

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 940 620**

51 Int. Cl.:

C12P 19/34 (2006.01)

C12Q 1/68 (2008.01)

C12Q 1/6806 (2008.01)

C12Q 1/6869 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **02.11.2017 PCT/US2017/059804**

87 Fecha y número de publicación internacional: **11.05.2018 WO18085599**

96 Fecha de presentación y número de la solicitud europea: **02.11.2017 E 17868046 (8)**

97 Fecha y número de publicación de la concesión europea: **21.12.2022 EP 3535405**

54 Título: **Métodos de preparación de muestras de ácido nucleico para la secuenciación del repertorio inmunitario**

30 Prioridad:

02.11.2016 US 201662416677 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
09.05.2023

73 Titular/es:

**ARCHERDX, LLC (100.0%)
1400 16th Street
San Francisco, CA 91403, US**

72 Inventor/es:

**MYERS, JASON;
STAHL, JOSHUA;
CULVER, BRADY;
KUDLOW, BRIAN y
EBERLEIN, JENS**

74 Agente/Representante:

VALLEJO LÓPEZ, Juan Pedro

Observaciones:

Véase nota informativa (Remarks, Remarques o Bemerkungen) en el folleto original publicado por la Oficina Europea de Patentes

ES 2 940 620 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

Métodos de preparación de muestras de ácido nucleico para la secuenciación del repertorio inmunitario

5 **Campo técnico**

La tecnología descrita en el presente documento se refiere a métodos y composiciones útiles en la preparación de moléculas de ácido nucleico para análisis.

10 **Antecedentes**

El sistema inmunitario adaptativo está compuesto por células altamente especializadas, células y procesos sistémicos destinados a eliminar o prevenir el crecimiento de patógenos. Los linfocitos T y los linfocitos B pueden considerarse los principales componentes celulares del sistema inmunitario adaptativo que impulsan la respuesta inmunitaria adaptativa al generar una amplia diversidad de moléculas de unión a antígeno a través de la recombinación genética y la mutación somática de sus loci receptores de antígenos. Los linfocitos T y B inmaduros se someten a procesos selectivos para producir poblaciones en gran medida desprovistas de autorreactividad. Armado con este repertorio inicial de moléculas de unión a antígeno (por ejemplo, receptores de linfocitos T e inmunoglobulinas), los linfocitos T y B vírgenes circulan por todo el cuerpo donde pueden entrar en contacto con el antígeno.

Tras la exposición al antígeno afín y junto con suficientes señales coestimuladoras, los linfocitos T o B específicos de antígeno se activan y proliferan, y en el caso de los linfocitos B, pueden sufrir más edición de secuencias de sus loci de receptores inmunitarios (por ejemplo, a través de hipermutación somática y/o recombinación adicional).

Como resultado de estos procesos selectivos, el repertorio de especificidades de unión en una muestra individual puede proporcionar un historial de exposiciones antigénicas pasadas, además de ser informativo de las capacidades y limitaciones inherentes al repertorio.

El documento US 2013/0303461 A1 desvela métodos para determinar secuencias de oligonucleótidos, por ejemplo, enriqueciendo las secuencias diana antes de secuenciar las secuencias. El documento WO 2014/008447 A1 desvela composiciones de oligonucleótidos modificados y métodos para reducir selectivamente los contaminantes de ácidos nucleicos no deseados y enriquecerlos para objetivos de ácidos nucleicos deseados a partir de mezclas de ácidos nucleicos genómicos complejos para aplicaciones de secuenciación. Liu *et al.*, PLOS One, 2016, 11(3): e0152464 compara dos métodos, PCR multiplexada y 5'RACE, que se utilizan predominantemente para enriquecer BCR y TCR reorganizados. Gazzola *et al.*, Ther Adv Hematol, 2014, 35-47 revisa los desarrollos recientes en las pruebas de clonalidad para los trastornos linfoproliferativos, con énfasis en sistemas altamente sensibles para mejorar los estudios de diagnóstico y las evaluaciones mínimas de enfermedad residual.

40 **Sumario**

La invención se define en las reivindicaciones adjuntas.

La tecnología desvelada en el presente documento se refiere a métodos de preparación y análisis de ácidos nucleicos. Los métodos proporcionados en el presente documento son útiles para detectar variantes de ácido nucleico de frecuencia ultrabaja (por ejemplo, variantes de corte y empalme, fusiones, variantes de un solo nucleótido, inserciones y deleciones, variantes de número de copia, ARNm de loci de receptores inmunitarios recombinados somáticamente y variantes del nivel de expresión) en preparaciones de ácidos nucleicos, incluyendo secuencias representativas de un repertorio inmunitario que comprende un panorama diverso de secuencias de ácidos nucleicos que codifican inmunoglobulinas y receptores inmunitarios. Los métodos proporcionados en el presente documento implican la captura basada en la ligadura que enriquece las moléculas de ácido nucleico que tienen secuencias de nucleótidos correspondientes a los ácidos nucleicos transcritos que reflejan eventos de recombinación y/o corte y empalme ocurridos previamente. La captura de moléculas únicas puede mejorarse enormemente con respecto a los métodos convencionales para los ácidos nucleicos extraídos de individuos, por ejemplo, individuos portadores de tumores o individuos inmunocomprometidos. La eficiencia de captura se puede al menos duplicar en comparación con los métodos convencionales para los ácidos nucleicos extraídos de individuos, por ejemplo, individuos portadores de tumores o individuos inmunocomprometidos. Se puede lograr una profundidad mejorada como resultado de una química de captura frontal mejorada.

Los métodos proporcionados en el presente documento son útiles para evaluar los repertorios inmunitarios de ARN a través de la secuenciación. En el presente documento se proporcionan métodos y composiciones útiles en la preparación de muestras de ácido nucleico para el análisis de secuencias (por ejemplo, usando secuenciación de última generación). Las técnicas descritas en el presente documento se refieren a métodos para determinar una secuencia de ácido nucleico. Los métodos y composiciones descritos en el presente documento pueden referirse a enriquecimiento de ácidos nucleicos que comprenden una o más secuencias de nucleótidos diana antes de la secuenciación. La divulgación proporciona métodos para preparar ácidos nucleicos (por ejemplo, para su uso en un análisis de secuenciación) que implican el uso de un cebador modificado con la resto de captura para sintetizar una

primera cadena de ácido nucleico o que implican la adición de uno o más nucleótidos modificados con la resto de captura a un ácido nucleico.

Los métodos pueden implicar además ligar un ácido nucleico adaptador al ácido nucleico al que se ha usado el cebador modificado con el resto de captura para sintetizar una primera cadena del ácido nucleico para producir un producto de ligadura. Los métodos pueden implicar además ligar un ácido nucleico adaptador al ácido nucleico al que se ha añadido el nucleótido modificado con el resto de captura para producir un producto de ligadura. Los métodos pueden implicar además capturar el producto de ligadura poniendo en contacto el producto de ligadura con un compañero de unión de un resto de captura del cebador modificado con el resto de captura o un compañero de unión de un resto de captura del nucleótido modificado con el resto de captura.

La divulgación proporciona métodos para preparar ácidos nucleicos para análisis, en el que los métodos implican: (a) poner en contacto una molécula de ácido nucleico que comprende una secuencia de nucleótidos diana con un cebador modificado con un resto de captura que se aparea específicamente con la secuencia de nucleótidos diana en condiciones de hibridación; (b) llevar a cabo una reacción de síntesis de la primera cadena que está cebada por un cebador modificado con un resto de captura hibridado y que usa la molécula de ácido nucleico como molde; (c) llevar a cabo una reacción de síntesis de la segunda cadena que usa un producto de la reacción de síntesis de la primera cadena como molde para generar un ácido nucleico bicatenario que comprende un resto de captura; (d) ligar un ácido nucleico adaptador al ácido nucleico bicatenario para producir un producto de ligadura que comprende el resto de captura; (e) capturar el producto de ligadura poniendo en contacto el producto de ligadura con un compañero de unión del resto de captura; y (f) amplificar el producto de ligadura capturado mediante la reacción en cadena de la polimerasa utilizando un cebador específico de la diana que comprende una porción en 3' que se aparea específicamente con la secuencia de nucleótidos diana y un primer cebador adaptador que se aparea específicamente con una secuencia complementaria del ácido nucleico adaptador, en donde el cebador específico de la diana comprende una parte de la cola en 5' que no se aparea específicamente con la secuencia de nucleótidos diana.

Los métodos proporcionados en el presente documento comprenden además (g) amplificar un producto de amplificación de la etapa (f) mediante la reacción en cadena de la polimerasa utilizando un cebador de cola que comprende una porción en 3' que se aparea específicamente con una secuencia complementaria de la porción de cola en 5' del cebador específico de la diana y un segundo cebador adaptador que se aparea específicamente con una secuencia complementaria del ácido nucleico adaptador, en donde el cebador de cola comprende una porción en 5' que no se aparea específicamente con una secuencia complementaria del cebador específico de la diana. La porción en 5' del cebador de cola puede comprender al menos una de una región índice de muestra, una región de unión al cebador de PCR, una región de código de barras molecular y una región de sitio de cebador de secuenciación. El primer cebador adaptador y el segundo cebador adaptador pueden ser iguales. Por ejemplo, los cebadores adaptadores primero y segundo pueden consistir en la misma secuencia de nucleótidos. El primer cebador adaptador y el segundo cebador adaptador pueden ser diferentes (por ejemplo, consistir en diferentes secuencias de nucleótidos y/o comprender uno o más restos diferentes). El segundo cebador adaptador puede anidarse en relación con el primer cebador adaptador. El segundo cebador adaptador no puede anidarse en relación con el primer cebador adaptador.

La etapa (d) puede comprender combinar el ácido nucleico adaptador, el ácido nucleico bicatenario y una ligasa en condiciones en las que la ligasa liga el ácido nucleico adaptador al ácido nucleico bicatenario, en donde el ácido nucleico adaptador que se combina con el ácido nucleico bicatenario comprende una porción dúplex y una secuencia saliente, en donde la secuencia saliente comprende una secuencia de nucleótidos que es complementaria a una secuencia saliente en el extremo 3' del ácido nucleico bicatenario. El ácido nucleico adaptador puede comprender al menos una de una región índice de la muestra, una región de unión al cebador de PCR, una región de código de barras molecular (por ejemplo, una región que identifica de forma única las moléculas de entrada) y una región del sitio del cebador de secuenciación.

La etapa (d) puede comprender combinar el ácido nucleico adaptador, el ácido nucleico bicatenario y una ligasa en condiciones en las que la ligasa liga el ácido nucleico adaptador al ácido nucleico bicatenario, en donde el ácido nucleico adaptador que se combina con el ácido nucleico bicatenario es monocatenario.

El cebador modificado con el resto de captura puede comprender al menos un nucleótido modificado con el resto de captura. En algunas realizaciones, el cebador modificado con el resto de captura comprende un primer grupo de acoplamiento químico configurado para unirse a un segundo grupo de acoplamiento químico unido a un resto de captura. En algunas realizaciones, el resto de captura es un resto de biotina. En algunas realizaciones, el resto de biotina comprende biotina-trietilenglicol, bis-biotina, biotina fotoescindible, destiobiotina, destiobiotina-trietilenglicol o azida de biotina. En algunas realizaciones, el compañero de unión del resto de captura es estreptavidina. En algunas realizaciones, la estreptavidina se une a un sustrato. En algunas realizaciones, el sustrato comprende una superficie sólida. En algunas realizaciones, la superficie sólida comprende una perla paramagnética.

Los métodos descritos en el presente documento pueden comprender además una etapa de liberación del producto de ligadura capturado del compañero de unión del resto de captura. El producto de ligadura capturado puede liberarse del compañero de unión del resto de captura poniéndolo en contacto con un reactivo químico y/o aplicando calor. El reactivo químico puede ser una base o una solución básica. En algunas realizaciones, el reactivo químico comprende

hidróxido de sodio (NaOH). Debe apreciarse que el contacto puede implicar la mezcla de dos soluciones (por ejemplo, una solución que comprende una base y una solución que comprende un ácido nucleico inmovilizado lavado), añadir un sólido a una solución o añadir una solución a un sólido. El producto de ligadura capturado puede liberarse del compañero de unión del resto de captura poniéndolo en contacto con NaOH y calentándolo (por ejemplo, calentando a temperatura superior a la ambiente, tal como una temperatura en un intervalo de 25 a 90 °C, de 25 a 70 °C, de 25 a 50 °C, de 35 a 65 °C, de 35 a 45 °C, de 30 a 40 °C, de 40 a 50 °C). El producto de ligadura capturado puede permanecer unido al compañero de unión del resto de captura, por ejemplo, para una mayor preparación para el análisis. El producto de ligadura capturado puede liberarse del compañero de unión del resto de captura antes de la preparación adicional para el análisis.

Los métodos proporcionados en el presente documento pueden comprender además una etapa de lavado después de la etapa (d) y antes de la etapa (e).

