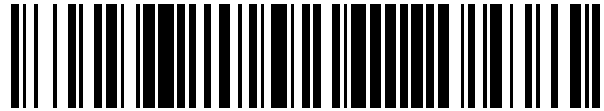


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 684 317**

51 Int. Cl.:

C12Q 1/68

(2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **13.08.2009 E 16194404 (6)**

97 Fecha y número de publicación de la concesión europea: **18.07.2018 EP 3162900**

54 Título: **Nanoinformadores estables**

30 Prioridad:

14.08.2008 US 88988 P

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

02.10.2018

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:

SÁEZ MAESO, Ana

ES 2 684 317 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

Nanoinformadores estables

Referencia cruzada

Antecedentes de la invención

- 5 Aunque todas las células en el cuerpo humano contienen el mismo material genético, los mismos genes no están activos 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 pueden ser el núcleo de procesos fisiológicos y patológicos alterados. Por lo tanto, identificar y cuantificar la expresión de genes en las células puede ayudar al descubrimiento de nuevos objetivos terapéuticos y de diagnóstico.
- 10 Hasta la fecha, existen varias técnicas disponibles que permiten la detección del nivel de expresión de múltiples genes en una muestra compleja a la vez. La mayoría de estas tecnologías emplean chips de ADN, dispositivos que consisten en miles de secuencias de ADN inmovilizadas presentes en una superficie miniaturizada que han hecho que este proceso sea más eficiente. Desafortunadamente, a pesar de la miniaturización de los formatos de chips, este método todavía requiere cantidades significativas de la muestra biológica. Sin embargo, en varios casos, tales como biopsias de tejidos enfermos o muestras de un tipo celular particular, la muestra biológica tiene un suministro limitado. Además, la cinética de hibridación en la superficie de un chip es menos eficiente que la hibridación en pequeñas cantidades de solución acuosa. Además, aunque existen métodos para estimar la cantidad de ácido nucleico presente en una muestra basada en el resultado de hibridación de chips, la tecnología de chips hasta ahora no permite la detección de moléculas objetivo a nivel individual, ni hay métodos basados en chips para cuantificar directamente la cantidad de molécula objetivo en una muestra dada.

- 20 Fortina P et al. "Digital mRNA Profiling" Nature Biotechnology, vol. 25, no. 3, 1 de marzo de 2008 presenta y resume los hallazgos de Geiss et al. Nature Biotechnology, vol. 26, no. 3, marzo de 2008. Geiss et al., describen el sistema de expresión génica NanoString nCounter®, que consiste en ADN de M13 de cadena sencilla linealizado, denominado cadena principal. Los documentos WO-A2-2007/139766 y WO-A2-2007/076128 proporcionan datos experimentales relacionados con nanoinformadores que tienen una cadena principal basada en ADN de M13 linealizado.

- 25 Por lo tanto, existe una necesidad de detección, identificación y cuantificación precisa y sensible de moléculas objetivo en mezclas complejas.

Sumario de la invención

- 30 Esta divulgación se refiere generalmente al campo de la detección, identificación y cuantificación de moléculas objetivo en una muestra e incluye nanoinformadores mejorados y estables, y poblaciones de estos nanoinformadores, con base en parte en ciertas características de diseño en las secuencias de polinucleótidos de las cadenas principales de nanoinformadores y secuencias complementarias de polinucleótidos unidas a las mismas.
- 35 En un primer aspecto de la invención, se proporciona una composición que comprende
- (a) una primera sonda nanoinformadora etiquetada de manera única que comprende: (i) una región específica de un objetivo único; (ii) una región que comprende un único nanoinformador designado en donde dicho nanoinformador comprende una cadena principal de ácido nucleico de diseño único de cadena sencilla, comprendiendo dicha cadena principal una pluralidad de regiones de fijación de la etiqueta diseñadas covalentemente unidas entre sí en una combinación lineal única, en donde cada región de fijación de la etiqueta de cada cadena principal se selecciona individualmente y es diferente de las otras regiones de fijación de la etiqueta en esa misma cadena principal, donde las regiones de fijación de la etiqueta se seleccionan de una población de secuencias de polinucleótidos diseñadas racionalmente, en donde la secuencia de polinucleótidos se hibrida con una secuencia complementaria de polinucleótidos que tiene unida a la misma una o más moléculas detectables, en donde cada región de fijación de la etiqueta tiene un contenido de G/C de aproximadamente 50%, y la secuencia complementaria de polinucleótidos tiene una relación de G/C de al menos aproximadamente 1/1 o aproximadamente 3/2, en donde cada nanoinformador tiene una señal detectable que la distingue de otros nanoinformadores en dicha composición; y (b) una segunda sonda que comprende (i) una región que se une selectivamente a la molécula objetivo y en un sitio diferente de la región específica objetivo única de la primera sonda; y (ii) una etiqueta de afinidad.
- 40
- 45
- 50 En un segundo aspecto de la invención, se proporciona un método para determinar la presencia de al menos una molécula objetivo en una muestra, que comprende: (1) formar al menos un complejo molecular que comprende: (a) al menos una molécula objetivo; (b) al menos una sonda que comprende i) una región específica objetivo única; y ii) una región que comprende un nanoinformador diseñado único, en donde dicho nanoinformador comprende una cadena principal de ácido nucleico de diseño único de cadena sencilla, comprendiendo dicha cadena principal una pluralidad de regiones de fijación de la etiqueta unidas covalentemente entre sí en una combinación lineal única; en donde cada región de fijación de la etiqueta se selecciona individualmente y es diferente de las otras regiones de
- 55

unión seleccionadas de la etiqueta en la cadena principal; en donde cada región de fijación de la etiqueta se hibrida con una secuencia complementaria de polinucleótidos que tiene unidas a la misma una o más moléculas detectables; en donde las regiones de fijación de la etiqueta se seleccionan de una población de secuencias de polinucleótidos diseñadas racionalmente; en donde cada secuencia complementaria de polinucleótidos ha designado una molécula detectable específica; en donde cada región de fijación de la etiqueta tiene un contenido de G/C de aproximadamente 50%, y la secuencia complementaria de polinucleótidos tiene una relación de G/C de al menos aproximadamente 1/1; y (c) una segunda sonda que comprende: (i) una región que se une selectivamente a la molécula objetivo y en un sitio diferente de la región específica objetivo única de la primera sonda; y (ii) una etiqueta de afinidad; 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 objetivo en la muestra.

Las realizaciones preferidas de la invención son cualquiera de sus diversos aspectos que se describen a continuación o como se define en las reivindicaciones dependientes.

También se proporciona una población de sondas nanoinformadoras etiquetadas de forma única, donde cada sonda nanoinformadora comprende: i) una región específica objetivo única; y ii) una región que comprende un nanoinformador diseñado único, donde el nanoinformador comprende una cadena principal de ácido nucleico de cadena sencilla, comprendiendo la cadena principal una pluralidad de regiones de fijación de la etiqueta unidas covalentemente entre sí en una combinación lineal única, donde las regiones de fijación de la etiqueta se seleccionan de una población de secuencias de polinucleótidos diseñadas, donde cada secuencia de polinucleótidos se hibrida con una secuencia complementaria de polinucleótidos que tiene unidas a la misma una o más moléculas detectables; donde cada secuencia complementaria de polinucleótidos ha designado una molécula detectable específica; y donde cada nanoinformador tiene una señal detectable que lo distingue de otros nanoinformadores en dicha población. Cada región de fijación de la etiqueta de cada cadena principal puede ser diferente de las otras regiones de fijación de la etiqueta en esa misma cadena principal. Cada una de las sondas nanoinformadoras puede comprender además una región constante, en donde la región constante comprende una pluralidad de secuencias repetidas de nucleótidos. La población de sondas nanoinformadoras puede comprender dos o más sondas nanoinformadoras.

También se proporcionan métodos para determinar la presencia de al menos una molécula objetivo en una muestra, que comprende: (1) formar al menos un complejo molecular que comprende: (a) al menos una molécula objetivo, y (b) al menos una sonda que comprende una región específica objetivo única y una región que comprende un nanoinformador diseñado único, en donde dicho nanoinformador comprende una cadena principal de ácido nucleico de cadena sencilla, comprendiendo la cadena principal una pluralidad de regiones de fijación de la etiqueta unidas covalentemente entre sí en una combinación lineal única, en donde cada región de fijación de la etiqueta se hibrida con una secuencia complementaria de polinucleótidos que tiene fijadas a la misma una o más moléculas detectables, en donde cada secuencia complementaria de polinucleótidos ha designado una molécula detectable específica; e individualmente contar la presencia de uno o más complejos moleculares o al menos parte de al menos un complejo molecular para determinar la presencia de dicha al menos una molécula objetivo en la muestra. El porcentaje de conteos moleculares válidos del complejo molecular puede ser mayor que aproximadamente 12,5%. Los números de conteos pueden ser al menos dos veces superiores a los conteos que se obtienen cuando se usa una sonda nanoinformadora que comprende ADN de M13. Un nanoinformador que comprende ADN de M13 puede comprender una cadena principal de cadena sencilla que comprende una pluralidad de regiones de ADN de M13 unidas covalentemente entre sí, en donde cada región se hibrida con una secuencia complementaria de polinucleótidos que tiene fijadas a la misma una o más moléculas detectables.

En los métodos y la composición, los números de conteos pueden estar por encima del valor base de cada complejo molecular después de que la normalización de la muestra sea al menos dos veces mayor cuando se compara con una sonda nanoinformadora que comprende ADN de M13. La T_m de las secuencias complementarias de polinucleótidos cuando se hibridan con sus regiones de fijación de la etiqueta puede ser de aproximadamente 80°C o más. La temperatura de fusión (T_m) de las secuencias complementarias de polinucleótidos cuando se hibrida con sus regiones de fijación de la etiqueta puede ser mayor que la T_m de secuencias complementarias de polinucleótidos al ADN de M13 cuando se hibrida con la sonda nanoinformadora que comprende ADN de M13.

Los métodos pueden comprender además la determinación de la presencia de una pluralidad de moléculas objetivo mediante un método que comprende formar complejos moleculares de pluralidad, comprendiendo cada complejo (a) al menos una molécula objetivo y (b) al menos una sonda que comprende un objetivo único región específica y una región que comprende un nanoinformador diseñado único, donde cada nanoinformador comprende una cadena principal de ácido nucleico de cadena sencilla, comprendiendo la cadena principal una pluralidad de regiones de fijación de la etiqueta unidas covalentemente entre sí en una combinación lineal única, en donde cada región de fijación de la etiqueta se hibrida con una secuencia complementaria de polinucleótidos que tiene unidas a la misma una o más moléculas detectables, y en donde cada sonda comprende una región nanoinformadora diferente. Preferiblemente, se determina la presencia de al menos 5, 10, 20, 30, 50, 100, 200, 300 o 500 moléculas objetivo diferentes. Preferiblemente, la molécula objetivo es un ácido nucleico. Preferiblemente, el ácido nucleico comprende al menos una mutación hereditaria, al menos una mutación somática, al menos un polimorfismo de un solo nucleótido (SNP), al menos una mutación puntual, al menos una mutación por eliminación, al menos una mutación

por inserción, al menos una translocación cromosómica, o combinaciones de los mismos. Preferiblemente, la molécula objetivo es un indicador de diagnóstico.

También se proporciona una sonda nanoinformadora etiquetada de forma única que comprende: i) una región específica objetivo única; y ii) una región que comprende una pluralidad de regiones de unión diseñadas de la etiqueta covalentemente unidas entre sí en una combinación lineal, donde cada región de fijación de la etiqueta comprende aproximadamente 800 a 1.300 bases nucleotídicas y tiene un contenido de G/C de aproximadamente 50%, en donde cada región de fijación de la etiqueta seleccionada es diferente de las otras regiones de unión seleccionadas de la etiqueta, y en donde cada región de fijación de la etiqueta se hibrida con una secuencia complementaria de polinucleótidos que tiene unidas una o más moléculas detectables, en donde la secuencia complementaria de polinucleótidos tiene una relación de G/C de al menos 1/1. La secuencia complementaria de polinucleótidos puede tener una relación de G/C de aproximadamente 3/2. La sonda nanoinformadora puede comprender además una región constante, donde la región constante comprende una pluralidad de secuencias repetidas de nucleótidos.

En los métodos y composiciones descritos en la presente memoria, la región de fijación de la etiqueta puede comprender aproximadamente 800 a 1.300 bases nucleotídicas y tiene un contenido de G/C de aproximadamente 50% y la secuencia complementaria de polinucleótidos tiene una relación de G/C de aproximadamente 1/1. La secuencia complementaria de polinucleótidos puede tener una relación de G/C de aproximadamente 3/2. Los ejemplos de plantillas que pueden usarse para generar las regiones de fijación de la etiqueta descritas en la presente memoria incluyen, pero no están limitadas a, secuencias seleccionadas del grupo que consiste en las SEQ ID NOS: 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 los métodos y composiciones, las regiones de fijación de la etiqueta pueden comprender un contenido similar de adenina. Las bases de adenina pueden estar separadas al menos en un promedio de cada 8 a 16 bases nucleotídicas. Las regiones de fijación de la etiqueta pueden comprender un patrón repetido regularmente de bases de adenina. Preferiblemente, las bases de adenina están separadas aproximadamente cada 8 a 16 bases nucleotídicas. Preferiblemente, las regiones de fijación de la etiqueta comprenden un contenido de timina de aproximadamente 35-45%.

En algunos métodos y composiciones, la secuencia complementaria de polinucleótidos comprende una secuencia de polinucleótidos de ARN. La secuencia de polinucleótidos de ARN puede comprender al menos una base de uracilo modificada con aminoalilo. La molécula detectable en la secuencia complementaria de polinucleótidos se puede unir a la base de uracilo modificado con aminoalilo. La secuencia de polinucleótidos de ARN puede comprender una pluralidad de bases de uracilo modificadas con aminoalilo que están separadas aproximadamente un promedio de cada 8 a 16 bases en dicha secuencia de polinucleótidos de ARN. Preferiblemente, la molécula detectable está unida a cada una de las bases de uracilo modificadas con alilo.

En algunas composiciones y métodos, las moléculas detectables son colorantes fluorescentes.

También se proporcionan métodos para preparar al menos un nanoinformador etiquetado de forma única que comprende: i) combinar una pluralidad de la región de fijación de la etiqueta que comprende cada una aproximadamente 800 a 1.300 bases nucleotídicas y un contenido de G/C de aproximadamente 50%, en donde cada región de fijación seleccionada de la etiqueta es diferente de las otras regiones de unión seleccionadas de la etiqueta, ii) unir covalentemente la pluralidad de regiones de fijación de la etiqueta entre sí en combinación lineal; y iii) hibridar una secuencia complementaria de polinucleótidos con dicha región de fijación de la etiqueta, en donde dicha secuencia complementaria de polinucleótidos está unida a la misma con una o más moléculas detectables. Los métodos pueden comprender además preparar una sonda nanoinformadora etiquetada uniendo el nanoinformador etiquetado a una región específica objetivo. Preferiblemente, la secuencia complementaria de polinucleótidos tiene una relación de G/C de al menos 1/1.

También se proporcionan kits para preparar al menos un nanoinformador etiquetado de forma única que comprende: a) al menos tres regiones de fijación de la etiqueta que comprenden cada una aproximadamente 800 a 1.300 bases nucleotídicas, un contenido de G/C de aproximadamente 50%; y b) al menos tres secuencias complementarias de polinucleótidos que tienen unida a la misma una molécula detectable, en donde las secuencias complementarias de polinucleótidos tienen una relación de G/C de al menos 1/1. Preferiblemente, los kits comprenden además al menos tres sondas específicas para el objetivo.

También se pueden proporcionar kits que comprenden una población de nanoinformadores como se describe en la presente memoria e instrucciones para su uso.

Breve descripción de los dibujos

Las nuevas características 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 haciendo referencia a la siguiente descripción detallada y a los dibujos adjuntos, de los cuales:

La Figura 1 es una ilustración esquemática de un nanoinformador dual con un código del nanoinformador de 6

posiciones, que utiliza una sonda de captura y un único componente nanoinformador de 6 posiciones. La flecha es ilustrativa de una etiqueta de afinidad, que se incluye opcionalmente y que se puede usar para purificar el nanoinformador o inmovilizar el nanoinformador (o complejo de nanoinformador-molécula objetivo) con el fin de obtener imágenes.

5 Figuras 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 nanoinformador etiquetado con S2-B y el ADN objetivo S2 (Figura 2B) o el ARN objetivo S2 (Figura 2C). La Figura 2E muestra un primer plano de complejos de nanoinformadores de la Figura 2B, conteniendo cada uno sonda de captura S2-A, un nanoinformador etiquetado con S2-B y ADN objetivo S2. La Figura 2D muestra una imagen de un experimento de control negativo, en donde la sonda de captura S2-A, el nanoinformador etiquetado con S2-B y el ARN objetivo S2 no se hibridaron.

15 Figuras 3A-3B: Muestran un nanoinformador etiquetado (basado en ácido nucleico) con una etiqueta de afinidad, A1, en un extremo. En la Figura 3, el nanoinformador etiquetado se inmoviliza mediante la unión de A1 a un compañero de afinidad inmovilizado. El otro extremo del nanoinformador está en solución (Figura 3A), pero puede inmovilizarse mediante hibridación con un oligonucleótido complementario que contiene otra etiqueta de afinidad (A2) utilizada para inmovilizar el nanoinformador (Figura 3B). A1 y A2 pueden ser iguales, por ejemplo, biotina, para la inmovilización sobre una superficie recubierta de avidina o estreptavidina. Tras la inmovilización de A1, el nanoinformador puede estirarse, o "alargarse" como se representa en la Figura 3, por ejemplo, mediante 20 contracción-relajación eléctrica, para la separación de las regiones de fijación de la etiqueta de una manera que permita la detección del código del nanoinformador. Opcionalmente, mientras el nanoinformador está en un estado alargado, A2 se introduce y une el extremo del nanoinformador que es complementario a A2 bajo la superficie.

25 Figuras 4A-4C. La Figura 4A ilustra la inmovilización de un extremo de una molécula de ADN en un dispositivo microfluído; 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 extendida mediante etiquetas de afinidad introducidas en el dispositivo después del alargamiento.

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

30 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.

35 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 promedio de 6 veces para el número absoluto de conteos medidos por los informadores DV1 en comparación con los informadores M13. Los datos se recogieron de experimentos representativos realizados de acuerdo con el Ejemplo 5.

40 La Figura 8 muestra un gráfico de "informadores válidos", que se refiere a cadenas de puntos que pueden ser interpretados por el software de análisis de imágenes como un informador "contable". Los resultados se describen como un porcentaje del número total de eventos vinculantes en un campo de visión. Un porcentaje significativamente mayor de los informadores DV1 son contables, en este experimento conduce a una ganancia de 3 veces en los datos (de 12,5% a 38%).

La Figura 9 muestra la secuencia de polinucleótidos de un vector plasmídico que puede utilizarse para clonar, propagar, y generar las cadenas principales de nanoinformador de cadena sencilla de la presente invención.

45 La Figura 10 muestra las secuencias de polinucleótidos de dos plantillas que se utilizaron en los experimentos de optimización del colorante para generar cadenas principales de nanoinformadores de cadena sencilla (véase el Ejemplo 6). En estas secuencias de plantilla, la base repetida regularmente es timina, que tras la transcripción produce una cadena principal de nanoinformador de cadena sencilla complementaria que tiene adenina como la base repetida regularmente. La Figura 10A muestra la secuencia de polinucleótidos de una plantilla que tiene una base regularmente repetida aproximadamente cada 8 nucleótidos. La Figura 10B muestra la secuencia de 50 polinucleótidos de una plantilla que tiene una base repetida regularmente aproximadamente cada 10 nucleótidos.

La Figura 11 muestra las secuencias de polinucleótidos de dos plantillas que se utilizaron en los experimentos de optimización del colorante para generar cadenas principales de nanoinformadores de cadena sencilla (véase el Ejemplo 6). La Figura 11A muestra la secuencia de polinucleótidos de una plantilla que tiene una base regularmente repetida aproximadamente cada 12 nucleótidos. La Figura 11B muestra la secuencia de polinucleótidos de una 55 plantilla que tiene una base regularmente repetida aproximadamente cada 14 nucleótidos.

La Figura 12 muestra las secuencias de polinucleótidos de dos plantillas que se utilizaron en los experimentos de optimización del colorante para generar cadenas principales de nanoinformadores de cadena sencilla (véase el

Ejemplo 6). La Figura 12A muestra la secuencia de polinucleótidos de una plantilla que tiene una base regularmente repetida aproximadamente cada 16 nucleótidos. La Figura 12B muestra la secuencia de polinucleótidos de una plantilla que tiene una base regularmente repetida aproximadamente cada 24 nucleótidos.

Descripción detallada de la invención

5 La presente divulgación proporciona composiciones y métodos para la detección y cuantificación de moléculas objetivo individuales en muestras biomoleculares. En particular, la divulgación proporciona nanoinformadores estables que son capaces de unirse a moléculas objetivo individuales, y proporcionan una detección mejorada de moléculas objetivo. A través de los códigos de etiqueta de los nanoinformadores, la unión de las sondas nanoinformadoras a las moléculas objetivo da como resultado la identificación de las moléculas objetivo. También se proporcionan métodos para elaborar y usar tales nanoinformadores. Los nanoinformadores se pueden utilizar en una amplia variedad de aplicaciones tales como diagnóstico, pronóstico, control de calidad y aplicaciones de detección.

10 La divulgación también se refiere a la selección de una biblioteca o población de nanoinformadores diseñados (por ejemplo, secuencias sintéticas). Más específicamente, ciertas características de secuencia optimizadas proporcionan una estabilidad molecular mejorada de un nanoinformador, así como una detección mejorada cuando el nanoinformador se combina con una secuencia específica objetivo. Por ejemplo, los métodos y composiciones proporcionan nanoinformadores que comprenden cadenas principales sintéticas únicas que no producen una estructura secundaria y producen un brillo constante.

15 La presente divulgación de la invención también puede proporcionar una población de nanoinformadores diseñada (por ejemplo, sintética) en donde dichos nanoinformadores comprenden una pluralidad de diferentes moléculas detectables en cada nanoinformador tienen una señal detectable que la distingue de otros nanoinformadores en dicha población. Por lo tanto, también se proporciona una población de nanoinformadores diseñados con estabilidad molecular mejorada, en donde cada nanoinformador en la población tiene una señal detectable que la distingue de otros nanoinformadores en dicha población.

20 También se proporcionan nanoinformadores que comprenden cadenas principales diseñadas únicas (por ejemplo, sintéticas) hibridadas con una secuencia complementaria de polinucleótidos única que tiene unida a la misma una etiqueta detectable. Cada nanoinformador generará una señal única que no cambiará durante el curso de un ensayo de detección. Esa es la señal o código de señal único y distinguible asociado con el nanoinformador que permanecerá igual durante el curso de un ensayo de detección. En algunas realizaciones, la secuencia complementaria de polinucleótidos que tiene unida a la misma la etiqueta detectable puede ser un segmento de ARN acoplado con colorante transcrito *in vitro* que tiene un brillo constante. El brillo constante como se describe en el presente documento se refiere a la fuerza, el tamaño y/o la intensidad de la señal producida por los segmentos de ARN acoplados al colorante. Es decir, la fuerza, el tamaño y/o la intensidad de la señal del colorante serán similares entre los segmentos de ARN acoplados al colorante que tienen 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 o señal similar con la misma intensidad y/o brillo. Esto es útil, entre otras cosas, porque permite que un software/algoritmo analizador de imágenes o un usuario defina cómo debe ser un punto de un color particular en términos de brillo y tamaño. Además, varios de los colorantes podrían drenar a los canales de los otros colorantes, si la intensidad de la señal es constante, entonces la señal del drenaje también será consistente, permitiendo así que se ignore el drenaje. Además, los ruidos en las imágenes que no se encuentran dentro de los parámetros del punto se pueden descartar. Como resultado, por ejemplo, se pueden definir estrechamente los parámetros asociados con las señales de los nanoinformadores, lo que permite que el software/algoritmo o el usuario ignoren un mayor porcentaje del ruido. Estos proporcionan lecturas más robustas y confiables.

25 Los nanoinformadores pueden comprender una cadena principal con una disposición de regiones de fijación de la etiqueta basadas en nucleótidos, en donde cada región de fijación de la etiqueta tiene una secuencia específica designada para una etiqueta específica. En este sistema, la secuencia única de la cadena principal dicta el código de color del nanoinformador. Cada cadena principal se hibridará solo con la secuencia complementaria de polinucleótidos a su secuencia, cada una de las cuales tiene una etiqueta específica designada (por ejemplo, color). Por lo tanto, cada cadena principal generará solo el código designado, por ejemplo, incluso si la secuencia de polinucleótidos se separa de la cadena principal durante el proceso de detección. Si la secuencia de polinucleótidos en el nanoinformador se separa durante la síntesis o durante un procedimiento de hibridación, solo pueden reemplazarse con una secuencia de polinucleótidos del mismo color, eliminando la posibilidad de etiquetas compartidas o intercambiadas. La secuencia complementaria de polinucleótidos que tiene unida a la misma la etiqueta detectable puede ser un segmento de ARN acoplado con colorante transcrito *in vitro*. Preferiblemente, al construir un cadena principal de nanoinformador dada a partir de una biblioteca de plantillas de la región de fijación de la etiqueta basada en polinucleótidos individualmente únicos, a cada región de fijación de la etiqueta se le asigna una etiqueta detectable (por ejemplo, una molécula detectable), y cada región de fijación de la etiqueta dentro de un cadena principal dada se selecciona para que sea diferente de las otras regiones de fijación de la etiquetas en esa misma cadena principal.

30 Las secuencias pueden diseñarse para tener una distribución uniforme de una base, permitiendo una distribución uniforme de etiquetas acopladas (por ejemplo, flúor) cuando esta base se introduce como un nucleótido modificado

5 en un ARN o ADN polimerizado *in vitro*. Las secuencias pueden diseñarse sin repeticiones directas o inversas significativas para hacer que los nanoinformadores sean tan desestructurados y únicos como sea posible. Las secuencias pueden diseñarse sin ninguna repetición directa o inversa de 9 nucleótidos o más en una secuencia de 1.100 pares de bases de longitud. Las secuencias pueden diseñarse sin ninguna repetición directa o inversa de 7 nucleótidos o más a través de cualquier región de 100 pares de bases. Las regiones de fijación de la etiqueta y/o la secuencia complementaria de polinucleótidos pueden comprender un contenido y una relación de G/C particulares.

Ejemplos de plantillas de polinucleótidos que pueden utilizarse para generar estas regiones de fijación de la etiqueta diseñada se exponen en las secuencias de polinucleótidos de las 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.

10 Por consiguiente, se proporciona una población de nanoinformadores únicos o sondas nanoinformadoras, cada uno de los cuales comprende un cadena principal única basada en polinucleótidos, en donde cada nanoinformador de la población no solo es distinto de los otros nanoinformadores de la población, sino que también tiene estabilidad molecular mejorada y una señal detectable que la distingue de otros nanoinformadores en dicha población. Cada sonda nanoinformadora puede comprender una pluralidad de regiones de fijación individuales de la etiqueta diseñada. Cada una de las regiones de fijación de la etiqueta puede ser diferente de las otras regiones de fijación de la etiqueta en ese mismo nanoinformador. Por lo tanto, se proporciona una población más estable de nanoinformadores únicos que tienen propiedades de detección mejoradas. Un ejemplo de nanoinformador que tiene tales características se describe en la presente memoria, denominado nanoinformador DV1.

Nanoinformadores

20 Una sonda nanoinformadora completamente ensamblada y etiquetada comprende dos porciones principales, una secuencia específica del objetivo que es capaz de unirse a una molécula objetivo, y un nanoinformador etiquetado que proporciona un "código" de señales (el "código del nanoinformador") asociado con la secuencia específica del objetivo. Tras la unión de la sonda nanoinformadora a la molécula objetivo, el código del nanoinformador identifica la molécula objetivo a la que se une el nanoinformador.

25 Los nanoinformadores son estructuras modulares. El nanoinformador puede comprender una pluralidad de diferentes moléculas detectables. Un nanoinformador etiquetado puede ser una entidad molecular que contiene ciertos elementos básicos: (i) una pluralidad de regiones de fijación únicas de la etiqueta unidas en una combinación lineal única particular, y (ii) secuencias complementarias de polinucleótidos unidas a las regiones de fijación de la etiqueta de la cadena principal. El nanoinformador etiquetado puede comprender 2, 3, 4, 5, 6, 7, 8, 9, 10 o más regiones de fijación únicas de la etiqueta unidas en una combinación lineal única particular, y secuencias complementarias de polinucleótidos unidas a las regiones de fijación de la etiqueta de la cadena principal. El nanoinformador etiquetado puede comprender 3 o más regiones de fijación únicas de la etiqueta unidas en una combinación lineal única particular, y unidas a las regiones de fijación de la etiqueta de la cadena principal. El nanoinformador etiquetado puede comprender 6 o más regiones de fijación únicas de la etiqueta unidas en una combinación lineal única particular, y secuencias complementarias de polinucleótidos unidas a las regiones de fijación de la etiqueta de la cadena principal. Una sonda nanoinformadora comprende además una secuencia específica del objetivo, también unida a la cadena principal.

40 El término región de fijación de etiqueta incluye una región de secuencia de polinucleótidos definida dentro de una cadena principal dada que puede servir como un punto de unión individual para una molécula detectable. Las regiones de fijación de la etiqueta pueden comprender secuencias diseñadas. Los ejemplos específicos de secuencias de plantilla definidas de polinucleótidos que pueden utilizarse para generar regiones de fijación de la etiqueta de cadena sencilla incluyen las secuencias de plantilla expuestas en las SEQ ID NOS: 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 60%, 70%, 80%, 90%, 95%, 98% o más, incluidos todos los enteros intermedios, de identidad de secuencia). La transcripción de las secuencias de plantilla de las SEQ ID NOS: 1-24 genera regiones de fijación de la etiqueta de cadena sencilla que tienen una secuencia de polinucleótidos que es complementaria a las secuencias de las SEQ ID NOS: 1-24.

50 El nanoinformador de la etiqueta también puede comprender una cadena principal que contiene una región constante. El término región constante incluye secuencias repetidas en tándem de aproximadamente 10 a aproximadamente 25 nucleótidos que están unidos covalentemente a un nanoinformador. La región constante puede unirse a la región 5' o a la región 3' de un nanoinformador, y puede utilizarse para la captura e inmovilización de un nanoinformador para formación de imágenes o detección, tal como mediante la unión a un sustrato sólido de una secuencia que es complementaria a la región constante. La región constante puede contener 2, 3, 4, 5, 6, 7, 8, 9, 10 o más secuencias repetidas en tándem, donde las secuencias repetidas comprenden cada una 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, incluidos aproximadamente 12-18, 13-17 o aproximadamente 14-16 nucleótidos.

55 Los nanoinformadores descritos en la presente memoria comprenden secuencias sintéticas diseñadas. Los nanoinformadores descritos en la presente memoria pueden proporcionar una mejor consistencia en el brillo de la etiqueta. El brillo consistente como se describe en el presente documento se refiere a la fuerza, el tamaño y/o la

intensidad de la señal producida por los segmentos etiquetados. Es decir, la fuerza, el tamaño y/o la intensidad de la señal de los segmentos etiquetados serán similares entre los segmentos que tienen la misma etiqueta unida a los mismos. Como se describió anteriormente, esta consistencia mejorada conduce a datos más sólidos. Un nucleótido puede estar separado al menos en un promedio de 8, 9, 10, 12, 15, 16, 20, 30 o 50 bases. Un nucleótido puede estar separado al menos un promedio de 8 a 16 bases. Un nucleótido puede estar separado al menos un promedio de 8 bases. Las secuencias pueden contener un patrón de separación muy regular de un residuo de nucleótido (por ejemplo, adenina) en la cadena principal. Esto permite un nucleótido complementario regularmente separado en la secuencia complementaria de polinucleótidos que tiene unida a la misma una molécula detectable. Por ejemplo, cuando las secuencias del nanoinformador contienen un patrón separado muy regularmente de residuos de adenina en la cadena principal, cuyo complemento es un patrón regularmente separado de residuos de uridina (U) en segmentos de ARN complementarios, la transcripción *in vitro* de los segmentos puede realizarse utilizando una base de uridina modificada con aminoalilo, que permite el acoplamiento covalente de amina de moléculas de colorante a intervalos regulares a lo largo del segmento. Las secuencias pueden contener aproximadamente el mismo número o porcentaje de un nucleótido (por ejemplo, adenina) que está separado al menos en un promedio de 8, 9, 10, 12, 15, 16, 20, 30 o 50 bases en las secuencias. Esto permite un número o porcentajes similares en la secuencia complementaria de polinucleótidos que tienen unida a la misma una molécula detectable. Por lo tanto, las secuencias pueden contener un nucleótido que no está separado regularmente pero que está separado al menos en un promedio de 8, 9, 10, 12, 15, 16, 20, 30 o 50 bases, el número de nucleótidos en las secuencias pueden variar dependiendo del brillo deseado para el nanoinformador. Preferiblemente, el 20%, 30%, 50%, 60%, 70%, 80%, 90% o 100% del nucleótido complementario está acoplado a una molécula detectable. Por ejemplo, cuando las secuencias del nanoinformador contienen un porcentaje similar de residuos de adenina en la cadena principal y la transcripción *in vitro* de los segmentos complementarios se realiza usando una base de uridina modificada con aminoalilo, 20%, 30%, 50%, 60%, 70%, 80%, 90% o 100% de la base de uridina modificada con aminoalilo se puede acoplar a una molécula detectable. Alternativamente, la proporción de bases de uridina modificadas con aminoalilo y bases de uridina puede cambiarse durante el proceso de transcripción *in vitro* para lograr el brillo deseado. Por ejemplo, el proceso de transcripción *in vitro* puede tener lugar en presencia de una mezcla con una relación de 1/1 de uridina con respecto a bases de uridina modificadas con aminoalilo, cuando algunas o todas las bases de uridina modificadas con aminoalilo se pueden acoplar a una molécula detectable. Por lo tanto, una persona con experiencia ordinaria en la técnica comprenderá que hay varios métodos en los que se puede lograr un brillo constante entre los nanoinformadores.

Los nanoinformadores descritos en la presente memoria pueden tener una temperatura de fusión bastante constante (T_m). La T_m de las secuencias complementarias de polinucleótidos cuando se hibridan con sus regiones de fijación de la etiqueta en los nanoinformadores descritos en la presente memoria puede ser mayor que la T_m de una secuencia complementaria de polinucleótidos a una plantilla de ADN de M13 cuando se hibrida con una sonda nanoinformadora que comprende el ADN de M13. Sin pretender limitarse a ninguna teoría, la T_m de los nanoinformadores descritos en la presente memoria permite enlaces más fuertes entre la cadena principal del nanoinformador y la secuencia complementaria de polinucleótidos que tiene unida a la misma una molécula detectable, evitando así la disociación durante los procedimientos de síntesis e hibridación. Además, la T_m consistente entre una población de nanoinformadores permite que los procedimientos de síntesis e hibridación estén estrechamente optimizados, ya que las condiciones óptimas son las mismas para todos los puntos y posiciones. En algunas realizaciones, las secuencias de los nanoinformadores tienen un 50% de guanina/citosina (G/C), con no más de tres G en una fila. Por lo tanto, puede proporcionarse una población de nanoinformadores en la que la T_m entre los nanoinformadores de la población sea bastante uniforme. Preferiblemente, se proporciona una población de nanoinformadores en los que la T_m de las secuencias complementarias de polinucleótidos cuando se hibridan con sus regiones de fijación de la etiqueta es de aproximadamente 80°C, 85°C, 90°C, 100°C o más. También se proporciona una población de nanoinformadores en la que la T_m de las secuencias complementarias de polinucleótidos cuando se hibrida con sus regiones de fijación de la etiquetaje es de aproximadamente 80°C o superior.

Los nanoinformadores descritos en la presente memoria pueden tener o no estructuras secundarias mínimas, tales como cualquier interacción intramolecular estable de apareamiento de bases (por ejemplo, horquillas). Sin pretender limitarse a ninguna teoría, la estructura secundaria mínima en los nanoinformadores proporciona una mejor hibridación entre la cadena principal del nanoinformador y la secuencia de polinucleótidos que tiene unida a la misma una molécula detectable. Además, la estructura secundaria mínima en los nanoinformadores proporciona una mejor detección de las moléculas detectables en los nanoinformadores. Los nanoinformadores descritos en la presente memoria pueden no tener emparejamiento intramolecular significativo en condiciones de hibridación de 75°C, 1XSSPE. Las estructuras secundarias pueden predecirse mediante programas conocidos en la técnica tales como MFOLD. Los nanoinformadores descritos en la presente memoria pueden contener menos del 1% de repeticiones invertidas en cada cadena, en donde las repeticiones invertidas son 9 bases o más. Los nanoinformadores descritos en la presente memoria pueden no contener repeticiones invertidas en cada cadena. Los nanoinformadores no pueden contener ninguna repetición invertida de 9 nucleótidos o más en una secuencia de 1.100 pares de bases de longitud. Los nanoinformadores no pueden contener ninguna repetición invertida de 7 nucleótidos o más en cualquier región de 100 pares de bases. Los nanoinformadores descritos en la presente memoria pueden contener menos del 1% de repeticiones invertidas en cada cadena, en donde las repeticiones invertidas son de 9 nucleótidos o más en una secuencia de 1.100 pares de bases de longitud. Los nanoinformadores

descritos en la presente memoria pueden contener menos del 1% de repeticiones invertidas en cada cadena, en donde las repeticiones invertidas son de 7 nucleótidos o más en cualquier región de 100 pares de bases. Los nanoinformadores descritos en la presente memoria pueden tener un contenido específico de cadena asimétrica de manera que una cadena es rica en CT y la otra es rica en GA.

