

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 372 427**

51 Int. Cl.:  
**C12Q 1/68** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **07814636 .2**  
96 Fecha de presentación: **31.08.2007**  
97 Número de publicación de la solicitud: **2069537**  
97 Fecha de publicación de la solicitud: **17.06.2009**

54 Título: **MÉTODO PARA PRODUCIR SONDAS DE ÁCIDO NUCLEICO.**

30 Prioridad:  
**01.09.2006 US 841896 P**  
**02.03.2007 US 892571 P**

45 Fecha de publicación de la mención BOPI:  
**19.01.2012**

45 Fecha de la publicación del folleto de la patente:  
**19.01.2012**

73 Titular/es:  
**VENTANA MEDICAL SYSTEMS, INC.**  
**1910 E. INNOVATION PARK DRIVE**  
**TUCSON, ARIZONA 85755, US**

72 Inventor/es:  
**FARRELL, Michael**

74 Agente: **Arias Sanz, Juan**

**ES 2 372 427 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

Método para producir sondas de ácido nucleico.

Referencia cruzada a solicitudes relacionadas

### CAMPO

- 5 La presente divulgación se refiere al campo de la detección molecular de secuencias diana de ácido nucleico (por ejemplo, ARN o ADN genómico). Más concretamente, la presente divulgación se refiere a las sondas de ácido nucleico que forman una o más redes detectables en la secuencia diana, métodos para la producción de la sonda y métodos para su uso. En algunas realizaciones, las sondas divulgadas están sustancialmente libres de secuencia de ácido nucleico repetitiva.

### 10 ANTECEDENTES

- 15 Las técnicas de citogenética molecular, como la hibridación fluorescente *in situ* (FISH), la hibridación cromogénica *in situ* (CISH) y la hibridación con plata *in situ* (SISH), combinan la evaluación visual de los cromosomas (análisis cariotípico) con las técnicas moleculares. Los métodos de la citogenética molecular se basan en la hibridación de una sonda de ácido nucleico con su ácido nucleico complementario dentro de una célula. Una sonda para una región cromosómica específica reconocerá y se hibridará con su secuencia complementaria en un cromosoma en metafase o dentro de un núcleo en interfase (por ejemplo, en una muestra de tejido). Se han desarrollado sondas con diversos fines de diagnóstico e investigación. Por ejemplo, determinadas sondas producen un diagrama de bandas en los cromosomas que imita los procedimientos de tinte citogenético tradicionales y permite la identificación de cromosomas individuales para el análisis cariotípico. Otras sondas se obtienen de un único cromosoma y cuando están etiquetadas se pueden utilizar como «pinturas del cromosoma» para identificar cromosomas específicos dentro de una célula. Sin embargo, otras sondas identifican estructuras del cromosoma particulares, como los centrómeros o telómeros de los cromosomas.

- 20 Las sondas de una secuencia única se hibridan con secuencias de ADN de copia única en un gen o una región cromosómica específica. Estas son las sondas empleadas para identificar el gen o la región crítica cromosómica asociada con un síndrome o condición de interés. En los cromosomas en metafase, estas sondas se hibridan con cada cromátida, normalmente dando dos pequeñas señales discretas por cromosoma.

- 25 La hibridación de sondas de secuencia única ha hecho posible la detección de anomalías cromosómicas asociadas con numerosas enfermedades y síndromes, incluyendo anomalías genéticas constitutivas, como los síndromes de microdelección, translocaciones cromosómicas, síndromes de aneuploidía y amplificación genética, enfermedades neoplásicas, así como infecciones patógenas. Más frecuentemente, estas técnicas se aplican a preparaciones citogenéticas estándar en los portaobjetos de los microscopios. Por otra parte, estos procedimientos se pueden utilizar en portaobjetos de tejido fijado en formalina, sangre o frotis de médula espinal, y células fijadas directamente u otros aislados nucleares.

- 30 Por ejemplo, estas técnicas se utilizan frecuentemente para caracterizar células tumorales, tanto para los fines del diagnóstico como del pronóstico del cáncer. Numerosas anomalías cromosómicas han sido asociadas con el desarrollo del cáncer (por ejemplo, aneuploidías, como la trisomía 8, asociadas con determinados trastornos mieloides; translocaciones, como el reordenamiento BCR/ABL, en la leucemia mielógena crónica; y amplificaciones de secuencias de ácido nucleico específicas asociadas con la transformación neoplásica). Las técnicas moleculares pueden aumentar las pruebas citogenéticas estándar en la detección y caracterización de estas anomalías cromosómicas adquiridas. Por ejemplo, la FISH se ha utilizado para buscar el relapso temprano y la enfermedad residual de células que no se dividen. La detección inmunocitoquímica de las células cancerosas y las técnicas de FISH se han combinado para estudiar las anomalías cromosómicas en poblaciones de células definidas.

- 35 La presente divulgación proporciona sondas y métodos mejorados para la producción de estas sondas para su uso en aplicaciones de diagnóstico e investigación de la hibridación *in situ*.

### 40 RESUMEN

- 45 La presente divulgación se refiere a las sondas de ácido nucleico y métodos para su uso y producción. Las sondas corresponden a una secuencia de ácido nucleico diana (por ejemplo, una secuencia de ácido nucleico diana de ARN o genómico) y son adecuadas para el análisis molecular de esta(s) diana(s), por ejemplo, los métodos de hibridación *in situ*, como FISH, CISH y SISH. En determinados ejemplos, las sondas divulgadas ofrecen una mayor sensibilidad y especificidad, y una menor señal de fondo, en comparación con las sondas convencionales.

- 50 Las sondas se proporcionan en la presente solicitud. En un ejemplo, una sonda de ácido nucleico incluye una pluralidad de moléculas de ácido nucleico. Prácticamente toda la pluralidad de las moléculas de ácido nucleico incluye al menos una primera región de unión y una segunda región de unión, donde la primera región de unión y la

segunda región de unión son porciones contiguas y complementarias de las porciones no contiguas y secuencias únicas de una molécula de ácido nucleico diana. De este modo, las regiones de unión se pueden posicionar en la pluralidad de las moléculas de ácido nucleico de forma que el orden y/o la orientación de las regiones de unión sean diferentes del orden y la orientación de las regiones de unión de la secuencia diana. Por tanto, la pluralidad de moléculas de ácido nucleico se denomina en el presente moléculas de ácido nucleico fragmentadas, permutadas, concatenadas (FPC). En determinados ejemplos, las regiones de unión están sustancialmente libres de la secuencia no deseada, como las secuencias que resultan en una unión no específica aumentada de una sonda de ácido nucleico con una secuencia de ácido nucleico diana (por ejemplo, secuencias de ácido nucleico repetitivas encontradas en secuencias de ácido nucleico diana del genoma de los mamíferos, secuencias que codifican dominios conservados en secuencias diana de ARN, o secuencias homólogas en secuencias de ácido nucleico de genomas virales diana). La pluralidad de moléculas de ácido nucleico puede ser etiquetada, produciendo así una sonda etiquetada. En un ejemplo, la pluralidad de ácido nucleico está etiquetada utilizando la traslación por muescas, fragmentando así la pluralidad de moléculas de ácido nucleico, donde las moléculas fragmentadas se pueden utilizar como una sonda.

En algunos ejemplos, la sonda incluye una pluralidad heterogénea de moléculas de ácido nucleico. Prácticamente toda la pluralidad heterogénea de moléculas de ácido nucleico incluye al menos una primera región de unión con una primera secuencia de nucleótido y una segunda región de unión que cuenta con una segunda secuencia de nucleótido. La primera región de unión y la segunda región de unión son porciones contiguas y complementarias de las porciones no contiguas y secuencias de nucleótidos únicas de la molécula de ácido nucleico diana. La primera secuencia de nucleótido y la segunda secuencia de nucleótido de cada pluralidad de moléculas de ácido nucleico pueden diferir de la primera secuencia de nucleótido y la segunda secuencia de nucleótido en otras pluralidades de moléculas de ácido nucleico.

También se divulgan métodos para producir las sondas de la presente divulgación, así como sondas producidas por el método. En un ejemplo, las sondas se producen por un método que incluye la unión al menos de una primera región de unión y una segunda región de unión, produciendo así una pluralidad de moléculas de ácido nucleico. Prácticamente toda la pluralidad de moléculas de ácido nucleico incluye al menos una primera región de unión y una segunda región de unión contiguas, donde al menos la primera región de unión y la segunda región de unión contiguas son complementarias de las porciones no contiguas y secuencias únicas de la molécula de ácido nucleico diana. En determinados ejemplos, las regiones de unión están sustancialmente libres de la secuencia no deseada presente en la secuencia diana, tales como secuencias que resultan en un aumento de la unión no específica de una sonda de ácido nucleico a una secuencia de ácido nucleico diana (por ejemplo, secuencias de ácido nucleico repetitivas encontradas en secuencias de ácido nucleico diana del genoma de los mamíferos, secuencias que codifican dominios conservados en secuencias diana de ARN, o secuencias homólogas en secuencias de ácido nucleico de genomas virales diana). La pluralidad resultante de moléculas de ácido nucleico forma la sonda. El método también puede incluir la amplificación de la pluralidad de moléculas de ácido nucleico para producir una pluralidad de amplímeros de moléculas de ácido nucleico para formar la sonda. El método también puede incluir el etiquetado de la pluralidad de moléculas de ácido nucleico o amplímeros de la misma para producir una sonda etiquetada.

En determinados ejemplos, las regiones de unión se generan mediante una o más de las siguientes acciones: aislamiento de las regiones de unión de la secuencia de ácido nucleico diana; obtención de regiones de unión de la secuencia de ácido nucleico diana mediante hibridación sustractiva; o amplificación de las regiones de unión de la secuencia de ácido nucleico diana. Por ejemplo, las regiones de unión se pueden aislar o amplificar de una secuencia de ácido nucleico diana presente en un vector.

Los métodos para utilizar las sondas divulgadas incluyen, por ejemplo, la detección (y, en algunos ejemplos, la cuantificación) de una secuencia de ácido nucleico diana. Por ejemplo, el método puede incluir la puesta en contacto de las sondas divulgadas con una muestra que contiene moléculas de ácido nucleico (por ejemplo, ADN o ARN) en unas condiciones suficientes como para permitir la hibridación entre las moléculas de ácido nucleico de la muestra y la pluralidad de moléculas de ácido nucleico de la sonda. Se detecta la hibridación resultante, donde la presencia de hibridación indica la presencia de la secuencia de ácido nucleico diana.

También se divulgan los juegos que incluyen las sondas y/o materiales de partida y/o reactivos para la elaboración de las sondas.

Lo anterior y otros objetos y características de la divulgación resultarán más evidentes en la descripción detallada siguiente.

#### BREVE DESCRIPCIÓN DE LAS FIGURAS

La FIG. 1A es una ilustración esquemática que muestra cómo se puede generar la sonda 300 de una secuencia de ácido nucleico diana 10.

5 La FIG. 1B es una ilustración esquemática que muestra cómo las moléculas de ácido nucleico FPC se pueden producir de las regiones de unión que están completa o sustancialmente libres de las secuencias de ácido nucleico no deseadas (por ejemplo, las secuencias de ácido nucleico repetitivas que se encuentran en las secuencias de ácido nucleico diana del genoma de los mamíferos, secuencias que codifican dominios conservados en secuencias diana de ARN, o secuencias homólogas en secuencias de ácido nucleico de genomas virales diana), posteriormente amplificadas y etiquetadas, para producir una sonda.

La FIG. 2A es una ilustración esquemática que muestra una sonda 510, que incluye una pluralidad de moléculas de ácido nucleico FPC 520, 530, 540, 550 hibridadas 538,548,558 con una secuencia de ácido nucleico diana 500.

10 La FIG. 2B es una ilustración esquemática que muestra cómo una pluralidad de moléculas de ácido nucleico FPC 600, 602, 604, 606, 608, 610 amplifica la señal resultante de la hibridación de moléculas de ácido nucleico FPC directa o indirectamente con la secuencia de ácido nucleico diana 500.

15 La FIG. 3 es una ilustración esquemática que muestra ejemplos de secuencias de ácido nucleico no deseadas (por ejemplo, secuencias de ácido nucleico repetitivas que se encuentran en las secuencias de ácido nucleico diana del genoma de los mamíferos) y regiones de unión correspondientes a una secuencia de ácido nucleico diana (por ejemplo, secuencia de ácido nucleico genómica diana). En el ejemplo mostrado, la secuencia de ácido nucleico diana tiene unos 200 kb. Las secuencias de ácido nucleico no deseadas intercaladas se muestran en negro y las regiones de unión únicas (por ejemplo, por genoma haploide) se muestran en gris.

20 La FIG. 4 es una ilustración esquemática que muestra la producción de moléculas de ácidos nucleicos FPC, añadiendo oligonucleótidos duplicados que incluyen una secuencia de nucleótidos fijada a regiones de unión de cadena doble antes de la unión.

La FIG. 5 es una ilustración esquemática que muestra la amplificación de una molécula de ácido nucleico FPC de muestra utilizando un cebador fijado (solo purina).

Las FIG. 6A-D son imágenes digitales que muestran la hibridación de un ejemplo de sonda de ácido nucleico FPC HER2 con muestras de tejido fijado

25 La FIG. 7 es una reproducción digital de una imagen de gel de agarosa que muestra la distribución del tamaño de una sonda de ácido nucleico FPC para el HER2 del genoma humano (banda A) y fragmentos de la misma, producidos mediante un calentamiento a 100°C durante 5 (banda 2), 30 (banda 3) o 60 (banda 4) minutos.

30 Las FIG. 8A-D son imágenes digitales que muestran secciones del xenógrafo BT474 HER2-positivas embebidas en parafina y fijadas con formalina (paneles de la izquierda), y tejidos de las secciones del xenógrafo MCF7 HER2-negativas (paneles de la derecha), teñidos con la sonda de ácido nucleico FPC para HER2 del genoma humano (panel A) y fragmentos de la misma, producidos mediante calentamiento a 100°C durante 5 (panel B), 30 (panel C) o 60 (panel D) minutos.

35 Las FIG. 9A y B son imágenes digitales que muestran secciones del xenógrafo de la célula Caski HPV16-positivas (A) y secciones del xenógrafo de la célula C3.3A HPVJ6-negativas (B) teñidas con una sonda de ácido nucleico FPC para ADN del genoma HPV16.

40 La FIG. 10 muestra imágenes de secciones de tejido renal humano embebidas en parafina y fijadas con formalina de cuatro individuos infectados por el virus BK («Riñón positivo») y dos individuos no infectados por el virus BK («Riñón negativo»), cada una de ellas teñidas con una sonda de ADN del genoma del virus BK permutado. La concentración de la sonda del virus BK, el aumento y el método de detección se muestran a la izquierda de las respectivas imágenes.

## LISTA DE SECUENCIAS

45 Las secuencias de ácido amino y nucleico recogidas en la lista de secuencias para acompañar esta solicitud se mostrarán utilizando abreviaturas de letras estándar para las bases de nucleótidos definidas en 37 C. F.R. 1.822. Solamente se mostrará una cadena de cada secuencia de ácido nucleico, pero la cadena complementaria se entiende que está incluida por cualquier referencia a la cadena mostrada (a menos que el contexto exija lo contrario). Todos los números de entrada en la base de datos de las secuencias mencionados en el presente se entiende que se refieren a la versión de la secuencia identificada por ese número de entrada, disponible el 1 de septiembre de 2006.

50 Las SEC ID N°: 1-356 son cebadores utilizados para amplificar 178 regiones de unión específicas para la región del genoma humano que contiene el gen HER2, donde las SEC ID N°: 1-178 son cebadores directos y las SEC ID N°: 179-356 son los cebadores inversos correspondientes, respectivamente (véase la Tabla 1).

La SEC ID N°: 357 es un ejemplo de oligonucleótido que se compone exclusivamente de purinas.

La SEC ID N°: 358 es el complemento inverso de la SEC ID N°: 357.

La SEC ID N°: 359 es un ejemplo de secuencia repetida en tándem que se encuentra en los telómeros.

La SEC ID N°: 360 es un ejemplo de secuencia de una unidad de repetición con cinco bases, que se encuentra en las secuencias de ácido nucleico repetitivas II y III de tipo satélite.

## 5 DESCRIPCIÓN DETALLADA

### INTRODUCCIÓN

La producción de sondas correspondientes a las secuencias de ácido nucleico diana seleccionadas (por ejemplo, secuencias de ácido nucleico genómico diana o secuencias de ácido nucleico diana del ARN) para el análisis molecular se puede complicar por la presencia de secuencias no deseadas que pueden incrementar potencialmente la cantidad de señal de fondo cuando está presente en una sonda. Algunos ejemplos de secuencias no deseadas incluyen, a título meramente enunciativo, las siguientes: elementos de ácido nucleico repetitivos intercalados presentes en todos los genomas eucarióticos (por ejemplo, humanos), dominios conservados codificados por secuencias de ARN, así como secuencias homólogas presentes en un genoma viral. Por ejemplo, si la secuencia de ácido nucleico diana es una secuencia HPV-1, las secuencias no deseadas pueden incluir secuencias homólogas encontradas en otras secuencias del genoma HPV (por ejemplo, secuencias específicas para la familia de virus HPV), pero que no son específicas para HPV-1. En otro ejemplo, la secuencia diana es una secuencia de ARN y las secuencias no deseadas pueden incluir dominios conservados presentes en esa secuencia de ARN, las secuencias encontradas en otras secuencias de ARN no diana y, por tanto, no específicas para la secuencia de ARN diana. La selección de sondas típicamente pretende equilibrar la fuerza de una señal específica diana con el nivel de fondo no específico. A la hora de seleccionar una sonda correspondiente a una diana, por lo general la señal se maximiza aumentando el tamaño de la sonda. Sin embargo, cuando el tamaño de una sonda (por ejemplo, en el caso de las secuencias de ácido nucleico genómico diana) aumenta, también lo hace la cantidad de secuencia de ácido nucleico (por ejemplo, repetitiva) no deseada incluida en la sonda. Cuando la sonda es etiquetada (sea directamente con una fracción detectable, como un fluoróforo, o indirectamente con una fracción como un hapteno, que se puede detectar indirectamente basándose en la unión y detección de componentes adicionales), los elementos de la secuencia de ácido nucleico (por ejemplo, repetitiva) no deseados son etiquetados junto con los elementos específicos diana dentro de la secuencia diana. Durante la hibridación *in situ*, la unión de las secuencias de ácido nucleico (por ejemplo, repetitivas) no deseadas etiquetadas resulta en una señal de fondo dispersa, que puede confundir la interpretación, por ejemplo cuando se desean datos numéricos o cuantitativos (como el número de copia de una secuencia).

La reducción de la señal de fondo debido a la hibridación de secuencias de ácido nucleico no deseadas o repetitivas etiquetadas en la sonda normalmente se ha conseguido añadiendo ADN bloqueante (por ejemplo, ADN repetitivo no etiquetado como ADN C<sub>6</sub>t-J™) a la reacción de la hibridación.

La presente divulgación proporciona un planteamiento para reducir o eliminar la señal de fondo provocada por la presencia de secuencias de ácido nucleico repetitivas o no deseadas en una sonda. Algunos ejemplos de las sondas divulgadas en el presente están sustancial o completamente libres de secuencias de ácido nucleico repetitivas o no deseadas, tales como las sondas libres de repetición o sustancialmente libres de repetición.

Las sondas de ácido nucleico se proporcionan en la presente divulgación. En determinados ejemplos, las sondas incluyen una pluralidad de moléculas de ácido nucleico, denominadas en el presente moléculas de ácido nucleico fragmentadas, permutadas y concatenadas (FPC). Las moléculas de ácido nucleico FPC incluyen porciones o segmentos de una secuencia de ácido nucleico diana seleccionada (por ejemplo, una secuencia de ácido nucleico diana de ARN o genómica) y, por tanto, se dice que están fragmentadas. Los segmentos de la secuencia de ácido nucleico diana se denominan en el presente regiones de unión, que pueden estar libres o sustancialmente libres de secuencias de ácido nucleico no deseadas (por ejemplo, repetitivas, de dominio conservadas u homólogas). Se dice que las moléculas de ácido nucleico FPC están permutadas porque el orden o la orientación de las regiones de unión puede ser diferente en una molécula de ácido nucleico FPC con respecto a la correspondiente secuencia de ácido nucleico diana (por ejemplo, secuencia de ácido nucleico diana de ARN o genómica). Se dice que las moléculas de ácido nucleico FPC están concatenadas porque las moléculas de ácido nucleico FPC incluyen una pluralidad de regiones de unión ligadas, formando así una molécula de ácido nucleico lineal de regiones de unión ligadas.

Sustancialmente toda la pluralidad de moléculas de ácido nucleico o las moléculas de ácido nucleico FPC de una sonda incluyen al menos una primera región de unión y una segunda región de unión, donde la primera región de unión y la segunda región de unión son contiguas y complementarias de las porciones no contiguas, y secuencias únicas de una molécula de ácido nucleico diana. Por ejemplo, la sonda puede incluir una pluralidad heterogénea de moléculas de ácido nucleico, incluyendo sustancialmente todas ellas al menos una primera región de unión con una primera secuencia del nucleótido y una segunda región de unión con una segunda secuencia del nucleótido, donde la primera región de unión y la segunda región de unión son contiguas y complementarias de las porciones no

contiguas y de las secuencias únicas del nucleótido de una molécula de ácido nucleico diana, y donde la primera secuencia del nucleótido y la segunda secuencia del nucleótido de cada pluralidad de moléculas de ácido nucleico pueden diferir de la primera secuencia de nucleótido y la segunda secuencia del nucleótido de otras pluralidades de moléculas de ácido nucleico. Por tanto, las moléculas de ácido nucleico FPC pueden incluir múltiples regiones de unión en un diferente orden u orientación, o ambas cosas, con respecto a la secuencia de ácido nucleico diana. Por ejemplo, la pluralidad de moléculas de ácido nucleico puede incluir al menos cinco o diez regiones de unión contiguas con la primera y la segunda regiones de unión, donde al menos las cinco o diez regiones de unión son complementarias con porciones no contiguas y secuencias únicas de la molécula de ácido nucleico diana. Por ejemplo, una molécula de ácido nucleico FPC puede incluir un pequeño número de regiones de unión diferentes, como dos, tres, cinco, diez o veinte. En otros casos, el número de regiones de unión diferentes correspondiente a una secuencia diana es relativamente grande, como al menos 50, al menos 100, al menos 150 o incluso más. En determinados ejemplos, las regiones de unión de las moléculas de ácido nucleico FPC están sustancial o completamente libres de las secuencias de ácido nucleico no deseadas (por ejemplo, repetitivas, conservadoras u homólogas) de la molécula de ácido nucleico diana (por ejemplo, ARN o molécula de ácido nucleico diana genómico).

El número de moléculas de ácido nucleico FPC en una sonda puede variar. En determinados ejemplos, la sonda incluye al menos 10 moléculas de ácido nucleico FPC diferentes, como al menos 25, al menos 50, al menos 100, al menos 250, al menos 500, al menos 1.000, al menos 2.000, al menos 5.000, al menos 10.000, o al menos 50.000 moléculas de ácido nucleico FPC diferentes.

La longitud de las regiones de unión puede variar, aunque por lo general va desde varios cientos a varios miles de pares de bases. En algunos ejemplos, las regiones de unión son más pequeñas, hasta de 10 a 50 nucleótidos. En otros ejemplos, las regiones de unión son más largas, como al menos de 100 nucleótidos, al menos de 200 nucleótidos, al menos de 1.000 nucleótidos, al menos de 2.000 nucleótidos, o al menos de 5.000 nucleótidos de longitud, como de 100 a 10.000 o de 100 a 6.000 nucleótidos de longitud.

La longitud de la pluralidad de moléculas de ácido nucleico de una sonda puede variar. Dentro de una sola molécula de ácido nucleico FPC, las diferentes regiones de unión pueden variar en un rango de diferentes tamaños, por ejemplo con una longitud media de 20 a 100 nucleótidos o con una longitud media de al menos 1.000 a 5.000 nucleótidos. En algunos ejemplos, sustancialmente todas las moléculas de la pluralidad de moléculas de ácido nucleico tienen al menos 1000 nucleótidos, como al menos 2.000 nucleótidos o al menos 5.000 nucleótidos. En otros casos, por ejemplo cuando la pluralidad de moléculas de ácido nucleico está etiquetada mediante traslación por muescas, la longitud de la pluralidad de moléculas de ácido nucleico es inferior, como de 100 a 500 nucleótidos o de 50 a 200 nucleótidos.

La pluralidad de moléculas de ácido nucleico puede estar etiquetada con un agente detectable, como fracciones ópticamente detectables, tales como una fracción fluorescente, un hapteno que se pueda detectar indirectamente a través de un socio de unión específica etiquetado (como un anticuerpo o avidina) o una enzima que sea capaz de convertir un sustrato en un producto ópticamente detectable.

En el presente se proporcionan los métodos para generar las sondas divulgadas, que están sustancial o completamente libres de la secuencia de ácido nucleico (por ejemplo, repetitiva, de dominio conservado u homóloga) no deseada. En determinados ejemplos, las sondas demuestran una señal intensificada con respecto a las sondas que incluyen cantidades importantes de las secuencias de ácido nucleico (por ejemplo, repetitivas, de dominio conservado u homólogas) no deseadas. En algunos ejemplos, las sondas se pueden producir de forma fiable y económica en cantidades suficientes para aplicaciones automatizadas de hibridación cromosómica *in situ*.

De este modo, un aspecto de la presente divulgación se refiere a métodos para producir sondas de ácido nucleico que incluyen una pluralidad de moléculas de ácido nucleico FPC. En un ejemplo, los métodos incluyen la unión al menos de una primera región de unión y una segunda región de unión, produciendo así una pluralidad de moléculas de ácido nucleico, en la que sustancialmente todas las moléculas de ácido nucleico incluyen al menos una primera región de unión y una segunda región de unión contiguas, donde al menos la primera región de unión y la segunda región de unión contiguas son complementarias de las porciones no contiguas y secuencias únicas de una molécula de ácido nucleico diana, y donde la pluralidad de moléculas de ácido nucleico forma la sonda. En algunos ejemplos, al menos la primera región de unión y la segunda región de unión están sustancialmente libres de las secuencias de ácido nucleico (por ejemplo, repetitivas, de dominio conservado u homólogas) no deseadas de la molécula de ácido nucleico diana. Los múltiples segmentos están unidos para formar una plantilla de ácido nucleico permutada lineal. En determinados ejemplos, las regiones de unión de una molécula de ácido nucleico FPC están unidas o ligadas entre sí enzimáticamente (por ejemplo, utilizando una ligasa). Por ejemplo, las regiones de unión se pueden unir en una unión de extremos romos o en un sitio de restricción. La unión química y la amplificación también se pueden utilizar para unir regiones de unión. En algunos ejemplos, las regiones de unión están separadas por enlaces.

El método también puede incluir la amplificación de la pluralidad de moléculas de ácido nucleico (es decir, moléculas de ácido nucleico FPC) para producir una pluralidad de amplímeros de la molécula de ácido nucleico (es decir, amplímeros de la molécula de ácido nucleico FPC) para formar la sonda. Básicamente, se puede utilizar cualquier

procedimiento de amplificación para producir amplímeros de la molécula de ácido nucleico FPC, como PCR, DOP-PCR, NASBA, RCA, amplificación T7/Primasa, SDA, LAMP, 3SR y MDA. En determinadas realizaciones, la amplificación se realiza utilizando un proceso de amplificación isotérmica, como MDA. Los amplímeros de la molécula de ácido nucleico FPC se pueden producir en múltiples reacciones de amplificación en serie (como en dos, tres o cuatro reacciones de amplificación en serie).

En algunos ejemplos, la pluralidad de moléculas de ácido nucleico FPC o amplímeros de las mismas son etiquetados, por ejemplo, para emplearlos para el análisis de hibridación *in situ* de núcleos en interfase o metafase. En determinados ejemplos, el etiquetado se realiza tras la amplificación de plantillas de moléculas de ácido nucleico FPC. La etiqueta puede ser cualquier fracción directa o indirectamente detectable que se pueda adjuntar a un ácido nucleico (o un componente del nucleótido), por ejemplo mediante el etiquetado químico o enzimático. La etiqueta puede ser, por ejemplo, un hapteno (como DNP o biotina), un ligando, una enzima, un radioisótopo o fracción fluorescente (como un fluoróforo o una nanopartícula fluorescente, tales como un nanocrystal semiconductor o «punto cuántico»).

La región o regiones de unión se pueden generar de diferentes maneras. Por ejemplo, las regiones de unión sustancial o completamente libres de secuencias de ácido nucleico no deseadas (como repetitivas, conservadoras u homólogas) y que corresponden a una secuencia de ácido nucleico diana (como ARN o una secuencia de ácido nucleico genómico) pueden ser producidas por uno o más de los siguientes métodos: el aislamiento de las regiones de unión de la molécula de ácido nucleico diana; la obtención de regiones de unión de la molécula de ácido nucleico diana mediante hibridación sustractiva; o la amplificación de las regiones de unión de la molécula de ácido nucleico diana. Por ejemplo, las regiones de unión pueden ser amplificadas de una secuencia de ácido nucleico diana (como una secuencia de ácido nucleico genómico diana), por ejemplo en múltiples reacciones de amplificación, generadas mediante la síntesis química de un ácido nucleico correspondiente a una secuencia de ácido nucleico diana, generadas por digestión de endonucleasa de restricción (u otras) de una secuencia de ácido nucleico diana, que en algunos casos puede estar seguida por la selección de regiones de unión mediante el corte físico (por ejemplo, mecánico) de una secuencia diana seguida de la eliminación de las secuencias de ácido nucleico repetitivas, o por cualquier combinación de estos métodos.

En efecto, los segmentos de la región de unión se pueden proporcionar por cualquier método que produzca un ácido nucleico que se corresponda en secuencia con una secuencia libre de ácido nucleico repetitivo (o sustancialmente libre de ácido nucleico repetitivo) de una secuencia de ácido nucleico diana (o similar o sustancialmente libre de otra secuencia no deseada presente en la secuencia de ácido nucleico diana). En realizaciones específicas, las regiones de unión se proporcionan aislando los fragmentos de restricción sustancialmente libres de ácido nucleico repetitivo (u otra secuencia no deseada) de una secuencia de ácido nucleico diana (como una secuencia de ácido nucleico genómico diana), obtenido una secuencia sustancialmente libre de ácido nucleico repetitivo (u otra secuencia no deseada) de una secuencia de ácido nucleico diana mediante hibridación sustractiva, mediante la amplificación de la secuencia sustancialmente libre de ácido nucleico repetitivo (u otra secuencia no deseada) de una secuencia de ácido nucleico diana (como una secuencia de ácido nucleico genómico diana), o mediante una combinación de estos métodos. En determinadas realizaciones, las secuencias de ácido nucleico repetitivas (u otra secuencia no deseada) se seleccionan de una secuencia de ácido nucleico diana (por ejemplo, ARN o una secuencia de ácido nucleico genómico diana) utilizando un programa o algoritmo generado por ordenador, como PRIME® o RepeatMasker.

Por ejemplo, las regiones de unión que están completa o sustancialmente libres de secuencias de ácido nucleico repetitivas (u otra secuencia no deseada), identificadas, por ejemplo, utilizando un algoritmo generado por ordenador) se pueden generar en un proceso de amplificación. Las diferentes regiones de unión se pueden amplificar en diferentes reacciones de amplificación de una molécula de ácido nucleico que contiene algunas o todas las regiones de unión de interés. Por ejemplo, las diferentes regiones de unión se pueden amplificar en una reacción de PCR utilizando al menos un par de cebadores de secuencia única. El cebador o cebadores empleados para la amplificación de las regiones de unión pueden incluir una secuencia de oligonucleótido fijada, como un tracto solo de purina. Opcionalmente, el cebador puede incluir un sitio de reconocimiento de la enzima de restricción. En algunos ejemplos, el cebador se selecciona utilizando un algoritmo generado por ordenador (como OLIGO™)

En determinados ejemplos, las regiones de unión se producen a partir de moléculas de ácido nucleico diana que no contienen naturalmente secuencias repetitivas características del ADN del genoma de los mamíferos. Estas moléculas de ácido nucleico diana incluyen ARN o moléculas de ácidos nucleicos genómicos virales. En algunas de estas realizaciones, las regiones de unión se pueden producir directamente de una molécula de ácido nucleico diana con la eliminación opcional de la molécula de ácido nucleico diana de las secuencias que razonablemente cabría esperar que provocasen una mayor señal de fondo (por ejemplo, secuencias homólogas o secuencias que codifican dominios conservados). Por ejemplo, si la molécula de ácido nucleico diana es un genoma viral o una secuencia de ARN, las regiones de unión se pueden generar directamente de la molécula diana (por ejemplo, mediante digestión con enzimas de restricción de un preparado de ácido nucleico celular aislado o bruto). Sin embargo, las posibles secuencias de ácido nucleico que generan señal de fondo y, por tanto, no deseadas, (por ejemplo, secuencias homólogas o dominios conservados de codificación de secuencias) se pueden eliminar de la secuencia diana de ARN o viral, antes o después de generar las regiones de unión de estas secuencias de ácido nucleico diana.

En otras realizaciones, las regiones de unión se proporcionan por amplificación, digestión o similares de uno o más vectores que incluyen la secuencia de ácido nucleico diana. Por ejemplo, se puede utilizar un único vector capaz de abarcar grandes porciones de ácido nucleico, como un cromosoma artificial (por ejemplo, BAC, YAC, PAC, *etc.*) o determinados vectores virales capaces de replicar gran cantidad de kilobases de ADN (como CMV).  
 5 Alternativamente, se pueden utilizar múltiples vectores (como plásmidos, cósmidos, fagos u otros virus) que incluyen porciones más pequeñas de la secuencia de ácido nucleico diana (por ejemplo, la secuencia de ácido nucleico genómico diana).

Una molécula de ácido nucleico diana puede ser cualquier ácido nucleico seleccionado, como ADN o ARN. En determinados ejemplos, la secuencia diana es una secuencia genómica diana o una subsecuencia genómica, por ejemplo de un genoma eucariótico, como el genoma humano. En algunos ejemplos, la molécula del ácido nucleico diana se selecciona de un patógeno, como un virus, una bacteria o un parásito intracelular, como de un genoma viral. Por ejemplo, la molécula de ácido nucleico diana puede ser una secuencia asociada con (por ejemplo, correlacionada con o implicada en, *etc.*) una enfermedad. En determinados ejemplos, la molécula de ácido nucleico diana seleccionada es una molécula de ácido nucleico diana asociada con una enfermedad neoplásica (o cáncer).  
 10 Por ejemplo, la secuencia diana genómica puede incluir al menos un gen asociado con el cáncer (por ejemplo, HER2, c-Myc, n-Myc, Abl, Bcl2, Bcl6, Rbl, p53, EGFR, TOP2A, MET, o genes que codifican otros receptores y/o moléculas de señalización, *etc.*) o región cromosómica asociada con un cáncer. Por ejemplo, la secuencia de ácido nucleico diana puede estar asociada con una anomalía estructural cromosómica, por ejemplo una traslocación, deleción o reduplicación (como una amplificación génica o polisomía) que ha sido correlacionada con un cáncer. En determinados ejemplos, la secuencia de ácido nucleico diana abarca una secuencia genómica con reduplicación o deleción en al menos algunas células neoplásicas. La secuencia del ácido nucleico diana puede variar notablemente de tamaño, como al menos 1.000 pares de bases de longitud, al menos 50.000, al menos 100.000 o incluso al menos 250.000, 500.000 o varios millones (por ejemplo, al menos tres millones) de pares de bases de longitud total.

En un método específico, las sondas se producen amplificando una pluralidad de regiones de unión que están sustancialmente libres de ácidos nucleicos no deseados (por ejemplo, repetitivos) de una secuencia de ácido nucleico diana para producir una pluralidad de amplímeros de la región de unión. Los componentes de la pluralidad de amplímeros de la región de unión se unen para producir una pluralidad de moléculas de ácido nucleico, incluyendo sustancialmente todas las regiones de unión contiguas que son complementarias de las porciones no contiguas y secuencias únicas de una molécula de ácido nucleico diana. Los componentes de la pluralidad resultante de las moléculas de ácido nucleico se amplifican para producir una pluralidad de amplímeros de la molécula de ácido nucleico, que pueden ser etiquetados para producir la sonda.  
 25 30

También se proporcionan métodos para utilizar las sondas divulgadas. Por ejemplo, las sondas se pueden utilizar para detectar una molécula de ácido nucleico diana. En un ejemplo, el método incluye la puesta en contacto de una o más de las sondas divulgadas con una muestra que incluye moléculas de ácido nucleico en condiciones suficientes para permitir la hibridación entre las moléculas de ácido nucleico de la muestra y la pluralidad de moléculas de ácido nucleico de la sonda. La hibridación resultante es detectada y, en algunos ejemplos, cuantificada, donde la presencia de hibridación indica la presencia de la molécula de ácido nucleico diana.  
 35

También se proporcionan juegos que incluyen los ácidos nucleicos divulgados en el presente. Por ejemplo, los juegos pueden incluir uno o más de los elementos siguientes: plantillas de la molécula de ácido nucleico FPC útiles para producir sondas de la presente divulgación y amplímeros de la molécula de ácido nucleico FPC (que pueden estar etiquetados o sin etiquetar). El juego también puede incluir uno o más reactivos auxiliares, como soluciones tampón, etiquetas, cebadores, enzimas y similares.  
 40

En el presente se proporciona un ejemplo de sonda útil para la detección de una región del genoma humano, incluyendo el gen HER2 (denominada en el presente sonda HER2 o sonda del HER2 humano) y métodos para realizar la sonda. A pesar de que la especificación describe una sonda HER2 con detalle, una persona con conocimientos en el campo apreciará que se pueden utilizar métodos similares para producir una sonda para cualquier secuencia de ácido nucleico diana (por ejemplo, ARN o una molécula de ácido nucleico genómico diana) de interés. Por otra parte, una persona con conocimientos en el campo apreciará que los métodos concretos proporcionados en el presente se pueden modificar, obteniéndose un resultado similar. Por ejemplo, se pueden amplificar entre 1 y 178 regiones de unión representativas (que están sustancialmente libres de la secuencia de ácido nucleico repetitiva) de la secuencia del genoma humano que incluye el gen HER2 (por ejemplo, como se describe en el Ejemplo J o de cualquier otra forma conocida para una persona con conocimientos en el campo), utilizando los pares de cebadores y plantillas del clon BAC disponibles en el mercado identificadas en la Tabla 1 (véase el Ejemplo 1). En algunos ejemplos, una sonda de HER2 humano incluye de 2 a 178 de estas regiones de unión de ejemplo (por ejemplo, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 100 o 150 de estas regiones de unión de ejemplo). En determinadas realizaciones de una sonda de HER2 humano, la secuencia de ácido nucleico de cada región de unión se compone de la secuencia de ácido nucleico amplificada del clon BAC aplicable, utilizando el par de cebadores aplicable identificado en la Tabla 1. En otras realizaciones, las regiones de unión representativas de la secuencia del genoma humano que contiene el gen HER2 están unidas (con o sin ligadores entre regiones de unión) en orden aleatorio o en una orientación aleatoria (o ambas cosas) para producir una pluralidad de ácidos nucleicos FPC de HER2. En otra realización, las regiones de unión de ejemplo de la secuencia del genoma humano  
 45 50 55 60

que contiene el gen HER2 están unidas (con o sin enlaces entre segmentos), en el orden en el que esos segmentos se encontrarían en la secuencia de ácido nucleico diana nativa (sin embargo, careciendo sustancialmente de las secuencias de ácido nucleico repetitivas que intervienen) en una orientación aleatoria u otra orientación no nativa.

## TÉRMINOS

5 A menos que se indique lo contrario, todos los términos técnicos y científicos empleados en el presente tienen el significado que entendería normalmente una persona con conocimientos en el campo al que se refiere esta divulgación. Las definiciones de términos comunes de biología molecular se pueden encontrar en Benjamin Lewin, *Genes V*, publicado por Oxford University Press, 1994 (ISBN 0-19-854287-9); Kendrew *et al.*, (eds.), *The Encyclopedia of Molecular Biology*, publicado por Blackwell Science Ltd., 1994 (ISBN 0-632-02182-9); y Robert A. Meyers (ed.), *Molecular Biology and Biotechnology: a Comprehensive Desk Reference*, publicado por VCH Publishers, Inc. 1995 (ISBN 1-56081-569-8).

10 Los términos en singular «un», «el» y «la» incluyen el plural, a menos que el contexto indique claramente lo contrario. Del mismo modo, la palabra «o» incluirá «y», a menos que el contexto indique claramente lo contrario. El término «pluralidad» se utiliza como sinónimo de la frase «más de uno», es decir dos o más. También se entenderá que todos los tamaños de las bases o los tamaños de los aminoácidos, así como todos los valores de los pesos o las masas moleculares indicados para los ácidos nucleicos o polipéptidos son aproximados y se ofrecen a título descriptivo. El término «comprende» significa «incluye». La abreviatura «p. ej.» viene del latín «*exempli gratia*» y se utiliza en el presente para dar un ejemplo, a título meramente enunciativo. Por tanto, la abreviatura «p. Ej.» es sinónimo del término «por ejemplo». A pesar de que se pueden emplear métodos y materiales similares o equivalentes a los descritos en el presente para la práctica o prueba de esta divulgación, los métodos y materiales adecuados se describen a continuación.

Al objeto de facilitar la revisión de las diversas realizaciones de esta divulgación, se proporcionan las siguientes explicaciones de términos específicos:

25 «**Amplificación de una molécula de ácido nucleico**» se refiere a los métodos utilizados para aumentar el número de copias de una molécula de ácido nucleico, como una región de unión de una molécula de ácido nucleico diana o una molécula de ácido nucleico FPC. Los productos resultantes se pueden denominar amplímeros o productos de la amplificación. Los métodos para amplificar las moléculas de ácido nucleico son conocidos en el campo e incluyen MDA, PCR, DOP-PCR, RCA, amplificación dependiente de T7/Primasa, SDA, 3SR, NASBA, y LAMP, entre otros.

30 «**Unión o unión estable**» se refiere a la asociación entre dos sustancias o moléculas, como la hibridación de una molécula de ácido nucleico (p. ej., una región de unión) con otra (o consigo misma) (p. ej., una molécula de ácido nucleico diana). Una molécula de ácido nucleico FPC se une o se une de forma estable a una molécula de ácido nucleico diana, si una cantidad suficiente de la molécula de ácido nucleico FPC forma pares de bases o se hibrida con su molécula de ácido nucleico diana para permitir la detección de esa unión.