Los métodos proporcionados en el presente documento pueden comprender además, después de la etapa (e) y antes de la etapa (f): i) inmovilizar el ácido nucleico bicatenario, que comprende el resto de captura, sobre un sustrato paramagnético; y ii) lavar el ácido nucleico bicatenario inmovilizado. En algunas realizaciones, los métodos proporcionados en el presente documento comprenden además, después de la etapa ii): iii) liberar el ácido nucleico bicatenario inmovilizado lavado del sustrato paramagnético.

Los métodos proporcionados en el presente documento pueden comprender además, después de la etapa (c) y antes de la etapa (d), 5' fosforilando el ácido nucleico bicatenario. La fosforilación en 5' puede comprender generar un grupo fosfato en el extremo 5' de una cadena del ácido nucleico bicatenario. Por ejemplo, se puede añadir un grupo fosfato a un grupo hidroxilo en el extremo 5' de la cadena (por ejemplo, mediante la acción de una polinucleótido cinasa).

Los métodos proporcionados en el presente documento comprenden además, después de la etapa (c) y antes de la etapa (d), reparación de los extremos del ácido nucleico bicatenario para producir un ácido nucleico bicatenario de extremos romos. La reparación de extremos puede comprender extremos romos. Los extremos romos se pueden lograr mediante la eliminación de nucleótidos no apareados terminales (por ejemplo, una secuencia saliente) de una cadena del ácido nucleico bicatenario. Por ejemplo, los nucleótidos no apareados terminales pueden eliminarse de una cadena de un ácido nucleico bicatenario mediante el uso de una enzima (por ejemplo, fragmento Klenow) con actividad de exonucleasa (por ejemplo, para hidrolizar un enlace fosfodiéster terminal, eliminando así la base saliente de una en una). Los extremos romos se pueden lograr rellenando un extremo rebajado con una enzima polimerizante (por ejemplo, una ADN polimerasa) en presencia de nucleótidos trifosfato.

Los métodos proporcionados en el presente documento pueden comprender además añadir (por ejemplo, ligando, enlazando) uno o más nucleótidos a un extremo 3' del ácido nucleico bicatenario de extremos romos. En algunas realizaciones, el uno o más nucleótidos comprenden nucleótidos de desoxiadenosina. Por ejemplo, los métodos pueden comprender además la cola dA de un extremo 3' del ácido nucleico bicatenario (por ejemplo, usando el fragmento Klenow). En algunas realizaciones, el ácido nucleico adaptador comprende una secuencia de nucleótidos en un extremo 3' que comprende uno o más nucleótidos complementarios al uno o más nucleótidos añadidos al extremo 3' del ácido nucleico bicatenario de extremos romos. En algunas realizaciones, la secuencia de nucleótidos en el extremo 3' del ácido nucleico adaptador comprende uno o más nucleótidos de desoxitimidina. En algunas realizaciones, el ácido nucleico adaptador comprende además una cadena de bloqueo apareada con una cadena de amplificación que comprende la secuencia de nucleótidos en el extremo 3' y en donde la secuencia de nucleótidos en el extremo 3' no está apareado de manera que forma una secuencia saliente.

Los métodos pueden comprender, además, después de la reparación de los extremos: i) inmovilizar el ácido nucleico bicatenario, que comprende el resto de captura, sobre un sustrato paramagnético; ii) lavar el ácido nucleico bicatenario inmovilizado; y iii) liberar el ácido nucleico bicatenario inmovilizado lavado del sustrato paramagnético.

Los métodos pueden comprender, además, después de la reparación de los extremos: i) inmovilizar el ácido nucleico bicatenario, que comprende el resto de captura, sobre un sustrato paramagnético; y ii) lavar el ácido nucleico bicatenario inmovilizado. Los métodos pueden comprender además una etapa de lavado después de la reparación de los extremos y antes de la etapa i).

Los métodos proporcionados en el presente documento pueden comprender además: (h) inmovilizar un producto de amplificación de la etapa (g) sobre un sustrato paramagnético; (i) lavar el producto de amplificación inmovilizado; y (j) liberar el producto de amplificación inmovilizado lavado del sustrato paramagnético. El método puede comprender además una etapa de lavado después de la etapa (g) y antes de la etapa (h).

En la etapa (d), el ácido nucleico bicatenario puede ligarse al ácido nucleico adaptador en presencia de un agente de agregación. El agente de agregación puede ser polietilenglicol en una cantidad que representa del 5 % al 50 % de una mezcla de ligadura.

Los métodos proporcionados en el presente documento pueden comprender además, después de la etapa (b) y antes de la etapa (c), poner en contacto la molécula de ácido nucleico con una enzima ribonucleasa. La enzima ribonucleasa

puede degradar porciones de la molécula de ácido nucleico de modo que los fragmentos permanezcan apareados con el producto de la reacción de síntesis de la primera cadena. La reacción de síntesis de la segunda cadena puede cebarse con un fragmento de la molécula de ácido nucleico hibridada con el producto de la reacción de síntesis de la primera cadena.

5 La síntesis de la segunda cadena puede cebarse aleatoriamente usando una pluralidad de cebadores aleatorios. La pluralidad de cebadores aleatorios puede tener entre 6 bases de longitud y 15 bases de longitud.

10 La molécula de ácido nucleico puede comprender ARNm. En algunas realizaciones, la molécula de ácido nucleico se obtiene de una muestra que comprende un linfocito T, un linfocito B o una mezcla de los mismos. La muestra puede comprender uno o más linfocitos T, uno o más linfocitos B o una mezcla de los mismos. La muestra puede comprender una mezcla de linfocitos T y linfocitos B. La mezcla de linfocitos T y linfocitos B puede comprender además uno o más tipos de células no sanguíneas. La mezcla de linfocitos T y linfocitos B puede comprender además uno o más tipos de leucocitos (por ejemplo, neutrófilos, eosinófilos, basófilos, linfocitos citotóxicos naturales (NK, *Natural Killer*), monocitos, histiocitos, células dendríticas, mastocitos, microglía, etc.).

20 En algunas realizaciones, la muestra se obtiene de un sujeto que tiene, o se sospecha que tiene, una neoplasia maligna de linfocitos T o una neoplasia maligna de linfocitos B. La muestra puede obtenerse de un sujeto que tiene, o se sospecha que tiene, linfoma o leucemia (por ejemplo, mieloma múltiple, leucemia linfoblástica aguda y leucemia linfocítica crónica). La muestra puede obtenerse de un sujeto que tiene, o se sospecha que tiene, un tumor sólido (por ejemplo, sarcoma, carcinoma, linfoma o cualquier tumor de origen no linfóide que puede contener o no células inmunitarias no malignas). La muestra puede obtenerse de un sujeto que ha sufrido o sufrirá un trasplante. La muestra se puede obtener de un sujeto cuya respuesta inmunitaria a un tratamiento se está evaluando. La muestra puede obtenerse de un sujeto que tiene, o se sospecha que tiene, una neoplasia maligna de glóbulos blancos. La muestra puede obtenerse de un sujeto que tiene, o se sospecha que tiene, una afección autoinmunitaria. Por ejemplo, la afección inmunitaria puede ser cualquier afección impulsada por linfocitos T autorreactivos, linfocitos B autorreactivos o una combinación de los mismos. La muestra puede obtenerse de un sujeto que tiene, o se sospecha que tiene, una neoplasia maligna de linfocitos T y/o una neoplasia maligna de linfocitos B (por ejemplo, linfoma y subtipos de los mismos, mieloma múltiple, etc.) La muestra se puede obtener de un sujeto que tiene un tumor sólido. En algunas realizaciones, el sujeto es un ser humano. El sujeto puede ser un cordado. La molécula de ácido nucleico se puede obtener de una muestra que comprende un leucocito. En algunas realizaciones, la secuencia de nucleótidos diana comprende una secuencia de nucleótidos correspondiente a una porción de un gen del receptor de linfocitos T (TCR) o un gen del receptor de linfocitos B (BCR).

35 El cebador modificado con el resto de captura puede comprender una secuencia de nucleótidos que es complementaria a un gen de receptor inmunitario o un gen de inmunoglobulina. En algunas realizaciones, el cebador modificado con el resto de captura se aparea específicamente con una región constante que está corriente abajo de una región determinante de la complementariedad 3 (CDR3). En algunas realizaciones, el cebador específico de la diana se aparea específicamente con una región constante o un segmento J que está corriente abajo de una CDR3. El cebador específico de la diana puede aparearse específicamente con una unión exón-exón formada entre una región constante y un segmento J, y donde la unión exón-exón está corriente abajo de una CDR3.

45 La divulgación proporciona métodos para preparar ácidos nucleicos para análisis, en el que los métodos implican: (a) poner en contacto una molécula de ácido nucleico que comprende una secuencia de nucleótidos diana conocida y una secuencia de nucleótidos diana desconocida con un cebador modificado con un resto de captura que se aparea específicamente con la secuencia de nucleótidos diana conocida en condiciones de hibridación; (b) llevar a cabo una reacción de síntesis de la primera cadena que está cebada por un cebador modificado con un resto de captura hibridado y que usa la molécula de ácido nucleico como molde; (c) llevar a cabo una reacción de síntesis de la segunda cadena que está cebada por un fragmento de la molécula de ácido nucleico y que usa un producto de la reacción de síntesis de la primera cadena como molde para generar un ácido nucleico bicatenario que comprende un resto de captura; (d) reparar los extremos del ácido nucleico bicatenario para producir un ácido nucleico bicatenario de extremos romos que comprende el resto de captura; (e) ligar un ácido nucleico adaptador al ácido nucleico bicatenario de extremos romos para producir un producto de ligadura, en donde el producto de ligadura comprende la secuencia de nucleótidos diana desconocida flanqueada por el ácido nucleico adaptador y el resto de captura; (f) capturar el producto de ligadura poniendo en contacto el producto de ligadura con un compañero de unión inmovilizado del resto de captura; (g) amplificar el producto de ligadura mediante reacción en cadena de la polimerasa utilizando un cebador específico de la diana que comprende una porción en 3' que se aparea específicamente con la secuencia de nucleótidos diana conocida y un primer cebador adaptador que se aparea específicamente con una secuencia complementaria del ácido nucleico adaptador, en donde el cebador específico de la diana comprende una parte de la cola en 5' que no se aparea específicamente con la secuencia de nucleótidos diana conocida; y (h) amplificar un producto de amplificación de la etapa (g) mediante reacción en cadena de la polimerasa utilizando un cebador de cola que se aparea específicamente con una secuencia complementaria de la porción de cola en 5' del cebador específico de la diana y un segundo cebador adaptador que se aparea específicamente con una secuencia complementaria del ácido nucleico adaptador. En algunas realizaciones, los métodos descritos en el presente documento comprenden además, después de la etapa (e) y antes de la etapa (f), lavar el producto de ligadura. En algunas realizaciones, los métodos descritos en el presente documento comprenden además, después de la etapa (f) y antes de la etapa (g), lavar un producto de ligadura

capturado. En algunas realizaciones, los métodos descritos en el presente documento comprenden además (i) lavar un producto de amplificación de la etapa (h). La reacción de síntesis de la segunda cadena de la etapa (c) puede cebarse con cualquier fragmento de ácido nucleico presente en una muestra que comprende la molécula de ácido nucleico. Por ejemplo, la muestra puede comprender una mezcla compleja de ácidos nucleicos que son capaces de disociarse de una cadena complementaria y volver a aparearse con una cadena diferente presente en la mezcla.

La divulgación proporciona métodos para preparar ácidos nucleicos para análisis, en el que los métodos implican: (a) poner en contacto una molécula de ácido nucleico que comprende una secuencia de nucleótidos diana con un primer cebador modificado con un resto de captura que se aparea específicamente con la secuencia de nucleótidos diana en condiciones de hibridación; (b) llevar a cabo una reacción de síntesis de la primera cadena que está cebada por un primer cebador específico de la diana hibridado y que usa la molécula de ácido nucleico como molde; (c) llevar a cabo una reacción de síntesis de la segunda cadena que está cebada por un fragmento hibridado (por ejemplo, una parte o porción hibridada) de la molécula de ácido nucleico y que utiliza un producto de la reacción de síntesis de la primera cadena como molde para generar un ácido nucleico bicatenario; (d) ligar un ácido nucleico adaptador al ácido nucleico bicatenario para producir un producto de ligadura; (e) amplificar el producto de ligadura mediante reacción en cadena de la polimerasa utilizando un segundo cebador específico de la diana que comprende una porción en 3' que se aparea específicamente con la secuencia de nucleótidos diana y un primer cebador adaptador que se aparea específicamente con una secuencia complementaria del ácido nucleico adaptador, en el que el segundo cebador específico de la diana comprende una parte de la cola en 5' que no se aparea específicamente con la secuencia de nucleótidos diana; y (f) amplificar un producto de amplificación de la etapa (e) mediante reacción en cadena de la polimerasa utilizando un cebador de cola que comprende una porción en 3' que se aparea específicamente con una secuencia complementaria de la porción de cola en 5' del segundo cebador específico de la diana y un segundo cebador adaptador que se aparea específicamente con una secuencia complementaria del ácido nucleico adaptador, en donde el cebador de cola comprende una porción en 5' que no se aparea específicamente con una secuencia complementaria del segundo cebador específico de la diana. La reacción de síntesis de la segunda cadena de la etapa (c) puede cebarse con cualquier fragmento de ácido nucleico presente en una muestra que comprende la molécula de ácido nucleico. Por ejemplo, la muestra puede comprender una mezcla compleja de ácidos nucleicos que son capaces de disociarse de una cadena complementaria y volver a aparearse con una cadena diferente presente en la mezcla.

