
 <221> Fuente
 <223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

<400> 32

gaattctcta gactgcacgg atccctcgag ggtaccaacc atggaagcta gcaagcatgc 60
 gaggaaagca aggaaaatga agaacgaagc gggctgaaca ggaagacat aaaaaacagg 120
 caccatgaca ccaacgcgga atgacacaac acaaaggtag cgaggcacgc acgtaagagg 180
 cacaagaaat gaaagaaaaa gagaatacgc agagagggca gtaaggcagg gaaacgatga 240
 gcagacgacg ggatgctagc agcgaaagct gagaagaaga gacctcaa aagcaaaaga 300
 ataagaaagc caaagcataa aagaaaagcg caatcaagga aacgggcgct agaagaccga 360
 gacagtgaga cgagcacacg ctaaaaaaga gagacaatcc agaaagaaga cagtggacaa 420
 gagaccgggt agcaccagag gacaatgaag gctggacggc ataaacgccc aaacgggtaa 480
 gacgggcaa cgataaagga agacgaaggt caacgcagaa ggaatccag aacacagac 540
 ataaagaaaa aaagaaatgg cagccggaga agctgagacg ggaccgcagt gcgagagaag 600
 agacgtgctg aaaagagaag gtcgcaag acgggctgg accggacaaa ggatgggaaa 660
 acgaccggat ggagagcgcc cagagtcaga aacggaagag atgaaacca gaaagactcg 720
 ccacaagaag gactcgggaa acacaagaat aaagcagctg ggaatagcg gacgaacgga 780
 atagaggccc aaagaactgg gagaggacag ccgtacagga ggaacgggat agcgaagaac 840

ES 2 614 810 T3

caacgtgcga aaagaccggc ataaaacaga aaggcaatag cagacagcaa aaatagagag 900
 acaacgacgt agaaagaggg aaacgtcgca aaaaaccggg ctagaggagg accaacatag 960
 cgagacacca agatgaagca acggagaagt cacacaaaga gacagtcgaa aagcaaagaa 1020
 atagcccagc aagcgaatac ccaacaagaa gagtcaacag gagaaagagt ccaaggagaa 1080
 gcaagtacac aagacaaaag gtagcagaca caagcaatgc acgggcaggg aaatgcgaaa 1140
 cagagaggat aagaagatct aaaatgcatg gtctgtgtga tgttggtctg tgtgatgtg 1200
 gtctgtgtga tgttggtctg tgtgatgttc tgcagaagct t 1241

5 <210> 33
 <211> 1106
 <212> ADN
 <213> Secuencia Artificial

10 <220>
 <221> Fuente
 <223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

<400> 33

ggatccgacg ggaacaaaga gaaagcaca gtaaaggca acgacgcgca cgaggtcaga 60
 acgcaaacag gaagaagaat gcaacaaagg agcggccaac ggatacggg gccaacacga 120
 agaggaatga aagggaaacg aaacaaaaa gtgaagaggc caaagagca cagcatggag 180
 gaagccacgg ccgaaaaact acaccgggaa ggaccaaacg cgatggcaga cagagcgag 240
 acagggatag agccaccag gaagcacgca gtcagggaga aaagggcaca gacactacga 300
 aaaagggcga caagagggct gggagagggg acgggcagaa acataaaacc caagagccga 360
 ccagaagtaa aaggcgatat caacacagac gtaaagcaag cagaacaaga gcacataagc 420
 aagaaggaga ggagacacct accagagaga aaaaccgaa ccgtgcgga aggcgggcga 480
 gaagggataa aagagcacga acccaagagg atgaaaggac aagagagaac gaaagtccaa 540
 cacagaaagg ccaaaggaat aaaaagcaac agccaaaaaa aagtaagaaa gcaagacaaa 600
 gcgagaatga aaagcgaagg cacaaaagaa atagacagag ggaacagcg aaagatccca 660
 cgggcgacag caaccggact ccacgggaac acgaggagaa acgtgaaaac cgagcggcaa 720
 cccagactca gcggaacaga agacaaaaaa atacgggcag ggcgaggaaa gaaaataggg 780

ES 2 614 810 T3

ccaacaaaa ggaagcagat gaaaaacaca aaaaccacc acatcacgcg ggccgagagc 840
agcaacgtgg cagcacgagg aggaggagcc ctagaacgca aaggcgaaaa gacgctgcga 900
ggtaccgggc caaagaaaat aaaggaacgc aaacaaagga gcatcaaagg gaagaacaa 960
aagaaaataa aaaaggagaa gaaaggaag ataaaaggaa acggaaaaaa gaaactcgcg 1020
acaagacgg cagcaaaggt agaggaaaag gaagaggaca cactgaaaga gcaccaggag 1080
aaagacctag cagaagcacg ctcgag 1106

REIVINDICACIONES

1. Una población de sondas de nanoindicador marcadas de forma única, en la que la sonda de nanoindicador comprende:

- 5
- i) una región específica de una diana única;
 - ii) una región que comprende un nanoindicador diseñado, único,

10 en la que dicho nanoindicador comprende una estructura principal de ácido nucleico monocatenario, diseñada de forma única, comprendiendo dicha estructura principal al menos tres regiones de unión al marcador unidas covalentemente entre sí en una combinación lineal única, en la que cada región de unión al marcador de cada estructura principal se selecciona individualmente y es diferente de las otras regiones de unión al marcador de esa misma estructura principal,

15 en la que las regiones de unión al marcador se seleccionan de una población de secuencias polinucleotídicas de diseño racional, en la que la secuencia polinucleotídica se hibrida a una secuencia polinucleotídica complementaria que tiene unidas a la misma una o más moléculas detectables;

20 en la que cada secuencia polinucleotídica complementaria tiene designada una molécula detectable específica; y en la que cada una de dichas regiones de unión al marcador tiene un contenido de G/C del aproximadamente 50 %, y en la que cada una de dichas secuencias polinucleotídicas complementarias tiene una proporción de G/C de al menos 1/1 u, opcionalmente, de aproximadamente 3/2, y

en la que cada nanoindicador tiene una señal detectable que lo distingue de otros nanoindicadores de dicha población.

25 2. La población de la reivindicación 1, en la que cada una de dichas regiones de unión al marcador se genera a partir de una secuencia molde seleccionada del grupo que consiste en SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23 y 24.