35 La unión se puede detectar por cualquier procedimiento conocido por una persona con conocimientos en el campo, como por las propiedades físicas o funcionales de la diana: el complejo del ácido nucleico FPC. Los métodos físicos para detectar la unión de cadenas complementarias de la molécula de ácido nucleico incluyen, a título meramente enunciativo, métodos como DNasa I o huella química, ensayos de retardo en gel y clivaje de afinidad, la técnica de *Northern blot* o *dot blot* y procedimientos de detección de absorción de luz. En otro ejemplo, el método implica la detección de una señal, como una etiqueta detectable, presente en una o en ambas moléculas de ácido nucleico (p. ej., una etiqueta asociada con la molécula de ácido nucleico FPC).

40 Una «**región de unión**» es un segmento o una porción de una molécula de ácido nucleico diana que es único en la molécula diana y en algunos ejemplos está libre o sustancialmente libre de secuencia de ácido nucleico repetitiva (u otra no deseada). La secuencia de ácido nucleico de una región de unión y su correspondiente molécula de ácido nucleico diana tienen una complementariedad de la secuencia de ácido nucleico suficiente como para que, cuando las dos se incuban en condiciones de hibridación apropiadas, ambas moléculas se hibridicen para formar un complejo detectable. Una molécula de ácido nucleico diana puede contener múltiples regiones de unión diferentes, como al menos 10, al menos 50, al menos 100 o al menos 1.000 regiones de unión únicas. En determinados ejemplos, una región de unión tiene típicamente de varios cientos a varios miles de pares de bases de longitud. No obstante, en algunos ejemplos la región de unión es menor, como de 50 a 200 pares de bases de longitud. Cuando se obtienen regiones de unión de una secuencia de ácido nucleico diana, la secuencia diana se puede obtener en su forma nativa en una célula, como una célula de mamífero, o en forma clonada (p. ej., en un vector).

55 Se dice que una molécula de ácido nucleico es «**complementaria**» con otra molécula de ácido nucleico, cuando las dos moléculas comparten un número suficiente de nucleótidos complementarios como para formar un dúplex o tríplex cuando las cadenas se unen (hibridan) entre sí, por ejemplo formando pares de bases de Watson-Crick, Hoogsteen o Hoogsteen inverso. La unión estable se produce cuando una molécula de ácido nucleico (p. ej., molécula de ácido nucleico FPC) continúa unida de forma detectable a una secuencia de ácido nucleico diana (p.ej., secuencia de ácido nucleico genómico diana), en las condiciones necesarias.

La complementariedad es el grado en el que las bases de una molécula de ácido nucleico (p.ej., molécula de ácido nucleico FPC) se emparejan con las bases de una segunda molécula de ácido nucleico (p.ej., secuencia de ácido nucleico genómico diana). La complementariedad se describe convenientemente por el porcentaje, es decir por la proporción de nucleótidos que forman pares de bases entre dos moléculas o dentro de un dominio o una región específica de dos moléculas. Por ejemplo, si 10 nucleótidos de una región de 15 nucleótidos contiguos de una molécula de ácido nucleico FPC forman pares de bases con una molécula de ácido nucleico diana, se dice que esa región de la molécula de ácido nucleico FPC tiene una complementariedad del 66,67% con la molécula de ácido nucleico diana.

En la presente divulgación, «complementariedad suficiente» significa que existe un número suficiente de pares de bases entre una molécula de ácido nucleico o una región de la misma (como la región de una molécula de ácido nucleico FPC) y una secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) como para conseguir una unión detectable. Una descripción detallada de las consideraciones cualitativas y cuantitativas implicadas para determinar las condiciones de unión se proporciona en Beltz *et.al. Methods Enzymol.* 100:266-285, 1983, y en Sambrook *et.al. (ed.), Molecular Cloning: A Laboratory Manual, 2ª ed.*, vol. 1 -3, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989

Un «**algoritmo generado por ordenador**» es un algoritmo o programa (un conjunto de códigos ejecutables en un medio legible por ordenador) que se realiza o se ejecuta a través de un dispositivo informático bajo un comando de un usuario. En el contexto de la presente divulgación, los algoritmos generados por ordenador se pueden utilizar para facilitar (p. ej., automatizar) la selección de secuencias de polinucleótidos con características concretas, tales como la identificación de secuencias de ácido nucleico repetitivas (o con otras características no deseadas, como señal de fondo) o regiones de unión únicas de una secuencia de ácido nucleico diana. Típicamente, un usuario inicia la ejecución del algoritmo introduciendo un comando y estableciendo uno o más criterios de selección en un ordenador que es capaz de acceder a una base de datos de secuencias. La base de datos de secuencias puede encontrarse dentro del medio de almacenamiento del ordenador o puede estar almacenada remotamente y resultar accesible a través de una conexión del ordenador con un medio de almacenamiento de una ubicación cercana o remota a través de una intranet o de Internet. Tras haber iniciado el algoritmo, el algoritmo o programa se ejecuta en el ordenador, por ejemplo para seleccionar una o más secuencias de polinucleótidos que satisfacen los criterios de selección. Habitualmente, a continuación se despliegan las secuencias de polinucleótidos seleccionadas (p. ej., en una pantalla) o se facilitan (p. ej., en formato impreso o soporte magnético).

El término «**correspondiente**» en referencia a un primer y un segundo ácido nucleico (por ejemplo, una región de unión y una secuencia de ácido nucleico diana) indica que el primer y el segundo ácido nucleico comparten una complementariedad o una identidad de la secuencia sustancial en al menos una porción de la secuencia total del primer y/o el segundo ácido nucleico. Así, una región de unión se corresponde con una secuencia de ácido nucleico diana, si la región de unión posee una complementariedad (p. ej., complementariedad inversa) o una identidad sustancial de la secuencia con al menos una porción de la secuencia de ácido nucleico diana (p. ej., si es al menos idéntica o complementaria en un 80%, un 85%, un 90%, un 95% o incluso un 100%). Por ejemplo, una región de unión se puede corresponder con una secuencia de ácido nucleico diana, si la región de unión posee una identidad sustancial de la secuencia con una cadena de una secuencia de ácido nucleico diana de doble cadena (p. ej., una secuencia de ADN genómico diana) o si la región de unión es sustancialmente complementaria con una secuencia de ácido nucleico diana de cadena única (p. ej., ARN o un genoma viral de ARN).

Una «**molécula de ácido nucleico fragmentada, permutada, concatenada (FPC)**» se refiere a uno o más ácidos nucleicos en los que los segmentos o regiones de unión constitutivas que se corresponden con una molécula de ácido nucleico diana están presentes en un orden o una orientación diferente al orden o la orientación de las regiones de unión de la secuencia de ácido nucleico diana (como una secuencia de ácido nucleico genómico diana). Es decir, que las regiones de unión contiguas de la molécula de ácido nucleico FPC pueden ser complementarias de las porciones no contiguas y de las secuencias únicas de una molécula de ácido nucleico diana. Las regiones de unión presentes en una o más de las moléculas de ácido nucleico FPC pueden incluir toda la secuencia de polinucleótidos presente en la molécula de ácido nucleico diana o un subconjunto de la secuencia de polinucleótidos presente en la molécula de ácido nucleico diana. Por ejemplo, en algunos ejemplos las moléculas de ácido nucleico FPC están completa o sustancialmente libres de secuencia de ácido nucleico repetitiva u otra no deseada. Las moléculas de ácido nucleico FPC se pueden utilizar como plantillas en una reacción de amplificación, produciendo así amplímeros de la molécula de ácido nucleico FPC.

Dado que las moléculas de ácido nucleico FPC incluyen segmentos (regiones de unión) de la secuencia de ácido nucleico diana correspondiente (p.ej., secuencia de ácido nucleico genómico diana), se dice que están fragmentadas. El orden o la orientación (o ambas cosas) de las regiones de unión es diferente en una molécula de ácido nucleico FPC con respecto a la secuencia de ácido nucleico diana correspondiente (p.ej., la secuencia de ácido nucleico genómico diana) y, por tanto, se dice que las moléculas están permutadas. Las moléculas de ácido nucleico FPC incluyen una pluralidad de regiones de unión ligadas o unidas, formando así una molécula de ácido nucleico lineal, por lo que se dice que están concatenadas.

Cuando dos (o más) moléculas de ácido nucleico FPC correspondientes a la misma molécula de ácido nucleico diana se comparan, la primera y la segunda (y cualquier otra) plantilla de ácido nucleico FPC puede incluir subconjuntos de regiones de unión de la secuencia diana que se solapan en gran medida, aunque en un orden o una orientación diferente entre la primera y la segunda molécula de ácido nucleico FPC, o la primera y la segunda molécula de ácido nucleico FPC puede incluir diferentes subconjuntos de subsecuencias (segmentos) de la secuencia de ácido nucleico diana.

Un «genoma» es la totalidad de los componentes genéticos de un organismo. En el caso de los organismos eucarióticos, el genoma se contiene en un conjunto haploide de cromosomas de una célula. En el caso de los organismos procarióticos, el genoma se contiene en un único cromosoma y, en algunos casos, uno o más elementos genéticos extracromosómicos, como los episomas (p. ej., plásmidos). Un genoma viral puede adoptar la forma de una o más moléculas de ARN o ADN de cadena única o doble, dependiendo del virus concreto.

El término «aislado» en referencia a un componente biológico (como una molécula de ácido nucleico, una proteína o célula) se refiere a un componente biológico que ha sido sustancialmente separado o purificado de otros componentes biológicos de la célula del organismo, o del organismo en sí, donde el componente está naturalmente presente, como otro ARN o ADN cromosómico y extracromosómico, proteínas, células y organelos. Las moléculas de ácido nucleico que han sido «aisladas» incluyen moléculas de ácido nucleico purificadas con métodos de purificación estándar. El término también abarca ácidos nucleicos preparados por amplificación o clonado, así como ácidos nucleicos sintetizados químicamente.

Una «etiqueta» es una composición o compuesto detectable que se conjuga directa o indirectamente con otra molécula (como una molécula de ácido nucleico FPC) para facilitar la detección de dicha molécula. Algunos ejemplos, a título meramente enunciativo, de etiquetas incluyen las fracciones fluorescentes y fluorogénicas, fracciones cromogénicas, haptenos, etiquetas de afinidad e isótopos radiactivos. La etiqueta puede ser directamente detectable (p. ej., ópticamente) o indirectamente detectable (p. ej., mediante la interacción con una o más moléculas adicionales que son, a su vez, detectables). A continuación se describen ejemplos de etiquetas en el contexto de las sondas divulgadas en el presente. Los métodos para el etiquetado de los ácidos nucleicos y las orientaciones relativas a la selección de etiquetas útiles para diversos propósitos se debaten, por ejemplo, en Sambrook y Russel, en *Molecular Cloning A Laboratory Manual*, 3ª Ed., Cold Spring Harbor Laboratory Press (2001) y *Ausubel et al.*, en *Current Protocols in Molecular Biology*, Greene Publishing Associates and Wiley-Intersciences (1987, e incluyendo actualizaciones),

Un «ácido nucleico» es un polímero de ribonucleótido o desoxirribonucleótido en forma de cadena única o doble cadena y, a menos que se indique lo contrario, abarca análogos de nucleótidos naturales que se hibridan con los ácidos nucleicos de forma similar a los nucleótidos presentes naturalmente. El término «nucleótido» incluye, a título meramente enunciativo, un monómero que incluye una base (como una pirimidina, purina o sintéticos análogos de las mismas) unida a un azúcar (como ribosa, desoxirribosa o análogos sintéticos de las mismas), o una base unida a un aminoácido, como en el ácido nucleico peptídico (PNA). Un nucleótido es un monómero de un polinucleótido. Una secuencia de nucleótidos se refiere a la secuencia de bases de un polinucleótido.

Un «segmento» de ácido nucleico es una subporción o subsecuencia de una molécula de ácido nucleico diana. Un segmento de ácido nucleico se puede obtener hipotéticamente o realmente de una molécula de ácido nucleico diana de diversas maneras. Por ejemplo, un segmento de una molécula de ácido nucleico diana (como una molécula de ácido nucleico genómico diana) se puede obtener mediante la digestión con una o más enzimas de restricción para producir un segmento de ácido nucleico que sea un fragmento de restricción. Los segmentos de ácido nucleico también se pueden producir de una molécula de ácido nucleico diana mediante amplificación, mediante hibridación (por ejemplo, hibridación sustractiva), mediante síntesis artificial o mediante cualquier otro procedimiento que produzca uno o más ácidos nucleicos que se correspondan en secuencia con una molécula de ácido nucleico diana. Un ejemplo concreto de segmento de ácido nucleico es una región de unión.

Una «sonda» o «sonda de ácido nucleico» es una molécula de ácido nucleico capaz de hibridarse con una molécula de ácido nucleico diana (p. ej., una molécula de ácido nucleico genómico diana) y, cuando se hibrida con el objetivo, puede ser detectada directa o indirectamente. Por tanto, la sonda permite la detección y, en algunos casos, la cuantificación de una molécula de ácido nucleico diana. En determinados ejemplos, una sonda incluye una pluralidad de moléculas de ácido nucleico, como una mezcla heterogénea de moléculas de ácido nucleico FPC, que incluyen regiones de unión obtenidas de la molécula de ácido nucleico diana y que, por tanto, son capaces de hibridarse específicamente con al menos una porción de la molécula de ácido nucleico diana. Por lo general, una vez que una porción de una molécula de ácido nucleico FPC se ha hibridado (y se mantiene así) con la molécula de ácido nucleico diana, otras porciones de la molécula de ácido nucleico FPC pueden (aunque no necesariamente) quedar impedidas físicamente para hibridarse con los sitios de unión de los cognados de esas otras posiciones (p. ej., aquellas otras porciones están demasiado lejos de los sitios de unión de sus cognados); sin embargo, otras moléculas de ácido nucleico FPC presentes en la sonda pueden unirse entre sí, amplificando así la señal de la sonda. Una sonda puede denominarse «sonda de ácido nucleico etiquetada», para indicar que la sonda está acoplada directa o indirectamente a una fracción detectable o «etiqueta», que hace que la sonda resulte detectable.

La frase «**sustancialmente libre de secuencia de ácido nucleico repetitiva**» en relación con un ácido nucleico (como una región de unión o una molécula de ácido nucleico FPC) indica que el ácido nucleico se compone exclusiva o predominantemente de regiones de unión complementarias de secuencias únicas de una molécula de ácido nucleico diana y no incluye una cantidad apreciable de secuencias de ácido nucleico repetitivas (p. ej., ADN) o «repeticiones». Las secuencias de ácido nucleico repetitivas son secuencias de ácido nucleico dentro de una secuencia de ácido nucleico (como un genoma, por ejemplo un genoma de mamífero o viral) que abarcan una serie de nucleótidos que se repiten muchas veces, normalmente en tándem. Las secuencias de ácido nucleico repetitivas pueden producirse en una secuencia de ácido nucleico (p. ej., un genoma de mamífero) en múltiples copias, que van desde dos a cientos de miles de copias, y que se pueden agrupar o intercalar en uno o más cromosomas de todo el genoma. En algunos ejemplos, la presencia de abundantes moléculas de ácido nucleico repetitivo en una sonda puede aumentar la señal de fondo. Las secuencias de ácido nucleico repetitivo incluyen, por ejemplo en humanos, repeticiones teloméricas, repeticiones subteloméricas, repeticiones de microsatélite, repeticiones de minisatélite, repeticiones de Alu, repeticiones de L1, repeticiones de ADN satélite Alpha, satélite 1, H y III, y ADN Cot-I™. De este modo, las regiones de unión o las moléculas de ácido nucleico FPC que están sustancialmente libres de secuencias de ácido nucleico repetitivas pueden incluir menos de un 10% de secuencias de ácido nucleico repetitivas, como menos de un 5%, menos de un 4%, menos de un 3%, menos de un 2% o incluso menos de un 1% de secuencias de ácido nucleico repetitivas. En determinados ejemplos, no hay ninguna secuencia de ácido nucleico repetitiva detectable presente en una región de unión o molécula de ácido nucleico FPC que está sustancialmente libre de secuencias de ácido nucleico repetitivas.

El término «**reduplicado**» se refiere a una secuencia de polinucleótidos genómicos que se encuentra típicamente en una única copia del genoma haploide de una célula. En determinadas condiciones, como un desarrollo o transformación de tipo neoplásico, la secuencia se replica en múltiples ocasiones, de forma que se encuentran múltiples copias (en ocasiones, numerosas) en la célula neoplásica. Con frecuencia, este fenómeno se denomina «amplificación» de la secuencia de polinucleótidos. Sin embargo, en el contexto de la presente divulgación, el término reduplicado se puede utilizar en lugar del término amplificación para distinguir el fenómeno genético de la replicación múltiple de una célula (es decir, reduplicación) de la amplificación artificial de una secuencia diana (p. ej., por PCR u otros métodos *in vitro*).

Una «**muestra**» es una muestra biológica que contiene ADN genómico, ARN (incluyendo ARNm), proteína o combinaciones de los mismos, obtenidos de un sujeto. Algunos ejemplos incluyen, a título meramente enunciativo, preparados cromosómicos, sangre periférica, orina, saliva, biopsia de tejido, muestra quirúrgica, médula espinal, muestras de amniocentesis y material de autopsia. En un ejemplo, una muestra incluye ADN genómico o ARN. En algunos ejemplos, la muestra es una preparación citogenética, por ejemplo, que se puede colocar en el portaobjetos de un microscopio. En determinados ejemplos, las muestras se utilizan directamente o pueden ser manipuladas antes del uso, por ejemplo, mediante fijación (p. ej., utilizando formalina).

Un «**sujeto**» incluye cualquier organismo vertebrado multicelular, como mamíferos humanos y no humanos (p. ej., sujetos veterinarios).

Una «**molécula o secuencia de ácido nucleico diana**» es una región definida o una secuencia concreta de una molécula de ácido nucleico, por ejemplo un genoma (como un gen o una región de ADN del genoma de mamífero que contiene un gen de interés) o una secuencia de ARN. En un ejemplo en el que la secuencia de ácido nucleico diana es una secuencia genómica diana, esta diana se puede definir por su posición en un cromosoma (p. ej., en una célula normal), por ejemplo, de acuerdo con la nomenclatura citogenética por referencia a una ubicación determinada en un cromosoma; por referencia a su ubicación en un mapa genético; por referencia a una secuencia contigua unida o hipotética; por su función o secuencia específica; por el nombre de su proteína o gen, o por cualquier otro medio que la identifique de forma única frente a otras secuencias genéticas de un genoma. En algunos ejemplos, la secuencia de ácido nucleico diana es una secuencia genómica viral o de mamífero. En otros ejemplos, la secuencia de ácido nucleico diana es una secuencia de ARN.

En algunos ejemplos, las alteraciones de una secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico) están «asociadas con» una enfermedad o condición. Por tanto, la detección de la secuencia de ácido nucleico diana se puede utilizar para interferir en el estado de una muestra con respecto a la enfermedad o condición. Por ejemplo, la secuencia de ácido nucleico diana puede existir en dos (o más) formas diferenciables, de forma que una primera forma está correlacionada con la ausencia de una enfermedad o condición y una segunda (o diferente) forma está correlacionada con la presencia de la enfermedad o la condición. Las dos formas diferentes pueden distinguirse cualitativamente, como por los polimorfismos de los polinucleótidos, y/o cuantitativamente, como por el número de copias de la secuencia de ácido nucleico diana presente en una célula.

Un «**cebador de secuencia única**» es un cebador, como un cebador oligonucleótido, que incluye una secuencia de polinucleótidos única, como una secuencia de ácido nucleico diana única. Un cebador de secuencia única puede incluir opcionalmente nucleótidos adicionales (típicamente en su extremo 5') que facilitan su posterior manipulación (p. ej., restricción, unión, clonado, etc.), como sitios de restricción o polímeros de nucleótidos cortos que contienen nucleótidos solamente de purina (o solamente de pirimidina). El término «cebador de secuencia única» se utiliza frecuentemente para distinguir un cebador que corresponde a una secuencia de ácido nucleico diana única de un

cebador que se compone de una secuencia que es compartida por una pluralidad de secuencias de ácido nucleico diana, como un «cebador universal» que corresponde a una secuencia de polinucleótidos que es común a una familia de moléculas de ácido nucleico diana (como ácidos nucleicos que incluyen una secuencia del adaptador o del enlace o ácidos nucleicos clonados en un vector común) o un «cebador aleatorio».

5 Un «vector» es un ácido nucleico que actúa como portador para otras secuencias de ácido nucleico («extrañas») que no son nativas del vector. Cuando se introduce en una célula huésped apropiada, un vector puede replicarse (y, de este modo, la secuencia de ácido nucleico extraña) o expresar al menos una porción de una secuencia de ácido nucleico extraña. En un contexto, un vector es un ácido nucleico circular o lineal en el que se introduce una  
10 secuencia de ácido nucleico diana de interés (por ejemplo, clonada) para los fines de la replicación (p. ej., producción) y/o manipulación utilizando técnicas de ácido nucleico recombinante estándar (p. ej., digestión de restricción. Un vector puede incluir secuencias de ácido nucleico que le permiten replicarse en una célula huésped, como un origen de replicación. Un vector también puede incluir uno o más genes marcadores seleccionables y otros elementos genéticos conocidos en el campo. Entre los vectores comunes se incluyen, por ejemplo, plásmidos, cósmidos, fagos, fagémidos, cromosomas artificiales (p. ej., BAC, PAC, HAC, YAC) e híbridos que incorporan  
15 características de más de uno de estos tipos de vectores. Típicamente, un vector incluye uno o más sitios de restricción únicos (y en algunos casos un sitio de multiclonado) para facilitar la inserción de una secuencia de ácido nucleico diana.

En un ejemplo debatido en el presente, una o más regiones de unión sustancialmente libres de secuencias de ácido nucleico repetitivas se introducen y replican en un vector, como un cromosoma artificial (p. ej., cromosoma artificial de levadura, cromosoma artificial derivado de P1, cromosoma artificial bacteriano (BAC)).  
20

### **SONDAS DE ÁCIDO NUCLEICO**

La presente divulgación proporciona sondas de ácido nucleico. Estas sondas se pueden utilizar para detectar una secuencia de ácido nucleico diana, como una secuencia de ácido nucleico genómico diana asociada con una enfermedad o con un patógeno, o una secuencia de ácido nucleico de ARN diana. Por ejemplo, las sondas  
25 se pueden utilizar en procedimientos de hibridación *in situ* que incluyen la hibridación de los ácidos nucleicos de cadena única etiquetados de una sonda con preparados cromosómicos, como núcleos en interfase o metafase, o secciones de tejido.

Las sondas de ácido nucleico divulgadas incluyen una pluralidad heterogénea de moléculas individuales de ácido nucleico fragmentadas, permutadas, concatenadas (moléculas de ácido nucleico FPC). Se dice que las moléculas  
30 de ácido nucleico FPC están fragmentadas porque incluyen porciones o segmentos de la secuencia de ácido nucleico diana seleccionada (p. ej., una secuencia de ácido nucleico genómico diana). Los segmentos de la secuencia de ácido nucleico diana se denominan en el presente regiones de unión. En algunas realizaciones (p. ej., las que implican secuencias de ácido nucleico diana del genoma humano), los segmentos de la región de unión de la secuencia de ácido nucleico diana están sustancial o completamente libres de secuencias de ácido nucleico no  
35 deseadas, como aquellas que incrementan la señal de fondo. Algunos ejemplos de secuencias de ácido nucleico no deseadas incluyen secuencias de ácido nucleico repetitivas (p. ej., las que se encuentran en secuencias del genoma de mamíferos), las secuencias de codifican dominios conservados (p. ej., las que se encuentran en secuencias de ARN) y secuencias homólogas (p. ej., las que se encuentran en secuencias de genomas virales). Las regiones de unión tienen una complementariedad suficiente con porciones de la secuencia de ácido nucleico diana seleccionada  
40 como para hibridarse con (y, por tanto, detectar) la secuencia diana. Se dice que las moléculas de ácido nucleico FPC están permutadas porque el orden o la orientación de las regiones de unión puede ser diferente en las moléculas de ácido nucleico FPC que en la secuencia de ácido nucleico diana correspondiente (p. ej., una secuencia de ácido nucleico genómico diana). Se dice que las moléculas de ácido nucleico FPC están concatenadas porque  
45 las moléculas de ácido nucleico FPC incluyen una pluralidad de regiones de unión ligadas o unidas, formando así una molécula de ácido nucleico lineal de regiones de unión ligadas.

Las moléculas de ácido nucleico FPC de las composiciones de sonda divulgadas están sustancial o complemente  
libres de secuencias de ácido nucleico repetitivas (es decir, secuencias de repetición) o de otras secuencias  
indeseadas e incluyen regiones de unión que se corresponden, cada una de ellas, con elementos únicos (de  
50 codificación o no codificación) de la secuencia de ácido nucleico diana (p. ej., la secuencia de ácido nucleico genómico diana, como un genoma haploide). Los genomas de mamífero (incluyendo, por ejemplo, genomas humanos) incluyen numerosos elementos o secuencias de ácido nucleico repetitivas que dan cuenta de casi la mitad del ADN genómico total. Estas secuencias repetitivas pueden ser específicas de un cromosoma, específicas de un elemento estructural de un cromosoma o estar intercaladas en todos los cromosomas. El ARN puede incluir regiones  
que codifican dominios conservados que no son específicos de la secuencia de ARN diana. Los genomas virales  
55 pueden incluir secuencias de ácido nucleico encontradas en otros virus no diana (p. ej., secuencias homólogas) y, por tanto, no son específicos para la secuencia viral diana. La presencia de estas secuencias de ácido nucleico no deseadas en una sonda puede complicar el análisis, por ejemplo incrementando la señal de fondo. Las sondas divulgadas en el presente proporcionan una baja señal no específica (o de fondo) y una elevada señal específica (o diana). Adicionalmente, en algunas realizaciones, las sondas divulgadas se pueden producir fácilmente en grandes  
60 cantidades (de miligramos a gramos), por ejemplo utilizando los métodos divulgados en el presente.

Una descripción general de las sondas divulgadas y cómo se generan con los ejemplos de moléculas de ácido nucleico diana se incluye en las FIG. 1A y B.

5 Como se muestra en la FIG. 1 A, la secuencia de ácido nucleico diana (p. ej., ADN genómico humano) que contiene subsecuencias no deseadas (p. ej., secuencias de ácido nucleico repetitivas) 10 o la secuencia de ácido nucleico diana de contenido completo (p. ej., libre de secuencia de ácido nucleico repetitiva) (p. ej., ARN o secuencias de ácido nucleico genómico viral) 40 se pueden utilizar para generar una sonda 300 que incluye una pluralidad de moléculas de ácido nucleico fragmentadas, permutadas, concatenadas (moléculas de ácido nucleico FPC) 310, 312, 314, 316, 318, 320, 322, 324, 326, 328, 330, 332. La secuencia de ácido nucleico genómico diana 10, 40 incluye una pluralidad de regiones de unión 14, 16, 18 o 42, 44, 46, que pueden ser específicas para la secuencia de ácido nucleico diana. Las regiones de unión 14, 16, 18, 42, 44, 46 pueden incluir una secuencia de codificación o no codificación. Por simplicidad, solamente se muestran tres regiones de unión específicas para cada secuencia de ácido nucleico genómico diana 10, 40 (14, 16, 18 y 42, 44, 46, respectivamente). No obstante, una persona con conocimientos en el campo apreciará que muchas más regiones de unión pueden estar presentes (por ejemplo, en las regiones 26 y 48 de la secuencia de ácido nucleico diana), como al menos 10, al menos 50, al menos 100, al menos 200, al menos 1.000, o incluso al menos 10.000 regiones de unión (p. ej., 100 a 500 o 100 a 1.000 regiones de unión). En la FIG. 1 A, las flechas mostradas en cada región de unión 14, 16, 18, 42, 44, 46 representan una dirección de referencia para la secuencia de ácido nucleico (p. ej., → indica 5' a 3' y ← indica 3' a 5'). La longitud de cada región de unión 14, 16, 18, 42, 44, 46 puede variar. En determinados ejemplos, cada región de unión 14, 16, 18, 42, 44, 46 tiene de varios cientos a varios miles de pares de base de longitud, como al menos 200, al menos 500, al menos 1.000, al menos 5.000, al menos 10.000, al menos 50.000, o al menos 100.000 nucleótidos de longitud.

Una secuencia de ácido nucleico diana puede (como se muestra en 10) o no (como se muestra en 40) incluir secuencias de ácido nucleico no deseadas (p. ej., secuencias de ácido nucleico repetitivas) 20, 22, 24. Por ejemplo, la secuencia de ácido nucleico diana 10 puede ser una secuencia humana asociada con la enfermedad; la secuencia de ácido nucleico diana 40 puede ser una secuencia de genoma viral o ARN de mamífero (p. ej., humana) (las secuencias de genomas virales y ARN de mamífero, por lo general, no tienen elementos de ácido nucleico repetitivos). Por simplicidad, solamente se muestran tres secuencias de ácido nucleico no deseadas (p. ej., secuencias de ácido nucleico repetitivas) 20, 22, 24 para la secuencia de ácido nucleico diana 10. No obstante, una persona con conocimientos en el campo apreciará que puede haber muchas más secuencias de ácido nucleico no deseadas (p. ej., secuencias de ácido nucleico repetitivas) presentes (por ejemplo, en la región 26), como al menos 10, al menos 50, al menos 100 (p. ej., entre unos 100 y unos 200), al menos 1.000, o al menos 10.000 secuencias de ácido nucleico no deseadas (p. ej., secuencias de ácido nucleico repetitivas).

Como se muestra en la FIG. 1A en el paso 60, la secuencia de ácido nucleico diana 10 o 40 se utiliza para generar una población de regiones de unión 100 y 140. Cada población de regiones de unión 100 y 140 incluye una pluralidad de regiones de unión 114, 116, 118 o 142, 144, 146 obtenidas de la correspondiente secuencia de ácido nucleico diana 10 o 40, respectivamente. El número de regiones de unión correspondiente a la secuencia de ácido nucleico diana 10, 40 puede variar en gran medida. Una persona con conocimientos en el campo apreciará que puede haber regiones de unión adicionales (p. ej., en las regiones diana 126 o 148) presentes en una población de regiones de unión 100 y 140. Por ejemplo, aunque solamente se muestran tres regiones de unión para la población 100 y 140 (114, 116, 118 y 142, 144, 146, respectivamente) para una mayor simplicidad, una persona con conocimientos en el campo apreciará que puede haber presentes muchas regiones de unión (por ejemplo, tal y como se representa en 126 y 148, respectivamente), como al menos 10, al menos 50, al menos 100, al menos 200, al menos 1.000, al menos 5.000, o incluso al menos 10.000 regiones de unión diferentes que corresponden a una única secuencia de ácido nucleico diana 10, 40.

La longitud de cada región de unión 114, 116, 118, 142, 144, 146 puede variar. En determinados ejemplos, las regiones de unión 114, 116, 118, 142, 144, 146 tienen de varios cientos a varios miles de pares de bases de longitud, como al menos 200, al menos 500, al menos 1.000, al menos 5.000, al menos 10.000, al menos 50.000, o al menos 100.000 nucleótidos de longitud (p. ej., de unos 100 a unos 6.000 nucleótidos de longitud). El término nucleótido no solamente se refiere a los nucleótidos de una molécula de ácido nucleico de cadena única, por ejemplo que ha sido desnaturalizada para permitir la hibridación con una diana, sino que también indica la longitud en pares de bases de las moléculas de ácido nucleico de cadena doble. Así, una región de unión individual puede tener al menos 20 nucleótidos, al menos 100 nucleótidos, al menos 200 nucleótidos, al menos 300 nucleótidos, al menos 400 nucleótidos, al menos 500 nucleótidos, al menos 1.000 nucleótidos (1 kb), o al menos 2.500 nucleótidos de longitud o más. Por supuesto, están permitidas las regiones de unión más largas. Sin embargo, debido a la presencia de secuencias repetitivas intercaladas 20, 22, 24, en el ADN del genoma de mamífero, es raro que la longitud sea superior en una secuencia de ácido nucleico del genoma humano. No obstante, se pueden obtener regiones de unión más largas, por ejemplo, de genomas virales mayores, que generalmente no incluyen secuencias de ácido nucleico repetitivas. Las regiones de unión individuales dentro de la misma molécula de la sonda pueden tener un tamaño relativamente uniforme o pueden variar de una región de unión a otra en todo el rango.

La población de regiones de unión 100 y 140 están sustancialmente libres de secuencias de ácido nucleico no deseadas (p. ej., secuencias de ácido nucleico repetitivas) 20, 22, 24. En algunos ejemplos, cada población de regiones de unión 100, 140 contiene menos del 10% de secuencias de ácido nucleico no deseadas (p. ej.,

secuencias de ácido nucleico repetitivas), como menos del 5%, menos del 1%, menos del 0,1%, o incluso menos del 0,01%. En algunos ejemplos, por ejemplo cuando la molécula de ácido nucleico diana incluye una secuencia de ácido nucleico no deseada (p. ej., una secuencia de ácido nucleico genómico diana), al menos el 80% (como, al menos el 85%, al menos el 90%, al menos el 95%, al menos el 98%, o al menos el 99%) de la secuencia de ácido nucleico no deseada se elimina de la secuencia de ácido nucleico diana.

La población de regiones de unión 100, 140 se puede someter opcionalmente a una o más rondas de amplificación de ácido nucleico 200. En algunos ejemplos, la población de regiones de unión 100, 140 se somete al menos a una ronda de amplificación, produciendo así una población de amplímeros de la región de unión. Opcionalmente, la población de regiones de unión 100, 140 se puede amplificar mediante amplificación específica para la secuencia (como PCR) hasta tres o más veces con diluciones intervinientes de las plantillas (p. ej., dilución 1:200). Además de aumentar el número de regiones de unión para cualquier paso posterior, este procedimiento también diluye cualquier molécula residual de la plantilla original por un amplio factor, reduciendo así los potenciales contaminantes, que se podrían incorporar de forma inadvertida a la sonda y conducir a la hibridación de fondo.

La pluralidad de regiones de unión 114, 116, 118 (y en algunos ejemplos las regiones de unión amplificadas) se somete a condiciones que permiten la unión 250 para formar ácido nucleico contiguo de las regiones de unión, resultando así en una población o pluralidad de moléculas de ácido nucleico fragmentadas, permutadas, concatenadas (moléculas de ácido nucleico FPC). En algunos ejemplos, se realizan múltiples reacciones de unión separadas, como reacciones de unión separadas que incluyen al menos cinco regiones de unión diferentes, al menos 10 regiones de unión diferentes, o al menos 20 regiones de unión diferentes (p. ej., de 5 a 50 o de 10 a 20 regiones de unión diferentes). En determinados ejemplos, se realizan al menos cinco o al menos 20 reacciones de unión diferentes. De forma similar, las regiones de unión 142, 144, 146 se pueden someter a la unión 250, pero eso no se muestra en la FIG. 1A. Esta mezcla de moléculas de ácido nucleico FPC individuales se denomina sonda 300. Por razones de simplicidad, solamente se muestran 12 ejemplos de productos de unión (denominados en el presente, moléculas de ácido nucleico FPC) 310, 312, 314, 316, 318, 320, 322, 324, 326, 328, 330 y 332. Las moléculas de ácido nucleico FPC concretas 310, 312, 314, 316, 318, 320, 322, 324, 326, 328, 330 y 332 mostradas en la FIG. 1A se ofrecen únicamente a título ilustrativo, dado que una persona con conocimientos en el campo reconocerá que hay muchas otras combinaciones de regiones de unión posibles 114, 116, 118. Las moléculas de ácido nucleico FPC individuales de la sonda 300 pueden estar compuestas de subconjuntos de regiones de unión solapadas en gran medida de la misma secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) o de regiones de unión predominantemente diferentes de la misma secuencia de ácido nucleico diana.

Las moléculas de ácido nucleico FPC resultantes 310, 312, 314, 316, 318, 320, 322, 324, 326, 328, 330 y 332 que forman la sonda 300 son ácidos nucleicos lineales que incluyen múltiples regiones de unión contiguas 114, 116, 118 (como 10 o más regiones de unión, 25 o más regiones de unión, 50 o más regiones de unión, 100 o más regiones de unión o 150 o más regiones de unión), donde cada región de unión 114, 116, 118 se corresponde al menos con una porción de la secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) 10. Las moléculas de ácido nucleico FPC 310, 312, 314, 316, 318, 320, 322, 324, 326, 328, 330 y 332 incluyen segmentos de la secuencia de ácido nucleico diana correspondiente (p. ej., una secuencia de ácido nucleico genómico diana) y, por tanto, se dice que están fragmentadas. El orden o la orientación de las regiones de unión 114, 116, 118 se puede cambiar en las moléculas de ácido nucleico FPC 310, 312, 314, 316, 318, 320, 322, 324, 326, 328, 330 y 332 con respecto a la secuencia de ácido nucleico diana correspondiente (p. ej., una secuencia de ácido nucleico genómico diana) 10 y, por tanto, se dice que están permutadas. Las moléculas de ácido nucleico FPC 310, 312, 314, 316, 318, 320, 322, 324, 326, 328, 330 y 332 incluyen una pluralidad de regiones de unión ligadas o unidas 114, 116, 118, formando así una molécula de ácido nucleico lineal, por lo que se dice que están concatenadas.

A título ilustrativo solamente, las moléculas de ácido nucleico FPC 310, 312, 314 incluyen tres regiones de unión, las moléculas de ácido nucleico FPC 316, 318, 320 incluyen cuatro regiones de unión, las moléculas de ácido nucleico FPC 322, 324, 326 incluyen dos regiones de unión, y las moléculas de ácido nucleico FPC 328, 330, 332 incluyen al menos cinco regiones de unión. Las moléculas de ácido nucleico FPC de la sonda 300 pueden incluir más de una de la misma región de unión (p. ej., la molécula de ácido nucleico FPC 316 incluye dos de las regiones de unión 114, y la molécula de ácido nucleico FPC 332 incluye tres regiones de unión 116 y dos regiones de unión 188). A pesar de que se muestran las moléculas de ácido nucleico FPC que tienen solamente de dos a cinco regiones de unión, se aprecia que, por lo general, las moléculas de ácido nucleico FPC tienen muchas más regiones de unión (aunque se muestran de dos a cinco, a título ilustrativo, de los conceptos generales proporcionados por esta divulgación). Por ejemplo, las moléculas de ácido nucleico FPC puede incluir al menos 50, al menos 80, al menos 100, o al menos 150 regiones de unión, como de 50 a 100, 50 a 80, 100 a 200, 100 a 150, o 150 a 200 regiones de unión.

La región o regiones de unión 114, 116, 118 que componen cada molécula de ácido nucleico FPC se pueden disponer en la misma orientación y orden en los que se encuentran en la secuencia de ácido nucleico genómico diana 30 (p. ej., como se muestra en la molécula de ácido nucleico FPC 310), donde la molécula de ácido nucleico diana incluye la secuencia de ácido nucleico repetitiva (y, por tanto, la secuencia de la molécula de ácido nucleico FPC y la molécula de ácido nucleico diana son diferentes). Alternativamente (y con una mayor frecuencia), la región o regiones de unión 114, 116, 118 que componen cada molécula de ácido nucleico FPC se encuentran en una

orientación o un orden diferente (o ambas cosas) con respecto a la orientación o el orden (o ambas cosas) de las regiones de unión que se encuentran en la secuencia de ácido nucleico diana seleccionada (p. ej., la secuencia de ácido nucleico genómico diana) 10 y entre sí. Así, los segmentos o elementos de la región de unión de las moléculas de ácido nucleico FPC están «permutadas» con respecto a la secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana).

Por ejemplo, la molécula de ácido nucleico FPC 314 muestra un ejemplo en el que la orientación (aunque no el orden) de la región de unión 118 está modificada con respecto a la secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) 10. En la molécula de ácido nucleico FPC 314, la orientación de la región de unión 118 está invertida (la flecha se encuentra en la dirección opuesta) con respecto a la secuencia de ácido nucleico genómico diana 10, pero el orden de las regiones de unión 114, 116, 118 se mantiene sin cambios con respecto a la secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) 10. La molécula de ácido nucleico FPC 312 muestra un ejemplo en el que el orden (aunque no la orientación) de las regiones de unión 114, 116, 118 está modificado con respecto al orden de la secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) 10. En la molécula de ácido nucleico FPC 312, el orden de las regiones de unión 114, 116, 118 es diferente con respecto a la secuencia de ácido nucleico genómico diana 10 (114, 118, 116 frente a 14, 16, 18), pero la orientación de las regiones de unión 114, 116, 118 se mantiene sin cambios con respecto a la secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) 10 (las flechas se encuentran en la misma dirección). Los restantes ejemplos de moléculas de ácido nucleico FPC 316, 318, 320, 322, 324, 326, 328, 330, 332 muestran ejemplos de dónde tanto la orientación como el orden de las regiones de unión 114, 116, 118 está modificado con respecto a las regiones de unión 14, 16, 18 de la secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) 10.

Las moléculas de ácido nucleico FPC 310, 312, 314, 316, 318, 320, 322, 324, 326, 328, 330 y 332 se pueden amplificar, etiquetar o ambas cosas, 350, para generar una sonda que contiene moléculas de ácido nucleico FPC etiquetadas y amplificadas. Como se muestra en la FIG. 1B., las regiones de unión se mezclan y unen para formar una población de moléculas de ácido nucleico FPC. Estas moléculas de ácido nucleico FPC pueden ser utilizadas como plantillas en una reacción de amplificación para generar mayores cantidades de moléculas de ácido nucleico FPC (denominadas amplímeros de la molécula de ácido nucleico FPC), etiquetadas con una etiqueta detectable, o ambas cosas. Los amplímeros de la molécula de ácido nucleico FPC se pueden agrupar para formar una sonda de la presente divulgación. En algunos ejemplos, las moléculas de ácido nucleico FPC son tanto amplificadas como etiquetadas. En un ejemplo concreto, la traslación por muescas se utiliza para etiquetar moléculas de ácido nucleico FPC, que también fragmenta las moléculas de ácido nucleico FPC. Esto resulta en la producción de una sonda compuesta de moléculas de ácido nucleico FPC etiquetadas más cortas, tales como moléculas de ácido nucleico FPC de menos de 1.000 nucleótidos de longitud, como de menos de 500 o menos de 100 nucleótidos de longitud (p. ej., de 50 a 1.000 o de 100 a 500 nucleótidos de longitud). En otro ejemplo, las moléculas de ácido nucleico FPC (o amplímeros de la molécula de ácido nucleico FPC) son etiquetadas químicamente. Típicamente este método no fragmenta las moléculas de ácido nucleico FPC.