El primer cebador específico de la diana puede comprender un resto de captura. Los métodos proporcionados en el presente documento pueden comprender además, después de la etapa (d) y antes de la etapa (e), capturar el producto de ligadura poniendo en contacto el producto de ligadura con un compañero de unión del resto de captura.

La reacción de síntesis de la primera cadena puede llevarse a cabo utilizando al menos un tipo de nucleótido modificado con el resto de captura y en donde el producto de la reacción de síntesis de la primera cadena comprende al menos un resto de captura. Los métodos proporcionados en el presente documento pueden comprender además, después de la etapa (d) y antes de la etapa (e), capturar el producto de ligadura poniendo en contacto el producto de ligadura con un compañero de unión inmovilizado del al menos un resto de captura.

Los métodos proporcionados en el presente documento pueden comprender además, después de la etapa (c) y antes de la etapa (d), capturar el ácido nucleico bicatenario poniendo en contacto el ácido nucleico bicatenario con un compañero de unión inmovilizado del al menos un resto de captura.

La divulgación proporciona métodos para determinar un repertorio inmunitario en una muestra, en el que los métodos implican: (a) obtener una muestra que comprende una molécula de ácido nucleico que codifica al menos uno de un receptor inmunitario y una inmunoglobulina; (b) poner en contacto la molécula de ácido nucleico de la muestra con un primer cebador específico de la diana que se aparea específicamente con una secuencia de nucleótidos diana de la molécula de ácido nucleico en condiciones de hibridación; (c) llevar a cabo una reacción de síntesis de la primera cadena que está cebada por un primer cebador específico de la diana hibridado y que usa la molécula de ácido nucleico como molde; (d) llevar a cabo una reacción de síntesis de la segunda cadena que usa un producto de la reacción de síntesis de la primera cadena como molde para generar un ácido nucleico bicatenario; (e) ligar un ácido nucleico adaptador al ácido nucleico bicatenario para producir un producto de ligadura; y (f) amplificar el producto de ligadura mediante reacción en cadena de la polimerasa utilizando un segundo cebador específico de la diana que comprende una porción en 3' que se aparea específicamente con la secuencia de nucleótidos diana y un primer cebador adaptador que se aparea específicamente con una secuencia complementaria del ácido nucleico adaptador, en el que el segundo cebador específico de la diana comprende una parte de la cola en 5' que no se aparea específicamente con la secuencia de nucleótidos diana; (g) amplificar un producto de amplificación de la etapa (f) mediante reacción en cadena de la polimerasa utilizando un cebador de cola que comprende una porción en 3' que se aparea específicamente con una secuencia complementaria de la porción de cola en 5' del segundo cebador específico de la diana y un segundo cebador adaptador que se aparea específicamente con una secuencia complementaria del ácido nucleico adaptador, en donde el cebador de cola comprende una porción en 5' que no se aparea específicamente con una secuencia complementaria del segundo cebador específico de la diana; y (h) secuenciar un producto de amplificación de la etapa (g) utilizando un primer y un segundo cebador de secuenciación.

En algunas realizaciones, el receptor inmunitario comprende un TCR. En algunas realizaciones, la inmunoglobulina comprende un BCR. En algunas realizaciones, la secuencia de nucleótidos diana corresponde a una secuencia

genéticamente recombinada. La secuencia de nucleótidos diana puede corresponder al menos a uno de un locus TCR (por ejemplo, TCRA, TCRB, TCRG o TCRD). La secuencia de nucleótidos diana puede corresponder al menos a uno de un locus BCR (por ejemplo, IGH, IGK o IGL). En algunas realizaciones, la muestra comprende un linfocito T, un linfocito B o una mezcla de los mismos. En algunas realizaciones, la muestra se obtiene de un sujeto. En algunas realizaciones, el sujeto es un ser humano. El sujeto puede ser un roedor. Por ejemplo, el sujeto puede ser un ratón, una rata, un jerbo, un hámster, una cobaya, una chinchilla, una ardilla o una forma humanizada de cualquiera de dichos roedores (por ejemplo, un roedor que expresa TCR humanos y/o BCR humanos).

La divulgación proporciona métodos para preparar ácidos nucleicos para análisis, en el que los métodos implican: (a) poner en contacto una molécula de ácido nucleico que comprende una secuencia de nucleótidos diana con un primer cebador específico de diana que se aparea específicamente con la secuencia de nucleótidos diana en condiciones para generar una molécula de ácido nucleico hibridada con cebador; (b) poner en contacto la molécula de ácido nucleico hibridada con cebador con una pluralidad de tipos de nucleótidos para su incorporación en una primera cadena que es complementaria a la molécula de ácido nucleico, en donde al menos uno de la pluralidad de tipos de nucleótidos comprende un resto de captura; (c) llevar a cabo una reacción de síntesis de la primera cadena que está cebada por el primer cebador específico de diana de la molécula de ácido nucleico hibridada con cebador y que utiliza la molécula de ácido nucleico de la molécula de ácido nucleico hibridada con cebador como molde, en donde un producto de la reacción de síntesis de la primera cadena comprende al menos un resto de captura; (d) llevar a cabo una reacción de síntesis de la segunda cadena que está cebada por un fragmento de la molécula de ácido nucleico y que usa el producto de la reacción de síntesis de la primera cadena como molde para generar un ácido nucleico bicatenario que comprende el al menos un resto de captura; (e) ligar un ácido nucleico adaptador al ácido nucleico bicatenario para producir un producto de ligadura que comprende el al menos un resto de captura; y (f) amplificar el producto de ligadura mediante reacción en cadena de la polimerasa utilizando un segundo cebador específico de la diana que comprende una porción en 3' que se aparea específicamente con la secuencia de nucleótidos diana y un primer cebador adaptador que se aparea específicamente con una secuencia complementaria del ácido nucleico adaptador, en donde el segundo cebador específico de la diana comprende una parte de la cola en 5' que no se aparea específicamente con la secuencia de nucleótidos diana. La reacción de síntesis de la segunda cadena de la etapa (d) puede cebarse con cualquier fragmento de ácido nucleico presente en una muestra que comprende la molécula de ácido nucleico. Por ejemplo, la muestra puede comprender una mezcla compleja de ácidos nucleicos que son capaces de disociarse de una cadena complementaria y volver a aparearse con una cadena diferente presente en la mezcla.

Los métodos proporcionados en el presente documento pueden comprender además, después de la etapa (e) y antes de la etapa (f), capturar el producto de ligadura poniendo en contacto el producto de ligadura con un compañero de unión inmovilizado del resto de captura.

Los métodos proporcionados en el presente documento pueden comprender además, después de la etapa (d) y antes de la etapa (e), capturar el ácido nucleico bicatenario poniendo en contacto el ácido nucleico bicatenario con un compañero de unión inmovilizado del resto de captura.

La divulgación proporciona métodos para preparar ácidos nucleicos para análisis, en el que los métodos implican: (a) poner en contacto una molécula de ácido nucleico que comprende una secuencia de nucleótidos diana con un cebador modificado con un resto de captura que comprende una porción en 3' que se aparea específicamente con la secuencia de nucleótidos diana en condiciones de hibridación y una porción de cola en 5' que no se aparea específicamente con la secuencia de nucleótidos diana; (b) llevar a cabo una reacción de síntesis de la primera cadena que está cebada por un cebador modificado con un resto de captura hibridado y que usa la molécula de ácido nucleico como molde; (c) poner en contacto un producto de la reacción de síntesis de la primera cadena con una pluralidad de cebadores aleatorios en condiciones de hibridación, en donde cada uno de la pluralidad de cebadores aleatorios comprende una porción de cola en 5' no aleatoria que comprende una secuencia común; (d) llevar a cabo una reacción de síntesis de la segunda cadena que está cebada por al menos uno de la pluralidad de cebadores aleatorios y que usa un producto de la reacción de síntesis de la primera cadena como molde para generar un ácido nucleico bicatenario que comprende el resto de captura; (e) capturar el ácido nucleico bicatenario que comprende el resto de captura poniendo en contacto el ácido nucleico bicatenario con un compañero de unión del resto de captura; (f) amplificar el ácido nucleico bicatenario mediante reacción en cadena de la polimerasa utilizando un primer cebador de cola y un primer cebador específico de la diana que se aparea específicamente con la secuencia de nucleótidos diana, en donde el primer cebador de cola comprende una porción en 3' que se aparea específicamente con una complementaria de la secuencia común y una porción de cola en 5' que no se aparea específicamente con una complementaria de la secuencia común; y (g) amplificar un producto de amplificación de la etapa (f) mediante la reacción en cadena de la polimerasa utilizando un segundo cebador específico de diana, un segundo cebador de cola que se aparea específicamente con una complementaria de la porción de cola en 5' del primer cebador de cola, en donde el segundo cebador específico de la diana comprende una porción en 3' que se aparea específicamente con la secuencia de nucleótidos diana y una porción de cola en 5' que no se aparea específicamente con la secuencia de nucleótidos diana.

Breve descripción de los dibujos

Las realizaciones no limitantes de la descripción se describirán a modo de ejemplo con referencia a las figuras adjuntas, que son esquemáticas y no están destinadas a ser dibujadas a escala. Por razones de claridad, no todos los

componentes están etiquetados en cada figura, ni se muestran todos los componentes de cada realización de la invención donde la ilustración no es necesaria para permitir que los expertos en la técnica entiendan la invención. En las figuras:

- 5 la Figura 1 es una ilustración de una estrategia para preparar una molécula de ácido nucleico para el análisis usando un cebador modificado con el resto de captura.
 La Figura 2 es una ilustración de una estrategia para preparar una molécula de ácido nucleico para el análisis usando nucleótidos modificados con el resto de captura.
 La Figura 3 ilustra una comparación de estrategias de preparación de ácidos nucleicos utilizando un cebador modificado con un resto de captura o nucleótidos modificados con un resto de captura.
 10 La Figura 4 es una ilustración que muestra la superposición de clonotipos entre réplicas en relación con la cantidad entrante. La intersección de todas las muestras produce sesenta y seis clonotipos superpuestos.
 La Figura 5 es un gráfico que ilustra que el análisis pareado de muestras duplicadas demuestra la naturaleza altamente reproducible del ensayo.
 15 Las Figuras 6A y 6B representan una comparación entre todos los clones (Figura 6A) y los clones superpuestos (Figura 6B).
 La Figura 7 ilustra que la cantidad entrante impulsa la complejidad y la diversidad de la observación, con la diversidad en relación con el tamaño de la muestra en la parte superior y un gráfico que representa el índice de diversidad de Shannon en la parte inferior.
 20 La Figura 8 es un gráfico que demuestra un seguimiento de clones cuantitativo y altamente reproducible.
 La Figura 9 ilustra que el seguimiento clonal a través de diluciones es altamente reproducible entre laboratorios independientes.
 La Figura 10 es un gráfico que representa la reproducibilidad intralaboratorio utilizando una serie de diluciones de muestras de Jurkat.
 25 La Figura 11 es un gráfico que representa la reproducibilidad entre laboratorios utilizando la serie de dilución de muestras de Jurkat.
 La Figura 12 es un esquema, no se muestra a escala, que generalmente representa las ubicaciones de los cebadores del receptor de linfocitos T.
 La Figura 13 es un esquema, no dibujado a escala, que generalmente representa ubicaciones de cebadores de cadena pesada de inmunoglobulina (IgH).
 30 La Figura 14 es un esquema que muestra la estructura de un cebador de transcriptasa inversa (RT).