3. La población de la reivindicación 1 o 2,

30 (a) en la que cada una de dichas regiones de unión al marcador comprende de aproximadamente 800 a 1.300 bases de nucleótidos o

(b) en la que cada una de las regiones de unión al marcador comprende un contenido de adenina similar y, opcionalmente, en la que las bases de adenina están separadas al menos cada 8 a 16 bases de nucleótido, o

35 (c) en la que cada una de las regiones de unión al marcador comprende un patrón que se repite de forma regular de bases de adenina y, opcionalmente, en la que las bases de adenina están separadas aproximadamente cada 8 a 16 bases de nucleótido, o

(d) en la que cada una de las regiones de unión al marcador comprenden un contenido de timina de aproximadamente 35-45 % o

40 (e) en la que cada una de las secuencias polinucleotídicas complementarias comprende una secuencia polinucleotídica de ARN o

(f) en la que las moléculas detectables son colorantes fluorescentes o

(g) en la que una temperatura de fusión (T_f) de cada una de dichas secuencias polinucleotídicas complementarias cuando están hibridadas a sus regiones de unión al marcador es de aproximadamente 80 °C o superior, o

45 (h) en la que dicha sonda de nanoindicador comprende además una región constante, en la que la región constante comprende una pluralidad de secuencias de nucleótidos repetidas.

4. La población de la reivindicación 3, en la que la secuencia polinucleotídica complementaria comprende una secuencia polinucleotídica de ARN, y

50 (a) en la que la secuencia polinucleotídica de ARN comprende al menos una base de uracilo modificada con aminoalilo o

55 (b) en la que la secuencia polinucleotídica de ARN comprende una pluralidad de bases de uracilo modificadas con aminoalilo que están separadas una media de aproximadamente cada 8 a 16 bases en dicha secuencia polinucleotídica de ARN; y

en la que la molécula detectable está unida a la base de uracilo modificada con aminoalilo o en la que la molécula detectable está unida a cada una de las bases de uracilo modificadas con alilo.

60 5. La población de la reivindicación 1, en la que cuando dicha población de sondas de nanoindicador se usa para detectar una población de moléculas diana en una muestra mediante un método que comprende:

(a) formar al menos un complejo molecular con cada sonda de nanoindicador que comprende al menos una molécula diana y al menos una sonda de nanoindicador y

65 (b) contar individualmente la presencia de una o más moléculas de cada dicho complejo molecular de dicha muestra, el número de recuentos por encima del fondo de cada dicho complejo molecular tras la normalización

de la muestra es al menos dos veces superior en comparación con una sonda de nanoindicador que comprende ADN de M13.

6. Un método de determinación de la presencia de al menos una molécula diana en una muestra, que comprende:

(1) formar al menos un complejo molecular que comprende:

- (a) al menos una molécula diana y
- (b) al menos una sonda que comprende:

- i) una región específica de la diana única,
- ii) una región que comprende un nanoindicador diseñado, único,

en la que dicho nanoindicador comprende una estructura principal de ácido nucleico monocatenario, diseñada de forma única, comprendiendo dicha estructura principal al menos tres regiones de unión al marcador unidas covalentemente entre sí en una combinación lineal única, en la que cada región de unión al marcador seleccionada se selecciona individualmente y es diferente del resto de regiones de unión al marcador de la estructura principal seleccionadas, en la que cada región de unión al marcador se hibrida a una secuencia polinucleotídica complementaria que tiene unidas a la misma una o más moléculas detectables, en la que las regiones de unión al marcador se seleccionan de una población de secuencias polinucleotídicas de diseño racional, en la que cada secuencia polinucleotídica complementaria tiene designada una molécula detectable específica; y en la que cada región de unión al marcador tiene un contenido de G/C del aproximadamente 50 % y la secuencia polinucleotídica complementaria tiene una proporción de G/C de al menos 1/1, y

(2) contar individualmente la presencia de uno o más complejos moleculares o al menos parte de dicho al menos un complejo molecular para determinar la presencia de dicha al menos una molécula diana de la muestra.

7. El método de la reivindicación 6, en el que cada región de unión al marcador se genera a partir de una secuencia molde seleccionada del grupo que consiste en SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23 y 24.

8. El método de la reivindicación 6 o la reivindicación 7

- (a) en el que el porcentaje de recuentos moleculares válidos de dicho complejo molecular es superior al aproximadamente 12,5 % o
- (b) en el que el número de recuentos es al menos dos veces superior a los recuentos obtenidos cuando se usa una sonda de nanoindicador que comprende ADN de M13 y, opcionalmente, en el que dicho nanoindicador que comprende ADN de M13 comprende una estructura principal monocatenaria que comprende una pluralidad de regiones de ADN de M13 unidas covalentemente entre sí, hibridándose cada región a una secuencia polinucleotídica complementaria que tiene unidas a la misma una o más moléculas detectables o
- (c) en el que la temperatura de fusión (Tf) de dicha secuencia polinucleotídica complementaria, cuando se hibrida a su región de unión al marcador es de aproximadamente 80 °C o superior, o
- (d) en el que cada región de unión al marcador comprende de aproximadamente 800 a 1.300 bases de nucleótidos o
- (e) en el que las regiones de unión al marcador comprenden un contenido de adenina similar, y, opcionalmente, en el que las bases de adenina están separadas al menos una media de cada 8 a 16 bases de nucleótidos, o
- (f) en el que las regiones de unión al marcador comprenden un patrón que se repite de forma regular de bases de adenina y, opcionalmente, en el que las bases de adenina están separadas aproximadamente cada 8 a 16 bases de nucleótido, o
- (g) en el que cada una de las regiones de unión al marcador comprende un contenido de timina del aproximadamente 35 al 45 % o
- (h) en el que la secuencia polinucleotídica complementaria comprende una secuencia polinucleotídica de ARN o
- (i) en el que las moléculas detectables son colorantes fluorescentes o
- (j) en el que dicha molécula diana es un ácido nucleico o
- (k) en el que dicha molécula diana es un indicador de diagnóstico o
- (l) en el que dicha sonda comprende además una región constante, comprendiendo la región constante una pluralidad de secuencias de nucleótidos repetidas.

9. El método de la reivindicación 8, en el que una temperatura de fusión (Tf) de dicha secuencia polinucleotídica complementaria, cuando se hibrida a su región de unión al marcador, es superior a la temperatura de fusión de una secuencia polinucleotídica complementaria al ADN de M13 cuando se hibrida a dicha sonda de nanoindicador que comprende ADN de M13.