Las moléculas de ácido nucleico FPC 310, 312, 314, 316, 318, 320, 322, 324, 326, 328, 330, 332 pueden tener diferentes longitudes. En un ejemplo, por ejemplo en el que las moléculas de ácido nucleico FPC 310, 312, 314, 316, 318, 320, 322, 324, 326, 328, 330, 332 de la sonda 300 no están fragmentadas, las moléculas de ácido nucleico FPC son por lo general largas, como al menos de 1.000 nucleótidos, al menos 2.000 nucleótidos, al menos 5.000 nucleótidos, al menos 10.000 nucleótidos, al menos 50.000 nucleótidos, o al menos 100.000 nucleótidos de longitud, como más de 150.000, o más de 200.000 nucleótidos (200 kb) de longitud. En un ejemplo, las moléculas de ácido nucleico FPC tienen de 2 a 50 kb de longitud, como de 2 a 10 kb de longitud. En algunos ejemplos, la longitud media de cada molécula de ácido nucleico FPC de la sonda 300 es al menos de 2 kb a 10kb.

En otro ejemplo, las moléculas de ácido nucleico FPC 310, 312, 314, 316, 318, 320, 322, 324, 326, 328, 330, 332 de la sonda 300 están fragmentadas. Por ejemplo, las moléculas de ácido nucleico FPC pueden ser fragmentadas (por ejemplo, en 350) utilizando la traslación por muescas para etiquetar las moléculas de ácido nucleico FPC, generando así una sonda de fragmentos de molécula de ácido nucleico FPC etiquetada. En este ejemplo, las moléculas de ácido nucleico FPC son más cortas, como al menos de 40 nucleótidos, al menos de 50 nucleótidos, al menos de 100 nucleótidos o al menos de 1.000 nucleótidos de longitud. En algunos ejemplos, la longitud media de cada fragmento de molécula de ácido nucleico FPC de una sonda es al menos de 50 nucleótidos, como de 50 a 1.000 nucleótidos. En algunas realizaciones, sustancialmente todos (al menos el 90%, al menos el 95% o al menos el 98%) los fragmentos de una molécula de ácido nucleico FPC incluirán al menos dos regiones de unión (como al menos tres, cinco o diez regiones de unión) con cada región de unión de un determinado fragmento de molécula de ácido nucleico FPC siendo complementaria de las secuencias únicas y no contiguas de la molécula de ácido nucleico diana y donde típicamente la mayoría de los fragmentos no incluyen todas las mismas (o incluso ninguna de las mismas) al menos dos regiones de unión.

Las FIG. 2A y 2B muestran cómo una sonda que incluye una pluralidad de moléculas de ácido nucleico FPC se puede utilizar para detectar una secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana). Como se muestra en la FIG. 2A, la secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) 500 (por ejemplo, en forma de un preparado cromosómico) se incubaba con la sonda

510 que contiene una pluralidad de moléculas de ácido nucleico FPC 520, 530, 550 (540 en la FIG. 2A incluye una pluralidad de moléculas de ácido nucleico FPC 600, 602, 604, 606, 608, 610 (véase la FIG. 2B para más información)) en condiciones que permiten la hibridación entre las moléculas de ácido nucleico FPC de la sonda 510 directamente (p. ej., 520, 530, 550) o indirectamente (p. ej., 540) y la secuencia de ácido nucleico diana (por ejemplo, una secuencia de ácido nucleico genómico diana) 500. Cada molécula de ácido nucleico FPC 520, 530, 550, 600, 602, 604, 606, 608, 610 incluye dos o más regiones de unión. Por ejemplo, la molécula de ácido nucleico FPC 520 incluye regiones de unión 522, 524, 526, 528. Las regiones de unión que comprenden una molécula de ácido nucleico FPC 520, 530, 550, 600, 602, 604, 606, 608, 610 tienen una complementariedad suficiente con una porción correspondiente de la secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) 500, de forma que cuando se incuban en las condiciones apropiadas, al menos una región de unión se hibrida con su secuencia complementaria inversa en la secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) 500. Por ejemplo, la región de unión 534 de la molécula de ácido nucleico FPC 530 tiene un grado de complementariedad suficiente con la secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) 500 como para permitir la hibridación 538 entre la molécula de ácido nucleico FPC 530 y la secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) 500. Se observa una hibridación similar entre las regiones de unión 542, 552 y la secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) 500 (véase, por ejemplo, 548, 558). La hibridación entre la región de unión 526 y la secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) 500 también ocurre, aunque no se muestra directamente en la FIG. 2A, para reducir la complejidad de la figura.

Dado que una molécula de ácido nucleico FPC 520, 530, 550, 600 puede incluir múltiples regiones de unión que son discontinuas (sea en orden, en orientación, o en ambas cosas) con respecto a la secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) 500, al menos una región de unión adyacente de la molécula de ácido nucleico FPC 520, 530, 550, 600 no se hibrida con la porción directamente contigua del cromosoma de la molécula de ácido nucleico diana (p. ej., una molécula de ácido nucleico genómico diana) 500, dejando una porción sin hibridar (p. ej., una burbuja, un bucle o una cola). Por ejemplo, como se muestra en la FIG. 2A, la molécula de ácido nucleico FPC 520 se hibrida con la molécula de ácido nucleico diana (p. ej., una molécula de ácido nucleico genómico diana) 500, a través de la región de unión 526, mientras que las otras regiones de unión adyacentes 522, 524, 528 no se hibridan con la molécula de ácido nucleico diana (p. ej., una molécula de ácido nucleico genómico diana) 500, porque son discontinuas (bien en orden, en orientación, o en ambas cosas) con respecto a la secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) 500. De un modo similar, la molécula de ácido nucleico FPC 530 se hibrida con la molécula de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) 500, a través de la región de unión 534, mientras que la región de unión adyacente 532 no se hibrida con la molécula de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) 500; la molécula de ácido nucleico FPC 550 se hibrida con la molécula de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) 500, mediante la región de unión 552, mientras que las regiones de unión restantes no se hibridan con la molécula de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) 500. Este resultado permite la amplificación de una señal para la detección de una secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana), por ejemplo en un preparado cromosómico.

La FIG. 2B muestra cómo una sonda que incluye una pluralidad de moléculas de ácido nucleico FPC heterogéneas puede generar una señal amplificada para la detección de una secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana). Como se ha descrito arriba, las moléculas de ácido nucleico FPC individuales se pueden hibridar directamente con una secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) 500 (FIG. 2A), debido a la complementariedad inversa de una región de unión con su región correspondiente de la secuencia diana. No obstante, como se muestra en la FIG. 2B, las regiones de unión de una molécula de ácido nucleico FPC (p. ej., 616 de ácido nucleico FPC 600) también pueden tener un grado de complementariedad suficiente con una región de unión de otra molécula de ácido nucleico FPC (p. ej. 620 de 602) como para unirse a esa otra molécula de ácido nucleico FPC, creando así una red de moléculas de ácido nucleico FPC originadas de un único sitio de unión de la secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) 500 y amplificando una señal generada por la sonda. Por ejemplo, al menos una región de unión (p. ej., 620) de una segunda molécula de ácido nucleico FPC (p. ej., 602) puede después hibridarse con una región de unión no hibridada (p. ej., 616) de la primera molécula de ácido nucleico FPC (p. ej., 600), dejando de nuevo al menos una segunda región de unión sin hibridar (p. ej., 626). De forma similar, una tercera, cuarta, etc., molécula de ácido nucleico FPC (p. ej., 604, 606, 608, 610) se puede hibridar con una región de unión no hibridada de moléculas de ácido nucleico FPC parcialmente hibridadas (p. ej., 600, 602). Esta «reacción en cadena» de hibridación secuencial produce una red de moléculas de ácido nucleico FPC, aumentando así de forma significativa la señal detectable de la sonda.

Por ejemplo, la incubación de una secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) 500 con la sonda 510 que incluye una pluralidad de moléculas de ácido nucleico FPC 520, 530, 550, 600, 602, 604, 606, 608, 610 puede resultar en la hibridación de moléculas de ácido nucleico FPC entre sí. Como se muestra en la FIG. 2B, la molécula de ácido nucleico FPC 600 incluye la región de unión 612 que tiene un grado de complementariedad suficiente con la secuencia de ácido nucleico genómico diana 500 como para permitir la hibridación 614 de la molécula de ácido nucleico FPC 600 con la secuencia diana (p. ej., una secuencia de ácido

nucleico genómico diana) 500. Otra amplificación de la señal se consigue cuando otras moléculas de ácido nucleico FPC se unen directa o indirectamente a la molécula de ácido nucleico FPC 600. Como se muestra en la FIG. 2B, la región de unión 616 de la molécula de ácido nucleico FPC 600 se puede hibridar con la región de unión 620 de la molécula de ácido nucleico FPC 602. La molécula de ácido nucleico FPC 602 se puede hibridar con la molécula de ácido nucleico FPC 604 mediante la interacción de las regiones de unión 622 y 624. La molécula de ácido nucleico FPC 600 se puede hibridar con la molécula de ácido nucleico FPC 606 a través de la interacción de las regiones de unión 626 y 628. La molécula de ácido nucleico FPC 606 se puede hibridar con la molécula de ácido nucleico FPC 608 a través de la interacción de las regiones de unión 630 y 632. La molécula de ácido nucleico FPC 608 se puede hibridar con la molécula de ácido nucleico FPC 610 a través de la interacción de las regiones de unión 634 y 636. Las interacciones de las moléculas de ácido nucleico FPC entre sí (p. ej., la hibridación de las moléculas de ácido nucleico FPC 602, 604, 606, 608, 610 directa o indirectamente con la molécula de ácido nucleico FPC 600) además de las interacciones de las moléculas de ácido nucleico FPC con la secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) 500 (p. ej., la hibridación de las moléculas de ácido nucleico FPC 520, 530, 550, 600 directamente con la secuencia de ácido nucleico diana 500) resultan en una señal incrementada de la sonda 510 en presencia de la secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) 500.

### ***Ejemplos de secuencias de ácido nucleico diana***

Las moléculas o secuencias de ácido nucleico diana incluyen tanto secuencias diana de ARN como de ADN. Se pueden generar moléculas de ácido nucleico FPC que se corresponden esencialmente con cualquier secuencia diana. En algunos ejemplos, se selecciona una secuencia diana que está asociada con una enfermedad o condición, para que la detección de la hibridación se pueda utilizar para obtener información (como información sobre el diagnóstico o el pronóstico para el sujeto del que se obtiene la muestra) asociada con la enfermedad o condición. En un ejemplo específico, la secuencia de ácido nucleico diana es una secuencia de ácido nucleico genómico diana, como una secuencia del genoma de los mamíferos o viral. En otro ejemplo específico, la secuencia diana es una secuencia de ácido nucleico diana de ARN.

En algunos ejemplos, la secuencia diana es una secuencia genómica, como una secuencia eucariótica (p. ej., de mamífero) o una secuencia genómica viral. Se pueden generar moléculas de ácido nucleico FPC que se corresponden básicamente con cualquier secuencia genómica diana que incluye al menos una porción de ADN único no repetitivo (p. ej., 14, 16, 18, 42, 44, 46 en la FIG. 1A). Se puede utilizar básicamente cualquier secuencia de ácido nucleico genómico diana para generar sondas que incluyen moléculas de ácido nucleico FPC con regiones de unión específicas para la secuencia de ácido nucleico genómico diana. Por ejemplo, la secuencia genómica diana puede ser una porción de un genoma eucariótico, como un genoma de mamífero (p. ej., humano), fúngico o de un parásito intracelular. Alternativamente, una secuencia genómica diana puede ser un genoma procariótico o viral (como un genoma bacteriano) o una porción del mismo. En un ejemplo específico, la secuencia genómica diana está asociada con un organismo infeccioso (p. ej., virus, bacterias, hongos). Las moléculas de ácido nucleico FPC y las sondas que incluyen estas moléculas se pueden corresponder con genes individuales (incluyendo las porciones de codificación y/o no codificación de los genes) o regiones de cromosomas (p. ej., que incluyen uno o más genes de interés).

La secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) puede tener cualquier número de pares de bases. En un ejemplo, como el de una secuencia de ácido nucleico genómico diana seleccionado de un mamífero u otro genoma con una secuencia de ácido nucleico repetitiva intercalada sustancial, la secuencia de ácido nucleico diana tiene al menos 1.000 pares de bases. En determinados ejemplos, una secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) tiene al menos 10.000, al menos 50.000, al menos 100.000, al menos 150.000, al menos 250.000, o al menos 500.000 pares de bases de longitud (como de 100 kb a 600 kb, de 200 kb a 500 kb, o de 300 kb a 500 kb). En ejemplos en los que la secuencia de ácido nucleico diana es de un genoma eucariótico (como un genoma de mamífero, por ejemplo, un genoma humano), la secuencia diana representa típicamente una pequeña porción del genoma (o una pequeña porción de un único cromosoma) del organismo (por ejemplo, menos del 20%, menos del 10%, menos del 5%, menos del 2%, o menos del 1 % del ADN genómico (o un único cromosoma) del organismo). En algunos ejemplos, donde la secuencia diana (p. ej., la secuencia de ácido nucleico genómico diana) es de un organismo infeccioso (como un virus), la secuencia diana puede representar una proporción mayor (por ejemplo, del 50% o más) o incluso la totalidad del genoma del organismo infeccioso.

En ejemplos específicos, mostrados a título meramente enunciativo, se selecciona una secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) asociada con una neoplasia (por ejemplo, un cáncer). Se han identificado numerosas anomalías cromosómicas (incluyendo traslocaciones y otras redistribuciones, reduplicación o delección) en las células neoplásicas, especialmente en las células cancerosas, como las leucemias de células B y T, linfomas, cáncer de mama, cáncer de colon, cánceres neurológicos y similares. Por tanto, en algunos ejemplos, al menos una porción de la secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) es objeto de la reduplicación o delección en al menos un subconjunto de células de una muestra.

Las traslocaciones en las que hay implicados oncogenes son conocidas por varias malignidades humanas. Por ejemplo, las reorganizaciones cromosómicas que afectan al gen SYT ubicado en la región del punto de inflexión del cromosoma 18q11.2 son frecuentes entre los tumores de tejidos blandos, como el sarcoma sinovial. Las traslocación de t(18q11.2) se puede identificar, por ejemplo, utilizando sondas con diferentes etiquetas: la primera sonda incluye moléculas de ácido nucleico FPC generadas de una secuencia de ácido nucleico diana que se extiende distalmente desde el gen SYT, y la segunda sonda incluye ácido nucleico FPC generado de una secuencia de ácido nucleico diana que se extiende 3' o proximal al gen SYT. Cuando las sondas correspondientes a estas secuencias de ácido nucleico diana (p. ej., una sonda de ácido nucleico genómico diana) se utilizan en un procedimiento de hibridación *in situ*, las células normales, que carecen de un t(18q11.2) en la región del gen SYT, exhiben dos señales de fusión (generadas por las dos etiquetas en estrecha proximidad, reflejando las dos copias intactas de SYT. Las células anómalas con un t(18q11.2) exhiben una única señal de fusión.

Se han observado numerosos ejemplos de reduplicación de genes implicados en la transformación neoplásica y pueden ser detectados citogenéticamente mediante hibridación *in situ*, utilizando las sondas divulgadas. En un ejemplo, se selecciona la secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) que incluye un gen (p. ej., un oncogen) que está reduplicado en una o más malignidades (p. ej., una malignidad humana). Por ejemplo, HER2, también conocido como c-erbB2 o HER2/neu, es un gen que participa en la regulación del crecimiento celular (una secuencia genómica HER2 humana representativa se proporciona en el n° de acceso GENBANK™ NCJ300017, nucleótidos 35097919-35138441). El gen codifica un receptor transmembrana de la superficie celular de 185 kd que es miembro de la familia de la tirosina quinasa. HER2 se amplifica en el cáncer de mama, de ovarios y otros cánceres humanos. Por tanto, un gen HER2 (o una región del cromosoma 17 que incluye el gen HER2) se puede utilizar como una secuencia de ácido nucleico genómico diana para generar sondas que incluyen moléculas de ácido nucleico FPC con regiones de unión específicas para HER2.

En otros ejemplos, se selecciona una secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) que es un gen supresor tumoral que está sujeto a delección (perdido) en las células malignas. Por ejemplo, la región p16 (incluyendo D9S1749, D9S1747, pl6(INK4A), pl4(ARF), D9S1748, pl5(INK4B), y D9S1752) ubicada en el cromosoma 9p21 está sujeta a delección en determinados cánceres de vejiga. Las delecciones cromosómicas que afectan a la región distal del brazo corto del cromosoma 1 (que abarca, por ejemplo, SHGC57243, TP73, EGFL3, ABL2, ANGPTL1, y SHGC-1322) y la región pericentromérica (p. ej., 19p13-19q13) del cromosoma 19 (que abarca, por ejemplo, MAN2B1, ZNF443, ZNF44, CRX, GLTSCR2, y GLTSCR1) son características moleculares típicas de determinados tipos de tumores sólidos del sistema nervioso central.

Los mencionados ejemplos se proporcionan únicamente y exclusivamente a título ilustrativo. Muchas otras anomalías citogenéticas que están correlacionadas con el crecimiento y/o la transformación neoplásica son conocidas para las personas con conocimientos en el campo. Las secuencias de ácido nucleico diana (p. ej., secuencias de ácido nucleico genómico diana), que han sido correlacionadas con la transformación neoplásica y que son útiles en los métodos divulgados y para las que se pueden preparar las sondas divulgadas también incluyen el gen EGFR (7p12; p. ej., n° de acceso GENBANK™ NC\_000007, nucleótidos 55054219-55242525) el gen C-MYC (8q24.21; p. ej., n° de acceso GENBANK NC\_000008, nucleótidos 128817498-128822856), D5S271 (5p15.2), gen lipoproteinlipasa (LPL) (8p22; p. ej., n° de acceso GENBANK™ NC\_000008, nucleótidos 19841058-19869049), RBI (13qJ4; p. ej., n° de acceso GENBANK™ NCJ)00013, nucleótidos 47775912-47954023), p53 (17p13.1; p. ej., n° de acceso GENBANK™ NC\_000017, complemento, nucleótidos 7512464-7531642)), N-MYC (2p24; p. ej., n° de acceso GENBANK™ NCJ300002, complemento, nucleótidos 151835231-151854620), CHOP (12q1.3; p. ej., n° de acceso GENBANK™ NC\_000012, complemento, nucleótidos 56196638-56200567), FUS (16p11.2; p. ej., n° de acceso GENBANK™ NC\_000016, nucleótidos 31098954-31110601), FKHR (13p14; p. ej., n° de acceso GENBANK™ NC\_000013, complemento, nucleótidos 40027817-40138734), así como, por ejemplo: ALK (2p23; p. ej., n° de acceso GENBANK™ NC\_000002, complemento, nucleótidos 29269144-29997936), cadena pesada de Ig, CCND1 (J lq13; p. ej., n° de acceso GENBANK™ NC\_000011, nucleótidos 69165054.69178423), BCL2 (18q21.3; p. ej., n° de acceso GENBANK™ NC\_000018, complemento, nucleótidos 58941559-59137593), BCL6 (3q27; p. ej., n° de acceso GENBANK™ NC\_000003, complemento, nucleótidos 188921859-188946169), MALF1, API (1p32-p31; p. ej., n° de acceso GENBANK™ NC\_000001, complemento, nucleótidos 59019051-59022373), TOP2A (17q21-q22; p. ej., n° de acceso GENBANK™ NC\_000017, complemento, nucleótidos 35798321-35827695), TMPRSS (21q22.3; p. ej., n° de acceso GENBANK™ NCJ) 00021, complemento, nucleótidos 41758351-41801948), ERG (21q22.3; p. ej., n° de acceso GENBANK™ NC\_000021, complemento, nucleótidos 38675671-38955488); ETV1 (7p21.3; p. ej., n° de acceso GENBANK™ NC\_000007, complemento, nucleótidos 13897379-13995289), EWS (22q12.2; p. ej., n° de acceso GENBANK™ NC\_000022, nucleótidos 27994271-28026505); FLU (1 lq24.1-q24.3; p. ej., n° de acceso GENBANK™ NC\_000011, nucleótidos 128069199-128187521), PAX.3 (2q35-q37; p. ej., n° de acceso GENBANK™ NC\_000002, complemento, nucleótidos 222772851-222871944), PAX7 (1p36.2-p36.12; p. ej., n° de acceso GENBANK™ NC\_000001, nucleótidos 18830087-18935219), PTEN (10q23.3; p. ej., n° de acceso GENBANK™ NC\_000010, nucleótidos 89613175-89716382), AKT2 (19q1.3.1-q13.2; p. ej., n° de acceso GENBANK™ NC\_000019, complemento, nucleótidos 45431556-454830.36), MYCL1 (1p34.2; p. ej., n° de acceso GENBANK™ NC\_000001, complemento, nucleótidos 40133685-40140274), REL (2p1.3-pl2; p. ej., n° de acceso GENBANK™ NC\_000002, nucleótidos 60962256-61003682) y CSF1R (5q33-q35; p. ej., n° de acceso GENBANK™ NC\_000005, complemento, nucleótidos 149413051-149473128). Una sonda o método divulgado puede incluir una región del respectivo cromosoma humano que contiene al menos uno (o más, según resulte aplicable) de los anteriores genes. Por

ejemplo, la secuencia de ácido nucleico diana para algunas sondas o métodos divulgados incluye cualquiera de los anteriores genes y una secuencia genómica contigua adicional suficiente (sea 5' del gen, 3' del gen o una combinación de los mismos) por un total de al menos 100.000 pares de bases (como al menos 250.000 o al menos 500.000 pares de bases) o un total de entre 100.000 y 500.000 pares de bases.

5 En determinadas realizaciones, la sonda específica para la molécula de ácido nucleico diana se somete a ensayo (en la misma muestra u otra diferente pero análoga) en combinación con una segunda sonda que proporciona una indicación del número de cromosoma, como una sonda específica para el cromosoma (p. ej., un centrómero). Por ejemplo, una sonda específica para una región del cromosoma 17 que contiene al menos el gen HER2 (una sonda HER2) se puede utilizar en combinación con una sonda CEP 17 que se hibrida con el ADN satélite alfa ubicado en el  
10 centrómero del cromosoma 17 (17pl 1 1-qll.l). La inclusión de la sonda CEP 17 permite determinar el número de copia relativa del gen HER2. Por ejemplo, las muestras normales tendrán un ratio HER2/CEP17 inferior a 2, mientras que las muestras en las que el gen HER2 está reduplicado tendrán un ratio HER2/CEP17 superior a 2.0. De forma similar, las sondas del centrómero CEP correspondientes a la ubicación de cualquier otra secuencia genómica diana seleccionada también se pueden utilizar en combinación con una sonda para una diana única en el mismo (o  
15 diferente) cromosoma.

En otros ejemplos, se selecciona una secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) de un virus u otro microorganismo asociado con una enfermedad o condición. La detección de la secuencia de ácido nucleico diana obtenida del virus o microorganismo (p. ej., una secuencia de ácido nucleico genómico diana) en una célula o muestra de tejido es indicativa de la presencia del organismo. Por ejemplo, la  
20 sonda se puede seleccionar del genoma de un virus oncogénico o patógeno, una bacteria o un parásito intracelular (como *Plasmodium falciparum* y otras *Plasmodium species*, *Leishmania* (sp.), *Cryptosporidium parvum*, *Entamoeba histolytica*, y *Giardia lamblia*, así como *Toxoplasma*, *Eimeria*, *TJieileria*, y *Babesia species*).

En algunos ejemplos, la secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) es un genoma viral. Algunos ejemplos de virus y las correspondientes secuencias genómicas (n° de acceso RefSeq  
25 GENBANK™, entre paréntesis) incluyen un adenovirus humano A (NC\_001460), adenovirus humano B (NC\_004001), adenovirus humano C (NC\_001405), adenovirus humano D (NC\_002067), adenovirus humano E (NC\_003266), adenovirus humano F (NC\_001454), astrovirus humano (NC\_001943), poliomavirus humano BK (V01109; G160851) bocavirus humano (NC\_007455), coronavirus humano 229E (NC\_002645), coronavirus humano HKU1 (NC\_006577), coronavirus humano NL63 (NC\_005831), coronavirus humano OC43 (NC\_005147), enterovirus humano A (NC\_001612), enterovirus humano B (NC\_001472), enterovirus humano C (NC\_001428), enterovirus humano D (NC\_001430), eritrovirus humano V9 (NCJ304295), espumavirus humano (NC\_001736), herpesvirus humano 1 (virus del herpes simple tipo I) (NC\_001806), herpesvirus humano 2 (virus del herpes simple tipo 2) (NC\_001798), herpesvirus humano 3 (virus varicella-zóster) (NC\_001348), herpesvirus humano 4 tipo 1 (virus Epstein-Ban tipo 1) (NC\_007605), herpesvirus humano 4 tipo 2 (virus Epstein-Ban tipo 2) (NCJ09334), herpesvirus humano 5 cadena AD169 (NC\_001347), herpesvirus humano 5 cadena Merlin Strain (NC\_006273), herpesvirus humano 6A (NC\_001664), herpesvirus humano 6B (NC\_000898), herpesvirus humano 7 (NC\_001716), herpesvirus humano 8 tipo M (NC\_003409), herpesvirus humano 8 tipo P (NC\_009333), virus de la inmunodeficiencia humana 1 (NC\_001802), virus de la inmunodeficiencia humana 2 (NC\_001722), metapneumovirus humano (NC\_004148), virus del papiloma humano 1 (NC\_001356), virus del papiloma humano 18 (NCL001357), virus del papiloma humano 2 (NC\_001352), virus del papiloma humano 54 (NC\_001676), virus del papiloma humano 61 (NC\_001694), virus del papiloma humano cand90 (NC\_004104), virus del papiloma humano RTRX7 (NC\_004761), virus del papiloma humano tipo 10 (NC\_001576), virus del papiloma humano tipo 101 (NCJ308189), virus del papiloma humano tipo 103 (NCL008188), virus del papiloma humano tipo 107 (NC\_009239), virus del papiloma humano tipo 16 (NC\_001526), virus del papiloma humano tipo 24 (NC.001683), virus del papiloma humano tipo 26 (NC\_001583), virus del papiloma humano tipo 32 (NC.001586), virus del papiloma humano tipo 34 (NC\_001587), virus del papiloma humano tipo 4 (NC\_001457), virus del papiloma humano tipo 41 (NCJXH354), virus del papiloma humano tipo 48 (NC\_001690), virus del papiloma humano tipo 49 (NC\_001591), virus del papiloma humano tipo 5 (NCJXH531), virus del papiloma humano tipo 50 (NC\_001691), virus del papiloma humano tipo 53 (NC\_001593), virus del papiloma humano tipo 60 (NC\_001693), virus del papiloma humano tipo 63 (NCJJ01458), virus del papiloma humano tipo 6b (NC\_001355), virus del papiloma humano tipo 7 (NC\_001595), virus del papiloma humano tipo 71 (NC\_002644), virus del papiloma humano tipo 9 (NCJ301596), virus del papiloma humano tipo 92 (NC\_004500), virus del papiloma humano tipo 96 (NC\_005134), virus de la parainfluenza humano I (NC\_003461), virus de la parainfluenza humano 2 (NC\_003443), virus de la parainfluenza humano 3 (NC\_001796), parecovirus humano (NC\_001897), parvovirus humano 4 (NC\_007018), parvovirus humano B19 (NC\_000883), virus sincitial respiratorio humano (NC\_001781),  
55 rinovirus humano A (NC\_001617), rinovirus humano B (NC\_001490), retrovirus espuma humano (NC\_001795), virus linfotrópico T humano I (NC\_001436), virus linfotrópico T humano 2 (NC\_001488).

En determinados ejemplos, la secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) es de un virus oncogénico, como el virus Epstein-Ban Virus (EBV) o un virus del papiloma humano (HPV, p. ej., HPV16, HPV18). En otros ejemplos, la secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) es de un virus patógeno, como un virus sincitial respiratorio, un virus de la hepatitis (p. ej., virus de la hepatitis C), un coronavirus (p. ej., virus SARS), un adenovirus, un poliomavirus, un citomegalovirus (CMV) o un virus del herpes simple (HSV).  
60

**Ejemplos de secuencias de ácido nucleico no deseadas**

5 Algunas moléculas de ácido nucleico diana pueden contener secuencias de nucleótidos que cabe esperar razonablemente que aumenten la unión no específica de una sonda de ácido nucleico a una secuencia no diana (p. ej., secuencias repetitivas en secuencias diana de ácido nucleico genómico, secuencias que codifican dominios conservados en secuencias diana de ARN o secuencias homólogas en secuencias diana de ácido nucleico del genoma viral). Estas regiones se denominan en el presente secuencias no deseadas. Como se ha descrito anteriormente, estas secuencias que potencialmente producen una señal de fondo (p. ej., secuencias de ácido nucleico repetitivas) son sustancialmente (o incluso completamente) eliminadas de las secuencias diana para generar una población de regiones de unión utilizadas para generar las sondas y moléculas de ácido nucleico FPC divulgadas. Por ejemplo, al menos el 80%, al menos el 90%, al menos el 95%, al menos el 98% o al menos el 99% de la secuencia de ácido nucleico que potencialmente produce una señal de fondo (p. ej., una secuencia repetitiva, una secuencia de codificación de dominios conservados o una secuencia homóloga) se puede eliminar de una secuencia diana para generar una población de regiones de unión utilizadas para generar las sondas y moléculas de ácido nucleico FPC divulgadas en el presente.

15 No obstante, en algunos ejemplos, una secuencia de ácido nucleico diana, como una secuencia genómica diana viral o una secuencia de ARN diana, puede no incluir ninguna o solamente una pequeña cantidad de secuencias de ácido nucleico no deseadas (p. ej., que potencialmente producen una señal de fondo). En tales ejemplos, las secuencias de ácido nucleico no deseadas (si las hubiere) pueden, aunque no es necesario, ser eliminadas de la secuencia diana antes de generar una población de regiones de unión de la secuencia de ácido nucleico diana.

20 En un ejemplo, la molécula de ácido nucleico diana es una molécula de ARN, que puede incluir secuencias que codifican dominios conservados (p. ej., el dominio de unión de ADN de los receptores de esteroides), donde uno o más dominios conservados son sustancialmente eliminados para reducir la unión no específica de una sonda que contiene moléculas de ácido nucleico FPC específicas para la secuencia diana de ARN.

25 Como se ha descrito anteriormente, las secuencias de ácido nucleico repetitivas están presentes en algunas secuencias de ácido nucleico diana (p. ej., secuencias de ácido nucleico genómico diana) y son sustancial (o incluso completamente) eliminadas de estas secuencias diana para generar una población de regiones de unión utilizadas para generar las sondas y moléculas de ácido nucleico FPC divulgadas. Las secuencias de ácido nucleico repetitivas (o elementos repetitivos) son conocidas en el campo. Las principales clases de secuencias de ácido nucleico repetitivas intercaladas del genoma humano incluyen las repeticiones Alu y Line (L1). Las repeticiones Alu son la secuencia de ácido nucleico repetitiva intercalada más abundante en el genoma humano con un número de copias totales de aproximadamente un millón. La secuencia Alu tiene aproximadamente 300 pares de bases de longitud y ocurre de media en uno de cada 3.300 pares de bases. Ocurren en toda la familia de los primates y son homólogas de un pequeño y abundante gen de ARN que codifica la molécula de ARN de 300 nucleótidos de largo conocida como 7SL. Las secuencias de ácido nucleico repetitivas L1 son secuencias de repetición intercaladas de entre 1.000 y 7.000 pares de bases. Las L1 tienen una secuencia común en el extremo 3', pero son variablemente acortadas en el extremo 5' (lo que da cuenta de la disparidad de longitud). Ocurren, de media, en uno de cada 28.000 pares de bases en el genoma humano, para un número de copia total de unos 100.000. A diferencia de las secuencias de ácido nucleico repetitivas Alu, que se limitan a los primates, las secuencias de ácido nucleico repetitivas L1 se encuentran en la mayor parte de las demás especies de mamíferos.

40 Las repeticiones de microsatélite incluyen una variedad de repeticiones en tándem de mono/di/tri/tetra/pentanucleótidos que están dispersas en los brazos eucromáticos de la mayoría de los cromosomas. La repetición dinucleótida (GT)<sub>n</sub> es la más común de estas secuencias de ácido nucleico repetitivas dispersas, que ocurre de media cada 30.000 bases en el genoma humano, para un número de copia total de 100.000. Las repeticiones GT tienen un tamaño de unos 20 a 60 pares de bases y aparecen en la mayoría de los genomas eucarióticos. Las repeticiones de minisatélite son una clase de repeticiones en tándem dispersas, en las que la unión de repetición tiene de 30 a 35 pares de bases de longitud y que tiene una secuencia variable, pero contiene una secuencia central de 10 a 15 pares de bases de longitud. El tamaño de las repeticiones de minisatélite varía de los 200 pares de bases hasta los miles pares de bases, y están presentes en números de copia inferiores que las repeticiones de microsatélite. Las repeticiones de minisatélite tienden a ocurrir en mayores números hacia los extremos teloméricos de los cromosomas.

55 Otras secuencias de ácido nucleico repetitivas están predominantemente limitadas a determinadas estructuras del cromosoma. Las repeticiones del telómero consisten en repeticiones en tándem de la secuencia «TTAGGG» (SEC ID N°: 359) y se encuentran ubicadas en los mismos extremos de las moléculas de ADN lineal en los cromosomas de humanos y vertebrados. Las repeticiones subteloméricas incluyen clases de secuencias repetitivas que están intercaladas en las últimas 500.000 bases de ADN no repetitivo ubicado junto al telómero. Algunas secuencias de ácido nucleico repetitivas son específicas de un cromosoma y otras parecen estar presentes cerca de los extremos de todos los cromosomas humanos.

El ADN satélite alfa es una familia de secuencias de ácido nucleico repetitivas relacionadas que aparecen como largas agrupaciones en tándem en la región centromérica de todos los cromosomas humanos. La unidad de

repetición es de unos 340 pares de bases y aparece como un dímero compuesto por dos subunidades, cada una de unos 170 pares de bases de largo. El ADN satélite alfa ocurre a ambos lados de la constricción centromérica y se extiende hasta 5.000 pares de bases desde el centrómero.

5 Las repeticiones satélite I, II y III son los tres ADN humanos satélite clásicos. Los ADN satélite pueden aislarse del grueso del ADN genómico mediante irrigación central en gradientes de densidad variables porque sus densidades difieren de las densidades de otras secuencias de ADN. La satélite I es rica en A y T y se compone de grupos alternos de una unidad de repetición de 17 y 25 pares de bases. Las satélites II y III se obtienen ambos de la unidad de repetición de cinco bases simple «ATTCC» (SEC ID N°: 360). La satélite II es más divergente de la unidad de repetición básica que la satélite III. Las satélites I, II y III aparecen como largas agrupaciones en tándem en las regiones heterocromáticas de los cromosomas humanos 1,9, 16, 17 e Y, y las regiones satélite de los brazos cortos (p) de los cromosomas humanos 13, 14, 15, 21 y 22.

15 El ADN C<sub>0</sub>t-1™ es una fracción de ADN repetitivo que es separable de otro ADN genómico basándose en la capacidad de reanillar más rápidamente tras la fusión para disociar las dos cadenas de la hélice. La definición del ADN C<sub>0</sub>t-1™ DNA se basa en las propiedades de hibridación (o de reanillar) del ADN más que en las características de la secuencia e incluye una población heterogénea de los elementos repetitivos específicos comentados anteriormente.

En determinados ejemplos, las moléculas de ácido nucleico FPC divulgadas no incluyen cantidades apreciables de las secuencias de ácido nucleico no deseadas (p. ej., repetitivas) (p. ej., 20, 22, 24 de la FIG. 1 A), incluyendo, por ejemplo, cualquiera de las clases expresamente recogidas anteriormente.

## 20 **Etiquetas detectables**

Las moléculas de ácido nucleico FPC divulgadas pueden incluir una o más etiquetas, por ejemplo para permitir la detección de una molécula de ácido nucleico diana, utilizando las sondas divulgadas. En diversas aplicaciones, como los procedimientos de hibridación *in situ*, una molécula de ácido nucleico FPC incluye una etiqueta (p. ej., una etiqueta detectable). Una «etiqueta detectable» es una molécula o material que se puede utilizar para producir una señal detectable que indica la presencia o concentración de la sonda (particularmente la sonda hibridada o unida) en una muestra. Así, una molécula de ácido nucleico FPC etiquetada proporciona un indicador de la presencia o concentración de una secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) (al que la molécula de ácido nucleico FPC etiquetada se une o hibrida) en una muestra. La divulgación no se limita al uso de etiquetas concretas, aunque se proporcionan ejemplos.

30 Una etiqueta asociada con una molécula de ácido nucleico FPC se puede detectar directa o indirectamente. Una etiqueta puede ser detectada por cualquier mecanismo conocido o por descubrir, incluyendo la absorción, emisión y/o difusión de un fotón (incluyendo fotones de radiofrecuencia, frecuencia de microondas, frecuencia infrarroja, frecuencia visible y frecuencia ultravioleta). Las etiquetas detectables incluyen materiales y moléculas coloreados, fluorescentes, fosforescentes y luminiscentes, catalizadores (como enzimas) que convierten una sustancia en otra sustancia para proporcionar una diferencia detectable (como la conversión de una sustancia incolora en una sustancia coloreada o viceversa, o produciendo una precipitación o aumentando la turbidez de la muestra), haptenos que pueden ser detectados por interacciones de unión de los anticuerpos, y materiales o moléculas magnéticos y paramagnéticos.

40 Algunos ejemplos concretos de etiquetas detectables incluyen las moléculas fluorescentes (o fluorocromos). Hay numerosos fluorocromos conocidos para aquellas con conocimientos en el campo y pueden ser seleccionados, por ejemplo, de Invitrogen, véase, p. ej., *The Handbook — A Guide to Fluorescent Probes and Labeling Technologies*, Invitrogen Detection Technologies, Molecular Probes, Eugene, OR). Algunos ejemplos concretos de fluoróforos que se pueden unir (por ejemplo, conjugándolos químicamente) a una molécula de ácido nucleico FPC se proporcionan en la Patente estadounidense n° 5.866.366 de Nazarenko *et al.*, como ácido 4-acetamido-4'-isotiocianatostilbeno-2,2'-disulfónico, acridina y derivados como isotiocianato de acridina, ácido 5-(2'-aminoetil)aminonaftaleno-1-sulfónico (EDANS), 4-amino-N-[3-vinylsulfonyl]phenyl]naftalimida-3,5 disulfonato (Lucifer Yellow VS), N-(4-anilino-1-naftil)maleimida, antranilamida, Brilliant Yellow, cumarina y derivados como 7-amino-4-metilcumarina (AMC, Coumarin 120), 7-amino-4-trifluorometilcumarina (Coumaran 151); cianosina; 4',6-diaminidino-2-fenilindol (DAPI); 5', 5"-dibromopirogalol-sulfoneftaleína (Bromopyrogallol Red); 7-dietilamino-3-(4'-isotiocianatofenil)-4-metilcumarina; dietilenetriamina pentaacetato; ácido 4,4'-diisotiocianatodihidro-estilbeno-2,2'-disulfónico; ácido 4,4'-diisotiocianatoestilbeno-2,2'-disulfónico; cloruro 5-[dimetilamino]naftaleno-1-sulfonil (DNS, cloruro de dansilo); ácido 4-(4'-dimetilaminofenilazo)benzoico (DABCYL); 4-dimetilaminofenilazofenil-4'-isotiocianato (DABITC); eosina y derivados, como el isotiocianato de eosina; eritrosina y derivados, como eritrosina B e isotiocianato de eritrosina; etidio; fluoresceína y derivados, como 5-carboxifluoresceína (FAM), 5-(4,6-diclorotriacín-2-il)aminofluoresceína (DTAF), 2'-7'-dirnetoxi-4'-5'-dicloro-6-carboxifluoresceína (JOE), fluoresceína, isotiocianato de fluoresceína (FITC), y QFITC (XRITC); 2', 7'-difluorofluoresceína (OREGON GREEN®); fluorescamina; IR144; IR1446; isotiocianato Malachite Green; 4-metilumbeliferona; orto-cresolfaleína; nitrotiosina; pararosahilina; Phenol Red; B-ficoeritrina; o-ftaldialdehído; pireno y derivados, como butirato de pireno y succinimidilo 1-pireno butirato; Reactive Red 4 (Cibacron RTM Brilliant Red 3B-A); rodamina y derivados, como 6-carboxi-X-rodamina (ROX), 6-carboxirodamina (R6G),

cloruro de lisamina-rodamina B sulfonilo, rodamina (Rhod), rodamina B, rodamina 123, rodamina X isotiocianato, rodamida-verde, sulforhodamina B, sulforhodamina 101 y el cloruro de sulfonilo derivado de la sulforhodamina 101 (Texas Red); N,N,N'N' Metrametil-6-carboxirodamina (TAMRA); tetrametil rodamina; isotiocianato de tetrametil-rodamina (TRITC); riboflavina; ácido rosólico y derivados de quelato de terbio.

5 Otros fluoróforos adecuados incluyen los quelatos de europio tiol-reactivos, que emiten a aproximadamente 617 nm (Heyduk y Heyduk, *Analyt. Biochem.* 248:216-27, 1997; *J. Biol Chem*, 274:3315-22, 1999), así como GFP, Lissamine™, dietilaminocumarina, fluoresceína clorotiacionilo, naftofluoresceína, 4,7-diclororodamina y xanteno (tal y como se describe en la Patente estadounidense nº 5.800.996 de Lee et al.) y derivados. También se pueden emplear otros fluoróforos conocidos para las personas con conocimientos en el campo, como los que ofrece  
 10 Invitrogen Detection Technologies, Molecular Probes (Eugene, OR) e incluye la serie de tintes ALEXA FLUOR® (por ejemplo, como se describe en las Patentes estadounidenses nº 5.696.157, 6.130.101 y 6.716.979), la serie de tintes BODIPY (tintes de difluoruro de boro-dipirrometano, por ejemplo como se describe en las Patentes estadounidenses nº 4.774.339, 5.187.288, 5.248.782, 5.274.113, 5.338.854, 5.451.663 y 5.433.896), Cascade Blue (un derivado amino-reactivo del pireno sulfonado descrito en la Patente estadounidense nº 5.132.432) y Marina Blue (Patente estadounidense nº 5.830.912).