Descripción detallada

- 35 Entre otros aspectos, la presente divulgación proporciona técnicas mejoradas relacionadas con la preparación de bibliotecas de ácidos nucleicos para el análisis del repertorio inmunitario. Tal como se describe en el presente documento, una molécula de ácido nucleico diana (por ejemplo, ARNm) que comprende una secuencia de nucleótidos diana conocida se pone en contacto con un cebador específico de la diana y se extiende en una reacción de síntesis de la primera cadena utilizando la molécula de ácido nucleico como molde. En algunas realizaciones, la reacción de
 40 síntesis de la primera cadena puede llevarse a cabo de manera que un producto de la reacción de síntesis de la primera cadena comprenda un resto de captura y un compañero de unión del resto de captura puede usarse para capturar (por ejemplo, aislar) el producto. En consecuencia, aspectos de la divulgación proporcionan técnicas útiles para enriquecer el producto de la síntesis de la primera cadena que es complementaria a la entrada inicial que comprende la molécula de ácido nucleico diana (por ejemplo, ARNm).
 45 En algunos aspectos, la divulgación se refiere al reconocimiento de que la incorporación de un resto de captura en una primera cadena sintetizada a partir de una molécula de ácido nucleico diana (por ejemplo, ARN) durante la preparación de la biblioteca minimiza la presencia de ácidos nucleicos no diana durante el enriquecimiento de ácidos nucleicos que comprenden una secuencia de nucleótidos diana. Esto puede ser particularmente ventajoso cuando se preparan bibliotecas de ácidos nucleicos para la secuenciación del repertorio inmunitario. Por ejemplo, aunque las secuencias genómicas de TCR y BCR estarán presentes en todas las células, el ARNm de TCR y BCR solo se expresará en linfocitos T y linfocitos B, respectivamente. La evaluación del repertorio inmunitario se basa en el análisis de estos genes después del procesamiento (por ejemplo, la recombinación, el corte y empalme, etc.) para evaluar el panorama de secuencias presente en un sistema. En consecuencia, en algunas realizaciones, las técnicas
 50 proporcionadas en el presente documento permiten la captura selectiva de ARNm expresados a partir de loci inmunitarios recombinados. Por ejemplo, como se describe en el presente documento, cuando una primera cadena se sintetiza directamente a partir de un ácido nucleico diana, el resto de captura permite el enriquecimiento del producto deseado mientras minimiza el arrastre de ácidos nucleicos no diana.
 55 En algunos aspectos, la divulgación proporciona un método para preparar ácidos nucleicos para análisis, que comprende sintetizar una primera cadena de ácido nucleico utilizando un cebador modificado con un resto de captura y una molécula de ácido nucleico que comprende una secuencia de nucleótidos diana como molde, ligar un ácido nucleico adaptador a la primera cadena de ácido nucleico y capturar la primera cadena de ácido nucleico ligada al adaptador con un compañero de unión del resto de captura.
 60 En algunas realizaciones, el cebador modificado con el resto de captura es un cebador modificado con un resto de
 65

biotina. Una representación de este método se muestra en la Figura 1, que proporciona un ejemplo no limitante de un método que implica un cebador modificado con un resto de biotina. En esta realización, una molécula de ARN (por ejemplo, ARNm) se aparea con un cebador de ADN que se modifica para comprender un resto de biotina. El cebador modificado con el resto de biotina se extiende en una reacción de síntesis de la primera cadena para generar un híbrido de ADN/ARN. El ARN del híbrido ADN/ARN se degrada mediante la acción de una ribonucleasa, y los fragmentos de ARN que permanecen apareados con el ADN sirven como cebadores en una reacción de síntesis de la segunda cadena para generar un ADNc bicatenario. En algunas realizaciones, la reacción de síntesis de la segunda cadena puede ser cebada por cualquier fragmento de ácido nucleico presente en una muestra que comprende la molécula de ácido nucleico. Por ejemplo, en algunas realizaciones, la muestra comprende una mezcla compleja de ácidos nucleicos que son capaces de disociarse de una cadena complementaria y reaparecer a una cadena diferente presente en la mezcla. El ADNc bicatenario se somete a reparación de extremos y colas de dA para generar salientes 3' adecuados para una reacción de ligadura. Después de la ligadura de un ácido nucleico adaptador al ADNc bicatenario, estas moléculas de la biblioteca se capturan o aíslan, del adaptador no ligado utilizando un sustrato recubierto de estreptavidina (por ejemplo, un sustrato paramagnético).

En algunas realizaciones, tras la captura, se realiza una primera ronda de amplificación por PCR ADNc bicatenario liado por adaptador capturado inmovilizado en sustrato. Incluso en otras realizaciones, el ADNc bicatenario ligado con adaptador capturado se eluye del sustrato paramagnético antes de la primera ronda de PCR. Se puede realizar la elución de ácidos nucleicos ligados a adaptadores capturados de sustratos paramagnéticos, a modo de ejemplo y no de limitación, usando un reactivo químico y/o calor. En algunas realizaciones, el reactivo químico es una base (por ejemplo, NaOH). En algunas realizaciones, el ADNc bicatenario ligado con adaptador capturado se eluye con una concentración baja (por ejemplo, menos de 1 M, menos de 0,5 M, menos de 0,1 M, menos de 0,05 M, menos de 0,01 M, menos de 0,001 M, menos de 0,0001 M) de NaOH. En algunas realizaciones, el ADNc bicatenario ligado con adaptador capturado se eluye con una baja concentración de NaOH y calor.

El ADNc bicatenario ligado al adaptador inmovilizado o eluido se somete a una primera ronda de amplificación por PCR utilizando un primer cebador adaptador que se aparea con una complementaria del adaptador y un cebador específico de la diana que se hibrida con una secuencia de nucleótidos diana y tiene un región de la cola hibridada que comprende una secuencia común. De esta manera, el primer cebador adaptador se cebe con la cadena generada por el cebador específico de la diana. Se lleva a cabo una segunda ronda de amplificación por PCR utilizando un cebador de cola que se aparea con una complementaria de la secuencia común y un segundo cebador adaptador que se aparea con una complementaria del adaptador. Tal como se muestra, el cebador de cola incluye un cebador con código de barras índice 1. En algunas realizaciones, el cebador de cola incluye un cebador con código de barras índice 2. En algunas realizaciones, el ácido nucleico adaptador incluye un cebador índice 1 y el cebador de cola incluye un cebador índice 2.

En algunos aspectos, la divulgación proporciona un método para preparar ácidos nucleicos para análisis, que comprende poner en contacto una molécula de ácido nucleico hibridada con cebador que comprende una secuencia de nucleótidos diana con un nucleótido modificado con un resto de captura para su incorporación en una cadena recién sintetizada, someter la mezcla a condiciones de extensión para sintetizar una primera cadena de ácido nucleico que comprende al menos un nucleótido modificado con resto de captura, ligar un ácido nucleico adaptador a la primera cadena de ácido nucleico y capturar la primera cadena de ácido nucleico ligada al adaptador con un compañero de unión del resto de captura.

En algunas realizaciones, el resto de captura del nucleótido modificado con el resto de captura es un resto de biotina. Por ejemplo, la Figura 2 representa una realización no limitante de un método que implica el uso de nucleótidos modificados con biotina. En esta realización, una molécula de ARN (por ejemplo, ARNm) se aparea con un cebador de ADN. Se lleva a cabo una reacción de síntesis de la primera cadena utilizando nucleótidos biotinilados para generar un híbrido de ADN/ARN biotinilado. El ARN del híbrido ADN/ARN se degrada mediante la acción de una ribonucleasa, y los fragmentos de ARN que permanecen apareados con el ADN sirven como cebadores en una reacción de síntesis de la segunda cadena para generar un ADNc bicatenario. En algunas realizaciones, la reacción de síntesis de la segunda cadena puede ser cebada por cualquier fragmento de ácido nucleico presente en una muestra que comprende la molécula de ácido nucleico. Por ejemplo, en algunas realizaciones, la muestra comprende una mezcla compleja de ácidos nucleicos que son capaces de disociarse de una cadena complementaria y volver a aparearse con una cadena diferente presente en la mezcla. El ADNc bicatenario se somete a reparación de extremos y colas de dA para generar salientes 3' adecuados para una reacción de ligadura. Después de la ligadura de un ácido nucleico adaptador al ADNc bicatenario, estas moléculas de la biblioteca se capturan o aíslan, del adaptador no ligado utilizando una perla paramagnética recubierta de estreptavidina.

Como se describe en lo anterior, los aspectos de la descripción proporcionan técnicas para preparar una molécula de ácido nucleico para análisis que puede implicar el uso de un resto de captura para capturar un producto de ácido nucleico generado durante un proceso de preparación. La Figura 3 es una ilustración que compara diferentes estrategias para la preparación de ácidos nucleicos usando restos de captura. Las etapas de preparación seleccionadas se muestran para un primer proceso 310 que incorpora un resto de captura en un producto de ácido nucleico usando un cebador modificado con el resto de captura (por ejemplo, como se ilustra en la Figura 1) y para un segundo proceso 320 que incorpora un resto de captura en un ácido nucleico. producto ácido que usa nucleótidos

modificados con resto de captura (por ejemplo, como se ilustra en la Figura 2). En cada uno del primer proceso 310 y segundo proceso 320, una molécula de ácido nucleico se expone a un cebador específico de la diana en condiciones para promover la hibridación del cebador (etapa i)). Tal como se muestra, el cebador específico de la diana del primer proceso 310 es un resto cebador 312 modificado con el resto de captura, mientras que el cebador específico de la diana del segundo proceso 320 utiliza un primer cebador 322 específico de la diana que no incluye una modificación del resto de captura.

Después de la hibridación del cebador, se realiza una reacción de síntesis de la primera cadena (etapa ii)) en el primer proceso 310 usando una pluralidad de tipos de nucleótidos 314 para incorporarlos en una primera cadena 316 recién sintetizada usando la molécula de ácido nucleico como molde. En comparación, se realiza una reacción de síntesis de la primera cadena (etapa ii)) en el segundo proceso 320 usando una pluralidad de tipos de nucleótidos 324, de los cuales al menos uno de la pluralidad es un nucleótido modificado con resto de captura. A modo de ejemplo y no de limitación, se muestra un nucleótido C que tiene una modificación del resto de captura. De esta manera, durante la síntesis de la primera cadena utilizando la molécula de ácido nucleico como molde, el nucleótido modificado con el resto de captura se incorpora en la primera cadena 326 recién sintetizada en una posición complementaria a un nucleótido G en la molécula de ácido nucleico.

Después de la síntesis de la primera cadena en el primer proceso 310 o en el segundo proceso 320, un ácido nucleico que comprende el resto de captura se somete opcionalmente a un procesamiento adicional (etapa iii)) que implica modificaciones adicionales en el ácido nucleico. Los ejemplos de procesamiento adicional pueden incluir, sin limitación, la síntesis de segunda cadena, reparación de extremos, ligadura del adaptador y otras etapas de procesamiento descritas en otra parte del presente documento. Después del procesamiento adicional opcional en el primer proceso 310 o el segundo proceso 320, un producto de ácido nucleico que comprende el resto de captura se pone en contacto con un compañero de unión del resto de captura con el fin de capturar el producto de ácido nucleico (etapa iv)) generado a partir de cualquiera de los procesos. Tal como se muestra, la captura del producto se puede realizar utilizando un compañero de unión 352 del resto de captura, que está opcionalmente inmovilizado en un sustrato 354. Cuando el sustrato 354 es un sustrato paramagnético, el producto de ácido nucleico puede aislarse por exposición a un campo magnético 356.

En algunos aspectos, la divulgación proporciona un método para preparar una biblioteca de ácidos nucleicos para el análisis de un repertorio inmunitario. Por ejemplo, en algunas realizaciones, la biblioteca de ácidos nucleicos se prepara a partir de una muestra que comprende una secuencia de ácidos nucleicos que codifica un receptor inmunitario (por ejemplo, un TCR), una inmunoglobulina (por ejemplo, un BCR) o una mezcla de los mismos.

Repertorio Inmunitario

Dado que el sistema inmunitario adaptativo funciona en parte mediante la expansión clonal de células que expresan moléculas únicas de unión a antígenos, medir con precisión los cambios en la abundancia total de cada clon de linfocitos T o B es importante para comprender la dinámica de una respuesta inmunitaria adaptativa. Utilizando los avances en la secuenciación de alto rendimiento, recientemente ha surgido un nuevo campo de la inmunología molecular para perfilar los vastos repertorios de TCR y BCR. En algunas realizaciones, las técnicas descritas en el presente documento son útiles para analizar un clonotipo de célula inmunitaria.

Como se utiliza en el presente documento, un "clonotipo" se refiere a una secuencia de nucleótidos recombinada con éxito que surge durante un proceso de reordenamiento para uno o más genes que codifican una cadena de receptor inmunitario o una porción de la misma. En algunas realizaciones, una "secuencia de nucleótidos recombinada con éxito" se refiere a una secuencia de nucleótidos que está compuesta por ARNm y ha sufrido una recombinación genética para producir un clonotipo único. En consecuencia, en algunas realizaciones, las técnicas proporcionadas en la presente divulgación pueden ser útiles para detectar un reordenamiento con éxito de un loci IR que se expresa como ARNm, pero no un reordenamiento *IR per se*. En algunas realizaciones, un clonotipo se refiere a una secuencia de nucleótidos que corresponde a un ARNm que codifica una sola cadena de receptor. En algunas realizaciones, las técnicas proporcionadas en la presente divulgación pueden ser útiles para detectar una o más hipermutaciones somáticas. En algunas realizaciones, las técnicas proporcionadas en la presente divulgación pueden ser útiles para detectar isotipos de TCR y/o BCR (por ejemplo, los isotipos IgH A, D, E, G y M, y subclases seleccionadas de los mismos). En algunas realizaciones, un clonotipo es una secuencia de nucleótidos recombinada de un linfocito T o un linfocito B. En algunas realizaciones, un clonotipo codifica un TCR o BCR, o una parte del mismo. En algunas realizaciones, los clonotipos pueden codificar todo o una parte de un reordenamiento VDJ de IgH, un reordenamiento DJ de IgH, un reordenamiento VJ de IgK, un reordenamiento VJ de IgL, un reordenamiento VDJ de TCR β , un reordenamiento DJ de TCR β , un reordenamiento VJ de TCR α , un reordenamiento VJ de TCR γ , un reordenamiento VDJ de TCR δ , un reordenamiento VD de TCR δ , un reordenamiento del elemento de delección kappa (KDE), o similar. En algunas realizaciones, los clonotipos tienen secuencias lo suficientemente largas como para representar o reflejar la diversidad de las moléculas inmunitarias de las que derivan. En consecuencia, en algunas realizaciones, los clonotipos pueden variar ampliamente en longitud. En algunas realizaciones, los métodos de la divulgación son útiles para determinar un perfil de clonotipo.