10. El método de la reivindicación 8, en el que la secuencia polinucleotídica complementaria comprende una secuencia polinucleotídica de ARN y
- 5 (a) en el que la secuencia polinucleotídica de ARN comprende al menos una base de uracilo modificada con aminoalilo o
 (b) en el que la secuencia polinucleotídica de ARN comprende una pluralidad de bases de uracilo modificadas con aminoalilo que están separadas al menos una media de cada 8 bases en dicha secuencia polinucleotídica de ARN; y
- 10 en el que la molécula detectable está unida a la base de uracilo modificada con aminoalilo o en el que la molécula detectable está unida a cada una de las bases de uracilo modificadas con alilo.
11. El método de la reivindicación 6 o la reivindicación 7, que comprende además determinar la presencia de una pluralidad de moléculas diana mediante un método que comprende formar una pluralidad de complejos moleculares, comprendiendo cada complejo
- 15 (a) al menos una molécula diana y
 (b) al menos una sonda que comprende:
- 20 i) una región específica de la diana única,
 ii) una región que comprende un nanoindicador diseñado, único,
- 25 en la que dicho nanoindicador comprende una estructura principal de ácido nucleico monocatenario, diseñada de forma única, comprendiendo dicha estructura principal al menos tres regiones de unión al marcador unidas covalentemente entre sí en una combinación lineal única,
 en la que cada región de unión al marcador seleccionada se selecciona individualmente y es diferente del resto de regiones de unión al marcador de la estructura principal seleccionadas,
 en la que las regiones de unión al marcador se seleccionan de una población de secuencias polinucleotídicas de diseño racional,
 30 en la que cada región de unión al marcador se hibrida a una secuencia polinucleotídica complementaria que tiene unidas a la misma una o más moléculas detectables y
 en la que cada sonda comprende una región de nanoindicador diferente y
 en la que cada región de unión al marcador tiene un contenido de G/C del aproximadamente 50 % y la secuencia polinucleotídica complementaria tiene una proporción de G/C de al menos 1/1.
- 35 12. El método de la reivindicación 11, en el que se determina la presencia de al menos 5, 10, 20, 30, 50, 100, 200, 300 o 500 moléculas diana diferentes.
- 40 13. El método de la reivindicación 6 o la reivindicación 7, en el que dicha molécula diana es un ácido nucleico y en el que dicho ácido nucleico comprende al menos una mutación heredable, al menos una mutación somática, al menos un polimorfismo de un solo nucleótido (SNP), al menos una mutación puntual, al menos una mutación de delección, al menos una mutación de inserción, al menos una translocalización cromosómica o combinaciones de las mismas.
- 45 14. Una sonda de nanoindicador marcada de forma única que comprende:
- 50 i) una región específica de la diana única y
 ii) una región de una estructura principal de ácido nucleico monocatenario, diseñada de forma única, que comprende al menos tres regiones diseñadas de unión al marcador unidas covalentemente entre sí en una combinación lineal,
- 55 en la que cada región de unión al marcador comprende de aproximadamente 800 a 1.300 bases de nucleótidos y tiene un contenido de G/C del aproximadamente 50 %,
 en la que cada región de unión al marcador de la sonda se selecciona individualmente y es diferente de las otras regiones de unión al marcador de la misma sonda,
 en la que las regiones de unión al marcador se seleccionan de una población de secuencias polinucleotídicas de diseño racional
 y
 en la que cada región de unión al marcador se hibrida a una secuencia polinucleotídica complementaria que tiene unidas a la misma una o más moléculas detectables y que tiene una proporción de G/C de al menos 1/1.
- 60 15. La sonda de nanoindicador de la reivindicación 14, en la que cada una de las al menos tres regiones de unión al marcador se genera a partir de una secuencia molde seleccionada del grupo que consiste en SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23 y 24.
- 65 16. La sonda de nanoindicador de la reivindicación 14 o la reivindicación 15,

- (a) en la que la secuencia polinucleotídica complementaria tiene una proporción de G/C de aproximadamente 3/2 o
- (b) en la que las regiones de unión al marcador comprenden un patrón que se repite de forma regular de bases de adenina y, opcionalmente, en la que las bases de adenina están separadas aproximadamente una media de cada 8 a 16 bases de nucleótidos o
- (c) en la que las regiones de unión al marcador comprenden un contenido de timina del aproximadamente 35-45 % o
- (d) en la que la secuencia polinucleotídica complementaria comprende una secuencia polinucleotídica de ARN o
- (e) en la que las moléculas detectables son colorantes fluorescentes o
- (f) en la que la sonda de nanoindicador comprende además una región constante, en la que la región constante comprende una pluralidad de secuencias de nucleótidos repetidas.
17. La sonda de nanoindicador de la reivindicación 14, en la que la secuencia polinucleotídica complementaria comprende una secuencia polinucleotídica de ARN y
- (a) en la que la secuencia polinucleotídica de ARN comprende al menos una base de uracilo modificada con aminoalilo o
- (b) en la que la secuencia polinucleotídica de ARN comprende una pluralidad de bases de uracilo modificadas con aminoalilo que están separadas al menos una media de cada 8 a 16 bases en dicha secuencia polinucleotídica de ARN y
- en la que la molécula detectable está unida a la base de uracilo modificada con aminoalilo o en la que la molécula detectable está unida a cada una de las bases de uracilo modificadas con alilo.
18. Un método de preparación de al menos un nanoindicador marcado de forma única que comprende una estructura principal de ácido nucleico monocatenario, diseñada de forma única, incluyendo el método las etapas de:
- i) combinar al menos tres regiones de unión al marcador, comprendiendo cada una de aproximadamente 800 a 1.300 bases de nucleótidos y un contenido de G/C del aproximadamente 50 %, en el que cada región de unión al marcador del nanoindicador se selecciona individualmente y es diferente de las otras regiones de unión al marcador del mismo nanoindicador;
- ii) unir covalentemente las al menos tres regiones de unión al marcador entre sí en combinación lineal; e
- iii) hibridar una secuencia polinucleotídica complementaria a dicha región de unión al marcador, en la que dicha secuencia polinucleotídica complementaria tiene unidas a la misma una o más moléculas detectables y en la que la secuencia polinucleotídica complementaria tiene una proporción de G/C de al menos 1/1; y
- en el que las regiones de unión al marcador se seleccionan de una población de secuencias polinucleotídicas de diseño racional.
19. El método de la reivindicación 18, que comprende además preparar una sonda de nanoindicador marcada mediante la unión del nanoindicador marcado a una región específica de la diana.
20. El método de la reivindicación 18, en el que dicha secuencia polinucleotídica complementaria tiene una proporción de G/C de al menos 1/1.
21. Un kit de preparación de al menos un nanoindicador marcado de forma única que comprende una estructura principal de ácido nucleico monocatenario, diseñada de forma única, incluyendo el kit:
- a) al menos tres regiones de unión al marcador, comprendiendo cada una de aproximadamente 800 a 1.300 bases de nucleótidos y un contenido de G/C del aproximadamente 50 %, en el que cada región de unión al marcador del nanoindicador se selecciona individualmente y es diferente de las otras regiones de unión al marcador del mismo nanoindicador; y
- b) al menos tres secuencias polinucleotídicas complementarias que tienen unida a las mismas una molécula detectable, teniendo las secuencias polinucleotídicas complementarias una proporción de G/C de al menos 1/1; y
- en el que las regiones de unión al marcador se seleccionan de una población de secuencias polinucleotídicas de diseño racional.
22. El kit de la reivindicación 21 que comprende además al menos tres sondas específicas de la diana.