Además de los fluorocromos anteriormente descritos, una etiqueta fluorescente puede ser una nanopartícula fluorescente, como un nanocrystal semiconductor, p. ej., un QUANTUM DOT™ (obtenido, por ejemplo, de QuantumDot Corp, Invitrogen Nanocrystal Technologies, Eugene, OR; véanse también las Patentes estadounidenses nº 6.815.064, 6.682.596 y 6.649.138). Los nanocristales semiconductores son partículas  
 20 microscópicas que tienen propiedades eléctricas y/u ópticas dependientes del tamaño. Cuando los nanocristales semiconductores se iluminan con una fuente de energía primaria, se produce una emisión secundaria de energía de una frecuencia que corresponde al salto de energía del material semiconductor utilizado en el nanocrystal semiconductor. Esta emisión puede ser detectada como una luz coloreada de una longitud de onda específica o fluorescencia. Los nanocristales semiconductores con diferentes características espectrales se describen, p. ej., en  
 25 la Patente estadounidense nº 6.602.671. Los nanocristales semiconductores que se pueden acoplar a diversas moléculas biológicas (incluyendo dNTP y/o ácidos nucleicos) o sustratos mediante las técnicas descritas en, por ejemplo, Bruchez *et. al.* (1998) *Science* 281:201, 3-6, Chan *et. at.* (1998) *Science* 281:2016-8, y en la Patente estadounidense nº 6.274.323.

La formación de nanocristales semiconductores de diversas composiciones se divulgan en, por ejemplo, las  
 30 Patentes estadounidenses nº 6.927.069; 6.914.256; 6.855.202; 6.709.929; 6.689.338; 6.500.622; 6.306.736; 6.225.198; 6.207.392; 6.114.038; 6.048.616; 5.990.479; 5.690.807; 5.571.018; 5.505.928; 5.262.357 y en la Publicación de la Patente estadounidense nº 2003/0165951, así como en la Publicación PCT nº 99/26299 (publicada el 27 de mayo de 1999). Se pueden producir poblaciones separadas de nanocristales semiconductores que sean identificables basándose en sus características espectrales diferentes. Por ejemplo, se pueden producir  
 35 nanocristales semiconductores que emiten luz de diferentes colores basándose en su composición y tamaño o tamaño y composición. Por ejemplo, los puntos cuánticos que emiten luz en diferentes longitudes de onda basándose en el tamaño (longitudes de onda de emisión de 565 nm, 655 nm, 705 nm o 800 nm), que son adecuados como etiquetas fluorescentes en las sondas divulgadas en el presente, están disponibles en Invitrosen.

Otras etiquetas adicionales incluyen, por ejemplo, radioisótopos (como <sup>3</sup>H), quelatos metálicos, como quelatos  
 40 DOTA y DPTA de iones metálicos paramagnéticos o radiactivos como Gd<sup>3+</sup> y liposomas.

Las etiquetas detectables que se pueden utilizar con las moléculas de ácido nucleico FPC divulgadas también incluyen las enzimas, como la peroxidasa de rábano, fosfatasa alcalina, fosfatasa ácida, glucosa oxidasa, β-  
 galactosidasa, β-glucuronidasa o β-lactamasa. Cuando la etiqueta detectable incluye una enzima, se puede emplear  
 45 un cromógeno, un compuesto fluorogénico o un compuesto luminogénico en combinación con la enzima para generar una señal detectable (muchos de estos compuestos están disponibles en el mercado, por ejemplo, de Invitrogen Corporation, Eugene OR). Algunos ejemplos concretos de compuestos cromogénicos incluyen la diaminobencidina (DAB), 4-nitrofenilfosfato (pNPP), rojo rápido, bromocloroindolil-fosfato (BC1P), nitroazul de tetrazolio (NBT), BCIP/NBT, rojo rápido, AP-Orange, AP-Blue, tetrametilbencidina (TMB), 2,2'-acino-di-[3-  
 50 etilbenzotiazolina sulfonato] (ABTS), o -dianisidina, 4-cloronaftol (4-CN), nitrofenil-β-D-galactopiranosida (ONPG), o-fenilenodiamina (OPD), 5-bromo-4-cloro-3-indolil-β-galactopiranosida (X-Gal), metilumbeliferil-β-D-galactopiranosida (MU-Gal), p-nitrofenil-α-D-galactopiranosida (PNP), 5-bromo-4-cloro-3-indolil-β -D-glucuronida (X-Gluc), 3-amino-9-etil carbazol (AEC), fucsina, yodonitrotetrazolio (INT), azul de tetrazolio y violeta de tetrazolio.

Alternativamente, se puede utilizar una enzima en un programa de detección metalográfica. Por ejemplo, los  
 55 procedimientos SISH implican programas de detección metalográfica para la identificación y localización de una secuencia de ácido nucleico genómico diana hibridada. Los métodos de detección metalográfica incluyen el uso de una enzima, como fosfatasa alcalina, en combinación con un ión metálico soluble en agua y un sustrato redox-inactivo de la enzima. El sustrato se convierte en un agente redox-activo con la enzima y el agente redox-activo reduce el ión metálico, haciendo que forme un precipitado detectable (véase, por ejemplo, la Publicación de la Solicitud de Patente estadounidense nº 2005/0100976, la Publicación PCT nº 2005/003777 y la Publicación de la

Solicitud de Patente estadounidense nº 2004/0265922). Los métodos de detección metalográfica incluyen el uso de una enzima oxido-reductasa (como peroxidasa de rábano), junto con un ión metálico soluble en agua, un agente oxidante y un agente reductor, una vez más para formar un precipitado detectable (véase, por ejemplo, la Patente estadounidense nº 6.670.113).

- 5 En algunos ejemplos descritos a continuación a título meramente enunciativo, las moléculas de ácido nucleico FPC están etiquetadas con dNTP acoplados covalentemente a moléculas de hapteno (como un compuesto nitroaromático (p. ej., dinitrofenilo (DNP)), biotina, fluoresceína, digoxigenina, etc.). Los métodos para conjugar los haptenos y otras etiquetas con los dNTP (p. ej., para facilitar la incorporación a las sondas etiquetas) son bien conocidos en el campo. Para consultar ejemplos de los procedimientos, véanse, p. ej., las Patentes estadounidenses nº 5.258.507, 4.772.691, 5.328.824, y 4.711.955. En efecto, hay muchos dNTP etiquetados disponibles en el mercado, por ejemplo de Invitrogen Detection Technologies (Molecular Probes, Eugene, OR). Una etiqueta se puede acoplar directa o indirectamente a un dNTP en cualquier ubicación del dNTP, como un fosfato (p. ej., fosfato  $\alpha$ ,  $\beta$  o  $\gamma$ ) o un azúcar. La detección de las moléculas de ácido nucleico FPC etiquetadas se puede realizar poniendo en contacto las moléculas de ácido nucleico FPC etiquetadas con hapteno unidas a la secuencia genómica diana con un anticuerpo anti-hapteno primario. En un ejemplo, el anticuerpo anti-hapteno primario (como un anticuerpo anti-hapteno murino) es etiquetado directamente con una enzima. En otro ejemplo, se utiliza un anticuerpo secundario (como un anticuerpo de cabra anti-IgG murino) conjugado con una enzima para la amplificación de la señal. En la CISH se añade un sustrato cromogénico; en la SISH, se añaden iones de plata y otros reactivos recogidos en las solicitudes/patentes mencionadas.
- 10
- 15
- 20 Se puede incorporar una etiqueta detectable a las moléculas de ácido nucleico FPC, conjugando la etiqueta detectable con un dNTP, que se incorpora a la molécula de ácido nucleico FPC. La incorporación de los dNTP etiquetados a las moléculas de ácido nucleico se comenta más detalladamente con respecto a los métodos para generar sondas.

#### **MÉTODOS PARA PRODUCIR SONIDAS**

- 25 En el presente se proporcionan los métodos para producir las sondas divulgadas que incluyen una población de moléculas de ácido nucleico FPC. A pesar de que se proporcionan métodos de ejemplo, la divulgación no se limita a estos métodos concretos. Una descripción general de los métodos que se pueden utilizar para generar las sondas divulgadas se proporciona en las FIG. 1A, 1B y 3. En pocas palabras, como se muestra en la FIG. 1A, se selecciona una molécula de ácido nucleico diana (p. ej., una molécula de ácido nucleico genómico diana) 10, 40. Si la molécula de ácido nucleico diana (p. ej., una molécula de ácido nucleico genómico diana) 10, 40 tiene o se piensa que tiene secuencias de ácido nucleico no deseadas (como secuencias que pueden incrementar la señal de fondo de una sonda, incluyendo secuencias repetitivas, secuencias homólogas, secuencias que codifican dominios conservados) (p. ej., 20, 22, 24), estas secuencias de ácido nucleico son identificadas y sustancialmente eliminadas. No obstante, si la molécula de ácido nucleico diana (p. ej., una molécula de ácido nucleico viral de un genoma viral o una molécula de ácido nucleico diana de ARN) 40 no tiene (o no tiene una cantidad significativa de) secuencias de ácido nucleico que producen una señal de fondo, se puede omitir este paso. Esta molécula de ácido nucleico diana 10, 40, se utiliza en el paso 60 para generar una población de regiones de unión 100, que está sustancialmente libre de secuencias de ácido nucleico que producen una señal de fondo (p. ej., repetitivas u otras secuencias no deseadas) 20, 22, 24. En algunos ejemplos, al menos el 80% (como al menos el 90% o al menos el 98%) de las secuencias que producen una señal de fondo (p. ej., repetitivas u otras secuencias no deseadas) 20, 22, 24 son eliminadas de la molécula de ácido nucleico diana.
- 30
- 35
- 40

- Las regiones de unión 114, 116, 118 de la población de regiones de unión está unidas 250. Las regiones de unión 114, 116, 118 se pueden unir directamente o incluir enlaces entre las regiones de unión 114, 116, 118. Opcionalmente, las regiones de unión 114, 116, 118 pueden someterse a una o más reacciones de amplificación 200 antes de la unión 250. La unión de las regiones de unión 114, 116, 118 produce la sonda 300 que incluye una población (como una población heterogénea) de moléculas de ácido nucleico FPC 310, 312, 314, 316, 318, 320, 322, 324, 326, 328, 330, 332.
- 45

- La FIG. 3 muestra una secuencia de ácido nucleico diana (p. ej., una molécula de ácido nucleico genómico diana) que contiene regiones de unión únicas (regiones grises) y secuencias de ácido nucleico repetitivas u otras no deseadas (regiones negras). Las regiones de unión únicas son aisladas (p. ej., amplificadas) de la secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana), resultando en una población de regiones de unión individuales que se corresponden con la secuencia de ácido nucleico diana. Posteriormente, las regiones de unión individuales se unen, resultando en la producción de moléculas de ácido nucleico FPC.
- 50

#### ***Identificación de secuencias de ácido nucleico no deseadas en una secuencia de ácido nucleico diana***

- 55 Como se ha comentado anteriormente, algunas secuencias de ácido nucleico diana útiles para las sondas y métodos divulgados contendrán elementos de ácido nucleico repetitivos u otros no deseados (p. ej., secuencias de ácido nucleico genómico diana). En esas realizaciones, se pueden producir sondas de ácido nucleico que están sustancialmente (p. ej., completamente) libres de secuencias no deseadas y que incluyen una población (p. ej., una

población heterogénea) de moléculas de ácido nucleico FPC, amplificando una o más de una (es decir, una pluralidad) regiones de unión de una secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) para producir una población de regiones de unión (p. ej., 100, 140 en la FIG. 1A). Como se ha comentado anteriormente, algunos ejemplos de moléculas de ácido nucleico FPC incluyen una pluralidad de regiones de unión contiguas, están sustancialmente libres de secuencias de ácido nucleico repetitivas u otras no deseadas y se corresponden con una secuencia diana. Las FIG. 1A, 1B y 3 ilustran esquemáticamente la producción de poblaciones de regiones de unión que están sustancialmente libres de secuencias de ácido nucleico repetitivas, y su incorporación a una población de moléculas de ácido nucleico FPC. Como se ha comentado anteriormente, una molécula de ácido nucleico FPC es un ácido nucleico que incluye dos o más regiones de unión que se corresponden con una secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana). Las regiones de unión individuales pueden estar dispuestas en la molécula de ácido nucleico FPC en un orden y/o una orientación diferentes a los de la secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana).

Los segmentos de la región de unión libres de repeticiones (es decir, sustancial o completamente libres de repeticiones o segmentos libres o sustancialmente libres de otra secuencia de ácido nucleico no deseada) que se corresponden con una secuencia de ácido nucleico diana seleccionada (como se ha señalado anteriormente) pueden ser identificados y/o aislados y/o producidos por diversos métodos. Siempre que las regiones de unión no contengan una secuencia de ácido nucleico repetitiva (u otra no deseada) apreciable, cualquier método para generar una población de regiones de unión correspondientes a una molécula de ácido nucleico diana (p. ej., una molécula de ácido nucleico genómico diana) resultará apropiado.

En un ejemplo, las regiones de unión de ácido nucleico sustancial o completamente libres de repetición (u otras no deseables, como por ejemplo las que generan una señal de fondo) correspondientes a la molécula de ácido nucleico diana (p. ej., una molécula de ácido nucleico genómico diana) se identifican utilizando un algoritmo generado por ordenador. Los programas o algoritmos generados por ordenador para identificar los elementos de ácido nucleico no deseados (p. ej., repetitivos) y para «eliminarlos» de una secuencia representada en una base de datos son bien conocidos en el campo (véase, p. ej., Queen y Korn, *Nucleic Acids Res.* 12:581-599, 1984).

Por ejemplo, RepeatMasker es un programa que busca bases de datos establecidas, utilizando el programa de alineación CrossMatch para identificar y «ocultar» los elementos de ácido nucleico repetitivos. De una forma similar, se puede utilizar el programa MaskeiAid, que sustituye al algoritmo de búsqueda WU-BLAST en el contexto del RepeatMasker1. Algunos programas alternativos incluyen, por ejemplo, FORRepeats (Lefebvre *et. al.*, *Bioinformatics* 19: 319-326, 2003), MUMmer (Delcher *et. al.*, *Nucleic Acids Res.* 27:2369-23 76, 1999) y REPuter (Kurtz *et. al.*, *Nucleic Acids Res* 29:4633-42, 2001), que también pueden ser utilizados para identificar elementos de ácido nucleico repetitivos dentro de una molécula de ácido nucleico diana (p. ej., una molécula de ácido nucleico genómico diana).

Se apreciará que los procesos implementados por el ordenador para identificar y eliminar las secuencias de ácido nucleico no deseadas (p. ej., repetitivas) se pueden realizar manualmente, sin la ayuda de un ordenador. Sin embargo, dado que con frecuencia se valoran grandes extensiones de ARN o ADN (que suelen superar los 100 kb) como moléculas de ácido nucleico diana (p. ej., una molécula de ácido nucleico genómico diana), el examen manual de una secuencia puede resultar difícil y llevar mucho tiempo. Por otra parte, dado que puede haber presentes múltiples clases de moléculas de ácido nucleico no deseadas (p. ej., repetitivas) en una molécula de ácido nucleico diana (p.ej., una molécula de ácido nucleico genómico diana) y dado que existe un cierto grado de variabilidad dentro de las diversas clases de moléculas de ácido nucleico no deseadas (p. ej., repetitivas), el examen manual puede ser menos preciso y exhaustivo que los procedimientos asistidos por ordenador. Por otra parte, las regiones de unión que están sustancial o completamente libres de secuencia de ácido nucleico no deseada (p. ej., repetitiva) se pueden determinar revisando una base de datos ya existente que contenga información de las secuencias relativas a moléculas de ácido nucleico diana previamente exploradas (p. ej., una molécula de ácido nucleico genómico diana) (p. ej., información de la secuencia obtenida a través de Masker Aid u otros programas implementados por ordenador).

#### **Generación de regiones de unión sustancialmente libres de la secuencia de ácido nucleico no deseada**

Una vez identificadas, las regiones de unión que están sustancial o completamente libres de secuencias de ácido nucleico repetitivas (u otras secuencias indeseables, como las que producen una señal de fondo) de una molécula de ácido nucleico diana (p. ej., una molécula de ácido nucleico genómico diana) se pueden generar utilizando métodos conocidos. Por ejemplo, las regiones de unión se pueden amplificar de una molécula de ácido nucleico diana, obtener de la molécula de ácido nucleico diana mediante hibridación sustractiva, aislar de una molécula de ácido nucleico diana o combinaciones de las mismas. En algunos ejemplos, la molécula de ácido nucleico diana está presente en su forma nativa o es parte de un vector que incluye la secuencia diana.

En un ejemplo, las regiones de unión son amplificadas de una secuencia de ácido nucleico diana (por ejemplo, mediante amplificación en una reacción en cadena de polimerasa o PCR). Básicamente, se puede emplear cualquier método de amplificación conocido en el campo para amplificar regiones de unión que están sustancial o completamente libres de secuencias de ácido nucleico repetitivas (u otras no deseables, como las que producen una

señal de fondo). En algunos ejemplos, se realizan reacciones de amplificación múltiples. Si se desea, los amplímeros de la región de unión resultantes pueden ser diluidos antes de la siguiente amplificación.

Un ejemplo de amplificación es la reacción en cadena de polimerasa (PCR), en la que una muestra biológica (como la que contiene moléculas de ácido nucleico genómico, por ejemplo un preparado cromosómico) se pone en contacto con un par de cebadores oligonucleótidos, en condiciones que permiten la hibridación de los cebadores con una plantilla de ácido nucleico de la muestra. Los cebadores se extienden en condiciones apropiadas, se disocian de la plantilla y posteriormente se reanillan, extienden y disocian para amplificar el número de copias del ácido nucleico. Hay numerosos procedimientos para la PCR conocidos en el campo y se pueden encontrar protocolos de ejemplo, p. ej., en Sambrook y Russell., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2001; y Ausubel *et. al.*, *Short Protocols in Molecular Biology*, 4<sup>a</sup> ed. John Wiley & Sons, Inc. 1999.

La amplificación de los segmentos de la región de unión que se corresponden con una secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) se puede realizar utilizando la reacción en cadena de polimerasa (PCR), que se describe en detalle en las Patentes estadounidenses nº 4.683.195, 4.683.202 y 4.800.159, así como en *PCR Protocols A Guide to Methods and Applications*, Innis *et. al.*, eds., Academic Press Inc., San Diego, CA, 1990, (véase también, Sambrook, Ausubel). La PCR utiliza pares de cebadores que tienen secuencias complementarias a las cadenas opuestas de los ácidos nucleicos diana (como una secuencia de ácido nucleico genómico diana) y se posicionan de forma que los cebadores sean convergentes. Los cebadores son incubados con el ácido nucleico de la plantilla en condiciones que permiten la hibridación selectiva. Los cebadores se pueden proporcionar en forma de doble cadena o cadena única. Si la secuencia o secuencias de ácido nucleico genómico diana están presentes en una muestra, los cebadores se hibridarán para formar un complejo ácido nucleico/cebador. Se añade un exceso de desoxirribonucleótidos trifosfatos (dNTP), junto con una polimerasa de ADN termoestable, p. ej., polimerasa Taq. Si se ha formado el complejo diana/cebador, la polimerasa extenderá el cebador a lo largo de la secuencia genómica diana añadiendo nucleótidos. Tras la polimerización, la cadena recién sintetizada de ADN se disocia de su cadena diana complementaria (plantilla), subiendo la temperatura de la mezcla de la reacción. Cuando posteriormente se baja la temperatura, nuevos cebadores se unirán a cada una de estas dos cadenas de ácido nucleico, y se repite el proceso. Se realizan múltiples ciclos de subida y bajada de temperatura, con una ronda de replicación en cada ciclo, hasta que se produce una cantidad suficiente de producto de la amplificación. Otras variantes de la PCR, como DOP-PCR (Feher *et. al.*, *Diagn. Mol Pathol* 15:43-48, 2006); reacción en cadena de reparación (divulgada en WO 90/01069); métodos de amplificación rápida descritos en la Patente estadounidense nº 6.638.722; amplificación lineal basada en T7 (p. ej., como en Liu *et al.*, *BMC Genomics* 4:19pp 1-11, 2003), también son adecuadas para generar regiones de unión correspondientes a una secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana).

Alternativamente, se pueden utilizar métodos como los sistemas de amplificación basados en la transcripción (TAS, por ejemplo, véase Kwoh *et. al.*, *Proc. Natl Acad. Sci.* 86:1173-7, 1989), o NASBA (amplificación basada en la secuencia de ácido nucleico; Malek *et. al.*, *Methods Mol Biol* 28:253-60, 1994; y la Patente Estadounidense nº 6.025.134) para amplificar porciones de una secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana). En estos métodos, la secuencia promotora de una polimerasa de ARN dependiente de ADN designada se añade al segmento deseado de la secuencia diana y se amplifica en múltiples rondas de transcripción mediante la polimerasa de ARN apropiada. Estas reacciones isotérmicas evitan la necesidad de desnaturalizar las cadenas de ADNc de sus plantillas de ARN, incluyendo RNAsa H para degradar el ARN hibridado con ADN. Otros métodos que utilizan la amplificación isotérmica, incluyendo, por ejemplo, los métodos descritos en la Patente estadounidense nº 6.251.639, también se pueden emplear en el contexto de la presente divulgación.

La amplificación también se puede realizar utilizando la reacción en cadena de ligasa (LCR), divulgada en la Solicitud de Patente europea nº 320.308, o la reacción de detección de ligasa (LDR), divulgada en la Patente estadounidense nº 4.883.750, o la amplificación por reacción en cadena de ligasa de llenado de huecos, como se divulga en la Patente estadounidense nº 5.427.930. En la LCR, se preparan dos pares de sondas, que son complementarias entre sí, y de las secuencias adyacentes de las dos cadenas de la secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana). Cada par se unirá a las cadenas opuestas de la secuencia diana de forma que queden adyacentes. Cada uno de los dos pares de sondas puede unirse después para formar una única unidad, utilizando una ligasa termoestable. Mediante ciclos de temperatura, como en la PCR, las unidades ligadas unidas se disocian de la secuencia diana (p. ej., una secuencia de ácido nucleico genómico diana), por lo que ambas moléculas pueden servir de «secuencias diana» para ligar los pares de sondas sobrantes, proporcionando una amplificación exponencial. La LDR es muy similar a la LCR. En esta variación, se utilizan oligonucleótidos complementarios solamente con una cadena de la secuencia genómica diana, resultando en una amplificación lineal de los productos de la unión, dado que solamente la molécula de ácido nucleico diana original puede servir como plantilla de hibridación. Se utiliza tras una amplificación por PCR de una secuencia de ácido nucleico diana, al objeto de aumentar la señal.

Adicionalmente, se pueden emplear métodos de amplificación isotérmicos para amplificar segmentos de una secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana). Algunos ejemplos de métodos incluyen: amplificación isotérmica libre de transcripción, como la divulgada en la Patente Estadounidense nº 6.033.881; o sistemas isotérmicos basados en la transcripción como 3SR (Self-Sustained Sequence Replication;

1990; Mueller *et. al.*, *Histochem Cell Biol.* 108:4310437, 1997) amplificación con desplazamiento de cadena (como se divulga en las Patentes estadounidenses nº 5.648.211; 5.744.311; 5.648.211; 5.744.311; Walker *et. al.*, *Proc Natl Acad Sci* 89:392-396, 1992; Walker *et. al.*, *Nucleic Acids Res* 20:1691-1696, 1992; Tsurumi *et. al.*, *Biochem Biophys Res Commun.*, 238:33-38, 1997), amplificación isotérmica lazo-mediada (Notomi *et. al.*, *Nucleic Acids Res.* 28: e63i-vii, 2000); amplificación con círculo rodante (Patente estadounidense nº 5.648.245; 5.714.320; 6.218.152; 6.291.187; y Mikawa *et. al.*, *Nucleic Acids Res* 34:e69pp1-9, 2006); amplificación isotérmica (como se divulga en la Publicación de Patente estadounidense nº 2005/0164213); amplificación independiente de una secuencia (SIA) (véase, por ejemplo, la Patente estadounidense nº 5.731.171); y desplazamiento de cadena múltiple (MDA) como se divulga, p. ej., en la Patente estadounidense nº 6.323.009, y modificaciones de los mismos, tal y como se describe, p. ej., en Panelli *et. al.*, *BioTechniques.* 39:174-180, 2005, así como amplificación dependiente de helicasa (HDA). La HDA utiliza una helicasa de ADN para generar plantillas de cadena única para la hibridación de un cebador y la posterior extensión del cebador mediante una polimerasa de ADN en condiciones isotérmicas (Vincent *et. al.*, *EMBOReports* 5:795-800, 2004).

En algunos ejemplos, el método de amplificación utilizado para producir regiones de unión sustancialmente libres de secuencia de ácido nucleico repetitiva, que corresponden a una secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana), se selecciona para generar un producto de la amplificación de la región de unión que sea relativamente uniforme en términos de tamaño y composición. Así, por ejemplo, la PCR es un método de amplificación adecuado para generar regiones de unión que estén libres o sustancialmente libres de secuencia de ácido nucleico repetitiva. Los cebadores se seleccionan basándose en la ubicación y la secuencia de las secuencias de ácido nucleico repetitivas identificadas.

Los cebadores se pueden seleccionar manualmente o con la ayuda de un algoritmo generado por ordenador que optimice la selección del cebador basándose en los parámetros deseados, como la temperatura de anillado, la longitud, el contenido de GC, etc. Hay numerosos programas o algoritmos generados por ordenador disponibles para utilizarlos vía Internet o en un ordenador personal, por ejemplo en la página web indicada. Una lista no exclusiva de estos programas (con sus correspondientes direcciones de Internet) incluye: CODEHOP ([blocks.fhrc.org/codehop.html](http://blocks.fhrc.org/codehop.html)); Gene Fisher ([bibiserv.techfak.uni-bielefeld.de/genefisher/](http://bibiserv.techfak.uni-bielefeld.de/genefisher/)); DoPrimer ([doprimer.interactiva.de/](http://doprimer.interactiva.de/)); Primer3 ([genome.wi.mit.edu/cgi-bin/primer/primer3\\_www.cgi](http://genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi)); Primer Selection ([alces.med.umn.edu/rawprimer.html](http://alces.med.umn.edu/rawprimer.html)); Web Primer ([genome-www2.stanford.edu/cgi-bin/SGD/web-primer](http://genome-www2.stanford.edu/cgi-bin/SGD/web-primer)); PCR Designer ([cedai.genetics.soton.ac.uk/public\\_html/primer.html](http://cedai.genetics.soton.ac.uk/public_html/primer.html)); Piimo Pro .3 4 ([changbioscience.com/primo/piimo.html](http://changbioscience.com/primo/piimo.html)); Primo Degenerate 3 4 ([changbioscience.com/primo/prmod.html](http://changbioscience.com/primo/prmod.html)); PCR Primer Design ([mgh.harvard.edu/servlet/org.mgh.proteome.Primer](http://mgh.harvard.edu/servlet/org.mgh.proteome.Primer)); The Primer Generator ([med.jhu.edu/medcenter/primer/primer.cgi](http://med.jhu.edu/medcenter/primer/primer.cgi)); EPRIMER3 ([bioweb.pasteur.fr/seqanal/interfaces/eprimer3.html](http://bioweb.pasteur.fr/seqanal/interfaces/eprimer3.html)); PRIMO ([bioweb.pasteur.fr/seqanal/interfaces/primo.html](http://bioweb.pasteur.fr/seqanal/interfaces/primo.html)); PrimerQuest ([idtdna.com/biotools/primer\\_quest/primer\\_quest.asp](http://idtdna.com/biotools/primer_quest/primer_quest.asp)); MethPrimer ([itsa.ucsf.edu/~urolab/methprimer/index.html](http://itsa.ucsf.edu/~urolab/methprimer/index.html)); Rawprimer ([alces.med.umn.edu/rawprimer.html](http://alces.med.umn.edu/rawprimer.html)); MEDUSA ([cgr.ki.se/cgr/MEDUSA](http://cgr.ki.se/cgr/MEDUSA)); The Primer Primer Project ([nrcabm.rutgers.edu/bioinformatics/Primer\\_Primer\\_Piject/Primer.html](http://nrcabm.rutgers.edu/bioinformatics/Primer_Primer_Piject/Primer.html)); Oligonucleotides for the PCR ([cit2.fr/bio2/01igo21ib.html](http://cit2.fr/bio2/01igo21ib.html)); GAP ([promoter.ics.uci.edu/Primers/](http://promoter.ics.uci.edu/Primers/)); Oligonucleotides pour la PCR ([cit2.fr/bio2/OligoTM.html](http://cit2.fr/bio2/OligoTM.html)); Oligonucleotide properties calculator ([asic.nwu.edu/biotools/oligoCalc.html](http://asic.nwu.edu/biotools/oligoCalc.html)); Oligonucleotide analyzer ([matuie.com/oligonucleotide.html](http://matuie.com/oligonucleotide.html)); Oligo Tm Determination ([alces.med.umn.edu/iawtm.html](http://alces.med.umn.edu/iawtm.html)); Poland ([biophys.uni-duesseldorf.de/iocal/POLAND/poland.html](http://biophys.uni-duesseldorf.de/iocal/POLAND/poland.html)); PROLIGO ([gensetoligos.com/Calculation/calculation.html](http://gensetoligos.com/Calculation/calculation.html)); PrimerSelect ([dnastai.com](http://dnastai.com)); DNASIS Max ([medprobe.com/no/dnasis.html](http://medprobe.com/no/dnasis.html)); Primer Premier 5 ([premier/biosoft.com/](http://premier/biosoft.com/)); Primer Premier ([piemier/biosoft.com/](http://piemier/biosoft.com/)); NetPrimer ([piemiebiosoft.com/NetPrimer.html](http://piemiebiosoft.com/NetPrimer.html)); Array Designer 2 ([premierbiosoft.com/Vdnamicroarray/dnamicroarray.html](http://premierbiosoft.com/Vdnamicroarray/dnamicroarray.html)); Beacon Designer 2.1 ([premierbiosoft.com/moleculai\\_beacons/taqman\\_moleculai\\_beacons.html](http://premierbiosoft.com/moleculai_beacons/taqman_moleculai_beacons.html)); GenomePRIDE 1 0 ([pride.molgen.mpg.de/genomepride.html](http://pride.molgen.mpg.de/genomepride.html)); Fast PCR ([biocenterhelsinki.fi/bi/bare-l.html/manual.htm](http://biocenterhelsinki.fi/bi/bare-l.html/manual.htm)); OLIGO 6 ([oligo.net/](http://oligo.net/)); Primer Designer 4 ([scied.com/ses\\_pd5.htm](http://scied.com/ses_pd5.htm)); GPRIME ([life.anu.edu.au/moleculai/software/gprimer.htm](http://life.anu.edu.au/moleculai/software/gprimer.htm)); Sarani Gold ([mail.standgenomics.com/products/sarani/](http://mail.standgenomics.com/products/sarani/)); PCR Help ([techne.com/CatMol/pcrhelp.htm](http://techne.com/CatMol/pcrhelp.htm)); Geniama chip Design Software ([asperbio.com/Chip\\_design\\_soft.htm](http://asperbio.com/Chip_design_soft.htm)); Primer Designer ([genamics.com/expression/primer.htm](http://genamics.com/expression/primer.htm)); Primer Premier ([biotechniques.com/fieesamples/itembtn21.html](http://biotechniques.com/fieesamples/itembtn21.html)); y PrimerDesign ([chemie.unimarburg.de/%7becker/pdhome.html](http://chemie.unimarburg.de/%7becker/pdhome.html)).

Por ejemplo, para generar múltiples regiones de unión de una secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana), las regiones de la secuencia libres de secuencia de ácido nucleico repetitiva (u otras no deseadas, como las que producen señal de fondo) son identificadas, por ejemplo manualmente o utilizando un algoritmo por ordenador seleccionado entre los programas indicados anteriormente (p. ej., RepeatMasker, Masker Aid, PRIME). Dentro de una secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) que tiene varios cientos de kilobases, normalmente se identifican numerosas regiones de unión que están sustancial o totalmente libres de secuencias de ácido nucleico repetitivas (u otras no deseadas, como las que producen señal de fondo). A continuación, se diseñan cebadores de secuencia única basándose en la secuencia de (o cercanas a) los extremos de cada región de unión que está sustancial o completamente libre de secuencia de ácido nucleico repetitiva (u otra no deseada, como la que produce señal de fondo). Por ejemplo, se pueden seleccionar cebadores de secuencia única que sean complementarios de las secuencias tan cercanas como sea posible a los extremos de una región de unión que esté sustancial o completamente libre de secuencia de ácido nucleico repetitiva (u otra no deseada, como la que produce señal de fondo), para maximizar la longitud del segmento. En algunos ejemplos, se seleccionan cebadores internos. Para

5 facilitar la amplificación de las regiones de unión individuales, puede ser conveniente seleccionar secuencias de cebador que compartan características de anillado y amplificación (p. ej., la temperatura de fusión o  $T_m$ ), para que se puedan amplificar múltiples regiones de unión en tándem, utilizando los mismos parámetros de reacción. Por ejemplo, se puede utilizar el programa PRIME para seleccionar secuencias de cebador únicas complementarias de las regiones de unión que comparten una  $T_m$  y que puedan ser amplificadas en tándem utilizando parámetros de reacción idénticos. Opcionalmente, al menos uno de los cebadores incluye un 5' fosfato para facilitar la posterior unión enzimática de las regiones de unión amplificadas. Por ejemplo, si se desean moléculas de ácido nucleico FPC que difieran de la secuencia de ácido nucleico diana solamente en el orden (pero no en la orientación, un único cebador (por ejemplo, complementario de la cadena «superior») de cada reacción incorpora un fosfato. Si las regiones de unión están unidas sin fosforilación adicional (p. ej., fosforilación enzimática con una quinasa), la unión resultante procede en una única dirección (por ejemplo, de cabeza a cola). Si se desean moléculas de ácido nucleico FPC que difieran tanto en orden como en orientación, tanto el cebador «superior» como el «inferior» pueden incorporar un grupo fosfato, de forma que la unión pueda producirse de cabeza a cola, de cabeza a cabeza, o de cola a cola.

15 Opcionalmente, los cebadores de la región de unión pueden incluir una secuencia adicional en el 5' extremo del cebador que está diseñada para facilitar la posterior manipulación (p. ej., la unión y/o posterior amplificación) de la región de unión amplificada resultante. Estas secuencias adicionales pueden incluir, por ejemplo, uno o más sitios de reconocimiento de la enzima de restricción, una o más secuencias de cebador «universal» (como los cebadores universales disponibles en el mercado T7 o T3), o una secuencia de nucleótido fijo (como un tracto completo de purina o completo de pirimidina) que puede servir como un sitio de anillado del cebador para las posteriores reacciones de amplificación. Por ejemplo, como se muestra en la FIG. 4, un oligonucleótido completo de purina puede ser añadido al 5' extremo (es decir, el no cebado) de un cebador de secuencia única que sea complementario de la cadena o cadenas 5' y/o 3' (superior y/o inferior) de una porción de ácido nucleico que vaya a ser amplificado (como una región de unión que esté sustancial o completamente libre de secuencia de ácido nucleico repetitiva, de una secuencia diana genómica). Por lo general, el 5' extremo de uno o ambos nucleótidos está fosforilado para facilitar la posterior unión. La amplificación con este cebador o cebadores resulta en un producto que incluye un tracto completo de purina en el terminal de una cadena (y un tracto de pirimidina en la cadena complementaria).

20 Alternativamente, estas secuencias (p. ej., sitios de reconocimiento de la enzima de restricción, cebadores universales, trectos de purina o pirimidina, etc.) se pueden añadir tras la amplificación mediante la unión de ADN duplicado que presenta la región de unión deseada con uno o ambos extremos de la región de unión amplificada. Por ejemplo, un oligonucleótido compuesto exclusivamente de purinas (p. ej., 5'-GAGGAG-3'; SEC ID N°: 357) se puede anillar con su complemento inverso (p. ej., 5'-CTCCTC-3'; SEC ID N°: 358). A menudo, uno o los dos oligonucleótidos están fosforilados en el 5' extremo para facilitar la posterior unión. Tras anillar los dos oligonucleótidos juntos para formar una molécula de oligonucleótido de doble cadena, el enlace o adaptador del oligonucleótido duplicado se une (típicamente en una reacción química o enzimática) con uno o los dos extremos de la región de unión. Este proceso resulta en una región de unión que incluye una secuencia fija en uno o ambos extremos (por ejemplo, dependiendo de si uno o ambos extremos del producto de amplificación incluyen un grupo fosforilo reactivo). Cuando se unen, las regiones de unión con estas secuencias fijadas en uno o ambos extremos producen un ácido nucleico lineal con tramos de purina/pirimidina de intervención que pueden servir, por ejemplo, como sitios de unión del cebador (FIG. 5). Opcionalmente (como se muestra en el ejemplo descrito anteriormente, a título meramente enunciativo) o alternativamente, el enlace/adaptador puede incluir un sitio de reconocimiento de la enzima de restricción.

30 A pesar de que la amplificación es un ejemplo de método para proporcionar regiones de unión de una secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) y puede aumentar (p. ej., aumentar geométricamente) la cantidad de ácidos nucleicos de la región de unión, también se pueden utilizar otros métodos para proporcionar regiones de unión que están sustancial o completamente libres de secuencia de ácido nucleico repetitiva (u otras no deseadas, como las que producen señal de fondo), en el contexto de la producción de las sondas divulgadas. Por ejemplo, una molécula de ácido nucleico diana puede ser digerida con uno o más endonucleasas de restricción (enzimas de restricción). Tras la digestión con una o más enzimas de restricción u otras endonucleasas, las regiones de unión específicas (p. ej., predeterminadas o seleccionadas) que están sustancial o completamente libres de secuencia de ácido nucleico repetitiva (u otras no deseadas, como las que producen señal de fondo) pueden ser recuperadas. Los métodos para digerir ácidos nucleicos son bien conocidos en el campo y hay suficientes ejemplos de métodos para guiar a las personas con conocimientos en el campo se encuentran, p. ej., en Sambrook y Ausubel.

55 Los elementos repetitivos (u otros no deseados, p. ej., los que producen señal de fondo) de una secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) se pueden eliminar de diferentes maneras. Por ejemplo, las moléculas de ácido nucleico no deseadas pueden ser fragmentadas en pequeñas porciones con una endonucleasa de restricción que se corta frecuentemente dentro de la secuencia de ácido nucleico no deseada, pero no (o menos frecuentemente) en la secuencia de ácido nucleico única deseada. La mezcla resultante de fragmentos puede ser fraccionada en gel de agarosa de acuerdo con procedimientos establecidos y solamente los elementos deseados de mayor tamaño seleccionados como fragmentos para la unión a las moléculas de ácido nucleico FPC (que pueden servir como plantillas de amplificación). Los elementos no deseados (p. ej., los que

5 producen señal de fondo) pueden ser eliminados mediante clivaje enzimático o físico (corte mecánico, p. ej.,  
sonicación) del ADN hasta convertirlo en porciones, desnaturalizando el ADN duplicado en cadenas únicas mediante  
calentamiento (por ejemplo, a 100°C durante 10 minutos) o tratamiento con álcali (0,1 de volumen 3.0 M NaOH, 10-  
30 minutos a temperatura ambiente), y reanillando en condiciones controladas. Dado que la secuencia de ácido  
nucleico repetitiva (u otra no deseada, como la que produce señal de fondo) se reanilla más rápidamente que la  
10 secuencia de ácido nucleico única, la secuencia de ácido nucleico repetitiva (u otra no deseada, p. ej., que produce  
señal de fondo) que se anilla más rápidamente (o fracción CDT-1™) puede ser separada de la secuencia de ácido  
nucleico única poniendo en contacto la mezcla parcialmente reanillada con hidroxiapatita (p. ej., en una columna). La  
secuencia de ácido nucleico repetitiva (u otra no deseada, como la que produce señal de fondo) de doble cadena  
(que se reanilla rápidamente) es retenida por la columna, mientras que las secuencias deseadas de cadena única  
permanecen sin unirse.

Las regiones de unión que están sustancial o completamente libres de secuencia de ácido nucleico no deseada (p.  
ej., las que producen señal de fondo) también se pueden preparar con una secuencia de ácido nucleico diana  
fragmentada (p. ej., una secuencia de ácido nucleico genómico diana) mediante hibridación sustractiva. En este  
15 método, la secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) es  
desnaturalizada (si se encuentra en forma de doble cadena) y las moléculas de ácido nucleico diana de cadena  
única resultantes hibridadas con moléculas de ácido nucleico no deseadas (p. ej., repetitivas) etiquetadas (p. ej.,  
ADN C<sub>0</sub>t-1™, o secuencia de ácido nucleico repetitiva específica, como ALU, LINE, etc.). El ácido nucleico  
etiquetado de doble cadena hibridado resultante se puede eliminar después, utilizando un socio de unión específico  
20 que se una a la etiqueta. Más detalles se proporcionan, p. ej., en las Patentes estadounidenses n° 5.643.761,  
6.569.621, y en Davison *et. al. Am J Pathol* 153:1401-1409, 1998. Por ejemplo, la secuencia de ácido nucleico  
repetitiva (u otra no deseada, como la que produce señal de fondo) puede ser retirada de una muestra de  
fragmentos de la molécula de ácido nucleico diana, desnaturalizando con calor las moléculas de ácido nucleico  
diana y mezclando con un exceso de molécula de ácido nucleico sustractora (como ADN CDT-1™) etiquetada con  
25 un hapteno como biotina, digoxigenina o DNP. Se permite que las moléculas de ácido nucleico sustractoras diana  
etiquetadas se anillen, período durante el cual las moléculas de ácido nucleico sustractoras se hibridan con los  
elementos de ácido nucleico repetitivos (u otros no deseados, como los que producen señal de fondo) en la  
secuencia de ácido nucleico diana. Las moléculas de ácido nucleico repetitivas anilladas (u otras no deseadas,  
como las que producen señal de fondo) pueden entonces ser extraídas, poniendo en contacto la mezcla con una  
30 matriz (como una columna, perlas magnéticas, etc.) conjugada con un socio de unión específico para el hapteno  
concreto (p. ej., avidina, estreptavidina para biotina, anticuerpos específicos para digoxigenina o DNP). Las  
moléculas de ácido nucleico repetitivas (u otras no deseadas, como las que producen señal de fondo) se retienen en  
la matriz, mientras que las regiones de unión de ácido nucleico al menos sustancialmente libres de ácido nucleico no  
deseado (p. ej., de repetición mermada) permanecen en la solución y pueden ser recuperadas para su posterior  
35 procesamiento.

Cualquiera de estos u otros métodos para la producción de regiones de unión de una molécula de ácido nucleico  
diana (p. ej., una molécula de ácido nucleico genómico diana) sustancialmente (o completamente) libres de  
secuencias no deseadas (p. ej., secuencias de ácido nucleico repetitivas) se pueden utilizar de forma aislada o en  
combinación para preparar regiones de unión para la producción de moléculas de ácido nucleico FPC. Por ejemplo,  
40 otros ejemplos de métodos se proporcionan en la Publicación de la Solicitud estadounidense n° 2006/0160116, así  
como en la Patente n° 6.280.929. Las personas con conocimientos en el campo apreciarán que si un  
enlace/adaptador u otra secuencia de nucleótidos especificada se desea en el extremo de las regiones de unión, la  
secuencia (como un dúplex completo de purina/pirimidina) se puede unir con uno o con los dos extremos de la  
región de unión, como se ha descrito anteriormente.