- 5 También se proporcionan nanoinformadores únicos. Los nanoinformadores descritos en la presente memoria pueden contener menos del 1% de repeticiones directas. Los nanoinformadores descritos en la presente memoria pueden no contener repeticiones directas. Los nanoinformadores no pueden contener ninguna repetición directa de 9 nucleótidos o más en una secuencia de 1.100 pares de bases de longitud. Preferiblemente, los nanoinformadores etiquetados no contienen ninguna repetición directa de 7 nucleótidos o más en ninguna región de 100 pares de bases. Preferiblemente, los nanoinformadores descritos en la presente memoria contienen menos del 1% de repeticiones directas en cada cadena, en donde las repeticiones directas son de 9 nucleótidos o más en una secuencia de 1.100 pares de bases de longitud. Preferiblemente, los nanoinformadores descritos en la presente memoria contienen menos del 1% de repeticiones directas en cada cadena, en donde las repeticiones directas son de 7 nucleótidos o más en cualquier región de 100 pares de bases. Preferiblemente, los nanoinformadores descritos en la presente memoria contienen menos de 85, 80, 70, 60, 50, 40, 30, 20, 10 o 5% de homología con cualquier otra secuencia utilizada en las cadenas principales o con cualquier secuencia descrita en la base de datos pública REFSEQ. Preferiblemente, los nanoinformadores descritos en la presente memoria contienen menos del 85% de homología con cualquier otra secuencia usada en las cadenas principales o con cualquier secuencia descrita en la base de datos pública REFSEQ. Preferiblemente, los nanoinformadores descritos en la presente memoria contienen menos de 20, 16, 15, 10, 9, 7, 5, 3, 2 bases contiguas de homología con cualquier otra secuencia usada en las cadenas principales o cualquier secuencia descrita en la base de datos pública REFSEQ. Preferiblemente, los nanoinformadores descritos en la presente memoria tienen no más de 15 bases contiguas de homología y no más de 85% de identidad en toda la longitud del nanoinformador con cualquier otra secuencia utilizada en las cadenas principales o cualquier secuencia descrita en la base de datos pública REFSEQ.
- 25 Preferiblemente, las características de secuencia de las sondas nanoinformadoras descritas en la presente memoria proporcionan una detección mejorada de una molécula objetivo. Por ejemplo, la unión de las sondas nanoinformadoras a moléculas objetivo que da como resultado la identificación de las moléculas objetivo se puede realizar detectando individualmente la presencia del nanoinformador. Esto puede realizarse contando individualmente la presencia de una o más de las moléculas de nanoinformador en una muestra. Cuando se usan tales métodos de conteo, las sondas nanoinformadoras descritas en la presente pueden permitir un aumento en el número de conteos. Preferiblemente, el número de conteos moleculares por encima del valor base de dicho complejo molecular después de la normalización de la muestra es superior a 300, 400, 450, 500, 600, 700, 800, 900 o 1.000 conteos moleculares. Preferiblemente, el número de conteos moleculares por encima del valor base de dicho complejo molecular después de la normalización de la muestra es superior a 400 conteos moleculares.
- 35 Preferiblemente, el porcentaje de conteos moleculares válidos de las sondas nanoinformadoras descritas en la presente memoria es mayor que aproximadamente 10, 11, 12, 12,5, 14, 15, 20, 25, 30, 35, 40, 50, 60, 70, 80 o 90%. Preferiblemente, el porcentaje de conteos moleculares válidos de las sondas nanoinformadoras descritas en la presente memoria es mayor que aproximadamente 10%. Preferiblemente, el porcentaje de conteos moleculares válidos de las sondas nanoinformadoras descritas en la presente memoria es mayor que aproximadamente 12,5%.
- 40 Preferiblemente, el número de conteos moleculares por encima del valor base del nanoinformador descrito en la presente memoria después de la normalización de dicha muestra es de al menos 2, 3, 5, 6, 10, 20, 30, 40, 50, 60, 70, 80, 90 o 100 veces más alta que una sonda nanoinformadora comparable que comprende ADN de M13. Un nanoinformador comparable que comprende ADN de M13 es un nanoinformador que comprende la misma región específica objetivo unida a una cadena principal de ADN de M13. Ejemplos de sondas nanoinformadoras comparables que comprenden ADN de M13 se describen en la sección de Ejemplos. Preferiblemente, el número de conteos moleculares por encima del valor base del nanoinformador descrito en la presente memoria después de la normalización de dicha muestra es al menos 2 veces mayor que un nanoinformador comparable que comprende ADN de M13. Preferiblemente, el número de conteos moleculares por encima del valor base del nanoinformador descrito en la presente memoria después de la normalización de dicha muestra es al menos 6 veces mayor que un nanoinformador comparable que comprende ADN de M13. Preferiblemente, el número de conteos moleculares por encima del valor base del nanoinformador descrito en la presente memoria después de la normalización de dicha muestra es al menos 20 veces mayor que un nanoinformador comparable que comprende ADN de M13. Preferiblemente, el número de conteos moleculares por encima del valor base del nanoinformador descrito en la presente memoria después de la normalización de dicha muestra es al menos 100 veces mayor que un nanoinformador comparable que comprende ADN de M13.

Los elementos de un nanoinformador se pueden encontrar en una única entidad molecular (un nanoinformador "singular"), o dos entidades moleculares distintas (un nanoinformador "dual"). Cada entidad molecular puede estar compuesta de una molécula o más de una molécula unidas entre sí por medios covalentes o no covalentes. Preferiblemente, cada componente de un nanoinformador dual tiene una secuencia específica del objetivo que se une a un sitio diferente en la misma molécula objetivo. Esto permite componentes de nanoinformadores más pequeños con una cinética de unión del nanoinformador más eficiente a la molécula objetivo y una mejor relación señal:ruido como resultado de la mayor especificidad de unión. Cuando se utiliza un sistema de nanoinformador dual, una de las sondas nanoinformadoras puede no estar etiquetada. Preferiblemente, la sonda nanoinformadora

sin etiquetar puede comprender una región de captura. En algunas realizaciones, la sonda nanoinformadora sin etiquetar puede comprender una región específica del objetivo y una cadena principal que puede ser de cadena sencilla. Preferiblemente, la sonda nanoinformadora no etiquetada puede comprender una región específica del objetivo y una cadena principal que puede ser de doble cadena.

5 Las secuencias complementarias de polinucleótidos unidas a una cadena principal del nanoinformador sirven para unir moléculas detectables, o monómeros etiqueta, a la cadena principal del nanoinformador. Las secuencias complementarias de polinucleótidos pueden etiquetarse directamente, por ejemplo, mediante la incorporación covalente de una o más moléculas detectables en la secuencia complementaria de polinucleótidos. Alternativamente, las secuencias complementarias de polinucleótidos pueden etiquetarse indirectamente, tal como
10 mediante la incorporación de biotina u otra molécula capaz de una interacción de ligando específica en la secuencia complementaria de polinucleótidos. En tales casos, el ligando (por ejemplo, estreptavidina en el caso de la incorporación de biotina en la secuencia complementaria de polinucleótidos) puede unirse covalentemente a la molécula detectable. Cuando las moléculas detectables unidas a una región de fijación de la etiqueta no se incorporan directamente en la secuencia complementaria de polinucleótidos, esta secuencia sirve como un puente
15 entre la molécula detectable y la región de fijación de la etiqueta, y puede denominarse una molécula puente, por ejemplo, un ácido nucleico puente.

El nanoinformador a base de ácido nucleico y los complejos objetivo de nanoinformador comprenden ácidos nucleicos, que pueden purificarse por afinidad o inmovilizarse usando un ácido nucleico, tal como un oligonucleótido, que es complementario a la región constante o el nanoinformador o el ácido nucleico objetivo. Como se indicó
20 anteriormente, los nanoinformadores pueden comprender al menos una región constante, que puede servir como una etiqueta de afinidad para la purificación y/o para la inmovilización (por ejemplo, a una superficie sólida). La región constante típicamente comprende dos o más regiones repetidas en tándem de nucleótidos repetidos, tales como una serie de repeticiones de 15 bases. En dichas realizaciones a modo de ejemplo, el nanoinformador, ya sea complejo a una molécula objetivo o de otro modo, puede purificarse o inmovilizarse mediante un reactivo de
25 afinidad recubierto con un oligonucleótido de 15 bases que es el complemento inverso de la unidad de repetición.

Los nanoinformadores, o complejos de nanoinformador-molécula objetivo, pueden purificarse en dos o más etapas de selección de afinidad. Por ejemplo, en un nanoinformador dual, una sonda puede comprender una primera etiqueta de afinidad y la otra sonda puede comprender una segunda etiqueta de afinidad (diferente). Las sondas se mezclan con moléculas objetivo, y los complejos que comprenden las dos sondas del nanoinformador dual se separan de los materiales no unidos (por ejemplo, el objetivo o las sondas individuales del nanoinformador)
30 mediante purificación por afinidad contra una o ambas etiquetas de afinidad individuales. En la primera etapa, la mezcla se puede unir a un reactivo de afinidad para la primera etiqueta de afinidad, de modo que solo se purifican las sondas que comprenden la primera etiqueta 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 la segunda etiqueta de afinidad, permitiendo la separación de complejos de sondas que comprenden la primera etiqueta de afinidad. En este punto, solo los complejos completos estarían unidos. Los complejos finalmente se liberan del reactivo de afinidad para la segunda etiqueta de afinidad y, a continuación, preferiblemente se estiran y se toman imágenes. El reactivo de afinidad puede ser cualquier superficie sólida recubierta con un compañero de unión para la etiqueta de afinidad, tal como una columna, una perla (por ejemplo, látex o perla magnética) o platina recubierta con la pareja de
35 unión. Los nanoinformadores de inmovilización y estiramiento que usan reactivos de afinidad se describen completamente en la solicitud provisional de Estados Unidos No. 60/753,816 de Sean M. Ferree y Dwayne L. Dunaway, titulada "Compositions Comprising Oriented, Immobilized Macromolecules and Methods for Their Preparation", presentada el 23 de diciembre de 2005.

La secuencia de señales proporcionada por los monómeros etiqueta asociados con las diversas regiones de fijación de la etiqueta de la cadena principal de un nanoinformador dado permite la identificación única del nanoinformador. Por ejemplo, cuando se usan etiquetas fluorescentes, un nanoinformador que tiene una identidad única o una firma espectral única se asocia con una secuencia específica del objetivo que reconoce una molécula objetivo específica o una porción de la misma. Cuando un nanoinformador se expone a una mezcla que contiene la molécula objetivo en condiciones que permiten la unión de la secuencia o secuencias específicas objetivo del nanoinformador a la
45 molécula objetivo, la secuencia o secuencias específicas se unen preferentemente a la molécula objetivo. La detección de la señal del nanoinformador, tal como el código espectral de un nanoinformador etiquetado fluorescentemente, asociada con el nanoinformador permite la detección de la presencia de la molécula objetivo en la mezcla (análisis cualitativo). Contando todos los monómeros etiqueta asociados con un código o firma espectral dado, permite el recuento de todas las moléculas en la mezcla asociada con la secuencia específica del objetivo acoplada al nanoinformador (análisis cuantitativo). Los nanoinformadores son por lo tanto útiles para el diagnóstico o pronóstico de diferentes estados biológicos (por ejemplo, enfermedad frente salud) por análisis cuantitativo de marcadores biológicos conocidos. Además, la exquisita sensibilidad de la detección y cuantificación de moléculas individuales proporcionada por los nanoinformadores permite la identificación de nuevos marcadores de diagnóstico y pronóstico, incluidos aquellos cuyas fluctuaciones entre los diferentes estados biológicos son demasiado pequeñas
50 para detectar una correlación con un estado biológico particular utilizando métodos moleculares tradicionales. La sensibilidad de la detección molecular basada en nanoinformadores permite un análisis farmacocinético detallado de agentes terapéuticos y de diagnóstico en pequeñas muestras biológicas.
60

Muchos nanoinformadores, denominados nanoinformadores singulares, están compuestos por una entidad molecular. Sin embargo, para aumentar la especificidad de un nanoinformador y/o mejorar la cinética de su unión a una molécula objetivo, un nanoinformador puede ser un nanoinformador dual compuesto por dos entidades moleculares, cada una con una secuencia específica del objetivo diferente que se une a una región diferente de la misma molécula objetivo. En un nanoinformador dual, al menos una de las dos entidades moleculares está etiquetada. La otra entidad molecular no necesita ser etiquetada necesariamente. Dichos componentes no etiquetados de nanoinformadores duales pueden usarse como sondas de captura (véanse las Figuras 1 y 2) y opcionalmente tienen etiquetas de afinidad unidas, tales como biotina, que son útiles para inmovilizar y/o estirar el complejo que contiene el nanoinformador dual y la molécula objetivo para permitir visualización y/o formación de imágenes del complejo. Por ejemplo, un nanoinformador dual con un código del nanoinformador de 6 posiciones, utilizando un nanoinformador codificado de 6 posiciones y una sonda de captura. Se puede utilizar un nanoinformador dual con un código del nanoinformador de 7 posiciones, utilizando un componente de nanoinformador de 8 posiciones y un componente de nanoinformador de posición única. Un nanoinformador dual con un código del nanoinformador de 6 posiciones, utilizando una sonda de captura con una etiqueta de afinidad y un componente de nanoinformador de 6 posiciones. Se incluye opcionalmente una etiqueta de afinidad y se puede usar para purificar el nanoinformador o inmovilizar el nanoinformador (o complejo de nanoinformador-molécula objetivo) con el fin de obtener imágenes.

Debido a sus estructuras modulares, los nanoinformadores se pueden ensamblar y etiquetar de varias formas diferentes. Por ejemplo, un cadena principal de nanoinformador puede unirse a una secuencia específica del objetivo (por ejemplo mediante hibridación y, opcionalmente, ligamiento), y la estructura que comprende la cadena principal y la secuencia específica del objetivo unida a una o más secuencias complementarias de polinucleótidos que tienen unidas a la misma, ya sea directa o indirectamente, una molécula detectable. Alternativamente, la cadena principal del nanoinformador se puede unir primero a una o más secuencias complementarias de polinucleótidos, y la estructura de la cadena principal se une a continuación a una secuencia específica objetivo. Por lo tanto, a menos que se indique lo contrario, una discusión o listado de etapas en el ensamblaje de nanoinformadores no implica que se deba seguir una ruta específica de ensamblaje.

Las síntesis de nanoinformadores pueden realizarse por cualquier método adecuado conocido en la técnica. Los plásmidos de cadena doble que portan secuencias de nanoinformador (por ejemplo, secuencias DV1) se pueden cultivar en algunas cepas bacterianas en condiciones de baja temperatura (no superior a 34°C). Las cadenas principales de cadena sencillas lineales de nanoinformadores pueden prepararse a partir de ADN plásmido de cadena doble usando un protocolo de cuatro etapas que incluye linealización con una enzima de restricción, desfosforilación con una fosfatasa termolábil, digestión con una segunda enzima de restricción para separar el vector de clonación de la secuencia principal, y una digestión con exonucleasa lambda específica de la cadena que deja intacta solo una cadena del fragmento de cadena principal. La Figura 9 muestra un ejemplo de un vector que se usa para la síntesis de nanoinformadores.

El ensamblaje y uso de un nanoinformador se ejemplifica en la presente memoria en gran parte a modo de descripción de una diversidad de nanoinformadores basados en ácido nucleico. Las realizaciones ilustrativas de nanoinformadores parcialmente y completamente ensamblados se enumeran a continuación.

En su forma más simple, se proporciona una cadena principal de ácido nucleico que tiene una pluralidad (por ejemplo, 3) de regiones de fijación de la etiqueta que son capaces de etiquetarse y resolverse, cada una de las cuales está hecha de una secuencia de nucleótidos diseñada de manera racional. Estas secuencias abarcan una o más, o todas las características descritas en la presente memoria que hacen que el nanoinformador sea más estable. Los ejemplos de plantillas de polinucleótidos que pueden utilizarse para generar estas regiones de fijación diseñadas de la etiqueta se exponen en las secuencias de polinucleótidos de las SEQ ID NOS: 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 fijación de la etiqueta de cada cadena principal individual están dispuestas en un orden o combinación lineal única, por lo que cada estructura individual es única, en comparación con las otras cadenas principales en una población. Dado que cada uno de los nanoinformadores de una población está formado por una estructura única, cada nanoinformador es similar en comparación con los otros nanoinformadores de una población. Cada región de fijación de la etiqueta puede ser única cuando se compara con la otra región de nanoinformador de etiqueta en la cadena principal.

Las secuencias de nucleótidos de las regiones de fijación de la etiqueta individual dentro de cada nanoinformador pueden ser diferentes de las secuencias de nucleótidos de las otras regiones de fijación de la etiqueta dentro de ese nanoinformador, evitando reordenamientos, tales como recombinación, compartición o intercambio de las secuencias de polinucleótidos de etiqueta, y mejorando de ese modo la estabilidad molecular. El número de regiones de fijación de la etiqueta que se forman en una cadena principal se basa en la longitud y naturaleza de la cadena principal, los medios de etiquetado del nanoinformador, así como el tipo de monómeros etiqueta que proporcionan una señal para unirse a las regiones de fijación de la etiqueta de la cadena principal. Preferiblemente, a la secuencia de nucleótidos complementaria de cada región de fijación de la etiqueta se le asigna una molécula detectable específica.

También se proporcionan nanoinformadores etiquetados en los que una o más regiones de fijación de la etiqueta están unidas a una molécula detectable correspondiente, proporcionando cada molécula detectable una señal. Por

ejemplo, puede obtenerse un nanoinformador etiquetado cuando al menos tres moléculas detectables están unidas a tres regiones de fijación de la etiqueta correspondientes de la cadena principal de forma que estas regiones de fijación de la etiqueta etiquetadas, o los puntos, se distingan con base en su disposición lineal única. Un "punto", en el contexto de detección de nanoinformadores, es la señal agregada detectada de los monómeros etiqueta unidos a un único sitio de unión de la etiqueta en un nanoinformador, y que, dependiendo del tamaño de la región de fijación de la etiqueta y la naturaleza (p. ej., longitud de onda de emisión primaria) del monómeros etiqueta, puede aparecer como una única fuente puntual de luz cuando se visualiza bajo un microscopio. Los puntos de un nanoinformador pueden superponerse o no superponerse. El código del nanoinformador que identifica esa molécula objetivo puede comprender cualquier permutación de la longitud de un punto, su posición relativa con respecto a otros puntos, y/o la naturaleza (por ejemplo, la longitud o longitudes de onda de emisión primaria) de su señal. Generalmente, para cada sonda o par de sondas de la invención, las regiones de fijación de la etiqueta adyacentes no se solapan, y/o los puntos de regiones de fijación de la etiqueta adyacentes son distinguibles espacial y/o espectralmente, al menos bajo las condiciones de detección (por ejemplo, cuando el nanoinformador se inmoviliza, estira y observa bajo un microscopio, como se describe en la solicitud de Estados Unidos serial No. 61/029.220).

Ocasionalmente, se hace referencia a un tamaño de punto como un cierto número de bases o nucleótidos. Como entenderá fácilmente un experto en la técnica, esto se refiere al número de bases o nucleótidos en la región de fijación de la etiqueta correspondiente.

El orden y la naturaleza (por ejemplo, la longitud o longitudes de onda de emisión primaria, opcionalmente también longitud) de puntos de un nanoinformador sirven como un código del nanoinformador que identifica la molécula objetivo capaz de ser unida por el nanoinformador a través de la secuencia o secuencias específicas objetivo del nanoinformador. Cuando el nanoinformador está unido a una molécula objetivo, el código del nanoinformador también identifica la molécula objetivo. Opcionalmente, la longitud de un punto puede ser un componente del código del nanoinformador.

Las moléculas detectables que proporcionan una señal asociada con diferentes regiones de fijación de la etiqueta de la cadena principal pueden proporcionar señales que son indistinguibles bajo las condiciones de detección (señales "similares"), o pueden proporcionar señales que son distinguibles, al menos bajo las condiciones de detección (por ejemplo, cuando el nanoinformador está inmovilizado, estirado y observado bajo un microscopio).

También se proporciona un nanoinformador en donde dos o más moléculas detectables están unidas a una región de fijación de la etiqueta. La señal proporcionada por las moléculas detectables asociadas con dicha región de fijación de la etiqueta produce una señal agregada que se detecta. La señal agregada producida puede estar formada por señales similares o constituidas por al menos dos señales distinguibles (por ejemplo, señales distinguibles espectralmente).

También se proporciona un nanoinformador en donde al menos tres moléculas detectables que proporcionan señales similares están unidas a tres regiones de fijación de la etiqueta correspondientes de la cadena principal y dichas tres moléculas detectables son espacialmente distinguibles. También se proporciona un nanoinformador en donde al menos tres moléculas detectables que proporcionan tres señales distinguibles están unidas a tres regiones de fijación de la etiqueta vecinas, por ejemplo tres regiones de fijación de la etiqueta adyacentes, por lo que dichos al menos tres monómeros etiqueta son distinguibles espectralmente.

También se proporciona un nanoinformador en donde los puntos que proporcionan señales similares o diferentes están separadas por una región separadora, por lo que la interposición de la región separadora permite la generación de puntos oscuros, que expanden la combinación posible de señales detectables de manera única. El término "punto oscuro" se refiere a la falta de señal de un sitio de unión de la etiqueta en un nanoinformador. Los puntos oscuros se pueden incorporar al código del nanoinformador para agregar más permutaciones de codificación y generar una mayor diversidad de nanoinformadores en una población de nanoinformadores. Preferiblemente, las regiones separadoras tienen una longitud determinada por la resolución de un instrumento empleado para detectar el nanoinformador.

También se proporciona un nanoinformador con uno o más "puntos dobles". Cada punto doble contiene dos o más (por ejemplo, tres, cuatro o cinco) puntos adyacentes que proporcionan señales similares sin estar separados por una región separadora. Los puntos dobles se pueden identificar por sus tamaños.

Una molécula detectable que proporciona una señal se puede unir de forma covalente o no covalente (por ejemplo, mediante hibridación) a una secuencia complementaria de polinucleótidos que está unida a la región de fijación de la etiqueta. Los monómeros etiqueta también se pueden unir indirectamente a la secuencia complementaria de polinucleótidos, tal como uniéndose covalentemente a una molécula de ligando (por ejemplo, estreptavidina) que se une mediante su interacción con una molécula incorporada en la secuencia complementaria de polinucleótidos (por ejemplo, biotina incorporada en la secuencia complementaria de polinucleótidos), que a su vez está unida mediante hibridación a la cadena principal.

También se proporciona un nanoinformador asociado con una señal detectable de forma única, tal como un código espectral, determinado por la secuencia de señales proporcionadas por los monómeros etiqueta unidos (por

ejemplo, indirectamente) para marcar regiones de unión en la cadena principal del nanoinformador, por lo que la detección de la señal permite la identificación del nanoinformador.

También se proporciona un nanoinformador que comprende además una etiqueta de afinidad unida a la cadena principal del nanoinformador, de manera que la unión de la etiqueta de afinidad a un soporte permite el estiramiento de la cadena principal y la resolución de las señales proporcionadas por monómeros etiqueta correspondientes a diferentes regiones de fijación de la etiquetas en la cadena principal. El estiramiento del nanoinformador puede implicar cualquier medio de estiramiento conocido en la técnica, que incluye, pero no se limita a, medios que implican medios físicos, hidrodinámicos o eléctricos. La etiqueta de afinidad puede comprender una región constante.

Un nanoinformador puede incluir además una secuencia específica del objetivo acoplada a la cadena principal. La secuencia específica del objetivo se selecciona para permitir que el nanoinformador reconozca, se enlace o se una a una molécula objetivo. Los nanoinformadores son adecuados para la identificación de moléculas objetivo de todos los tipos. Por ejemplo, las secuencias apropiadas específicas del objetivo se pueden acoplar a la cadena principal del nanoinformador para permitir la detección de una molécula objetivo. Preferiblemente, la molécula objetivo es ADN (incluido ADNc), ARN (incluyendo ARNm y ARNc), un péptido, un polipéptido o una proteína.

También se proporciona una mayor flexibilidad en la detección de moléculas objetivo con monómeros etiqueta. Preferiblemente, un nanoinformador dual que comprende dos entidades moleculares diferentes, cada una con una región separada específica del objetivo, al menos una de las cuales está etiquetada, se une a la misma molécula objetivo. Por lo tanto, las secuencias específicas del objetivo de los dos componentes del nanoinformador dual se unen a diferentes porciones de una molécula objetivo seleccionada, por lo que la detección del código espectral asociado con el nanoinformador dual permite la detección de la molécula objetivo seleccionada en una muestra biomolecular contactada con dicho nanoinformador dual.

También se proporciona un método para detectar la presencia de una molécula objetivo específica en una muestra biomolecular que comprende: (i) poner en contacto dicha muestra con un nanoinformador como se describe en la presente memoria (por ejemplo, un nanoinformador singular o doble) en condiciones que permiten la unión de las secuencias específicas del objetivo en el nanoinformador dual a la molécula objetivo y (ii) la detección del código espectral asociado con el nanoinformador dual. Dependiendo de la arquitectura del nanoinformador, el nanoinformador dual puede etiquetarse antes o después de unirse a la molécula objetivo.

La unicidad de cada sonda nanoinformadora en una población de sondas puede permitir el análisis multiplexado de una pluralidad de moléculas objetivo. Por ejemplo, cada sonda nanoinformadora contiene seis regiones de fijación de la etiqueta, donde cada región de fijación de la etiqueta de cada cadena principal es diferente de las otras regiones de fijación de la etiqueta en esa misma cadena principal. Si las regiones de fijación de la etiqueta van a etiquetarse con uno de cuatro colores y hay 24 secuencias únicas posibles para las regiones de fijación de la etiqueta y cada región de fijación de la etiqueta tiene asignado un color específico, cada región de fijación de etiqueta en cada cadena principal consistirá en una de cuatro secuencias. Habrá 4.096 posibles nanoinformadores en este ejemplo. El número de posibles nanoinformadores se puede aumentar, por ejemplo, aumentando el número de colores, aumentando el número de secuencias únicas para las regiones de fijación de la etiqueta y/o aumentando el número de regiones de fijación de la etiqueta por cadena principal. Asimismo, el número de posibles nanoinformadores puede disminuirse disminuyendo el número de colores, disminuyendo el número de secuencias únicas para las regiones de fijación de la etiqueta y/o disminuyendo el número de regiones de fijación de la etiqueta por cadena principal.

Los métodos de detección se pueden llevar a cabo en ensayos múltiplex, por lo que se detecta una pluralidad de moléculas objetivo en el mismo ensayo (una sola mezcla de reacción). Preferiblemente, el ensayo es un ensayo de hibridación en donde la pluralidad de moléculas objetivo se detecta simultáneamente. La pluralidad de moléculas objetivo detectadas en el mismo ensayo puede ser, al menos 2, al menos 5 moléculas objetivo diferentes, al menos 10 moléculas objetivo diferentes, al menos 20 moléculas objetivo diferentes, al menos 50 moléculas objetivo diferentes, al menos 75 moléculas objetivo diferentes, al menos 100 moléculas objetivo diferentes, al menos 200 moléculas objetivo diferentes, al menos 500 moléculas objetivo diferentes, o al menos 750 moléculas objetivo diferentes. La pluralidad de moléculas objetivo detectadas en el mismo ensayo es de hasta 50 moléculas objetivo diferentes, hasta 100 moléculas objetivo diferentes, hasta 150 moléculas objetivo diferentes, hasta 200 moléculas objetivo diferentes, hasta 300 moléculas objetivo diferentes, hasta 500 moléculas objetivo diferentes, hasta 750 moléculas objetivo diferentes, hasta 1.000 moléculas objetivo diferentes, hasta 2.000 moléculas objetivo diferentes o hasta 5.000 moléculas objetivo diferentes. La pluralidad de moléculas objetivo detectadas puede ser cualquier intervalo entre los números anteriores de moléculas objetivo diferentes, tales como, pero sin limitación, de 20 a 50 moléculas objetivo diferentes, de 50 a 200 moléculas objetivo diferentes, de 100 a 1.000 moléculas objetivo diferentes, de 500 a 5000 moléculas objetivo diferentes, y así sucesivamente.

La detección de diferentes formas de empalme del mismo ARN también se describe en la presente memoria. Las diferentes formas de empalme se pueden detectar usando una pluralidad de sondas nanoinformadoras, cada una con una secuencia diferente específica del objetivo complementaria a un exón diferente del mismo gen.

Además de las capacidades analíticas cualitativas proporcionadas por los nanoinformadores y las técnicas analíticas basadas en ellas, los nanoinformadores son especialmente adecuados para realizar análisis cuantitativos. Al proporcionar una unión uno a uno entre los nanoinformadores (ya sean nanoinformadores singulares o duales) y sus moléculas objetivo en una muestra biomolecular, se pueden identificar y contar todas o una porción representativa de las moléculas objetivo 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 objetivo en la muestra biomolecular. Además, la capacidad de abordar cada molécula en una mezcla individualmente aprovecha los beneficios de la miniaturización que incluyen alta sensibilidad, requerimientos mínimos de cantidad de muestra, altas velocidades de reacción que son proporcionadas por la cinética de la fase de solución en un volumen pequeño, y en última instancia muy bajos costos de reactivos.

Como se apreciará a partir de la descripción y ejemplos proporcionados a continuación, se proporcionan numerosas ventajas. Por ejemplo, la modularidad compleja en la formación de nanoinformadores permite la creación sistemática de bibliotecas de nanoinformadores únicos que tienen un grado muy alto de diversidad (por ejemplo, millones de nanoinformadores reconocibles de forma única). Esta modularidad permite flexibilidad en la personalización de poblaciones de nanoinformadores para aplicaciones específicas que a su vez proporciona eficiencias de fabricación significativas. Otra ventaja que se apreciará a través de la siguiente divulgación se deriva de la flexibilidad en el montaje de los nanoinformadores. Es decir, debido a su estructura modular, los nanoinformadores se pueden ensamblar antes del envío a un punto de uso o ensamblarse en el punto de uso.

Nanoinformadores duales

Las Figuras 1 y 2 ilustran nanoinformadores duales. Cada uno de los dos componentes del nanoinformador puede estar etiquetado, de modo que el código espectral del nanoinformador se forma solo cuando los dos componentes del nanoinformador se unan al unirse el nanoinformador dual a su molécula objetivo. Sin embargo, en un nanoinformador dual, no es necesario que ambos componentes estén etiquetados. Por ejemplo, como se representa en las Figuras 1 y 2, un componente de un nanoinformador doble está etiquetado con el código del nanoinformador y el otro componente está unido a una etiqueta de afinidad (flecha) que es útil para inmovilizar el nanoinformador para el estiramiento y la visualización.

Definiciones

A menos que se defina lo contrario, todos los términos técnicos y científicos usados en la presente memoria tienen el mismo significado que los expertos en la técnica entienden comúnmente. Aunque cualquier método y materiales similares o equivalentes a los descritos en la presente memoria se puede usar en la práctica o prueba, se describen métodos y materiales preferidos. Los siguientes términos se definen a continuación.

Los artículos "un" y "uno, una" se usan en la presente memoria para referirse a uno o a más de uno (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 "complementario" con respecto a polinucleótidos se refieren a polinucleótidos relacionados por las reglas de emparejamiento 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 coinciden de acuerdo con las reglas de emparejamiento de bases. O bien, 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 eficacia y la fuerza de la 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 no limitantes de polinucleótidos: regiones codificantes o no codificantes de un gen o fragmento de gen, loci (locus) definidos a partir del análisis de ligamiento, exones, intrones, ARN mensajero (ARNm), ARN de transferencia, ARN ribosómico, ribozimas, ADNc, polinucleótidos recombinantes, polinucleótidos ramificados, plásmidos, vectores, ADN aislado de cualquier secuencia, ARN aislado de cualquier secuencia, sondas de ácidos nucleicos y cebadores. Un polinucleótido puede comprender nucleótidos modificados, tales como nucleótidos metilados y análogos de nucleótidos. Si está presente, pueden impartirse modificaciones a la estructura de nucleótidos antes o después del ensamblaje del polímero. La secuencia de nucleótidos puede ser interrumpida por componentes no nucleotídicos. Un polinucleótido puede modificarse adicionalmente después de la polimerización, tal como mediante conjugación con un componente de etiquetado.

La expresión "identidad de secuencia" o, por ejemplo, que comprende una "secuencia 50% idéntica a", como se usa en el presente documento, se refiere a la extensión en que las secuencias son idénticas en una base de nucleótido por nucleótido a través de una ventana de comparación. Así, un "porcentaje de identidad de secuencia" puede calcularse comparando dos secuencias óptimamente alineadas sobre la ventana de comparación, determinando el número de posiciones en las que está presente la base de ácido nucleico idéntica (por ejemplo, A, T, C, G, I) en ambas secuencias para obtener el número de posiciones coincidentes, dividiendo el número de posiciones coincidentes por el número total de posiciones en la ventana de comparación (por ejemplo, el tamaño de ventana) y

multiplicando el resultado por 100 para obtener el porcentaje de identidad de secuencia.

Los términos usados para describir 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) secuencia (por ejemplo, solo una porción de la secuencia de polinucleótidos completa) que es similar entre los dos polinucleótidos, y (2) secuencia que es divergente entre los dos polinucleótidos, las comparaciones de secuencia entre dos (o más) polinucleótidos se realizan típicamente 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, habitualmente de aproximadamente 50 hasta aproximadamente 100, más habitualmente de aproximadamente 100 hasta aproximadamente 150 en donde una secuencia se compara con una secuencia de referencia del mismo número de posiciones contiguas después de que las dos secuencias están alineadas de manera óptima. La ventana de comparación puede comprender adiciones o eliminaciones (es decir, huecos) de aproximadamente 20% o menos en comparación con la secuencia de referencia (que no comprende adiciones o eliminaciones) para la alineación óptima de las dos secuencias. La alineación óptima de secuencias para alinear una ventana de comparación puede realizarse mediante implementaciones computarizadas de algoritmos (GAP, BESTFIT, FASTA y TFASTA en el paquete de software Wisconsin Genetics Release 7.0, Genetics Computer Group, 575 Science Drive Madison, WI, Estados Unidos) o mediante inspección y la mejor alineación (es decir, que da como resultado la homología porcentual más alta en la ventana de comparación) generada por cualquiera de los diversos métodos seleccionados. También se puede hacer referencia a la familia de programas BLAST tal como se describe, por ejemplo, por Altschul et al., 1997, Nucl. Acids Res. 25: 3389. Se puede encontrar una discusión detallada del análisis de secuencias en la Unidad 19.3 de Ausubel et al., "Current Protocols in Molecular Biology", John Wiley & Sons Inc., 1994-1998, Capítulo 15.

Estructura de la cadena principal

La cadena principal del nanoinformador es una molécula de ácido nucleico que contiene una pluralidad de regiones de fijación de la etiqueta (por ejemplo, al menos tres regiones de fijación de la etiqueta) dispuestas en una combinación lineal a la que pueden unirse directa o indirectamente monómeros etiqueta. La cadena principal del nanoinformador puede ser una cadena principal de ácido nucleico en la que las regiones de fijación de la etiqueta son regiones de cadena sencilla a las que pueden unirse por hibridación otros ácidos nucleicos, tales como oligonucleótidos complementarios, secuencias de ARN complementarias o secuencias de ADN complementarias. La cadena principal del nanoinformador puede ser una molécula de ácido nucleico de cadena sencilla.

La cadena principal de la presente invención puede ser ADN. Las estructuras basadas en ADN ofrecen numerosas ventajas debido, al menos en parte, al vasto universo de técnicas y metodologías existentes que permiten la manipulación de construcciones de ADN. Como se indicó anteriormente, la cadena principal puede ser de cadena sencilla.

Cada cadena principal está compuesta por una disposición única de regiones de fijación de la etiqueta. Las regiones de fijación de la etiqueta de una estructura principal de nanoinformador variarán en tamaño dependiendo del método de etiquetado. En diversas realizaciones, una región de fijación de la etiqueta puede tener una longitud en cualquier lugar de 10 nm a 10.000 nm, pero es más preferiblemente de 50 nm a 5.000 nm, y es más preferiblemente de 100 nm a 1.000 nm. En diversas realizaciones, la región de fijación de la etiqueta 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 fijación de la etiqueta corresponde estrechamente al tamaño de un punto limitado por difracción, por ejemplo, el punto más pequeño que se puede detectar con una óptica estándar, que es de aproximadamente 300 nm.

Las regiones de fijación de la etiqueta de la presente invención pueden estar constituidas por secuencias de polinucleótidos. Para las cadenas principales de ácido nucleico de la presente invención, 1 nm corresponde a aproximadamente 3 nucleótidos; por lo tanto, una región de fijación de la etiqueta de aproximadamente 300 nm corresponde a una secuencia de nucleótidos de aproximadamente 900 bases. La región de fijación de la etiqueta puede ser 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 fijación de la etiqueta 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, incluidos todos los números enteros intermedios. La región de fijación de la etiqueta puede ser una secuencia de nucleótidos de aproximadamente 1.100 nucleótidos de longitud.

Una cadena 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 fijación de la etiquetas o más, incluidos todos los números enteros intermedios. La cadena principal puede tener 6 regiones de fijación de la etiqueta. Las cadenas principales individuales pueden tener un conjunto de regiones de fijación de la etiqueta distintas, cada una de las cuales tiene una secuencia de polinucleótidos que es diferente de las secuencias de las otras regiones de fijación de la etiqueta en esa misma cadena principal.

La cadena principal puede tener 6 regiones únicas de unión de la etiqueta. La cadena principal puede comprender una pluralidad de diferentes moléculas detectables en donde dicha pluralidad de diferentes moléculas detectables en cada cadena principal tiene una señal detectable que la distingue de otras cadenas principales en la población.

5 Cada cadena principal en una población de cadenas principales se distingue de las otras por tener una combinación lineal única, orden, o disposición de regiones de fijación de la etiqueta. Además, la secuencia de polinucleótidos de cada región de fijación de la etiqueta en una cadena principal dada puede ser diferente de la secuencia de polinucleótidos de las otras regiones de fijación de la etiqueta en esa misma cadena principal. En una disposición lineal de múltiples regiones de fijación de la etiqueta en la misma cadena principal, la repetición de más de una de las mismas regiones de fijación de la etiqueta puede provocar 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, difícil de propagar. Además, la introducción de segmentos de ARN de complemento inverso en dicha cadena principal puede provocar la formación de informadores "anudados" en los que un segmento de ARN une y atrae dos regiones idénticas de fijación de la etiqueta, interrumpiendo la secuencia lineal del informador e interfiriendo con la precisión y legibilidad del código del informador, entre otros problemas. Cuando la inestabilidad del plásmido y la legibilidad del informador potencialmente pobre son un problema, estos problemas se evitan construyendo cada cadena principal individual para tener un conjunto de regiones de fijación de la etiqueta distintas, cada una de las cuales tiene una secuencia de polinucleótidos que es diferente de las secuencias de las otras regiones de fijación de la etiqueta en esa misma cadena principal.

20 Además, estas regiones de fijación de la etiqueta pueden seleccionarse de una población de secuencias de nucleótidos únicas, diseñadas racionalmente (por ejemplo, secuencias sintéticas), cada una de las cuales tiene una secuencia de polinucleótidos única representada por ciertas características (por ejemplo, contenido de G/C, relación de G/C, repeticiones de adenina), como se detalla en el presente documento, y cada una de las cuales se designa como una molécula detectable dada. Estas secuencias diseñadas racionalmente no solo mejoran la estabilidad evitando la formación de estructuras secundarias (por ejemplo, manteniendo la secuencia lineal del informador para mantener la precisión y legibilidad del código del informador), sino que optimizan la concentración y separación de las moléculas detectables unidas a cada región de fijación de la etiqueta, mejorando de este modo la legibilidad de una cadena principal etiquetada dada entre una población de cadenas principales etiquetadas de forma única.

Regiones de fijación de etiquetas

30 La presente descripción proporciona cadenas principales de nanoinformadores que son moléculas de ácido nucleico sintéticas (ADN, ARN o híbridos de ADN/ARN), diseñadas racionalmente para tener características que optimizan el etiquetado y la detección del nanoinformador, y que previenen la formación de estructuras secundarias. En algunas realizaciones de la invención, una cadena principal de nanoinformador es una secuencia de polinucleótidos diseñada que comprende una o más secuencias de 50 a 50.000 bases de longitud.