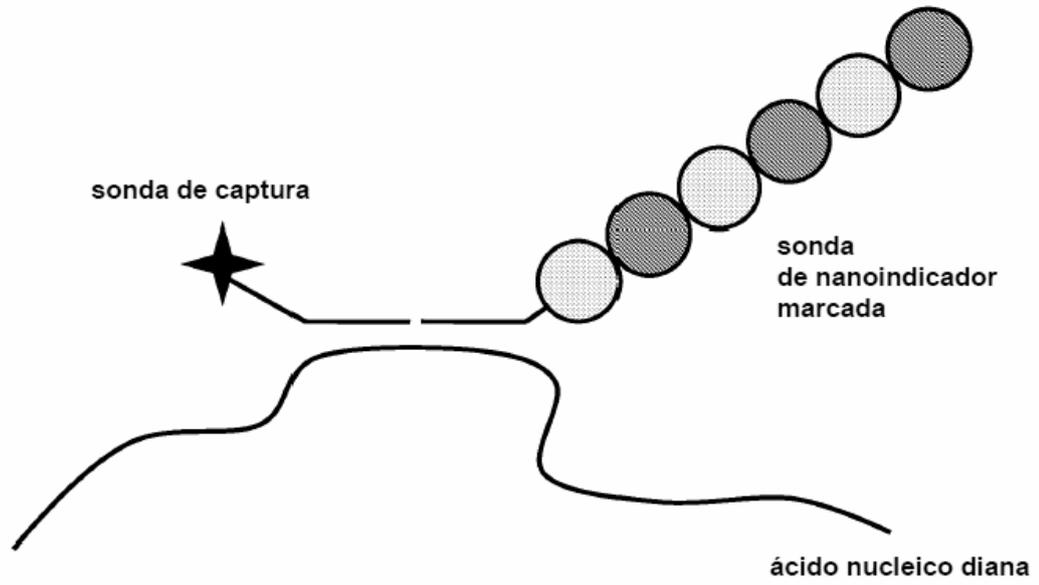


Figura 1

Figura 2A

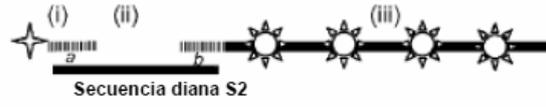


Figura 2E

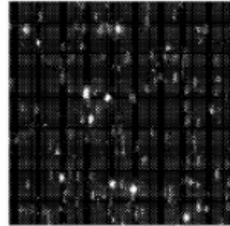
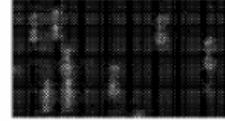