#### 45 **Unión de regiones de unión para formar moléculas de ácido nucleico FPC**

Las moléculas de ácido nucleico FPC (que pueden servir como plantillas para una posterior amplificación) se  
producen uniendo regiones de unión individuales (por ejemplo, presentes en una población de regiones de unión)  
que se corresponden con la secuencia de ácido nucleico diana seleccionada (p. ej., una secuencia genómica diana).  
El término «unir» (o «unido» o «uniendo») se utiliza en el presente para indicar que dos secuencias o segmentos de  
50 ácido nucleico de la región de unión se unen para formar una molécula de ácido nucleico FPC lineal contigua. No  
obstante, puede haber enlaces presentes entre los segmentos de la región de unión. Típicamente, dos regiones de  
unión se unen enzimáticamente mediante una ligasa en una reacción de unión. Sin embargo, dos regiones de unión  
también se pueden unir químicamente (p. ej., incorporando los nucleótidos modificados apropiados, tal y como se  
describe en Dolinnaya *et. al., Nucleic Acids Res* 16:3721-38, 1988; Mattes y Seitz, *Chem. Commun.* 2050-2051,  
2001; Mattes y Seitz, *Agnew. Chem Int.* 40:3178-81, 2001; Ficht *et al., J Am. Chem. Soc* 126:9970-81, 2004.  
55 Alternativamente, dos regiones de unión diferentes se pueden unir en una reacción de amplificación o utilizando una  
recombinasa. Por ejemplo, las regiones de unión que están sustancialmente libres de secuencias de ácido nucleico  
repetitivas (u otras no deseables, como las que producen señal de fondo) se pueden incubar en una mezcla que  
contiene una ligasa de ADN codificada por una bacteria o bacteriófago, como la ligasa de ADN T4 y ATP, en un  
60 tampón adecuado, e incubadas a una temperatura adecuada y durante un período de tiempo suficiente para que se  
produzca la unión. El tiempo y la temperatura se optimizan típicamente dependiendo de si las regiones de unión a  
unir tienen el extremo romo o poseen proyecciones compatibles (por ejemplo, producidas por digestión con una

endonucleasa de restricción o añadiendo oligonucleótidos de adaptador/enlace proyectado). Por tanto, las reacciones de unión solapadas y de extremo romo se pueden utilizar para producir moléculas de ácido nucleico FPC. Los detalles relativos a la unión enzimática suficientes para guiar a una persona con conocimientos en el campo se pueden encontrar, p. ej., en Sambrook y Ausubel.

5 La unión es dependiente de la presencia de los grupos 5' fosforilo y 3' hidroxilo en el terminal adyacente de las moléculas de la región de unión a unir. Por tanto, la orientación de la unión se puede predeterminar mezclando regiones de unión con solo un grupo 5' fosforilo. En la fabricación de las sondas divulgadas en el presente, las regiones de unión se pueden unir de una manera que sea diferente en orden, diferente en orientación, o tanto en orden como en orientación, a la secuencia genómica diana a la que corresponden los segmentos. Mezclando regiones de unión que estén todas fosforiladas en el 5' terminal de la cadena «superior», la orientación se puede fijar en la misma orientación que la secuencia de ácido nucleico genómico diana. Esto se puede conseguir incorporando un grupo 5' fosforilo al cebador antes de la amplificación. Si se desean moléculas de ácido nucleico FPC que difieran tanto en orden como en orientación de la secuencia genómica diana, tanto los cebadores de amplificación de cadena 5' (superior) como 3' (inferior) pueden ser fosforilados. Alternativamente, las regiones de unión amplificadas pueden ser fosforiladas enzimáticamente (p. ej., con una quinasa) o químicamente, antes de mezclarlas con la ligasa. De un modo similar, los enlaces pueden ser fosforilados antes de añadirlos a los extremos de las regiones de unión deseadas. Típicamente, las regiones de unión aisladas tras la restricción u otra digestión de endonucleasa o corte mecánico poseen grupos fosforilo y no precisan ningún procesamiento adicional, aunque la fosforilación se puede realizar para garantizar una unión completa y eficiente.

## 20 **Amplificación de las moléculas de ácido nucleico FPC**

Las moléculas de ácido nucleico FPC resultantes se pueden utilizar como una sonda (por ejemplo etiquetando las moléculas de ácido nucleico FPC) y pueden servir como plantillas para la amplificación. Por ejemplo, las plantillas de moléculas de ácido nucleico FPC se pueden amplificar en una o más reacciones (por ejemplo, en una serie de reacciones secuenciales) para producir una población de amplímeros de ácido nucleico FPC que sean adecuados como sondas. El número de reacciones viene típicamente determinado por la cantidad de sonda deseada; así, para una sola aplicación, una reacción de amplificación puede ser suficiente. Por el contrario, cuando se van a someter a ensayo múltiples muestras, se van a realizar múltiples ensayos o los ensayos se van a realizar en un proceso automatizado o semiautomatizado, puede resultar recomendable realizar varias reacciones de amplificación secuencial para aumentar el rendimiento de los amplímeros de las moléculas de ácido nucleico FPC.

30 Cualquiera de los procedimientos de amplificación anteriormente descritos o conocidos en el campo puede ser adaptado para ser utilizado por una persona con conocimientos en el campo para la producción de moléculas de ácido nucleico FPC amplificado de las moléculas de ácido nucleico FPC de la plantilla. En determinados ejemplos, se emplean reacciones de amplificación isotérmicas. Como se ha comentado anteriormente, las reacciones de amplificación isotérmicas incluyen: amplificación isotérmica libre de transcripción, como se ha divulgado en la Patente Estadounidense nº 6.033.881; o sistemas isotérmicos basados en la transcripción como 3SR (Self-Sustained Sequence Replication; 1990; Mueller *et. al.*, *Histochem Cell Biol.* 108:4310437, 1997) amplificación con desplazamiento de cadena (como se divulga en las Patentes estadounidenses nº 5.648.211; 5.744.311; Walker *et al.*, *Proc. Natl Acad.*, Set 89:392-396, 1992; Wallcei *et.al. Nucleic Acids Res* 20:1691-1696, 1992; Tsurumi *et al.*, *Biochem Biophys Res Commun.*, 238:33-38, 1997), amplificación isotérmica lazo-mediada (Notomi *et. al.*, *Nucleic Acids Res* 28:e63i-vii, 2000); amplificación con círculo rodante (Patentes estadounidenses nº 5.648.245; 5.714.320; 6.218.152; 6.291.187; y Mikawa *et. al.*, *Nucleic Acids Res* 34:e69ppl-9, 2006); amplificación isotérmica (como se divulga en la Publicación de Patente estadounidense nº 2005/0164213); y desplazamiento de cadena múltiple (MDA) como se divulga, p. ej., en la Patente estadounidense nº 6.323.009, y modificaciones de los mismos, tal y como se describe, p. ej., en Panelli *et. al.*, *BioTechniques* 39:174-180, 2005, así como amplificación dependiente de helicasa (HDA). La HDA utiliza una helicasa de ADN para generar plantillas de cadena única para la hibridación de un cebador y la posterior extensión del cebador mediante una polimerasa de ADN en condiciones isotérmicas (Vincent *et. al.*, *EMBO Reports* 5:795-800, 2004).

Por ejemplo, las reacciones de amplificación isotérmicas se pueden emplear utilizando cebadores de hexámeros al azar. Los hexámeros al azar se unen a su secuencia complementaria dispersada por toda la molécula de ácido nucleico FPC de la plantilla. El hexámero anillado se extiende posteriormente para formar un amplímero de ácido nucleico FPC complementario de la molécula de ácido nucleico FPC de la plantilla. Este proceso se repite en toda la longitud de la molécula de ácido nucleico FPC de la plantilla en las numerosas cadenas del producto, resultando en la amplificación de las moléculas de ácido nucleico FPC de la plantilla. La amplificación mediante este método genera el solapamiento de moléculas de amplímeros de ácido nucleico FPC que se extienden por toda la longitud de la molécula de ácido nucleico FPC de la plantilla. Dado que el tamaño de la molécula de ácido nucleico FPC de la plantilla puede exceder el tamaño generalmente razonable o practicable para métodos como la PCR que confían en cebadores específicos y dado que el orden y la orientación de los segmentos son variables (y típicamente impredecibles sin un análisis laborioso), los métodos de amplificación isotérmica, que son altamente procesivos, pueden ofrecer un mejor resultado. En determinadas realizaciones, por ejemplo, cuando se ha incorporado una secuencia fija a la región o regiones de unión (por ejemplo, añadiendo un enlace/adaptador o mediante la

incorporación de un enlace al cebador utilizado para amplificar la región de unión), los cebadores que se hibridan con la secuencia fija (o su complemento) también se pueden utilizar.

### **Etiquetado de las moléculas de ácido nucleico**

5 En algunos ejemplos, las moléculas de ácido nucleico FPC están etiquetadas. El etiquetado de las moléculas de ácido nucleico FPC se puede realizar antes, durante o después de la amplificación de las plantillas de moléculas de ácido nucleico FPC. Por ejemplo, las moléculas de ácido nucleico FPC se pueden etiquetar «antes» de la síntesis, incorporando uno o más nucleótidos etiquetados en el cebador utilizado para la amplificación de las moléculas de ácido nucleico FPC. Los nucleótidos etiquetados también se pueden incorporar durante la amplificación, incluyendo uno o más nucleótidos etiquetados en la mezcla de reacción de la amplificación (p. ej., amplificación de desplazamiento múltiple). En algunos ejemplos, una sonda es etiquetada incorporando uno o más dNTP etiquetados, utilizando una reacción enzimática (polimerización), tras la amplificación de las moléculas de ácido nucleico FPC. Por ejemplo, la sonda de ácido nucleico amplificada puede ser etiquetada mediante traslación de muescas (utilizando, por ejemplo, Bio-11-dUTP, 2,4-dinitrofenol, digoxina, *etc.*), mediante extensión del cebador al azar con (p. ej., cola de extremo 3').

15 Se pueden utilizar procedimientos de etiquetado enzimático para producir sondas etiquetadas. También se pueden emplear procedimientos de etiquetado químico. Hay numerosos reactivos (incluyendo hapteno, fluoróforo y otros nucleótidos etiquetados) y otros juegos disponibles en el mercado para el etiquetado enzimático de los ácidos nucleicos, incluyendo los cebadores y las moléculas de ácido nucleico FPC amplificadas (y, si se desea, las moléculas de ácido nucleico FPC de la plantilla). Como resultará evidente para una persona con conocimientos en el campo, cualquiera de las etiquetas y procedimientos de detección divulgados en las secciones anteriores son aplicables en el contexto del etiquetado de una sonda, p. ej., para el uso en reacciones de hibridación *in situ*. Por ejemplo, el sistema de etiquetado de ADN Amersham MULTIPRIME®, diversos reactivos específicos y juegos disponibles de Invitrogen Detection Technologies (Molecular Probes, Eugene, OR) o cualquier otro reactivo o juego similar se pueden utilizar para etiquetar los ácidos nucleicos divulgados en el presente. En determinados ejemplos, las moléculas de ácido nucleico FPC (incluyendo amplímeros de la molécula de ácido nucleico FPC) pueden ser directa o indirectamente etiquetadas con un hapteno, un ligando, una fracción fluorescente (p. ej., un fluoróforo o un nanocrystal semiconductor), una fracción cromogénica o un radioisótopo. Por ejemplo, para el etiquetado indirecto, la etiqueta se puede añadir a las moléculas de ácido nucleico FPC a través de un enlace (p. ej., PEG o biotina).

20 Otros métodos adicionales que se pueden utilizar para etiquetar moléculas de ácido nucleico FPC se proporcionan en la Publicación de Solicitud estadounidense nº 2005/0158770.

### **MÉTODOS PARA UTILIZAR LAS SONDAS**

Las sondas divulgadas, que incluyen una pluralidad de moléculas de ácido nucleico FPC (o fragmentos de las mismas), se pueden utilizar para la detección de ácido nucleico, como los procedimientos de hibridación *in situ* (p. ej., hibridación fluorescente *in situ* (FISH), hibridación cromogénica *in situ* (CISH) e hibridación con plata *in situ* (SISH)). La hibridación entre las moléculas de ácido nucleico complementarias viene mediada por el enlace de hidrógeno, que incluye el enlace de hidrógeno Watson-Crick, Hoogsteen o Hoogsteen invertido entre unidades de nucleótidos complementarias. Por ejemplo, la adenina y timina son nucleobases complementarias que se emparejan a través de la formación de enlaces de hidrógeno. Si una unidad de nucleótido en una determinada posición de una sonda de la presente divulgación se puede unir mediante un enlace de hidrógeno con una unidad de nucleótido en la misma posición de una molécula de ADN o ARN (p. ej., una secuencia de ácido nucleico diana), entonces los oligonucleótidos son complementarios entre sí en esa posición. La sonda y el ADN o ARN son complementarios entre sí cuando un número suficiente de posiciones correspondientes de cada molécula está ocupado por unidades de nucleótidos que se pueden unir mediante un enlace de hidrógeno entre sí y, por tanto, producir una unión detectable. Una sonda no tiene que ser 100% complementaria con su secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) para ser específicamente hibridable. No obstante, se necesita una complementariedad suficiente para que la sonda se una, se duplique o se hibride solo o prácticamente solo con una secuencia de ácido nucleico diana cuando esa secuencia está presente en una mezcla compleja (p. ej., ARN o ADN celular total).

50 La hibridación *in situ* implica poner en contacto una muestra que contiene una secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana) en el contexto de un preparado cromosómico en metafase o interfase (como una muestra celular o tisular sobre un portaobjetos) con una sonda etiquetada específicamente hibridable o específica para la secuencia de ácido nucleico diana (p. ej., una secuencia de ácido nucleico genómico diana). Los portaobjetos son opcionalmente pretratados, por ejemplo para eliminar la parafina u otros materiales que pueden interferir en una hibridación uniforme. La muestra cromosómica y la sonda son tratados, por ejemplo mediante calentamiento para desnaturalizar los ácidos nucleicos de doble cadena. La sonda (formulada en un tampón de hibridación adecuado) y la muestra se combinan, en unas condiciones y durante un tiempo suficiente como para permitir que se produzca la hibridación (típicamente para alcanzar el equilibrio). El preparado cromosómico se lava para eliminar la sonda sobrante y la detección del etiquetado específico de la diana cromosómica se realiza utilizando técnicas estándar.

Por ejemplo, una sonda biotinilada se puede detectar utilizando avidina etiquetada con fluoresceína o fosfatasa avidin-alcalina. Para la detección del fluorocromo, el fluorocromo se puede detectar directamente o las muestras se pueden incubar, por ejemplo, con DCS avidina conjugada con isotiocianato de fluoresceína (FITC). La amplificación de la señal FITC se puede efectuar, si es necesario, mediante incubación con anticuerpo de cabra anti-avidina D biotin-conjugado, lavado y una segunda incubación con avidina FITC-conjugado. Para la detección por actividad de la enzima, las muestras pueden ser incubadas, por ejemplo, con estreptavidina, lavadas, incubadas con fosfatasa alcalina biotin-conjugada, lavada de nuevo y preequilibrada (p. ej., en tampón de fosfatasa alcalina (AP)). La reacción de la enzima se puede realizar, por ejemplo, en un tampón de AP, que contiene tetrazolio nitroazul y 5' bromo-4-cloro-3-indoilo fosfato y detenida mediante incubación en 2 X SSC. Para una descripción general de los procedimientos de hibridación *in situ*, véase, p. ej., la Patente estadounidense nº 4.888.278.

Numerosos procedimientos para la hibridación fluorescente *in situ* (FISH), la hibridación cromogénica *in situ* (CISH) y la hibridación con plata *in situ* (SISH) son conocidos en el campo. Por ejemplo, los procedimientos para realizar la FISH se describen en las Patentes estadounidenses nº 5.447.841, 5.472.842, 5.427.932, y por ejemplo, en Pinkel *et al.*, *Proc. Natl. Acad. Set* 83:2934-2938, 1986; Pinkel *et al.*, *Proc Natl Acad. Set* 85:91.38-9142, 1988, y Lichter *et al.*, *Proc Natl. Acad. Set* 85:9664-9668, 1988. La CISH se describe, p. ej., en Tanner *et al.*, *Am. J Pathol* 157:1467-1472, 2000, y en la Patente estadounidense nº 6.942.970. Los métodos de detección adicionales se proporcionan en la Patente estadounidense nº 6.280.929. Algunos ejemplos de procedimientos para la detección de virus mediante hibridación *in situ* se pueden encontrar en Poddighe *et al.*, *J Clin. Pathol* 49:M340-M344, 3996.

Se pueden utilizar numerosos reactivos y programas de detección conjuntamente con los procedimientos FISH, CISH y SISH para mejorar la sensibilidad, la resolución u otras propiedades deseables. Como se ha comentado anteriormente, las sondas etiquetadas con fluoróforos (incluyendo tintes fluorescentes y QUANTUM DOTS®) se pueden detectar ópticamente de forma directa, cuando se realiza la FISH. Alternativamente, la sonda puede ser etiquetada con una molécula no fluorescente, como un hapteno (tales como los siguientes ejemplos, dados a título meramente enunciativo: biotina, digoxigenina, DNP, y varios oxazoles, pinazoles, tiazoles, nitroarilos, benzofuranos, triterpenos, ureas, tioureas, rotenonas, cumarina, compuestos a base de cumarina, podofilotoxina, compuestos a base de podofilotoxina, y combinaciones de los mismos), un ligando u otra fracción indirectamente detectable. Las sondas etiquetadas con estas moléculas no fluorescentes (y las secuencias de ácido nucleico diana a las que se unen) pueden entonces ser detectadas poniendo en contacto la muestra (p. ej., la muestra de célula o tejido a la que está unida la sonda) con un reactivo de detección etiquetado, como un anticuerpo (o receptor, u otros socio de unión específico) específico para el hapteno o ligando seleccionado. El reactivo de detección puede estar etiquetado con un fluoróforo (p.ej., QUANTUM DOT®) o con otra fracción indirectamente detectable, o se puede poner en contacto con uno o más agentes de unión específicos adicionales (p. ej., anticuerpos secundarios o específicos), que puede, a su vez, estar etiquetados con un fluoróforo. Opcionalmente, la etiqueta detectable se añade directamente al anticuerpo, receptor (u otro agente de unión específico). Alternativamente, la etiqueta detectable se añade al agente de unión a través de un enlace, como un enlace de tiol hidracida, un enlace de glicol de polietileno, o cualquier otra fracción de unión flexible con reactividades comparables.

Por ejemplo, un agente de unión específico, como un anticuerpo, un receptor (u otro anti-ligando), avidina, o similares, puede ser modificado covalentemente con un fluoróforo (u otra etiqueta) a través de un enlace de polialquilénoglicol heterobifuncional, como un enlace de polietilénoglicol heterobifuncional (PEG). Un enlace heterobifuncional combina dos grupos reactivos diferentes seleccionados, p. ej., de un grupo carbonil-reactivo, un grupo amino-reactivo, un grupo tiol-reactivo y un grupo foto-reactivo, uniéndose el primero a la etiqueta y el segundo al agente de unión específico.

En otros ejemplos, la sonda, o el agente de unión específico (como un anticuerpo, p. ej., un anticuerpo primario, receptor u otro agente de unión) está etiquetado con una enzima que es capaz de convertir una composición fluorogénica o cromogénica en una señal detectable fluorescente, coloreada o de otro modo (p. ej., como en la deposición de partículas metálicas detectables de la SISH). Como se ha comentado anteriormente, la enzima se puede añadir directa o indirectamente a través de un enlace al reactivo de detección o a la sonda correspondiente. Algunos ejemplos de reactivos adecuados (p. ej., reactivos de unión) y procesos químicos (p. ej., procesos químicos de unión y enlace) se describen en la Publicación de las Solicitudes de Patentes estadounidenses nº 2006/0246524; 2006/0246523, y en la Solicitud de Patente Provisional estadounidense nº 60/739.794. [0180] Las personas con conocimientos en el campo apreciarán que seleccionando adecuadamente los pares de agentes de unión específicos para la sonda etiquetada, se pueden producir múltiples programas de detección para facilitar la detección de múltiples secuencias de ácido nucleico diana (p. ej., secuencias de ácido nucleico genómico diana) en un único ensayo (p. ej., en una única muestra celular o de tejido o en más de una muestra celular o de tejido). Por ejemplo, una primera sonda que se corresponde con una primera secuencia diana puede ser etiquetada con un primer hapteno, como biotina, mientras que una segunda sonda que se corresponde con una segunda secuencia diana puede estar etiquetada con un segundo hapteno, como DNP. Tras la exposición de la muestra a las sondas, las sondas unidas se pueden detectar poniendo en contacto la muestra con un primer agente de unión específico (en este caso avidina etiquetada con un primer fluoróforo, por ejemplo, un primer QUANTUM DOT® espectralmente distinto, p. ej., que emite a 585 nm) y un segundo agente de unión específico (en este caso un anticuerpo anti-DNP, o un fragmento de anticuerpo, etiquetado con un segundo fluoróforo (por ejemplo, un segundo QUANTUM DOT® espectralmente distinto, p. ej., que emite a 705 nm). Se pueden añadir pares de agentes de unión/sondas

adicionales para el programa de detección múltiple, utilizando otros fluoróforos espectralmente distintos. Numerosas variaciones de directo, e indirecto (un paso, dos pasos más) se pueden prever, todas ellas adecuadas en el contexto de los ensayos y las sondas divulgadas.

- 5 Se pueden encontrar más detalles relativos a determinados métodos de detección, p. ej., como los utilizados en los procedimientos CISH y SISH, en Bourne, *The Handbook of Immunoperoxidase Staining Methods*, publicado por Dako Corporation, Santa Barbara, CA.

### **JUEGOS**

- 10 Los juegos que incluyen al menos un ácido nucleico divulgado en el presente (como una molécula de ácido nucleico FPC o una población de moléculas de ácido nucleico FPC) son también una característica de esta divulgación. Por ejemplo, los juegos para procedimientos de hibridación *in situ*, como FISH, CISH y/o SISH, incluyen al menos una sonda como la descrita en el presente y/o un ácido nucleico que puede servir como plantilla para realizar esta sonda (p. ej., una plantilla de molécula de ácido nucleico FPC). Por consiguiente, los juegos pueden incluir una o más moléculas de ácido nucleico FPC de la plantilla; una o más moléculas de ácido nucleico FPC amplificadas; o una o más sondas etiquetadas que incluyen moléculas de ácido nucleico FPC amplificadas etiquetadas.

- 15 Los juegos también pueden incluir uno o más reactivos para realizar un ensayo de hibridación *in situ* o para producir una sonda. Por ejemplo, un juego puede incluir al menos una molécula de ácido nucleico FPC (o población de estas moléculas), junto con uno o más tampones, dNTP etiquetados, una enzima de etiquetado (como una polimerasa), cebadores, agua libre de nucleasa e instrucciones para producir una sonda etiquetada.

- 20 En un ejemplo, el juego incluye una o más moléculas de ácido nucleico FPC (no etiquetadas o etiquetadas), junto con tampones y otros reactivos para realizar la hibridación *in situ*. Por ejemplo, si una o más moléculas de ácido nucleico FPC amplificadas no etiquetadas se incluyen en el juego, también se pueden incluir reactivos de etiquetado, junto con agentes de detección específicos y otros reactivos para realizar un ensayo de hibridación *in situ*, como un tampón de pretratamiento de parafina, proteasa(s) y un tampón de proteasa, un tampón de prehibridación, un tampón de hibridación, un tampón de lavado, contracoloración(es), medio de montaje o combinaciones de los mismos. El juego puede incluir también opcionalmente portaobjetos de control para valorar la hibridación y la señal de la sonda.

- 30 En determinados ejemplos, los juegos incluyen avidina, anticuerpos y/o receptores (u otros anti-ligandos). Opcionalmente, uno o más de los agentes de detección (incluyendo un agente de detección primario y, opcionalmente, secundario, terciario o reactivos de detección adicionales) están etiquetados, por ejemplo, con un hapteno o fluoróforo (como tinte fluorescente o QUANTUM DOT®). En algunos ejemplos, los reactivos de detección están etiquetados con diferentes fracciones detectables (por ejemplo, diferentes tintes fluorescentes, QUANTUM DOT® espectralmente diferenciables, diferentes haptenos, etc.). Por ejemplo, un juego puede incluir dos o más regiones de unión diferentes que están sustancial o completamente libres de secuencia de ácido nucleico repetitiva (u otra no deseada, como la que produce señal de fondo), sondas (p. ej., una molécula de ácido nucleico FPC amplificada o mezclas de moléculas de ácido nucleico FPC amplificadas) que se corresponden y son capaces de hibridarse con diferentes secuencias de ácido nucleico diana (por ejemplo, cualquiera de las secuencias diana divulgadas en el presente). La primera sonda puede ser etiquetada con una primera etiqueta detectable (p. ej., hapteno, fluoróforo, etc.); la segunda sonda puede estar etiquetada con una segunda etiqueta detectable; y cualquier sonda adicional (p. ej., tercera, cuarta, quinta, etc.) puede estar etiquetada con etiquetas detectables adicionales. La primera, segunda y cualquier sonda posterior pueden estar etiquetadas con diferentes etiquetas detectables, a pesar de que hay otros programas de detección posibles. Si la sonda o sondas están etiquetadas con etiquetas indirectamente detectables, como haptenos, los juegos pueden incluir agentes de detección (como avidina etiquetada, anticuerpos u otros agentes de unión específicos) para algunas o la totalidad de las sondas. En una realización, el juego incluye sondas y reactivos de detección adecuados para múltiples ISH.

- 45 En un ejemplo, el juego también incluye un conjugado de anticuerpo, como un conjugado de anticuerpo para una etiqueta (p. ej., una enzima un fluoróforo, una nanopartícula fluorescente). En algunos ejemplos, el anticuerpo se conjuga con la etiqueta a través de un enlace, como PEG, 6X-His, estreptavidina y GST.

### **EJEMPLOS**

#### **EJEMPLO 1: Producción de una sonda de HER2 sustancialmente libre de repetición**

- 50 Este ejemplo describe métodos utilizados para generar una sonda que incluye moléculas de ácido nucleico FPC heterogéneas de una secuencia diana del genoma humano, que incluye el gen HER2. Se pueden emplear métodos similares para generar una sonda que incluye moléculas de ácido nucleico FPC heterogéneas de otras secuencias de ácido nucleico diana, utilizando los cebadores adecuados.

- 55 Se obtuvieron tres clones de cromosomas artificiales bacterianos (BAC) que contenían secuencias de ácido nucleico del cromosoma 17 humano que cruzan por el gen HER2 de Invitrogen (Carlsbad, CA). Los tres clones se designaron

CTD-2019C10 (C10; nº de acceso GENBANK™ AC040933), RP11-94L15 (94L; nº de acceso GENBANK™ AC079199), y RPI1-.387H17 (H17; nº de acceso GENBANK™ AC090844). La secuencia de codificación HER2 está contenida dentro de 94L, donde se transcribe de izquierda a derecha (es decir, de C10 hacia H17). El DNA BAC se realizó con cultivos de cuatro litros y se purificó utilizando juegos de preparación de estructura grande de Qiagen (Valencia, CA), de acuerdo con las instrucciones del fabricante.

Las secuencias de polinucleótido de estos tres clones BAC se obtuvieron de GENBANKIM y las secuencias repetitivas se identificaron utilizando el programa RepeatMasker. El programa RepeatMasker sustituye todos los nucleótidos en elementos repetitivos por «N». Tras la identificación de elementos repetitivos, todos los N menos uno se eliminaron en cada elemento repetitivo. Los cebadores de amplificación se seleccionaron entonces para amplificar secuencias no repetidas entre cada N, utilizando el programa de ordenador OLIGO™. Los cebadores de los oligonucleótidos se seleccionaron para una Tm lo más cercana posible a 69°C y para una posición lo más cercana posible a cada extremo del segmento de la región de unión única, para maximizar el tamaño de los productos PCR. Los cebadores para la amplificación de las regiones de unión de la región del cromosoma 17, incluyendo el gen HER2, que estaban sustancialmente libres de segmentos de ácido nucleico repetitivo se muestran en la Tabla 1.

Tabla 1. Cebadores utilizados para generar regiones de unión para una sonda HER2.

SEQ ID NO:	Forward Sequence	SEQ ID NO:	Reverse Sequence	Product B.P.
<b>BAC C10 End to 120K</b>				
1	AAAIGAI IAGCAAGGCCAGAAGTC	179	GGGAAAAAATCAGAAAACCTACACI	
2	AACIGGACAAGCTCIIIGGGA	180	GAACCGCCICIGICIIIGAIACC	797
3	CCAGCICCAAAAAIGAAAAAG	181	IGIGCAICAGCIAICCAACAA	1368
4	AACCAGGCAGGCAACIIAITA	182	CCACGICCAGGCIGIIIIAIII	1334
5	IICAAIGACCAGACICCIIGC	183	CCAAGGCACIGIIIIIIIGAAG	1204
6	IAAIGCAIGGIAGGACCGAAI	184	IAI IAGGGIGGGIGGGICIIIC	1489
7	ATIAGCCAGCAIIIIIGIGACC	185	CAAGCIGACAGAAIGGAGAGG	3016
8	GAACCAACAGGAIGIGCGAIA	186	GAI IATGCAGIAACCACAAGG	737
9	IICAAACIGCAAAAACCCIGIG	187	TGGAAACIC IGGGACACICAA	565
10	GAGAGAGACAGGCACACAIGG	188	AAIGIIACCIIIGAGGGGIGG	3319
11	ACCCIGCCCCACACAICTACT	189	GCTCCAGGICIICCCICICIC	1595
12	CCCIGCICIAAGCCIIIGIICI	190	GICCAGICIGCAACAICCAAC	1482
13	CIGAACI ICCACCCCCIIITAC	191	IGGIICCCIIICI IGAIIICAGC	618
14	GCAGIACGIGGCAGAIGIGAA	192	GIIGCIGGGAGICCTGIGICT	1691
15	GAGGAGGIAAAGAGGICCCAG	193	IACAACATAGAGGGGAGGCAC	1157
16	CICICCIIGCCIIITCIGACICC	194	AGCACAAAGIIGCICACAGGA	305
17	GCCICCCACI IITCICIIITC	195	ACCIGICCAIICACCCCAIIT	2705
<b>BAC C10 120K to 64K</b>				
18	CCAGAGCTTTCTCCAGGICAC	196	AGGAGTAGCAGGACACCCGTT	792
19	CCCCAGAGICIGGIGCIACTI	197	GCCCCACCACI IICICTIICI	4400
20	AIGGCIGIGGIIIGIGATGGI	198	ACAAGAAGGIIIIIGAGGCICC	1356
21	ACACCAIGAAIIGIIGAAGCC	199	AGGIIIGCGGGAGICATAICI	1085
22	IAATGCGTIIITCCICICIGGG	200	GGGAGAGIIGGICCCCTIITA	503
23	GGAGIGAIGICACCCIGIIC	201	AGCIGGGICTGAAICCAGGIA	1674
24	AIGCGIGGIAGGGCAIITTAAG	202	CICIGGICTCCCAICIGCIII	785
25	AIGGACAACIACICCCICCCI	203	GIIGAAAGAACAAGGCAGCIC	1522

26	CCACICCCCAI IGI IGI IGI I	204	AGIGGGAGAGGGAIAGIGGCT	2259
27	AIICCCAGCCAACAATAAIGGG	205	GCAGTACCTGCAACTIGGIGA	708
28	AAGIIGTGGACAGICGAGACG	206	AAIGCACACAGGIGGACAGAT	774
29	ICGATIGGACIGICTCCICCC	207	CICCAACIGCAIICCAACAAG	479
30	GGACACCICIAACCCIGAICC	208	AACIIAIIICCIIGGACCGCIG	414
31	AGTICICCAITGGCIGGICAAIG	209	IACCAAGAGGGGAGACAGAGG	764
32	CCAICAGAAACGAAIIGIICC	210	GCACCAAAGICICCCICCCICT	1085
33	ICCAGGGCTGIAAAAICAICA	211	GGICAACICCAGGGGACACIA	471
34	IGGTICIIIGCCCACIAIGGI	212	GGTGGIGACAGIAAACAGCCC	922
35	AAACTIGGCCTCGCIAGACAA	213	CCCCAIIAIGCICICICIII	548
36	TGGAATIGAGATIGCTCCAAG	214	TGIIGCTTCAGCAIGICAAGA	627
<b>BAC C10 64K to 1</b>				
37	TGAAAAATCCAAGAATCAGGG	215	AGGGIITAGCACTTGTGGAGG	771
38	CACCACIIACCCICCCICICI	216	AGAFGI IAGAIGI IGGGGIGG	638
39	AGGGICICICCAIICCAGAAC	217	ICCACCICIGICICCCIIICIC	694
40	CAIACICCCICCCAGIGCICCI	218	IGGGICICIGIGAGIGGAAGG	2035
41	CAGACCAGAACGAGGGAGAGI	219	GTCCCTAAAGCCIIIGIICCI	3457
42	IIIGAGGACATCACCAIGACA	220	IGACCTIGGCCIIICIIAGII	935
43	CGGACACIAAGGGAGATGGAT	221	AIACCIACCAGCCAGGCICAG	723
44	ACCACIIACCTGACCACIGGC	222	AAICIIITAGACCCCTCACCC	3413
45	CGTIGIAGGAGGAIICAAGCA	223	IIAATACAAAGGICCCCCAGG	1424
46	GCGACCIGITCCAAAAGTICIC	224	ATGGGGAAGAGIGGGGICIAI	881
47	CAGGIGGGAGAGGGGAAGATAA	225	GICITGIIICACAGCACCAIC	2936
48	CIGGAAAGAGGAAGGAGGACA	226	IAIGCIGCCAAAAGAGAACCC	1279
49	CIICCCICCAGGICICAIIGCII	227	GICIGCCAAGGGAAAACAICAI	595
50	ICICCAIAGCICCAAGCACAC	228	CCIGITGIIICCCICCCAGICII	944
51	CCCIACCCICCCACICICACIC	229	CCAGCACCAGGGAGIAGI IIG	1757
52	GGCIAGGAAACGCCIIACIGAG	230	GAGIGCAGGGGCTGAICICIA	313
53	CAGAI GAATGCI AAGCCCAAA	231	CICTGACIGACIGGCGAGATG	522
54	CCACAICAGI GGGACAAAAGA	232	CCCI GGAGAAGI GGGAGIGIA	596
55	CCTGGCICIIIGCCAATAAAI	233	CCCAAACACCAGGTACIAGGC	515
56	CGCAGAGCCIGIGIICITTAIC	234	CGGTACGAAGAAAACCAGGAG	2238
57	GGCIGAAGICCI GAAGGICAI	235	ICICACCCCTCIIICCACTIGI	828
58	CAGGICCAAGAAGAGGGAAGA	236	CTGCAGGIGAGACIAGCAAT	2579
59	IGAACAGGAGICAAAGCIGGA	237	GAAAGGGAAGCAGGAAAGAGG	1188
60	GACACGCAGAGACACTCAGGA	238	CCCIICCIAICTIICCCACC	1833
61	AIAAGIICAGCAGAGCCTCCG	239	AAIGAGCAIGGAGAATCGTGI	1808
62	AAACACAICTIGCIIGGGAGG	240	GAAIGGGACICCTIGAGAGCIG	587
63	GICCCIIIGGAACIIGCAGAT	241	TAACACAIICAGGAIGGACGC	1237
64	CCAICICGCICCCIIACAAAGI	242	AGAGCACIGACCCICCIITAGG	1135
<b>BAC 94L 1 to 64K</b>				
65	TAGTGACTGAGGGTGAAGGG	243	CCACTCACGAAGATGTGGAAG	1166
66	CCIGIGCAAGGIIACAI CCAA	244	CCIIAAGAGGCAGCCAGACTG	2955
67	GCGACCIGITCCAAAAGTICIC	245	CACAGCCTGACIGGACAAAAG	838
68	CGTIGIAGGAGGAIICAAGCA	246	CTGGGAGAGGCAGAGATTCAT	1379
69	CAGGIGTIGGGGTAGAACIGG	247	CICITTCCTGAIITCGAGGIGG	3386
70	CGGACACIAAGGGAGAIIGGAT	248	AIACCIACCAGCCAGGCTCAG	723
71	ICCAGAI GAGACACAI IIGC	249	IGACCIIGGCCITCCIIAGI I	1147
72	AACAGIGCAGACIGCII CAGA	250	IGIICCI GGGCACAI I I I IAA	3461
73	AGIGCICCTCAGAGGGAGITG	251	CCAAAICIGAGGAAAGGGIGA	2116
74	IAGACTGACICICACCACGCC	252	CATCCACCICITGICICCCIIIC	785

ES 2 372 427 T3

75	CACCACIICACCCICCCICICI	253	AGAIGIIGGGGIGGIAGAAGG	631
76	CCAAGAAGGAAGCIGAGIIGC	254	GGIGGAGACAGGGIITAGCAC	815
77	ACTAIGGCIGIGACICCCAC	255	AICIACCCIGACCCATCIGGI	833
78	GGTACCAGCAAACIGIGCCIC	256	ICCAACIACAGGCCICIIIG	597
79	GGIICIIIGCCACIAGGTC	257	IIIGCTCAIIAACCAICAGG	908
80	ICCAGGGCIGIAAAAICAICA	258	AGGGGAGACICCCIAIIIGIC	729
81	AGGTGGGIICAGIGAAIGAAA	259	ACCIACIICACCAGCCAGCTI	1085
82	ICCAIGGCIGGICAAIGAI	260	ATIACCCAAGAGCCCCAGI	787
83	AIGGCIGCCACAGIAGCIIII	261	AGGGIAGGGGIGAGGAICAGG	382
84	CCIIICGAIGIGACIGICICC	262	CCAACIGCAIICCAACAAGIC	481
85	CGAAGIIGTGGACAGICGAGA	263	AAIGCACACAGGIGGACAGAI	776
86	ACACIGCICCCIGAGICACIG	264	GCTCAGCCIGACAGCICAGIA	488
87	CCACICCCCAIIGIIGIIGII	265	GCTGGCAAGAGAGCACAAGAI	2210
88	GAGAAGGAAGGAGAGAGCIGC	266	GCIGIIGAAAGAACAAGGCAG	1572
89	AIGCGIGGIAGGGCAIIAAG	267	CIGGICCCCATCIGCIIICI	783
90	AGIGAIGICACCCITGITCCI	268	AGCIGGGICTGAAICCAGGIA	1672
91	GTAATGCGIICCCICICIGG	269	GIGGGGAIAGAACIGCTAGGG	604
<b>BAC 94L 64K to 120K</b>				
92	ACACCATGAATTGTTGAAGCC	270	AACCCCAATGAAGAGAGACCA	1059
93	GIGIGIGGICICCCAIACCCI	271	GIICCAAGAGIGGCIIIGG	1352
94	CICIAACCACCIGAGGGCIIIG	272	CCAACAIGAGIICCCIIICAI	3633
95	ICCAGGGIGGACCICIIAICA	273	CACCIGTCCIAICACCCAI	4080
96	GCACACAACIGGIICCGIIAA	274	AGAGGGGAGGCACAGGACIIA	1300
97	AAGTCCIGCTCACICAIGCIG	275	GCIGGGAGICCIIGIGICICAI	1578
98	AAIAAAAGGAAIIGGIGGGGC	276	IGGIICCCIIICIIGAIICAGC	679
99	CAITCCICAGCCACAGIGACA	277	CIGGACAIGCIGAAGAGGIGA	576
100	CCIGICICIAGCCIIIGIICI	278	CTCTCCICCCCAACICAACI	1435
101	CCCACACATCTACTGGAGGAA	279	CICCIAAGACAGGCCICAACC	1625
102	AAGGAGAGAGACAGGCACACA	280	AIGITACCIIGAGGGGIGGI	3321
103	CCITGAAGAACCAACAGGAIG	281	GAAGAAGGIGGIGGAGAGGAA	593
104	IGIGACCICCAIATCCAGIGC	282	CACCICAAGAACAGGAACIGG	3179
105	IAATGCAIGGIAGGACCGAAI	283	IATAGGGIGGIGGGICIIIC	1489
106	ICICCCACCAACIICAAIGA	284	CCICICCAAGGCACIGIIIII	1223
<b>BAC 94L 120K to end</b>				
107	GAGGGGATCTCCCTAAACTGA	285	GCACCAIGCTITITIIICAAA	1697
108	GGAAACATCAITCCIGGIICCA	286	ICAAAGGGCACICAIITICAI	901
109	AACIGGACAAGCTCITIGGGA	287	AACCIGCCTCIGICIIIGAIACC	796
110	AGICCTIGITGCCAAIACCIG	288	GAGGGGAAAAAICAGAAAAC	820
111	AAAAAGCAGAAGCACTGCAAG	289	ATIGAGGGIAGAGGAGGIIIG	634
112	ICIGAAGAAATGIACGGCAGC	290	CGAATGGCIAACICCCACAAA	535
113	ATGAAGGAAAGGIITCACCCA	291	GGGACCICAIGIICITIGAIGIIA	643
114	CCAAICAAGAGGAAAGIIGGA	292	AAATAGCCIIGCTIAGIGGGIGC	875
115	ICCIACAAGGIGIIGAGGGGC	293	CAIGGGGAGGIACAACIIIG	1935
116	IIGGATIIIGACCTCAIGCACI	294	CCCCAGTAACCAIGCAGAGAG	1148
117	CAGAGGCAGGAAAAGAIGIC	295	GGCAITIIICAAAATIAACAGACG	1377
118	CTGGCCATCAGIAAATCACATCA	296	GGCAACAICTAAAACIICAGCCI	551
119	ATTAGCCGAAGACAGAGGGAA	297	IIGCIIGCAGCCITGAAGIAT	435
120	IGIIGAAAAACACAAGGGCAC	298	AGGTGICTIIAAGCCIGAGCC	1876
121	IATITCCAGAACTIIGGCACA	299	AAAATCCITGICCIIGGCICAIC	572
122	CCCGIIGACACAGIACAAGC	300	TACACCCITCATCTCCIGGCI	766
123	IAAAIGTGCAACICAGGCAGG	301	AAGATIIIGGCACITGAAAGGA	1456

<b>BAC C17 1 to 64K</b>				
124	CCCATCGTTTCAGTGTTCITCTT	302	CACGAACACACACACACACCC	1007
125	IGCAAGGGAIIIIIGAATGAA	303	CATCACIICAAACACAAGAGCAII	628
126	AIAGAIIGCAGIIIGTIIGGCC	304	IAIITAGCCGAAGACAGAGGGA	528
127	AGGIGICIIIAAGCCIGAGCC	305	CCIGATIGGACAIITICCCAIA	1801
128	IACACCCIIICATICCCIGGCI	306	ACAAGCIGTAAAGCICIGCCC	751
129	ICAAAAGGGCAICAGIGAGGA	307	IAAATIGIGCAAATCAGGCAGG	1415
130	CIGGCACAGGAIGIGGICAI	308	GIGCIITACCACIIGGGIIII	407
131	GAGICCAIGGIGACCACAIIII	309	GACCCCAIIIAACIIICAGC	1064
132	AICACAGCCAICAIGGICAAC	310	GGCIGAGGCIIIAIIIIIGGAG	1476
<b>BAC C17 180 to end</b>				
133	IGCCACAIIIACAGICCCAGG	311	GCGGIIITICACTGACGCAGTA	1386
134	CCATIIITCIAATCCAIGCCC	312	CACCIITIGIGGAACIGCCIAAC	533
135	IGCCCCIACACCAAACAIAACC	313	AAACCTTCGCAIITICAIACCIC	218
136	CIICIGICIGIIAIGGICGGG	314	GCCIIITCIGIGGCIIIIIGIII	506
137	IGGGGGICACAIIGAIICAIA	315	IGIICTIACCAGGACGCCACGG	832
138	GGIGIIGGGAGAAGAIGIIGA	316	ACGCAAGGACCIGACAIIAAA	2502
139	CCIIIGGAAATCACICCCIIIGC	317	IGGIIAACIGAAAAIGGGIGGA	1849
140	IGAIGGIIICACIGCAICIGG	318	ITAAITGCIICCACCAACCCI	2525
141	GCIICICAIIAAGCCAIGCACA	319	AGAICCCIGGICITIIIGIICCC	243
142	AIGAIGCICIGGGAIGIGAAA	320	CCIGGAAAGCAGAGAACIAGG	529
143	GIGGIGAIIAAICTIIGGGGG	321	CAITIIIGGCAIGIAIGIGGIG	671
144	CCIIICAGACCIGCAAAACICC	322	CCCCICTCCIAGAAAAIICC	636
145	CIATCAGCICAGCAGCAAGGA	323	GCGAAIGGGAIACAICAAAGA	388
<b>BAC C17 120K to 180K</b>				
146	GCAAGGGCCAAATAACCAAGT	324	CAGAGCCIAAAGAACCCACCA	1330
147	AICCIICCAIGIICCIIGGCI	325	IGCACACIIIAACTGCACCAI	908
148	IGGCIIICIGIITCIGAGIIGG	326	CAGIAAGCAAACIGCCCIICAI	598
149	CCITAAICIGCCICCAGCICA	327	AGCCAAAGGITCCAGGACAAI	913
150	IAAGTIIGGGGIIIGGGGAGAI	328	CCAGCACCCICACCCIGACIAI	4665
151	CGCACCAAAATICTAICATCC	329	IGTGAIGCIACCCACICCAI	2847
152	CTIICTIGCCITAAIGCIGGG	330	GGICICCTCGGIACICCCAI	581
153	GAAGAAGGIGATCIGGGAACG	331	GCAAGIGCAAGGAAGGAAAGA	590
154	ICACTCIGIIGGIIGIGICGG	332	GGAGGCCIGCTGAACITCIII	462
155	GGACCAGACCAGATGGTAGGA	333	GCACATAAGGCICACAGGAAA	618
156	ICAAACAICCTACAGCGAAGC	334	ATAGTIGIGGGIAGGAIIGTGC	3185
157	CCICICCCICICCICTAICA	335	AAAAGCAAGAGCAGAGAAGGC	1182
158	ICICAAGIGCAIACCAGCIC	336	AGGCGCIGCAACIACAAAGAT	562
159	IAAGCCIACCACACCAGCCAC	337	CIGACCIICACAGACAICCCC	920
160	AGAGCACCCAGCAGGTACAGA	338	GGACTACAAGGGACGAGAGGC	564
161	GGGCAAGGAGAAIGIIGTAGG	339	CIGGGAGGATCACCTGACAGA	589
162	IGICGIIAGICAICIGGICCC	340	CACCACCAAGACAAGCCIAAA	695
<b>BAC C17 64K to 120K</b>				
163	GACACACTTCCTCCATCTGGC	341	CCATAAATGAAGCCTCCTGCC	737
164	GICICIGCTCACCCACICAT	342	CITGGAAAAGCCAAIGGAAAT	606
165	TIGAGGACAAAGGTCICAGGC	343	AAGGICCIAGCCCTIAGCAGA	809
166	TTACCIGITAGGGCICCAACG	344	ACACACACATCIGCAACTCGG	1410
167	CCAGCGGIIIGAIAGAIIGI	345	CCGIITCAGCAAAACIGAGAA	684
168	AAATAAAICACAGCCGAGGI	346	AGTIGICIITAGIACCCCCIGC	1679
169	CACCAICACAICCICAAAGC	347	CCCAGGAIGIGGAGAIAGAAAG	1099
170	CAGAAGAAAGAGGCAGCAGCA	348	GGTGGGGGACIAGGAGTGAAA	574

171	GGACAGGGCIGAACGAAATAA	349	ACGAAGTCAGI I IGGIGGIGG	493
172	CACATGCACAICCAIGCTCTC	350	CACCCACACTTTTCIGCCCTC	1693
173	AGCIGGIGAI GGACACAIGGI	351	GGIGAGCCCI I AICCCAGIT	1413
174	I I AACCACCIGAACCIGICC	352	CICCAGCCCI GGICACAATAI	348
175	AGAACI I ICCICCCICCCCI	353	CGIGICCACI I CAAGGIGAAI	3389
176	IGIGAGGGAAAIC I ACCI I CG	354	CACCAGGCI I GICAI I I ACCA	951
177	GAICICAGGGICI I ICIGGG	355	GAIGCCICAICI I ICCICACI	1048
178	IGI I I A I I I AIGIGGCAGGI I GG	356	GIGAACICACI CACI I GGGI I AGC	

Los cebadores directos se sintetizaron con un 5' fosfato, mientras que los cebadores inversos no. Los productos de amplificación resultantes poseían 5' fosfatos en el extremo único.