35 La cadena principal del nanoinformador puede diseñarse para minimizar las estructuras secundarias predecibles. La cadena principal del nanoinformador puede estar desprovista de cualquier estructura secundaria. Se pueden predecir estructuras secundarias putativas (por ejemplo, horquillas, plegado o emparejamiento de bases internas) mediante métodos conocidos en la técnica tales como MFOLD. Sin pretender limitarse a ninguna teoría, la estructura secundaria predecible en la estructura del nanoinformador puede minimizarse evitando las repeticiones invertidas y sesgando el contenido específico de la cadena principal de forma tal que la cadena principal sea rica en CT o GA. La cadena principal del nanoinformador puede no tener ninguna paridad intramolecular significativa en condiciones de hibridación de 75°C, 1XSSPE. La cadena principal del nanoinformador puede tener menos de aproximadamente 20, 15, 10, 5, 4, 3, 2 o 1% de repeticiones invertidas, en donde las repeticiones invertidas son 9 nucleótidos o mayores. La cadena principal puede tener menos de aproximadamente 1% de repeticiones invertidas, en donde las repeticiones invertidas son 9 nucleótidos o mayores. Los nanoinformadores no pueden contener ninguna repetición invertida de 9 nucleótidos o más en una secuencia de 1.100 pares de bases de longitud. Los nanoinformadores no pueden contener ninguna repetición invertida de 7 nucleótidos o más en cualquier región de 100 pares de bases. La estructura principal del nanoinformador puede diseñarse para sesgar el contenido específico de cadena de la cadena principal de modo que la cadena principal sea rica en CT o GA. Una región de CT o GA es una región compuesta por cualquier combinación de los nucleótidos C y T, o G y A. Por ejemplo, la región de 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. Aproximadamente 30, 35, 40, 45, 50, 55, 60, 65 o 70% de la cadena principal del nanoinformador puede estar compuesto por regiones de CT o GA. Aproximadamente del 50 al 65% de la estructura principal del nanoinformador puede estar compuesto por regiones de CT o GA. Aproximadamente el 60% de la estructura principal del nanoinformador puede estar compuesta por regiones de CT o GA. La cadena principal del nanoinformador puede diseñarse para sesgar el contenido específico de la cadena de la cadena principal de manera que la cadena principal sea rica en CT. Aproximadamente el 60% de la cadena principal del nanoinformador puede estar compuesto por regiones de CT.

60 La cadena principal del nanoinformador puede diseñarse para maximizar la singularidad de toda la secuencia en una población de nanoinformadores evitando las repeticiones directas y descartando toda secuencia que tuviera una homología significativa con cualquier otra secuencia utilizada en otras cadenas principales en una población de nanoinformadores o con cualquier secuencia descrita en la base de datos pública REFSEQ. La cadena principal del

nanoinformador puede tener menos de aproximadamente 20, 15, 10, 5, 4, 3, 2 o 1% de repeticiones directas. La cadena principal del nanoinformador puede tener menos de aproximadamente 1% de repeticiones directas, donde las repeticiones directas son 9 nucleótidos o más en una secuencia de 1.100 pares de bases de longitud o 7 nucleótidos o más en cualquier región de 100 pares de bases. Los nanoinformadores descritos en la presente memoria pueden contener menos de 85, 80, 70, 60, 50, 40, 30, 20, 10 o 5% de homología con cualquier otra secuencia utilizada en otras cadenas principales en una población de nanoinformadores o en cualquier secuencia descrita en la base de datos pública REFSEQ. Los nanoinformadores descritos en la presente memoria pueden contener menos del 60% de homología con cualquier otra secuencia utilizada en otras cadenas principales en una población de nanoinformadores o en cualquier secuencia descrita en la base de datos pública REFSEQ. Los nanoinformadores descritos en la presente memoria pueden contener menos del 50% de homología con cualquier otra secuencia utilizada en otras cadenas principales en una población de nanoinformadores o en cualquier secuencia descrita en la base de datos pública REFSEQ. La cadena principal del nanoinformador puede diseñarse para que contenga menos del 85% de homología con cualquier otra secuencia utilizada en la cadena principal o en cualquier secuencia descrita en la base de datos pública de REFSEQ. La cadena principal de los nanoinformadores puede diseñarse para que contenga menos de 20, 16, 15, 10, 9, 7, 5, 3, 2 bases contiguas de homología con cualquier otra secuencia utilizada en la cadena principal o en cualquier secuencia descrita en la base de datos pública de REFSEQ. La cadena principal de nanoinformadores puede diseñarse para tener no más de 15 bases contiguas de homología y no más del 85% de identidad en toda la longitud del nanoinformador con cualquier otra secuencia utilizada en las cadenas principales o en cualquier secuencia descrita en la base de datos pública REFSEQ.

Las regiones de fijación del etiquetador de ácido nucleico preferiblemente no tienen repeticiones directas o invertidas que sean mayores de 12 bases de longitud. Las regiones de fijación de la etiqueta de ácido nucleico pueden no tener repeticiones directas o invertidas que sean mayores de aproximadamente 7, 8, 9, 10, 11, 12, 13, 14, 15, 16 o más bases nucleotídicas de longitud. Las regiones de fijación de la etiqueta de ácido nucleico pueden no contener ninguna repetición invertida de 9 nucleótidos o más en una secuencia de 1.100 pares de bases de longitud. Las regiones de fijación de la etiqueta de ácido nucleico pueden no contener ninguna repetición invertida de 7 nucleótidos o más en ninguna región de 100 pares de bases.

Las secuencias también pueden cribarse para evitar sitios de reconocimiento de enzimas de restricción cortadoras de seis bases. Las secuencias seleccionadas pueden someterse adicionalmente al análisis de estructura secundaria predicha, y las que tienen la estructura secundaria menor se pueden elegir para una evaluación adicional. Cualquier programa conocido en la técnica se puede usar 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 cadena principal del nanoinformador, que puede ser una molécula de ácido nucleico de cadena sencilla, está diseñada para tener una o más regiones de fijación de la etiqueta, que comprende un patrón regular de una base particular (la "base repetida regularmente"). En tales regiones, la base regularmente repetida típicamente se presenta con una periodicidad mínima de aproximadamente cada n ésimo residuo, donde n es cualquier número, y preferiblemente de 4 a 25. Preferiblemente, no más del 25% de la base repetida regularmente en una región de fijación de la etiqueta aparece en un intervalo diferente a dichos intervalos regulares. Por ejemplo, si en una región de fijación de la etiqueta de 100 nucleótidos hay 12 bases de timina, y la timina es la base repetida regularmente, en este aspecto de la invención, no más del 25% de estas, es decir, 3 bases de timina, aparecen fuera del patrón regular de timinas. De manera similar, si en una región de fijación de la etiqueta de 100 nucleótidos hay 12 bases de adenina y la adenina es la base repetida regularmente, entonces no más del 25% de estas, es decir, 3 bases de adenina, aparecen fuera del patrón regular de adeninas. Preferiblemente, no más de 20%, no más de 15%, no más de 10%, no más de 9%, no más de 8%, no más de 7%, no más de 6%, no más de 5%, no más de 4%, no más del 3%, no más del 2% o no más del 1% de dicha base aparece en un intervalo diferente a dichos intervalos regulares en dicha región. Las bases pueden no separarse regularmente, pero las bases tienen al menos un promedio de 8, 9, 10, 12, 15, 16, 18, 20, 30 o 50 nucleótidos de separación. Las bases pueden no estar separadas regularmente, pero las bases están separadas por al menos 8 nucleótidos en promedio.

La base repetida regularmente en una región dada de fijación de la etiqueta puede estar 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 nucleotídicas. La base regularmente repetida puede estar separada aproximadamente cada 5 a 15 bases nucleotídicas, aproximadamente cada 6 a 12 bases nucleotídicas, aproximadamente cada 7 a 10 bases nucleotídicas, o aproximadamente cada 8 a 10 bases nucleotídicas. La base repetida regularmente puede estar separada aproximadamente cada 8 bases nucleotídicas. Una base de adenina puede repetirse aproximadamente cada 8 bases nucleotídicas. Una base de adenina puede repetirse aproximadamente cada 15 a 20 bases nucleotídicas. Una base de adenina puede repetirse aproximadamente cada 16 bases nucleotídicas.

La base repetida regularmente puede estar presente en una secuencia de polinucleótidos de una región de fijación de la etiqueta dada aproximadamente 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20% del contenido de nucleótidos. La presencia de una base repetida regularmente a un promedio de aproximadamente cada 8 a 16 nucleótidos en una región de fijación de la etiqueta representa aproximadamente 12,5% del contenido de nucleótidos. El contenido de nucleótidos de la base repetida regularmente puede ser de aproximadamente 12,5%. La región de fijación de la etiqueta puede comprender un número seleccionado o

porcentaje seleccionado de bases repetidas regularmente por región, que se distribuyen aleatoriamente, excepto 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 separadas.

5 La base repetida regularmente puede ser cualquier base nucleotídica de adenina, timina, guanina, citosina o uracilo. La base repetida regularmente en la región de fijación de la etiqueta de cadena sencilla puede ser adenina.

10 La secuencia de polinucleótidos de las regiones de fijación de la etiqueta también se puede diseñar para que tenga un cierto contenido de guanina/citosina (G/C). Por ejemplo, ciertas regiones de fijación de la etiqueta están diseñadas para tener un contenido general de G/C de aproximadamente 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, más preferiblemente aproximadamente 50%, ya que esto proporciona una T_M bastante consistente entre las numerosas secuencias de polinucleótidos únicas de las regiones de fijación de la etiqueta. En ciertos aspectos, las secuencias de polinucleótidos no tienen más de tres guaninas en una fila, para minimizar los problemas de síntesis. El contenido total de GC también es preferiblemente consistente en tramos relativamente cortos para hacer que las T_m locales sean similares. Sin pretender limitarse a ninguna teoría, cuando el nanoinformador es un ácido nucleico de cadena doble, la T_m de los nanoinformadores descritos en la presente memoria proporciona enlaces más fuertes entre la cadena principal del nanoinformador y la secuencia complementaria de polinucleótidos que tiene una molécula detectable unida, evitando así la disociación durante la síntesis y procedimientos de hibridación. Además, la T_m consistente entre la población de nanoinformadores permite que los procedimientos de síntesis e hibridación estén estrechamente optimizados, ya que las condiciones óptimas son las mismas para todos los puntos y posiciones. Las secuencias de los nanoinformadores pueden tener aproximadamente un 50% de guanina/citosina (G/C), con no más de tres G en una fila.

20 La secuencia de polinucleótidos de las regiones de fijación de la etiqueta también puede diseñarse para tener una cierta relación de G/C en una cadena dada. Por ejemplo, la secuencia de polinucleótidos de las regiones de fijación de la etiqueta se diseña de modo que las secuencias complementarias de polinucleótidos tengan una relación de G/C de aproximadamente 1/1, 2/1, 3/2, aproximadamente 5/3 o aproximadamente 7/5. La relación de G/C puede ser de al menos 1/1. La relación de G/C puede ser aproximadamente 1/1, 2/1, 3/2, 5/3 o aproximadamente 7/5 en una cadena dada. La relación de G/C puede ser al menos 1/1 en la secuencia complementaria de polinucleótidos. La relación de G/C puede ser de aproximadamente 3/2 en la secuencia complementaria de polinucleótidos.

25 Cuando la adenina es la base repetida regularmente, separada al menos aproximadamente cada 8 nucleótidos (aproximadamente 12,5% del contenido de adenina), y el contenido de G/C es aproximadamente 50%, el contenido de timina puede ser de aproximadamente 37-38% o aproximadamente 37,5%. Cuando la adenina es la base que se repite regularmente, y el contenido de G/C es de aproximadamente 50%, el contenido de timina puede ser de aproximadamente 35-45%. En una región de ejemplo de fijación de la etiqueta en la que el nucleótido repetido regularmente es adenina y un contenido de GC de aproximadamente 50%, se pueden utilizar adeninas en exceso para compensar la pérdida en abundancia de T. Para generar la secuencia seleccionada, se pueden crear y seleccionar secuencias aleatorias con patrones fijos de A que van desde cada cuarta base hasta cada vicesimoquinta base para minimizar la presencia de repeticiones invertidas y directas.

30 También se proporcionan poblaciones de nanoinformadores o unidades de etiquetas de nanoinformadores, por ejemplo nanoinformadores o bibliotecas de unidades de etiquetas de nanoinformadores, 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 nanoinformadores únicos, unidades únicas de etiquetas de nanoinformador o nanoinformadores etiquetados únicos. Como se usa en el presente documento, "único" cuando se usa en referencia a un nanoinformador o unidades de etiquetas de nanoinformadores dentro de una población pretende significar un nanoinformador que tiene un código que lo distingue de otros nanoinformadores en la misma población, por ejemplo, cada nanoinformador tiene una señal detectable que lo distingue de otros nanoinformadores en dicha población. Típicamente, un nanoinformador se vuelve "único" por la combinación lineal particular de las regiones de fijación de la etiqueta, cada una de las cuales tiene un monómero etiqueta seleccionado.

35 También se proporcionan poblaciones de nanoinformadores con al menos 1.000, 5.000, al menos 10.000, al menos 20.000 o al menos 50.000 nanoinformadores únicos o unidades de etiquetas de nanoinformadores.

Los nanoinformadores en una población de nanoinformadores pueden ser nanoinformadores singulares, nanoinformadores duales o una combinación de los mismos. Los nanoinformadores pueden etiquetarse o no etiquetarse. La población de nanoinformadores aquí descritos puede comprender dos o más nanoinformadores.

40 También se proporciona una población de cadenas principales de nanoinformadores en la que una base (por ejemplo, adenina) está regularmente separada, los nanoinformadores no tienen ninguna formación significativa de estructuras secundarias y la T_m entre los nanoinformadores en la población es bastante uniforme. También se proporciona una población de nanoinformadores en los que la T_m de las secuencias complementarias de polinucleótidos cuando se hibridan con sus regiones de fijación de la etiqueta es de aproximadamente 80°C o superior.

También se proporciona una población de cadenas principales de nanoinformadores en la que cada cadena principal del nanoinformador comprende una región que comprende una pluralidad de regiones de fijación de la etiqueta unidas covalentemente en una combinación lineal, en donde cada región de fijación de la etiqueta comprende aproximadamente 800 a 1.300 bases nucleotídicas y tiene un contenido de G/C de aproximadamente 50%, donde cada región de fijación de la etiqueta seleccionada es diferente de las otras regiones de fijación seleccionadas de la etiqueta, y en donde cada región de fijación de la etiqueta se hibrida con una secuencia complementaria de polinucleótidos que tiene una o más moléculas detectables, en donde la secuencia complementaria de polinucleótidos tiene una relación de G/C de al menos 1/1. La secuencia complementaria de polinucleótidos puede tener una relación de G/C de aproximadamente 3/2. La cadena principal del nanoinformador puede diseñarse para sesgar el contenido específico de la cadena de la cadena principal de manera que la cadena principal sea rica en CT. Aproximadamente el 60% de la cadena principal del nanoinformador puede estar compuesto por regiones de CT.

También se proporciona una población de cadenas principales de nanoinformadores en la que cada cadena principal de los nanoinformadores comprende una cadena principal de ácido nucleico de cadena sencilla, comprendiendo dicha cadena principal una pluralidad de regiones de fijación de la etiqueta unidas covalentemente entre sí en una combinación lineal única, donde dicha cadena principal de nanoinformador comprende una base (por ejemplo, adenina) separadas regularmente, las cadenas principales de nanoinformadores no tienen ninguna formación significativa de estructuras secundarias y contienen menos del 85% de homología con cualquier otra secuencia utilizada en otras cadenas principales en la población o en cualquier secuencia descrita en la base de datos pública REFSEQ.

También se proporciona una población de cadenas principales de nanoinformadores en la que cada cadena principal de nanoinformadores comprende una cadena principal de ácido nucleico de cadena sencilla, comprendiendo dicha cadena principal una pluralidad de regiones de fijación de la etiqueta unidas covalentemente entre sí en una combinación lineal única, en donde dicha cadena principal del nanoinformador tiene un contenido de G/C de aproximadamente 50%, las cadenas principales de nanoinformadores no tienen ninguna formación significativa de estructuras secundarias y contienen menos del 85% de homología con cualquier otra secuencia utilizada en otras cadenas principales de la población o en cualquier secuencia descrita en la base de datos pública de REFSEQ.

También se proporciona una población de cadenas principales de nanoinformadores en la que cada cadena principal del nanoinformador comprende una cadena principal de ácido nucleico de cadena sencilla, comprendiendo dicha cadena principal una pluralidad de regiones de fijación de la etiqueta unidas covalentemente entre sí en una combinación lineal única, en la que dicha cadena principal de nanoinformador tiene un contenido de G/C de aproximadamente el 50%, las cadenas principales de los nanoinformadores no tienen ninguna formación significativa de estructuras secundarias y contienen menos del 1% de secuencias de repetición directa.

También se proporciona una población de cadenas principales de nanoinformadores en la que cada cadena principal de nanoinformadores comprende una cadena principal de ácido nucleico de cadena sencilla, comprendiendo dicha cadena principal una pluralidad de regiones de fijación de la etiqueta unidas covalentemente entre sí en una combinación lineal única, en donde las regiones de fijación de la etiqueta se seleccionan de una población de secuencias de polinucleótidos únicas diseñadas, en donde cada secuencia de polinucleótidos está hibridada con una secuencia complementaria de polinucleótidos única que tiene fijadas a la misma una o más moléculas detectables, en donde cada nanoinformador tiene una señal detectable que la distingue de otros nanoinformadores en dicha población, y en donde cada nanoinformador es más estable que un nanoinformador comparable que comprende ADN de M13.

También se proporciona una población de cadenas principales de nanoinformadores en donde cada cadena principal de nanoinformadores comprende una cadena principal de ácido nucleico de cadena sencilla, comprendiendo dicha cadena principal una pluralidad de regiones de fijación de la etiqueta unidas covalentemente entre sí en una combinación lineal única, en donde las regiones de fijación de la etiqueta se seleccionan de una población de secuencias de polinucleótidos únicas diseñadas, en donde cada secuencia de polinucleótidos se hibrida con una secuencia complementaria de polinucleótidos única que tiene fijadas a la misma una o más moléculas detectables, en donde cada nanoinformador tiene una señal detectable que lo distingue de otros nanoinformadores en dicha población, en donde la T_m de las secuencias complementarias de polinucleótidos cuando se hibridan con sus regiones de fijación de la etiqueta en los nanoinformadores descritos en la presente memoria son mayores que la T_m de una secuencia complementaria de polinucleótidos a un molde de ADN de M13 cuando se hibrida con una sonda nanoinformadora que comprende el ADN de M13.

También se proporciona una población de cadenas principales de nanoinformadores en donde las características de secuencia de las sondas nanoinformadoras descritas en la presente memoria proporcionan una detección mejorada de una molécula objetivo. Por ejemplo, cuando la unión de las sondas nanoinformadoras a moléculas objetivo se realiza contando individualmente la presencia de una o más moléculas del nanoinformador en una muestra, las sondas nanoinformadoras descritas en la presente memoria permiten un aumento en el número de conteos. El número de conteos moleculares por encima del valor base de dicho complejo molecular después de la normalización de la muestra puede ser superior a 300, 400, 450, 500, 600, 700, 800, 900 o 1.000 conteos moleculares. Las sondas nanoinformadoras descritas en la presente memoria pueden permitir un aumento en el número de conteos. El

- número de conteos moleculares por encima del valor base de dicho complejo molecular después de la normalización de la muestra puede ser superior a 400 conteos moleculares. El porcentaje de conteos moleculares válidos de las sondas nanoinformadoras descritas en la presente memoria puede ser mayor que aproximadamente 10, 11, 12, 12,5, 14, 15, 20, 25, 30, 35, 40, 50, 60, 70, 80 o 90%. El porcentaje de conteos moleculares válidos de las sondas nanoinformadoras descritas en la presente memoria puede ser mayor que aproximadamente 12,5%. El número de conteos moleculares por encima del valor base de las sondas nanoinformadoras descritas en la presente memoria después de la normalización de dicha muestra puede ser de al menos 2, 5, 6, 10, 15, 20, 30, 40, 50, 60, 70, 80 o 90 veces mayor que un nanoinformador comparable que comprende ADN de M13. El número de conteos moleculares por encima del valor base de las sondas nanoinformadoras descritas en la presente memoria después de la normalización de dicha muestra puede ser al menos 6 veces mayor que un nanoinformador comparable que comprende ADN de M13. El número de conteos moleculares por encima del valor base de las sondas nanoinformadoras descritas en la presente memoria después de la normalización de dicha muestra puede ser al menos 2 veces mayor que un nanoinformador comparable que comprende ADN de M13.
- Una sonda nanoinformadora M13 comparable es un nanoinformador que comprende la misma región específica objetivo que las sondas nanoinformadoras descritas en la presente memoria unidas a una cadena principal de cadena sencilla que comprende una pluralidad de regiones de ADN de M13 unidas covalentemente entre sí en donde cada región se hibrida con una secuencia complementaria de polinucleótidos habiendo unido a la misma una o más moléculas detectables.
- El tamaño de una población de sondas nanoinformadoras y la naturaleza de las secuencias específicas del objetivo de los nanoinformadores dentro de ella dependerán del uso previsto del nanoinformador. Pueden elaborarse poblaciones de sondas nanoinformadoras en las que las secuencias específicas del objetivo corresponden a marcadores de un tipo de célula dado, incluido un tipo de célula enferma. La célula o tejido puede ser una célula o tejido de mamífero, y más preferentemente es una célula o tejido humano.
- Una secuencia apropiada se puede dividir en regiones de fijación de la etiqueta que varían de 50 bases a 2 kilobases de longitud (podrían ser más largas). Las regiones de fijación de la etiqueta también pueden variar de aproximadamente 200 bases a aproximadamente 1.800 bases, aproximadamente 400 bases a aproximadamente 1.600 bases, aproximadamente 600 bases a aproximadamente 1.500 bases, aproximadamente 800 bases a aproximadamente 1.300 bases, aproximadamente 900 bases a aproximadamente 1.200 bases, aproximadamente 1.000 bases a aproximadamente 1.150 bases, aproximadamente 1.050 bases a aproximadamente 1.150 bases, incluyendo combinaciones de estos ejemplos de intervalos y todos los números enteros intermedios (por ejemplo, 850, 950, etc.). Las regiones de fijación de la etiqueta pueden ser de aproximadamente 1.100 bases.
- Cada región de fijación de la etiqueta puede ser una secuencia única, pero contiene un número y espaciamiento constante y mínimo de adeninas, u otra base seleccionada, en relación con las otras regiones de fijación de la etiqueta en una secuencia indicadora dada. Estas regiones de fijación de la etiqueta pueden intercarse con otras regiones cuya secuencia no importa. Las regiones de fijación de la etiqueta en una cadena principal de nanoinformador pueden ser de diferentes longitudes y/o tener diferentes bases repetidas regularmente.
- Los ejemplos específicos de secuencias de polinucleótidos que pueden utilizarse como plantillas para sintetizar regiones de fijación de la etiqueta de cadena sencilla incluyen las secuencias de polinucleótidos expuestas en las SEQ ID NOS: 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 mencionada regularmente repetida es timina, que produce adenina como la regularmente repetida con base en la síntesis de las regiones de fijación de la etiqueta de cadena sencilla, típicamente en la forma de una cadena principal de nanoinformador. También se contemplan variantes de estas secuencias de polinucleótidos, que incluyen secuencias de polinucleótidos que tienen al menos 60%, 70%, 80%, 90%, 95%, 98% o más de identidad de secuencia con las secuencias de polinucleótidos indicadas en las SEQ ID NOS: 1-24, en donde tales variantes retienen las características deseadas o adecuadas de una región de fijación de la etiqueta, como se describe en mayor detalle en la presente memoria (por ejemplo, base repetida regularmente, contenido de G/C, relación de G/C en una cadena dada, Tm, falta de repeticiones invertidas). Tal como apreciará un experto habitual en la técnica, se pueden diseñar otras secuencias para que posean las características deseadas o adecuadas de una región de fijación de la etiqueta descrita en la presente memoria.
- Se puede añadir una secuencia de inicio optimizada para la transcripción por la ARN polimerasa T7, T3 o SP6 (que comienza en la posición +1 del transcrito) en el extremo 5' de cada región de fijación de la etiqueta. Los sitios de restricción se añaden opcionalmente en los límites de cada región de fijación de la etiqueta para permitir la adición o eliminación específica de regiones de fijación de la etiqueta individuales a la secuencia usando técnicas de clonación convencionales.
- El número de regiones de fijación de la etiqueta en una estructura principal de nanoinformador oscila preferiblemente de 1 a 50. El número de regiones de fijación de la etiqueta sintética en una cadena principal de nanoinformador puede variar de 1, 2, 3, 4, 5, 6, 7, 8, 9, o 10 regiones de fijación de la etiqueta diseñadas racionalmente a 15, 20, 30, 40 o 50 regiones de fijación de la etiqueta, o cualquier intervalo entre estos valores. El número de regiones de fijación de la etiqueta en una cadena principal puede ser seis.

La base repetida regularmente en la región de fijación de la etiqueta en una cadena principal de nanoinformador, o su base complementaria regularmente repetida en una secuencia (o segmento) polinucleotídica complementaria hibridada puede usarse para unir monómeros etiqueta, preferiblemente monómeros etiqueta de emisión de luz, al nanoinformador en un patrón regular y uniformemente separado para una mejor distribución de la señal del nanoinformador. Preferiblemente, cuando se marca una región de fijación de la etiqueta, al menos 20%, al menos 30%, al menos 40%, al menos 50%, al menos 60%, al menos 70%, al menos 80%, al menos 90%, al menos el 95% o al menos el 98% de las ocurrencias de la base repetida regularmente se une a al menos un monómero etiqueta emisor de luz, ya sea por unión covalente de un monómero etiqueta a una base, o por hibridación con un ácido nucleico en donde los complementos de la base regularmente repetida están etiquetados así.

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

$$\frac{\text{no. de monómeros etiquetadores}}{1.000 \text{ nucleótidos}} = \frac{A_{260}}{A_{LM}} \times \frac{9010}{EC_{LM}} \times 1000$$

donde EC_{LM} es el coeficiente de extinción para el monómeros etiqueta. A partir de esta fórmula, se puede calcular el porcentaje de ocurrencias de la base repetida regularmente que están unidas a un monómero etiqueta de emisión de luz.

La base de repetición regular preferida en una región de fijación de la etiqueta puede ser adenina, de modo que la región puede etiquetarse por hibridación con una o más secuencias complementarias de polinucleótidos (por ejemplo, segmentos de ARN) en las que la base repetida regularmente es uracilo. La adenina puede no repetirse regularmente, pero está separada al menos a un promedio de 8 bases. Esto permite el uso de UTP modificados con aminoalilo, que se encuentran fácilmente disponibles comercialmente, como sitios de unión de monómeros etiqueta, en una secuencia por lo demás aleatoria. Preferiblemente, además de la periodicidad regular de las regiones de fijación de la etiqueta, las regiones (y el ácido nucleico que las comprende) contienen una estructura secundaria mínima y una T_m bastante consistente.

Los ácidos nucleicos se pueden sintetizar químicamente usando nucleótidos naturales o nucleótidos modificados de distinta forma diseñados para aumentar la estabilidad biológica de las moléculas o para aumentar la estabilidad física del dúplex formado entre la región de fijación de la etiqueta y las secuencias o segmentos de polinucleótidos complementarios hibridados, por ejemplo, se pueden usar derivados de fosforotioato y nucleótidos sustituidos con acridina. Los ejemplos de nucleótidos modificados que pueden usarse para generar el ácido nucleico sintético incluyen 5-fluorouracilo, 5-bromouracilo, 5-clorouracilo, 5-yodouracilo, hipoxantina, xantina, 4-acetilinosina, 5-(carboxihidroximetil)uracilo, 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-metoxiuracilo, 2-metil-N6-isopenteniladenina, ácido uracil-5-oxiacético (v), wybutoxosina, pseudouracilo, queosina, 2-tiocitosina, 5-metil-2-tiouracilo, 2-tiouracilo, 4-tiouracilo, 5-metiluracilo, metiléster del ácido uracil-5-oxiacético, ácido uracilo-5-oxiacético (v), 5-metil-2-tiouracilo, 3-(3-amino-3-N-2-carboxipropil)uracilo, (acp3)w, y 2,6-diaminopurina.

Alternativamente, el ácido nucleico sintético se puede producir biológicamente usando un vector en el que se ha subclonado un ácido nucleico. Como un ejemplo, una cadena principal de ADN de cadena sencilla lineal puede prepararse a partir de un ADN de plásmido de cadena doble usando un protocolo de cuatro etapas que incluye (i) linealización del ADN de cadena doble con una enzima de restricción, (ii) desfosforilación con una fosfatasa termolábil, (iii) digestión con una segunda enzima de restricción para separar el vector de clonación de la secuencia principal, y (iv) digestión con exonucleasa lambda específica de cadena, dejando intacta solo una cadena del fragmento de cadena principal.

Las moléculas de ácido nucleico de la invención se pueden modificar en la porción de la base, porción de azúcar o cadena principal de fosfato para mejorar, por ejemplo, la estabilidad, hibridación o solubilidad de la molécula. Por ejemplo, la cadena principal de fosfato de desoxirribosa de los ácidos nucleicos se puede modificar para generar ácidos nucleicos peptídicos (véase Hyrup et al., 1996, Bioorganic & Medicinal Chemistry 40: 5-23). Como se usa en el presente documento, los términos "ácidos nucleicos peptídicos" o "PNA" se refieren a imitadores de ácidos nucleicos, por ejemplo, imitadores de ADN, en los que la estructura de fosfato de desoxirribosa está reemplazada por una estructura pseudopeptídica y solo se retienen las cuatro nucleobases naturales. Se ha demostrado que la

cadena principal neutra de los PNA permite la hibridación específica de ADN y ARN en condiciones de baja fuerza iónica. La síntesis de oligómeros de PNA puede realizarse usando protocolos estándar de síntesis de péptidos en fase sólida como se describe en Hyrup et al., 1996, *Bioorganic & Medicinal Chemistry* 4 (1): 5-23; Perry-O'Keefe et al., 1996, *Proc. Natl. Acad. Sci. USA* 93: 14670-675.

5 Preferiblemente, la secuencia nueva seleccionada puede construirse sintéticamente como ADN de cadena doble por parte de una compañía comercial de síntesis de genes y clonarse de forma orientada en un "fagémido", un vector plasmídico que contiene una región intergénica (IG) del fago M13 o f1 que contiene las secuencias que actúan en cis necesarias para la replicación del ADN y la encapsidación del fago, tales como pUC119. La orientación apropiada del inserto clonado con relación al origen de replicación del fago permite la generación de una cadena principal de
10 ADN de cadena sencilla que es el complemento inverso de las moléculas de ARN generadas por la transcripción *in vitro* para cada región de fijación de la etiqueta.

Para generar la cadena principal de ADN de cadena sencilla del nuevo informador, el fagémido se transforma en una cepa de *E. coli* que contiene un episoma F'. La infección posterior de las bacterias transformadas con un fago auxiliar tal como el mutante K07 de M13 da como resultado la secreción del fagémido que porta la nueva secuencia
15 indicadora como un fago empaquetado de cadena sencilla a partir del cual se prepara el ADN circular de cadena sencilla usando un protocolo estándar. Este ADN se linealiza y la parte del vector se escinde recortando oligonucleótidos complementarios cortos a cada extremo de la nueva secuencia informadora para generar sitios de restricción de doble cadena, seguido de tratamiento con las enzimas de restricción apropiadas. La Figura 9 muestra un ejemplo de vector plasmídico que puede usarse para clonar, propagar y generar los nanoinformadores descritos
20 en la presente memoria.

Secuencias complementarias de polinucleótidos

Las moléculas detectables, o monómeros etiqueta, que proporcionan o emiten señales que constituyen la totalidad o parte del código del nanoinformador pueden unirse a la región o regiones de fijación de la etiqueta de la cadena principal del nanoinformador a través de una estructura denominada aquí como una secuencia complementaria de
25 polinucleótidos. Las moléculas detectables pueden estar unidas directamente (por ejemplo, covalentemente o no covalentemente) a una secuencia complementaria de polinucleótidos, o pueden estar unidas indirectamente a dicha secuencia (por ejemplo, a través de un componente intermedio, tal como un ligando).

Las secuencias complementarias de polinucleótidos pueden ser de 25 nucleótidos hasta varias kilobases (por ejemplo, 5 kb) de longitud, y son preferiblemente de 50 nucleótidos a 2 kb de longitud. Las secuencias complementarias de polinucleótidos pueden tener una longitud de aproximadamente 25 a 250, 50 a 200, 50 a 150 o 50 a 100 nucleótidos. Las secuencias complementarias de polinucleótidos pueden ser de aproximadamente 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. Las secuencias complementarias de polinucleótidos pueden tener una longitud de aproximadamente 800, 850, 900, 950, 1.100, 1.150 o 1.100 nucleótidos, incluidos todos los números enteros intermedios. Las secuencias complementarias de
35 polinucleótidos pueden ser un polinucleótido complementario de ARN o un polinucleótido complementario de ADN. En una realización preferida, las secuencias complementarias de polinucleótidos son polinucleótidos complementarios de ARN.

Una molécula detectable o monómeros etiqueta puede unirse covalentemente a una secuencia complementaria de polinucleótidos antes o después de que la secuencia se una a la región de fijación de la etiqueta de una cadena principal de nanoinformador. Por ejemplo, al unir una molécula detectable a una secuencia complementaria de polinucleótidos, la etiqueta puede unirse covalentemente mediante la incorporación de un nucleótido que contiene una molécula detectable en el ácido nucleico durante su síntesis, pero antes de hibridarse con la región de fijación de la etiqueta de la cadena principal. Alternativamente, durante la síntesis de una secuencia complementaria de polinucleótidos, puede incluirse un nucleótido que contiene un grupo aceptor de molécula detectable, y la molécula detectable añadida a la secuencia complementaria de polinucleótidos después de su síntesis, antes o después de hibridarse con la región de fijación de la etiqueta de la cadena principal. Alternativamente, el monómeros etiqueta puede unirse indirectamente a la secuencia complementaria de polinucleótidos, por ejemplo, incorporando un nucleótido que contiene una molécula de unión al ligando (por ejemplo, biotina) en la secuencia complementaria de polinucleótidos durante la síntesis, y añadiendo un ligando (por ejemplo, estreptavidina) que está unido
50 covalentemente a la molécula detectable.

La secuencia complementaria de polinucleótidos puede oscilar entre 20 nucleótidos y más de 5 kb de longitud, dependiendo del método de ensamblaje del nanoinformador. Por ejemplo, cuando se ha incorporado una secuencia complementaria de polinucleótidos covalentemente en ella, uno o más monómeros etiqueta que proporcionan señales que son parte del código del nanoinformador en el contexto del nanoinformador etiquetado, la secuencia complementaria de polinucleótidos es preferiblemente de aproximadamente 100 a aproximadamente 10.000 bases, más preferiblemente de 200 a aproximadamente 2.000 bases, y aún más preferiblemente de 700 a aproximadamente 1.200 nucleótidos de longitud, o aproximadamente 1.100 pares de bases de longitud, y en general se denomina en el presente documento segmento, siendo un segmento oscuro la secuencia complementaria de polinucleótidos antes de la incorporación del monómeros etiqueta (pero, en una realización preferida, que contiene
60 sitios aceptores de monómeros etiqueta, tales como nucleótidos de aminoalilo), y un segmento coloreado que es

uno que contiene el monómeros etiqueta o monómeros etiqueta etiquetados. La T_m de un segmento cuando se hibrida con su región de fijación a la etiqueta preferiblemente es $> 80^\circ\text{C}$, más preferiblemente $> 90^\circ\text{C}$, en 825 mM de Na^+ (5X SSC).

5 Como con las regiones de fijación de la etiqueta, una secuencia complementaria de polinucleótidos puede comprender un patrón de bases repetidas regularmente o puede comprender un porcentaje similar de bases, tal como una base de repetición de guanina, adenina, timina, citosina o uracilo. Cuando la base repetida en la región de fijación de la etiqueta de cadena sencilla es adenina, la base repetida en la secuencia complementaria de polinucleótidos puede ser timina si la secuencia complementaria de polinucleótidos es un polinucleótido complementario de ADN, o uracilo si la secuencia complementaria de polinucleótidos es un polinucleótido complementario de ARN.

10 La base repetida regularmente en una secuencia complementaria de polinucleótidos dada puede separarse 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 nucleotídicas. La base regularmente repetida puede estar separada aproximadamente cada 5 a 15 bases nucleotídicas, aproximadamente cada 6 a 12 bases nucleotídicas, aproximadamente cada 7 a 10 bases nucleotídicas, o aproximadamente cada 8 a 10 bases nucleotídicas. La base repetida regularmente puede estar separada aproximadamente cada 8 bases nucleotídicas. La base repetida regularmente puede estar separada aproximadamente cada 15 a 20 bases nucleotídicas. La base repetida regularmente puede estar separada aproximadamente cada 16 bases nucleotídicas.

20 La base repetida regularmente puede ser modificada por aminoalilo. Esta modificación permite la unión fácil (por ejemplo, directa o indirecta, como anteriormente) de una molécula detectable o monómeros etiqueta a la base repetida regularmente. La base modificada con aminoalilo puede ser una base de guanina, adenina, timina, citosina o uracilo. La base repetida regularmente puede ser una base de uracilo modificada con aminoalilo.

25 La molécula detectable o el monómeros etiqueta puede unirse a la base repetida regularmente, tal como a través de la modificación con aminoalilo de esa base. La molécula detectable se puede unir a todas o casi todas las bases repetidas regularmente en una secuencia complementaria dada de polinucleótidos. La molécula detectable puede unirse a aproximadamente 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% o 100% de las bases repetidas regularmente en una secuencia complementaria dada de polinucleótidos.

30 La molécula detectable o monómeros etiqueta puede estar unida a una base que no se repite regularmente pero que está separada al menos 8 bases, tal como a través de la modificación con aminoalilo de esa base. El brillo de la secuencia complementaria de polinucleótidos puede variar según el número de bases en la secuencia o el número de bases etiquetadas en la secuencia.

35 Como un ejemplo específico, cuando la adenina se repite regularmente en la región de fijación de la etiqueta de cadena sencilla y está presente en aproximadamente cada 8 bases nucleotídicas, también puede estar presente un uracilo modificado con aminoalilo en el polinucleótido complementario de ARN a aproximadamente cada 8 bases nucleotídicas. Además, los uracilos modificados con aminoalilo pueden incorporarse aleatoriamente en las secuencias complementarias de polinucleótidos, tal como ajustando la relación 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, puede incorporarse en todos o casi todos los uracilos modificados con aminoalilo (por ejemplo, aproximadamente cada 8 bases) o puede incorporarse en las bases de uracilo modificadas con aminoalilo al azar.

40 Modificando el número de bases repetidas regularmente (por ejemplo, cada 5, 6, 7, 8, 9, 10 o más bases), e incorporando la molécula detectable en cada base repetida regularmente, en cada una de las otras bases repetidas regularmente, en cada tercera base regularmente repetida, o al azar, etc., o combinaciones de las mismas, la molécula detectable puede incorporarse en aproximadamente cada 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 o más bases nucleotídicas, dependiendo del nivel deseado de señal detectable.