Figura 2B

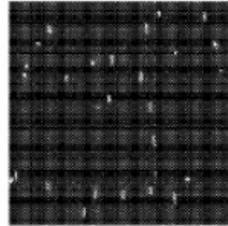


Figura 2C

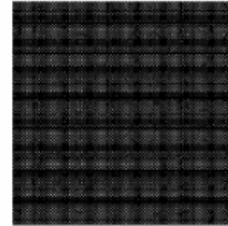


Figura 2D

Figura 2

Figura 3A

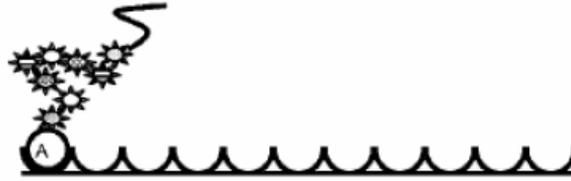


Figura 3B

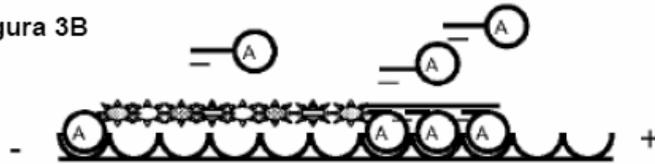


Figura 3

Figura 4A

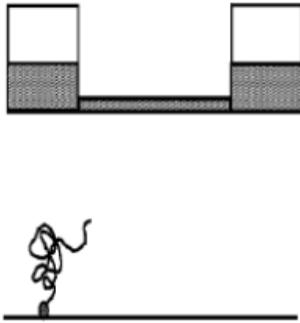


Figura 4B

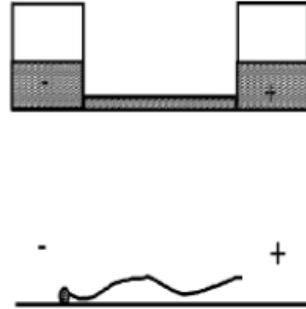


Figura 4C

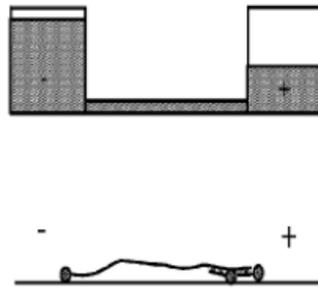


Figura 4

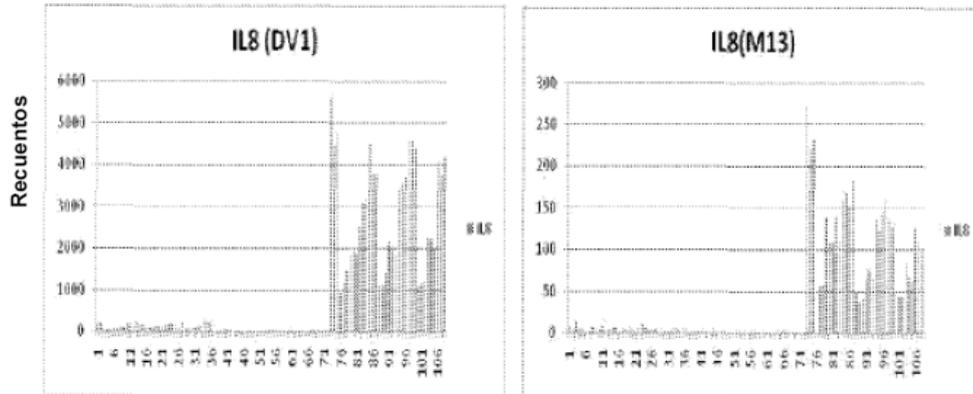
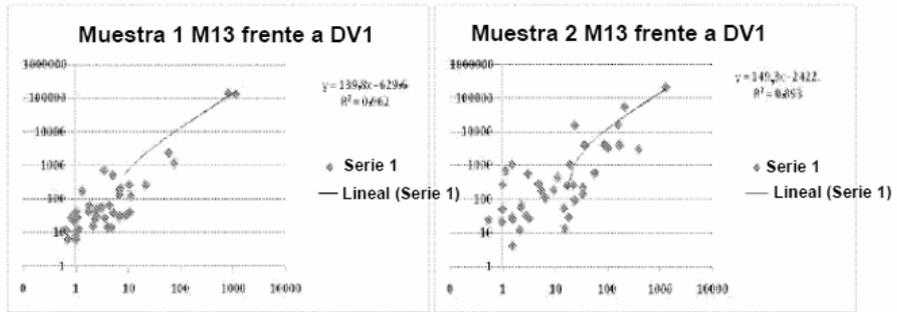


Figura 5

40 indicadores de 2 muestras "representativas":



Correlación del 89-96 % entre los dos sistemas, con un aumento superior a 100 veces en la señal con DV1