Como se utiliza en el presente documento, un "perfil de clonotipo" se refiere a una lista de clonotipos distintos y sus

abundancias relativas que se derivan de una población de linfocitos. En algunas realizaciones, la población de linfocitos se obtiene a partir de una muestra de tejido. Un perfil de clonotipo está relacionado con el concepto inmunológico de "repertorio inmunitario". En algunas realizaciones, un perfil de clonotipo incluye una amplia variedad de listas y cantidades de ácidos nucleicos que codifican receptores inmunitarios reordenados, que pueden derivarse de subconjuntos seleccionados de linfocitos (por ejemplo, linfocitos infiltrantes de tejidos, subconjuntos inmunofenotípicos o similares) o que pueden codificar porciones de receptores inmunitarios que tienen una diversidad reducida en comparación con los receptores inmunitarios completos. En algunas realizaciones, los perfiles de clonotipo pueden comprender al menos 10^3 clonotipos distintos, al menos 10^4 clonotipos distintos, al menos 10^5 clonotipos distintos, al menos 10^6 clonotipos distintos o al menos 10^7 clonotipos distintos. En algunas realizaciones, los perfiles de clonotipos pueden comprender entre aproximadamente 1 y 500.000 clonotipos distintos. En algunas realizaciones, los clonotipos pueden comprender entre aproximadamente 1 y 1.000.000 clonotipos distintos (por ejemplo, entre aproximadamente 1 y aproximadamente 100.000, entre aproximadamente 100.000 y aproximadamente 200.000, entre aproximadamente 200.000 y aproximadamente 300.000, entre aproximadamente 300.000 y aproximadamente 400.000, entre aproximadamente 400.000 y aproximadamente 500.000, entre aproximadamente 500.000 y aproximadamente 600.000, entre aproximadamente 600.000 y aproximadamente 700.000, entre aproximadamente 700.000 y aproximadamente 800.000, entre aproximadamente 800.000 y aproximadamente 900.000, entre aproximadamente 900.000 y aproximadamente 1.000.000 de clonotipos distintos). En algunas realizaciones, tales perfiles de clonotipos pueden comprender además abundancias o frecuencias relativas de cada uno de los distintos clonotipos. En un aspecto, un perfil de clonotipo es un conjunto de distintas secuencias de nucleótidos recombinadas (con sus abundancias) que codifican TCR o BCR, o fragmentos de los mismos, respectivamente, en una población de linfocitos de un individuo, en donde las secuencias de nucleótidos del conjunto tienen una correspondencia biunívoca con distintos linfocitos o sus subpoblaciones clonales para sustancialmente todos los linfocitos de la población. En un aspecto, los segmentos de ácido nucleico que definen los clonotipos se seleccionan de modo que su diversidad (por ejemplo, el número de secuencias de ácido nucleico distintas en el conjunto) sea lo suficientemente grande como para que sustancialmente cada linfocito T o linfocito B o clon de las mismas en un individuo lleve una secuencia de ácido nucleico única de tal repertorio. Esto es, preferentemente cada clon diferente de una muestra tiene un clonotipo diferente. En otros aspectos, la población de linfocitos correspondiente a un repertorio puede ser de linfocitos B circulantes, o puede ser de linfocitos T circulantes, o puede ser subpoblaciones de cualquiera de las poblaciones anteriores, incluyendo pero sin limitación, linfocitos T CD4+ o linfocitos T CD8+, u otras subpoblaciones definidas por marcadores de superficie celular, o similares. En algunas realizaciones, tales subpoblaciones pueden adquirirse tomando muestras de tejidos particulares, por ejemplo, de médula ósea, ganglios linfáticos o similares, o clasificando o enriqueciendo células de una muestra (como sangre periférica) en función de uno o más marcadores de superficie celular, tamaño, morfología o similares. En otros aspectos más, la población de linfocitos correspondiente a un repertorio puede derivarse de tejidos enfermos, tal como un tejido tumoral, un tejido infectado o similar. En algunas realizaciones, un perfil de clonotipo que comprende ácidos nucleicos correspondientes a cadenas de TCR y/o BCR o fragmentos de las mismas comprende una serie de secuencias de nucleótidos distintas en el rango de aproximadamente 1 a aproximadamente 25, de aproximadamente 25 a aproximadamente 50, de aproximadamente 50 a aproximadamente 100, de aproximadamente 11 a aproximadamente 250, de aproximadamente 250 a aproximadamente 500, de aproximadamente 500 a aproximadamente 1000, de aproximadamente 1000 a aproximadamente 2500, de aproximadamente 2500 a aproximadamente 5000, de aproximadamente 5000 a aproximadamente 7500, de aproximadamente 7500 a aproximadamente 10000, de aproximadamente 10000 a aproximadamente 25000, de aproximadamente 25000 a aproximadamente 50000, de aproximadamente 50000 a aproximadamente 100000, de aproximadamente 100000 a aproximadamente 250000, de aproximadamente 250000 a aproximadamente 500000, de aproximadamente 500000 a aproximadamente 750000, de aproximadamente 750000 a aproximadamente 1000000.

En algunas realizaciones, un perfil de clonotipo comprende un conjunto de secuencias de nucleótidos que codifican sustancialmente todos los segmentos de la región V(D)J de un BCR (por ejemplo, una cadena de IgH). En un aspecto, "prácticamente todo" como se usa en el presente documento significa cada segmento que tiene una abundancia relativa del 0,0001 por ciento o más, 0,0005 por ciento o más, 0,001 por ciento o más, 0,005 por ciento o más, 0,01 por ciento o más, 0,05 por ciento o más o 0,1 por ciento o más; o en otro aspecto, "prácticamente todo" como se usa en el presente documento significa cada segmento que tiene una abundancia relativa del 0,0001 por ciento o más. En algunas realizaciones, no todos los segmentos V, D y J están representados. En otra realización, un perfil de clonotipo comprende un conjunto de secuencias de nucleótidos que codifican sustancialmente todos los segmentos de la región V(D)J de un TCR (por ejemplo, cadena β del TCR, cadena δ del TCR). En otra realización, un perfil de clonotipo comprende un conjunto de secuencias de nucleótidos que tienen longitudes en el intervalo de 1-25, 1-50, 1-100, 1-200, 1-300, 1-400, 1-450, 1-500, 25-100, 25-200, 25-300, 25-400, 25-450, 100-200, 100-300, 100-400, 100-450, 100-500, 200-300, 200-400, 200-450, 200-500, 300-400, 300-450, 300-500, 400-450, 400-500, 450-500 o más nucleótidos e incluyendo segmentos de la las regiones V, D y J de un TCR (por ejemplo, cadena β del TCR, cadena δ del TCR). En otra realización, un perfil de clonotipo comprende un conjunto de secuencias de nucleótidos que tienen longitudes en el intervalo de 1-25, 1-50, 1-100, 1-200, 1-300, 1-400, 1-450, 1-500, 25-100, 25-200, 25-300, 25-400, 25-450, 25-500, 100-200, 100-300, 100-400, 100-450, 100-500, 200-300, 200-400, 200-450, 200-500, 300-400, 300-450, 300-500, 400-450, 400-500, 450-500 o más nucleótidos e incluyendo segmentos de la las regiones V, D y J de un BCR (por ejemplo, una cadena de IgH). En otra realización, un perfil de clonotipo comprende una serie de secuencias de nucleótidos distintas que es sustancialmente equivalente a la cantidad de linfocitos que expresan un BCR distinto (por ejemplo, cadena de IgH, cadena de IgK, cadena de IgL). En otra realización, un perfil de clonotipo

comprende una serie de secuencias de nucleótidos distintas que es sustancialmente equivalente a la cantidad de linfocitos que expresan un TCR distinto (por ejemplo, cadena β del TCR, cadena δ DEL TCR, cadena α del TCR, cadena γ del TCR). En aún otra realización, "sustancialmente equivalente" significa que con un 99 por ciento de probabilidad, un perfil de clonotipo incluirá una secuencia de nucleótidos que codifica una BCR (por ejemplo, IgH, IgK, IgL) o TCR (por ejemplo, TCR β , TCR δ , TCR α , TCR γ) o parte del mismo transportado o expresado por cada linfocito de una población de un individuo. En aún otra realización, "sustancialmente equivalente" significa que, con una probabilidad del 99 por ciento, un repertorio de secuencias de nucleótidos incluirá una secuencia de nucleótidos que codifica un BCR (por ejemplo, IgH, IgK, IgL) o TCR (por ejemplo, TCR β , TCR δ , TCR α , TCR γ) o parte del mismo transportado o expresado por cada linfocito presente en una muestra.

En algunas realizaciones, los perfiles de clonotipo se obtienen a partir de muestras de células inmunitarias, que están presentes en una amplia variedad de tejidos. En algunas realizaciones, las células inmunitarias de interés incluyen linfocitos T y/o linfocitos B. En algunas realizaciones, los linfocitos T (linfocitos T) incluyen, por ejemplo, células que expresan TCR. En algunas realizaciones, Los linfocitos B (linfocitos B) incluyen, por ejemplo, células que expresan BCR. En algunas realizaciones, los linfocitos T incluyen linfocitos T colaboradores (linfocitos T efectoras o células Th), linfocitos T citotóxicos (CTL), linfocitos T de memoria y linfocitos T reguladores, que pueden distinguirse por marcadores de superficie celular. En algunas realizaciones, una muestra de células inmunitarias también puede comprender linfocitos B. En algunas realizaciones, los linfocitos B incluyen, por ejemplo, plasmocitos, linfocitos B de memoria, linfocitos B1, linfocitos B2, linfocitos B de la zona marginal y linfocitos B foliculares. Los linfocitos B pueden expresar inmunoglobulinas (también denominadas en el presente documento anticuerpos o receptores de linfocitos B).

Receptores de linfocitos T

El sistema inmunitario adaptativo emplea varias estrategias para generar un repertorio de receptores de antígenos de linfocitos T y B (por ejemplo, receptores inmunitarios adaptativos) con diversidad suficiente para reconocer el universo de patógenos potenciales. La capacidad de las linfocitos T para reconocer el universo de antígenos asociados a diversos tipos de cáncer u organismos infecciosos se la confiere su TCR, que es un heterodímero de una cadena α (alfa) del locus TCRA y una cadena β (beta) del locus TCRB o un heterodímero de una cadena γ (gamma) del locus TCRG y una cadena δ (delta) del locus del TCRD. Las proteínas que componen estas cadenas están codificadas por ADN, que en las células linfoides emplea un mecanismo de reordenamiento único para generar la tremenda diversidad del TCR. Este receptor de reconocimiento inmunitario de múltiples subunidades se asocia con el complejo CD3 y se une a los péptidos presentados por las proteínas del complejo principal de histocompatibilidad (MHC) de clase I o MHC de clase II en la superficie de las células presentadoras de antígenos (CPA). La unión de TCR al péptido antigénico en la CPA es el evento central en la activación de los linfocitos T, que ocurre en una sinapsis inmunológica en el punto de contacto entre el linfocito T y la CPA.

Cada péptido del TCR contiene regiones determinantes de complementariedad variables (CDR), así como regiones marco (FR) y una región constante. La diversidad de secuencias de los linfocitos $\alpha\beta$ T está determinada en gran medida por la secuencia de aminoácidos de los bucles de la tercera región determinante de la complementariedad (CDR3) de los dominios variables de las cadenas α y β , cuya diversidad es el resultado de la recombinación entre la variable ($V\beta$), diversidad ($D\beta$) y los segmentos génicos de unión ($J\beta$) en el locus de la cadena β y entre segmentos génicos $J\alpha$ y $J\alpha$ análogos en el locus de la cadena α , respectivamente. La existencia de múltiples segmentos génicos de este tipo en los loci de la cadena α y β del TCR permite codificar un gran número de secuencias distintas de CDR3. La diversidad de secuencias de CDR3 aumenta aún más mediante la adición y delección independientes de nucleótidos en las uniones $V\beta$ - $D\beta$, $D\beta$ - $J\beta$ y $V\alpha$ - $J\alpha$ durante el proceso de reordenamiento del gen TCR. En este sentido, la inmunocompetencia se deriva de la diversidad de TCR.

El heterodímero $\gamma\delta$ TCR se distingue del $\alpha\beta$ TCR porque codifica un receptor que interactúa estrechamente con el sistema inmunitario innato y reconoce el antígeno de una manera no dependiente de HLA. El TCR $\gamma\delta$ se expresa temprano en el desarrollo y tiene una distribución anatómica especializada, especificidades únicas de patógenos y moléculas pequeñas, y un amplio espectro de interacciones celulares innatas y adaptativas. Un patrón sesgado de expresión de los segmentos V y J de TCRRy se establece temprano en la ontogenia. Como consecuencia, el diverso repertorio de TCRRy en tejidos adultos es el resultado de una extensa expansión periférica después de la estimulación por exposición ambiental a patógenos y moléculas tóxicas.