45 Las secuencias complementarias de polinucleótidos pueden ser un polinucleótido de ARN que comprende una secuencia que es complementaria a las regiones de fijación de la etiqueta de cadena sencilla generadas por las secuencias de plantilla de polinucleótidos expuestas en las SEQ ID NOS: 1-24. Un complementario puede ser complementario a la secuencia de longitud completa de cualquiera de las regiones de fijación de la etiqueta generadas a partir de las secuencias de plantilla de SEQ ID NO: 1-24, o puede ser complementario a una región dentro de esa secuencia de polinucleótidos, tal como una región que comprende 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000 o más bases nucleotídicas, incluidos todos los números enteros intermedios.

50 Cuando la secuencia complementaria de polinucleótidos es un polinucleótido de ARN, los nucleótidos complementarios de ARN (por ejemplo, segmentos) para cada región de fijación de la etiqueta pueden prepararse usando métodos conocidos en la técnica, tales como protocolos de transcripción estándar *in vitro* en (IVT). Para sintetizar los nucleótidos complementarios de ARN, los cebadores de reacción en cadena de la polimerasa ("PCR") pueden diseñarse para generar una plantilla de cadena doble que comience con un promotor de ARN polimerasa (T7, T3 o SP6) directamente secuencia arriba (5') del sitio de inicio de la transcripción y terminando después del sitio

de la enzima de restricción 3'. Las plantillas para estas reacciones de IVT pueden incluir los productos de ADN sintéticos de las plantillas descritas en las SEQ ID NOS: 1-24, o secuencias variantes de las mismas, que se han clonado secuencia abajo de los promotores de la ARN polimerasa T7, T3 o SP6. Usando tales plantillas, en la transcripción *in vitro* de moléculas de ARN se puede realizar en presencia de una base repetida regularmente 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 cada base regularmente repetida (por ejemplo, U) se modifica para permitir el acoplamiento covalente de un monómero etiqueta en esa posición en la molécula de ARN.

Los transcritos modificados con aminoalilo (segmentos de aa) resultantes se pueden acoplar luego a diversos fluoróforos (por ejemplo, Alexa 488 (azul), Alexa 546 (verde), Alexa 594 (amarillo) y Alexa 647 (rojo); también pueden usarse otros colorantes). Los fluoróforos Alexa se pueden adquirir como ésteres de succinimidilo o de TFP (Invitrogen^{MR}) y se pueden acoplar a los transcritos mediante la formación de enlaces amida usando protocolos estándar. El nivel de incorporación de colorante en un segmento se correlaciona con el número de posibles sitios de unión (porciones de aminoalilo (aa)). Para hacer los segmentos de colores lo más brillantes posibles, se usa 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 etiqueta

Los nanoinformadores pueden etiquetarse con cualquiera de una variedad de monómeros etiqueta, tales como un radioisótopo, fluorocromo, colorante, enzima, nanopartícula, marcador quimioluminiscente, biotina u otro monómero conocido en la técnica que pueda detectarse directamente (por ejemplo, por emisión de luz) o indirectamente (por ejemplo, mediante unión de un anticuerpo etiquetado en forma fluorescente). Generalmente, una o más de las regiones de fijación de la etiqueta en el nanoinformador se marcan con uno o más monómeros etiqueta, y las señales proporcionadas por los monómeros etiqueta unidos a las regiones de fijación de la etiqueta de un nanoinformador constituyen un código detectable que identifica el objetivo al que se une la región específica del objetivo del nanoinformador. En ciertas realizaciones, la falta de una señal dada desde la región de fijación de la etiqueta (por ejemplo, un punto oscuro) también puede constituir parte del código del nanoinformador.

Preferiblemente, tal como cuando una cadena principal de nanoinformador se construye a partir de una biblioteca (por ejemplo, las plantillas de regiones de fijación de la etiqueta a modo de ejemplo de las SEQ ID NOS: 1-24), cada región de fijación de la etiqueta única puede asignarse a una molécula detectable dada. Simplemente a modo de ilustración específica, en una biblioteca de las 24 regiones únicas de fijación de etiquetas, por ejemplo, productos para plantillas de las SEQ ID NOS: 1-24, los productos de plantillas de las SEQ ID NOS: 1, 5, 9, 13, 17 y 21 pueden asignarse a una primera molécula detectable, tal como un fluoróforo azul, los productos de plantillas de las SEQ ID NOS: 2, 6, 10, 14, 18 y 22 pueden asignarse a una segunda molécula detectable, tal como un fluoróforo verde, los productos de las plantillas de las SEQ ID NOS: 3, 7, 11, 19 y 23 pueden asignarse a una tercera molécula detectable, tal como un fluoróforo amarillo, y los productos de plantillas de las SEQ ID NOS: 4, 8, 12, 20 y 24 pueden asignarse a una cuarta molécula detectable, tal como un fluoróforo rojo. Una persona experta en la técnica apreciará que este ejemplo es meramente ilustrativo, y que se pueden utilizar numerosas combinaciones o alternativas para lograr el objetivo de crear una población de nanoinformadores detectables únicos, incluido el uso de bibliotecas de diferentes tamaños y que tengan secuencias distintas de las expuestas en las SEQ ID NOS: 1-24.

Otro ejemplo de monómeros etiqueta que pueden utilizarse son fluoróforos. Se pueden usar varios fluoróforos como monómeros etiqueta para etiquetar nucleótidos que incluyen, por ejemplo, fluoresceína, tetrametilrodamina y Rojo Texas. Se conocen varios fluoróforos diferentes, y se siguen produciendo más, que abarcan todo el espectro. Además, se han producido diferentes formulaciones del mismo fluoróforo para diferentes aplicaciones. Por ejemplo, la fluoresceína puede usarse en su forma de isotiocianato (FITC), como isómero mixto o en formas de isómero único de succinimidil éster de carboxifluoresceína (FAM), o como formas isoméricas de diclorotriazina 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 que la fluoresceína o muy similares. Por ejemplo, el colorante Verde Oregón tiene espectros de excitación y emisión virtualmente superponibles en comparación con la fluoresceína. Otros fluoróforos, tales como Verde Rhodol y Verde Rodamina, se desplazan solo ligeramente en sus picos de emisión y también sirven funcionalmente como sustitutos de la fluoresceína. Además, se han desarrollado diferentes formulaciones o colorantes relacionados alrededor de otros fluoróforos que emiten luz en otras partes del espectro.

También están disponibles monómeros etiqueta no radiactivos y no fluorescentes. Por ejemplo, la biotina se puede unir directamente a nucleótidos y detectarse mediante unión específica y de alta afinidad a avidina o estreptavidina que se ha acoplado químicamente a una enzima que cataliza una reacción colorimétrica (tal como fosfatasa, luciferasa o peroxidasa). Los nucleótidos etiquetados con digoxigenina se pueden usar de manera similar para la detección no isotópica de ácidos nucleicos. Los nucleótidos etiquetados con biotina y digoxigenina están disponibles comercialmente.

Partículas muy pequeñas, denominadas nanopartículas, también pueden usarse como monómeros etiqueta para

marcar ácidos nucleicos. Estas partículas tienen un tamaño de 1-1.000 nm e incluyen diversas estructuras químicas tales como partículas de oro y plata y puntos cuánticos.

5 Cuando se irradian con luz blanca incidente en ángulo, las nanopartículas de plata o de oro que varían de 40-120 nm dispersarán la luz monocromática con alta intensidad. La longitud de onda de la luz dispersa depende del tamaño de la partícula. Cuatro a cinco partículas diferentes muy próximas dispersarán cada una luz monocromática que, al superponerse, dará un color específico y único. Las partículas están siendo fabricadas por compañías como Genicon Sciences. Las partículas de plata u oro que forman derivados se pueden unir a una amplia gama de moléculas, incluidas proteínas, anticuerpos, moléculas pequeñas, ligandos de receptores y ácidos nucleicos. Por ejemplo, la superficie de la partícula puede formar químicamente derivados para permitir la unión a un nucleótido.

10 Otro tipo de nanopartículas que se pueden usar como monómeros etiqueta son puntos cuánticos. Los puntos cuánticos son cristales fluorescentes de 1-5 nm de diámetro que son excitables por un amplio rango de longitudes de onda de luz. Estos cristales emiten luz, como la luz monocromática, con una longitud de onda que depende de su composición química y tamaño. Los puntos cuánticos como CdSe, ZnSe, InP o InAs poseen propiedades ópticas únicas.

15 Se pueden crear muchas docenas de clases de partículas según el número de clases de tamaño de los cristales de puntos cuánticos. Las clases de tamaño de los cristales se crean 1) mediante un control estricto de los parámetros de formación de cristales para crear cada clase de partículas de tamaño deseado, o 2) mediante la creación de lotes de cristales bajo parámetros de formación de cristales poco controlados, seguido de clasificación según el tamaño deseado y/o longitudes de onda de emisión. El uso de puntos cuánticos para etiquetar partículas, en el contexto de
20 la presente invención, es nuevo, pero es antiguo en la técnica de los semiconductores. Dos ejemplos de referencias anteriores en las que los puntos cuánticos están incrustados dentro de las capas epitaxiales de silicio intrínsecas de los dispositivos semiconductores emisores/detectores de luz son las patentes de Estados Unidos Nos. 5.293.050 y 5.354.707 de Chapple Sokol, et al.

25 Una o más de las regiones de fijación de la etiqueta en el nanoinformador pueden etiquetarse con uno o más colorantes emisores de luz, conteniendo cada región de fijación de la etiqueta, directa o indirectamente, una o más moléculas detectables o monómeros etiqueta. La luz emitida por los colorantes puede ser luz visible o luz invisible, tal como luz ultravioleta o infrarroja. En realizaciones a modo de ejemplo, el colorante es un colorante de transferencia de energía de 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-naftaleno-sulfonato y 2-p-6-naftaleno-sulfonato); un colorante que tiene 3-fenil-7-isocianatocumarina; una acridina, tal como 9-isotiocianatoacridina y naranja de acridina; un pireno, un bensoxadiazol y un estilbeno; un colorante que tiene 3-(ε-carboxipentil)-3'-etil-5,5'-dimetiloxacarbocianina (CYA); 6-carboxi fluoresceína (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);
35 ALEXA Fluor^{MR}; 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'-tetracloro fluoresceína (ZOE); NAN; NED; Cy3; Cy3.5; Cy5; Cy5.5; Cy7; y Cy7.5; Alexa Fluor 350; Alexa Fluor 488; Alexa Fluor 532; Alexa Fluor 546; Alexa Fluor 568; Alexa Fluor 594; o Alexa Fluor 647.

40 Los monómeros etiqueta pueden incorporarse en un nanoinformador en diferentes etapas de su ensamblaje, o en un componente (por ejemplo, nucleótidos complementarios de ARN del nanoinformador antes de su ensamblaje en el nanoinformador).

45 Un monómero etiqueta puede unirse directamente a un nucleótido utilizando métodos bien conocidos en la técnica. Los nucleótidos también se pueden modificar químicamente o formar derivados para unir un monómero etiqueta. Por ejemplo, un monómero fluorescente tal como una molécula de fluoresceína se puede unir a dUTP (desoxiuridina-trifosfato) usando un grupo aminoalquilino de cuatro átomos. Cada monómero etiqueta está unido a un nucleótido que forma un monómero etiqueta: complejo nucleotídico.

50 Este complejo de monómero etiqueta/nucleótido se puede incorporar en ácidos nucleicos (por ejemplo, un parche de ADN o un oligonucleótido de detección) en una variedad de formas. Por ejemplo, un complejo de monómero etiqueta/nucleótido puede incorporarse en solo una ubicación dentro de un ácido nucleico o en dos o más ubicaciones dentro de un ácido nucleico.

55 Los fluoróforos reactivos con tiol y reactivos con amina están disponibles y se usan para marcar nucleótidos y biomoléculas. Generalmente, los nucleótidos se etiquetan de forma fluorescente durante la síntesis química, por ejemplo, la incorporación de aminas o tioles durante la síntesis de nucleótidos permite la adición de fluoróforos. Los nucleótidos etiquetados de forma fluorescente están disponibles comercialmente. Por ejemplo, están disponibles uridina y trifosfatos de desoxiuridina que están conjugados con diez fluoróforos diferentes que cubren el espectro.

Un nucleótido se puede unir a un monómero etiqueta en primer lugar y luego incorporarse a un ácido nucleico. Alternativamente, un ácido nucleico existente puede etiquetarse uniendo un monómero etiqueta a un nucleótido dentro del ácido nucleico. Por ejemplo, se pueden incorporar nucleótidos de UTP modificados con aminoalilo ("AA")

5 en el producto de ARN durante la transcripción. En diversas realizaciones, el 20% o más de nucleótidos de UTP en una reacción de transcripción para generar nucleótidos complementarios de ARN se modifican con AA. En diversas realizaciones, aproximadamente 10% o aproximadamente 20% a 100%, 20% a 80%, 30% a 80%, 40% a 60% o 50% a 75% de UTP en una reacción de transcripción se modifican con AA, en una realización preferida, aproximadamente 40% o aproximadamente 50% de UTP en una reacción de transcripción están modificados con AA.

Además, por ejemplo, se pueden incorporar diferentes tipos de complejos monómeros etiqueta:nucleótido en un único ácido nucleico ácido, donde un componente del código del nanoinformador comprende más de un tipo de señal.

10 Los colorantes fluorescentes que se pueden unir directamente a nucleótidos también se pueden utilizar como monómeros etiqueta. Por ejemplo, FAM, JOE, TAMRA y ROX son colorantes fluorescentes reactivos con aminos que se han unido a nucleótidos y se usan en la secuenciación automatizada de ADN. Estos nucleótidos etiquetados en forma fluorescente, por ejemplo, ROX-ddATP, ROX-ddCTP, ROX-ddGTP y ROX-ddUTP, están disponibles comercialmente.

15 Otros tipos de monómeros etiqueta que pueden usarse para marcar un nanoinformador son puntos cuánticos. Debido a su tamaño muy pequeño, los puntos cuánticos se pueden acoplar en oligonucleótidos directamente sin afectar la solubilidad o el uso del oligonucleótido. En una realización preferida, solo una molécula de oligonucleótido está acoplada a cada nanopartícula. Para sintetizar un complejo oligonucleótido-nanopartícula en una relación 1:1 por química convencional de lotes, tanto el oligonucleótido como la nanopartícula requieren un solo grupo reactivo de diferentes tipos que puedan 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. Un oligonucleótido puede formar un derivado para unir un solo amino u otro grupo funcional usando química bien conocida en la técnica. Sin embargo, cuando se forma un derivado de una nanopartícula, se cubre con un reactivo químico que da como resultado el recubrimiento de toda la superficie de la nanopartícula con varios grupos funcionales.

Etiquetas de afinidad

Se puede usar una variedad de etiquetas de afinidad conocidas en la técnica para purificar y/o inmovilizar nanoinformadores.

30 Cuando se usa una etiqueta de afinidad para inmovilizar un nanoinformador con el fin de detección o formación de imágenes, se puede mencionar aquí como un "ancla". (Véase la Figura 3). En algunas realizaciones, un ancla de biotina se une al nanoinformador, permitiendo la inmovilización del nanoinformador en un portaobjetos recubierto con estreptavidina.

Un nanoinformador etiquetado puede contener una etiqueta de afinidad en cada extremo, A1 y A2. El nanoinformador etiquetado puede inmovilizarse en una superficie a través de la unión de A1 a un compañero de afinidad inmovilizado. En ausencia de un compañero de unión de afinidad para A2, el extremo A2 del nanoinformador permanece en solución, pero en presencia de un compañero de unión de afinidad (A2'), el extremo A2 del nanoinformador también está inmovilizado. Un nanoinformador etiquetado contendrá una sola etiqueta de afinidad, A1. Otra etiqueta de afinidad, A2, puede unirse al nanoinformador por unión directa del nanoinformador a una molécula que contiene A2 (por ejemplo, si el nanoinformador es o comprende un ácido nucleico, puede hibridar directamente con otro ácido nucleico al que está unido A2). Alternativamente, cualquier etiqueta de afinidad puede unirse al nanoinformador etiquetado a través de una molécula puente, tal como el ácido nucleico puente. La Figura 3 ilustra otra realización más en la que un nanoinformador etiquetado (basado en ácido nucleico) contiene una etiqueta de afinidad, A1, en un extremo. En la Figura 3, el nanoinformador etiquetado se inmoviliza mediante la unión de A1 a un compañero de afinidad inmovilizado. El otro extremo del nanoinformador está en solución (Figura 3A), pero puede inmovilizarse mediante hibridación con un oligonucleótido complementario que contiene otra etiqueta de afinidad (A2) utilizada para inmovilizar el nanoinformador (Figura 3B). Tras la inmovilización de A1, el nanoinformador puede estirarse, o "alargarse" como se representa en la Figura 3, por ejemplo, mediante contracción-relajación eléctrica, para la separación de las regiones de fijación de la etiqueta de una manera que permita la detección del código del nanoinformador. Opcionalmente, mientras el nanoinformador está en un estado alargado, A2 se introduce y une el extremo del nanoinformador que es complementario a A2 hacia la superficie.

Se puede usar una etiqueta de afinidad para unirse a perlas u otras matrices para una variedad de aplicaciones útiles que incluyen, pero no se limitan a, la purificación.

55 Ejemplos no limitantes de etiquetas de afinidad adecuadas se proporcionan a continuación. Debe entenderse que la mayoría de las etiquetas de afinidad pueden servir para dos propósitos: como anclajes para la inmovilización de los nanoinformadores y etiquetas para la purificación de los nanoinformadores (ya sea total o solamente parcialmente ensamblados) o sus componentes.

La etiqueta de afinidad puede ser un monómero de proteína. Los ejemplos de monómeros de proteína incluyen, pero no se limitan a, las regiones constantes de inmunoglobulina (véase Petty, 1996, Metal-chelate affinity

chromatography, in *Current Protocols in Molecular Biology*, Vol. 2, Ed. Ausubel et al., Greene Publish. Assoc. & Wiley Interscience), glutatión 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 los Estados Unidos Nos. 5.496.934; 5.202.247; 5.137.819; Tomme et al., 1994, *Protein Eng.* 7: 117-123), etc.

5 Otras etiquetas de afinidad son reconocidos por compañeros de unión específicos y de esta manera facilitan el aislamiento y la inmovilización por unión de afinidad al compañero de unión, que puede inmovilizarse en un soporte sólido. Por ejemplo, la etiqueta de afinidad puede ser un epítipo y el compañero de unión un anticuerpo. Los ejemplos de tales epítipos incluyen, pero no se limitan a, el epítipo FLAG, el epítipo myc en los aminoácidos 408-439, el epítipo de hemaglutinina (HA) del virus de la gripe o digoxigenina ("DIG"). En otras realizaciones, la etiqueta de afinidad es una secuencia de proteína o aminoácido que es reconocida por otra proteína o aminoácido, por ejemplo, la avidina/estreptavidina y la biotina.

15 La etiqueta de afinidad puede ser una secuencia de nucleótidos. Se puede usar una gran variedad de secuencias de aproximadamente 8 a aproximadamente 30 bases, más preferiblemente de aproximadamente 10 a aproximadamente 20 bases, para la purificación e inmovilización de nanoinformadores, y la secuencia puede repetirse en tándem (por ejemplo, de 1 a 10 repeticiones en tándem). Dicha secuencia preferiblemente no está ampliamente representada (es decir, presente en menos del 5% de los genes, más preferiblemente, presente en menos del 3% de los genes y, lo más preferiblemente, presente en menos del 1% de los genes) en la muestra que se está ensayando (por ejemplo, cuando el nanoinformador se usa para la detección de ARN celular humano, la secuencia preferiblemente no está ampliamente representada en el genoma humano); tienen poca o ninguna estructura secundaria o autocomplementariedad ya sea internamente o con copias de sí mismo cuando están multimerizadas (es decir, todas las estructuras secundarias de la etiqueta multimerizada tienen preferiblemente una Tm menor de 25°C con NaCl 1 M); no tienen una identidad o complementariedad significativa con la cadena principal o las secuencias del segmento (es decir, la Tm de las secuencias complementarias es preferiblemente inferior a 25°C con NaCl) 0,2 M); y tienen una Tm de aproximadamente 35-65°C, más preferiblemente de aproximadamente 40-50°C, en Na⁺ 50 mM.

30 En ciertas realizaciones, se usan diferentes secuencias como etiquetas de purificación e inmovilización. En este caso, por ejemplo, la etiqueta de purificación puede ser como se describió anteriormente, pero la etiqueta de inmovilización puede estar en el intervalo de 10 a 100 bases, con una Tm hasta 95°C en Na⁺ 50 mM. Una realización alternativa sería tener la etiqueta de purificación anidada dentro de la etiqueta de inmovilización (por ejemplo, la etiqueta de afinidad comprendería una secuencia de 25 bases, de las cuales 15 bases se usan como una etiqueta de purificación y las 25 bases completas se usan como la etiqueta de inmovilización).

En ciertos casos, la etiqueta de afinidad puede usarse para marcar un nanoinformador además de purificar o inmovilizar el nanoinformador.

35 Como apreciarán los expertos en la técnica, se pueden usar muchos métodos para obtener la región codificante de las etiquetas de afinidad, que incluyen, pero no se limitan a, clonación de ADN, amplificación de ADN y métodos sintéticos. Algunas de las etiquetas de afinidad y reactivos para su detección y aislamiento están disponibles comercialmente.

Región constante

40 Los nanoinformadores pueden comprender al menos una región constante. Una región constante puede comprender una secuencia de polinucleótidos. La secuencia de polinucleótidos de una región constante puede comprender una pluralidad de secuencias repetidas de nucleótidos individuales, tales 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.

45 Una secuencia de repetición individual puede comprender 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 nucleotídicas. Una secuencia de repetición individual puede comprender aproximadamente 15 bases nucleotídicas.

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

50 La región constante se puede ligar a un cadena principal de nanoinformador después del aislamiento y la preparación de la secuencia de polinucleótidos de la cadena principal, como se describe en la presente memoria. La región constante puede representar una parte permanente de la cadena principal del polinucleótido, como si estuviera fusionada a la secuencia "codificante" de la cadena principal (es decir, la región constante se clona en la cadena principal). La región constante puede estar adyacente a un sitio de enzima de restricción para la incorporación inmediata de las regiones de unión seleccionadas de la etiqueta.

Regiones específicas del objetivo

La expresión "secuencia específica del objetivo" se refiere a una entidad molecular que es capaz de unirse a una molécula objetivo. En el contexto de un nanoinformador, la secuencia específica del objetivo se une a la cadena principal del nanoinformador.

5 La secuencia específica del objetivo generalmente es una secuencia de aminoácidos (es decir, una secuencia de polipéptidos o péptidos) o una secuencia de ácidos nucleicos.

Cuando la secuencia específica del objetivo es una secuencia de aminoácidos, la secuencia específica del objetivo puede comprender un fragmento de anticuerpo, tal como un fragmento Fab' de anticuerpo, un anticuerpo Fv de cadena sencilla.

10 La secuencia específica del objetivo es preferiblemente una secuencia de ácido nucleico, y lo más preferiblemente está dentro de un oligonucleótido que está unido covalentemente (por ejemplo, mediante ligación) o unido de forma no covalente (por ejemplo, mediante hibridación) a la cadena principal del nanoinformador. Una secuencia de ácido nucleico específica del objetivo es preferiblemente de al menos 15 nucleótidos de longitud, y más preferiblemente es de al menos 20 nucleótidos de longitud. La secuencia específica del objetivo puede ser de aproximadamente 10 a 500, de 20 a 400, de 30 a 300, de 40 a 200 o de 50 a 100 nucleótidos de longitud. La secuencia específica del objetivo puede ser de aproximadamente 30 a 70, 40 a 80, 50 a 90 o 60 a 100, 30 a 120, 40 a 140 o 50 a 150 nucleótidos de longitud.

Una secuencia de nucleótidos específica del objetivo preferiblemente tiene una Tm de aproximadamente 65-90°C para cada sonda en Na⁺ 825 mM (5X SSC), más preferiblemente aproximadamente 78-83°C.

20 La secuencia específica del objetivo de cada sonda de un nanoinformador doble puede ser de aproximadamente 35 a 100 nucleótidos (para una secuencia objetivo total de aproximadamente 70 a 200 nucleótidos, cubierta por 2 sondas), lo más preferiblemente de aproximadamente 40 a 50 nucleótidos para cada sonda (para un total de aproximadamente 80 a 100 nucleótidos).

Moléculas objetivo

25 El término "molécula objetivo" es la molécula detectada o medida por la unión de un nanoinformador etiquetado cuya secuencia o secuencias específicas del objetivo reconocen (es decir, son compañeras de unión específicas) a la misma. Preferiblemente, una molécula objetivo puede ser, pero no se limita a, cualquiera de los siguientes: ácido nucleico, péptido, un polipéptido/proteína (por ejemplo, una proteína bacteriana o viral o un anticuerpo), un lípido, un carbohidrato, 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 de cadena doble, ADN de cadena sencilla, horquillas de ADN de cadena sencillas, híbridos ADN/ARN, ARN (por ejemplo, ARNm o miARN) y horquillas de ARN. Generalmente, una molécula objetivo es una molécula que se produce de forma natural o un ADNc de una molécula que se produce de forma natural o el complemento de dicho ADNc.

35 Una molécula objetivo puede ser parte de una muestra biomolecular que contiene otros componentes o puede ser el único o principal componente de la muestra. Una molécula objetivo puede ser un componente de una célula o tejido completo, una célula o extracto de tejido, un lisado fraccionado del mismo o una molécula sustancialmente purificada. La molécula objetivo se puede unir en solución o en fase sólida, incluyendo, por ejemplo, a una superficie sólida tal como un chip, microarreglo o perla. Además, la molécula objetivo puede tener una estructura o secuencia conocida o desconocida.

40 Esa molécula objetivo puede no ser un cromosoma. La molécula objetivo no puede tener un tamaño superior a 1.000 kb (o 1 mb) de tamaño, no más de 500 kb de tamaño, no más de 250 kb de tamaño, no más de 175 kb de tamaño, no más de 100 kb de tamaño, no más de 50 kb de tamaño, no más de 20 kb de tamaño, o no más de 10 kb de tamaño. La molécula objetivo puede aislarse de su medio celular.

Muestras biomoleculares

45 Los sistemas de nanoinformador pueden usarse para detectar moléculas objetivo en cualquier muestra biomolecular. Como apreciarán los expertos en la técnica, la muestra puede comprender cualquier cantidad de elementos, que incluyen, entre otros: muestras biológicas, tales como células (que incluyen tanto células primarias como líneas celulares cultivadas), lisados celulares o extractos (incluyendo, pero sin limitarse a, extractos de ARN; ARNm purificado), tejidos y extractos de tejidos (que incluyen pero no se limitan a extractos de ARN; ARNm purificado); fluidos corporales (que incluyen, entre otros, sangre, orina, suero, linfa, bilis, líquido cefalorraquídeo, líquido intersticial, humor acuoso o vítreo, calostro, esputo, líquido amniótico, saliva, secreciones anales y vaginales, sudoración y semen, transudado, un exudado (por ejemplo, fluido obtenido de un absceso o cualquier otro sitio de infección o inflamación) o fluido obtenido de una articulación (por ejemplo, una articulación normal o una articulación afectada por una enfermedad como artritis reumatoide, osteoartritis, gota o artritis séptica) de prácticamente cualquier organismo, siendo preferidas las muestras de mamíferos y siendo particularmente preferidas las muestras humanas; muestras ambientales (que incluyen, entre otras, muestras aéreas, agrícolas, de agua y de suelo); muestras de agentes de guerra biológica; muestras de investigación que incluyen fluidos extracelulares, sobrenadantes extracelulares de cultivos celulares, cuerpos de inclusión en bacterias, compartimientos celulares,

periplasma celular, compartimiento de mitocondrias, etc.

- Las muestras biomoleculares pueden derivarse indirectamente de muestras biológicas. Por ejemplo, cuando la molécula objetivo de interés es un transcrito celular, por ejemplo, un ARN mensajero, la muestra biomolecular de la invención puede ser una muestra que contiene ADNc producido por una transcripción inversa de ARN mensajero.
- 5 En otro ejemplo, la muestra biomolecular de la invención se genera sometiendo una muestra biológica a fraccionamiento, por ejemplo, fraccionamiento de tamaño o fraccionamiento de membrana.

Las muestras biomoleculares pueden ser nativas, por ejemplo, no están sujetas a manipulación o tratamiento, o pueden tratarse, lo que puede incluir cualquier cantidad de tratamientos, incluida la exposición a agentes candidatos, incluidos fármacos, ingeniería genética (por ejemplo, la adición o eliminación de un gen), etc.

- 10 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 terrestres.

Métodos

- 15 También se proporcionan métodos para la detección y/o cuantificación de una o más moléculas objetivo. En particular, se proporcionan nanoinformadores estables que son capaces de unirse a moléculas objetivo individuales y proporcionar una detección mejorada de las moléculas objetivo. A través de los códigos de etiqueta de los nanoinformadores, la unión de las sondas nanoinformadoras a las moléculas objetivo da como resultado la identificación de las moléculas objetivo. La detección y/o cuantificación se puede realizar usando un único sistema de sonda nanoinformadora o un sistema de sonda nanoinformadora dual como se describe en la presente memoria.

- 20 También se proporcionan métodos para detectar la presencia de una molécula objetivo específica en una muestra biomolecular que comprende poner en contacto la muestra con una sonda nanoinformadora como se describe en la presente memoria bajo condiciones que permiten la unión de las secuencias específicas de objetivo en las sondas a una molécula objetivo en la muestra y detectar la señal asociada con la sonda nanoinformadora etiquetada de manera única. La muestra biomolecular puede ser una muestra biológica. La muestra biológica puede seleccionarse de una célula, un lisado celular, una muestra de tejido, un extracto de tejido o un fluido corporal. La muestra biomolecular puede ser una muestra ambiental.
- 25

- También se proporcionan métodos de detección de la presencia de una pluralidad de moléculas objetivo específicas en una muestra biomolecular que comprende poner en contacto la muestra con una pluralidad de sondas nanoinformadoras como se describe en el presente documento en condiciones que permiten la unión de las secuencias específicas del objetivo en las sondas a una molécula objetivo en la muestra, y detectar la señal asociada con las sondas nanoinformadoras etiquetadas de forma única.
- 30

- También se proporcionan métodos para determinar la presencia de al menos una molécula objetivo en una muestra, que comprende formar al menos un complejo molecular que comprende (a) al menos una molécula objetivo y (b) al menos una sonda nanoinformadora que comprende una región específica única del objetivo y una región que comprende un nanoinformador único diseñado, en el que dicho nanoinformador 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 objetivo en la muestra mediante un método que comprende contar individualmente la presencia de una o más moléculas nanoinformadoras. Los métodos pueden proporcionar la determinación de la presencia de una pluralidad de moléculas objetivo mediante un método que comprende formar una pluralidad de complejos moleculares, comprendiendo cada complejo (a) al menos una molécula objetivo y (b) al menos una sonda nanoinformadora que comprende una región específica única del objetivo y una región que comprende un nanoinformador único diseñado, en el que cada nanoinformador comprende una pluralidad de diferentes moléculas detectables; y en el que cada sonda nanoinformadora comprende una región diferente de nanoinformador. El número de conteos moleculares por encima del valor base de dicho complejo molecular después de la normalización de la muestra puede ser superior a 300, 400, 450, 500, 600, 700, 800, 900 o 1.000 conteos moleculares. El número de conteos moleculares por encima del valor base del complejo molecular después de la normalización de la muestra puede ser superior a 400 conteos moleculares. El porcentaje de conteos moleculares válidos de las sondas nanoinformadoras puede ser mayor que aproximadamente 10, 11, 12, 12,5, 14, 15, 20, 25, 30, 35, 40, 50, 60, 70, 80 o 90%. El porcentaje de conteos moleculares válidos de las sondas nanoinformadoras puede ser mayor que aproximadamente 12,5%. El número de conteos moleculares por encima del valor base del nanoinformador descrito en la presente memoria después de la normalización de dicha muestra puede ser al menos 6 veces mayor que un nanoinformador comparable que comprende ADN de M13. El número de conteos moleculares por encima del valor base del nanoinformador descrito en la presente memoria después de la normalización de dicha muestra puede ser al menos 2 veces mayor que un nanoinformador comparable que comprende ADN de M13.
- 35
- 40
- 45
- 50
- 55

También se proporcionan métodos para determinar la presencia de al menos una molécula objetivo en una muestra, que comprende (1) formar al menos un complejo molecular que comprende (a) al menos una molécula objetivo y (b) al menos una sonda que comprende una región específica única del objetivo y una región que comprende un

nanoinformador único diseñado, donde dicho nanoinformador comprende una cadena principal de ácido nucleico de cadena sencilla, comprendiendo dicha cadena principal una pluralidad de regiones de fijación de la etiqueta unidas covalentemente entre sí en una combinación lineal única, en donde cada región de fijación de la etiqueta se hibrida con una secuencia única complementaria de polinucleótidos que tiene fijadas a la misma una o más moléculas detectables, y (2) detectar individualmente 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 objetivo en la muestra. El porcentaje de conteos moleculares válidos de las sondas nanoinformadoras puede ser mayor que aproximadamente 10, 11, 12, 12,5, 14, 15, 20, 25, 30, 35, 40, 50, 60, 70, 80 o 90%. El porcentaje de conteos moleculares válidos de las sondas nanoinformadoras puede ser mayor que aproximadamente 12,5%.

El número de conteos moleculares por encima del valor base del nanoinformador descrito en la presente memoria después de la normalización de dicha muestra puede ser al menos 2 veces mayor que un nanoinformador que comprende ADN de M13. La T_m de las secuencias complementarias de polinucleótidos cuando se hibrida con sus regiones de fijación a la etiqueta puede ser mayor que la T_m de una secuencia complementaria de polinucleótidos a una plantilla de ADN de M13 cuando se hibrida con una sonda nanoinformadora que comprende el ADN de M13; la sonda del nanoinformador de M13 comprende la misma secuencia específica del objetivo de la sonda nanoinformador, y comprende una cadena principal de cadena sencilla que comprende una pluralidad de regiones de ADN de M13 unidos covalentemente entre sí en donde cada región se hibrida con una secuencia complementaria de polinucleótidos que tiene unida a la misma una o más moléculas detectables. También se proporcionan métodos para determinar la presencia de una pluralidad de moléculas objetivo mediante un método que comprende la formación de una pluralidad de complejos moleculares, comprendiendo cada complejo (a) al menos una molécula objetivo y (b) al menos una sonda nanoinformadora, en donde cada sonda nanoinformadora tiene una temperatura de fusión más alta que un nanoinformador comparable que comprende ADN de M13.

Los ejemplos se muestran en las Figuras 1 y 2.

Además de detectar una señal total generada a partir de un nanoinformador etiquetado, también se proporciona la determinación de la ubicación espacial de las señales que emanan de los monómeros etiqueta (por ejemplo, puntos) en un nanoinformador, representando cada punto la señal agregada de los monómeros etiqueta unidos a una región de fijación de la etiqueta dada. Un punto puede contener señales de la misma longitud de onda o de diferentes longitudes de onda. Por lo tanto, la naturaleza de los puntos en un nanoinformador y su ubicación constituye el código del nanoinformador.

El nanoinformador etiqueta puede capturarse y estirarse en una ubicación (véase la Figura 4). Antes de estirar un nanoinformador, el nanoinformador puede inmovilizarse en una superficie sólida usando una etiqueta de afinidad, tal como se describió anteriormente (véase la Figura 3). Un extremo de un nanoinformador se puede inmovilizar, ya sea mediante unión específica o no específica a una superficie sólida, el nanoinformador se estira, y luego el otro extremo del informador se inmoviliza, también a través de unión específica o no específica a una superficie sólida (véase la Figura 3). Los métodos para capturar, estirar e inmovilizar ácidos nucleicos son bien conocidos en la técnica. Los ejemplos de métodos que se pueden usar en los métodos descritos en la presente memoria se describen en la patente de los Estados Unidos No. 7.473.767 titulada "Methods for detection and quantification of analytes in complex mixtures", la publicación de patente de Estados Unidos No. 2007/0166708 titulada "Methods for detection and quantification of analytes in complex mixtures", la solicitud de patente de Estados Unidos No. 11/645,270 titulada "Compositions comprising oriented, immobilized macromolecules and methods for their preparation", la solicitud PCT No. US06/049274 titulada "Nanoreporters and methods of manufacturing and use thereof", y la solicitud provisional de patente de Estados Unidos No. 60/088.988 titulada "Stable nanoreporter".

El nanoinformador puede estar "congelado" en su estado alargado o extendido para facilitar la resolución del código del nanoinformadores detectando y/o formando imágenes de las señales proporcionadas (por ejemplo, emitidas) por los monómeros etiqueta unidos a un nanoinformador y sus ubicaciones relativas entre sí. Esto se describe en la patente de Estados Unidos No. 7.473.767 titulada "Methods for detection and quantification of analytes in complex mixtures", la publicación de patente de Estados Unidos No. 2007/0166708 titulada "Methods for detection and quantification of analytes in complex mixtures", la solicitud de patente de Estados Unidos No. 11/645,270 titulada "Compositions comprising oriented, immobilized macromolecules and methods for their preparation", la solicitud PCT No. US06/049274 titulada "Nanoreporters and methods of manufacturing and use thereof", y la solicitud provisional de patente de Estados Unidos No. 60/088.988 titulada "Stable nanoreporter".

El nanoinformador se puede inmovilizar al sustrato. En los métodos, el sustrato para la inmovilización puede ser cualquier sustrato capaz de unirse selectivamente al nanoinformador evidente para aquellos expertos en la técnica. Una primera porción del nanoinformador se puede inmovilizar al sustrato mediante un par de unión avidina-biotina. El nanoinformador puede comprender un porción de biotina en su primera porción. Por ejemplo, un nanoinformador de polinucleótido puede comprender un residuo de nucleótido biotinilado. En un sistema de sonda nanoinformadora dual, una de las sondas nanoinformadoras (por ejemplo, sonda nanoinformadora de captura) puede comprender un porción de biotina. Esta sonda puede estar etiquetada o sin etiquetar. Preferiblemente, cuando se usa un sistema de nanoinformador dual, la sonda nanoinformadora que comprende una porción de biotina no está etiquetada. El sustrato que comprende avidina puede ser cualquier sustrato que comprenda avidina conocido por los expertos en la técnica. Los sustratos útiles que comprenden avidina están disponibles comercialmente, incluidos TB0200 (Accelr8),

SAD6, SAD20, SAD100, SAD500, SAD2000 (Xantec), SuperAvidin (Array-It), platina de estreptavidina (catálogo #MPC 000, Xenopore) y STREPTAVIDINnslide (catálogo No. 439003, Greiner Bio-one).