- 5 Los cebadores se disolvieron en agua hasta 20 X de la concentración de la reacción final. La concentración final era de 0,5  $\mu$ M por cada cebador. Las regiones de unión resultantes fueron amplificadas mediante la reacción en cadena de polimerasa (PCR) utilizando TOPOTAQ™1 (Fidelity Systems Inc., Gaithersburg, MD). En cada reacción inicial, se utilizaron 0,2  $\mu$ g de ADN BAC como plantilla en un volumen final de 100  $\mu$ l. Las PCR se realizaron en una mezcla que contenía: 0,5mM cada dNTP, 0,5  $\mu$ M cada cebador, 1 Unidad TOPOTAQ™, en un tampón a un pH final de aproximadamente 8.0. Las condiciones de la PCR fueron las siguientes: 98°C, 5 minutos; 30 ciclos de 98°C, 10 segundos; 55°C, 20 segundos; 72°C, 3 minutos, seguidas de una incubación de 10 minutos a 72°C. Las reacciones completadas se almacenaron a 4°C.

- 15 Los productos de amplificación de cada reacción inicial se pusieron en geles de agarosa de bajo punto de fusión y, después de fotografiarlos, fueron cortados y purificados. Los fragmentos de gel se digirieron con GELASE™ (Epicentre, Madison, WI) de acuerdo con las instrucciones del fabricante y se recuperaron mediante precipitación con isopropanol. Los productos de ADN precipitados se disolvieron en 100  $\mu$ l de agua. Estos productos se designaron «SEED 1».

- 20 Se prepararon las reacciones en cadena de polimerasa (100  $\mu$ l) conteniendo 1  $\mu$ l de SEED 1, y se amplificaron utilizando las mismas condiciones y cebadores que anteriormente. El producto de esta reacción se diluyó hasta 200  $\mu$ l y se designó «SEED 2».

Cada SEED 2 (5  $\mu$ l) se utilizó para preparar reacciones, cuyos productos se denominaron «SEED 3». Cada reacción SEED 3 se realizó 10 veces y las 10 reacciones resultantes se combinaron en una placa de microtitulación de un único pozo profundo. Estas se denominaron «productos PCR combinados».

- 25 Una alícuota (10  $\mu$ l) de cada uno de los productos PCR combinados se puso en gel de agarosa para testar el éxito de las reacciones. Las alícuotas de cada reacción combinada se combinaron y purificaron mediante extracción de fenol y se recuperaron mediante precipitación de isopropanol. El gránulo precipitado se disolvió en 100  $\mu$ l TE.

- 30 Este proceso se realizó para cada uno de los tres BCS resultantes en productos de 100  $\mu$ l para cada BAC. Las concentraciones finales se determinaron espectrofotométricamente como: C10 828 ng/ $\mu$ l; 94L 525 ng/ $\mu$ l; and H17 717 ng/ $\mu$ l. Los volúmenes correspondientes a 10  $\mu$ l cada uno (es decir, 12  $\mu$ l, 19  $\mu$ l, y 14  $\mu$ l, respectivamente) se combinaron y trataron con polimerasa de ADN T4 y 4 dNTP en una reacción de 500  $\mu$ l, para generar extremos romos.

- 35 Los productos de la región de unión amplificados fueron extraídos con fenol y precipitados; a continuación se unieron como sigue. Los amplímeros de la región de unión se disolvieron en 40  $\mu$ l de agua y se trataron con polinucleótido quinasa T4 (NEB, de acuerdo con las instrucciones del fabricante); para después aplicar ligasa de ADN T4 durante una noche a temperatura ambiente en el mismo tampón. Al día siguiente se puso una alícuota de 1  $\mu$ l en gel de agarosa para confirmar la unión. La reacción de unión produjo una banda difusa de unos 10 kb, sin ningún fragmento desunido residual visible. Este material se extrajo con fenol y se recuperó en un volumen de 30  $\mu$ l de TE. El producto resuspendido, designado SEED H-ZERO (las moléculas de ácido nucleico FPC resultantes) se almacenó congelado.

- 40 SEED H-ZERO (las moléculas de ácido nucleico FPC) se amplificó después con polimerasa Phi29 utilizando hexámeros al azar como cebadores, en las siguientes condiciones de reacción: aproximadamente 20  $\mu$ g de plantilla en un tampón que contenía 37 mM trisHCl pH 7.5, 50 mM KCL, 10 mM MgCl<sub>2</sub>, 5 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 1 mM dNTP's, 50  $\mu$ M de cebadores, 1  $\mu$ g/ml de pirofosfatasa (levadura) y 0,5  $\mu$ g/ml de polimerasa de ADN phi29.

SEED H-ZERO (1  $\mu$ l) se amplificó en una reacción de 100  $\mu$ l a 30 °C durante una noche. Esta reacción produjo una cantidad muy grande de ADN (aproximadamente 1 mg/ml), que se designó H1. De los 100  $\mu$ l de H1, 10  $\mu$ l se utilizaron para preparar una reacción de 1 ml para producir H2. H2 fue idéntico a H1 en apariencia sobre un gel de agarosa. Se emplearon 100  $\mu$ l de H2 en una reacción de amplificación de 10 ml para producir H3, que, una vez más, era idéntico en apariencia sobre gel. Se utilizó un ml de H3 en una reacción de 50 ml (2 X 25 ml) que se dejó activar durante unas 60 horas para generar H4. El producto de H4 se precipitó con isopropanol y se disolvió en un total de 10 ml de tampón TE. Aproximadamente 28,9 mg de ADN (H4) se produjeron mediante esta serie de reacciones.

ADN H4 (100  $\mu$ g) (los amplímeros de la molécula de ácido nucleico FPC resultantes) se etiquetó con dinitrofenol (DNP), utilizando el juego MIRUS de acuerdo con las instrucciones del fabricante. El ADN se purificó mediante precipitación con isopropanol y se lavó con 80% EtOH. Después del etiquetado y la purificación, el ADN se ajustó a 50 mM NaPO<sub>4</sub> de tampón pH6 8, 10 mM EDTA, y se calentó a 100 °C en un baño de agua durante 45 minutos para fragmentar el ADN. Esta sonda que contenía amplímeros de la molécula de ácido nucleico FPC etiquetados se utilizó en los experimentos iniciales.

### **Ejemplo 2: Síntesis de la sonda utilizando secuencias de oligonucleótido fijadas**

Este ejemplo describe un método alternativo utilizado para generar una sonda que incluye moléculas de ácido nucleico FPC heterogéneas específicas para una región del cromosoma humano 17 que contiene el gen HER2. En este ejemplo, se utilizó una secuencia de oligonucleótido fijada (un hexámero compuesto de purinas) para minimizar los productos de amplificación derivados del cebador. Sin embargo, una persona con conocimientos en el campo apreciará que se pueden emplear métodos similares con otras secuencias de oligonucleótidos fijadas. Se pueden emplear métodos similares para generar una sonda que incluye moléculas de ácido nucleico FPC heterogéneas de otras secuencias de ácido nucleico diana, utilizando los cebadores adecuados.

La secuencia de oligonucleótido fijada era un hexámero compuesto exclusivamente de purinas, GAGGAG (SEC ID N°: 357). El hexámero GAGGAG (SEQ ID NO: 357) es el sitio de reconocimiento para la endonucleasa de restricción Bsel, que puede utilizarse después para el análisis y la posterior manipulación de las sondas. Se describen dos planteamientos alternativos para incorporar secuencias de oligonucleótido fijas a un amplímero de la región de unión.

En el primer método, las regiones de unión sustancialmente libres de secuencia de ácido nucleico repetitiva de la región del cromosoma 17, incluyendo el gen HER2, se generaron y amplificaron, tal y como se describe en el Ejemplo 1. Los amplímeros de la región de unión resultantes se unieron en uno de los extremos (o ambos) a los oligonucleótidos (por ejemplo, hexámeros) duplicados (de doble cadena) compuestos solo por purina en una cadena y pirimidinas complementarias en la otra. Dado que una cadena del oligonucleótido fijado duplicado se compone de una secuencia solo de purina, p. ej., CTCCTC (SEC ID N°: 358), la amplificación posterior se puede realizar utilizando solamente un único cebador, p. ej., GAGGAG (SEC ID N°: 357). Posteriormente, los amplímeros de la región de unión que contenían el hexámero se unieron para formar una molécula de ácido nucleico FPC, tal y como se describe más detalladamente a continuación.

En un segundo planteamiento, las regiones de unión sustancialmente libres de secuencia de ácido nucleico repetitiva de la región del cromosoma 17, que incluye el gen HER2, se amplificaron para formar amplímeros de la región de unión, utilizando un cebador específico para la secuencia que se incorporó a una secuencia compuesta exclusivamente por purinas. La secuencia de oligonucleótidos fijada (purinas) se incorporó en el cebador 5' a la región complementaria de la secuencia diana de la que se inició el cebador y la extensión. Los amplímeros de la región de unión resultantes incluyen la secuencia de un hexámero en el 5' extremo. Posteriormente, los amplímeros de la región de unión que contienen el hexámero se ligan para formar una molécula de ácido nucleico FPC como la que se describe más detalladamente a continuación.

Los métodos descritos anteriormente presentan cada uno la secuencia hexamérica de los extremos de cada fragmento PCR (es decir, el amplímero de la región de unión). Posteriormente, los amplímeros de las regiones de unión que contenían el hexámero se unieron entre sí (tal y como se describe en el Ejemplo 1) para generar una población de moléculas de ácido nucleico FPC que incluía la secuencia del hexámero entre cada región de unión. Las moléculas de ácido nucleico FPC que contenían el hexámero se utilizaron como plantillas para la amplificación, utilizando polimerasa de ADN Phi29 cebada con el hexámero completamente de purina. Las moléculas de ácido nucleico FPC amplificadas resultantes se utilizaron como sonda, después del etiquetado con DNP, tanto por métodos químicos como enzimáticos (véase el Ejemplo 1).

La amplificación de la molécula de ácido nucleico FPC que contenía el hexámero con un cebador completamente de purina reduce (o elimina) la aparición de productos compuestos solamente del cebador (p. ej., dímeros del cebador) en la mezcla de amplificación final. Por otra parte, en la concentración del cebador estándar utilizada para los hexámeros, la concentración molar del cebador hexámero fijado está disponible en más de 4.000 veces la concentración de cualquier hexámero al azar en una mezcla. Esto resulta en una hibridación más completa del cebador con su secuencia complementaria en toda la secuencia de ácido nucleico diana, reduciendo el efecto de

reducción del cebador selectivo durante la amplificación, con una mejora resultante de la retención de la heterogeneidad de la secuencia como resultado de la amplificación.

**Ejemplo 3: Análisis de la hibridación *in situ***

5 Este ejemplo describe métodos que demuestran que la sonda que contiene moléculas de ácido nucleico FPC etiquetadas y amplificadas se puede utilizar para la hibridación *in situ*. Una persona con conocimientos en el campo apreciará que se pueden utilizar métodos similares para sondas específicas para otras secuencias de ácido nucleico diana.

10 La sonda etiquetada generada en el Ejemplo 1 se formuló en un tampón de hibridación (50% formamida, 2x SSC, 10% sulfato de dextrano) para uso en ensayos. La CISH y SISH se realizaron en secciones de tejido de cuatro micrones de grosor embebidas en parafina y fijadas con formalina, montadas en un portaobjetos del microscopio utilizando un protocolo ISH automatizado (véase la Tabla 2), disponible conjuntamente con el instrumento Ventana Benchmark™ XT (Ventana Medical Systems, Tucson, AZ).

Tabla 2 Ejemplo de protocolo de teñido SISH automatizado

Paso	Procedimiento
1	***** Seleccionar EZ Prep *****
2	***** Iniciar pasos programados *****
3	***** Mezcladores desactivados*****
4	Si *desparafinización* está *seleccionado*
5	Calentar portaobjetos a *75 °C* e incubar durante *4 minutos*
6	Aplicar el ajuste de volumen EZPrep
7	Incubar durante *4 Minutos*
8	Enjuagar portaobjetos
9	Aplicar el ajuste de volumen EZPrep
10	Incubar durante *4 Minutos*
11	Enjuagar portaobjetos
12	Aplicar el ajuste de volumen EZPrep
13	Aplicar cubreobjetos
14	Calentar portaobjetos a *76 ° C* e incubar durante *4 Minutos*
15	Desactivar calentador de portaobjetos
16	Incubar durante *4 minutos*
17	***** Mezcladores activados*****
18	Enjuagar portaobjetos
19	Aplicar 900ul de tampón de enjuague
20	Aplicar cubreobjetos
21	Calentar portaobjetos a *37 ° C* e incubar durante *4 minutos*
22	Aplicar acondicionador de la célula #2
23	Calentar portaobjetos a *90 ° C* e incubar durante *8 minutos*
24	Aplicar acondicionador de la célula #2
25	Calentar portaobjetos a *90 ° C* e incubar durante *12 minutos*
26	Si *CC2 extendido* está *seleccionado*
27	Aplicar acondicionador de la célula #2
28	Aplicar cubreobjetos corto
29	Calentar portaobjetos a *90 ° C* e incubar durante *8 minutos*
30	Desactivar calentador de portaobjetos
31	Incubar durante *4 minutos*
32	Enjuagar portaobjetos con tampón de reacción
33	Aplicar 900ul de tampón de reacción
34	Aplicar cubreobjetos
35	***** Seleccionar lavado SSC *****
36	Enjuagar portaobjetos con tampón de reacción
37	Ajustar volumen de portaobjetos con tampón de reacción
38	Aplicar cubreobjetos
39	Calentar portaobjetos a *37 ° C* e incubar durante *4 minutos*
40	Si *1SH-Proteasa 2* está *seleccionado*
41	Enjuagar portaobjetos con tampón de reacción
42	Ajustar volumen de portaobjetos con tampón de reacción
43	Aplicar una gota de *ISH-PROTEASA 2*. Aplicar cubreobjetos e incubar durante #tiempo de incubación#

ES 2 372 427 T3

44	Si *ISH-Proteasa 2* está *no seleccionado*
45	Si *ISH-Proteasa 3* está *seleccionado*
46	Enjuagar portaobjetos con tampón de reacción
47	Ajustar volumen de portaobjetos con tampón de reacción
48	Aplicar una gota de *ISH-PROTEASA 3*. Aplicar cubreobjetos e incubar durante #tiempo de incubación#
49	Desactivar calentador de portaobjetos
50	Enjuagar portaobjetos con tampón de reacción
51	Aplicar 300ul de tampón de reacción
52	Aplicar cubreobjetos
53	Enjuagar portaobjetos con tampón de reacción
54	Aplicar 300ul de tampón de reacción
55	Aplicar cubreobjetos
56	Enjuagar portaobjetos
57	Ajustar volumen de portaobjetos
58	Aplicar cubreobjetos
59	Calentar portaobjetos a *37 ° C* e incubar durante *4 minutos*
60	Enjuagar portaobjetos
61	Ajustar volumen de portaobjetos
62	Si *sonda ADN HER2* está ^seleccionado*
63	Aplicar una gota de *SISH DET HYB*. Aplicar cubreobjetos e incubar durante *4 minutos*
64	Aplicar dos gotas de *sonda ADN HER2*, e incubar durante *4 minutos*
65	Calentar portaobjetos a *95 ° C* e incubar durante *12 minutos*
66	Calentar portaobjetos a *52 ° C* e incubar durante *4 minutos*
67	Aplicar cubreobjetos corto
68	Incubar durante *2 horas*
69	Enjuagar portaobjetos
70	Ajustar volumen de portaobjetos
71	Aplicar cubreobjetos
72	Enjuagar portaobjetos
73	Aplicar 900ul de tampón de enjuague
74	Aplicar cubreobjetos
75	Calentar portaobjetos a *72 ° C* e incubar durante *8 minutos*
76	Enjuagar portaobjetos
77	Aplicar 900ul de tampón de enjuague
78	Aplicar cubreobjetos
79	Calentar portaobjetos a *72 ° C* e incubar durante *8 minutos*
80	Enjuagar portaobjetos
81	Aplicar 900ul de tampón de enjuague
82	Aplicar cubreobjetos
83	Calentar portaobjetos a *72 ° C* e incubar durante *8 minutos*
84	Enjuagar portaobjetos
85	Si *sonda ADN HER2* está *no seleccionado*
86	Si *sonda Chrl 7* está *seleccionado*
87	Aplicar una gota de *SISH DET HYB*. Aplicar cubreobjetos e incubar durante *4 minutos*
88	Aplicar dos gotas de *sonda Chrl7*, e incubar durante *4 minutos*
89	Calentar portaobjetos a *95 ° C* e incubar durante *12 minutos*
90	Calentar portaobjetos a *44 ° C* e incubar durante *4 minutos*
91	Aplicar cubreobjetos corto
92	Incubar durante *2 horas*
93	Enjuagar portaobjetos
94	Ajustar volumen de portaobjetos
95	Aplicar cubreobjetos
96	Enjuagar portaobjetos
97	Aplicar 900ul de tampón de enjuague
98	Aplicar cubreobjetos
99	Calentar portaobjetos a *59 ° C* e incubar durante *8 minutos*
100	Enjuagar portaobjetos
101	Aplicar 900ul de tampón de enjuague
102	Aplicar cubreobjetos
103	Calentar portaobjetos a *59 ° C* e incubar durante *8 minutos*
104	Enjuagar portaobjetos

ES 2 372 427 T3

105	Aplicar 900ul de tampón de enjuague
106	Aplicar cubreobjetos
107	Calentar portaobjetos a *59 ° C* e incubar durante *8 minutos*
108	Enjuagar portaobjetos
109	Si *sonda Chrl 7* está *no seleccionado*
110	Si *control negativo ISH* está "seleccionado"
111	Aplicar una gota de *SISH DET HYB*. Aplicar cubreobjetos e incubar durante *4 minutos*
112	Aplicar una gota de *ISH NEG CTRL*, e incubar durante *4 minutos*
113	Calentar portaobjetos a *95 ° C* e incubar durante *12 minutos*
114	Calentar portaobjetos a *52 ° C* e incubar durante *4 minutos*
115	Aplicar cubreobjetos corto
116	Incubar durante *2 horas*
117	Enjuagar portaobjetos
118	Ajustar volumen de portaobjetos
119	Aplicar cubreobjetos
120	Enjuagar portaobjetos
121	Aplicar 900ul de tampón de enjuague
122	Aplicar cubreobjetos
123	Calentar portaobjetos a *72 ° C* e incubar durante *8 minutos*
124	Enjuagar portaobjetos
125	Aplicar 900ul de tampón de enjuague
126	Aplicar cubreobjetos
127	Calentar portaobjetos a *72 ° C* e incubar durante *8 minutos*
128	Enjuagar portaobjetos
129	Aplicar 900ul de tampón de enjuague
130	Aplicar cubreobjetos
131	Calentar portaobjetos a *72 ° C* e incubar durante *8 minutos*
132	Enjuagar portaobjetos
133	Ajustar volumen de portaobjetos
134	Aplicar cubreobjetos
135	Desactivar calentador de portaobjetos
136	Enjuagar portaobjetos con tampón de reacción
137	Ajustar volumen de portaobjetos con tampón de reacción
138	Aplicar cubreobjetos
139	***** Sincronización del procedimiento *****
140	Calentar portaobjetos a *37 ° C* e incubar durante *4 minutos*
141	Enjuagar portaobjetos con tampón de reacción
142	Ajustar volumen de portaobjetos con tampón de reacción
143	Aplicar una gota de *anti-DNP de conejo*. Aplicar cubreobjetos e incubar durante *20 minutos*
144	Enjuagar portaobjetos con tampón de reacción
145	Ajustar volumen de portaobjetos con tampón de reacción
146	Aplicar cubreobjetos
147	Enjuagar portaobjetos con tampón de reacción
148	Ajustar volumen de portaobjetos con tampón de reacción
149	Aplicar una gota de *SISH DET HRP*. Aplicar cubreobjetos e incubar durante *16 minutos*
150	Enjuagar portaobjetos con tampón de reacción
151	Ajustar volumen de portaobjetos con tampón de reacción
152	Aplicar cubreobjetos
153	Enjuagar portaobjetos con tampón de reacción
154	Ajustar volumen de portaobjetos con tampón de reacción
155	Aplicar cubreobjetos
156	Desactivar calentador de portaobjetos
157	***** Seleccionar lavado opcional *****
158	Enjuagar portaobjetos
159	Ajustar volumen de portaobjetos
160	Aplicar cubreobjetos
161	Enjuagar portaobjetos
162	Desagüe del chorro
163	Aplicar una gota de *Silver A;.. Aplicar cubreobjetos e incubar durante *4 minutos*
164	Enjuagar portaobjetos
165	Ajustar volumen de portaobjetos
166	Aplicar cubreobjetos

167	Enjuagar portaobjetos
168	Ajustar volumen de portaobjetos
169	Aplicar cubreobjetos
170	Enjuagar portaobjetos
171	Ajustar volumen de portaobjetos
172	Aplicar cubreobjetos
173	Enjuagar portaobjetos
174	Desagüe del chorro
175	Aplicar una gota de *Silver A* Aplicar cubreobjetos e incubar durante *4 minutos*
176	Aplicar una gota de *Silver B* e incubar durante *4 minutos*
177	Aplicar una gota de *Silver C* e incubar durante *12 minutos*
178	Enjuagar portaobjetos
179	Ajustar volumen de portaobjetos
180	Aplicar cubreobjetos
181	Enjuagar portaobjetos con tampón de reacción
182	Ajustar volumen de portaobjetos con tampón de reacción
183	Aplicar cubreobjetos
184	Si *contracolorar* está -""seleccionado*
185	Enjuagar portaobjetos con tampón de reacción
186	Ajustar volumen de portaobjetos con tampón de reacción
187	Aplicar una gota de #Counterstain# ( *Contracolorar* ). Aplicar cubreobjetos e incubar durante #tiempo de incubación#
188	Enjuagar portaobjetos con tampón de reacción
189	Ajustar volumen de portaobjetos con tampón de reacción
190	Aplicar cubreobjetos
191	Si *post contracolorar* está *seleccionado*
192	Enjuagar portaobjetos con tampón de reacción
193	Ajustar volumen de portaobjetos con tampón de reacción
194	Aplicar una gota de #Connterstain# ( *post contracolorar* ), Aplicar cubreobjetos e incubar durante #tiempo de incubación#
195	Enjuagar portaobjetos con tampón de reacción
196	Ajustar volumen de portaobjetos con tampón de reacción
197	Aplicar cubreobjetos
198	***** Seleccionar lavado SSC *****
199	***** Iniciar pasos temporizados *****
200	Enjuagar portaobjetos con tampón de reacción
201	Ajustar volumen de portaobjetos con tampón de reacción
202	Aplicar cubreobjetos

- 5 En resumen, tras la eliminación de parafina y el tratamiento de proteasa, la hibridación con la sonda etiquetada DNP se realizó durante dos horas a 50 °C en 2XSSC y 23% de formamida. Después del lavado con 2XSSC, se aplicó un anticuerpo anti-DNP (conejo 75 µg/ml), seguido de una incubación de 20 minutos a 37 °C. Después del lavado, se aplicó un anticuerpo de cabra anti-conejo conjugado con peroxidasa de rábano (20 fig./ml) y se incubó otros 20 minutos. Después del lavado con 100 mM de tampón de citrato de pH 3.6, se aplicó una solución de acetato de plata (3 68 mg/ml). Cuatro minutos más tarde, sin lavado, se aplicó una solución de hidroquinona (1,78 mg/ml en 0,1 M citrato de pH 3.8, seguida de una solución de 0,09% peróxido de hidrógeno, Después de 12 minutos, se lavó el portaobjetos y se secó para el montaje.
- 10 Algunos ejemplos de resultados de SISH se ilustran en la FIG. 6. Las FIG. 6A y C son imágenes de campo claro, las FIG. 6B y D son imágenes negativas que muestran la señal como puntos blancos. Las FIG. 6A y B muestran una muestra en la que la secuencia diana HER2 está amplificada (diploide). Las células de esta muestra exhiben dos o menos señales de hibridación (que aparecen como una mancha oscura en el luminoso y brillante teñido de las imágenes de campo oscuro). Las FIG. 6C y D muestran una muestra en la que la secuencia diana HER2 está amplificada demasiadas veces el número de copia diploide. Las señales de hibridación aparecen como un agregado multifocal.
- 15

**Ejemplo 4: Las moléculas de ácido nucleico FPC sustancialmente no fragmentadas o fragmentadas detectan las moléculas de ácido nucleico diana**

- 20 Este ejemplo describe métodos utilizados para demostrar que las moléculas de ácido nucleico FPC no fragmentadas, así como las moléculas de ácido nucleico FPC fragmentadas, se pueden utilizar como una sonda para detectar una secuencia de ácido nucleico diana. A pesar de que se describen métodos concretos para fragmentar las moléculas

de ácido nucleico FPC, una persona con conocimientos en el campo apreciará que se pueden emplear otros métodos.

5 La sonda que contiene moléculas de ácido nucleico FPC específicas para HER2 descritas en el Ejemplo 1 se produjo por un método que no reduce el tamaño de la sonda (p. ej., mediante fragmentación o de otro modo). Este Ejemplo demuestra que una sonda que contiene moléculas de ácido nucleico FPC sustancialmente no fragmentadas o que contiene una pluralidad de fragmentos de moléculas de ácido nucleico FPC resulta útil para la detección de una secuencia de ácido nucleico diana.

10 Después del etiquetado, tres alícuotas de las moléculas de ácido nucleico FPC específicas para HER2 descritas en el Ejemplo 1 se colocaron en tubos separados y se calentaron a 100°C durante 5, 30 o 60 minutos, respectivamente, para producir una pluralidad de fragmentos de moléculas de ácido nucleico FPC en diferentes rangos de tamaño. Las muestras del ADN etiquetado resultante se analizaron mediante electroforesis de gel de agarosa, para establecer el rango de tamaño aproximado de los fragmentos de molécula de ácido nucleico FPC de las alícuotas. Como se muestra en la FIG. 7, la sonda que contiene moléculas de ácido nucleico FPC no fragmentadas (banda 1) apenas entró en la agarosa, lo que es coherente con su gran tamaño (sustancialmente >23Kb en comparación con los marcadores más ligeros). El tamaño de los fragmentos de la sonda se redujeron como una función de la duración del tratamiento de calor a 100°C. El tamaño estimado (aproximado) de la mayoría de los fragmentos tratados con calor durante 5 (banda 2), 30 (banda 3) o 60 (banda 4) minutos era aproximadamente de 500, 300, o 100 pares de bases, respectivamente.

20 La sonda que contenía la molécula de ácido nucleico FPC sustancialmente no fragmentada y la sonda que contenía la molécula de ácido nucleico FPC fragmentada específica para HER2 se formularon ambas en tampón de hibridación y se utilizaron en el ensayo de hibridación *in situ* descrito en el Ejemplo 3. Como se muestra en la FIG. 8, el ADN genómico HER2 diana se detectó con cada una de las formulaciones de la sonda (paneles izquierdos de las FIG. 8A-8D) y ninguna de las formulaciones de la sonda tñieron sustancialmente las secciones de tejido HER2-negativas (paneles de la derecha de las FIG. 8A-8D).

25 La fuerza de la señal detectada tenía una relación inversa con la longitud de la sonda, con la sonda sustancialmente no fragmentada proporcionando la señal más fuerte (véase el panel izquierdo de la FIG. 8A en comparación con los paneles izquierdos de las FIG. 8B-D, cada uno de ellos teñidos con unos fragmentos de sonda progresivamente menores). Este resultado es contrario a la opinión convencional (véase, p. ej., Angerer y Angerer, *Nuc. Acids Res.*, 9:2819, 1981), pero se puede explicar, por ejemplo, por la formación de una red de sonda (véase la FIG. 2B) que es única para la estructura de las sondas de ácido nucleico divulgadas que incluyen moléculas de ácido nucleico FPC.

#### **Ejemplo 5: Sondas de genoma viral y detección de virus utilizando las mismas**

Este ejemplo describe métodos utilizados para generar una sonda que incluye moléculas de ácido nucleico FPC heterogéneas de una secuencia diana del genoma viral. Se pueden emplear métodos similares para generar una sonda, incluyendo moléculas de ácido nucleico FPC heterogéneas de otras secuencias de ácido nucleico diana viral.

35 La detección de virus en muestras biológicas tiene muchas aplicaciones en el campo de la investigación y clínico. El genoma viral es una diana para la detección en una muestra biológica y se conocen muchas sondas de ácido nucleico complementarias de parte o la totalidad de los genomas de diversos virus. Un ensayo de detección que utiliza una sonda de ácido nucleico complementaria de la totalidad o prácticamente de la totalidad de un genoma viral completo típicamente será más sensible que un ensayo que emplea una sonda específica para una porción del genoma viral, porque se cubre una mayor parte de la diana con fracciones detectables. Por otra parte, una sonda de ácido nucleico complementaria de la totalidad o de prácticamente la totalidad de un genoma viral completo puede tener una infectividad residual, que, dependiendo de la naturaleza del virus, puede suponer un riesgo para el usuario. En un ejemplo, una molécula de ácido nucleico FPC divulgada y las sondas que contienen estas moléculas pueden contener la totalidad o prácticamente la totalidad del contenido de un genoma viral completo, pero, dado que los segmentos del genoma viral contenido en la molécula de ácido nucleico FPC están combinados en una orientación y un orden aleatorios, existen pocas posibilidades o ninguna de que una sonda específica para un virus fabricada con la molécula de ácido nucleico FPC sea o resulte inefectiva.

#### **A Sonda de ADN del genoma HPV16**

50 El ADN de HPV16 en el vector pGEM2 se obtuvo de Ventana Medical Systems. Alternativamente, el ADN del genoma HPV16 de pBluescript SK" se puede obtener de ATCC™ (Cat. Nº 4511.3). El ADN de HPV 16 se liberó del vector pGEM2 mediante digestión con BamHI. Después de la electroforesis en gel de agarosa, el ADN de HPV se separó del gel y se purificó. El ADN viral purificado se sometió a una doble digestión con AluI y DpnI, para ser posteriormente incubado con la polimerasa de ADN T4 en presencia de dATP, dTTP, dCTP y dGTP para obtener un ADN con extremos romos. La mezcla de la reacción de polimerasa se extrajo con fenol y los fragmentos de restricción de ADN de extremos romos se precipitaron con etanol. Se resuspendieron 20 ug de los fragmentos de restricción del ADN precipitado en un pequeño volumen (20 ug) de tampón de unión y se unieron utilizando ligasa de

ADN T4 (durante aproximadamente 24 horas a temperatura ambiente) para generar moléculas de ácido nucleico FPC específicas para los ácidos nucleicos HPV 16.

- 5 Las moléculas de ácido nucleico FPC específicas para los ácidos nucleicos HPV 16 se sometieron a una amplificación secuencial repetida, utilizando cebadores de hexámeros al azar (terminados en 2-fosforotioatos en el 3' extremo) para generar cantidades en miligramos de ADN de HPV 16 permutado. El producto de esta amplificación fue etiquetado con DNP utilizando el juego de etiquetado Label IT™ DNP (MIRUS, Madison, WI, EE.UU.) de conformidad con las instrucciones del fabricante. Las moléculas de ácido nucleico FPC etiquetadas resultantes específicas para ácidos nucleicos de HPV16 se utilizaron como sonda en secciones de xenógrafo de la célula C33A HPV16-negativa o Caski HPV16-positiva fijadas en formalina y embebidas en parafina sobre una Ventana Benchmark™ XT utilizando un protocolo de hibridación *in situ* estándar proporcionado por el fabricante (véase la Tabla 3). La sonda que contenía las moléculas de ácido nucleico FPC específicas para ácidos nucleicos de HPV 16 se detectó con un anticuerpo anti-DNP murino, anticuerpo anti-murino de cabra de un conjugado fosfatasa-alciano y los correspondientes reactivos de detección CISH (p. ej., sustratos cromogénicos NBT/BCIP)
- 10

Tabla 3 Ejemplo de protocolo de teñido CISH automatizado

1	***** Seleccionar EZ Prep *****
2	***** Iniciar pasos programados *****
3	***** Mezcladores desactivados*****
4	Calentar portaobjetos a *65 ° C* e incubar durante *12 minutos*
5	Calentar portaobjetos a *75 ° C* e incubar durante *4 minutos*
6	Enjuagar portaobjetos con tampón de reacción
7	Aplicar 600ul de tampón de reacción
8	Aplicar cubreobjetos
9	Desactivar calentador de portaobjetos
10	***** Mezcladores activados*****
11	Calentar portaobjetos a *37 ° C* e incubar durante *4 minutos*
12	Enjuagar portaobjetos
13	Aplicar 900ul de tampón de enjuague
14	Aplicar cubreobjetos
15	Aplicar Acondicionador de la célula #2
16	Calentar portaobjetos a *90 ° C* e incubar durante *16 minutos*
17	Aplicar 900ul de tampón de reacción
18	Aplicar cubreobjetos
19	Desactivar calentador de portaobjetos
20	Enjuagar portaobjetos con tampón de reacción
21	Ajustar volumen de portaobjetos con tampón de reacción
22	Aplicar cubreobjetos
23	***** Seleccionar lavado SSC *****
24	Calentar portaobjetos a *37 ° C* e incubar durante *4 minutos*
25	Enjuagar portaobjetos con tampón de reacción
26	Aplicar 300ul de tampón de reacción
27	Aplicar una gota de *1SH-PROTEASA 3*. Aplicar cubreobjetos e incubar durante *4 minutos*
28	Enjuagar portaobjetos con tampón de reacción
29	Aplicar 300ul de tampón de reacción
30	Aplicar cubreobjetos
31	Desactivar calentador de portaobjetos
32	Enjuagar portaobjetos
33	Ajustar volumen de portaobjetos
34	Aplicar cubreobjetos
35	Calentar portaobjetos a *37 ° C* e incubar durante *4 minutos*
36	Enjuagar portaobjetos
37	Ajustar volumen de portaobjetos
38	Aplicar una gota de *EW+HybReady* Aplicar cubreobjetos e incubar durante *4 minutos*
39	Si *lote de referencia* está *seleccionado*
40	Aplicar dos gotas de *HPV III Fam6(C)*, e incubar durante *4 minutos*
41	Si *lote de referencia* está *no seleccionado*
42	Si *lote de prueba* está *seleccionado*
43	[ Sonda 2 ]
44	Aplicar dos gotas de *SONDA 2*, e incubar durante *4 minutos*
45	Si *lote de prueba* está *no seleccionado*
46	Si *control negativo 1SH* está "seleccionado"

ES 2 372 427 T3

47	Aplicar una gota de *ISH NEG CTRL*, e incubar durante *4 minutos*
48	Calentar portaobjetos a *95 ° C* e incubar durante *12 minutos*
49	Calentar portaobjetos a *52 ° C* e incubar durante *4 minutos*
50	Aplicar cubreobjetos corto
51	Incubar durante *2 horas*
52	Enjuagar portaobjetos
53	Aplicar 900ul de tampón de enjuague
54	Aplicar cubreobjetos
55	Calentar portaobjetos a *72 ° C* e incubar durante *8 minutos*
56	Enjuagar portaobjetos
57	Aplicar 900ul de tampón de enjuague
58	Aplicar cubreobjetos
59	Calentar portaobjetos a *72 ° C* e incubar durante *8 minutos*
60	Enjuagar portaobjetos
61	Aplicar 900ul de tampón de enjuague
62	Aplicar cubreobjetos
63	Calentar portaobjetos a *72 ° C* e incubar durante *8 minutos*
64	Desactivar calentador de portaobjetos
65	Enjuagar portaobjetos con tampón de reacción
66	Ajustar volumen de portaobjetos con tampón de reacción
67	Aplicar cubreobjetos
68	***** Sincronización del procedimiento *****
69	Calentar portaobjetos a *37 ° C* e incubar durante *4 minutos*
70	Enjuagar portaobjetos con tampón de reacción
71	Ajustar volumen de portaobjetos con tampón de reacción
72	Aplicar una gota de *iVIEW + Anti-DNP* Aplicar cubreobjetos e incubar durante *20 minutos*
73	Enjuagar portaobjetos con tampón de reacción
74	Ajustar volumen de portaobjetos con tampón de reacción
75	Aplicar una gota de *iVIEW + Amp* Aplicar cubreobjetos e incubar durante *8 minutos*
76	Enjuagar portaobjetos con tampón de reacción
77	Ajustar volumen de portaobjetos con tampón de reacción
78	Aplicar una gota de *iVIEW+ Biotin-Ig* Aplicar cubreobjetos e incubar durante *12 minutos*
79	Enjuagar portaobjetos con tampón de reacción
80	Ajustar volumen de portaobjetos con tampón de reacción
81	Aplicar una gota de *iVIEW+ SA-AP* Aplicar cubreobjetos e incubar durante *8 minutos*
82	Enjuagar portaobjetos con tampón de reacción
83	Aplicar 900ul de tampón de reacción
84.	Aplicar cubreobjetos
85	Enjuagar portaobjetos con tampón de reacción
86	Ajustar volumen de portaobjetos con tampón de reacción
87	Aplicar una gota de *iVIEW+Enhancer* Aplicar cubreobjetos e incubar durante *4 minutos*
88	Aplicar una gota de *iVIEW+ NBT* y una gota de *iVIEW+ BCIP* e incubar durante *24 minutos*
89	Enjuagar portaobjetos con tampón de reacción
90	Ajustar volumen de portaobjetos con tampón de reacción
91	Aplicar cubreobjetos
92	Desactivar calentador de portaobjetos
93	Calentar portaobjetos a *37 ° C* e incubar durante *4 minutos*
94	Enjuagar portaobjetos con tampón de reacción
95	Ajustar volumen de portaobjetos con tampón de reacción
96	Aplicar una gota de *Red Stain 11*. Aplicar cubreobjetos e incubar durante *4 minutos*
97	Enjuagar portaobjetos con tampón de reacción
98	Ajustar volumen de portaobjetos con tampón de reacción
99	Aplicar cubreobjetos
100	Desactivar calentador de portaobjetos
101	***** Seleccionar lavado opcional *****
102	***** Seleccionar lavado SSC *****
103	***** Iniciar pasos temporizados *****
104	Enjuagar portaobjetos con tampón de reacción
105	Ajustar volumen de portaobjetos con tampón de reacción
106	Aplicar cubreobjetos

Como se muestra en la FIG. 9A, la sonda que contiene moléculas de ácido nucleico FPC específicas para ácidos nucleicos HPV 16 etiquetó fuertemente las secciones del xenógrafo de la célula de Caski, mientras que, como se muestra en la FIG. 9B, esta sonda no mostró ningún etiquetado detectable de las secciones del xenógrafo de la célula C33A HPV16-negativas.