Los procesos para generar diversidad de un TCR son similares a los descritos para las inmunoglobulinas. La cadena α del TCR se genera por recombinación de VJ, mientras que la cadena β se genera por recombinación de V(D)J. De modo similar, la generación de la cadena γ de TCR implica la recombinación de VJ, mientras que la generación de la cadena δ del TCR ocurre por recombinación de V(D)J. La intersección de estas regiones específicas (V y J para la cadena α o γ , V D y J para la cadena β o δ) corresponde a la región CDR3 que es importante para el reconocimiento del antígeno-MHC. Es la combinación única de los segmentos en esta región, junto con adiciones palindrómicas y aleatorias de N- y P-nucleótidos, lo que explica el repertorio de unión de TCR. Adicionalmente, la región CDR3 comienza con la segunda cisteína conservada en la región 3' del gen $v\beta$ y termina con la fenilalanina conservada codificada por la región 5' del gen $l\beta$. Por lo tanto, las secuencias amplificadas se pueden traducir informáticamente para localizar la cisteína conservada, obtener la secuencia peptídica intermedia y tabular los recuentos de cada clon

único en la muestra.

Receptores de los linfocitos B

5 Las inmunoglobulinas (Ig), también denominadas en el presente documento receptores de linfocitos B (BCR), son proteínas expresadas por los linfocitos B que consisten en cuatro cadenas polipeptídicas, dos cadenas pesadas (cadenas H) del locus IGH y dos cadenas ligeras (cadenas L) del locus IGK o IGL, formando una estructura H2L2. Las cadenas H y L contienen cada una tres CDR implicadas en el reconocimiento de antígenos, así como regiones marco y un dominio constante, análogo al TCR. Las cadenas H de las Ig se expresan inicialmente como isoformas unidas a la membrana utilizando los exones de la región constante IGM o IGD, pero después del reconocimiento del antígeno, la región constante puede cambiar de clase a varios isotipos adicionales, incluyendo la IGG, IGE e IGA. Al igual que con el TCR, la diversidad de Ig vírgenes dentro de un individuo está determinada principalmente por las CDR hipervariables. Análogamente a la TCRB, el dominio CDR3 de las cadenas H se crea mediante la unión combinatoria de los segmentos génicos VH, DH y JH. La diversidad de secuencias de dominio hipervariable aumenta aún más mediante la adición y eliminación independientes de nucleótidos en las uniones VH-DH, DH-JH y VH-JH durante el proceso de reordenamiento del gen de Ig. A diferencia del TCR, la diversidad de secuencias de Ig aumenta aún más por la hipermutación somática (SHM) en todo el gen de IG reorganizado después de que una célula B ingenua reconoce inicialmente un antígeno. El proceso de SHM no está restringido a la CDR3 y, por lo tanto, puede introducir cambios en la secuencia de la línea germinal en regiones marco, CDR1 y CDR2, así como en la CDR3 reorganizada somáticamente.

En algunos aspectos, el ADN y el ARN analizados en los métodos descritos en el presente documento pueden corresponder a secuencias que codifican inmunoglobulinas de cadena pesada (IgH) con regiones constantes (α , δ , ϵ , y μ) o inmunoglobulinas de cadena ligera (IgK o IGL) con regiones constantes (λ o κ). Cada anticuerpo tiene dos cadenas ligeras idénticas y dos cadenas pesadas idénticas. Cada cadena está compuesta por una región constante (C) y una región variable. Para la cadena pesada, la región variable está compuesta por segmentos variable (V), de diversidad (D) y unión (J). Varias secuencias distintas que codifican para cada tipo de estos segmentos están presentes en el genoma. Un evento de recombinación VDJ específico ocurre durante el desarrollo de un linfocito B, marcar esa célula para generar una cadena pesada específica. La diversidad en la cadena ligera se genera de manera similar, excepto que no hay región D, por lo que solo hay recombinación VJ. La mutación somática a menudo ocurre cerca del sitio de la recombinación, causando la adición o delección de varios nucleótidos, aumentando aún más la diversidad de cadenas pesadas y ligeras generadas por los linfocitos B. La posible diversidad de anticuerpos generados por una célula B es entonces el producto de las diferentes cadenas pesadas y ligeras. Las regiones variables de las cadenas pesada y ligera contribuyen a formar la región o sitio de reconocimiento (o unión) de antígeno. A esta diversidad se suma un proceso de hipermutación somática que puede ocurrir después de que se monta una respuesta específica contra algún epítipo.

En algunas realizaciones, los anticuerpos son producidos por secuencias de Ig genómicas recombinadas en células de linaje B. Las cadenas ligeras de inmunoglobulina derivan de los genes κ o λ . Los genes λ se componen de cuatro genes constantes (C) y aproximadamente treinta genes variables (V). Por lo contrario, los genes κ están compuestos por un gen C y 250 genes V. La familia de genes de cadenas pesadas está compuesta por varios cientos de genes V, quince genes D y cuatro genes de unión (J). La recombinación somática durante la diferenciación de linfocitos B elige aleatoriamente una combinación V-D-J en la cadena pesada y una combinación V-J en la cadena ligera κ o λ . Porque hay tantos genes, son posibles millones de combinaciones únicas. Los genes V también sufren una hipermutación somática después de la recombinación, generando mayor diversidad. A pesar de esta complejidad subyacente, es posible usar docenas de cebadores dirigidos a secuencias conservadas para secuenciar el complemento completo de cadena pesada y ligera en varias reacciones multiplexadas.

Análisis del repertorio inmunitario

En algunos aspectos, las técnicas descritas en el presente documento pueden usarse para determinar la presencia de una condición de interés. En algunas realizaciones, la determinación de la presencia de una condición de interés puede relacionarse con aplicaciones de diagnóstico, cuando un sujeto tiene o se sospecha que tiene la afección. En algunas realizaciones, la determinación de la presencia de una condición de interés puede ser útil para medidas predictivas con el fin de un tratamiento preventivo. En algunas realizaciones, el análisis de un repertorio inmunitario puede indicar la presencia de una condición de interés. Por ejemplo, en algunas realizaciones, un historial de cáncer puede reflejarse en la presencia de secuencias de receptores inmunitarios que se unen a uno o más antígenos del cáncer. En algunas realizaciones, la presencia de enfermedad autoinmunitaria puede reflejarse en la presencia de secuencias de receptores inmunitarios que se unen a autoantígenos. En algunas realizaciones, las afecciones relacionadas con la autoinmunidad pueden evaluarse en un punto particular en el tiempo o rastrearse durante un período de tiempo usando las técnicas descritas en el presente documento. En algunas realizaciones, las afecciones relacionadas con la autoinmunidad incluyen esclerosis múltiple (EM), enfermedad celíaca, diabetes mellitus de tipo 1, sarcoidosis, lupus eritematoso sistémico (LES), síndrome de Sjögren, granulomatosis eosinofílica con poliangéftis, tiroiditis de Hashimoto, enfermedad de Graves, púrpura trombocitopénica idiopática, enfermedad de Addison, artritis reumatoide (AR), espondilitis anquilosante, polimiositis (PM) y dermatomiositis (DM). En consecuencia, en algunas realizaciones, los métodos descritos en el presente documento pueden usarse para determinar si el tratamiento de

una afección es apropiado. Por ejemplo, en algunas realizaciones, la abundancia de linfocitos T y linfocitos B malignos puede rastrearse a lo largo del tiempo a través de secuencias específicas de CDR3.

En algunos aspectos, los métodos proporcionados por la divulgación pueden usarse para determinar un tratamiento terapéutico óptimo. En algunas realizaciones, se puede determinar un tratamiento terapéutico óptimo analizando el repertorio inmunitario en una muestra y, sobre la base de esa información, seleccionar una terapia apropiada, dosis o modalidad de tratamiento que sea óptima para estimular o suprimir una respuesta inmunitaria específica mientras se minimiza la toxicidad indeseable. En algunas realizaciones, un tratamiento se optimiza mediante la selección de un tratamiento que minimiza la toxicidad indeseable mientras proporciona una actividad eficaz. Por ejemplo, en algunas realizaciones, un sujeto (por ejemplo, un paciente) puede evaluarse para determinar el repertorio inmunitario relevante para una enfermedad autoinmunitaria, y puede seleccionarse un régimen inmunosupresor sistémico o dirigido basándose en esa información.

En algunos aspectos, las técnicas proporcionadas por la divulgación pueden usarse para evaluar la progresión de una afección en un sujeto. Por ejemplo, en algunas realizaciones, el análisis de un repertorio inmunitario se puede utilizar para evaluar la progresión, el estancamiento o la regresión de una afección. En tales realizaciones, el repertorio inmunitario puede evaluarse ventajosamente antes, durante y/o después del tratamiento con un agente terapéutico para evaluar la eficacia del agente terapéutico en el tratamiento de la afección. Por ejemplo, en algunas realizaciones, los métodos descritos en el presente documento pueden ser útiles para detectar los cambios más tempranos a lo largo de la vía de una enfermedad (por ejemplo, una vía de carcinogénesis, vía inflamatoria, etc.), y/o para monitorizar la eficacia de varias terapias e intervenciones preventivas.

En algunos aspectos, los métodos desvelados en el presente documento también se pueden utilizar para analizar los efectos de los agentes sobre las células del sistema inmunitario. Por ejemplo, en algunas realizaciones, se puede realizar un análisis de los cambios en el repertorio inmunitario después de la exposición a uno o más compuestos de prueba para analizar el(los) efecto(s) de los compuestos de prueba en un individuo. En tales realizaciones, estos análisis pueden ser útiles para múltiples propósitos, por ejemplo, en el desarrollo de terapias inmunosupresoras o inmunoestimulantes. En algunas realizaciones, los agentes a analizar para determinar su valor terapéutico potencial pueden ser cualquier compuesto, molécula pequeña, proteína, lípido, hidrato de carbono, ácido nucleico u otro agente apropiado para uso terapéutico. En algunas realizaciones, se realizan pruebas *in vivo*, por ejemplo, utilizando un modelo animal, para determinar los efectos sobre el repertorio inmunitario.

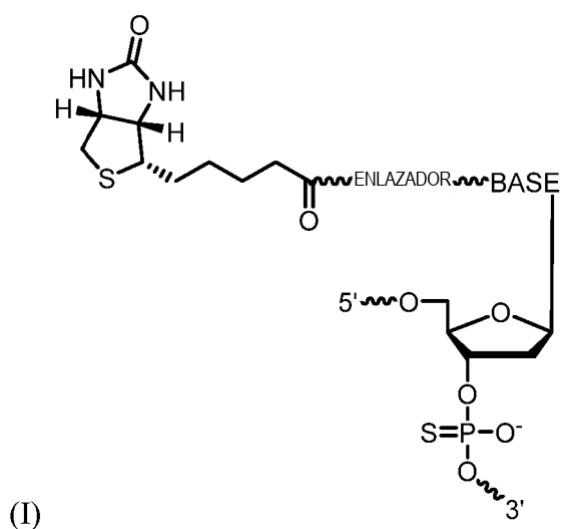
En algunas realizaciones, el análisis de un repertorio inmunitario se puede utilizar para determinar los efectos de una exposición a antígeno en un organismo. En algunas realizaciones, los ácidos nucleicos se obtienen de un organismo después de que el organismo haya sido expuesto a un antígeno (por ejemplo, después de la vacunación). En algunas realizaciones, los ácidos nucleicos se obtienen de un organismo antes de que el organismo haya sido expuesto a un antígeno. En algunas realizaciones, comparar la diversidad del repertorio inmunitario presente antes y después del desafío puede ayudar al análisis de la respuesta del organismo al desafío.

Resto de captura

Los aspectos de las técnicas descritas en el presente documento se refieren al uso de un resto de captura para aislar una molécula de interés (por ejemplo, un ácido nucleico, un producto de ligadura, etc.). Como se utiliza en el presente documento, un "resto de captura" se refiere a un resto que está configurado para interactuar selectivamente con un compañero de unión con el fin de capturar (por ejemplo, aislar/purificar) la molécula de interés.

Un resto de captura y un compañero de unión del resto de captura pueden comprender cualquier par de unión adecuado. En algunas realizaciones, un par de unión puede interactuar selectivamente mediante unión covalente o no covalente. En algunas realizaciones, un par de unión puede interactuar selectivamente por hibridación, enlaces iónicos, enlaces de hidrógeno, interacciones de van der Waals o cualquier combinación de estas fuerzas. En algunas realizaciones, un resto de captura y/o un compañero de unión puede comprender, por ejemplo, biotina, avidina, estreptavidina, digoxigenina, inosina, avidina, secuencias GST, secuencias GST modificadas, secuencias de reconocimiento de biotina ligasa (BiTag), marcas S, marcas SNAP, sitios de enterocinasa, sitios de trombina, anticuerpos o dominios de anticuerpo, fragmentos de anticuerpo, antígenos, receptores, dominios de receptor, fragmentos de receptores o combinaciones de los mismos.