La primera porción del nanoinformador puede comprender una secuencia de nucleótidos que es capaz de unirse selectivamente a una secuencia de nucleótidos en el sustrato (por ejemplo, biotina). En ciertos aspectos, la región constante (por ejemplo, copias múltiples de una secuencia de repetición de 15 bases) del nanoinformador es capaz de unirse selectivamente a una secuencia complementaria de polinucleótidos que está inmovilizada en el sustrato. Si el nanoinformador comprende cuatro copias del ejemplo de secuencia de repetición de 15 bases de 5'-GGTCTGTGTGATGTT-3' (SEQ ID NO: 25), un oligonucleótido que tiene la secuencia 5'-AACATCACACAGACC AACATCACACAGACC AACATCACACAGACC AACATCACACAGACC AGCCCTTTG-3' (SEQ ID NO: 26) se puede utilizar para inmovilizar el nanoinformador a un sustrato.

Cuando se usa un sistema de sonda nanoinformadora dual, puede usarse una primera sonda nanoinformadora de captura sin etiquetar que comprende biotina y una segunda sonda nanoinformadora etiquetada que comprende una región constante para la detección y/o cuantificación de una molécula objetivo (véanse las Figuras 1 y 2). Ambas sondas nanoinformadoras se unen a la molécula objetivo a través de su secuencia específica objetivo formando un complejo que comprende la molécula objetivo y las sondas nanoinformadoras. El complejo luego se captura e inmoviliza en un sustrato que contiene avidina tal como los descritos en la presente memoria. El complejo puede luego estirarse y extenderse como se describe en la presente memoria. La región constante en la sonda nanoinformadora se puede unir a una secuencia complementaria de polinucleótidos que se inmoviliza en el sustrato. Los ejemplos de métodos que pueden usarse en los métodos descritos en la presente memoria con respecto a sistemas de nanoinformadores duales se describen en la solicitud PCT No. US06/049274 titulada "Nanoreporters and methods of manufacturing and use thereof", y la solicitud provisional de patente de Estados Unidos No. 60/088.988 titulada "Stable nanoreporter".

Los ejemplos se muestran en las Figuras 2-4.

También se proporcionan métodos para preparar una población adecuada de cadenas principales únicas de ácido nucleico, que comprende (a) seleccionar un conjunto de regiones de fijación de la etiqueta de una biblioteca de regiones únicas de fijación de la etiqueta diseñadas, y (b) unir covalentemente el conjunto de regiones de fijación de la etiqueta entre sí en una combinación lineal que es diferente de las otras cadenas principales en la población, y repetir las etapas (a)-(b) hasta que se haya preparado una población adecuada de cadenas principales únicas de ácidos nucleicos. Las regiones de fijación de la etiqueta pueden comprender aproximadamente 800 a 1.300 bases nucleotídicas. Las regiones de fijación de la etiqueta pueden tener un contenido de G/C de aproximadamente 50%. Las bases de adenina pueden separarse en promedio a aproximadamente 8 a 16 bases en las regiones de fijación de la etiqueta. Las regiones de fijación de la etiqueta pueden estar desprovistas de estructuras secundarias. Cada región de fijación de la etiqueta seleccionada puede ser diferente de las otras regiones de unión seleccionadas de la etiqueta en el conjunto. La etapa de selección puede (a) comprender seleccionar un conjunto de 3, 4, 5, 6, 7, 8 regiones de fijación de la etiqueta. La etapa de selección puede (a) comprender seleccionar un conjunto de 6 regiones de fijación de la etiqueta. Las regiones de fijación de la etiqueta se pueden hibridar con un polinucleótido complementario de ARN que tiene unida a la misma una molécula detectable. El polinucleótido complementario de ARN puede tener una relación de G/C de al menos 1/1. El polinucleótido complementario de ARN puede tener una relación de G/C de aproximadamente 3/2.

También se proporcionan métodos para preparar una población adecuada de cadenas principales únicas de ácido nucleico, que comprende (a) seleccionar un conjunto de regiones de fijación de la etiqueta de una biblioteca de regiones únicas de fijación de la etiqueta diseñadas, en donde cada región de fijación de la etiqueta comprende aproximadamente 800 a 1.300 bases nucleotídicas y tiene un contenido de G/C de aproximadamente 50%, en donde cada región de fijación de la etiqueta comprende un patrón separado regularmente de bases de adenina, y (b) unión covalente del conjunto de regiones de fijación de la etiqueta entre sí en una combinación lineal que es diferente de las otras cadenas principales de la población, y la repetición de las etapas (a)-(b) hasta que se haya preparado una población adecuada de cadenas principales únicas de ácidos nucleicos. Las bases de adenina se pueden espaciar a aproximadamente cada 8 a 16 bases en las regiones de fijación de la etiqueta. Las regiones de fijación de la etiqueta pueden estar desprovistas de estructuras secundarias. Cada región de fijación de la etiqueta seleccionada puede ser diferente de las otras regiones de fijación de la etiqueta seleccionada en el conjunto. La etapa de selección puede (a) comprender seleccionar un conjunto de 3, 4, 5, 6, 7, 8 regiones de fijación de la etiqueta. La etapa de selección puede (a) comprender seleccionar un conjunto de 6 regiones de fijación de la etiqueta. Las regiones de fijación de la etiqueta se pueden hibridar con un polinucleótido complementario de ARN que tiene unida al mismo una molécula detectable. En algunas realizaciones, el polinucleótido complementario de ARN tiene una relación de G/C de aproximadamente 3/2.

Métodos de uso de nanoinformadores selectivamente inmovilizados, extendidos u orientados

También se proporcionan nanoinformadores alargados selectivamente inmovilizados que se pueden usar para crear códigos de barras macromoleculares con el fin de separar y detectar secuencialmente las etiquetas. Estas etiquetas separadas a lo largo de la molécula proporcionan un código único que se puede leer cuando el nanoinformador se extiende e inmoviliza. La extensión y la inmovilización selectiva pueden facilitar la decodificación del código de

barras macromolecular.

Los nanoinformadores alargados selectivamente inmovilizados pueden usarse en cualquier contexto en donde la detección o formación de imágenes de un nanoinformador podría ser útil. Se pueden usar con fines terapéuticos de y de detección de diagnóstico y pronóstico. Por ejemplo, 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. Se pueden usar para diagnosticar infecciones de patógenos, por ejemplo, infecciones por bacterias y virus intracelulares, determinando la presencia y/o cantidad de marcadores de bacterias o virus, respectivamente, en la muestra. Las composiciones y métodos se pueden usar para cuantificar moléculas objetivo cuya abundancia es indicativa de un estado biológico o condición de enfermedad, por ejemplo, marcadores sanguíneos que están sobreexpresados o subexpresados como resultado de un estado de enfermedad. Además, las composiciones y métodos pueden usarse para proporcionar información de pronóstico que ayude a determinar un curso de tratamiento para un paciente.

Detección de nanoinformadores

Los nanoinformadores se detectan por cualquier medio disponible en la técnica que sea capaz de detectar las señales específicas en un nanoinformador dado. Cuando el nanoinformador está etiquetado fluorescentemente, se puede investigar la consideración adecuada de las fuentes de excitación apropiadas. Las posibles fuentes pueden incluir, pero no están limitadas a, lámpara de arco, lámpara 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 fluorescente invertido, un microscopio epifluorescente o un microscopio confocal. Preferiblemente, se usa un microscopio que puede permitir la detección con suficiente resolución espacial para determinar la secuencia de los puntos en el nanoinformador. Por ejemplo, en una realización, se puede obtener una imagen de un nanoinformador dual hibridado con una molécula objetivo. Si, por ejemplo, los nanoinformadores están etiquetados con tres colores diferentes, Alexa 488, Cy3 y Alexa 647 (etiquetados 1, 2 y 3, respectivamente). Los colores 1, 2 y 3 se adquieren cada uno en diferentes canales y el primer y segundo registros, que se pueden ver como filas de puntos, se desplazan varios píxeles para poder mostrar cada registro individualmente.

Los ejemplos de métodos para la detección de nanoinformadores que se pueden usar en los métodos se describen en la patente de los Estados Unidos 7.473.767 titulada "Methods for detection and quantification of analytes in complex mixtures", la publicación de patente de Estados Unidos No. 2007/0166708 titulada "Methods for detection and quantification of analytes in complex mixtures", la solicitud de patente de Estados Unidos No. 11/645,270 titulada "Compositions comprising oriented, immobilized macromolecules and methods for their preparation", la solicitud PCT No. US06/049274 titulada "Nanoreporters and methods of manufacturing and use thereof", y la solicitud provisional de patente de Estados Unidos No. 60/088.988 titulada "Stable nanoreporter".

Selección del microscopio y lente objetivo

La principal consideración con respecto a la lente del objetivo del microscopio es con la resolución óptica, que está determinada por su apertura numérica (NA). En general, cuanto mayor es la NA, mejor es la resolución óptica. El NA requerido es preferiblemente al menos 1,07 basado en la relación de $\delta = 0,61 \lambda/NA$ (δ = resolución óptica y λ = longitud de onda). La cantidad de luz que se recoge por un objetivo está determinada por NA^4/Mag^2 (Mag = aumento del objetivo). Por lo tanto, con el fin de recoger la mayor cantidad de luz posible, se deben seleccionar objetivos con NA alta y aumentos bajos.

Selección de cámara CCD y técnicas de captura de imágenes

Cuando se selecciona una cámara CCD, la primera consideración es el tamaño de píxel, que determina parcialmente la resolución final del sistema de formación de imágenes. De manera óptima, la resolución óptica no debe verse comprometida por la cámara CCD. Por ejemplo, si la resolución óptica es 210-300 nm, que corresponde a 12,6-18 μm en un chip CCD después de una ampliación de 60x, para resolver y mantener la resolución óptica debe haber al menos dos píxeles para muestrear cada punto. O el tamaño de píxel del chip CCD debe ser como máximo de 6,3-9 μm .

La segunda consideración es la sensibilidad de detección que puede determinarse por muchos factores que incluyen, pero no están limitados a tamaño de píxel, eficiencia cuántica, ruido de lectura y ruido oscuro. Para lograr una alta sensibilidad, seleccione una cámara cualitativa con un gran tamaño de píxel (lo que puede proporcionar una gran área de recolección), alta eficacia cuántica y bajo nivel de ruido. Un ejemplo de cámara con estos criterios es la cámara Orca-Ag de Hamamatsu Inc. El tamaño del chip es de 1344 x 1024 píxeles; cuando se usa el objetivo 60x, el campo de visión es 144 x 110 μm^2 .

Aplicaciones de la tecnología Nanoinformador

Las composiciones y métodos pueden usarse con fines terapéuticos de detección de pronóstico y diagnósticos. La ventaja es que muchas moléculas objetivo diferentes pueden analizarse al mismo tiempo a partir de una sola muestra biomolecular usando los métodos. Esto permite, por ejemplo, realizar varias pruebas de diagnóstico en una muestra.

Los métodos descritos en la presente memoria discriminan entre secuencias de nucleótidos. La diferencia entre las secuencias de nucleótidos objetivo puede ser, por ejemplo, una única diferencia de base de ácido nucleico, una eliminación de ácido nucleico, una inserción de ácido nucleico o un reordenamiento. Tales diferencias de secuencia que implican más de una base también se pueden detectar. Los conjuntos de sondas de oligonucleótidos pueden tener sustancialmente la misma longitud para que se hibriden con las secuencias de nucleótidos objetivo en condiciones de hibridación sustancialmente similares. Como resultado, el proceso puede detectar enfermedades infecciosas, enfermedades genéticas y cáncer. También es útil en monitoreo ambiental, análisis forense y ciencia de los alimentos. Los ejemplos de análisis genéticos que pueden realizarse sobre ácidos nucleicos incluyen, por ejemplo, detección de SNP, detección de STR, análisis de expresión de ARN, metilación de promotor, expresión génica, detección de virus, subtipificación viral y resistencia a fármacos.

En el área de monitoreo ambiental, puede usarse para la detección, identificación y monitoreo de microorganismos patógenos y autóctonos en ecosistemas naturales y modificados y microcosmos, tales como en sistemas de purificación de aguas residuales municipales y depósitos de agua o en áreas contaminadas sometidas a biorremediación. También es posible detectar plásmidos que contienen genes que pueden metabolizar xenobióticos, monitorear microorganismos objetivo específicos en estudios dinámicos de población, o bien para detectar, identificar o monitorear microorganismos genéticamente modificados en el medio ambiente y en plantas industriales.

También se puede usar en una variedad de áreas forenses, incluida la identificación humana para personal militar y la investigación criminal, pruebas de paternidad y análisis de relaciones familiares, tipificación de compatibilidad HLA y análisis de sangre, espermatozoides u órganos de trasplante por contaminación.

En la industria de alimentos y piensos, la presente invención tiene una amplia variedad de aplicaciones. Por ejemplo, puede usarse para la identificación y caracterización de organismos de producción como la levadura para la producción de cerveza, vino, queso, yogurt, pan, etc. Otra área de uso es con respecto al control de calidad y certificación de productos y procesos (por ejemplo, ganado, pasteurización y procesamiento de carne) por contaminantes. Otros usos incluyen la caracterización de plantas, bulbos y semillas con fines de fitomejoramiento, la identificación de la presencia de patógenos específicos de plantas y la detección e identificación de infecciones veterinarias.

Métodos de diagnóstico/pronóstico

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

Los métodos descritos en la presente memoria se pueden usar en el diagnóstico de una afección. Como se usa en la presente memoria, el término "diagnosticar" o "diagnóstico" de una condición incluye predecir o diagnosticar la afección, determinar la predisposición a la afección, controlar el tratamiento de la enfermedad, diagnosticar una respuesta terapéutica de la enfermedad y el pronóstico de la afección, progresión de la condición y respuesta al tratamiento particular de la condición. Por ejemplo, una muestra de sangre puede analizarse de acuerdo con cualquiera de los métodos descritos en la presente memoria para determinar la presencia y/o cantidad de marcadores de un tipo de célula cancerosa en la muestra, diagnosticando o estadificando así el cáncer.

Los cánceres que pueden detectarse por el proceso generalmente implican oncogenes, genes supresores de tumores o genes implicados en la amplificación, replicación, recombinación o reparación de ADN. Ejemplos de estos incluyen: gen BRCA1, gen p53, gen APC, amplificación Her2/Neu, Bcr/Ab1, gen K-ras y virus del papiloma humano tipos 16 y 18. Se puede usar para identificar amplificaciones, eliminaciones grandes así como mutaciones puntuales y eliminaciones/inserciones pequeñas 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, carcinomas de testículos y ovarios, tumores de cabeza y cuello y neoplasmas de cuello uterino.

Las enfermedades genéticas también se pueden detectar mediante el proceso. Esto puede llevarse a cabo mediante exámenes de detección prenatales o postnatales de aberraciones cromosómicas y genéticas o enfermedades genéticas. 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, enfermedad cardíaca, enfermedades de genes únicos, tipificación de HLA, fenilcetonuria, anemia de células falciformes, enfermedad de Tay-Sachs, talasemia, síndrome de Klinefelter, enfermedad de Huntington, enfermedades autoinmunes, lipidosis, defectos de la obesidad, hemofilia, errores innatos del metabolismo y diabetes.

Alternativamente, los métodos descritos en la presente memoria pueden usarse para diagnosticar infecciones patógenas, por ejemplo infecciones por bacterias y virus intracelulares, determinando la presencia y/o cantidad de marcadores de bacterias o virus, respectivamente, en la muestra.

El proceso puede detectar una amplia variedad de enfermedades infecciosas. Típicamente, estas son causados por agentes infecciosos bacterianos, virales, parasitarios y fúngicos. La resistencia de diversos agentes infecciosos a los fármacos también se puede determinar usando la presente invención.

Los agentes infecciosos bacterianos que se pueden detectar incluyen *Escherichia coli*, *Salmonella*, *Shigella*, *Klebsiella*, *Pseudomonas*, *Listeria monocytogenes*, *Mycobacterium tuberculosis*, *Mycobacterium aviumintracellulare*, *Yersinia*, *Francisella*, *Pasteurella*, *Brucella*, *Clostridia*, *Bordetella pertussis*, *Bacteroides*, *Staphylococcus aureus*, *Streptococcus neumonia*, *B-streptococos hemolíticos*, *Corynebacteria*, *Legionella*, *Mycoplasma*, *Ureaplasma*,
 5 *Chlamydia*, *Neisseria gonorrhoea*, *Neisseria meningitidis*, *Hemophilus influenza*, *Enterococcus faecalis*, *Proteus vulgaris*, *Proteus mirabilis*, *Helicobacter pylori*, *Treponema paladio*, *Borrelia burgdorferi*, *Borrelia recurrentis*, patógenos *Rickettsial*, *Nocardia* y *Actinomicetos*.

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

Los agentes infecciosos virales que pueden detectarse incluyen virus de inmunodeficiencia humana, virus linfocitotrófico de células T humanas, 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, rhabdovirus, virus de la polio, virus toga, virus bunya, virus arena, virus de rubéola y reovirus.

15 Los agentes parasitarios que se pueden detectar incluyen *Plasmodium falciparum*, *Plasmodium malaria*, *Plasmodium vivax*, *Plasmodium ovale*, *Onchocerca volvulus*, *Leishmania*, *Trypanosoma spp.*, *Schistosoma spp.*, *Entamoeba histolytica*, *Cryptosporidium*, *Giardia spp.*, *Trichomonas spp.*, *Balantidium coli*, *Wuchereria bancrofti*, *Toxoplasma spp.*, *Enterobius vermicularis*, *Ascaris lumbricoides*, *Trichuris trichiura*, *Dracunculus medinensis*, *Trematodos*, *Diphyllobothrium latum*, *Taenia spp.*, *Pneumocystis carinii* y *Necator americanis*.

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

25 De este modo, las moléculas objetivo detectadas usando las composiciones y métodos pueden ser marcadores de pacientes (tales como un marcador de cáncer) o marcadores de infección con un agente extraño, tales como marcadores bacterianos o virales.

Debido a la naturaleza cuantitativa de los nanoinformadores, las composiciones y los métodos se pueden usar para cuantificar moléculas objetivo cuya abundancia es indicativa de un estado biológico o enfermedad, por ejemplo, marcadores sanguíneos que se sobrerregulan o subregulan como resultado de un estado de enfermedad.

30 Además, las composiciones y métodos pueden usarse para proporcionar información de pronóstico que ayude a determinar un curso de tratamiento para un paciente. Por ejemplo, la cantidad de un marcador particular para un tumor puede cuantificarse con precisión incluso a partir de una muestra pequeña 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 tratamiento más agresivo.

35 Análisis de muestras de patología

El ARN extraído de muestras de tejido incluidas en parafina fijadas con formaldehído o paraformaldehído es típicamente de baja calidad (fragmentado) y de bajo rendimiento. Esto hace que el análisis de expresión génica de genes de baja expresión en muestras de histología o tejidos de patología de archivo sea extremadamente difícil y, a menudo, totalmente inviable. La tecnología de nanoinformadores puede cubrir esta necesidad no satisfecha al
 40 permitir el análisis de cantidades muy pequeñas de ARN total de baja calidad.

Para utilizar la tecnología del nanoinformador en una aplicación de este tipo, se puede extraer ARN total de muestras de tejido incluidas en parafina fijadas con formaldehído o paraformaldehído (o similares) usando kits comercialmente disponibles tales como el kit de aislamiento de ácido nucleico total RecoverAll (Ambion) de acuerdo con los protocolos del fabricante. El ARN en tales muestras se degrada con frecuencia en fragmentos pequeños (de
 45 200 a 500 nucleótidos de longitud), y muchas muestras de histología incluidas en parafina solo producen decenas de nanogramos de ARN total. Se pueden usar pequeñas cantidades (de 5 a 100 ng) de este ARN total fragmentado directamente como material objetivo en una hibridación de nanoinformadores siguiendo las condiciones de ensayo descritas en la presente memoria.

Métodos de detección

50 Los métodos pueden usarse, entre otros, para determinar el efecto de una perturbación, que incluye compuestos químicos, mutaciones, cambios de temperatura, hormonas de crecimiento, factores de crecimiento, enfermedad o un cambio en las condiciones de cultivo en diversas moléculas objetivo, por lo tanto, identificar moléculas objetivo cuya presencia, ausencia o niveles son indicativos de estados biológicos particulares. Los métodos pueden usarse para dilucidar y descubrir componentes y rutas de estados de enfermedad. Por ejemplo, la comparación de cantidades de
 55 moléculas objetivo presentes en un tejido enfermo con tejido "normal" permite la elucidación de importantes moléculas objetivo implicadas en la enfermedad, identificando de este modo objetivos para el

descubrimiento/detección de nuevos candidatos a fármacos que pueden usarse para tratar enfermedades.

Kits que comprenden nanoinformadores

5 Se proporcionan adicionalmente kits que comprenden uno o más componentes descritos en la presente memoria. Los kits pueden comprender, por ejemplo, un sustrato y uno o más nanoinformadores extendidos u orientados, o ambos, selectivamente inmovilizados sobre el sustrato. Los kits se pueden usar para cualquier propósito evidente para los expertos en la técnica, incluidos los descritos anteriormente.

10 También se proporcionan kits para preparar al menos un nanoinformador etiquetado de forma única que comprende al menos tres regiones de fijación de la etiqueta que comprenden cada una de aproximadamente 800 a 1.300 bases nucleotídicas, un contenido de G/C de aproximadamente el 50%; y al menos tres secuencias complementarias de polinucleótidos que tienen una molécula detectable unida, en la que las secuencias complementarias de polinucleótidos tienen una relación de G/C de al menos 1/1. Las secuencias complementarias de polinucleótidos pueden tener una relación de G/C de aproximadamente 3/2. El kit puede comprender además al menos tres sondas específicas para el objetivo.

15 También se proporcionan kits útiles para la extensión y la inmovilización selectiva de nanoinformadores. Los kits pueden comprender un sustrato para inmovilización y uno o más compañeros de unión para facilitar la extensión o la inmovilización de un nanoinformador. Los compañeros de unión podrían comprender una porción útil para la extensión del nanoinformador en una fuerza apropiada. Los compañeros de unión podrían facilitar la inmovilización o la inmovilización selectiva del nanoinformador a la superficie. El kit podría comprender un nanoinformador para extensión e inmovilización. El kit podría comprender un dispositivo capaz de extender el nanoinformador.

20 Los kits pueden contener una población de nanoinformadores como se describe en la presente memoria.

Los kits pueden contener nanoinformadores previamente etiquetados, o nanoinformadores no etiquetados con uno o más componentes para etiquetar los nanoinformadores. Además, los nanoinformadores proporcionados en un kit pueden o no tener secuencias específicas del objetivo unidas previamente. Las secuencias objetivo se pueden proporcionar en el kit sin unirse a la cadena principal del nanoinformador.

25 El kit puede incluir también otros reactivos, por ejemplo, reguladores para llevar a cabo reacciones de hibridación, enlazadores, endonucleasas de restricción y ADN I ligasas.

El kit también incluirá instrucciones para usar los componentes del kit, y/o para hacer y/o usar los nanoinformadores etiquetados.

La presente invención se puede entender mejor mediante los ejemplos no limitativos proporcionados a continuación.

30 Ejemplos

Ejemplo 1: Protocolo de diseño y fabricación para una biblioteca de la cadena principal del nanoinformador De Novo 1 (DV1)

35 Para construir una biblioteca de diversas cadenas principales de nanoinformadores, las regiones de fijación de la etiqueta se seleccionaron de una biblioteca de secuencias de polinucleótidos únicas diseñadas racionalmente y se clonaron en diversas combinaciones en un vector plasmídico que tenía la secuencia de polinucleótidos indicada en la SEQ ID NO: 27. Esta secuencia del vector no termina en la cadena principal del informador final, sino que se utiliza para clonar y propagar las secuencias de la cadena principal.

40 Específicamente, cada plásmido DV1 específico de código único se construyó de seis regiones de fijación a la etiqueta de aproximadamente 1.100 pares de bases, que se seleccionaron de las secuencias de polinucleótidos de las SEQ ID NOS: 1-24. Las regiones de fijación de la etiqueta definidas por estas secuencias de polinucleótidos se clonaron en diversas combinaciones en el vector de la SEQ ID NO: 27, correspondiendo cada región de fijación de la etiqueta a una posición fija en una cadena principal dada, es decir, posición 1, posición 2, hasta la posición 6. Las regiones de fijación de la etiqueta en cada posición en una cadena principal dada eran diferentes de las otras regiones de fijación de la etiqueta en esa misma cadena principal. El extremo 3' de cada secuencia de 6 posiciones se clonó junto a 4 copias de una repetición común de 15 bases conocida como la repetición G-4 (SEQ ID NO: 26), utilizada en la purificación e inmovilización de los informadores. Se diseñaron y clonaron 972 cadenas principales de nanoinformadores, cada una con una combinación lineal única de regiones de fijación de la etiqueta, de acuerdo con las técnicas biológicas moleculares de rutina.

50 Cuando a cada región de fijación de la etiqueta se le asigna una de cuatro moléculas detectables, el uso de regiones de unión de seis etiquetas en un sistema informador de cuatro colores proporciona 4096 nanoinformadores únicos posibles. En este ejemplo, las SEQ ID NOS: 1, 5, 9, 13, 17 y 21 se les asignó un fluoróforo azul, las SEQ ID NOS: 2, 6, 10, 14, 18 y 22 se les asignó un fluoróforo verde, las SEQ ID NOS: 3, 7, 11, 15, 19 y 23 se les asignó un fluoróforo amarillo, y las SEQ ID NOS: 4, 8, 12, 16, 20 y 24 se les asignó un fluoróforo rojo. Por lo tanto, en este ejemplo, incluso si una cadena principal dada comprende 6 regiones de fijación de la etiqueta asignadas a un fluoróforo azul

(por ejemplo, las SEQ ID NOS: 1, 5, 9, 13, 17 y 21 en las posiciones 1-6), cada región individual de fijación a la etiqueta tendría una secuencia de polinucleótidos diferente de las otras regiones de fijación de la etiqueta en esa misma cadena principal.

5 Los plásmidos individuales se amplificaron en bacterias, se aislaron y se convirtieron en cadenas principales lineales de cadena sencillas. Específicamente, se construyeron cadenas principales de ADN de cadena sencilla lineales a partir de un ADN plasmídico de cadena doble usando un protocolo de cuatro etapas: (i) se linealizó ADN de cadena doble con una enzima de restricción, (ii) el ADN linealizado se desfosforiló con una fosfatasa termolábil, (iii) el ADN se digirió con una segunda enzima de restricción para separar el vector de clonación de la secuencia de la cadena principal, y (iv) la mezcla se digirió con una digestión con exonucleasa lambda específica de la cadena, dejando
10 intacta solo una cadena del fragmento de la cadena principal.

Las cadenas principales y de cadena sencilla se ligaron a secuencias de nucleótidos específicas del objetivo y se incubaron con secuencias complementarias de nucleótidos de ARN teñidas con colorante para producir nanoinformadores etiquetados.

Ejemplo 2: generación de polinucleótidos complementarios de ARN teñidos con colorante

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

20 Los plásmidos que contienen un promotor de ARN polimerasa y la secuencia de polinucleótidos de interés se linealizaron por digestión de restricción, se precipitaron con etanol y se usaron como plantillas. En este ejemplo, las secuencias expuestas en las SEQ ID NOS: 1-24 se usaron como plantillas en la reacción de IVT, generando 24 polinucleótidos únicos complementarios de ARN.

25 Se sustituyó el amino-alil-UTP (aaUTP) (Fermentas) por el UTP suministrado en el kit, o se usó una mezcla de aaUTP y el UTP suministrado. El nivel de incorporación de colorante en un polinucleótido complementario de ARN se correlaciona con el número de posibles sitios de unión de aminoalilo (aa). Para hacer los segmentos de color más brillantes posibles, se debe usar 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 polinucleótido complementario de ARN. Por ejemplo, para incorporar colorantes al 50% de la base repetida regularmente, se utilizó una mezcla 1:1 de aaUTP y UTP. La concentración final de aaUTP + UTP juntas fue la misma que la de cada uno de los otros tres nucleótidos (es decir, ATP, CTP, GTP).

30 Se dejó que la reacción de IVT transcurriera a 37°C durante 22 horas. Después de la reacción, los transcritos de ARN modificados con aminoalilo se purificaron usando un kit RNeasy (Qiagen®) siguiendo las instrucciones del fabricante.

35 La reacción de acoplamiento del colorante descrita a continuación es para una reacción de 1 mg. La precipitación con etanol se realizó sobre 1 mg de un polinucleótido de ARN modificado con aminoalilo (aa). Los colorantes sólidos (éster de TFP Alexa 488, éster de succinimidilo Alexa 546, éster de succinimidilo Alexa 594 y éster de succinimidilo Alexa 647; Invitrogen) se resuspendieron en 11 µg/µL en DMSO anhidro. Los polinucleótidos de ARN modificados con aa se resuspendieron en un volumen final de 90 µL de Na₂B₄O₇ 100 mM, pH 8,5 y se calentaron a 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. La mezcla se incubó en la oscuridad a temperatura ambiente durante 30 min-2 h.

40 Los polinucleótidos de ARN modificados con aa teñidos con colorante se purificaron a partir 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 para una concentración final de 40 nM de cada polinucleótido ("mezcla de segmento de color"). Esta mezcla se puede combinar con una población de cadenas principales de nanoinformadores para generar una población de nanoinformadores únicos, como se describe en el Ejemplo 3 a
45 continuación.

Ejemplo 3: generación de una biblioteca de nanoinformadores DV1 etiquetados

Para producir una biblioteca de moléculas de nanoinformador de DV1 etiquetados, cada cadena principal de ADN específica del código del Ejemplo 1 se ligó por separado a una región de sonda, es decir, secuencia de nucleótidos específica del objetivo, para un gen específico de interés. La sonda se ligó a la cadena principal a través de un
50 oligonucleótido, que sirve como puente entre la cadena principal y la sonda específica. Específicamente, se añadió una mezcla maestra que contenía un oligonucleótido universal que sirvió como "puente" de ligación más regulador de ligasa a pozos individuales de placas de 96 pozos que contenían sondas oligonucleotídicas específicas del objetivo (10 µM) normalizadas (35-50 bases). Después de una incubación corta a 37°C para hibridar el oligonucleótido de la sonda con la porción complementaria del oligonucleótido puente, se inició la ligación mediante
55 la adición de 1,2 pmoles de la cadena principal del nanoinformador individual de cadena sencilla por pozo, regulador de ligación adicional y ligasa T4. Las placas se incubaron a 37°C en un termociclador de 96 pozos durante 2 h. Las reacciones de ligación se desalaron mediante centrifugación a través de columnas de Sephadex G-50 en un formato

de 96 pozos. Las cadenas principales ligadas individuales se juntaron, se precipitaron con etanol y se resuspendieron en Tris 10 mM pH 8,0 a 20 nM.

Para producir nanoinformadores etiquetados con colorante, se mezcló lo siguiente: 2 mL de SSPE 20X de pH 6,5, 28,05 mL de H₂O, 5 mL de las cadenas principales anteriores reunidas del nanoinformador de DV1, 1,2 mL de etanol, y 3,75 mL de mezcla del segmento coloreado del Ejemplo 2. La mezcla se incubó a 75°C durante 2 horas. Para eliminar segmentos en exceso, la mezcla se purificó sobre una columna de oligonucleótidos acoplada al complemento inverso de la secuencia de repetición G común a todas las secuencias de cadena principal (véase la SEQ ID NO: 26 para la secuencia de complemento inversa). Para producir sondas de "captura" no etiquetadas, las cadenas principales de nanoinformadores se biotinilaron cerca o en la región 3'.

10 Ejemplo 4: Comparación de los sistemas de nanoinformadores DV1 y M13

Se elaboró una biblioteca informadora DV1 de 40 genes como se describe en los Ejemplos 1 a 3, y también se elaboró una biblioteca comparable de M13, como se describe en la solicitud de patente de Estados Unidos No. 12/100.990, incorporada aquí como referencia. Las bibliotecas de DV1 y M13 contenían secuencias de sonda idénticas (es decir, nucleótidos específicos del objetivo) y se usaron para analizar las mismas muestras, 36 lisados celulares, ensayados por triplicado. Se siguieron los protocolos estándar de hibridación, purificación y formación de imágenes, como se describe en Geiss et al., *Nature Biotechnology* 26: 317-324, 2008, incorporado aquí como referencia en su totalidad. Generalmente, las bibliotecas de nanoinformadores se hibridaron con lisados celulares que contenían ARN, los informadores en exceso se eliminaron mediante lavado, los informadores se unieron a una superficie sobre la que se inmovilizaron y estiraron los informadores, se obtuvieron imágenes de la superficie y se analizaron las imágenes.

Reacciones de hibridación

La detección de transcritos celulares se llevó a cabo en reacciones de hibridación multiplexadas, que utilizaban un sistema nanoinformador doble que tenía tanto sondas nanoinformadoras etiquetadas como sondas nanoinformadoras no etiquetadas, es decir, sondas de "captura". Cada muestra se hibridó por triplicado con las concentraciones finales de los reactivos de hibridación de la siguiente manera: 200 pM cada sonda de captura biotinilada no etiquetada (sonda de captura), 40 pM de cada sonda indicadora etiquetada, 5X SSPE (pH 7,5), 5X reactivo de Denhardt (Sigma), 100 ng/μL de ADN de esperma de salmón cortado (Sigma) y 0,1% de Tween-20. Cada 30 μL de reacción de hibridación también contenía 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.

Purificación posterior a la hibridación

Para eliminar informadores no hibridados, las reacciones se purificaron sobre perlas magnéticas (Invitrogen^{MR}) acopladas a oligonucleótidos complementarios a la secuencia de repetición 3' contenida en cada sonda de captura. Las reacciones se diluyeron primero a 1X SSPE en Tween-20/TE al 0,1% y se dejaron unir a las perlas a 22,5°C durante 30 minutos con rotación continua. Las perlas se lavaron tres veces en 150 μL de 0,1X SSPE /Tween-20 al 0,1% y los complejos hibridados se eluyeron en 100 μL de 0,1X SSPE /Tween-20 al 0,1% durante 15 minutos a 45°C. Después de la elución, las muestras se purificaron una segunda vez para eliminar las sondas de captura en exceso mediante la unión a perlas magnéticas acopladas a oligonucleótidos complementarios a la secuencia de repetición 5' contenida en cada sonda indicadora. Las eluciones de las perlas antirepetición 3' se llevaron a una concentración final de 1X SSPE mediante la adición de 50 μL de 3X SSPE/Tween-20 al 0,1% y se unieron durante 15 minutos a 22,5°C con rotación. Las perlas se lavaron como anteriormente y se eluyeron en 30 μL de 0,1X SSPE/Tween-20 al 0,1% a 45°C. Las muestras doblemente purificadas se prepararon luego para la captura como se describe a continuación.

Captura del informador NanoString, estiramiento y formación de imágenes

Se añadió a cada muestra un microlitro de dilución 1/5000 de una solución de sólidos al 0,1% de una formulación personalizada de microesferas fluorescentes Tetraspeck (Invitrogen^{MR}). Las muestras se cargaron en un dispositivo de fluidos NanoString elaborado mediante laminación de acrílico fundido maquinado 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 profundos de microfluidos 30i.tm. Las muestras fueron conducidas a través del canal por presión hidrostática y unidas específicamente por el extremo 3' biotinilado de la sonda de captura. Después de la captura, la superficie se lavó una vez con 90 μL de 1X TAE y se preparó para el estiramiento mediante la adición de 40 μL de TAE a cada pozo. Las sondas informadoras se estiraron y alinearon aplicando 160 V/cm durante 1 minuto a lo largo del canal de fluido. Los informadores estirados se inmovilizaron luego a 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 informadoras. La corriente permaneció activa durante 5 minutos, durante todo el proceso de inmovilización. Después de la inmovilización, la solución TAE se eliminó y se reemplazó por una formulación personalizada del reactivo antifotoblanqueo SlowFade (Invitrogen) para formación de imágenes.

Se obtuvieron imágenes de las platinas en un Nikon Eclipse TE2000E equipado con Perfect Focus, una lente de inmersión en aceite Apo VC 60X Plan 1.4 NA (Nikon), una fuente de luz de haluro metálico X-cite 120 (Exfo Corporation), una etapa H117 automatizada (Prior Scientific) y 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 Orca Ag CCD (Hamamatsu) bajo el control de Metamorph (Universal Imaging Corporation) o software personalizado.

Procesamiento de imágenes

El procesamiento de imágenes se realizó sobre 4 imágenes (una para cada longitud de onda) en forma FOV por FOV. El algoritmo personalizado trata cada FOV como un bloque fundamental en donde se realizan los siguientes etapas básicas: 1) identificación de puntos, 2) registro de imágenes, 3) agrupamiento espacial para producir cadenas, y 4) clasificación de las cadenas.

En la primer etapa del algoritmo, se identificaron puntos. El nivel de intensidad del valor base de cada canal se calculó y utilizó para ajustar la imagen en la señal y el valor base, donde las regiones de la señal son el resultado de una longitud de onda de luz específica observada como la función de dispersión del punto (PSF). La máscara de la señal se segmentó utilizando un algoritmo de Watershed personalizado. Las regiones segmentadas fueron luego etiquetadas, parametrizadas y filtradas para eliminar puntos no PSF. Los puntos restantes fueron archivados centralmente para su uso en el registro y llamadas de informadores.

Se realizó el registro de la imagen en cada FOV basándose en puntos archivados que corresponden a perlas fluorescentes (referencia) que se unieron a la superficie de formación de imágenes (véase captura, estiramiento y formación de imágenes del informador NanoString). Los puntos archivados se cotejaron para identificar las agrupaciones entre canales de puntos que cumplen con los requisitos de referencia (umbrales de intensidad entre canales y proporciones). Las agrupaciones que cumplían con los requisitos se archivaron como referencias. La lista final de referencias representó las transformaciones espaciales que ocurrieron entre canales durante la adquisición de imágenes. Los desplazamientos espaciales eran tan grandes como 5-6 píxeles. La transformada espacial fue resuelta para usar los centroides de referencia observados y sus centroides coincidentes (supuestos) previamente transformados ($X_2 = X_1 * T$). La transformación inversa se aplicó a todos los puntos identificados para restaurar sus centroides originales.

Después de la identificación del punto y el registro de la imagen, los puntos se ensamblaron en "cuerdas" a través del agrupamiento. En este punto, cada cuerda se filtró para eliminar cualquier punto atribuido a la señal de purga. Las cadenas filtradas se clasificaron entonces como informadoras o no informadoras. Para clasificarse como informadora, la cadena debe contener el número correcto de puntos, cumplir con los umbrales específicos de separación punto a punto (1,2-2,9 píxeles) y cumplir los requisitos aceptables de linealidad y orientación. Las agrupaciones que se clasificaron como informadores se contaron y sumaron para cada gen sobre todos los FOV.

Normalización y análisis de datos de NanoString

Para explicar las ligeras diferencias en la eficacia de hibridación y purificación, los datos se normalizaron a los conteos promedio para todos los picos de control en cada muestra. Para determinar si un gen fue "detectado" por el sistema NanoString, las mediciones por triplicado obtenidas para cada gen experimental se compararon con mediciones por triplicado para los controles negativos. Para que un gen se categorice como detectado, el recuento promedio del gen experimental debe ser mayor que el promedio de los dos controles negativos, y el valor P de la prueba T de Student debe ser inferior a 0,05.