5 B Sonda del genoma del virus BK humano

10 El ADN genómico (cadena prototipo) del poliomavirus BK en pBR322 se obtuvo de ATCC™ (Cat. N° 45024). El ADN del virus BK se liberó del vector mediante digestión con electrofóresis de gel de agarosa BamHI After. El ADN del virus BK se separó del gel y se purificó. El ADN viral purificado se sometió a una doble digestión, con AluI y DpnI, para después ser incubado con polimerasa de ADN T4 en presencia de dATP, dTTP, dCTP y dGTP para obtener un ADN con extremos romos. La mezcla de la reacción de polimerasa se extrajo con fenol y los fragmentos de restricción de ADN de extremos romos se precipitaron con etanol. Se resuspendieron 20 ug de los fragmentos de restricción del ADN precipitado en un pequeño volumen (20 ug) de tampón de unión y se unieron utilizando ligasa de ADN T4 (durante aproximadamente 24 horas a temperatura ambiente) para generar la plantilla de ácido nucleico del virus BK permutado.

15 Las moléculas de ácido nucleico FPC específicas para los ácidos nucleicos del virus BK se sometieron a una amplificación secuencial repetida, utilizando cebadores de hexámeros al azar (terminados en 2-fosforotioatos en el 3' extremo) para generar cantidades en miligramos de ADN del virus BK permutado. El producto de esta amplificación fue etiquetado con DNP utilizando el juego de etiquetado Label IT™ DNP (MIRUS, Madison, WI, EE.UU.) de conformidad con las instrucciones del fabricante. Las moléculas de ácido nucleico FPC etiquetadas resultantes  
20 específicas para ácidos nucleicos del virus BK se utilizaron como sonda en una Ventana Benchmark™ XT utilizando un protocolo de hibridación *in situ* estándar proporcionado por el fabricante (véase la Tabla 3) para teñir secciones de tejido renal humano fijadas en formalina y embebidas en parafina pertenecientes a individuos cuyos riñones habían o no habían estado infectados con el virus BK. En una serie de experimentos, la sonda que contenía las moléculas de ácido nucleico FPC específicas para ácidos nucleicos del virus BK se detectó con un anticuerpo anti-DNP murino, anticuerpo anti-murino de cabra de un conjugado fosfatasa-alcálico y los correspondientes reactivos de detección CISH (p. ej., sustratos cromogénicos NBT/BCIP). También se utilizó una detección alternativa utilizando anticuerpo anti-DNP murino, anticuerpo anti-murino de cabra de un conjugado HRP y los correspondientes reactivos de detección SISH (incluyendo iones de plata, peróxido de hidrógeno e hidroquinona) (véase, p. ej., las Patentes Estadounidenses n° 6.670. 113 y 7.183.072).

30 Como se muestra en la FIG. 10, la sonda que contenía las moléculas de ácido nucleico FPC específicas para ácidos nucleicos del virus BK etiquetaron fuertemente en el tejido renal humano infectado, pero no etiquetaron sustancialmente las secciones renales humanas no infectadas. La detección CISH y SISH fue igualmente útil para detectar la sonda.

35 En vista de las numerosas realizaciones posibles a las que se pueden aplicar los principios de la divulgación, es necesario reconocer que las realizaciones ilustradas son solamente ejemplos de una divulgación y no se deberá interpretar que limitan el alcance de la misma.

LISTA DE SECUENCIAS

<110> Ventana Medical Systems Inc  
 Farrell, Michael  
 <120> MÉTODO PARA LA PRODUCCIÓN DE SONDAS DE ÁCIDO NUCLEICO  
 5 <130> 7668-75979-03  
 <150> 60/841,896  
 <151> 01-09-2006  
 <150> 60/892,571  
 <151> 02-03-2007  
 10 <160> 360  
 <170> Patente en versión 3.3  
  
 <210> 1  
 <211> 24  
 15 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 1  
 20 aaatgattag caaggccaga agtc 24  
  
 <210> 2  
 <211> 21  
 <212> ADN  
 25 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 2  
 30 aactggacaa gctctttggg a 21  
  
 <210> 3  
 <211> 21  
 <212> ADN  
 35 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 3  
 ccagctcaa aaatgaaaa g 21  
  
 40 <210> 4  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 45 <223> Cebador oligonucleótido sintético  
 <400> 4  
 aaccaggcag gcaactatt a 21  
  
 50 <210> 5  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 55 <400> 5  
 ttcaatgacc agactccttg c 21  
  
 60 <210> 6  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 6

5      **taatgcatgg taggaccgaa t      21**  
       <210> 7  
       <211> 21  
       <212> ADN  
       <213> Secuencia artificial  
       <220>  
       <223> Cebador oligonucleótido sintético  
 10     <400> 7  
       **attagccagc atttgtgac c      21**  
       <210> 8  
       <211> 21  
       <212> ADN  
 15     <213> Secuencia artificial  
       <220>  
       <223> Cebador oligonucleótido sintético  
       <400> 8  
 20     **gaaccaacag gatgtgcat a      21**  
       <210> 9  
       <211> 21  
       <212> ADN  
       <213> Secuencia artificial  
 25     <220>  
       <223> Cebador oligonucleótido sintético  
       <400> 9  
       **ttcaaactgc aaaaccctgt g      21**  
 30     <210> 10  
       <211> 21  
       <212> ADN  
       <213> Secuencia artificial  
       <220>  
 35     <223> Cebador oligonucleótido sintético  
       <400> 10  
       **gagagagaca ggcacacatg g      21**  
       <210> 11  
       <211> 21  
       <212> ADN  
       <213> Secuencia artificial  
       <220>  
 45     <223> Cebador oligonucleótido sintético  
       <400> 11  
       **accctgcccc acacatctac t      21**  
       <210> 12  
       <211> 21  
       <212> ADN  
       <213> Secuencia artificial  
       <220>  
       <223> Cebador oligonucleótido sintético  
 50     <400> 12  
       **ccctgctcta gcctttgttc t      21**  
       <210> 13  
       <211> 21  
       <212> ADN  
       <213> Secuencia artificial  
       <220>  
       <223> Cebador oligonucleótido sintético  
 60     <400> 13  
       **ctgaacttcc acccccttta c      21**

<210> 14  
 <211> 21  
 <212> ADN  
 5 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 14  
 gcagtacgtg gcagatgtga a 21  
 10  
 <210> 15  
 <211> 21  
 <212> ADN  
 15 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 15  
 gaggaggtaa agaggtcca g 21  
 20  
 <210> 16  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 25 <223> Cebador oligonucleótido sintético  
 <400> 16  
 ctctcctgcc tttctgactc c 21  
 30  
 <210> 17  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 35 <400> 17  
 gcctcccact tttctcttt c 21  
 40  
 <210> 18  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 18  
 45 ccagagcttt ctccaggtca c 21  
 50  
 <210> 19  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 19  
 55 cccagagtc tgggtgact t 21  
 60  
 <210> 20  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 20  
 atggctgtgg tttgtgatg t 21  
 65 <210> 21

<211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 5 <223> Cebador oligonucleótido sintético  
 <400> 21  
 acaccatgaa ttgttgaagc c 21

<210> 22  
 10 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 15 <400> 22  
 taatgcgttt tcctctctgg g 21

<210> 23  
 20 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 25 <400> 23  
 ggagtgatgt ccaccctgtt c 21

<210> 24  
 30 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 35 <400> 24  
 atgcgtggta gggcatttaa g 21

<210> 25  
 40 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 45 <400> 25  
 atggacaact cactcctccc t 21

<210> 26  
 50 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 55 <400> 26  
 ccactcccca ttgttgtgt t 21

<210> 27  
 60 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 65 <400> 27  
 attccagcca acaataatgg g 21

<210> 28  
 <211> 21  
 <212> ADN

<213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 28  
 5 aagttgtgga cagtcgagac g 21  
  
 <210> 29  
 <211> 21  
 <212> ADN  
 10 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 29  
 15 tcgatgtgac tgtctctcc c 21  
  
 <210> 30  
 <211> 21  
 <212> ADN  
 20 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 30  
 ggacacctct aacctgatc c 21  
  
 25 <210> 31  
 <211> 21  
 <212> ADN  
 <213> Sequenza artificiale  
 <220>  
 30 <223> Cebador oligonucleótido sintético  
 <400> 31  
 agtctccatg gctggcaat g 21  
  
 <210> 32  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 40 <223> Cebador oligonucleótido sintético  
 <400> 32  
 ccatcagaaa cgaattgtcc c 21  
  
 <210> 33  
 <211> 21  
 <212> ADN  
 45 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 33  
 50 tccagggctg taaatcatc a 21  
  
 <210> 34  
 <211> 21  
 <212> ADN  
 55 <213> Sequenza artificiale  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 34  
 60 tggttctttg cccactatgg t 21  
  
 <210> 35  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 65 <220>

<223> Cebador oligonucleótido sintético  
 <400> 35  
 aaactgtgcc tcgctagaca a 21

5

<210> 36  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>

10

<223> Cebador oligonucleótido sintético  
 <400> 36  
 tggaattgag attgctcaa g 21

15

<210> 37  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>

20

<223> Cebador oligonucleótido sintético  
 <400> 37  
 tgaaaaatcc aagaatcagg g 21

25

<210> 38  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>

30

<223> Cebador oligonucleótido sintético  
 <400> 38  
 caccactca ccctcctc t 21

35

<210> 39  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>

40

<223> Cebador oligonucleótido sintético  
 <400> 39  
 aggttctctc cattccagaa c 21

45

<210> 40  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>

50

<223> Cebador oligonucleótido sintético  
 <400> 40  
 catactctc ccagtgctcc t 21

55

<210> 41  
 <211> 21  
 <212> ADN  
 <213> Sequenza artificiale  
 <220>

60

<223> Cebador oligonucleótido sintético  
 <400> 41  
 cagaccagaa cgaggagag t 21

65

<210> 42  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>

<223> Cebador oligonucleótido sintético  
 <400> 42

tttgaggaca tcaccatgac a 21  
 <210> 43  
 <211> 21  
 5 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 43  
 10 cggacactaa gggagatgga t 21  
 <210> 44  
 <211> 21  
 <212> ADN  
 15 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 44  
 20 accacttacc tgaccactgg c 21  
 <210> 45  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 25 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 45  
 cgtttagga ggattcaagc a 21  
 30 <210> 46  
 <211> 21  
 <212> ADN  
 <213> Sequenza artificiale  
 <220>  
 35 <223> Cebador oligonucleótido sintético  
 <400> 46  
 gcgacctgtt ccaaaagtct c 21  
 40 <210> 47  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 45 <400> 47  
 caggtagggag aggaagata a 21  
 <210> 48  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 50 <400> 48  
 ctggaaagag gaaggaggac a 21  
 <210> 49  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 60 <400> 49  
 cttcctccag gtctcatgct t 21

<210> 50  
 <211> 21  
 <212> ADN  
 5 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 50  
 tctccatagc tccaagcaca c 21  
 10  
 <210> 51  
 <211> 21  
 <212> ADN  
 15 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 51  
 ccctacctcc cactctcact c 21  
 20  
 <210> 52  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 25 <223> Cebador oligonucleótido sintético  
 <400> 52  
 ggctaggaaa cgctactga g 21  
 30  
 <210> 53  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 35 <400> 53  
 cagatgaatg ctaagccaa a 21  
 40  
 <210> 54  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 54  
 45 ccacatcagt gggacaaaag a 21  
 50  
 <210> 55  
 <211> 21  
 <212> ADN  
 <213> Sequenza artificiale  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 55  
 cctggctctt tgccaataa t 21  
 55  
 <210> 56  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 60 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 56  
 cgcagagcct gtgttctat c 21  
 65 <210> 57

<211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 5 <223> Cebador oligonucleótido sintético  
 <400> 57  
 ggctgaagtc ctgaaggtca t 21

<210> 58  
 10 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 15 <400> 58  
 caggtccaag aagaggggaag a 21

<210> 59  
 20 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 25 <400> 59  
 tgaacaggag tcaaagctgg a 21

<210> 60  
 30 <211> 21  
 <212> ADN  
 <213> Sequenza artificiale  
 <220>  
 <223> Cebador oligonucleótido sintético  
 35 <400> 60  
 gacacgcaga gacactcagg a 21

<210> 61  
 40 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 61  
 ataagttcag cagagcctcc g 21

45 <210> 62  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 50 <223> Cebador oligonucleótido sintético  
 <400> 62  
 aaacacatct tgcttgggag g 21

55 <210> 63  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 60 <400> 63  
 gtcccttgg aactgcaga t 21

<210> 64  
 65 <211> 21  
 <212> ADN

<213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 64  
 5 ccatctcgct ccctacaag t 21  
  
 <210> 65  
 <211> 21  
 <212> ADN  
 10 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 65  
 15 tagtgactga ggggtgaagg g 21  
  
 <210> 66  
 <211> 21  
 <212> ADN  
 20 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 66  
 cctgtgcaag gttacatcca a 21  
  
 25 <210> 67  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 30 <223> Cebador oligonucleótido sintético  
 <400> 67  
 gcgacctgtt caaaagtct c 21  
  
 <210> 68  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 40 <223> Cebador oligonucleótido sintético  
 <400> 68  
 cgttgtagga ggattcaagc a 21  
  
 <210> 69  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 45 <223> Cebador oligonucleótido sintético  
 <400> 69  
 50 caggtgttg gtagaactg g 21  
  
 <210> 70  
 <211> 21  
 <212> ADN  
 55 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 70  
 60 cggacactaa gggagatgga t 21  
  
 <210> 71  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 65 <220>

<223> Cebador oligonucleótido sintético  
 <400> 71  
 tccagatgga gacacattg c 21

5 <210> 72  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>

10 <223> Cebador oligonucleótido sintético  
 <400> 72  
 aacagtgcag actgcttcag a 21

15 <210> 73  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>

20 <223> Cebador oligonucleótido sintético  
 <400> 73  
 agtgctcctc agaggagtt g 21

25 <210> 74  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>

30 <223> Cebador oligonucleótido sintético  
 <400> 74  
 tagactgact ctaccacgc c 21

35 <210> 75  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>

40 <223> Cebador oligonucleótido sintético  
 <400> 75  
 caccacttca ccctcctc t 21

45 <210> 76  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>

50 <223> Cebador oligonucleótido sintético  
 <400> 76  
 ccaagaagga agctgagttg c 21

55 <210> 77  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>

60 <223> Cebador oligonucleótido sintético  
 <400> 77  
 actatggctg tgactcccca c 21

65 <210> 78  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>

<223> Cebador oligonucleótido sintético  
 <400> 78

ggtaccagca aactgtgcct c 21  
 <210> 79  
 <211> 21  
 5 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 79  
 10 ggttctttgc ccactatggt c 21  
 <210> 80  
 <211> 21  
 <212> ADN  
 15 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 80  
 20 tccagggctg taaatcatc a 21  
 <210> 81  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 25 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 81  
 agtggggttc agtgaatgaa a 21  
 30 <210> 82  
 <211> 20  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 35 <223> Cebador oligonucleótido sintético  
 <400> 82  
 tccatggctg gtcaatgatt 20  
 <210> 83  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 45 <400> 83  
 atggcgtcca cagtagcttt t 21  
 <210> 84  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 84  
 55 cctttc gatg tgactgtctc c 21  
 <210> 85  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 85  
 60 cgaagttgtg gacagtcgag a 21

<210> 86  
 <211> 21  
 <212> ADN  
 5 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 86  
 acactgctcc ctgagtcact g 21  
 10  
 <210> 87  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 15 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 87  
 ccactcccca ttgtgtgtg t 21  
 20  
 <210> 88  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 25 <223> Cebador oligonucleótido sintético  
 <400> 88  
 gagaaggaag gagagagctg c 21  
 30  
 <210> 89  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 35 <400> 89  
 atgcgtggta gggcattta g 21  
 40  
 <210> 90  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 90  
 45 agtgatgtcc acctgttcc t 21  
 50  
 <210> 91  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 91  
 gtaatgcggt ttctctctg g 21  
 55  
 <210> 92  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 92  
 60 acaccatgaa ttgtgaagc c 21  
 65 <210> 93

<211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 5 <223> Cebador oligonucleótido sintético  
 <400> 93  
 gtgtgtggtc tcccatacc t 21

<210> 94  
 10 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 15 <400> 94  
 ctctaccacc tgaggcctt g 21

<210> 95  
 20 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 25 <400> 95  
 tccagggtgg acctctatc a 21

<210> 96  
 30 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 35 <400> 96  
 gcacacaact ggtccgta a 21

<210> 97  
 40 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 45 <400> 97  
 aagtcctgct cactcatgct g 21

<210> 98  
 50 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 55 <400> 98  
 aataaaagga aatggtgggg c 21

<210> 99  
 60 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 65 <400> 99  
 cattcctcag ccacagtgac a 21

<210> 100  
 <211> 21  
 <212> ADN

<213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 100  
 5 ccctgctcta gcctttgttc t 21  
  
 <210> 101  
 <211> 21  
 <212> ADN  
 10 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 101  
 15 cccacacatc tactggagga a 21  
  
 <210> 102  
 <211> 21  
 <212> ADN  
 20 <213> Sequenza artificiale  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 102  
 aaggagagag acaggcacac a 21  
  
 25 <210> 103  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 30 <223> Cebador oligonucleótido sintético  
 <400> 103  
 ccttgaagaa ccaacaggat g 21  
  
 <210> 104  
 <211> 21  
 <212> ADN  
 35 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 104  
 40 tgtgacctcc tattccagt g c 21  
  
 <210> 105  
 <211> 21  
 <212> ADN  
 45 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 105  
 50 taatgcatgg taggaccgaa t 21  
  
 <210> 106  
 <211> 21  
 <212> ADN  
 55 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 106  
 60 tctccacca actttcaatg a 21  
  
 <210> 107  
 <211> 21  
 <212> ADN  
 65 <213> Secuencia artificial  
 <220>

<223> Cebador oligonucleótido sintético  
 <400> 107  
 gaggggatct ccctaaactg a 21

5

<210> 108  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>

10

<223> Cebador oligonucleótido sintético  
 <400> 108  
 ggaacattca tctggttcc a 21

15

<210> 109  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>

20

<223> Cebador oligonucleótido sintético  
 <400> 109  
 aactggacaa gctcttggg a 21

25

<210> 110  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>

30

<223> Cebador oligonucleótido sintético  
 <400> 110  
 agtctgttg cccaatact g 21

35

<210> 111  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>

40

<223> Cebador oligonucleótido sintético  
 <400> 111  
 aaaaagcaga agcactgcaa g 21

45

<210> 112  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>

50

<223> Cebador oligonucleótido sintético  
 <400> 112  
 tctgaagaaa tgtacggcag c 21

55

<210> 113  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>

60

<223> Cebador oligonucleótido sintético  
 <400> 113  
 atgaagaaa ggttcaccc a 21

65

<210> 114  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>

<223> Cebador oligonucleótido sintético  
 <400> 114

ccaatcaaga ggaaagttgg a 21  
 <210> 115  
 <211> 21  
 5 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 115  
 10 tccttacaag gtgttgaggg c 21  
 <210> 116  
 <211> 21  
 <212> ADN  
 15 <213> Sequenza artificiale  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 116  
 20 ttggattga cctcatgcac t 21  
 <210> 117  
 <211> 21  
 <212> ADN  
 <213> Sequenza artificiale  
 25 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 117  
 cagaggcagg gaaaagatgt c 21  
 30 <210> 118  
 <211> 23  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 35 <223> Cebador oligonucleótido sintético  
 <400> 118  
 ctggccatca gtaaatcaca tca 23  
 <210> 119  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 119  
 45 attagccgaa gacagaggga a 21  
 <210> 120  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 120  
 55 tgttgaaaa cacaagggca c 21  
 <210> 121  
 <211> 21  
 <212> ADN  
 <213> Sequenza artificiale  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 121  
 60 tattcccaga acttggcac a 21

<210> 122  
 <211> 21  
 <212> ADN  
 5 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 122  
 cccgttgac acagtacaag c 21  
 10  
 <210> 123  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 15 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 123  
 taaatgtgca actcaggcag g 21  
 20  
 <210> 124  
 <211> 23  
 <212> ADN  
 <213> Sequenza artificiale  
 <220>  
 25 <223> Cebador oligonucleótido sintético  
 <400> 124  
 cccatcgttt cagtgttctt ctt 23  
 30  
 <210> 125  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 35 <400> 125  
 tgcaaggga tttgaaatga a 21  
 40  
 <210> 126  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 126  
 45 atagattgca gttgttggc c 21  
 50  
 <210> 127  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 127  
 aggtgtctt aagcctgagc c 21  
 55  
 <210> 128  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 60 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 128  
 tacacccttc atctcctggc t 21  
 65 <210> 129

<211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 5 <223> Cebador oligonucleótido sintético  
 <400> 129  
 tcaaaagggc atcagtgagg a 21

<210> 130  
 10 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 15 <400> 130  
 ctggcacagg atgtggtcat t 21

<210> 131  
 <211> 22  
 20 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 131  
 25 gagtccatgg tgaccacatt tt 22

<210> 132  
 <211> 21  
 <212> ADN  
 30 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 132  
 35 atcacagcca tcatggtaa c 21

<210> 133  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 40 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 133  
 tgccacattt acagtcccag g 21

<210> 134  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 50 <223> Cebador oligonucleótido sintético  
 <400> 134  
 ccatttttct aatccatgcc c 21

<210> 135  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 60 <400> 135  
 tgcccctaca ccaacatac c 21

<210> 136  
 <211> 21  
 65 <212> ADN

<213> Sequenza artificiale  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 136  
 5 cttctgtctg ttatggtcgg g 21  
  
 <210> 137  
 <211> 21  
 <212> ADN  
 10 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 137  
 15 tgggggtcac attgattcat a 21  
  
 <210> 138  
 <211> 21  
 <212> ADN  
 20 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 138  
 ggtgttgga gaagatgttg a 21  
  
 25 <210> 139  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 30 <223> Cebador oligonucleótido sintético  
 <400> 139  
 ccttgaaat cactccttg c 21  
  
 <210> 140  
 <211> 21  
 <212> ADN  
 35 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 140  
 40 tgatggttc actgcatctg g 21  
  
 <210> 141  
 <211> 21  
 <212> ADN  
 45 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 141  
 50 gcttctcata agccatgcac a 21  
  
 <210> 142  
 <211> 21  
 <212> ADN  
 55 <213> Sequenza artificiale  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 142  
 atgatgctct gggatgtgaa a 21  
  
 60 <210> 143  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 65 <220>

<223> Cebador oligonucleótido sintético  
 <400> 143  
 gtggtgatta attcttgggg g 21

5

<210> 144  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>

10

<223> Cebador oligonucleótido sintético  
 <400> 144  
 ccttcagacc tgcaaaactc c 21

15

<210> 145  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>

20

<223> Cebador oligonucleótido sintético  
 <400> 145  
 ctatcagctc agcagcaagg a 21

25

<210> 146  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>

30

<223> Cebador oligonucleótido sintético  
 <400> 146  
 gcaagggcca aataaccaag t 21

35

<210> 147  
 <211> 21  
 <212> ADN  
 <213> Sequenza artificiale  
 <220>

40

<223> Cebador oligonucleótido sintético  
 <400> 147  
 atcctttcat gttccttggc t 21

45

<210> 148  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>

50

<223> Cebador oligonucleótido sintético  
 <400> 148  
 tggcttctgt ttctgagtg g 21

55

<210> 149  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>

60

<223> Cebador oligonucleótido sintético  
 <400> 149  
 ccttaatctg cctccagctc a 21

65

<210> 150  
 <211> 21  
 <212> ADN  
 <213> Sequenza artificiale  
 <220>

<223> Cebador oligonucleótido sintético  
 <400> 150

taagttgggg tttggggaga t 21  
 <210> 151  
 <211> 21  
 5 <212> ADN  
 <213> Sequenza artificiale  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 151  
 10 cgcaccaaaa ttctatcatc c 21  
 <210> 152  
 <211> 21  
 <212> ADN  
 15 <213> Sequenza artificiale  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 152  
 20 ctttctgcc ttaatgctgg g 21  
 <210> 153  
 <211> 21  
 <212> ADN  
 <213> Sequenza artificiale  
 25 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 153  
 gaagaaggtg atctgggaac g 21  
 30 <210> 154  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 35 <223> Cebador oligonucleótido sintético  
 <400> 154  
 tcactctgtt ggttgtgtcg g 21  
 40 <210> 155  
 <211> 21  
 <212> ADN  
 <213> Sequenza artificiale  
 <220>  
 <223> Cebador oligonucleótido sintético  
 45 <400> 155  
 ggaccagacc agatgtagg a 21  
 <210> 156  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 55 <400> 156  
 tcaaacatcc tacagcgaag c 21  
 <210> 157  
 <211> 21  
 <212> ADN  
 <213> Sequenza artificiale  
 60 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 157  
 cctctccctc tcctctcatc a 21

<210> 158  
 <211> 21  
 <212> ADN  
 5 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 158  
 tctcaagtgc atcaccagct c 21

10 <210> 159  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 15 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 159  
 taagcctacc acaccagcca c 21

20 <210> 160  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 25 <223> Cebador oligonucleótido sintético  
 <400> 160  
 agagcaccca gcaggtacag a 21

30 <210> 161  
 <211> 21  
 <212> ADN  
 <213> Sequenza artificiale  
 <220>  
 <223> Cebador oligonucleótido sintético  
 35 <400> 161  
 gggcaaggag aatgtttag g 21

40 <210> 162  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 162  
 45 tgtcgttagt catctgttc c 21

50 <210> 163  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 163  
 gacacattc ctccatctgg c 21

55 <210> 164  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 60 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 164  
 gtctctgctc acccactca t 21

65 <210> 165

<211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 5 <223> Cebador oligonucleótido sintético  
 <400> 165  
 ttgaggacaa aggtctcagg c 21

<210> 166  
 10 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 15 <400> 166  
 ttacctgtta gggctccaac g 21

<210> 167  
 20 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 25 <400> 167  
 ccagcggttt gatgagattg t 21

<210> 168  
 30 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 35 <400> 168  
 aaataaatcc acagccgagg t 21

<210> 169  
 40 <211> 21  
 <212> ADN  
 <213> Sequenza artificiale  
 <220>  
 <223> Cebador oligonucleótido sintético  
 45 <400> 169  
 caccatcaca tcctcaaag c 21

<210> 170  
 50 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 170  
 cagaagaaag aggcagcagc a 21

<210> 171  
 55 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 60 <400> 171  
 ggacagggct gaacgaata a 21

<210> 172  
 65 <211> 21  
 <212> ADN

<213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 172  
 5 cacatgcaca tccatgctct c 21  
  
 <210> 173  
 <211> 21  
 <212> ADN  
 10 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 173  
 15 agctggtgat ggacacatgg t 21  
  
 <210> 174  
 <211> 21  
 <212> ADN  
 20 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 174  
 ttaaccacc tgaacctgct c 21  
  
 25 <210> 175  
 <211> 21  
 <212> ADN  
 <213> Sequenza artificiale  
 <220>  
 30 <223> Cebador oligonucleótido sintético  
 <400> 175  
 agaacttcc tctcctccc t 21  
  
 <210> 176  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 40 <223> Cebador oligonucleótido sintético  
 <400> 176  
 tgtgagggaa atctaccttc g 21  
  
 <210> 177  
 <211> 21  
 <212> ADN  
 <213> Sequenza artificiale  
 <220>  
 45 <223> Cebador oligonucleótido sintético  
 <400> 177  
 50 gatctcaggg tcttctctgg g 21  
  
 <210> 178  
 <211> 23  
 <212> ADN  
 55 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 178  
 60 tgtttatta tgtggcaggt tgg 23  
  
 <210> 179  
 <211> 24  
 <212> ADN  
 <213>Secuencia artificial  
 65 <220>

<223> Cebador oligonucleótido sintético  
 <400> 179  
 ggggaaaaat cagaaaacta cact      24

5    <210> 180  
      <211> 24  
      <212> ADN  
      <213>Secuencia artificial  
      <220>

10   <223> Cebador oligonucleótido sintético  
      <400> 180  
      gaacctgcct ctgtctttga tacc      24

15   <210> 181  
      <211> 21  
      <212> ADN  
      <213>Secuencia artificial  
      <220>

20   <223> Cebador oligonucleótido sintético  
      <400> 181  
      tgtgcatcag ctatccaaca a          21

25   <210> 182  
      <211> 21  
      <212> ADN  
      <213> Sequenza artificiale  
      <220>

30   <223> Cebador oligonucleótido sintético  
      <400> 182  
      ccacgtccag gctgtttatt t          21

35   <210> 183  
      <211> 21  
      <212> ADN  
      <213>Secuencia artificial  
      <220>

40   <223> Cebador oligonucleótido sintético  
      <400> 183  
      ccaaggcact gtttttgaa g          21

45   <210> 184  
      <211> 21  
      <212> ADN  
      <213>Secuencia artificial  
      <220>

50   <223> Cebador oligonucleótido sintético  
      <400> 184  
      tattagggtg gtgggtcttc c          21

55   <210> 185  
      <211> 21  
      <212> ADN  
      <213>Secuencia artificial  
      <220>

60   <223> Cebador oligonucleótido sintético  
      <400> 185  
      caagctgaca gaatggagag g          21

65   <210> 186  
      <211> 21  
      <212> ADN  
      <213>Secuencia artificial  
      <220>

     <223> Cebador oligonucleótido sintético  
      <400> 186

gattatgcag taaccacaag g 21

<210> 187  
 <211> 21  
 5 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 187  
 10 tggaaactct gggacactca a 21

<210> 188  
 <211> 21  
 <212> ADN  
 15 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 188  
 20 aatgttacct ttgaggggtg g 21

<210> 189  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 25 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 189  
 gctccaggtc ttcctctct c 21

<210> 190  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 30 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 190  
 35 gtccagtctg caacatcaa c 21

<210> 191  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 40 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 191  
 45 tggttccctt ctgattcag c 21

<210> 192  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 50 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 192  
 55 gttgctggga gtctgtgtc t 21

<210> 193  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 60 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 193  
 tacaacatag aggggaggca c 21

<210> 194  
 <211> 21  
 <212> ADN  
 5 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 194  
 agcacaagt tgctcacagg a 21  
 10  
 <210> 195  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 15 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 195  
 acctgtccta tccaccatt t 21  
 20  
 <210> 196  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 25 <223> Cebador oligonucleótido sintético  
 <400> 196  
 aggagtagca ggacaccgt t 21  
 30  
 <210> 197  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 35 <400> 197  
 gcccaccac tttctttc t 21  
 40  
 <210> 198  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 198  
 45 acaagaaggt tttgaggctc c 21  
 50  
 <210> 199  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 199  
 aggttgagg gagtcatatc t 21  
 55  
 <210> 200  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 200  
 60 gggagagttg gtccccttt a 21  
 65 <210> 201

<211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 5 <223> Cebador oligonucleótido sintético  
 <400> 201  
 agctgggtct gaatccaggt a 21

<210> 202  
 10 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 15 <400> 202  
 ctctgggtctc ccatctgctt t 21

<210> 203  
 20 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 25 <400> 203  
 gttgaagaa caagcagct c 21

<210> 204  
 30 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 35 <400> 204  
 agtgggagag ggatagtggc t 21

<210> 205  
 40 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 205  
 gcagtacctg caactgggtg a 21

45 <210> 206  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 50 <223> Cebador oligonucleótido sintético  
 <400> 206  
 aatgcacaca ggtggacaga t 21

55 <210> 207  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 60 <400> 207  
 ctccaactgc attccaaca g 21

<210> 208  
 65 <211> 21  
 <212> ADN

<213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 208  
 5 aacttattcc ttggaccgct g 21

<210> 209  
 <211> 21  
 <212> ADN  
 10 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 209  
 15 taccaagagg ggagacagag g 21

<210> 210  
 <211> 21  
 <212> ADN  
 20 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 210  
 gcaccaaagt ctctccctc t 21

25 <210> 211  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 30 <223> Cebador oligonucleótido sintético  
 <400> 211  
 ggtaactcc aggggacact a 21

35 <210> 212  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 40 <223> Cebador oligonucleótido sintético  
 <400> 212  
 ggtggtgaca gtaaacagcc c 21

45 <210> 213  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 50 <223> Cebador oligonucleótido sintético  
 <400> 213  
 cccattcat gctctctt t 21

55 <210> 214  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 214  
 60 tgttgcttca gcatgtcaag a 21

65 <210> 215  
 <211> 21  
 <212> ADN  
 <213> Sequenza artificiale  
 <220>

<223> Cebador oligonucleótido sintético  
 <400> 215  
 agggtttagc acttgtggag g 21

5

<210> 216  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>

10

<223> Cebador oligonucleótido sintético  
 <400> 216  
 agatgtaga tggggggg g 21

15

<210> 217  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>

20

<223> Cebador oligonucleótido sintético  
 <400> 217  
 tccacctctg tctccctt c 21

25

<210> 218  
 <211> 21  
 <212> ADN  
 <213> Sequenza artificiale  
 <220>

30

<223> Cebador oligonucleótido sintético  
 <400> 218  
 tgggtctctg tgagtgaag g 21

35

<210> 219  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>

40

<223> Cebador oligonucleótido sintético  
 <400> 219  
 gtcctaaag cctgttct g 21

45

<210> 220  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>

50

<223> Cebador oligonucleótido sintético  
 <400> 220  
 tgacctggc cttccttagt t 21

55

<210> 221  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>

60

<223> Cebador oligonucleótido sintético  
 <400> 221  
 atacctacca gccaggctca g 21

65

<210> 222  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>

<223> Cebador oligonucleótido sintético  
 <400> 222

aatctttaga cccctcacc c 21  
 <210> 223  
 <211> 21  
 5 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 223  
 10 ttaatacaaa ggtcccccag g 21  
 <210> 224  
 <211> 21  
 <212> ADN  
 15 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 224  
 20 atggggaaga gtggggtcta t 21  
 <210> 225  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 25 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 225  
 gtctgttcc acagaccat c 21  
 30 <210> 226  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 35 <223> Cebador oligonucleótido sintético  
 <400> 226  
 tatgctgcca aaagagaacc c 21  
 40 <210> 227  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 45 <223> Cebador oligonucleótido sintético  
 <400> 227  
 gtctgccaag ggaacatca t 21  
 <210> 228  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 50 <223> Cebador oligonucleótido sintético  
 <400> 228  
 55 cctgtgttc ctccagtct t 21  
 <210> 229  
 <211> 21  
 <212> ADN  
 60 <213> Sequenza artificiale  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 229  
 ccagcaccag ggagtagttt g 21

<210> 230  
 <211> 21  
 <212> ADN  
 5 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 230  
 gagtgcaggg gctgatctct a 21  
 10  
 <210> 231  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 15 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 231  
 ctctgactga ctggcgagat g 21  
 20  
 <210> 232  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 25 <223> Cebador oligonucleótido sintético  
 <400> 232  
 ccctggagaa gtgggagtgt a 21  
 30  
 <210> 233  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 35 <400> 233  
 cccaacacc aggtactagg c 21  
 40  
 <210> 234  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 234  
 45 cggtacgaag aaaaccagga g 21  
 50  
 <210> 235  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 235  
 tctcaccct ctccactgt t 21  
 55  
 <210> 236  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 236  
 60 ctgcaggtga gactcagcaa t 21  
 65 <210> 237

<211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 5 <223> Cebador oligonucleótido sintético  
 <400> 237  
 gaaaggaag caggaaagag g 21

<210> 238  
 10 <211> 21  
 <212> ADN  
 <213> Sequenza artificiale  
 <220>  
 <223> Cebador oligonucleótido sintético  
 15 <400> 238  
 cccttctat cttctccac c 21

<210> 239  
 <211> 21  
 20 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 239  
 25 aatgagcatg gagaatcgtg t 21

<210> 240  
 <211> 21  
 <212> ADN  
 30 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 240  
 35 gaatgggact cctgagagct g 21

<210> 241  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 40 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 241  
 taacacattc aggatggacg c 21

<210> 242  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 50 <223> Cebador oligonucleótido sintético  
 <400> 242  
 agagcactga ccctccttag g 21

<210> 243  
 <211> 21  
 <212> ADN  
 <213> Sequenza artificiale  
 <220>  
 <223> Cebador oligonucleótido sintético  
 60 <400> 243  
 ccactcacga agatgtcgaa g 21

<210> 244  
 <211> 21  
 65 <212> ADN

<213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 244  
 5 ccttaagagg cagccagact g 21  
  
 <210> 245  
 <211> 21  
 <212> ADN  
 10 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 245  
 15 cacagcctga ctggacaaaa g 21  
  
 <210> 246  
 <211> 21  
 <212> ADN  
 20 <213> Sequenza artificiale  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 246  
 ctgggagagg cagagattca t 21  
  
 25 <210> 247  
 <211> 21  
 <212> ADN  
 <213> Sequenza artificiale  
 <220>  
 30 <223> Cebador oligonucleótido sintético  
 <400> 247  
 ctctttcctg attcgaggtg g 21  
  
 <210> 248  
 <211> 21  
 <212> ADN  
 35 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 248  
 40 atacctacca gccaggctca g 21  
  
 <210> 249  
 <211> 21  
 <212> ADN  
 45 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 249  
 50 tgacctggc cttccttagt t 21  
  
 <210> 250  
 <211> 21  
 <212> ADN  
 55 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 250  
 60 tgttcctggg cacatttta a 21  
  
 <210> 251  
 <211> 21  
 <212> ADN  
 65 <213>Secuencia artificial  
 <220>

<223> Cebador oligonucleótido sintético  
 <400> 251  
 ccaaatctga ggaaaggtg a 21

5

<210> 252  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>

10

<223> Cebador oligonucleótido sintético  
 <400> 252  
 catccacctc tgtctccctt c 21

15

<210> 253  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>

20

<223> Cebador oligonucleótido sintético  
 <400> 253  
 agatgttggg gtggtagaag g 21

25

<210> 254  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>

30

<223> Cebador oligonucleótido sintético  
 <400> 254  
 ggtggagaca gggtttagca c 21

35

<210> 255  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>

40

<223> Cebador oligonucleótido sintético  
 <400> 255  
 atctaccctg acccatctgg t 21

45

<210> 256  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>

50

<223> Cebador oligonucleótido sintético  
 <400> 256  
 tccaacttac aggcctctt g 21

55

<210> 257  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>

60

<223> Cebador oligonucleótido sintético  
 <400> 257  
 tttgctcatt aaccatcag g 21

65

<210> 258  
 <211> 22  
 <212> ADN  
 <213>Secuencia artificial  
 <220>

<223> Cebador oligonucleótido sintético  
 <400> 258

aggggagact ccctatttg tc 22  
 <210> 259  
 <211> 21  
 5 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 259  
 10 acctacttca ccagccagct t 21  
 <210> 260  
 <211> 21  
 <212> ADN  
 15 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 260  
 20 attacctca agagccccag t 21  
 <210> 261  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 25 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 261  
 agggtaggg tgaggatcag g 21  
 30 <210> 262  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 35 <223> Cebador oligonucleótido sintético  
 <400> 262  
 ccaactgcat tccaacaagt c 21  
 <210> 263  
 40 <211> 21  
 <212> ADN  
 <213> Sequenza artificiale  
 <220>  
 <223> Cebador oligonucleótido sintético  
 45 <400> 263  
 aatgcacaca ggtggacaga t 21  
 <210> 264  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 50 <400> 264  
 gctcagcctg acagctcagt a 21  
 <210> 265  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 55 <400> 265  
 gctggcaaga gagcacaaga t 21

<210> 266  
 <211> 21  
 <212> ADN  
 5 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 266  
 gctgttgaaa gaacaaggca g 21  
 10  
 <210> 267  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 15 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 267  
 ctggctccc atctgcttc t 21  
 20  
 <210> 268  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 25 <223> Cebador oligonucleótido sintético  
 <400> 268  
 agctgggtct gaatccaggt a 21  
 30  
 <210> 269  
 <211> 21  
 <212> ADN  
 <213> Sequenza artificiale  
 <220>  
 <223> Cebador oligonucleótido sintético  
 35 <400> 269  
 gtggggatag aactgctagg g 21  
 40  
 <210> 270  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 270  
 45 aacccaatg aagagagacc a 21  
 50  
 <210> 271  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 271  
 gttcctcaag agtggcttg g 21  
 55  
 <210> 272  
 <211> 21  
 <212> ADN  
 <213> Sequenza artificiale  
 60 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 272  
 ccaacatgag ttccctcca t 21  
 65 <210> 273

<211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 5 <223> Cebador oligonucleótido sintético  
 <400> 273  
 cacctgtcct atccacccat t 21