En algunas realizaciones, un resto de captura comprende un resto de biotina. En algunas realizaciones, las técnicas descritas en el presente documento son útiles para preparar muestras de ácido nucleico para análisis. En consecuencia, en algunas realizaciones, una molécula de ácido nucleico comprende un resto de captura de biotina. En algunas realizaciones, la molécula de ácido nucleico comprende al menos un nucleótido modificado con resto de captura que comprende un resto de biotina. En algunas realizaciones, el nucleótido modificado con resto de captura comprende la estructura general de fórmula (I):

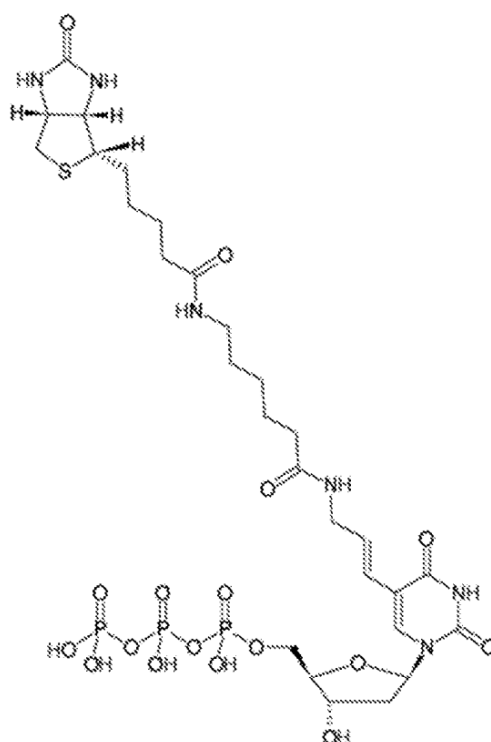


5 Tal como se muestra en la fórmula (I), un nucleótido modificado con resto de captura puede comprender un resto de biotina unido a una nucleobase de un nucleótido. Por ejemplo, en algunas realizaciones, el resto de biotina comprende biotina-trietilenglicol, bis-biotina, biotina fotoescindible, destiobiotina, destiobiotina-trietilenglicol o azida de biotina. En la Tabla 1 se muestran ejemplos no limitantes de nucleótidos modificados con resto de captura.

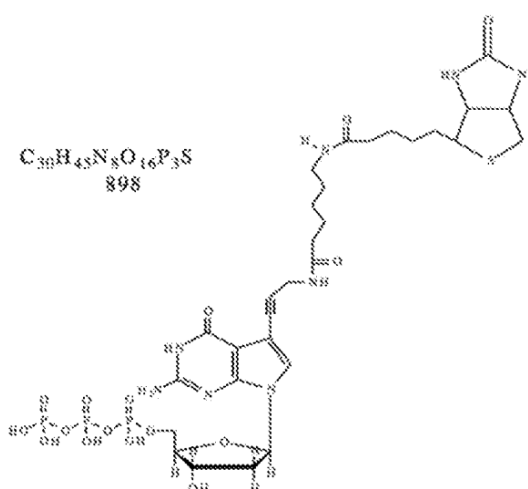
Tabla 1. Estructuras de ejemplo de nucleótidos modificados con resto de captura

<p>Biotina-11-dATP</p>	
<p>Biotina-11-dCTP</p>	

(continuación)

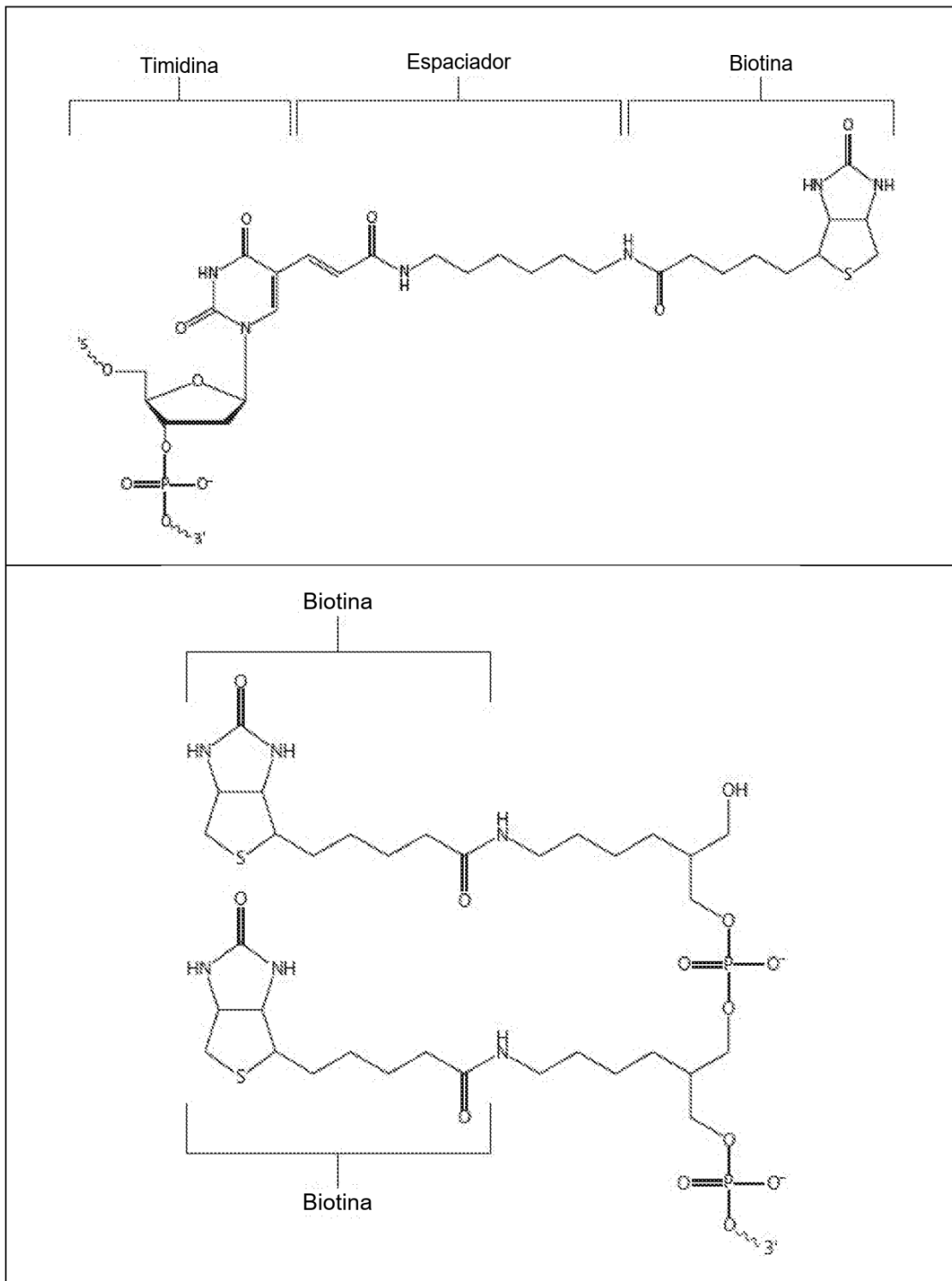


Biotina-11-dUTP

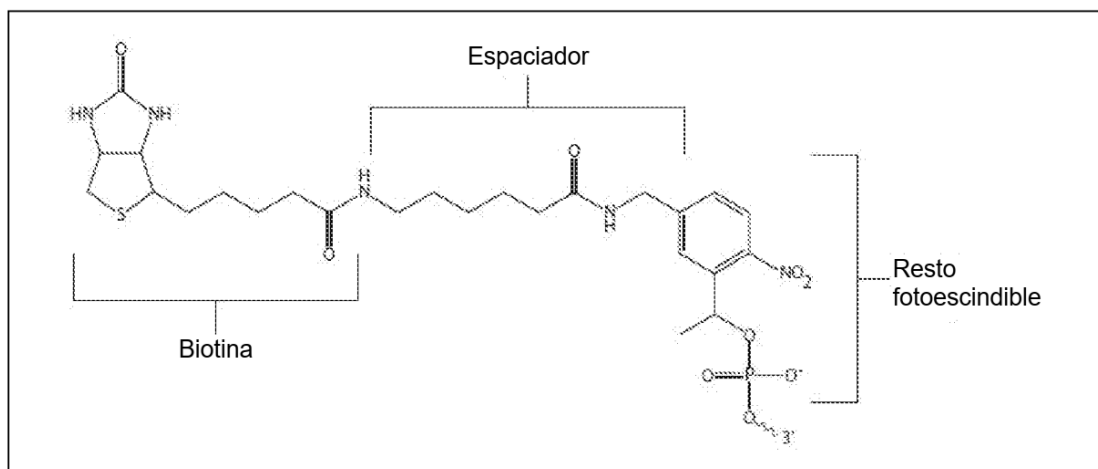


Biotina-11-dGTP

(continuación)



(continuación)



En algunas realizaciones, un nucleótido modificado con un resto de captura comprende un enlazador entre el resto de
 5 captura y una nucleobase del nucleótido. En algunas realizaciones, el resto de captura se une covalentemente a la nucleobase a través de un enlazador de cualquier longitud adecuada. En algunas realizaciones, el resto de captura se une covalentemente a la nucleobase a través de un enlazador de 5 a 20 átomos de longitud. En algunas realizaciones, el enlazador comprende una cadena alifática. En algunas realizaciones, un enlazador comprende $-(CH_2)_n-$, en donde n es un número entero de 1 a 20, ambos inclusive. En algunas realizaciones, n es un número entero de 1 a 10, ambos
 10 inclusive. En ciertas realizaciones, un enlazador comprende una cadena heteroalifática. En algunas realizaciones, un enlazador comprende un resto de polietilenglicol. En algunas realizaciones, un enlazador comprende un resto de polipropilenglicol. En algunas realizaciones, un enlazador comprende $-(CH_2CH_2O)_n-$, en donde n es un número entero de 1 a 20, ambos inclusive. En algunas realizaciones, un enlazador comprende $-(CH_2CH_2O)_n-$, en donde n es un número entero de 1 a 10, ambos inclusive. En ciertas realizaciones, un enlazador comprende uno o más arilenos. En
 15 algunas realizaciones, un enlazador comprende uno o más fenilenos (por ejemplo, fenileno sustituido en para). En ciertas realizaciones, un enlazador comprende un centro quiral. En ciertas realizaciones, un enlazador comprende uno o más fosfatos, una cadena alifática, una cadena heteroalifática y una o más amidas (por ejemplo, $-C(=O)NH-$).

En algunas realizaciones, un nucleótido modificado con resto de captura es biotina-n-dNTP, en donde n es un número
 20 entero de 5 a 20 que representa el número de átomos de enlace entre un grupo carbonilo del resto de biotina y la posición de unión en una nucleobase del NTP.

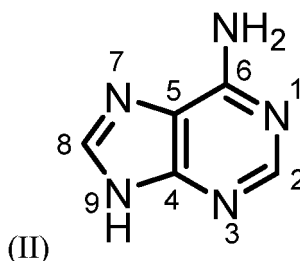
En algunas realizaciones, un compañero de unión está unido a un soporte insoluble. Por lo tanto, en algunas
 25 realizaciones, la molécula de interés puede inmovilizarse en un soporte insoluble a través de una interacción de unión selectiva formada entre un resto de captura y un compañero de unión del resto de captura unido al soporte insoluble.

En algunas realizaciones, el soporte insoluble comprende una perla u otra superficie sólida. Por ejemplo, en algunas
 30 realizaciones, la perla es una perla paramagnética. El uso de perlas para el aislamiento es bien conocido en la técnica, y se puede usar cualquier método de aislamiento de perlas adecuado con las técnicas descritas en el presente documento. En algunas realizaciones, las perlas pueden ser útiles para el aislamiento porque las moléculas de interés se pueden unir a las perlas y las perlas se pueden lavar para eliminar los componentes de la solución que no están adheridos a las perlas, permitiendo la purificación y el aislamiento. En algunas realizaciones, las perlas se pueden separar de otros componentes en la solución en función de propiedades TALES como el tamaño, la densidad o las propiedades dieléctricas, iónicas y magnéticas.

En algunas realizaciones, el soporte insoluble es una perla magnética. El uso de perlas permite que el resto de captura
 35 de ácido nucleico derivatizado se separe de una mezcla de reacción mediante centrifugación o filtración, o, en el caso de perlas magnéticas, por aplicación de un campo magnético. En algunas realizaciones, se pueden introducir perlas magnéticas, mezclar, eliminar y liberar en solución usando campos magnéticos. En algunas realizaciones, los procesos que utilizan perlas magnéticas pueden automatizarse. En algunas realizaciones, las perlas se pueden funcionalizar usando química bien conocida para proporcionar una superficie que tenga una funcionalización adecuada para unir un compañero de unión de un resto de captura. La derivatización de superficies para permitir la unión del resto de captura es convencional en la técnica. Por ejemplo, el recubrimiento de superficies con estreptavidina permite la unión de un resto de captura biotinilado. El recubrimiento de superficies con estreptavidina se ha descrito en, por
 40 ejemplo, el documento de Patente de Estados Unidos N.º 5.374.524 de Miller. En algunas realizaciones, se pueden usar superficies sólidas que no sean perlas. En algunas realizaciones, las superficies sólidas pueden ser superficies planas, tales como los que se utilizan para micromatrices de hibridación, o las superficies sólidas pueden ser el relleno de una columna de separación.