Se muestran los resultados para un representante de tres de los 40 genes analizados, 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 general 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 conteos para cada ensayo es > 100 veces mayor usando DV1 (observe la diferencia de escala para los gráficos de M13 y DV1).

Los resultados de una comparación de la expresión de los 40 genes entre los sistemas también se muestran en dos muestras representativas (véase la Figura 6). En la Figura 6, el valor R^2 muestra una correlación de aproximadamente 90% entre la expresión relativa de todos los genes en los dos sistemas, mientras que la pendiente de la línea indica un aumento de más de 100 veces en el número real de conteos usando los informadores DV1 en comparación con los informadores M13.

Ejemplo 5: Comparación de sistemas de nanoinformadores DV1 y M13

Se elaboró una biblioteca informadora de DV1 de 148 sondas como se describe en los Ejemplos 1 a 3, y también se elaboró una biblioteca de M13 comparable, como se describe en la solicitud de patente de Estados Unidos No. 12/100.990, incorporada aquí como referencia. Las bibliotecas DV1 y M13 contenían secuencias de sonda idénticas (es decir, nucleótidos específicos del objetivo) y se usaron para analizar las mismas muestras, 26 muestras de ARN de ratón. Se siguieron los protocolos estándar de hibridación, purificación y formación de imágenes, como se describe en el Ejemplo 4.

Los resultados para la detección de la expresión de GusB se muestran en la Figura 7, respectivamente. Los resultados para la detección de expresión de Modi, Acot4 y Atg7 no se muestran. La medición de los niveles de expresión relativa de un gen dado a través de las diversas muestras fue comparable entre los dos tipos de informadores. Sin embargo, se observó un aumento promedio de 6 veces para el número absoluto de conteos medidos por los informadores DV1 en comparación con los informadores M13. La Figura 8 proporciona un gráfico que muestra los "informadores válidos", que se refiere a cadenas de puntos que el software de análisis de imágenes puede interpretar como un informador "contable", descrito como un porcentaje del número total de eventos de unión en un campo de visión. Un porcentaje significativamente mayor de informadores DV1 son contables, que conduce en este experimento a una ganancia de datos de 3 veces (de 12,5% a 38%).

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

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

Para determinar la concentración óptima de colorante fluorescente en una molécula de nanoinformador dada, se incorporaron colorantes fluorescentes en las secuencias complementarias de polinucleótidos de ARN en diversas cantidades. Las moléculas de nanoinformador que tienen diversas cantidades de incorporación de colorante se analizaron a continuación para la intensidad de señal en un ensayo de detección de nanoinformador estándar.

Específicamente, para determinar el grado óptimo de espaciamiento de nucleótidos entre cada molécula de colorante incorporada, se incorporaron moléculas de colorante fluorescente en los polinucleótidos complementarios de ARN en cada una de cada 8, 10, 12, 16 o 24 bases nucleotídicas. Esto se logró generando 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 tenían aproximadamente 1.100 bases de longitud y se ensamblaron en un único plásmido de cadena principal informador de 6.600 bases, que se convirtió en una cadena principal de cadena sencilla. Cada plantilla se transcribió para producir el ARN complementario inverso, y las alícuotas de cada ARN se acoplaron por separado a cuatro colores: azul (Alexa 488), verde (Cy3 (Amersham)), amarillo (Alexa 594) y rojo (Alexa 647). Se generó una serie de indicadores etiquetados que usan los diversos colores mediante hibridación de segmentos de ARN etiquetados con la cadena principal. Se obtuvieron imágenes de estos informadores tal como se describe en la presente memoria y se determinó el brillo del punto resultante del acoplamiento del colorante en cada espaciamiento en cada color (datos no mostrados). Las mediciones de intensidad de los puntos están en unidades arbitrarias.

Estos experimentos muestran que el brillo de los puntos se puede manipular diseñando la secuencia subyacente para que contenga una mayor cantidad o menos bases repetidas regularmente. En esta memoria, una separación de 8 nucleótidos produce los puntos más brillantes. Las separaciones de menos de 8 no se probaron debido a las predicciones de interferencia esteárica de las moléculas de colorante vecinas con separaciones menores; sin embargo, es probable que algunos puntos se hagan aún más brillantes con separaciones menores. La separación más cercana posible variará de colorante a colorante, pero 8 bases proporcionan una separación que debe acomodar todos los colorantes adecuados comercialmente. Las bases no necesitan estar separadas regularmente.

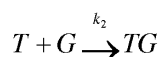
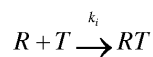
Ejemplo 7: consideraciones sobre la cinética de hibridación de los nanoinformadores

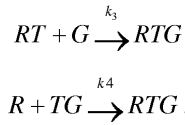
Antecedentes

Las hibridaciones en solución con un gran exceso de sonda sobre el objetivo siguen una cinética de pseudo primer orden. En este régimen, la velocidad de la reacción depende únicamente de la concentración de la sonda y no de la concentración objetivo. Para una estrategia de dos sondas y un objetivo para proporcionar información precisa sobre la concentración de un objetivo en solución, las sondas deberían estar presentes ambas en exceso con respecto al objetivo. El posible intervalo de concentración preferiblemente está limitado en el extremo inferior por la concentración del objetivo. Sin embargo, el intervalo de concentración útil para la tecnología del nanoinformador descrita en el presente documento está prácticamente limitado en el extremo inferior por la cantidad de tiempo necesaria para realizar la hibridación.

Cinética de hibridación

En realizaciones preferidas, se realizan ensayos de cuantificación y detección de objetivo en los que el objetivo (T) debe hibridarse tanto con una sonda indicadora (R) como con una sonda de captura (G) (por ejemplo, mediante 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 sean irreversibles, hay cuatro posibles reacciones elementales que ocurren.





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

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

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

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

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

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

10 en las que C_R , C_T , C_G , C_{RT} , C_{TG} y C_{RTG} son las concentraciones de las diversas especies, y k_1 - k_4 son las constantes cinéticas de las cuatro reacciones elementales. Valores de estas constantes cinéticas cuando las sondas y los objetivos son moléculas complementarios de cadena sencillas (es decir, cuando no hay etiqueta de purificación en la sonda de captura y no hay informador) pueden calcularse a partir de los datos disponibles en la literatura (Wetmur, J. Annu. Rev. Biophys. Bioeng. 1976; 5: 337-361).

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

15 En la ecuación anterior, k_N es la constante de 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 correcciones para la concentración de sal (Britten et al., 1974, Methods in Enzymology 29E: 363-406). En los sistemas de nanoinformadores descritos en la presente memoria, las constantes cinéticas dependerán de los tamaños de las etiquetas de sonda de captura y de la sonda informadora unidas. Sin estar limitados por ninguna teoría, los inventores creen que las constantes cinéticas tendrán la misma dependencia de la longitud que una reacción elemental en las constantes de difusión de los reactivos.

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

20 En la ecuación anterior D_1 y D_2 son las constantes de difusión de las dos especies que reaccionan (véanse las reacciones anteriores) y D_{50} es la constante de difusión de una molécula de ADN de cadena sencilla de 50 mer. Suponiendo un objetivo de cadena sencilla de 100 bases, una sonda de captura de cadena sencilla de 100 bases y un informador de cadena doble 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 Resolviendo numéricamente el sistema de ecuaciones diferenciales con estas constantes cinéticas (asumiendo al menos un exceso de 10 veces de sondas sobre el objetivo) se obtiene la predicción de que el informador de 5 pM y la sonda de captura de 5 pM conducirán la hibridación hasta un 10% de completarse en un reacción durante la noche (16-18 horas). A concentraciones inferiores a 5 pM, la cantidad de moléculas completamente hibridadas probablemente no sea práctica de medir. Por lo tanto, en una realización preferida, la concentración más baja de un componente nanoinformador (sonda de captura y/o sonda informadora) es de 5 pM.

40 Enlace de informadores

A medida que aumentan las concentraciones de sonda, la teoría predice que la cinética de hibridación se acelerará sin límite, siendo el único límite la solubilidad de las sondas. Sin embargo, la sonda informadora puede ser muy grande en comparación con la secuencia específica del objetivo en los sistemas nanoinformadores de la invención. Sin estar limitados por ninguna teoría, los inventores creen que, por su unión a la sonda del informador, las cinéticas de la secuencia específica del objetivo se alteran a partir de la cinética de hibridación de la solución clásica. Debido a que la sonda del informador es una molécula polimérica grande, puede tener interacciones de larga duración (entrelazamientos) con otros nanoinformadores cuando entran en contacto. A baja concentración, la probabilidad de que dos polímeros se entrelacen es pequeña, pero a medida que aumenta la concentración y/o el tamaño de un polímero en solución, estas interacciones se vuelven más y más comunes. En el caso extremo de moléculas muy largas a una concentración muy alta, los polímeros forman una red permanente, o gel, en solución. Para que se produzca la hibridación en solución, una sonda (por ejemplo, una sonda nanoinformadora)/par objetivo debe difundirse a través de la solución hasta que entren en contacto entre sí y se forme un núcleo de hibridación. Clásicamente, las reacciones de hibridación no están limitadas por difusión porque la difusión translacional de las moléculas es más rápida que la nucleación de la hibridación (es decir, la sonda y el objetivo se difunden e interactúan muchas veces antes de que ocurra una nucleación). En la solución diluida, su gran tamaño ralentizará la difusión translacional de la sonda informadora, pero puede no afectar significativamente la cinética. En alguna concentración intermedia, las sondas informadoras ocupan casi todo el espacio de la solución, formando efectivamente un gel interrelacionado permanentemente y ya no pueden difundirse en la solución. Sin embargo, la sonda de captura y los objetivos son moléculas más pequeñas que se cree que aún se difunden a través de las sondas informadoras entrelazadas, lo que permite que tenga lugar la hibridación (aunque posiblemente a un ritmo más lento). Los inventores también creen que, a una concentración más alta, la sonda indicadora en solución también impedirá el movimiento de la sonda de captura y los objetivos hasta el punto de que la reacción se vuelva limitada por 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 del objetivo) es el límite superior del intervalo de concentración útil en el sistema nanoinformador, y puede ser determinado empíricamente por un experto en la técnica, guiado por los principios descritos en la presente memoria.

Dependencia de la longitud con la cinética

Dado que la concentración superior limitante para hibridación depende tanto de la estructura del informador como de la estructura de la sonda de captura (de la que existen muchas variaciones posibles), un marco teórico para predecir las permutaciones de intervalos de concentración útiles es útil en la práctica de la invención. La teoría clásica predice que la cinética de hibridación depende únicamente del tamaño de la sonda más pequeña. Por lo tanto, la teoría predeciría que el tamaño del informador no jugará un papel en la cinética de hibridación, siempre que tanto la molécula objetivo como la sonda de captura sean significativamente más pequeñas. La teoría luego predice que la velocidad de hibridación (para una longitud de objetivo constante) depende de $1/L^{1/2}$, donde L es la longitud de la sonda de captura, debido a la inhibición estérica de la hibridación. En consecuencia, 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 debería disminuir de acuerdo con $1/L^{1/2}$. Si se supone una longitud de sonda de captura constante, entonces se puede definir el intervalo de longitudes y concentraciones del informador que darán lugar a una cantidad mensurable de eventos de hibridación. Una vez que se ha definido el tamaño de un informador, se puede determinar el rango aproximado de tamaños de la sonda de captura. Este es un proceso iterativo, pero puede proporcionar buenos puntos de partida para recopilar datos para generar pautas empíricas detalladas, dado que las teorías basadas en los fundamentos de los inventores se generaron a partir de datos de hibridación en sistemas que no emplean una sonda informadora.

Umbral de entrelazamiento

Una sonda indicadora es esencialmente un polímero en solución libre, que se comporta como una bobina aleatoria. El volumen ocupado por un único informador, V_p , se puede calcular a partir de las teorías de la física de polímeros según el modelo de cadena libremente articulada (FJC, para un polímero flexible, tal como ADN o ARN de cadena sencilla) o el modelo de cadena de tipo gusano (WLC, para un polímero rígido tal como ADN de cadena doble o un informador). Para cualquier modelo

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

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

$$R_g = b \left(\frac{N}{6} \right)^{0.6}$$

en el que b es la longitud del segmento y N es el número de segmentos en la cadena. Para el 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 del umbral de entrelazamiento definida como concentración, se define como la concentración en la que el volumen total de la solución está ocupado por los informadores.

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

5 en la que N_A es el número de Avogadro. Por encima de esta concentración, se supone que la difusión traduccional de los informadores está severamente restringida. La concentración del umbral de entrelazamiento varía con la estructura del informador. A medida que la longitud del informador aumenta, el umbral de entrelazamiento disminuye (de acuerdo con $1/L^{1.5}$). A partir de las ecuaciones anteriores, se puede calcular el umbral de entrelazamiento teórico para las sondas informadoras con diferentes tamaños de punto y diferentes longitudes. El resultado de tales cálculos muestra que para una sonda informadora híbrida de ARN/ADN de 7.200 pb con 8 regiones de fijación de la etiqueta de aproximadamente 900 pb cada una, el umbral de entrelazamiento es de aproximadamente 70 nM.

10 Si tanto el objetivo como la sonda de captura son mucho más pequeños que los informadores, entonces lo más probable es que sean libres de difundirse a través de la solución incluso a estas altas concentraciones de informadores. Los datos iniciales indican que la cinética de hibridación no disminuye apreciablemente hasta una concentración de 80 nM con una sonda informadora de 7.200 pb, un objetivo de 100 bases y una sonda de captura de 100 bases.

Efecto del umbral de entrelazamiento en multiplexación

20 Suponiendo que la concentración máxima para informadores en una reacción de hibridación es C^* , entonces la concentración de cada informador (especifica para un objetivo particular) es igual a C^*/M , donde M es el múltiplex de la reacción (número de diferentes objetivos abordados simultáneamente). Por el contrario, el posible nivel múltiplex para una estructura informadora particular se puede calcular a partir del límite inferior de la concentración de la sonda (C_p de cinética ~ 10 nM) y el umbral de entrelazamiento

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

25 Si el número de códigos de nanoinformador disponibles no depende del tamaño de la sonda informadora, entonces la multiplexación del nanoinformador depende principalmente del tamaño y la concentración de la sonda informadora (dado que es mucho más grande que la sonda de captura). Debido a que la sonda de captura contribuye de manera insignificante al entrelazamiento durante la hibridación, los inventores creen que la concentración de la sonda de captura puede incrementarse muy por encima de la concentración de la sonda informadora. En la Tabla 4 a continuación, la concentración máxima de la sonda de captura total ([G]) se establece en 1.000 nM para todas las concentraciones del informador. Esta diferencia en la concentración de la sonda de captura y la sonda informadora es un parámetro ajustable. Los experimentos preliminares muestran que en una reacción de hibridación múltiplex con un informador de 7.200 pb y sonda de captura de 100 b, 40 pM de cada sonda indicadora y 200 pM de cada sonda de captura da como resultado una hibridación casi completa en una reacción durante la noche.

Intervalos de tamaño y concentración óptimos

35 A continuación, en la Tabla 4, se presenta un resumen de los intervalos óptimos de tamaño y concentración útiles de la sonda de captura y la sonda informadora a diferentes multiplexaciones, aproximados por 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: Tamaños óptimos y rangos de concentración de sonda informadora, sonda de captura y objetivo, así como también multiplicidad de sondas, en los sistemas de nanoinformador de la invención.

Longitud del informador (pb)	Longitud de captura (b)	Mínimo [R] (pM)	Mínimo [G] (pM)	Máximo [R] (nM)	Máximo [G] (nM)	Múltiplex Máx
2.000	100	5	5	603	1.000	114.417
2.000	50	4	4	603	1.000	161.811

(continuación)

Longitud del informador (pb)	Longitud de captura (b)	Mínimo [R] (pM)	Mínimo [G] (pM)	Máximo [R] (nM)	Máximo [G] (nM)	Múltiplex Máx
2.000	200	7	7	603	1.000	80.905

ES 2 684 317 T3

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
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

<110> NANOSTRING TECHNOLOGIES, INC.

<120> NANOINFORMADORES ESTABLES

5 <130> NATE-006/D01EP 321329-2354

<140> 09791508.6

<141> 2009-08-13

<150> 61/088.988

<151> 2008-08-14

10 <160> 33

<170> PatentIn versión 3.5

<210> 1

<211> 1140

<212> ADN

<213> Secuencia Artificial

<220>

<221> fuente

5 <223> /nota = "Descripción de la secuencia artificial: polinucleótido sintético"

<400> 1

gaattcaaag ggctgggtga agcgggtgaaa cggctcgagcg ctgaaggaat gagaaagtga	60
ggcggtagac aaaagtaagc cagtggcaca gtgaggaaga tgagcgagct gaggacaatg	120
acggagtcgg aggaatcaga gcggtgagac aagtggagga tatcaaagat aagagcatag	180
ggaaatgcaa caatggaaac gtcccaaggt ggaagcgtgg gagaatgaag aggtaagcaa	240
atagaagacg tagggaacat gaaacctgc agaagataag aaaatgccag aatacgacgg	300
tgagagaaat caaccagtac aagcgtgaa cagctaccga ggtagcgaga tgaacaagat	360
gcgaacctca ggaactcaag aagtagcgaa atcgaccggg tcgggaaagt cgagaaataa	420
gaacgtacca gggatacaga actagggacg taggagggtg ggacgatacg gcgctgaaac	480
gggtgggag gtaacagggt ggaaaagtaa gagactaagg aactgaaaca gctaacaggc	540
taaggaaca tggagaaata aagacactgg agcgcagctg gaagatagag aaaatgagag	600
cgtgaaacca tgaagggat caagagtgga cggagcatag aaagctgaaa caaataggga	660
agctgaagac cataagcggg ctgccaaaga taagagagtg acaagatacg cgccgtggag	720
aagtgcagga cataaaacaa tggccgcatc aggccgggtg agggcaatac aagagctaga	780
agagtaccgc gataggaagg tggcaccagt aaggaaataa gcccatgagg acatacacga	840
gtcgaaaaat aagcagtgca aacgctaggc caactggcgg catgggacgg tgcgagggtc	900
gacagaggtg tacaagtgac aggatgaaag cataagaagg tgacgcaact agggccatac	960
aaagagtgga ccaatccaaa cctgcgaaga taggaggata acaccggtag ggcaactaca	1020
aggatcaaag gatgaaagaa taaaacta agggcgtcca acagtaccga agtcagggcg	1080
tcacaggtcg aacagaactc aaccgaagtc tagaaggatc cactcgagtt tagatctttt	1140

<210> 2

10 <211> 1141

<212> ADN

<213> Secuencia Artificial

<220>

<221> fuente

15 <223> /nota = "Descripción de la secuencia artificial: polinucleótido sintético"

<400> 2

ES 2 684 317 T3

gaattcaaag ggctgaacgc aggtacgaga gtaccgggaa tagcggagtg aagccctaaa 60
gaaataacag gaatagcaca agtaggaaca taaacagatg cacaccaata aagagatagc 120
gaaataaaga catagagaca tccgcaataa aagggatgaa acaataaccg ggtcgcgaag 180
ctaacaccca tcgccagctc gggcaaaactg atacctcga aggactagca agatggaaac 240
atggacaagt aaaagggatga agaagtacca caaactcaac agctaagaca agtcgggaag 300
taaaaaaatc caagggtaac agaatggcga agtgagcgaa agtagagcaa tgacggcatg 360
gagaaatggg aaaaaatcac gagatagacg agtgggcagc taggcgggtg accggaatac 420
gacaatgaga gactgcccga gtgcgaaagt ggcgggataa aaaagctgaa gggagtcaaa 480
ccatcagcgg gatacagaag tagaacaatg cacagatgag caggtgaagcg ggtgcaagac 540
tagaaaagct cggcaaatcg ccgaatagac aaatcagctg gtgagcgggt acgaagactg 600
gcggccataa cgccatgagg gcatagcaaa agtaccctaaa taaggcagtc agagagtgac 660
gggctacagc ggctaggccg actacagcag tgcgggaagt gaaaaaagct cgaaggatgc 720
aaccatgagc caataaaggc gtcgagagct gccaaagatgc 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 aggcatggaa acggtccggg cagtggaacg atccaagagt aggaaaaatca 1020
gaagcgtacg aaactacaag cgtcagggag tagaaaagta acacaatgaa gaaactaagc 1080
acgtgaaaga gtagaggaac taggaaacat ctagaaggat cactcagatg ttagatcttt 1140
t 1141

<210> 3

<211> 1142

<212> ADN

5 <213> Secuencia Artificial

<220>

<221> fuente

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

<400> 3

ES 2 684 317 T3

gaattcaaag ggctgacaac cactgaaggc ataaacaagt accaaagtaa cagcatggag 60
 agctgagggc gtaacaaagg tcgaaaactc aaacactaca aaaatgggac aaatagcgcc 120
 catcggaagc ctgagcggat agcggactac ggtagtaaa agatagcggg gctgaaagcc 180
 tcaggacgtg aacaaatggg aaggataacc gggtaggaggt gtgaaacaga tgagacggta 240
 agaaaagtaa gagaaggtgc gatatcgcgg gcgtcacgca acgtgcaaaa atgacggaat 300
 aaaagaatcg aggaggtcaa ggcgataagc gcgtgagaggt atagaaagat cgagccatag 360
 cgggccatca gcggcgtgcg aggagtcgca cactcaaga gctaaccgca tcagcagta 420
 cagcgggtag aagcgtcgcg ggatagagga agtccaaaga tcccgaactg ccagcgtagg 480
 aacactgacc acatagcacc atcaaaagct gaacgagatg agacactacg caggatgaga 540
 acgtcgaag catgaaccgg gtgcagagct aggacagctg cgccgatcag gccacgtagg 600
 aagatccaag cctggcacag agtcaagacg ctagaaaaat gaagaagtca gcaaagtagg 660
 gagggtgagg gcatgaacga gtgaggacat aggacgatcc caaagtgacg gaatgaccag 720
 gtgagaaagt agggcaaatc aagcagtcaa agcgtgaaga aactagaag gcgtaaaaga 780
 gtggagaagc tccgacagat acaaaagtag aggcctgaga gggtcgggca aaatcccaga 840
 ctcgagaaat cgcgacaatg caaacgtggc gcggtgggcg aggtgccgaa atcacgcgaa 900
 tggacggata tgtactactga aaaagctcac aaaataagcg gatcgggaca tcggaacctg 960
 agacgaatac agacgataaa gcaataaccg actagacgag tgccagggtg ggaacgtacc 1020
 aaaagtccag aagtccacgg gtggacagac tagagaacag tacagaaaat gccaccctaa 1080
 acgggtaagg gaatgcggga gtggacaaac tctagaagga tccactcgag tttagatctt 1140
 tt 1142

<210> 4

<211> 1140

<212> ADN

5 <213> Secuencia Artificial

<220>

<221> fuente

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

<400> 4

gaattcaaag ggctgaaaga atagacgggt aaacaagata gaggcattggg aggcctgcga 60
 aagtggacaa atgacgggat gccagggatg cacgactggc aggatggcaa cgtaagaggg 120
 10 taccacgctc cgggaggtaa ggcgatcgaa agagtagaga agcatacccg cataggacaa 180

ES 2 684 317 T3

taaaagggtgta aggggaagtgg gagcatccga aaatagaacc ctacaggcaa gtagacaggt	240
aaacgcatga gaagacctgg gcgcggtcg gaacactgcg ccagtaaggc ccatgaaaa	300
cggtgagaga tatcacagaa ctggaacaac tccggacata cagaaataca cgaatagggc	360
aatagacgag taacaacgat aacgcgagta gcgcgagata cgaagagtag caagactcgg	420
aggagtccaa gaaaataagc agatggcaaa gatagggaa aataaacgac tggagccagt	480
gaggagctaa aagggtcagc agactggacg ggtaaggaga taaaaactc agcgaatcaa	540
aaaatacagc caagtaccag ccatccacgg gtaacagctg gccaccgtcc ggcaataaa	600
agaatccagc agctagagcc ggtaaccxaa tcccgagagt cgaagagatg gcagaatgag	660
aacaatacaa gggataagca gggtagggga atgaaacggt agaggacata cgaggataac	720
aacactaaag gactagagag aataaacagc gtgacacgac taaaaaagg tacgggctg	780
caacaagtg gaagactcca aaagataacg cggtgagcga atacggaggt cgaaaaaat	840
agacgggatg gcgggatgcc agagtagaag ccctaagaga atagaaagca taaaacacta	900
ggcaacttg tacaatagac gggtgaaaag aatgaggcga tcaaagcata gcaaatcag	960
gaactaagca gggtcgcccg catggaaaca tagccaaata gggaggacta gcagggtgga	1020
aggatgacga actagaagcg atcaacggat aaggaggtgc acagatagca cgaataacgc	1080
gggtaagaaa gtacaagaat gagcaacgct tagaaggatc cactcgagtt tagatctttt	1140

<210> 5

<211> 1106

<212> ADN

5 <213> Secuencia Artificial

<220>

<221> fuente

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

<400> 5

tctagagaag gatacaggaa tcaagacatg cagggaacta aagacatacc ggaataagag	60
aggctaaagg gataaggaga gtacgaggag taaagacaga tgcacgggaa taacaagagt	120
caagacgtga ccgacctgag cgcacgaga ccgcgggaga ctcgccgggt gaacaaatgg	180
aaggataaga acgggtagaa gaactacgcg aatgaaggga tgtacaaggt agggcagata	240
agagggtggc gagatgaaaa ggctcgaaa actcaaccaa ctggagaact aggccggtcg	300
aaggaataga aacgtaagcg agataagagg atggcgcaga ctgggacaat cgggaggtcg	360
agagatcacg ggaatagaca cgtccaagaa atgaacaagg taaggacatg cgggaataag	420
aaactcaaac cctaggccag tcacagaagt aaccagtac gcaagtcacc gagtcggaac	480
gatgacaaag taaaaaagg gaaagaatga gagcactaaa aaagtaaaag agatcaggcg	540
cgtgggctgag atgcagaggt aacgaaataa aaagatcgca gaatccaaag actacggagg	600

10

ES 2 684 317 T3

atcaggcaaa taagaagata aaaaagatcc gaagataacg aggtaggcca atgagagaat 660
 accgagcgta gagccatacg agagtaaaga gtagtaggca gtaaaaagat ccagcggtag 720
 gagcatcggc gcctcaagaa gatccagaga tggaaacgcta ggagaataaa gcggtgaaag 780
 cgtacaaaacg tagggaagtc gggagagtag agagatacag agagtaagag ccatagacac 840
 ctgagacgat cggcaactgg gcaacatcca gagatgggca cctaggcaca tcaccgggtg 900
 caaaggatca cagagtagaa cgctcaaaga agtcaccaag tgcacgggta agggacatgc 960
 gaaggtgaga gcggtggaa aactcgacag ctcaacacat gagggaatgc cagagatgga 1020
 gcaataaga gacatgccgg gcgatgagga actacaacag tagcccaatg caaggctaag 1080
 gaagtacaaa agtcgggaaa ggatcc 1106

<210> 6

<211> 1107

<212> ADN

5 <213> Secuencia Artificial

<220>

<221> fuente

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

<400> 6

tctagagaat gaaacagtgg aaaggtcggg aactgaggcg gagtagaaac aagtcgctgag 60
 atacgagaga tgaaagagcg tcaaagaatg gaggactcag aagatagaaa catcgggaca 120
 ggtgagaaca ctgggaaaga atcgaaaaaa ctacccaaaac taaaaaaact gcaaaaatga 180
 aaggagtgga ccaactggaca gccgcggaag ctcaagaact gtgtacaggt gcagcggta 240
 gaaacatcaa aaggtaacag catcgcggga tcggaaaatc cgggagtcaa aaagtgcaga 300
 aatgacacgg gtacgaaacc tcagaaacat aaacagctac gggcatgaga gcataaagag 360
 ataaaaaat gcagagcgat accacacgta aggagatcgc aaggtgccac gatcagaaaa 420
 tagagacatc gacgggtggc gaactgagca acatccggga ataagaacct caaggcctg 480
 caaccggtac acgaatgaga ggatgcggga gtcagcggag tgggaaggtg aggaaagtag 540
 aaaaaataca gccaggtaca aaagtaaagc gggtaaccg gtagccaagg tgcgaaagta 600
 ggcccgtgag aaaatccggc aatagaacag taaaaggtg agacgcataa cacgggtgca 660
 aagctaagag agatccacca gtaggaagat gggaaagcat acaaagata gaagaataag 720
 gcagtaggag aatgaaaaca atacagaaa gtagggcaga taaagcacta acgagctgac 780
 aggtggtcag gatgaagaca tgacacggat caaggcatgg acgagtgcac aggtcaagcg 840
 ctccggcgat gagcgaatag cccgcgtaac acagtccgac aaggtgggag catagccoga 900
 tcacgagggg gcaggcgtga gaaaagatcc agggctaaaa cactagaaac atggcacagt 960
 acagggactg aggagctgac aagatgcaca cgttcaaagg ataagaaacg tcacgaggat 1020
 accacatca cgaaataagc gggtcgaaa catgccagac tgggcaagtc accaggatga 1080
 agaatgaag aagtagcaag aggatcc 1107

10

<210> 7

ES 2 684 317 T3

<211> 1106

<212> ADN

<213> Secuencia Artificial

<220>

5 <221> fuente

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

<400> 7

tctagagaaa aacataagca gcatggagcg ctacgggagg atggcacggt cgcgggaatc	60
gaacgctgaa gcaatcaaag ggtagggagg tcgggcaata cagcgcacgg aggatgagac	120
gggtcacaag aaatgggacc atccgcaaat caagagataa ggaaatacaa ggaataagaa	180
ggatgaaaac gtgcggagat gaagacgatg acgacaataa tgtacaatga gcagatacac	240
acggtgaacc ccgcggaacg ctaacaaaaat accgaaataa aaaaatagcg agatacaggg	300
ctaaagaagt gcaagggagt aaggccctgg cagggtaaca 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
aaagacgtgc gcggaataag aaaggataac gagatcaaga caatggaacg agtacaaaag	840
tcaaaacgtc gaaagggtaa agcactgacg aacatagagc gataaacacc tgggaagatc	900
cggaactagc aaaatcacga caaatgcaa gatcagccga tcgaggcata ccagggtgaa	960
aaaatcaaaa aagtaaccgg acctaaccac acgtgagaaa atgcgagcgt acagacctga	1020
cgaaatcaac aatgaagaa gtaacagaga tccaaaactg aaaaggtaaa agcatggaca	1080
cgtaagcaag atagacaagg ggatcc	1106

<210> 8

10 <211> 1107

<212> ADN

<213> Secuencia Artificial

<220>

<221> fuente

15 <223> /nota = "Descripción de la secuencia artificial: polinucleótido sintético"

<400> 8

ES 2 684 317 T3

tctagagagg tagcacacgg tgaaaagcta agaacctcaa acgatcgcac catgacgcga 60
atgagacaaa tgccgaata cacacatacg aaacctcaag accctaacgg gctaaaagag 120
tcggagccat agaaaggtaa cagcgatacc gaaatggaac gggcgcagca 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 caaggcgtcg 420
cagaagtaag ggaatggcaa aactaagcaa agtagagcac taacaaaggc aggacgaagt 480
acgaaggta ggaggtggaa agctggcgca catgagagaa tacaagaact agcgaaggta 540
cagcactgga cggatcaaaa cggcgcgaac gtaagcggc agggcggtag aaaagctcaa 600
cgactgagag gcatgaggac gctaaaggga gtagcaagct gagaaaatgc cggaaagtga 660
aaactggaaa gataagcaag tccgccaaagt cgagaacgta gaagagtacg ggagtacaga 720
acctggcgga agatcaggcg agtagcagag tcaaaaactg gcaagatcag gagctaacca 780
aggctcaggc aatgcaagac tgaacgggtc acaagcgtca ggacatagaa gaatgaaaga 840
gtaccagaat acacgggata gcagacgtaa ggacgatagg agcgtgagcc aatccgagga 900
tagaggaata aagcacataa gcgggtgaga ccatgagagg atcgaagaat acgggaataa 960
ggcgggtaag aacgtgacgc aagtgaggcg ataagaacaa tagcgcgaat gacggagtac 1020
ccaaatcaag ggatcgccgg gtgaaaaagt aagagaatcg ccagcgtcca aggaagtaga 1080
ggaatacaac aagtcacgga aggatcc 1107

<210> 9

<211> 1108

<212> ADN

5 <213> Secuencia Artificial

<220>

<221> fuente

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

<400> 9

ggatccgatc gggagataaa gggagtaaac gcatcaaacg atggcgagat gaacaggatg 60
acaggggtccc gaacataaccg ggatggagcg ctgggaacct ggaggaacta gaaacctccg 120
aagctccacg ggtgaacagg tggacaggta aaagaatata acaagtataa ccgtgaaaaa 180
ccctcaggcg gtagggaact cgaagaaatg agcgggtaag gaaggtaaag ggatgaacgg 240

10

ES 2 684 317 T3

gtggaaagat gggaaagtaa agagaatggg acaccataac cgcatacgcg gaatggcaag 300
 agtgcacaac taggagaagt cacaaaaggg aaaactcgca ccgtaaggac ctggacggga 360
 atagacggga tcgaacggat atcgagtacg gaaagtccaa gagctgcgac actagaacac 420
 atgcacgact acgccagggt gaaagcgtaa acgccgtgca aagctcgccc ggaatcagac 480
 agtgggcgga taacgaaagt aaccagggt gacgcgctgc cagactaaag gagtaaggga 540
 gatacaggca ctaaaaggga taagcggatg acacggatag caaacgtaag gagatgacag 600
 gcatagcaga actacgaaaa tgaacgaata gagaggctaa gaggggtggag acgatgcaaa 660
 gaataagaga ataaagacg tgcaggcatc gaaaactgca agagtagggc catcggcaac 720
 tcgaccgggt acaacgctaa cgaacgtaca gggctcggcg aatcgacaaa ctogaaggat 780
 cagaccacat ggaaacgtcg ggacaataac ggaataagca actacgacgg tgaaggcgtc 840
 aacaagtgga cgagtaaagg cgtgaagacg tccagagctg gaaccctgcg ggcataaaga 900
 gagagctcaa tgcgcagact gaaacaataa ccaggtagag gagatccacc ggtcaggcga 960
 taaagaaagt cgacagggtga cgagggtgac aggtacgcga ataaaccaat accggaatga 1020
 gaaggtacag gcgtagaaaa gatgaggaga tcagagcgat gaagagatag cgaagtaag 1080
 aaaatacгаа agtccgggag aactcgag 1108

<210> 10

<211> 1107

<212> ADN

5 <213> Secuencia Artificial

<220>

<221> fuente

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

<400> 10

ggatccgagc tacaagacta caagagtcgc gcgccgtaaa aaaagtggcg caggatggaa 60
 acgtgaggag gtagcgaagg tacagacctg gagagatcag gaactggcag aatgaggaca 120
 tcagccgatg aaccctggg caaataacgc aatagaagag taggcagggt agagcgctaa 180
 cggaatccgc acgataagag cgtgggaggc taacagaatg gcccaatgac gggacgtaag 240
 cagatgaaaa cagtggcgga gctggaagaa tcagagcagt ccgaggggtgc agaggttaagg 300
 agatgggaga aatgcaaaga atgcggaagt aaacaaataa gaggatcgga gggctaaaac 360
 gggtcggcga gtgaaaaaat cgggaagtca gcgcataaag gcctgaaaaa gataagccga 420
 tgaagaaact gggcggatga cgcgggtgaa agatatccgg gaatggcacg gtggcgggat 480
 ggaaaagtgc gggaatcggg ccgtgaacga atagaacggg tacacgcatg aggcaagata 540
 accggatcga gagcaactca gcgatacaca gctcgcaaga gtaggagcgt accgggagtg 600
 acaagataag agagtaggga gatccagaag tcacgacatc gggagctgcc gggatcagag 660

10

ES 2 684 317 T3

ggtcagggaa tagcgaaaaa tgcaagaata gaaacatgca gacatgaacc gggtaagcca 720
atgaaccact gacagaggtg agagcatagc gagctaggac ggtaaaagag gtccaacgcg 780
gtccagagat ccgaaaatcg aaagagtaaa aagataggaa ggtgaccac tgcgagggat 840
gacaacgtga aaaaaatggc gcgataacaa cagtacgaaa gctacgggaa ctgacagggg 900
cgagctctag caaaaatgcg ggagtggaac ggtagaggag tagggcaata gaaacatgaa 960
gcaatgaaag ccgtagggaa actccaccag tcggagaata cgacaggtca cgcaataaa 1020
gccgtcaacg ggtgaccggc tacaagcgtg caacaaaata agaacagtgg gacagagtag 1080
gccactagaa ccataaccaa actcgag 1107

<210> 11

<211> 1107

<212> ADN

5 <213> Secuencia Artificial

<220>

<221> fuente

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

<400> 11

ggatccgact acaaaaatgc acagatggcg agtataaaag gtagccaagc tgccgagggtg 60
agcggatggc gaagtgcaca ggtaagagag tacacagctc aagcagtagg gcaagtaaaag 120
aactcacaag ataagccgat caaggcgtcc aagcgtggcg aggtgggaga ctacaacacg 180
atagagcagc taaaaaagta aaaagataac aaaaatcccg gaagtgaag aatcaaacaa 240
tggagaagtg aggcaactgg agaactcaga gcatacggca gatggaacg tccgggaaat 300
caagcgagtg agagggatcg ggaaatagc agaatcaaac aagtgggaag atcaaccggt 360
gaggagacta aacgcataac cggatagagc cgtacggca ctaagaccg tagggagctg 420
cgggagtgcg agacatggac gagtggacac atcgagaggt caaggagata gagaggactg 480
agggactccc aggataaaaa ggagtaagag agtccggcag gtcccagata tcggagctca 540
aacgataaaa aatggggcg atacagacgg gtcaagaaat cgcacggctc ggacgatgac 600
ggaataggaa aatgagggcg taaaaaacta gagaaatgag gaagtacgag cgtcgggaa 660
tgaggaaata caagatagc caaatgagac gatggaggcg catcgacaag tcagaaaaat 720
gccgcatcg aggacgtcga cagagtagga cactcgggga 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 gtcgggaag tagagagaat acaaggtgga 1020
cggagtgagg aaataagaga gtggacagag tacgagagag tgcgaccaat aaacgggtag 1080
gcgaatcaac ggataggaga actcgag 1107

<210> 12

<211> 1107

10

ES 2 684 317 T3

<212> ADN

<213> Secuencia Artificial

<220>

<221> fuente

5 <223> /nota = "Descripción de la secuencia artificial: polinucleótido sintético"