Figura 6

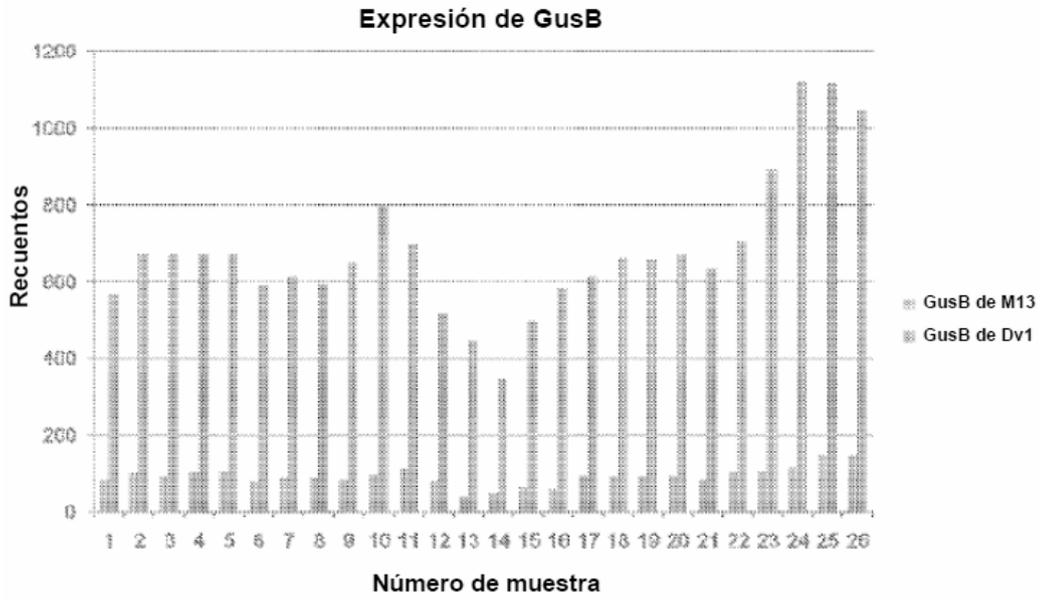


Figura 7

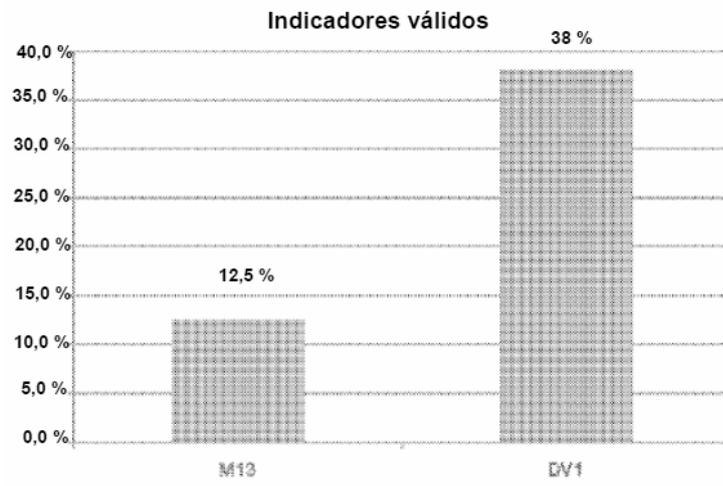


Figura 8

AAGCTTGGCCACACAGCTTGAGTATTCTATAGTGTACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGT
 TTCCTGTGTGAAATTGTTATCCGCTCACAATTCACACAAACATACGAGCCGGGAAGCATAAAGTGTAAAGCCT
 GGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTCCAGTCCGGAAACC
 TGTCTGCCCAGCTGCATTAATGAATCGGCCAACCGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTTTCCG
 CTTCCTCGCTCACTGACTCGCTGCGCTCGGTGCTTCCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGG
 TAATACGGTTATCCACAGAATCAGGGGATAACCGCAGGAAAGAACATGTGAGCAAAAGGCCAGAAAAGGCCA
 GGAACCGTAAAAAGGCCGCTTGCCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATC
 GACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCC
 TCGTGGCTCTCCTGTTCCGACCCTGCCGTTACCCGATACCTGTCCGCTTCTCCCTTCGGGAAGCGTGG
 CGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTGCTCCAAGCTGGGCTGTGTGC
 ACGAACCCCGCTTCCAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGAC
 ACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAG
 AGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCCCTCTGCTGAAGC
 CAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGTTTTT
 TTGTTTGAAGCAGCAGATTACGGCGAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGT
 CTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTACCT
 AGATCCTTTTAAATTAATAATGAAGTTTTAAATCAATCTAAAATATATATGAGTAAACTGGTCTGACAGTT
 ACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCCGTTCAATCCATAGTTGCCCTGACTCC
 CCGTGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATAACCGGAGACC
 CACGCTCACCGGCTCCAGATTTATCAGCAATAAACACCCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCTCTG
 CAATTTATCCGCTCCATCCAGTCTATTAATTGTTCCGGGAAGCTAGAGTAAGTAGTTCCGCACTAATA
 GTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTACGCTCGTCTTGGTATGGCTTCAATCA
 GCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCG
 GTCCTCCGATCGTTGTGAGAAGTAAGTTGGCCGAGTGTATCACTCATGGTTATGGCAGCACTGCATAATT
 CTCTTACTGTATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGTAGTACTCAACCAAGTCACTTCTGAGAAT
 AGTGTATGCGCGCAGCCGAGTTGCTCTTCCCGCGCTCAATACGGGATAATACCGCGCCACATAGCAGAACTT
 TAAAAGTGCTCATCATTGGAAAACGTTCTTCCGGGGCGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCCA
 GTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCCAGCATCTTTACTTTACCAGCGTTTCTGGGTGAG
 CAAAAACAGGAAGGCAAAATGCCGAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCT
 TCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATAATTTGAATGTATTT
 AGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTA
 TTATCATGACATTAACCTATAAAAAATAGGCGTATCACGAGGCCCTTTCGTCTCGCGCGTTTCGGTGTGACG
 GTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGAC
 AAGCCGTCAGGGCGCGTACGCGGTGTTGGCGGGTGTGGGGCTGCTTAACTATGCGGCATCAGAGCAGA
 TTGTACTGAGAGTGACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGA
 AATGTAAGCGTTAATATTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAAACCAATA
 GGCCGAAATCGGCAAAATCCCTTATAAATCAAAGAATAGACCGAGATAGGGTTGAGTGTGTTCCAGTTTG
 GAACAAGAGTCCACTATTAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGG
 CCCACTACGTGAACCATCACCTAATCAAGTTTTTGGGGTTCGAGGTGCCGTAAAGCACTAAATCGGAACCC
 TAAAGGAGCCCGGATTTAGAGCTTGACGGGGAAAGCCGGCAACGTGGCGAGAAAGGAAGGGAAGAAAGC
 GAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACGCTGCGCGTAACCACACCCCGCCGCGCT
 TAATGCGCCGCTACAGGGCGGTCCATTGCGCATTGAGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGG
 GCCTCTCGCTATTACGCCAGCTGGCGAAAGGGGATGTGCTGCAAGCGGATTAAGTTGGGTAACGCCAGGG
 TTTTCCAGTACGACGTTGTAACGACGGCCAGTGAATTGTAATACGACTCACTATAGGGCGAATTAGTG
 GATGAGAAGACCTGCAGAAAGAATTC (SEQ ID NO:27)

Figura 9

Espacio de 8

TCTAGAGAGAAAGGCTACGAAGATCCGGACATCAAAGATCAAACATGCGGAGATCGAC
ACATCCAAAATGAACAGATAACAGGATCAGAGGGTCAAGAAAGTGAACCATCGAGGAGT
CGCAGACTACGAAGATGGACAAGTGAGCAACAGCTGCGATGGAGAACTAAAGGAGTCGGG
AGGTACGGCAGTGAAGAAATGTGCACATGAAAAGCTGCCCCGGTGGCCTAACTGGAGGACT
AGAAGAATGGCACAGTGACCCGGTGAGAGCCTGAGGAGGTGCGAGAATAGGGCGATAGGG
AGGTAACCCGATGGCGGACTAAAGAGATACGGACATGGACGGATAGGAGAGTCCGAAGAT
GGAAGGATCACAAGTGAAGCCGTTCAAGCATAACGGGTTAAACGGATCAGAACGTAAAG
CGGTGGAAAAATACCCGAGTAGGAAGATAGCAAAAATACCCAGGTGAGCGGATAGCCACT
CGGCGAGTCCAAAATGAAAAATGGAGAGATGAGACACTAGCGGCATCGGACAGTAGAG
CAGTGAACGCATCAGAGAGTAGAAAAGTGAAGAAGTGCACGGTGAACCATAACAGAGT
CAGAGGGTACACCCATACGACCCACCGCAATACAAGGCTAAGACGGTCAGAGGATGGAC
GGCTCGCGAACTCACGCCGTGCCAAAGTGAAGAAGTACCAAAATAGGAGCGTAACGGCGT
AGAGAGGTAGACGAGTGAGACGGTGGGCAGCTCCCAAGTCCGAGCATGGACAGTGGCA
AAATCCAGCGGTACGAACTACGAGACTGGCCGGATGAAGAAAATAAGGCGCTAGAGAGAT
ACGGGAGTGACAAGATGGACCACAGCGAAGTGGAGGGTGAAGACGTGAACAACAGAC
CAGTACGGCACTAGCAGAATAGACCGATGAAAACCTGAAAAGATAACAGAGTGAAGGCAT
CGGAGGGTGGAGAAGTAAGAAGTGGAGCCATACCGAATAAAGCAGTACAAGGGTGGAC
AAATGAGGAAATCGGGAGGTCCACGGTGAACCGTAGGGAGGTGAACCAGTGGGAGAGT
CAACCAATACGAAGATGCGAGGATCC (SEQ ID NO:28)