<210> 274  
 10 <211> 21  
 <212> ADN  
 <213> Sequenza artificiale  
 <220>  
 <223> Cebador oligonucleótido sintético  
 15 <400> 274  
 agagggggagg cacaggactt a 21

<210> 275  
 20 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 25 <400> 275  
 gctgggagtc ctgtgtctca t 21

<210> 276  
 30 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 35 <400> 276  
 tggttccctt ctgattcag c 21

<210> 277  
 40 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 45 <400> 277  
 ctggacatgc tgaagagtg a 21

<210> 278  
 50 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 55 <400> 278  
 ctctcctccc ccaactcaac t 21

<210> 279  
 60 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 65 <400> 279  
 ctctaagac aggcctcaac c 21

<210> 280  
 <211> 21  
 <212> ADN

<213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 280  
 5 atgttacctt tgagggtgg t 21  
  
 <210> 281  
 <211> 21  
 <212> ADN  
 10 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 281  
 15 gaagaagggtg gtggagagga a 21  
  
 <210> 282  
 <211> 21  
 <212> ADN  
 20 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 282  
 cacctcaaga acaggaactg g 21  
  
 25 <210> 283  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 30 <223> Cebador oligonucleótido sintético  
 <400> 283  
 tattagggtg gtgggtcttc c 21  
  
 <210> 284  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 40 <223> Cebador oligonucleótido sintético  
 <400> 284  
 cctctcaag gcaactgttt t 21  
  
 <210> 285  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 285  
 50 gcaccatgct tttttcaa a 21  
  
 <210> 286  
 <211> 21  
 <212> ADN  
 55 <213> Sequenza artificiale  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 286  
 tcaaagggca ctcatctat g 21  
  
 60 <210> 287  
 <211> 23  
 <212> ADN  
 <213> Sequenza artificiale  
 65 <220>

<223> Cebador oligonucleótido sintético  
 <400> 287  
 aacctgcctc tgtctttgat acc 23

5 <210> 288  
 <211> 21  
 <212> ADN  
 <213> Sequenza artificiale  
 <220>

10 <223> Cebador oligonucleótido sintético  
 <400> 288  
 gagggggaaa aatcagaaaa c 21

15 <210> 289  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>

20 <223> Cebador oligonucleótido sintético  
 <400> 289  
 attgagggta gaggagggtg g 21

25 <210> 290  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>

30 <223> Cebador oligonucleótido sintético  
 <400> 290  
 cgaatggcta actcccacaa a 21

35 <210> 291  
 <211> 23  
 <212> ADN  
 <213>Secuencia artificial  
 <220>

40 <223> Cebador oligonucleótido sintético  
 <400> 291  
 gggacctcat gttcttgatg tta 23

45 <210> 292  
 <211> 21  
 <212> ADN  
 <213> Sequenza artificiale  
 <220>

50 <223> Cebador oligonucleótido sintético  
 <400> 292  
 aatagccttg ctagtgggtg c 21

55 <210> 293  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>

60 <223> Cebador oligonucleótido sintético  
 <400> 293  
 catggggagg tacaactttt g 21

65 <210> 294  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>

<223> Cebador oligonucleótido sintético  
 <400> 294

ccccagtaac catgcagaga g 21  
 <210> 295  
 <211> 23  
 5 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 295  
 10 ggcattttca aaattaacag acg 23  
 <210> 296  
 <211> 23  
 15 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 296  
 20 ggcaacatct aaaacttcag cct 23  
 <210> 297  
 <211> 21  
 25 <212> ADN  
 <213> Sequenza artificiale  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 297  
 ttgcttcag ccttgaagta t 21  
 30 <210> 298  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 35 <223> Cebador oligonucleótido sintético  
 <400> 298  
 aggtgtcttt aagcctgagc c 21  
 40 <210> 299  
 <211> 21  
 <212> ADN  
 <213> Sequenza artificiale  
 <220>  
 45 <223> Cebador oligonucleótido sintético  
 <400> 299  
 aaaatcctgt cctggctcat c 21  
 <210> 300  
 <211> 21  
 50 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 300  
 55 tacacccttc atctcctggc t 21  
 <210> 301  
 <211> 21  
 60 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 301  
 aagattggc acttgaaggg a 21

<210> 302  
 <211> 21  
 <212> ADN  
 5 <213> Sequenza artificiale  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 302  
 cacgaacaca cacacacacc c 21

10 <210> 303  
 <211> 24  
 <212> ADN  
 <213> Sequenza artificiale  
 15 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 303  
 catcacttca aacacaagag catt 24

20 <210> 304  
 <211> 21  
 <212> ADN  
 <213> Sequenza artificiale  
 <220>  
 25 <223> Cebador oligonucleótido sintético  
 <400> 304  
 tattagccga agacagaggg a 21

30 <210> 305  
 <211> 21  
 <212> ADN  
 <213> Sequencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 35 <400> 305  
 cctgattgga cattcccat a 21

40 <210> 306  
 <211> 21  
 <212> ADN  
 <213> Sequencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 45 <400> 306  
 acaagctgta aagctctgcc c 21

50 <210> 307  
 <211> 21  
 <212> ADN  
 <213> Sequencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 55 <400> 307  
 taaatgtgca actcaggcag g 21

60 <210> 308  
 <211> 21  
 <212> ADN  
 <213> Sequencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 65 <400> 308  
 gtgcttcacc acttgggtt t 21

<210> 309

<211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 5 <223> Cebador oligonucleótido sintético  
 <400> 309  
 gacccccatt taacttcag c 21

<210> 310  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 10 <223> Cebador oligonucleótido sintético  
 <400> 310  
 ggctgaggct ttatttga g 21

<210> 311  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 15 <223> Cebador oligonucleótido sintético  
 <400> 311  
 gcggtttca ctgacgcag a 21

<210> 312  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 20 <223> Cebador oligonucleótido sintético  
 <400> 312  
 cacctttgtg gaactgccta c 21

<210> 313  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 25 <223> Cebador oligonucleótido sintético  
 <400> 313  
 aaaccttcgc attcatcct c 21

<210> 314  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 30 <223> Cebador oligonucleótido sintético  
 <400> 314  
 gcctttctgt ggctttgt t 21

<210> 315  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 35 <223> Cebador oligonucleótido sintético  
 <400> 315  
 tgttctacca ggacgccag g 21

<210> 316  
 <211> 21  
 <212> ADN

<213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 316  
 5 acgcaaggac ctgacattaa a 21  
  
 <210> 317  
 <211> 21  
 <212> ADN  
 10 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 317  
 15 tggtaactga aaatgggtgg a 21  
  
 <210> 318  
 <211> 21  
 <212> ADN  
 20 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 318  
 ttaattgctt ccaccaacct t 21  
  
 25 <210> 319  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 30 <223> Cebador oligonucleótido sintético  
 <400> 319  
 agatccctgg tcttgttc c 21  
  
 <210> 320  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 40 <223> Cebador oligonucleótido sintético  
 <400> 320  
 cctggaaagc agagaactag g 21  
  
 <210> 321  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 45 <223> Cebador oligonucleótido sintético  
 <400> 321  
 50 catttggca tgtatgtgt g 21  
  
 <210> 322  
 <211> 21  
 <212> ADN  
 55 <213> Sequenza artificiale  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 322  
 60 cccctctcct tagaaaatcc c 21  
  
 <210> 323  
 <211> 21  
 <212> ADN  
 65 <213>Secuencia artificial  
 <220>

<223> Cebador oligonucleótido sintético  
 <400> 323  
 gcgaatggga tacatcaaag a 21

5

<210> 324  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>

10

<223> Cebador oligonucleótido sintético  
 <400> 324  
 cagagcctaa agaaccacc a 21

15

<210> 325  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>

20

<223> Cebador oligonucleótido sintético  
 <400> 325  
 tgcacacttt aactgacca t 21

25

<210> 326  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>

30

<223> Cebador oligonucleótido sintético  
 <400> 326  
 cagtaagcaa actgccctca t 21

35

<210> 327  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>

40

<223> Cebador oligonucleótido sintético  
 <400> 327  
 agccaaaggt tccaggacaa t 21

45

<210> 328  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>

50

<223> Cebador oligonucleótido sintético  
 <400> 328  
 ccagcacctc accctgacta t 21

55

<210> 329  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>

60

<223> Cebador oligonucleótido sintético  
 <400> 329  
 tgtgatgcta cccactccac t 21

65

<210> 330  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>

<223> Cebador oligonucleótido sintético  
 <400> 330

ggtctcctcg gttactcca t 21  
 <210> 331  
 <211> 21  
 5 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 331  
 10 gcaagtgcaa ggaaggaaag a 21  
 <210> 332  
 <211> 21  
 <212> ADN  
 15 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 332  
 20 ggaggcctgc tgaacttctt t 21  
 <210> 333  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 25 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 333  
 gacataagg ctacaggaa a 21  
 30 <210> 334  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 35 <223> Cebador oligonucleótido sintético  
 <400> 334  
 atagtggtgg taggatggtg c 21  
 <210> 335  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 40 <223> Cebador oligonucleótido sintético  
 <400> 335  
 45 aaaagcaaga gcagagaagg c 21  
 <210> 336  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 336  
 55 aggcgctgca actacaaga t 21  
 <210> 337  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 337  
 60 ctgacctca cagacatccc c 21

<210> 338  
 <211> 21  
 <212> ADN  
 5 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 338  
 ggactacaag ggacgagagg c 21  
 10  
 <210> 339  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 15 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 339  
 ctgggaggat cacctgacag a 21  
 20  
 <210> 340  
 <211> 21  
 <212> ADN  
 <213> Sequenza artificiale  
 <220>  
 25 <223> Cebador oligonucleótido sintético  
 <400> 340  
 caccaccaag acaagcctaa a 21  
 30  
 <210> 341  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 35 <400> 341  
 ccataaatga agcctcctgc c 21  
 40  
 <210> 342  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 342  
 45 cttggaaaag ccaatggaaa t 21  
 50  
 <210> 343  
 <211> 21  
 <212> ADN  
 <213> Sequenza artificiale  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 343  
 aaggtcctag cccttagcag a 21  
 55  
 <210> 344  
 <211> 21  
 <212> ADN  
 <213> Sequenza artificiale  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 344  
 acacacacat ctgcaactcg g 21  
 60  
 65 <210> 345

<211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 5 <223> Cebador oligonucleótido sintético  
 <400> 345  
 ccgtttcagc aaaactgaga a 21  
  
 <210> 346  
 10 <211> 21  
 <212> ADN  
 <213> Sequenza artificiale  
 <220>  
 <223> Cebador oligonucleótido sintético  
 15 <400> 346  
 agtgtcttag tcaccccctg c 21  
  
 <210> 347  
 <211> 21  
 20 <212> ADN  
 <213> Sequenza artificiale  
 <220>  
 <223> Cebador oligonucleótido sintético  
 25 <400> 347  
 cccaggatgt ggagatgaaa g 21  
  
 <210> 348  
 <211> 21  
 <212> ADN  
 30 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 348  
 35 ggtggggggac taggagtgaa a 21  
  
 <210> 349  
 <211> 21  
 <212> ADN  
 <213> Sequenza artificiale  
 40 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 349  
 acgaagtcag tttggtggtg g 21  
  
 <210> 350  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 50 <223> Cebador oligonucleótido sintético  
 <400> 350  
 cacccacact tttctgcctc t 21  
  
 <210> 351  
 <211> 21  
 <212> ADN  
 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 60 <400> 351  
 ggtgagccct taccctcagt t

21

<213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 352  
 5 ctccagcct ggtcacaata t 21  
  
 <210> 353  
 <211> 21  
 <212> ADN  
 10 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 353  
 15 cgtgtccact tcaagtgaa t 21  
  
 <210> 354  
 <211> 21  
 <212> ADN  
 20 <213>Secuencia artificial  
 <220>  
 <223> Cebador oligonucleótido sintético  
 <400> 354  
 caccaggctt gtcattacc a 21  
  
 25 <210> 355  
 <211> 21  
 <212> ADN  
 <213> Sequenza artificiale  
 <220>  
 30 <223> Cebador oligonucleótido sintético  
 <400> 355  
 gatgcctcat ctttctcac t 21  
  
 <210> 356  
 <211> 23  
 <212> ADN  
 <213> Sequenza artificiale  
 <220>  
 40 <223> Cebador oligonucleótido sintético  
 <400> 356  
 gtgaactcac tcaactgggt agc 23  
  
 <210> 357  
 <211> 6  
 <212> ADN  
 45 <213>Secuencia artificial  
 <220>  
 <223> Esempio di oligonucleotide consistente esclusivamente di purine  
 <400> 357  
 50 gaggag 6  
  
 <210> 358  
 <211> 6  
 <212> ADN  
 55 <213>Secuencia artificial  
 <220>  
 <223> Complemento inverso della SEQ ID N° 357  
 <400> 358  
 60 ctcttc 6  
  
 <210> 359  
 <211> 6  
 <212> ADN  
 65 <213>Secuencia artificial  
 <220>

<223> Esempio di ripetizioni telomeriche  
<400> 359  
ttaggg 6

5 <210> 360  
<211> 5  
<212> ADN  
<213>Secuencia artificial  
<220>

10 <223> Esempio di unità ripetuta a cinque basi nei satelliti II e III  
<400> 360  
attcc

**REIVINDICACIONES**

- 5 1. Un método para producir una sonda de ácido nucleico que comprende: la identificación de secuencias de ácido nucleico repetitivas de una molécula de ácido nucleico diana; la generación al menos de una primera región de unión y una segunda región de unión donde al menos la primera región de unión y la segunda región de unión están sustancialmente libres de secuencias de ácido nucleico repetitivas de la molécula de ácido nucleico diana, y
- 10 la unión de al menos la primera región de unión y la segunda región de unión en una orientación y un orden aleatorios, produciendo así una pluralidad de moléculas de ácido nucleico, en la que sustancialmente todas las moléculas de ácido nucleico comprenden al menos una primera región de unión y una segunda región de unión contiguas, donde la primera región de unión y la segunda región de unión contiguas son complementarias de las porciones no contiguas y secuencias únicas de la molécula de ácido nucleico diana, y donde la pluralidad de moléculas de ácido nucleico forma la sonda.
- 15 2. El método de la reivindicación 2, que comprende asimismo: la amplificación de la pluralidad de moléculas de ácido nucleico para producir una pluralidad de amplímeros de la molécula de ácido nucleico para formar la sonda, preferiblemente comprendiendo también el etiquetado de la pluralidad de amplímeros de la molécula de ácido nucleico, para producir una sonda etiquetada, preferiblemente en la que la pluralidad de amplímeros de la molécula de ácido nucleico está etiquetada utilizando la traslación por muescas.
- 20 3. El método de la reivindicación 1, donde la molécula de ácido nucleico diana es de un genoma eucariótico o viral, o de una molécula de ARN.
4. El método de la reivindicación 1, donde al menos la primera región de unión y la segunda región de unión se generan:
- (a) aislando al menos la primera región de unión y la segunda región de unión de la molécula de ácido nucleico diana;
- 25 (b) obteniendo al menos la primera región de unión y la segunda región de unión de la molécula de ácido nucleico diana mediante hibridación sustractiva;
- (c) amplificando al menos la primera región de unión y la segunda región de unión de la molécula de ácido nucleico diana; o
- (d) cualquier combinación de (a), (b) y (c).
- 30 5. El método de la reivindicación 1, donde al menos la primera región de unión y la segunda región de unión son aisladas o amplificadas de una molécula de ácido nucleico diana presente en un vector.
6. El método de la reivindicación 1, que comprende asimismo: la generación y amplificación de una pluralidad de regiones de unión que están sustancialmente libres de secuencias de ácido nucleico repetitivas de una secuencia de ácido nucleico diana para producir una pluralidad de amplímeros de la región de unión;
- 35 la unión de la pluralidad de los amplímeros de la región de unión para producir una pluralidad de moléculas de ácido nucleico, en la que sustancialmente todas las moléculas de ácido nucleico comprenden regiones de unión contiguas que son complementarias de las porciones no contiguas y secuencias únicas de una molécula de ácido nucleico diana.
- 40 la amplificación de la pluralidad de moléculas de ácido nucleico para producir una pluralidad de amplímeros de la molécula de ácido nucleico; y
- el etiquetado de la pluralidad de amplímeros de la molécula de ácido nucleico para producir la sonda.

FIG. 1A

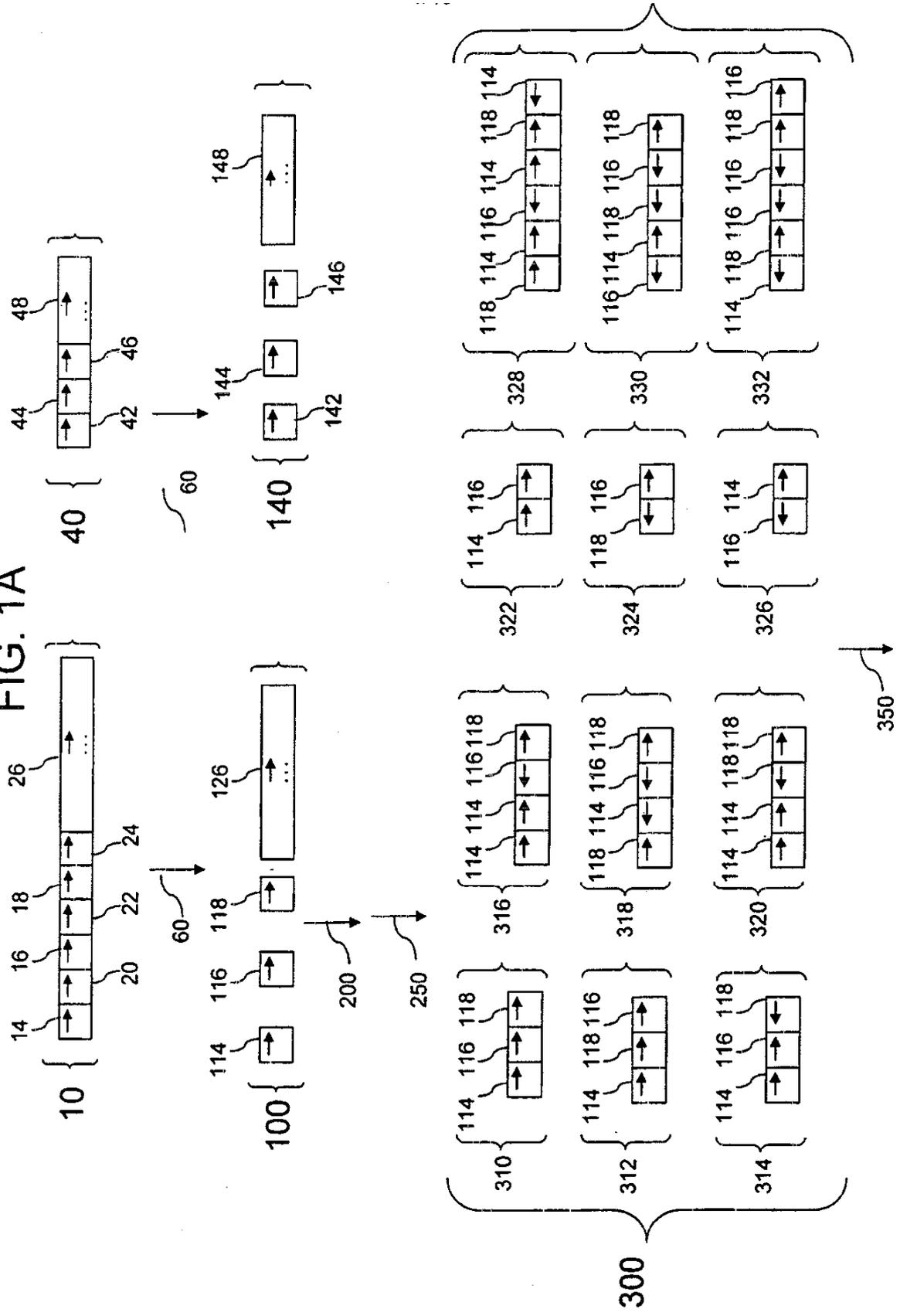
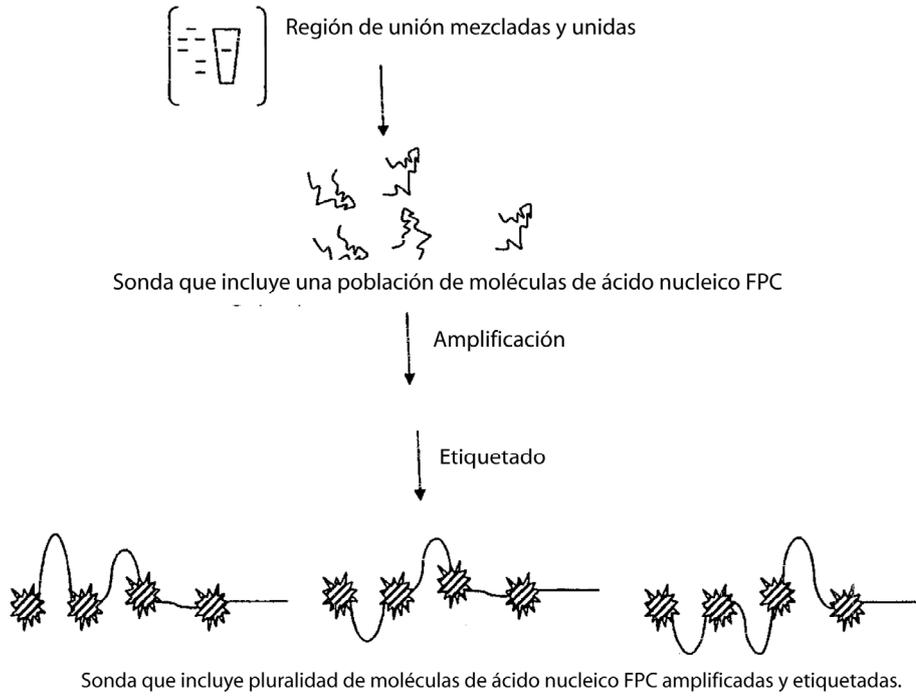


FIG. 1B



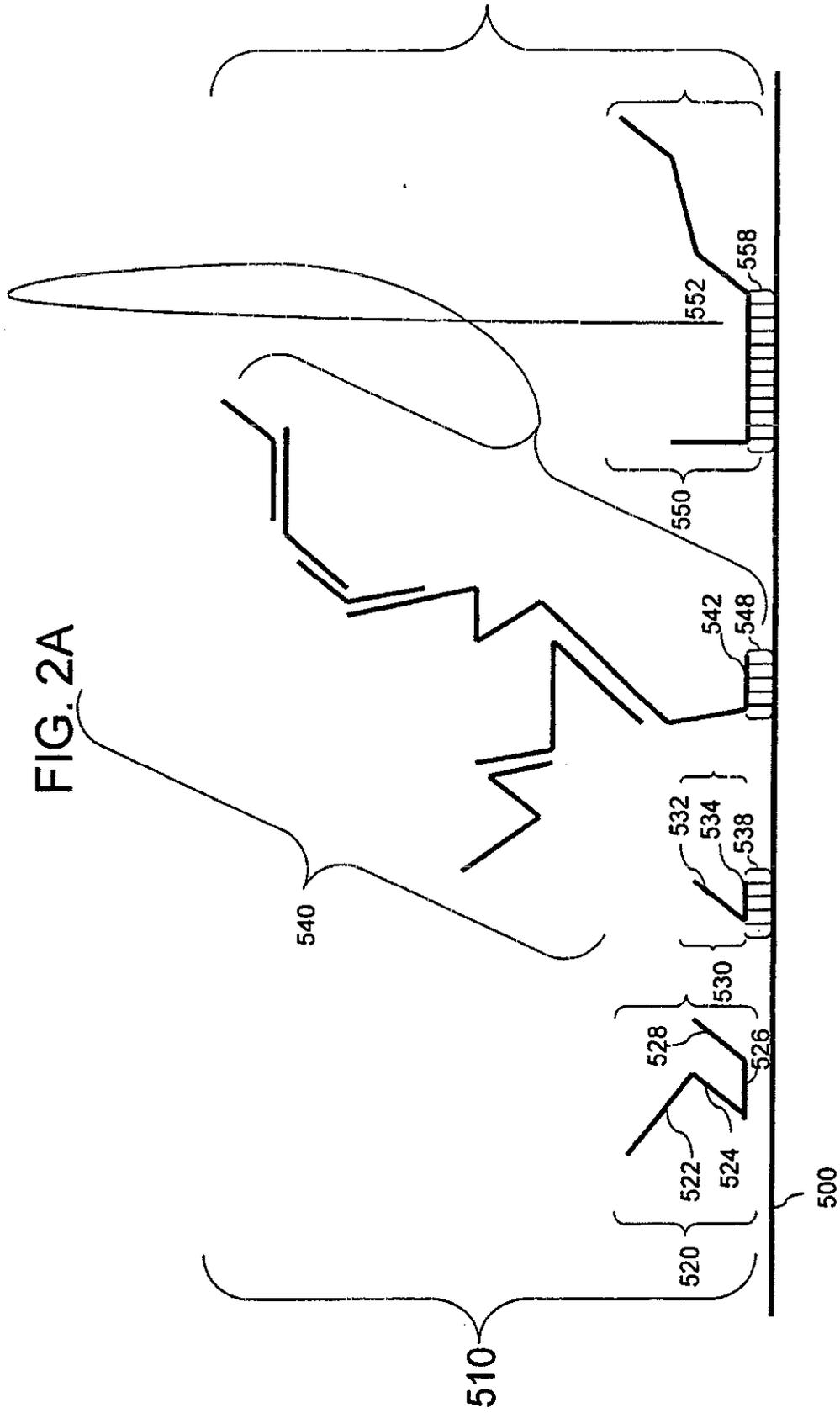


FIG. 2B

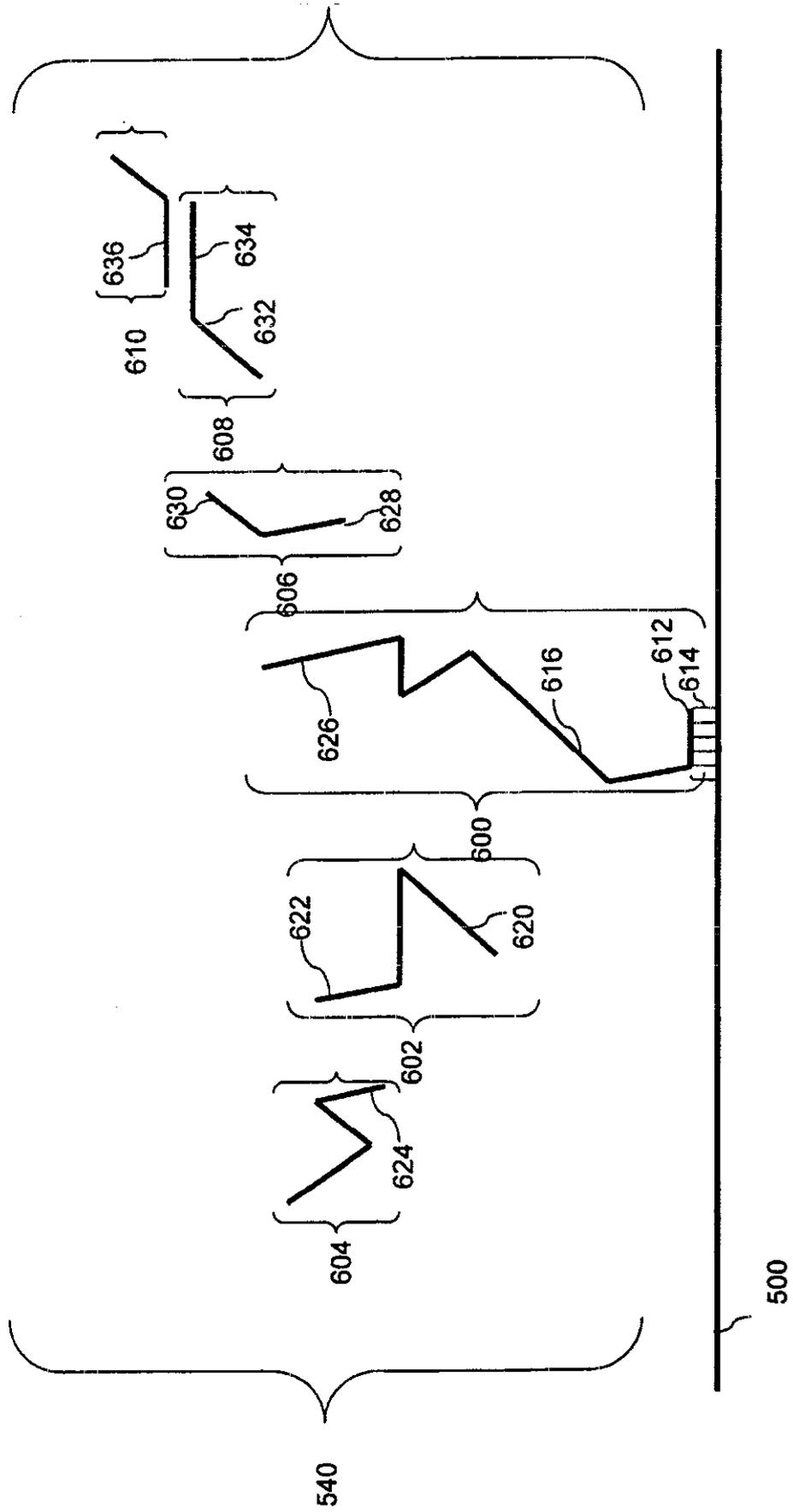


FIG. 3

~ 200 kb

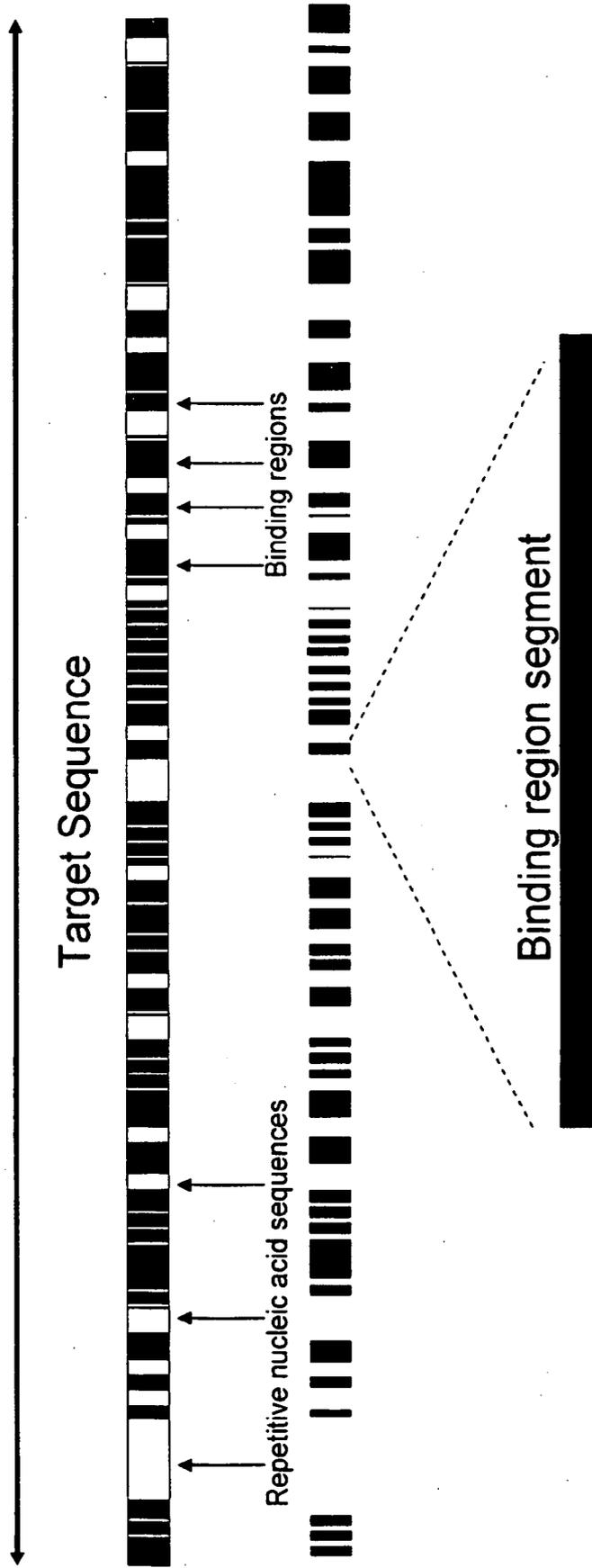


FIG. 4

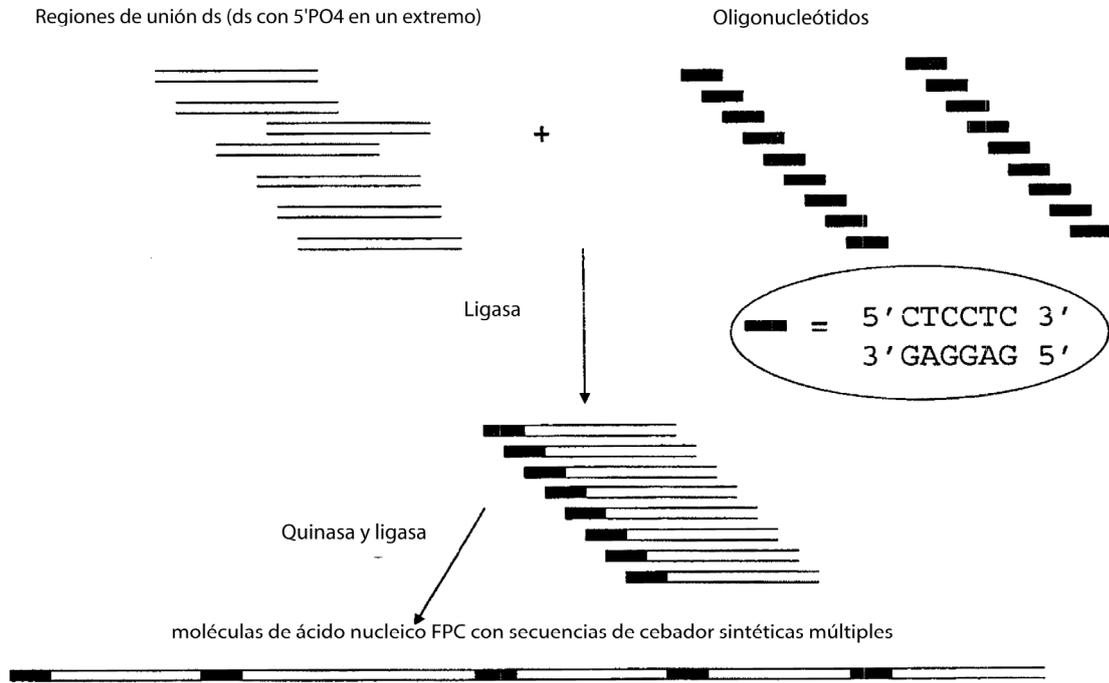


FIG. 5

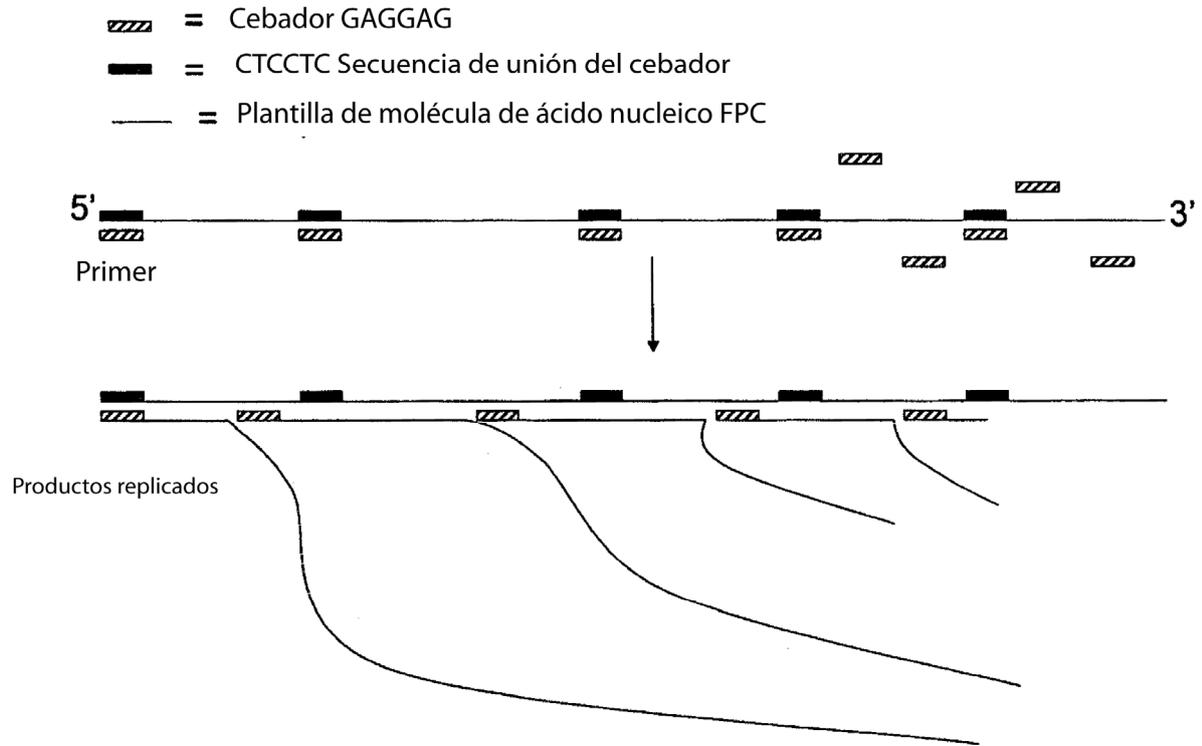


FIG. 6B



FIG. 6D



FIG. 6A

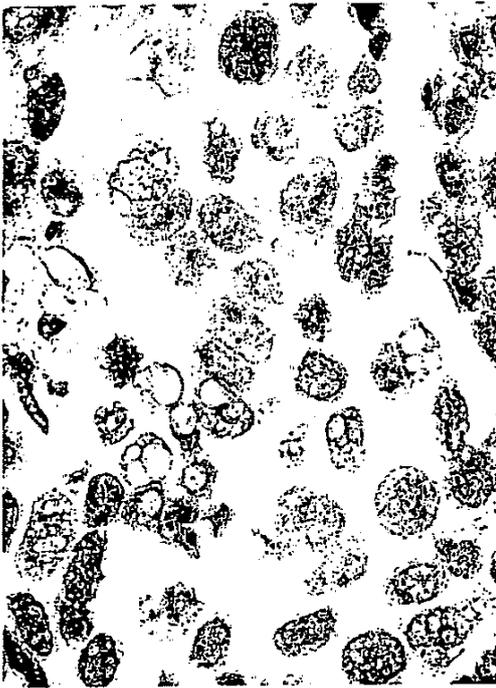
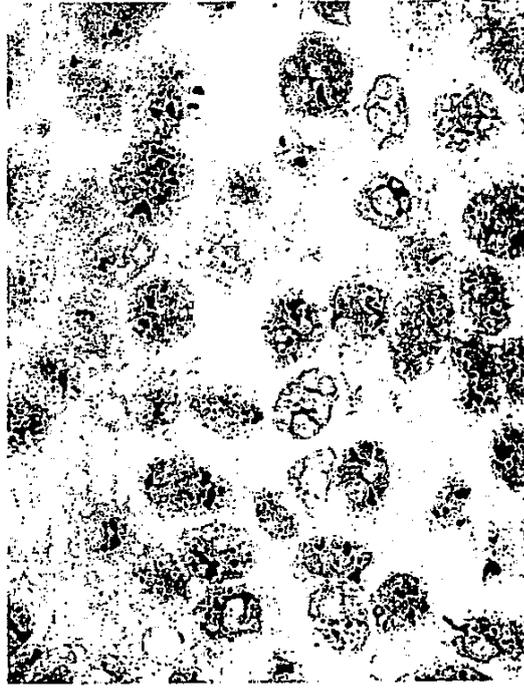


FIG. 6C



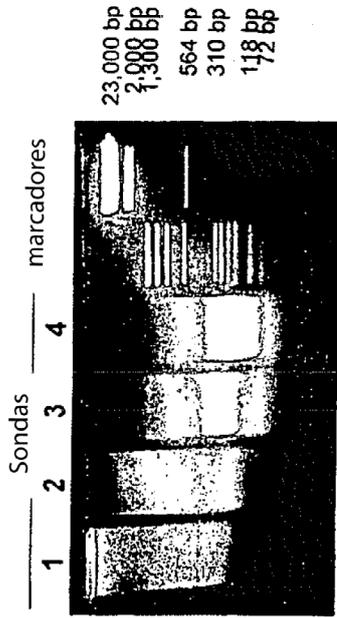


FIG. 7

FIG. 8A

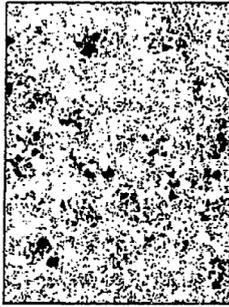


FIG. 8B



FIG. 8C



FIG. 8D

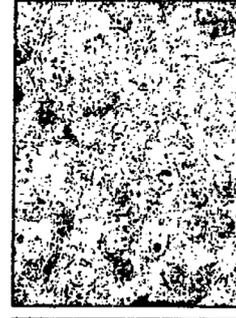
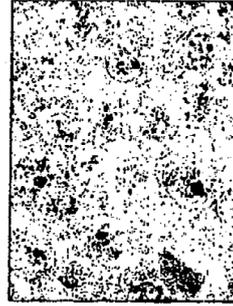


FIG. 9A

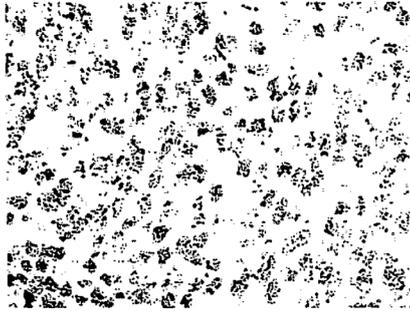


FIG. 9B

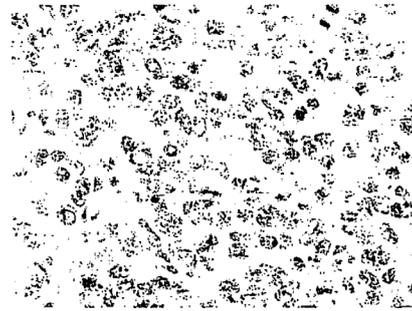


FIG. 10

	Riñón negativo	Riñón positivo		
Sonda BKV 1 ug/ml 10x detección de CISH"				
Sonda BKV 2 ug/ml 10X detección de SISH"				