<400> 12

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

<210> 13

<211> 1125

10 <212> ADN

<213> Secuencia Artificial

<220>

<221> fuente

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

15 <400> 13

ES 2 684 317 T3

ctcgaggacc cactaaggca atgaaagcct agaggcatga cgggataaac gggctcggga 60
 agctgccggg ctgacgagaa tggaggccta aagaaactgg agcgatcaga cagggtacga 120
 cgctcacgca ggatgacagc aatacgacgc tacaggaatg gaaagaatgc gagagctaaa 180
 acagtccaag ggtagagcgg tagagctcta gggaaactaag aagagatgga gaaatagaga 240
 cgatagaaac ctagcaaaat ccgaagctcg ggagatccag cgagtgagga gatacacgaa 300
 tgcgagcgat ggcgaactgc caacggctga acacaatgag caaatggaga aataagcgaa 360
 catagggcga taagagaccg cggcaacgtc gggaggtaa caagtagagg aataaaccag 420
 tgggaaaatc agaaaataaa gaggataaag gcgtcaggaa atgggaggaa tcgggaaatg 480
 aacgcgtaag aagatagaaa ggatgccgag gactcaaccg aatgacagac gtagggaaag 540
 atacaacaat caccaaatca aaaagtgcga gcagtcgagg aatcgggagg tagaaggaat 600
 gacaacgatg aagaacatca aaacgtgaca ccctaaggcc ctagagcgat aaaagagtga 660
 ggcaatcgag ggatgcgagc gtagacaagt gcgaacgtag accactaaga aagtcagcag 720
 aataaacaga gtaaaccact agcaagatca aagactaaaa aactacgcac cgtagccaga 780
 ctcgggcaat gaacggatca aaagataagg gaaatgacag gatgcagaag aataacggga 840
 ctacaaggct aaggcagtca 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 agggtgaaga gaccatggaa gcatgcaaaa gatct 1125

<210> 14

<211> 1125

<212> ADN

5 <213> Secuencia Artificial

<220>

<221> fuente

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

<400> 14

ctcgaggacg ataagcaaaa tagggccgtg agacaaatgg cgacatgaag caatacccaa 60
 gtgacaagct agagaggtaa cagcatagac aaccctaacg ggactagccc aatagaaga 120
 gtaaacgggt agggaactga ggacctgaaa acctgcaaag actgggcagg gataggaaca 180
 ataagaagat gaacagatga gcgagctcgg gaaatgagaa aataaaaaggc gtacgggagt 240
 gggacagtca cgagaataaa ggcgctcgca gatcaacgga tgcagggcat gggacagtac 300

10

ES 2 684 317 T3

gggagagtac gcgctgaag agctgacacc ctgaggggaag taggcaaata aaagggtagc 360
 ccactagcga gcgtcaccgg gtggaaagct aggaacgtcg gaaactagga gagtcagcag 420
 ctccgaagct gagaaactcg gcagatcacg gaccgcggga agatgaaaag ctgaggaggg 480
 tggacgggtc agaactggtg aaactagacg acgtgggca atacgcacgt aaaggagtac 540
 gacacatcag ggcctgggaa gataccaaga tgccggaaga tcgaaaacta aagcagtgga 600
 acagtcaaca aatcaagggc gataagcgaa taaggaggtc agcaggtgga aaccgctaag 660
 accgtgggaa actcaggaaa tacgggcagt aaggcggctg gcagactagc gacatgagcg 720
 ggtcacagaa ggtaacgcaa taacaaaatc gcgcagtggc acagtaaagg cctcgggaa 780
 tgccggagatg caagagtacc aaaggctaag gcaactggca cgatacggga actagggcga 840
 ctgacagcat ccacgcagta agagaatggc gggatggagc gctaagacgg tgaaccaata 900
 agggcctaac aggatgacaa ggatgcggga atgagccact aaaggaagta gaggagtgaag 960
 ccgggtgcga agctggaggg aatagggaaa tacgaggatg ggcaggtgag gagaaatcgg 1020
 acggatgaac ggctggcaag gtaaaagggt agcggaatag agagcagtcc gcaggtcaaa 1080
 cgggtgcacg ggctcagacg ggccatggaa gcatgcaaaa gatct 1125

<210> 15

<211> 1123

<212> ADN

5 <213> Secuencia Artificial

<220>

<221> fuente

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

<400> 15

ctcagggaaa aatggcaagc taagaggaat caagaactgc ccacctaaga ccaatgagggc 60
 gataaccgaa atcgggcaat agccgaataa agggaatgag acgggtgcgc gactacacaa 120
 gtgcgcaact aaaaaataa cgaagtggga agcgtacaa agatgggcag atcggcaggt 180
 aaggacgtag ggaagtacag gagctcccg actaggacca tccaacactg gcaggataga 240
 gcggtgaaaa gctggagaaa cgtcgggacg tgggcaggtc agagggtgca gaagtgcgg 300
 gcataagcac atacaacggt agcagagtcg aaaacataaa gagactggac gaatcagaga 360
 ctacgggact agccacctgg gagagtacc gggatacaag ggataagagg aataggcgag 420
 tggacagatg gaagcatggg aggatcacag aactaggag atagcgagat accagcgtgg 480
 agaagtaaaa aatgagggga ctaagggaat gaaaaagtaa gaaacccgcg gggtacacca 540
 gtgcagcagt aggagaatac acgaatgcag ccagtcaaga gaaatgaaga actaagagag 600
 tgcgagagta cagagctacg caagtgccca agtgagagaa tagcgggcct caagcgatca 660
 acgacataac aggagtccag agaatcgcaa agtcacggga tgcgagcagt ggacaatca 720

10

ES 2 684 317 T3

gcaaaatcaa agactcacia gatccgacia tagagggata aacaagatac aaacatccag 780
 agactggcag gataaagaaa gtacaaagcg tgccaagcta aggaagtaga gaaatccaag 840
 aatacagagg tgacgccgtg aagacatgcg caggtgagca ggataggaga ctaaaggcgt 900
 acgggaatgc gaaactagaa ggctgaaagg atcgaccac tcgcagcgta gagggctacg 960
 acaactaaag acataagcag atgaagccca tcaaggacat ggcgcgatgg gaagatccaa 1020
 gagatccaag ccctaggaaa gatgaacgcc tgaacagcta agagcgggtc caaggataag 1080
 aacatgcaga aatggacgag ccatggaagc atgcaaaaaga tct 1123

<210> 16

<211> 1125

<212> ADN

5 <213> Secuencia Artificial

<220>

<221> fuente

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

<400> 16

ctcgaggacc aaaagtcaag aaagtaccgg gctagaagag ctgaagccat aagcgagtag 60
 cagaatagaa agatcccaaa agtcccaggg atagacgagt aggaaggtga aaaaatggcg 120
 aggtagcgac atgcaaaagt aaaacgatga aaaactacga ggggtggaaga ataagcaggt 180
 gacgaagtaa acgggtcaaa ccgagctcag atcgaacgat aggaaacatg accgggtcac 240
 acgatcgaga ggtcccaaaag gatagaagaa gtgaaaaggt gagcaaggtg gcgaaaaataa 300
 aaagataaaa gaataagaca gtagcgggaa tacgacacta gaaggatcgg gacatgcagc 360
 agtaaaccba taggaggata acagggcatg gaagagtggg acggttaagac cctacacgaa 420
 tacaagcagt gccaggatgg cgcgagtgac aaaaagtaga agggtgaccg agtcgagaga 480
 tagagacgta aggaagtagg gaaggtggga cgatcgaaaag atcgaagagt aggagcgctc 540
 gaaaaatccg gaaaataggg aagatggacg gatcgggacg gtgaggagga atagcaaaag 600
 tcccgcgggt acgaaaagtc gggaaggtca agcaatcagg cgctcaacgg gactgaacaa 660
 ataaggacat acacaagtcg gcacgtcacg aactcaaaag gtggagaagg tagcgaatc 720
 gaggagtgga gaaggtaaag aatgggaag gctaaagaaa tggcagggtg gagaactggg 780
 acggtaaacg catgaaagaa tcagggagta gaagaacgtg aaggatagg agaactcaac 840
 aggtagcag aagtggaaag catggcaaga atggcagcat gaaaagatcc aggagtaagc 900
 gagctgaaga aatggagacg taacaacata gcgggagtag gcgcgtgaca agataacgcg 960
 aatgcggagg tcgaggaatc cgcaagtga cacgtcaacg caatgaacgg atgaacacat 1020
 gcacgaagtc gacaagtaaa aaacgtgga gccatgacaa cataacggga tggcaggata 1080
 agagagtaga acgatgcacg agccatggaa gcatgcaaaa gatct 1125

<210> 17

<211> 1108

<212> ADN

10

<213> Secuencia Artificial

<220>

<221> fuente

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

5 <400> 17

ccatgggaca gcgatcagag ggtaaaacgg gatgaagcag tgaaggacc tcagcgaatg	60
aaaaacgatg gccagatcca aagataaaaa actgaaagac tacggaaata caagaataga	120
agggtaaacg actgagaaag tacgaagcct agacgggtaa aaaaggtcgg gaagggtaac	180
gccatagaca aatgagaagg taaaggcatg gaaaaaatgg aggcacgac gaatgcccg	240
ctcaaaggat aacggactag cgcggtaaaa ggaatgcgg acgatcgaag aagctccgga	300
ccgatccacg gaatagagac atacgacagt gcgccaaatg gaccgataaa agggtagacg	360
aaataacagg atgaacagga ctgggagaaa taacaaagtg gagaagtac aaaagtcaac	420
gaataggcca gtggcaaaag tgagcgagtg aacaggtaga ggagtggaaa agtacaagg	480
atgcaaaaat gaaaaggtag aaaactaagg cagtacaggc ataaacgact cacaactag	540
aaaactacag cagatcgaag cataaaggaa ataggagaga ataaaaacgc ctaacaaact	600
acaagaacta cgcggagtgc gagacgtcag gcaagtgagg acctgaaaca atgcaagaat	660
ggcgaagtgg acgcggtagc gggataagca aatacaccgg tagatatcat aggaaggcca	720
cgcaaatgga ggagtcaaga aactggccaa gtgaagccct cggcagatac gcaaagtacg	780
aaaataaga ggctcagaag atccagacga gtgcaggaat aagacaatca agagaatgaa	840
cgcatcggaa cactaggcag cagtgggacc ggtaaaagca tagctagctc gggcgatgga	900
ggcacgtaca aagggtgaaa aagtaacggg aatacccacc gtaccaggat gaccagggat	960
cgcgaagata gcggaatcga gccctcagga gctaggccag catagagacg tcacgagatg	1020
aagggataaa ggaagtcaaa gaatgggaga actccggacg tacgaggcta cagaggtgaa	1080
aagataaagc aggggtgagga aggcacg	1108

<210> 18

<211> 1107

<212> ADN

10 <213> Secuencia Artificial

<220>

<221> fuente

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

<400> 18

ES 2 684 317 T3

ccatgggaca agagatagaa agcagtgaaa gaataggacg gtcagaggat ggagggcctc 60
 cgggcgtggc acgactagga cgatgcggaa ataacgacaa gtggaaacct agccagctaa 120
 ggaagctaag ggcaatggca aagtaaggaa gtacggaaat aggaccatag aagactggac 180
 cgatacagcg ctagggcggg taaacgagtg aaaggggtga acaataagga cagtgcagca 240
 ggtaagacca ctaaaagact ccacgacgta cagagactcc ggcctggaa cgatcgaagc 300
 gtaacgggca taaggaaaat agagaaggtc gaggaaataa agggaaatgg agaaactaag 360
 cggatagga aataaacgaa cctcaggaa tcccaaagtc cgaccaatga cagactgcac 420
 gcatccgagc gtaaaaacat gagaccaata acgagatcgg caagtcgaga agtcgcagaa 480
 tcaaacaacc tgagaacctg cgggagtga cggataagac ggtaagagaa ataagagcat 540
 gagaaactgg gacgatagaa cgatagccaa gtaaaaggt aagagaatgg aacgaagtgc 600
 aaaaagtggc agaatgaaca gataggcaga tcagaagaat gaggaagtgc cagaaataag 660
 aggggtgggag cgatgccggg atgcccgaat gcgaaggtaa gagaatgcag gagtaaagag 720
 gactgaaaag atcgggacat gaaacgatag gaaggtacgg cgatgagaca gtacaaaagt 780
 gaaaggtga cagcctgccc ggatatcagg gatgccacga tgggcacatg cccaaaatga 840
 aaaaatcacc aaataccaaa atgaagccga tgcgggaatg cgctagctaa cgagcataaa 900
 acggtagga aatggaaagc taaccgcagt ggaagaataa ggagctacgc aaatacggc 960
 aataaggaag tagcggacta aggaggtaga cgaataggcg aataacgcga gtcgagaaat 1020
 ccaagactac aggactcagc agatgaaaaa actagaaacg tgaaggccta agaaacataa 1080
 gacactgaaa gtagtagcga ggcacg 1107

<210> 19

<211> 1106

<212> ADN

5 <213> Secuencia Artificial

<220>

<221> fuente

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

<400> 19

ccatgggaga gtaaggaaat gagaacagtg aagacatccc aagaaatgaa aaaagtggag 60
 aggtcgaacg gtagagcagt ggagaaggat acgccgatcg ccgggataac cgggctaaac 120
 acaatgaaac acgtggccga atacggagga atcagaggag gtggcaggac tgaacgagtg 180
 ggcgggatag aaaaactaca gcgatacggc cgataggaac ctacgagaac taagaggata 240
 aagacataag ggcctacgca cgagtaaaag agtaccgacg tcagacaata gaagggtaaa 300
 aagatgaacc gatgagcaaa atccaggcgt cgcaaggtca cgcagtcaag acataagaga 360

10

ES 2 684 317 T3

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
 aacaagtcgg gaggtgaacg aataaggacg ggtaaggac ctggaccgga atggcaacat 720
 gcaaacataa gaggtcaac caatggaagg ctgaaaagat cgaaaaatgg gcgaatacaa 780
 aaggtaaagg gatagcggga tcagaagggt ggacgatgaa agaataaga aataccaagc 840
 gataacacga tccggaactg cgggagctga acaagtcacc gctagccaga ggggtgcaggg 900
 gatatacaat gggcagatac ggagcgataa aaacatgaaa ggagtgggcc actggaagga 960
 tcagcacgta acggcctaaa ggactaagag cacatgagcg aagtgaagag agtgaagaaa 1020
 ctaagaaaga gtagagagat gcgagagata gcgaaaatcg acaacatcgc gggagtggaa 1080
 aactgacggg atgacgagaa gcatgc 1106

<210> 20

<211> 1107

<212> ADN

5 <213> Secuencia Artificial

<220>

<221> fuente

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

<400> 20

ccatgggaac agactaaaag aatcaacaga tagaggaatg aggaagtgca ggaagtaagg 60
 aaaactggaa gagtaacaca atgggagaat acaaaagtca accagatgga cagatagaga 120
 aatgacgaga tggaacagct cacacgctaa ggaatggac gcgtgaccag atcgggaaga 180
 tccgggcaat aggacagtag aaaggtgcag gaatgacaag atggccaaat cacagcatag 240
 agccaataag acggtaaaag gcgtagcacg ataggacggg tcacgagagt gagagaggtg 300
 ccaacaacta ggacaataag ccgataaaag cgtcgaacaa tagagacgtc cagagaatgg 360
 acccaatcca ccggatagca agagtagcgg agataagaac gtcaggagat agcgaagtca 420
 cgaaatgaga gagtcacgcg gtagcacaat aaagacgtac aaaagtacaa caaagtgaga 480
 acatcaaaac gataagcagg ataaaaaggt aaaacgggat cgggacgctc gaaaactgac 540
 gaactagaaa agtaagaacc tagaaaaata gcggcataga aaactaaggg aatggcgaac 600
 ataagaggaa taggaaggtg gcgaagtgga gcaataaagg aggtgggacg gtcaagagct 660
 agagaaatcg caacgatacc ggaatcggga agtagaacia atcagcggcg taggacaagt 720
 ccgggaatcc acgagtcgaa aaaataagac actcagagag tgcggaggct aaacgggtgg 780

10

ES 2 684 317 T3

aaagaactgg aaagatccag agcatcgag aataagacga gtgcgaggat caacggataa 840
 ccacctaagg gcgtcgggag ataacggcag taccgaaatg ggctagccac gggatgggag 900
 aatggaaccg taggaccggt agagaaaagtg gaacactacg gagaatgcaa cagtaggaaa 960
 gctgacgggc taaggccggt aggacaagct ccgcgcgtga agatatcgcg gagtagagga 1020
 ataccgggcc atgacgggca ctcaacggct gaaagagctg aagcaggtgc aggactggga 1080
 ggagtaggga cgatcacagg agcatgc 1107

<210> 21

<211> 1107

<212> ADN

5 <213> Secuencia Artificial

<220>

<221> fuente

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

<400> 21

gcatgcgaag aaaagtggaa cactagggcg gtgaacggga tgaaaaggat agagcgggtgc 60
 gaaagtcaag aaataggaga actaacgaag tgggagagac taaagacgtc agaagatcaa 120
 aaggtcggag gcaatggaga agtgcagcaa tcagaccgtc cggcgcgtgaa aaccaatgaa 180
 gacagatcca agagctgagc agctagcgaag gaggggcggag cgtaaagcag tagggaggta 240
 aaagaactac aggcctcaga caatccagac gtaagaggga tacagccgtc agggaggtaa 300
 gagaatgaaa ggatgcgccc atgagccaac tgggagggct aaaaaactag aagagtgaaa 360
 aactggcacc gtgagaaaat agcacaatac gaaaggtcga gaaatcggga aagtgcggaa 420
 gtggaccact agcgagatca acagagtagg gacatcgaag aataaggcag aatgcgacag 480
 tacgggagat cgaaagactg cgagcgtgac gagatgaaga ggtaaaaaac tgaaaacatg 540
 aggcgggtgaa agggaataac caggtgggac acagtgcgag gcatggcggg cctcaagaga 600
 tgcacgactg agcaggtgag aaggtacaca gatacagaaa tggaacggct caaaaacaata 660
 agaaggtcgg cccgtgagac caggtaagag agctggagga cccgcggagg tgagcgggtc 720
 aaaaaatcga gagataagga gaggtaacgg gtgaagacag tagaaaaatg agagaaatcc 780
 ggcaaatagc aggactggac gcgtacgaaa gaggggcaaa ataagcggat ggcgagatgg 840
 cgggctcggc gagatagcaa gatgaagacg taagaacggt aaaaggctca agagatgaac 900
 aaataaagag atacgcggct aaaagacctt aaaggataag aaactcaagg cagtgcgaa 960
 acgtgcaagc agtaacagaa tgagaaagga tgaacaccgt ggcgggatag gagagtggag 1020
 aatggaaga atacaagaat aagaacggtc ggacgcataa acaggtgagc caataagaga 1080
 ggatcagcaca cgtacgaagg aagatct 1107

10

<210> 22

<211> 1107

<212> ADN

<213> Secuencia Artificial

ES 2 684 317 T3

<220>

<221> fuente

<223> /nota = "Descripción de la 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 ggctagcgct ccgggagata ggaaggatga caggccgtcc	240
	gaaagtacgg aaatcaaaaa gtgaggcact cagagagtga aagcgtgaag acgggtaggg	300
	agctagcggg gtaggccaca tcacgggata aagagatgac aagcgtgaag gaatccacgg	360
	agtgcgcaga cgtccaaagg tccacaactc cgccgggtac aacggtgaag caaatcacgg	420
	gctcgaaaag gtggcgagat gggaccatga aagaatcgaa gggtaagacg atccagagat	480
	gagcccataa caaggtaca acctaacaga agtaccagaa tagagcaact gaaaaagtac	540
	agcactggca acgtccaagg ctgcggcct gcgcgcgtgc gccgaaataa ggacaataac	600
	gacctaggac cgaatacaaa agctaacaga ctcagagcat ccaacgctga gccagtcaga	660
	aggtagagaa gtaagcgaaa taaagaaaat aaagaaatgc cgcgatcgga aggtgaagaa	720
	ctaccaggct acggagaatg aaaggtaag agggtaagga catacaagaa taaagaagct	780
	aggacaataa acgccctaaa accgcggaga gaatagcccg atacaagcgt ccgggcatag	840
	aaacctggag aaataggaca ggtgaaagcg tgggcacatg accgactgga gagcatggaa	900
	acatacacc acgtcaaaaag tcgaggaata gccgggtggc acgtaacaa aatcaagaaa	960
	ataccggagg gtgcagaagg gtacgaccgg tacaggacct aagaacgatg gcgagagtcg	1020
	ggccgtgagg acagtaagga cgtggaggaa gtagcagaat agcgggatag ccagctacga	1080
5	agaaatgggc aagtgcggag aagatct	1107

<210> 23

<211> 1107

<212> ADN

<213> Secuencia Artificial

10 <220>

<221> fuente

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

<400> 23

	gcatgcgatc gaaaaatagg gaggtggccg gctgcgaaag tcgggcccgt gaaagcaact	60
--	---	----

ES 2 684 317 T3

aaaaggatcg aagcatgaaa gagtaggaaa gtggaagaat gagaagataa cgaaataaagg 120
 aaggtaaaac ggtggagaga tagaggacat aacaaagtgg aaacactcaa gagctaaggg 180
 aactagaagc ataacggaat ggctagcatc gggagagtag cggcatgaag cgatagggaa 240
 cgctgacaag aaataaccga atcgggaaat caaacatag aagactagac gggtaacaga 300
 atggaggcaa taggaaacgt aagacactaa cgggatacca cgagtgacaa gatcggagga 360
 tggcaccatg aaaagataga gagctagaag ggtacgggaa tagaaaaatg caaagctagg 420
 gccgtcgacc ggtgagaagg gtaaaaaggg tgacaaaatg agagaataac cagataggga 480
 cgtgaaaggc taggcacctg gagacaatga ggcagtacac gcgtaccaag atccagagaa 540
 tcagacggtg agacactgga caccataaga agatgaggag gtgagggaca tgaaacaata 600
 acaacaataa ccacataagg gcctcgaac gtggagagca tacagccggt gcaaaagtga 660
 gacggatgcg aaaaatgaac aggctgggac gatacccaga atgccaagat ggcggcctgc 720
 cgggatcacc gagtagcaa cctgacaaaa agtgccgaaa tacgggcatg agaaggtgga 780
 gcactcagac gatgaagacg atacggacgt accaaaatgg aaaacaatgg gagcgtaaagc 840
 aaaatcagac ggtagaacgg taaagagatg cacaagatga agagcatcaa cacatgacag 900
 cccgcgggag taggcggagt aagcgggtaa cgaggtgagc acataaaaag gtcgcaaagt 960
 acgaaggtaa gggagtggag agagtgaagg gcctggaggg agtcggaaca atgccgacgt 1020
 aagcaagtca aacgctaagg caagataacg gaatgcgaaa ctggcgcctt aagaacgtag 1080
 gaccataaca aggtagaaaa aagatct 1107

<210> 24

<211> 1107

<212> ADN

5 <213> Secuencia Artificial

<220>

<221> fuente

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

<400> 24

gcatgcatg gaaagctgag agaaagtcag aagatcacag actcaggagg atcgggcagt 60
 agacacgtaa gaagtagag gccaatgaca ggtgaagag gtgaaggcct aaaagaatac 120
 gcgggtcagg aacaatgca agcctgggac gatcgggaga taaaagggat agagaaatgg 180
 ccgaatagac cggtagcggc cgctagcgat aaggaactag acgggtaaaa ccgtaggaag 240
 ctcagaaaca taacgaaagc tacagcaggt aagcaagcta agaacctgag caggataacg 300
 cagtaaggac actacgggaa ctaagccaga tgaccgggta cgaacgtcaa acgaatagca 360
 aaggtcgggg acgtggccga agatagaaaa catacccaaa tgccggaatg ggagaatgcg 420

10

ES 2 684 317 T3

caagtaaaaa cataggaaca gtaaggcaat cgggaggtga aagggaatga gacaactaac 480
aaggtgggaa aaatccagca gtaaacagca tagggcaatg agaaggtaga acaataaaga 540
cgatgaggaa gtaaaaacgg gtgggcaaat gacacgatga aaaagtaacg gagtaaacc 600
actcagaaac tggaaaaagt cgaaacatgg gaagaatata ccacatccag cgaataacgc 660
gactcccaac caatagacga gtgaagagat ggaagccctg gcgaacatgg agacaatagg 720
agggtcaagg 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 gtagaaaggt ggaccgataa cccgcgggca caataaacga 1020
actcgcgagca tcagagaggt gcagaagata ccacgagtag aggcattggga acgtgaaaac 1080
aataacgcaa gtcaggacga gagatct 1107

<210> 25

<211> 15

<212> ADN

5 <213> Secuencia Artificial

<220>

<221> fuente

<223> /nota= "Descripción de Secuencia Artificial: oligonucleótido sintético "

<400> 25

10 ggtctgtgtg atgtt 15

<210> 26

<211> 69

<212> ADN

<213> Secuencia Artificial

15 <220>

<221> fuente

<223> /nota="Descripción de Secuencia Artificial: oligonucleótido sintético"

<400> 26

aacatcacac agaccaacat cacacagacc aacatcacac agaccaacat cacacagacc 60

agcccttg 69

20 <210> 27

<211> 3194

<212> ADN

<213> Secuencia Artificial

<220>

25 <221> fuente

ES 2 684 317 T3

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

<400> 27

aagcttggcc acacagcttg agtattctat agtgtcacct aaatagcttg gcgtaatcat	60
ggtcatagct gtttcctgtg tgaattgtt atccgctcac aattccacac aacatacgag	120
ccggaagcat aaagtgtaaa gcctggggtg cctaatagagt gagctaactc acattaattg	180
cgttgcgctc actgcccgct ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa	240
tcggccaacg cgcggggaga ggcggttgc gtattgggcg ctcttcgct tcctcgctca	300
ctgactcgct gcgctcggtc gttcggctgc ggcgagcggc atcagctcac tcaaaggcgg	360
taatacggtt atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc	420
agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggctccgcc	480
cccctgacga gcatcacaaa aatcgacgct caagtcagag gtggcgaaac cgcacaggac	540
tataaagata ccaggcgttt ccccctggaa gctccctcgt gcgctctcct gttccgacct	600
tgccgcttac cggataacctg tccgcctttc tcccttcggg aagcgtggcg ctttctcata	660
gctcacgctg taggtatctc agttcgggtg agtgogttcg ctccaagctg ggctgtgtgc	720
acgaaccccc cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca	780
acccggtaa acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag	840
cgaggtatgt aggcggtgct acagagtctt tgaagtggcg gcctaactac ggctacacta	900
gaagaacagt atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg	960
gtagctcttg atccggcaaa caaaccaccg ctggtagcgg tggttttttt gtttgcaagc	1020
agcagattac gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt	1080
ctgacgctca gtggaacgaa aactcacggt aagggatfff ggtcatgaga ttatcaaaaa	1140
ggatcttcac ctagatcctt ttaaattaa aatgaagttt taaatcaatc taaagtatat	1200
atgagtaaac ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga	1260
tctgtctatt tcgttcatcc atagttgcct gactccccgt cgtgtagata actacgatac	1320
gggagggcct accatctggc cccagtgctg caatgatacc gcgagacca cgctcaccgg	1380
ctccagattt atcagcaata aaccagccag ccggaaggcc cgagcgcaga agtggtcctg	1440
caactttatc cgcctccatc cagtctatta attggtgccg ggaagctaga gtaagtgtt	1500
cgccagttaa tagtttgccg aacggtgttg ccattgctac aggcacgtg gtgtcacgct	1560
cgtcgtttg tatggcttca ttcagctccg gttcccaacg atcaaggcga gttacatgat	1620
ccccatggt gtgcaaaaaa gcggttagct ccttcgggcc tccgatcgtt gtcagaagta	1680
agttggccgc agtgttatca ctcatggtta tggcagcact gcataattct cttactgtca	1740
tgccatccgt aagatgcttt tctgtgactg gtgagtactc aaccaagtca ttctgagaat	1800

ES 2 684 317 T3

agtgtatgcg gcgaccgagt tgctcttgcc cggcgtcaat acgggataat accgcgccac 1860
 atagcagaac tttaaaagtg ctcatcattg gaaaacgttc ttcggggcga aaactctcaa 1920
 ggatcttacc gctgttgaga tccagttcga tgtaaccac tctgtcaccc aactgatctt 1980
 cagcatcttt tactttcacc agcgtttctg ggtgagcaaa aacaggaagg caaaatgccg 2040
 caaaaaaggg aataagggcg acacggaaat gttgaatact catactcttc ctttttcaat 2100
 attattgaag catttatcag ggttattgtc tcatgagcgg atacatattt gaatgtattt 2160
 agaaaaataa acaaataggg gttccgcgca catttccccg aaaagtgcc actgacgtct 2220
 aagaaacat tattatcatg acattaacct ataaaaatag gcgtatcacg aggccctttc 2280
 gtctcgcgcg tttcggatg gacgggtgaa acctctgaca catgcagctc cgggagacgg 2340
 tcacagcttg tctgtaagcg gatgccggga gcagacaagc ccgtcagggc gcgtcagcgg 2400
 gtgttgccgg gtgtcggggc tggcttaact atgcggcatc agagcagatt gtactgagag 2460
 tgcaccatat gcggtgtgaa ataccgcaca gatgcgtaag gagaaaatac cgcacagga 2520
 aattgtaagc gttaatattt tgtaaaatt cgcgttaaat ttttgtaaa tcagctcatt 2580
 ttttaaccaa taggccgaaa tcggcaaat ccctataaa tcaaaagaat agaccgagat 2640
 agggttgagt gttgttccag tttggaacaa gaggccacta ttaaagaacg tggactccaa 2700
 cgtcaaaggg cgaaaaaccg tctatcaggg cgatggccca ctacgtgaac catcacccta 2760
 atcaagtttt ttggggtcga ggtgccgtaa agcactaaat cggaaacct aagggagccc 2820
 ccgatttaga gcttgacggg gaaagccggc gaacgtggcg agaaaggaag ggaagaaagc 2880
 gaaaggagcg ggcgctaggg cgctggcaag tgtagcggtc acgctgcgcg taaccaccac 2940
 acccgcgcg cttaatgcbc cgctacaggg cgcgtccatt cgccattcag gctgcgcaac 3000
 tgttgggaag ggcgatcggg gcgggcctct tcgctattac gccagctggc gaaaggggga 3060
 tgtgctgcaa ggcgattaag ttgggtaacg ccagggtttt cccagtcacg acgttgtaaa 3120
 acgacggcca gtgaattgta atacgactca ctatagggcg aattagtgga tgagaagacc 3180
 tgcagaaaga attc 3194

<210> 28

<211> 1106

<212> ADN

5 <213> Secuencia Artificial

<220>

<221> fuente

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

<400> 28

tctagagaga aaggctacga agatccggac atcaaaagat caaaacatgc ggagatcgac 60
 acatccaaaa atgaacagat aacaggatca gagggtaacg aagtgaacc atcgaggagt 120

10

ES 2 684 317 T3

cgcagactac gaagatggac aagtgagcaa cagctgcgat ggagaactaa aggagtcggg 180
 aggtacggca gtgaagaaat gtgcacatga aaagctgccc ggggtcgcaa ctggaggact 240
 agaagaatgg cacagtgacc cgggtgagagc ctgaggaggt gcgagaatag ggcgataggg 300
 aggtaacccg atggcggact aaagagatac ggacatggac ggataggaga gtcgcaagat 360
 ggaaggatca caaagtgagc cgggtccaagc atacgcgggt aaacggatca gaacgtaaag 420
 cgggtgaaaa atacccgagt aggaagatag caaaatacc aggtcgagcg atagcccact 480
 cggcgagtcc aaaactgaaa aaatggagag atgagacact agcggcatcg gacagtagag 540
 cagtgaacgc atcagagagt agaaaagtga agaactgcga cggtgaaacc ataacagagt 600
 cagagggtag acccatacga ccctaccgca atacaaggct aagacggca gaggatggac 660
 ggctcgcgaa ctcacgccgt gccaaagtga agaagtacca aaataggagc gtaacggcgt 720
 agagaggtag acgagtgaga cgggtggcag ctcccaaggc cggagcatgg acagtgggca 780
 aaatccagcg gtcagaaact acgagactgg cgggatgaag aaataaggcg ctagagagat 840
 acgggagtga caagatggac cactagcgaa gtcggagggt gaagacgtga acaactagac 900
 cagtacggca ctagcagaat agaccgatga aaacctgaaa agataacaga gtgaaggcat 960
 cggagggtcg agaagtaaga acgtggagcc atacacgaat aaagcagtac aagggtggac 1020
 aaatgaggaa atcgggaggt cccacgggta aaacgtaggg aggtgaacca gtgggagagt 1080
 caaccaatac gaagatgcga g gatcc 1106

<210> 29

<211> 1106

<212> ADN

5 <213> Secuencia Artificial

<220>

<221> fuente

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

<400> 29

ccatgggaaa atgagagccg ctcgacaagg ataaagaaaa gtcgccaac atccgccacg 60
 ctaagcggcc ctaggagccg gtaaacaagc atgaagaaac atgcgaggca atgacacaaa 120
 ctccaggcgg gtgaaaaga ctcgacggag atgccacgca atagcgaac ctggaagaga 180
 gtggaagcaa ataggaaca gtagagcaa atgccgagc atagaaaaca gtggcgcgca 240
 ataagcagga ataaagcaca ataaaaagaa gtaaaccgca ctaaacagc atccagaccg 300
 ctgaaaagaa atggcggcga atgaaaaaaa gtcacaaccg gtgcaacaac ctgaaaaggg 360
 gtagagagcc ctcccgggca ataaaagacg gtaagaggaa atcccgaaa gtgcaaaaag 420
 gtcaaaacgg gtgcaacaaa atgccaagc gtgcaagaaa ctgaggacgg ataaaagaca 480

10

ES 2 684 317 T3

atgCGGacgG gtacgacacg ataaggaag ataaggCGaa gtaaaagcCG gtgCGGcaaa 540
 atacaacagG ctgcccagaa gtaccgagcG ctaggCGgga atgagcGagG gtaaaagaga 600
 atacacaacc ctgcaagaaa gtccGGgaaG ctaggaccgG atacagagga ataaggaagG 660
 gtacacagca atacgagcag atgaacgagc ataccagaga gatatcagGc ctaaaaaacG 720
 gtaacaagaa ctaggagGag gtgCGaaGgC ctggaccaaa atgggacGaa gtaggagGga 780
 atacacaacG atgggacGag atagagcCGg gtaggCGcaaa gtagcacGgG ataccacaga 840
 atacagcGac gtaccGgaaa gtaggagaga ctggaaaccC gtaaggcaag ataagcGagG 900
 gtagcagaa ctaaaacaga ctgaacacGg atcaaggCGa gtaggacGg atacacGccG 960
 atacgGcaaa gtaggCGcCa gtacgGagaa gtaaacagGg atgagagaca gtgaacaagG 1020
 ctgagacGaa atcacagca atacaacaaa gtagCGgca ataaacGgga gtggaacGcG 1080
 atagcagGaa gtgacGgaca gcatGc 1106

<210> 30

<211> 1169

<212> ADN

5 <213> Secuencia Artificial

<220>

<221> fuente

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

<400> 30

gaattcGagT accacctgat gAgcagcaag gtaacGgCGa agatGaaaag aaaaatGgag 60
 cGgCacGtac aaaaagGat gGcaccagac atagaagagG agctcagcaaa acGgctGgac 120
 agCGaaGtcc caggaagcct accaagGaca gatatcaaaG aaataagaaa caaaactGaga 180
 aagCGcGtCG gaagCGcagT acaagCGgCa atggcagaga gAgtaaaaca aacagTgaaa 240
 ccaacaataa acGcccagGt agggCGaaag gtaaacaaaa gGatCGaaag acGagTaaGg 300
 cGggagatGg aaacaacGgT gCGaaaacca gTgaaaaagG gAgTcCaag caccatGaaG 360
 accacaatac gacacaccct gGgagGacac gTgGagGagc cGgtacGaaG cagGgtacaa 420
 gacaacatGc aaacagagGt aaaggaaccG ataaagaacG acatGgaagG aaacGtagGa 480
 acccGagtag aggaacGgat agCGaacGac gTgcaacGcc acctagaacG agaagTcGca 540
 agggaaataa acGgagagGt cccacGgCGg atgagcagaa aagTgGaaaa gGacctGaaG 600
 agaaaaatGa gGaccacaat aaagGacGga ataaaaaaca cactccCaga gGaaatacaa 660
 cGagaaatag gGaacagGat gcaagaagc atcagaaagG cagTgCcaaa gGgaatcGga 720
 cGcaagctaa gacagGccct agaccCGgaa gTaaacGaac aactCGgagG gCGaatcagG 780
 cccaccatGc agagcaaaat aagagaccag atgCGCGgCa acctgagacc gaacctaaag 840
 gaaagGatag agggagagGt ccaagaaacG atggCGgGca ccatGagaaa gGaaGtcaaa 900

10

ES 2 684 317 T3

gtgcacctag caacacagat accacgagac atcgaagcga agatacggaa aaagctggag 960
 agggcaatga acagaaagct acggaaagaa gtcaaggaca caatgaggga gcaagtcgag 1020
 agcgaggtgg cccgccaaat aggagcaaca atcgagggcc aagtgggcac aacgatgaaa 1080
 gaccaactaa cacagccact tctagagtgc acggatccct cgaggggtacc aacctaggaa 1140
 gctagcaagc atgcaaactg cagaagctt 1169

<210> 31

<211> 1106

<212> ADN

5 <213> Secuencia Artificial

<220>

<221> fuente

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

<400> 31

ctcgaggaaa gacagagggt agcaggccac agatcaaggc ggcgaaatgg cccggagcac 60
 gtaagcacca cacggtcacc aagagagaat cagaagaggc aagtaaaaaa aagcgcgtac 120
 ggcacccaag gtggacgcac aacagtcacg ggaagaagt cgaaggcggg cgctcagaag 180
 agcgaagtac acaaccagac ggtaccggaa cacgatgcgc caacaagaat cagcggaaag 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 aggaagggca aagtgggaga gaaaggctcg 540
 gaaccgacga ctgcaccgaa caacgtggga gcccacggct acaaacgcc agatcgaagc 600
 ccagaggtcg aagaacaaaa gtaacggcgc agaaatagag agaggaggct aaagcggaac 660
 cagtgaagga aaagaaataa cgggcccaag gtgaagaaaa agaagtggca ggaacggagt 720
 acagcccagc aactgggcaa aagcggctga cgagaacaag atagagaacg aggaatgacg 780
 agacacagat agaagacaaa gactgcggac aggcaagtaa gagaaggaaa ataaagcgag 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 ggacctaaaa aaaacaaagt cagggaaaaa 1080
 10 ggctaacgga acaaagctgg ccatgg 1106