Figura 10A

Espacio de 10

CCATGGGAAAATGAGAGCCGCTCGACAAGGATAAAGAAAAGTCCGCCAACATCCGCCACG
CTAAGCGGCCCTAGGAGCCGGTAAACAAGCATGAAGAAACATGCGAGGCAATGACACAAA
CTCCAGGCGGGTGGAAAAGACTCGACGGAGATGCCACGCAATAGCGAAACCTGGAAAGAGA
GTGGAAGCAATAGGAAACAGTAGAGCAAAAATGCCCGCAGGATAGAAAACAGTGGCGCGCA
ATAAGCAGGAATAAAGCACAATAAAAAGAAGTAAACCGCACTAAAACAGCATCCAGACCG
CTGAAAAGAAATGGCGGCGAATGAAAAAAGTCAACCCGGTGAACAACCTGAAAAAGG
GTAGAGAGCCCTCCCGGGCAATAAAAGACGGTAAGAGGAAATCCCGAAAAGTGCAAAAAG
GTCAAACGGGTGCAACAAAATGCCAAGGCGTGCAAGAACTGAGGACGGATAAAAAGACA
ATGCGGACGGGTACGACACGATAAGGAAAGATAAGGCGAAGTAAAAGCCGGTGGCGCAAA
ATACAACAGGCTGCCCAGAAGTACCGAGCGCTAGGCGGGAATGAGCGAGGGTAAAAGAGA
ATACACAACCCTGCAAGAAAGTCCGGGAAGCTAGGACCGGATACAGAGGAATAAGGAAGG
GTACACAGCAATACGAGCAGATGAACGAGCATAACCAGAGAGATATCAGGCCTAAAAACG
GTAACAAGAAGTGGAGGAGGTGCGAAGGCCTGGACCAAAATGGGACGAAGTAGGAGGGA
ATACACAACGATGGGACGAGATAGAGCCGGTAGGCCAAAGTAGCACGGGATACCACAGA
ATACAGCGACGTACCGGAAAGTAGGAGAGACTGGAACCCGTAAGGCAAGATAAGCGAGG
GCTAGCAGAACTAAAACAGACTGAACACGGATCAAGGCGAGTAGGCACGAGTACACGCCG
ATACGGACAAGTAGGCCGAGTACGGAGAAGTAAACAGGGATGAGAGACAGTGAACAAGG
CTGAGACGAAATCACAGCAAAATACAACAAAGTAGCGGCAAAATAAACGGGAGTGGAAACGCG
ATAGCAGGAAGTGACGGACAGCATGC (SEQ ID NO:29)

Figura 10B

Figura 10

Espacio de 12

GAATTCGAGTACCACCTGATGAGCAGCAAGGTAACGGCGAAGATGAAAAGAAAAATGGA
 CGGCACGTACAAAAAGGATGGCACCAGACATAGAAGAGGAGCTCAGCAAACGGCTGGA
 AGCGAAGTCCCAGGAAGCCTACCAAGGACAGATATCAAAGAAATAAGAAAACAACTGAC
 AAGCGCGTCGGAAGCGACGTACAAGCGGCAATGGCAGAGAGAGTAAAACAAACAGTGAA
 CCAACAATAAACGCCAGGTAGGGCGAAAGGTAACAAAAGGATCGAAAGACGAGTAAC
 CGGGAGATGGAAACACGGTGGCAAACCAGTGAAAAGGGAGTCCAAAGCACCATGAA
 ACCACAATACGACACACCCCTGGGAGGACACGTGGAGGAGCCGGTACGAAGCAGGGTACA
 GACAACATGCAAACAGAGGTAAGGAACCGATAAAGAACGACATGGAAGGAAACGTAGC
 ACCCGAGTAGAGGAACGGATAGCGAACGACGTGCAACGCCACCTAGAACGAGAAGTCGC
 AGGGAAATAAACGGAGAGGTCCCACGGCCGATGAGCAGAAAAGTGGAAAAGGACCTGAA
 AGAAAAATGAGGACCACAATAAAGGACGGAATAAAAAACACACTCCCAGAGGAAATACA
 CGAGAAATAGGGAACAGGATGCAAAGAAGCATCAGAAAGGCAGTGCCAAAGGGAATCGC
 CGCAAGCTAAGACAGGCCCTAGACCAGCAAGTAAACGAACAACCTCGGAGGGCGAATCAG
 CCCACCATGCAGAGCAAAAATAAGAGACCAGATGCGCGGCAACCTGAGACCGAACCTAAA
 GAAAGGATAGAGGGAGAGGTCCAAGAAACGATGGCGGGCACCATGAGAAAGGAAGTCAA
 GTGCACCTAGCAACACAGATACCACGAGACATCGAAGCGAAGATACGGAAAAAGCTGGA
 AGGGCAATGAACAGAAAGCTACGGAAAGAAGTCAAGGACACAATGAGGGAGCAAGTCGA
 AGCGAGGTGGCCCGCAAATAGGAGCAACAATGCGAGGCCAAGTGGGCACAACGATGAA
 GACCAACTAACACAGCCACTTCTAGAGTGCACGGATCCCTCGAGGGTACCAACCATGGA
 GCTAGCAAGCATGCAAACCTGCAGAAGCTT (SEQ ID NO:30)

Figura 11A

Espacio de 14

CTCGAGGAAAGACAGAGGGTAGCAGGCCACAGATCAAGGCGGCGAAATGGCCCGGAGCA
 GTAAGCACCACACGGTCACCAAGAGAGAATCAGAAGAGGCAAGTAAAAAAAAAGCGCGTA
 GGCACCCAAGGTGGACGCACAACAGTCACGGGAAAGAAGTCAAGGCGGGCGCTCAGAA
 AGCGAAGTACACAACCAGACGGTACCGGAAACACGATGCGCCAACAAGAATCAGCGGAAA
 ACATACCGGCGGACGGTTCGAAACACGAAAATGAAACCGCAGAAGTAAAGAGAAAAAGGA
 CGCAGGAAGAGGATCAAGACAGAAAAATCGACAAAGCAAAGTGGACCGAAGAAGATCGA
 AAGAAGCACTCAGGACAGCTGAATCGCACAGACACGGTGACAAGACGAAAATGGACGAC
 ACCGATAAGCGAAAAAAGATACCGGCGAGCGAGTCGAGAACACGAAGTCAAGGCGCAAG
 ATAGAGCAAACGAAGTAAACAGGACAGAATAGGAAGGGCAAAGTGGGAGAGAAAGGCTC
 GAACCGACGACTGCACCGAACAACTGGGAGCCCACGGCTACAAAACGCCAGATCGAAG
 CCAGGAGTCGAAGAACAAGTAACGGCGCAGAAATAGAGAGAGGAGGCTAAAGCGGAA
 CAGTGAAGGAAAAGAAATAACGGGCCAAGGTGAAGAAAAAGAAGTGGCAGGAACGGAC
 ACAGCCCAGCAACTGGGCAAAAGCGGCTGACGAGAACAAGATAGAGAACGAGGAATGAC
 AGACACAGATAGAAGACAAAGACTGCGGACAGGCAAGTAAGAGAAGGAAAAATAAGCGA
 AGAAGTAGGCCGCCGAAATGCAAAAGAAAAAGTACACAGAGAAAAATGACCGGCGGGA
 GTAACACGCCCAGGATGACGGAGCACCAATAGAAAAAACACGTAGGAGAGGGAAACATC
 GGACAAGACAATCCCAACCCGAAGATAAACACCGAGAGGTAGACAGCGCAGACTGGAGC
 ACAGCCGTGGAAAGGGAAAGGTACGCAAACGGACCTAAAAAAAACAAAGTCAGGGAAAA
 GGCTAACGGAACAAAGCTGGCCATGG (SEQ ID NO:31)