<210> 32

<211> 1241

<212> ADN

<213> Secuencia Artificial

15 <220>

ES 2 684 317 T3

<221> fuente

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

<400> 32

```

gaattctcta gagtgcacgg atccctcgag ggtaccaacc atggaagcta gcaagcatgc      60
gaggaaagca aggaaaatga agaacgaagc gggtcgaaca ggaagacat aaaaaacagg      120
caccatgaca ccaacgcgga atgacacaac acaaaggtag cgaggcacgc acgtaagagg      180
cacaagaaat gaaagaaaaa gagaatacgc agagagggca gtaaggcagg gaaacgatga      240
gcagacgacg ggatgctagc agcgaagct  gagaagaaga gaccctcaa  aagcaaaaga      300
ataagaaagc caaagcataa aagaaaagcg caatcaagga aacggcgct  agaagaccga      360
gacagtgaga cgagcacacg ctaaaaaaga gagacaatcc agaaagaaga cagtggacia      420
gagaccgggt agcaccagag gacaatgaag gcgggacggc ataacgccc  aaacgggtaa      480
gacgggccaa cgataaagga agacgaaggt caacgcagaa gggaatccag aaacacagac      540
ataaagaaaa aaagaaatgg cagccggaga agctgagacg ggaccgcagt gcgagagaag      600
agacgtgcgg aaaagagaag gtcgcgcaag acyggcgctg accygacaaa ggatgggaaa      660
acgaccggat ggagagcgcc cagagtcaga aacggaagag atgaaacca  gaaagactcg      720
ccacaagaag gagtcgggaa acacaagaat aaagcagctg gggaatagcg gacgaacgga      780
atagaggccc aaagaactgg gagaggacag cgtacagga ggaacgggat agcgaagaac      840
caacgtgcga aaagaccggc ataaaacaga aaggcaatag cagacagcaa aaatagagag      900
acaacgacgt agaaagaggg aaacgtcgca aaaaaccggg ctagaggagg accaacatag      960
cgagacacca agatgaagca acggagaagt cacacaaaga gacagtcgaa aagcaagaa      1020
atagcccagc aagcgaatac ccaacaagaa gagtcaacag gagaagagt  ccaaggagaa      1080
gcaagtacac aagacaaaag gtagcagaca caagcaatgc acgggcaggg aaatgcgaaa      1140
cagagaggat aagaagatct aaaatgcatg gtctgtgtga tgttggctg  tgtgatgtt      1200
gtctgtgtga tgttggctg tgtgatgttc tgcagaagct t      1241

```

5 <210> 33

<211> 1106

<212> ADN

<213> Secuencia Artificial

<220>

10 <221> fuente

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

<400> 33

ES 2 684 317 T3

ggatccgacg ggaacaaaga gaaagcacia gtaaaaggca acgacgcgca cgaggtcaga	60
acgcaaacag gaagaagaat gcaacaaagg agcgccgaac ggatacgcgg gccaacacga	120
agaggaatga aagggaaacg aaacaaaaaa gtgaagaggc caaaagagca cagcatggag	180
gaagccacgg ccgaaaaact acaccgggaa ggaccaaaacg cgatggcaga cagagcgcag	240
acagggatag agccaccacg gaagcacgca gtcagggaga aaagggcaca gacactacga	300
aaaagggcga caagagggct gggagaggga acgggcagaa acataaaacc caagagccga	360
ccagaagtaa aaggcgatat caacacagac gtaaagcaag cagaacaaga gcacataagc	420
aagaaggaga ggagacacct accagagaga aaaacccgaa ccgtgcggaaggaggcgga	480
gaagggataa aagagcacga acccaagagg atgaaaggac aagagagaac gaaagtccaa	540
cacagaaagg ccaaaggaat aaaaagcaac agccaaaaaa aagtaagaaa gcaagacaaa	600
gcgagaatga aaagcgaagg cacaaaagaa atagacagag ggaacagcgg aaagatccca	660
cgggcgcag caaccggact ccacgggaac acgaggagaa acgtgaaaac cgagcggcaa	720
cccagactca gcggaacaga agacaaaaaa atacgggcag ggcgaggaaa gaaaataggg	780
ccaacaaaaa ggaagcagat gaaaaacaca aaaacccacc acatcacgcg ggccgagagc	840
agcaacgtgg cagcacgagg aggaggagcc ctagaacgca aaggcgaaaa gacgctgcga	900
ggtaccgggc caaagaaaat aaaggaacgc aaacaaagga gcatcaaagg gaagaacaa	960
aagaaaataa aaaaggagaa gaaagggag ataaaaggaa acggaaaaaa gaaactcgcg	1020
acaaagacgg cagcaaaggt agaggaaaag gaagaggaca cactgaaaga gcaccaggag	1080
aaagacctag cagaagcacg ctcgag	1106

REIVINDICACIONES

1. Una composición que comprende:
 - (a) una primera sonda nanoinformadora etiquetada de manera única que comprende:
 - (i) una región específica de un objetivo único;
 - 5 (ii) una región que comprende un único nanoinformador designado

en la que dicho nanoinformador comprende una cadena principal de ácido nucleico de diseño único de cadena sencilla, comprendiendo dicha cadena principal una pluralidad de regiones de fijación de la etiqueta diseñadas covalentemente unidas entre sí en una combinación lineal única, en donde cada región de fijación de la etiqueta de cada cadena principal se selecciona individualmente y es diferente de las otras regiones de fijación de la etiqueta en esa misma cadena principal,

10 en la que las regiones de fijación de la etiqueta se seleccionan de una población de secuencias de polinucleótidos diseñadas racionalmente,

en la que la secuencia de polinucleótidos se hibrida con una secuencia complementaria de polinucleótidos que tiene unida a la misma una o más moléculas detectables,

15 en la que cada región de fijación de la etiqueta tiene un contenido de G/C de aproximadamente 50%, y la secuencia complementaria de polinucleótidos tiene una relación de G/C de al menos aproximadamente 1/1 o aproximadamente 3/2,

en la que cada nanoinformador tiene una señal detectable que la distingue de otros nanoinformadores en dicha composición; y
 - 20 (b) una segunda sonda que comprende
 - (i) una región que se une selectivamente a la molécula objetivo y en un sitio diferente de la región específica objetivo única de la primera sonda; y
 - (ii) una etiqueta de afinidad.
- 25 2. La composición de la reivindicación 1, en la que dicha pluralidad de regiones de fijación de la etiqueta se genera a partir de las secuencias de plantilla seleccionadas del grupo que consiste en las SEQ ID NOS: 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 composición de la reivindicación 1 o 2, en la que,
 - (a) las regiones de fijación de la etiqueta comprenden un contenido similar de adenina, y, opcionalmente, las bases de adenina están separadas al menos cada 8 a 16 bases nucleotídicas, o
 - 30 (b) las regiones de fijación de la etiqueta comprenden un patrón regularmente repetido de bases de adenina y, opcionalmente, las bases de adenina están separadas al menos cada 8 a 16 bases nucleotídicas, o
 - (c) las regiones de fijación de la etiqueta comprenden un contenido de timina de aproximadamente 35-45%.
4. La composición de la reivindicación 1 o 2, en la que la secuencia complementaria de polinucleótidos comprende una secuencia de polinucleótidos de ARN.
- 35 5. La composición de la reivindicación 4, en la que,
 - (a) la secuencia de polinucleótidos de ARN comprende al menos una base de uracilo modificado con aminoalilo,
 - (b) la secuencia de polinucleótidos de ARN comprende una pluralidad de bases de uracilo modificadas con aminoalilo que están separadas a aproximadamente un promedio de cada 8 a 16 bases en dicha secuencia de polinucleótidos de ARN, o
 - 40 (c) la molécula detectable está unida a la base de uracilo modificado con aminoalilo o en la que la molécula detectable está unida a cada una de las bases de uracilo modificadas con alilo.
6. La composición de la reivindicación 1 o 2, en la que las moléculas detectables son colorantes fluorescentes.
7. La composición de la reivindicación 1 o 2 que comprende además una región constante, en la que la región constante comprende una pluralidad de secuencias repetidas de nucleótidos.

8. La composición de las reivindicaciones 1 o 2, en la que la temperatura de fusión (T_m) de dichas secuencias complementarias de polinucleótidos cuando se hibridan con sus regiones de fijación de la etiqueta es de aproximadamente 80°C o superior.
9. La composición de la reivindicación 1,
- 5 en la que cada región de fijación de la etiqueta comprende aproximadamente 800 a 1.300 bases nucleotídicas.
10. Un método para determinar la presencia de al menos una molécula objetivo en una muestra, que comprende:
- (1) formar al menos un complejo molecular que comprende:
- (a) al menos una molécula objetivo;
- (b) al menos una sonda que comprende
- 10 i) una región única específica del objetivo; y
- ii) una región que comprende un único, nanoinformador diseñado,
- en la que dicho nanoinformador comprende una cadena principal de ácido nucleico de diseño único de cadena sencilla, comprendiendo dicha cadena principal una pluralidad de regiones de fijación de la etiqueta unidas covalentemente entre sí en una combinación lineal única;
- 15 en la que cada región de fijación de la etiqueta se selecciona individualmente y es diferente de las otras regiones seleccionadas de fijación de la etiqueta en la cadena principal;
- en la que cada región de fijación de la etiqueta se hibrida con una secuencia complementaria de polinucleótidos que tiene unida a la misma una o más moléculas detectables;
- 20 en la que las regiones de fijación de la etiqueta se seleccionan de una población de secuencias de polinucleótidos racionalmente diseñadas;
- en la que cada secuencia complementaria de polinucleótidos ha designado una molécula detectable específica;
- en la que cada región de fijación de la etiqueta tiene un contenido de G/C de aproximadamente 50%, y la secuencia complementaria de polinucleótidos tiene una relación de G/C de al menos aproximadamente 1/1; y
- (c) una segunda sonda que comprende:
- 25 (i) una región que se une selectivamente a la molécula objetivo y en un sitio diferente de la región específica única del objetivo de la primera sonda; y
- (ii) una etiqueta de afinidad; 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 objetivo en la muestra.
- 30 11. El método de la reivindicación 10, en el que cada región de fijación de la etiqueta se genera a partir de las secuencias de plantilla seleccionadas del grupo que consiste en las SEQ ID NOS: 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.
12. La composición de cualquiera de las reivindicaciones 1 a 11, en la que la primera sonda comprende además una etiqueta de afinidad.
- 35 13. La composición de cualquiera de las reivindicaciones 1 a 12, en la que la etiqueta de afinidad se selecciona del grupo que consiste en biotina, avidina/estreptavidina, digoxigenina, epítipo HA, epítipo myc y epítipo FLAG.
14. La composición de la reivindicación 13, en la que la etiqueta de afinidad es biotina.
15. La composición de la reivindicación 13, en la que la etiqueta de afinidad es digoxigenina.

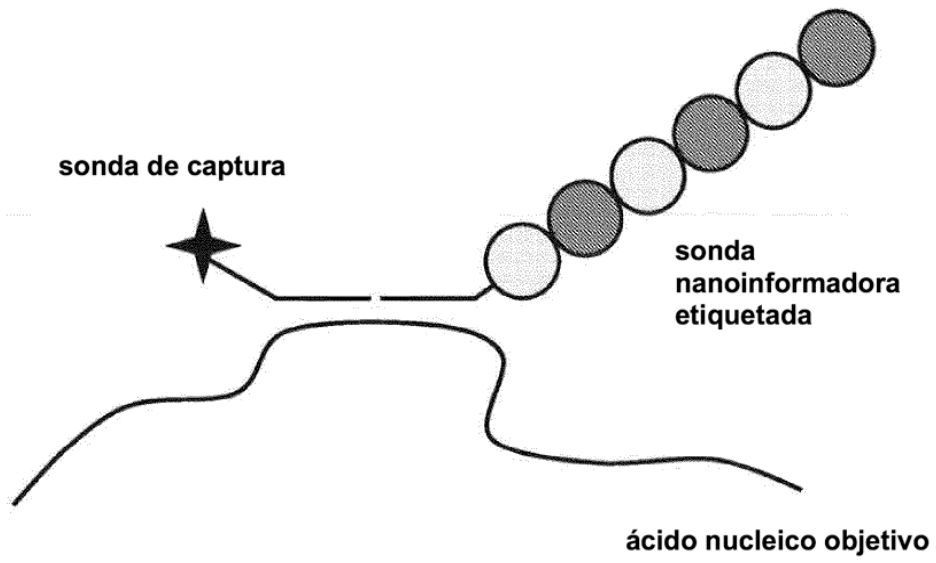


Figura 1

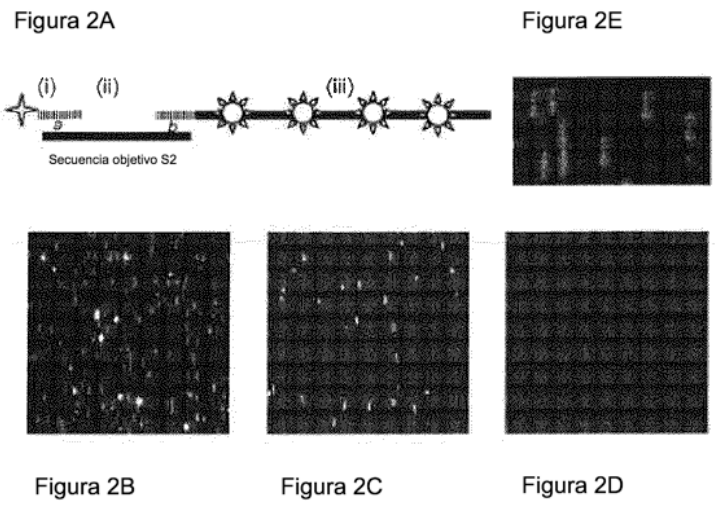


Figura 2

Figura 3A



Figura 3B

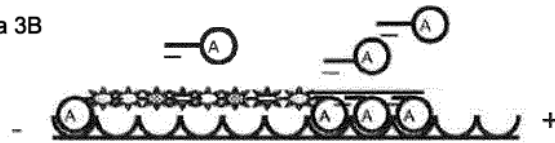


Figura 3

Figura 4A

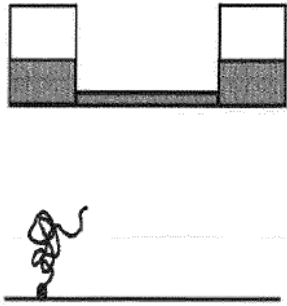


Figura 4B

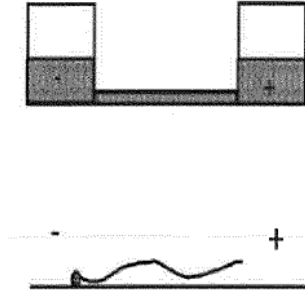


Figura 4C

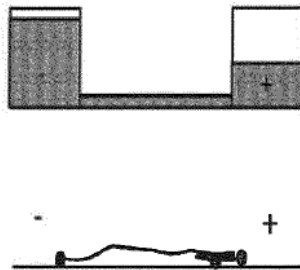


Figura 4

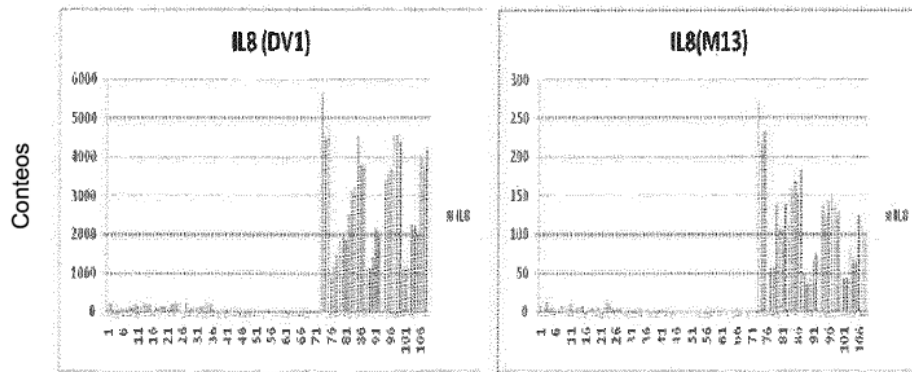
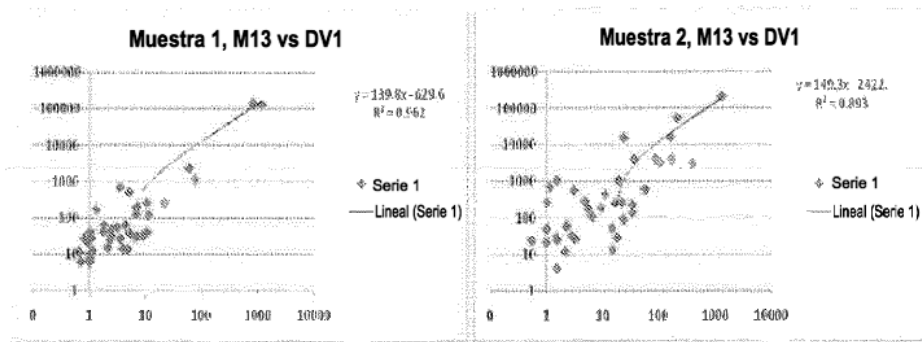


Figura 5

40 informadores de 2 muestras 'representativas':



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

Figura 6

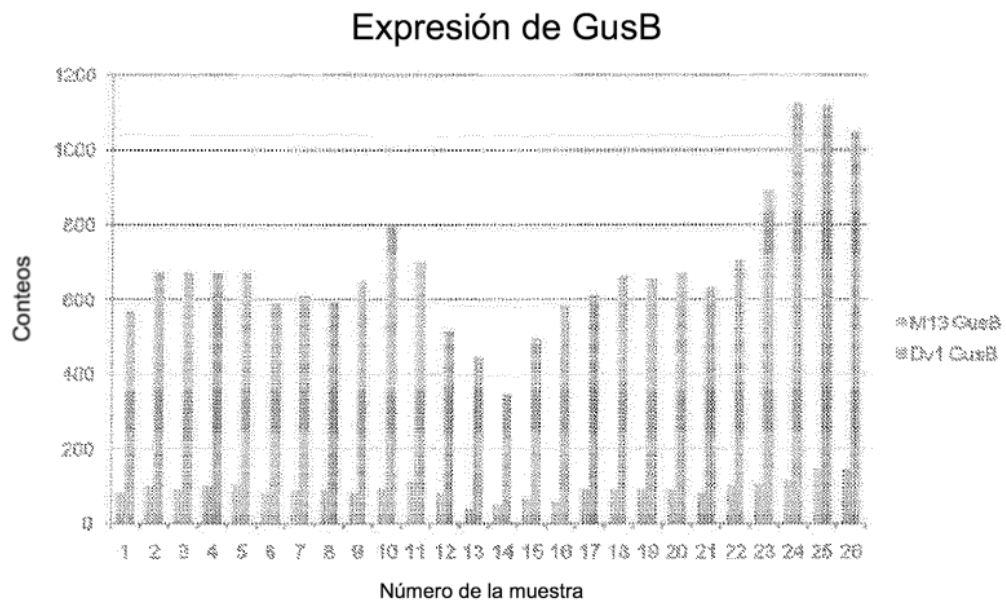


Figura 7

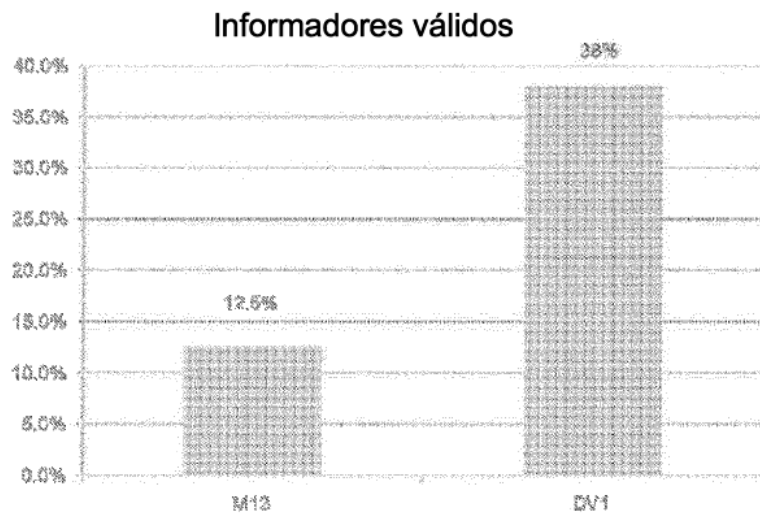


Figura 8

AAGCTTGGCCACACAGCTTGAGTATTCTATAGTGTACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGT
TTCCTGTGTGAAATTTGTTATCCGCTCACAATCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCT
GGGTGCCTAATGAGTGAGCTAACTCACAATTAATGCGTTGCGCTCACTGCCCGCTTCCAGTCGGGAAACC
TGTCGTGCCAGCTGCATTAATGAATCGGCCAACCGCGGGGAGAGGCGGTTTTCGCTATTGGGCGCTCTCCG
CTTCCCTCGCTCACTGACTCGCTGCGCTCGGTGCTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGG
TAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCA
GGAACCGTAAAAAGCCCGCTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCAGAAAAATC
GACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATAACCAGGCGTTTCCCCCTGGAAGCTCC
TCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAAGCGTGG
CGTTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGGTAGGTGCTTCGCTCCAAGCTGGGCTGTGTGC
ACGAATCCCGCTTACGCGGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCCGTAAGAC
ACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAG
AGTTCTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGC
CAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGTAGCGGTGGTTTTT
TTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAGATCTTTGATCTTTTCTACGGGGT
CTGACGCTCAGTGGAAACAAAACTCAGTTAAGGGATTTTGGTCAAGAGTATCAAAAAGGATCTTACCT
AGATCCTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTT
ACCAATGCTTAATCAGTCAGGCACCTATCTCAGCGATCTGTCTATTTGTTTATCCATAGTGCCTGACTCC
CCGTGCTGATAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCGAGTGTGCAATGATACCGGAGACC
CACGCTCACCGCTCCAGATTTATCAGCAATAAACAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCTG
CAACTTTTATCCGCCCTCCATCCAGTCTATTAATTTGTTGCCGGGAAGCTAGAGTAAGTAGTTCCGCCAGTTAATA
GTTTTCGCAACGTTGTTCCATTGCTACAGGCATCGTGGTGTACCGCTCGTCTGTTGGTATGGCTTCATTCA
GCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCG
GTCCTCCGATCGTTGTGAGAAGTAAGTTGGCCGCGAGTGTATCACTCATGGTTATGGCAGCACTGCATAAT
CTCTTACTGTGATGCCATCCGTAAGATGCTTTTTCTGTGACTGGTGTACTCAACCAAGTCACTCTGAGAAAT
AGTGTATGCGGCGACCGAGTTGCTCTTGCCCGCGCTCAATACGGGATAATACCGCGCCACATAGCAGAACTT
TAAAAGTGTCTCATCATTGAAAAAGCTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCA
GTTGATGTAACCCACTCGTGCACCCAACTGATCTTACGATCTTTTACTTTTACCAGCGTTTTCTGGGTGAG
CAAAAACAGGAAGGCAAAATGCCGCAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCT
TCCTTTTTCAATATATTGAGCATTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTT
AGAAAAATAACAAATAGGGTTCCGCGCACATTTCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTA
TTATCATGACATTAACCTATAAAAATAGGCGTATCAGAGGCCCTTTTCGCTCAGCGCGTTTCGGTGTGACG
GTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCAAGCTTGTCTGTAAGCGGATGCCGGGAGCAGAC
AAGCCCGTACAGGCGCGTACAGCGGTGTTGGCGGGTGTTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGA
TTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAAATCCGCATCAGGA
AATTGTAAGCGTTAATATTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAAACCAATA
GGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGGATAGGGTTGAGTGTGTTCCAGTTTG
GAACAAGAGTCCACTATTAAGAAGCGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGG
CCCCTACGTGAACCATCACCTAATCAAGTTTTTTGGGGTTCGAGGTGCCGTAAAGCACTAAATCGGAACCC
TAAAGGGAGCCCGGATTTAGAGCTTGACGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGC
GAAAGGAGCCGGGCTACGGCGCTGGCAAGTGTAGCGTCAAGCTGCGCGTAAACCAACACCCCGCGCT
TAATGCGCGCTACAGGCGCGTCCATTGCGCATTGAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGG
GCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGG
TTTTCCAGTACGACGTTGTAACGACGGCCAGTGAATTGTAATACGACTCACTATAGGGCGAATTAGTG
GATGAGAAGACCTGCAGAAAGAATTC (SEQ ID NO:27)

Figura 9

espaciado de 8

TCTAGAGAGAAAAGGCTACGAAGATCCGGACATCAAAAGATCAAAACATGCGGAGATCGAC
 ACATCCAAAAATGAACAGATAACAGGATCAGAGGGTCAAGAAGTAAAACCATCGAGGAGT
 CGCAGACTACGAAGATGGACAAGTGAAGAACAGCTGCGATGGAGAACATAAGGAGTCGGG
 AGGTACGGCAGTGAAGAAATGTGCACATGAAAAGCTGCCCGGTGCGCAACTGGAGGACT
 AGAAGAATGGCACACTGACCCGGTGAAGAGCTGAGGAGGTGCGAGAATAGGGCGATAGGG
 AGGTAACCCGATGCGGGACTAAAGAGATACGGACATGGACGGATAGGAGAGTCCGCAAGAT
 GGAAGGATCAAAAGTGAAGCCGGTCCAAAGCATAACGGGGTAAACGGATCAGAACGTAAAG
 CGGTGGAAAAATACCCGAGTAGGAAGATAGCAAAATACCCAGGTTCGAGCGATAGCCACT
 CGGCGAGTCCAAAACTGAAAAAATGGAGAGATGAGACACTAGCGGCATCGGACAGTAGAG
 CAGTGAACGCATCAGAGAGTAGAAAAGTGAAGAAGTGGACGGTAAAACATAACAGAGT
 CAGAGGGTACACCCATACGACCTACCGCAATACAGGCTAAGACGGTACAGAGGATGGAC
 GGCTCGGGAACCTACGCCGTGCCAAAGTGAAGAAGTACCAAAATAGGAGCGTAAACGGCGT
 AGAGAGGTAGACGAGTGAAGCAGGTGGGCAGCTCCCAAGGTCCGAGCATGGACAGGTGGCA
 AAATCCAGCGGTGAGAACTACGAGACTGGCCGGATGAAGAAATAAGGCGTACAGAGAT
 ACGGGAGTGACAAGATGGACCACTAGCGAAGTCCGAGGGTGAAGACGTGAACAACCTAGAC
 CAGTACGGCACTAGCAGAATAGACCGATGAAAACCTGAAAAGATAACAGAGTGAAGGCAT
 CGGAGGGTTCGAGAAGTAAGAAGCTGGAGCCATACCGAATAAAGCAGTACAAGGGTGGAC
 AAATGAGGAAATCCGGAGGTCCCACGGTGAAAACGTAGGGAGGTGAACCAAGTGGGAGAGT
 CAACCAATACGAAGATGCGAGGATCC (SEQ ID NO:28)

Figura 10A

espaciado de 10

CCATGGGAAAATGAGAGCCGCTCGACAAGGATAAAGAAAAGTGCGCCAACATCCGCCACG
 CTAAGCGGCCCTAGGAGCCGTTAAACAAGCATGAAGAAACATGCGAGGCAATGACACAAA
 CTCCAGGCGGGTGGAAAAGACTCGACGGAGATGCCACGCAATAGCGAAACCTGGAAAGAGA
 GTGGAAGCAAATAGGAAACAGTAGAGCAAAATGCCCGAGGATAGAAAACAGTGGCGCGCA
 ATAAGCAGGAATAAAGCACAATAAAAAGAAGTAAACCGCACTAAAACAGCATCCAGACCG
 CTGAAAAGAAAATGGCGGCAATGAAAAAAGTCAACAACCGGTGCAACAACCTGAAAAAGG
 GTAGAGAGCCCTCCCGGCAATAAAAAGACGGTAAGAGGAAATCCCGAAAAGTGCAAAAAG
 GTCAAAACGGGTGCAACAAAATGCCAAGCGTGAAGAAACTGAGGACGGATAAAAAGACA
 ATGCGGACGGGTACGACACGATAAGGAAAGATAAGGCGAAGTAAAAGCCGGTGCGCCAAA
 ATACAACAGGCTGCCAGAAAGTACCGAGCGCTAGGCGGGAATGAGCGAGGGTAAAAGAGA
 ATACACAACCCTGCAAGAAAGTCCGGGAAGCTAGGACCGGATACAGAGGAATAAGGAAGG
 GTACACAGCAATACGAGCAGATGAACGAGCATAACAGAGAGATATCAGGCCATAAAAAACG
 GTAACAAGAAGTAGGAGGAGGTGCGAAGGCCCTGGACCAAAATGGGACGAAGTAGGAGGGA
 ATACACAACGATGGGACGAGATAGAGCCGGTGGCCAAAGTAGCACGGGATACACAGAG
 ATACAGCGACGTACCGGAAAGTAGGAGAGACTGGAAACCGTAAGGCAAGATAAGCGAGG
 GCTAGCAGAATAAAACAGACTGAACACGGATCAAGGCGAGTAGGCACGAGTACACGCCG
 ATACGGACAAGTAGGCCCGCAGTACGGAGAAGTAAACAGGGATGAGAGACAGTGAACAAGG
 CTGAGACGAAATCACAGCAATACAACAAGTAGCGGCAATAAACGGGAGTGAACCGCG
 ATAGCAGGAAGTACGGACAGCATGC (SEQ ID NO:29)

Figura 10B

Figura 10

espaciado de 12

GAATTCGAGTACCACCTGATGAGCAGCAAGGTAACGGCGAAGATGAAAAGAAAAATGGA
 CGGCACGTACAAAAAGGATGGCACCAGACATAGAAGAGGAGCTCAGCAAACGGCTGGA
 AGCGAAGTCCCAGGAAGCCTACCAAGGACAGATATCAAAGAAATAAGAAACAACTGAC
 AAGCGCGTCGGAAGCGACGTACAAGCGGCAATGGCAGAGAGAGTAAAACAAACAGTGAA
 CCAACAATAAACGCCAGGTAGGGCGAAAGGTAACAAAAGGATCGAAAAGCAGTAAAG
 CGGGAGATGGAACAACGGTGCGAAAACAGTGAAAAAGGGAGTCCAAAGCACCATGAA
 ACCACAATACGACACACCCTGGGAGGACACGTGGAGGAGCCGGTACGAAAGCAGGGTACA
 GACAACATGCAACAGAGGTAAGGAACCGATAAAGAACGACATGGAAGGAAACGTAGC
 ACCCGAGTAGAGGAACGGATAGCGAACGACGTGCAACGCCACCTAGAACGAGAAGTGGC
 AGGGAAATAAACGGAGAGGTCCCACGGCCGATGAGCAGAAAAAGTGGAAAAGGACCTGAA
 AGAAAAATGAGGACCACAATAAAGGACGGAAATAAAAAACACACTCCCAGAGGAAATACA
 CGAGAAATAGGGAAACAGGATGCAAAGAAGCATCAGAAAGGCAGTGCCAAAGGGAAATCGC
 CGCAAGCTAAGACAGGCCCTAGACCGCAAGTAAACGAACTCGGAGGGCGAATCAG
 CCCACCATGCAGAGCAAAAATAAGAGACCAGATGCGCGGCAACCTGAGACCGAACCTAAA
 GAAAGGATAGAGGGAGAGGTCCAAGAAACGATGGCGGGCACCATGAGAAAGGAAGTCAA
 GTGCACCTAGCAACACAGATACCACGAGACATCGAAGCGAAGATACGGAAAAAGCTGGA
 AGGGCAATGAACAGAAAGCTACGGAAAGAGTCAAGGACACAATGAGGGAGCAAGTCGA
 AGCGAGGTGGCCCGCAAATAGGAGCAACAATGCGAGGCAAGTGGGCACAACGATGAA
 GACCAACTAACACAGCCACTTCTAGAGTGCACGGATCCCTCGAGGGTACCAACCATGGA
 GCTAGCAAGCATGCAAACTGCAGAAGCTT (SEQ ID NO:30)

Figura 11A

espaciado de 14

CTCGAGGAAAGACAGAGGGTAGCAGGCCACAGATCAAGGCGGCGAAATGGCCCGGAGCA
 GTAAGCACCACACGGTCACCAAGAGAGAATCAGAAGAGGCAAGTAAAAAAGCGCGTA
 GGCACCCAAAGGTGGACGCACAACAGTACGCGGAAAGAAGTCAAGGCGGGCGCTCAGAA
 AGCGAAGTACACAACCAGACGGTACCGGAACACGATGCGCCAACAAGAAATCAGCGGAAA
 ACATACCGGCGGACGGGTGCAAAACACGAAAATGAAACCGCAGAAGTAAAGAGAAAAGGA
 CGCAGGAAGAGGATCAAGACAGAAAAATCGACAAAGCAAAGTGGACCGAAGAAGATCGA
 AAGAAGCACTCAGGACAGCTGAATCGCACAGACACGGTGAACAAGACGAAAATGGACGAC
 ACCGATAAGCGAAAAAAGATACCGGCGAGCGAGTCAAGAACACGAAGTCAAGGCGCAAG
 ATAGAGCAAACGAAGTAAACAGGACAGAAATAGGAAGGGCAAAGTGGGAGAGAAAAGGCTC
 GAACCGACGACTGCACCCGAACAACGTGGGAGCCCACGGTACAAAACGCCAGATCGAAG
 CCAGGAGTCAAGAAACAAAAGTAAACGGCGCAGAAATAGAGAGAGGAGGCTAAAGCGGAA
 CAGTGAAGGAAAAGAAATAACGGGCCAAGGTGAAGAAAAGAAGTGGCAGGAACGGAC
 ACAGCCCAGCAACTGGGCAAAAGCGGCTGACGAGAACAAGATAGAGAACGAGGAATGAC
 AGACACAGATAGAAGACAAGACTGCGGACAGGCAAGTAAAGAGAAGGAAAATAAAGCGA
 AGAAGTAGGCCCGCCGAAATGCAAAAGAAAAAGTGAACAAGAGAAAATGACCGGCGGGA
 GTAACACGCCCAGGATGACGGAGCACCATAAGAAAAAACACGTAGGAGAGGGAACATC
 GGACAAGACAATCCCAACCCGAAGATAAACACCGAGAGGTAGACAGCGCAGACTGGAGC
 ACAGCCGTGGAAAGGGAAAGGTACGCAAAACGGACCTAAAAAAAACAAAGTCAAGGAAAA
 GGCTAACGGAAACAAAGCTGGCCATGG (SEQ ID NO:31)

Figura 11B

Figura 11

espaciado de 16

GAATTCCTCTAGAGTGCACGGATCCCTCGAGGGTACCAACCATGGAAGCTAGCAAGCATGC
 GAGGAAAGCAAGGAAAATGAAGAACGAAGCGGGTCCGAACAGGGAAGACATAAAAAACAGG
 CACCATGACACCAACCGCGGAATGACACACACAAAGGTAGCGAGGCACGCACGTAAGAGG
 CACAAGAAATGAAAGAAAAAGAGAAATACGCAGAGAGGGCAGTAAGGCCAGGGAACGATGA
 GCAGACGACGGGATGCTAGCAGCGAAAGCTGAGAAGAAGAGACCCCAAAAAGCAAAAAGA
 ATAAGAAAGCCAAAGCATAAAAAGAAAAGCGCAATCAAGGAAACGGCGCTAGAAGACCGA
 GACAGTGAGACGAGCACACCGCTAAAAAGAGAGACAATCCAGAAAAGAGACAGTGGACAA
 GAGACCGGGTAGCACCAGAGGACAATCAACCGGGACGGCATAAACGCCAAAACGGGTAA
 GACGGGCCAACGATAAAGGANGACGAAGGTCAACGCAGAAAGGGAATCCAGAAACACAGAC
 ATAAAGAAAAAAAGAAATGGCAGCCGAGAAAGCTGAGACGGGACCGCAGTCCGAGAGAAAG
 AGACCTGCCGAAAAGAGAAGGTCCGCGCAAGACGGGCGTGGACCCGACAAAGGATGGGAAA
 ACGACCGGATGGAGAGCGCCAGAGTCAGAAAACGGAAAGATGAAACCCAGAAAAGACTCG
 CCACAAGAAGGAGTCCGGAAACACACAATAAAGCAGCTGGGGAAATAGCCGGACGAACGGA
 ATAGAGGCCCAAAGACTGGGAGAGGACAGCCGTACAGGAGGAAACGGGATAGCGAAGAAC
 CAACGTGCCAAAAGACCGGCATAAAACAGAAAAGGCAATAGCAGACAGCAAAAATAGAGAG
 ACAACGACCTAGAAGAGGGGAAACGTCGCAAAAAACCGGGCTAGAGGAGGACCAACATAG
 CGAGAACCAAGATGAAGCAACGGAGAAGTACACAAAAGAGACAGTCCGAAAAGCAAAAGAA
 ATAGCCCAAGCAAGCGAATACCCAAAGAAAGAGTCAACAGGAGAAAGAGTCCAAGGAGAA
 GCAAGTACACAAGACAAAAGGTAGCACACACAAGCAATGCACGGCAGGGAAAATGCCAAA
 CAGAGAGGATAAGAAGATCTAAAATGCATGGTCTGTGTGATGTTGGTCTGTGTGATGTTG
 GTCTGTGTGATGTTGGTCTGTGTGATCTTCTGCAGAAGCTT (SEQ ID NO:32)

Figura 12A

espaciado de 24

GGATCCGACGGGAACAAGAGAAAGCACAAGTAAAAGGCAACGACGCGCACGAGGTCAGA
 ACGCAAAACAGGAAGAAGAAATGCAACAAAGGAGCGCCGAACGGATACGCGGGCCAACACGA
 AGAGGAATGAAAGGGAAACGAACAAAAAAGTGAAGAGGCCAAAAGAGCACAGCATGGAG
 GAAGCCACGGCCGAAAACACTACACCGGGAAGGACCAAACGGATGGCAGACAGAGCGCAG
 ACAGGGATAGAGCCACCCAGGAAGCACGCAGTCAGGGAGAAAAGGGCACAGACACTACGA
 AAAAGGGCGACAAGAGGGCTGGGAGAGGGAACGGGCAGAAACATAAAACCCAAGAGCCGA
 CCAGAAGTAAAAGGCGATATCAACACAGACGTAAAGCAAGCAGAACAAAGAGCACATAAGC
 AAGAAGGAGAGGAGACACCTACCAGAGAGAAAACCCGAACCGTCCGGAAAGGGCGGCGA
 GAAGGGATAAAGAGCACGAACCCAAAGAGGATGAAGGACAAAGAGAAACGAAAAGTCCAA
 CACAGAAAGGCCAAAGGAAATAAAAAGCAACAGCCAAAAAAAAGTAAAGAAAGCAAGACAAA
 GCGAGAATGAAAAGCGAAGGCCAAAAAGAAATAGACAGAGGGAAACAGCGGAAAGATCCCA
 CGGGCGACAGCAACCGGACTCCACGGGAACACGAGGAGAAACGTTGAAAACCGAGCGGCAA
 CCCAGACTCAGCGGAACAGAAAGACAAAAAATACGGGCGAGGGCGAGGAAAAGAAAATAGGG
 CCAACAAAAACGAAGCAGATGAAAAACACAAAAACCCACCACATCACCGGGCCCGAGAGC
 AGCAACGTGGCAGCACGAGGAGGAGGAGCCCTAGAACGCAAAAGCGAAAAGACGCTGCCA
 GGTACCGGGCCAAGAAAATAAAGGAACGCAACAAAGGAGCATCAAAGGGAAAGAAACAA
 AAGAAAATAAAAAGGAGAAAGAAAGGAAAGATAAAAAGGAAACGGAAAAAGAAAACTCGCG
 ACAAGACGGCAGCAAGGTAGAGGAAAAGGAAGAGGACACACTGAAAAGAGCACCAGGAG
 AAAGACCTAGCAGAAGCACGCTCGAG (SEQ ID NO:33)

Figura 12B

Figura 12