Figura 11B

Figura 11

Espacio de 16

GAATTCCTCTAGAGTGCACGGATCCCTCGAGGGTACCAACCATGGAAGCTAGCAAGCATGC
GAGGAAAGCAAGGAAAATGAAGAACGAAGCGGGTCGAACAGGGAAGACATAAAAAACAGG
CACCATGACACCAACCGGGAATGACACACACAAAGGTAGCGAGGCACGCACGTAAAGAGG
CACAAAGAAATGAAAGAAAAAGAGAAATACGCAGAGAGGGCAGTAAAGCCAGGGAACCGATGA
GCAGACGCAGGGATGCTAGCAGCGAAAAGCTGAGAAGAAGAGACCCTCAAAAAGCAAAAGA
ATAAGAAAGCCAAAGCATAAAAGAAAAGCGCAATCAAGGAAACGGCGCTAGAAAGACCGA
GACAGTGAAGACGAGCACACGCTAAAAAGAGAGACAATCCAGAAAAGACAGTGGACAA
GAGACCGGGTAGCACAGAGGACAAATGAAGCGGGACCGCATAAACGCCAAACCGGGTAA
GACGGGCCAACGATAAAGGAAGACGAAGGTCAACGCAGAAGGGAATCCAGAAAACACAGAC
ATAAGCAAAAAGAAATGCCAGCCGGAGAAGCTGAGACGGGACCAGTGCAGAGAGAAAG
AGACGTGCCGAAAAGAGAAGGTCCGCGCAAGACGGGCGTGGACCCGACAAAGGATGGGAAA
ACGACCCGGATGGAGAGCCGCCAGAGTCAGAAAACGGAAGAGATGAAACCCAGAAAAGACTCG
CCACAAGAAGGAGTCCGGAAAACACAAGAATAAAGCAGCTGGGGAATAGCGGACGAAACCGGA
ATAGCCGCCAAAGAACTGGGAGAGGACAGCCGTACAGGAGAAACGGGATAGCGGAAGAAC
CAACGTGCCAAAAGACCGGCATAAAACAGAAAGGCAATAGCAGACAGCAAAAATAGAGAG
ACAACGACGTAGAAAGAGGGAAACGTCGCAAAAACCGGGCTAGAGGAGGACCAACATAG
CGAGACACCAAGATGAAGCAACGGAGAAGTCACACAAAGAGACAGTCCGAAAAGCAAAGAA
ATAAGCCAGCAAGCGAATACCCACAAGAAAGAGTCAACAGGAGAAAAGATCCAAAGGAGAA
GCAAGTACACAAGACAAAAGGTAGCAGACACAAGCAATGCACGGGCAGGGAATGCGAAA
CAGAGAGGATAAGAAGATCTAAAAATGCATGGTCTGTGTGATGTTGGTCTGTGTGATGTTG
GTCTGTGTGATGTTGGTCTGTGTGATGTTCTGCAGAAGCTT (SEQ ID NO:32)

Figura 12A

Espacio de 24

GGATCCGACGGGAACAAAGAGAAAAGCACAAAGTAAAAGGCAACGACGCGCACGAGGTCAGA
ACGCAAAACAGGAAGAAAGAAATGCAACAAAGGAGCGCCGAACGGATACGCGGGCCAAACAG
AGAGGAATGAAAGGGAAACGAAACAAAAAGTGAAGAGGCCAAAAGAGCACAGCATGGAG
GAAGCCACGGCCGAAAAACTACACCGGGAAAGGACCAAACGCGATGGCAGACAGAGCGCAG
ACAGGGATAGAGCCACCCAGGAAGCACGCGAGTCAGGGAGAAAAGGGCACAGACTACGA
AAAAGGGCGACAAGAGGGCTGGGAGAGGGAAACGGGCAGAAACATAAAAACCCAAAGAGCCGA
CCAGAAGTAAAAGGCGATATCAACACAGACGTAAGCAAGCAGAAACAAAGAGCACATAAGC
AAGAAGGAGAGGAGACACCTACCAGAGAGAAAACCCGAACCGTGCAGGAAAGGCGGGCGA
GAAGGGATAAAAGAGCACGAACCCAAAGAGGATGAAAGGACAAGAGAGAACGAAAGTCCAA
CACAGAAAGGCCAAAGGAATAAAAAGCAACAGCCAAAAAAAGTAAAGAAAGCAAGACAAA
CGGAGAAATGAAAAGCGAAGGCACAAAAGAAATAGACAGAGGGAACAGCGGAAAGATCCCA
CGGGCCAGCAGCAACCGGACTCCAGGGAAACAGAGGAGAAACGTGAAAACCGAGCGGCA
CCCAGACTCAGCGGAACAGAAAGCAAAAAATACGGGCAGGGCAGGAAAGAAAATAGGG
CCAACAAAAGGAAGCAGATGAAAACACAAAACCCACCACATCACCGGGCCGAGAGC
AGCAACGTGGCAGCACGAGGAGGAGCCCTAGAACGCAAGGCGAAAAGACGCTGCCA
GGTACCGGGCCAAAGAAAATAAAGGAACGCAAAACAAAGGAGCATCAAAGGGAAGAAACAA
AAGAAAATAAAAAGGAGAAGAAAGGGAAGATAAAGGAAACGGAAAAAAGAACTCGCG
ACAAAGACCGCAGCAAGGTAGAGGAAAAGGAAGAGGACACACTGAAAAGAGCACCCAGGAG
AAAGACCTAGCAGAAGCACGCTCGAG (SEQ ID NO:33)

Figura 12B

Figura 12