En algunas realizaciones, un compañero de unión de un resto de captura se puede unir a un soporte insoluble antes, simultáneamente con, o después de, unir el resto de captura. En algunas realizaciones, puede ser preferible poner en contacto un resto de captura con un compañero de unión del resto de captura mientras ambos están en solución. En tales realizaciones, el complejo de resto de captura:compañero de unión puede entonces inmovilizarse sobre un soporte insoluble poniendo en contacto el complejo con una superficie apropiadamente derivatizada. Por lo tanto, en algunas realizaciones, la molécula de interés puede aislarse a través de un complejo formado entre un resto de captura unido a la molécula de interés y un compañero de unión del resto de captura.

En algunas realizaciones, puede ser deseable unir el resto de captura a una nucleobase de un nucleótido. De este modo, el extremo 3' permanece libre para ligarse opcionalmente a un ácido nucleico adaptador mientras que el resto de captura está disponible para ser capturado por un compañero de unión. En algunas realizaciones, el nucleótido modificado con el resto de captura comprende una nucleobase seleccionada del grupo que consiste en adenina, guanina, timina, uracilo y citosina, o un derivado de los mismos. Por ejemplo, en algunas realizaciones, el nucleótido modificado con el resto de captura comprende una nucleobase de adenina o un derivado de la misma. En algunas realizaciones, el resto de captura está unido covalentemente a la nucleobase de adenina o derivado de la misma en la posición 5, 6, 7 u 8. En algunas realizaciones, el resto de captura está unido covalentemente a la nucleobase de adenina en la posición 7. En la fórmula (II) se representa un esquema de numeración para un anillo de adenina:



En algunas realizaciones, puede ser deseable modificar una o más posiciones en una base nitrogenada que está unida a un resto de captura. Por ejemplo, en algunas realizaciones, la posición 7 de la nucleobase de adenina es un átomo de carbono. Sin embargo, debe apreciarse que cualquier átomo capaz de formar un enlace covalente adicional (por ejemplo, C, O, N, S, etc.) puede sustituirse en una posición en una nucleobase adecuada para la unión de un resto de captura. En algunas realizaciones, después de capturar los fragmentos ligados con adaptador, la biblioteca se somete a amplificación para enriquecer las secuencias de nucleótidos diana.

Preparación de ácidos nucleicos para análisis

Los aspectos de la divulgación proporcionan métodos mejorados para determinar la secuencia de nucleótidos contigua a una secuencia de nucleótidos diana conocida (por ejemplo, una secuencia de nucleótidos diana conocida de un receptor inmunitario). Los métodos de secuenciación tradicionales generan información de secuencia aleatoriamente (por ejemplo, secuenciación de "escopeta") o entre dos secuencias conocidas que se utilizan para diseñar cebadores. Por lo contrario, algunos de los métodos descritos en el presente documento, en algunas realizaciones, permiten determinar la secuencia de nucleótidos (por ejemplo, secuenciar) aguas arriba o aguas abajo de una región única de la secuencia conocida con un elevado nivel de especificidad y sensibilidad.

En algunas realizaciones, las técnicas descritas en el presente documento permiten el enriquecimiento de secuencias de nucleótidos diana a partir de una muestra de ácido nucleico. En algunas realizaciones, la muestra de ácido nucleico comprende ADN genómico. En algunas realizaciones, la muestra de ácido nucleico comprende ADNc. En algunas realizaciones, el ADNc se puede preparar realizando una reacción de síntesis de la primera cadena utilizando un cebador modificado con resto de captura que se hibrida con un ácido nucleico diana y realizando una reacción de síntesis de la segunda cadena utilizando un fragmento del ácido nucleico diana como cebador.

Purificación de muestras

En algunas realizaciones, los ácidos nucleicos y/o sus productos de amplificación pueden aislarse de enzimas, cebadores o componentes tampón antes y/o después de cualquier paso apropiado de un método. Puede usarse cualquier método adecuado para aislar ácidos nucleicos. En algunas realizaciones, el aislamiento puede comprender limpieza por inmovilización reversible en fase sólida (SPRI). Los métodos para la limpieza de SPRI son bien conocidos en la técnica, por ejemplo, Purificación con PCR de Agencourt AMPure XP (n.º de cat. A63880, Beckman Coulter; Brea, CA). En algunas realizaciones, las enzimas pueden inactivarse mediante tratamiento térmico. En algunas realizaciones, los dNTP no marcados se eliminan mediante tratamiento enzimático. En algunas realizaciones, se lleva a cabo una etapa de limpieza (por ejemplo, una limpieza SPRI) para eliminar los cebadores no extendidos o en exceso (por ejemplo, cebadores modificados con resto de captura, cebadores específicos de la diana, cebadores adaptadores, etc.).

En algunas realizaciones, la limpieza SPRI se refiere al uso de perlas paramagnéticas que se unen al ADN. Por ejemplo, en algunas realizaciones, la limpieza SPRI utiliza perlas que tienen un núcleo de poliestireno rodeado por una fina capa de magnetita, lo que hace que las perlas sean paramagnéticas (es decir, las perlas se agregan solo cuando se exponen a un campo magnético). En algunas realizaciones, la perla está recubierta por moléculas que comprenden grupos carboxilo que proporcionan grupos cargados para la unión al ADN. En algunas realizaciones, la limpieza SPRI se lleva a cabo en presencia de polietilenglicol (PEG) y sal, que trabajan juntos como agentes de agregación para activar las perlas para unirse de forma reversible al ADN. En algunas realizaciones, un SPRI comprende mezclar una muestra de ADN con perlas paramagnéticas y permitir que las perlas se unan al ADN, aplicar un campo magnético para agregar las perlas unidas al ADN, aclarar las perlas con etanol (por ejemplo, etanol al 70 %) y eluir el ADN de las perlas paramagnéticas.

En algunas realizaciones, los cebadores no hibridados se pueden eliminar de una preparación de ácido nucleico usando métodos apropiados (por ejemplo, purificación, digestión, etc.). En algunas realizaciones, se utiliza una nucleasa (por ejemplo, exonucleasa I) para eliminar los cebadores de una preparación. En algunas realizaciones, dichas nucleasas se inactivan por calor después de la digestión del cebador. Una vez inactivadas las nucleasas, se puede añadir un conjunto adicional de cebadores junto con otros componentes apropiados (por ejemplo, enzimas, tampones) para realizar una reacción de amplificación adicional.

En algunas realizaciones, las etapas de los métodos proporcionados en el presente documento comprenden opcionalmente una etapa intermedia de purificación de la muestra. En algunas realizaciones, una etapa de purificación de muestra comprende una etapa de lavado. En algunas realizaciones, una etapa de purificación de muestra comprende la limpieza SPRI (por ejemplo, AMPure). Por ejemplo, un método de preparación de ácidos nucleicos para análisis puede comprender: (i) lavar un ácido nucleico inmovilizado en sustrato; y (ii) liberar el ácido nucleico inmovilizado lavado del sustrato o superficie paramagnética.

Adaptador de ácido nucleicos

Como se utiliza en el presente documento, la expresión "ácido nucleico adaptador", "adaptador de ácido nucleico", o "adaptador" se refiere a una molécula de ácido nucleico que puede ligarse a un ácido nucleico que comprende una secuencia de nucleótidos diana para proporcionar uno o más elementos útiles durante la amplificación y/o secuenciación de la secuencia de nucleótidos diana. En algunas realizaciones, un adaptador es monocatenario. En algunas realizaciones, un adaptador es bicatenario. En algunas realizaciones, un adaptador bicatenario comprende un primer extremo dúplex ligable y un segundo extremo no apareado. En algunas realizaciones, un adaptador comprende una cadena de amplificación y una cadena de bloqueo. En algunas realizaciones, la cadena de amplificación comprende una porción no apareada en 5' y una porción dúplex en 3'. En algunas realizaciones, la cadena de amplificación comprende además un saliente en 3'. En algunas realizaciones, el saliente en 3' es un saliente en T en 3'. En algunas realizaciones, la cadena de amplificación comprende secuencias de nucleótidos idénticas a un primer y segundo cebador adaptador. En algunas realizaciones, la cadena de bloqueo del adaptador comprende una porción dúplex en 5' y una porción en 3' no extensible. En algunas realizaciones, la cadena de bloqueo comprende además una parte no apareada en 3'. En algunas realizaciones, las partes dúplex de la cadena de amplificación y la cadena de bloqueo son sustancialmente complementarias y la parte dúplex tiene una longitud suficiente para permanecer en forma dúplex a la temperatura de ligadura.

En algunas realizaciones, la porción de la cadena de amplificación que comprende una secuencia nucleotídica idéntica a un primer y un segundo cebador de secuenciación puede estar comprendida, al menos en parte, por la porción no apareada en 5' de la cadena de amplificación.

En algunas realizaciones, el adaptador puede tener forma de "Y", es decir, el segundo extremo no apareado comprende una porción en 5' no apareada de una cadena de amplificación y una porción en 3' de una cadena de bloqueo. La porción desapareada en 3' de la cadena de bloqueo puede ser más corta, más larga o igual a la longitud de la porción desapareada en 5' de la cadena de amplificación. En algunas realizaciones, la porción no apareada en 3' de la cadena de bloqueo puede ser más corta que la porción no apareada en 5' de la cadena de amplificación. Los adaptadores en forma de Y tienen la ventaja de que la porción no apareada de la cadena de bloqueo no estará sujeta a la extensión 3' durante un régimen de PCR.

En algunas realizaciones, la cadena de bloqueo del adaptador puede comprender además una porción desapareada en 3' que no es sustancialmente complementaria a la porción desapareada en 5' de la cadena de amplificación, en donde la porción no apareada en 3' de la hebra de bloqueo no es sustancialmente complementaria a o sustancialmente idéntica a cualquiera de los cebadores. En algunas realizaciones, la cadena de bloqueo puede comprender además una porción no apareada en 3' que no se aparea específicamente con la porción no apareada en 5' de la cadena de amplificación a la temperatura de apareamiento, en donde la porción desapareada en 3' de la cadena de bloqueo no se apareará de forma específica con ninguno de los cebadores o sus complementarios a la temperatura de apareamiento. En algunas realizaciones, un ácido nucleico adaptador comprende, como mínimo, una secuencia índice de la muestra para multiplexación. Sin embargo, en algunas realizaciones, el adaptador nucleico comprende además un código de barras molecular aleatorio.

Extensión y amplificación

Los aspectos de la presente divulgación se refieren a técnicas que pueden comprender una o más reacciones de extensión (por ejemplo, síntesis de la primera cadena, síntesis de la segunda cadena) y/o una o más rondas de amplificación. Tal como se describe en el presente documento, las reacciones de extensión y amplificación se pueden realizar utilizando uno o más cebadores específicos de la diana.

Como se utiliza en el presente documento, un "cebador específico de la diana" se refiere a un cebador que comprende una secuencia que es complementaria a una secuencia de nucleótidos diana. En algunas realizaciones, se utiliza un cebador específico de la diana para cebar una reacción de síntesis de la primera cadena. Por ejemplo, en algunas realizaciones, el cebador específico de la diana es un cebador de transcriptasa inversa que se aparea con una molécula de ARNm que comprende una secuencia de nucleótidos diana. En algunas realizaciones, como se describe en el presente documento, un cebador modificado con resto de captura es un cebador específico de la diana que se puede usar para cebar una reacción de síntesis de la primera cadena. En algunas realizaciones, se utiliza un cebador específico de la diana para cebar una reacción de amplificación. Por ejemplo, en algunas realizaciones, los métodos descritos en el presente documento pueden incluir un paso de amplificación que utiliza un cebador específico de la diana que comprende una porción en 3' que se aparea específicamente con la secuencia de nucleótidos diana. En algunas realizaciones, la divulgación proporciona métodos que pueden incluir el uso de cebadores específicos de diana (por ejemplo, cebadores específicos de diana idénticos o diferentes) en más de una etapa.

En consecuencia, en algunas realizaciones de los métodos descritos en el presente documento, cuando el término cebador específico de la diana aparece en más de una etapa y se refiere a cebadores separados, se puede incluir terminología adicional para aclaración. Por ejemplo, en algunas realizaciones, se puede utilizar un cebador específico de la diana inicial en una reacción de síntesis de la primera cadena para generar un ADNc que se amplifica posteriormente utilizando un cebador específico de la diana diferente. En tales realizaciones, el cebador específico de la diana inicial puede denominarse "cebador modificado con resto de captura" mientras que el último cebador específico de la diana puede denominarse "cebador específico de la diana". Como alternativa, en algunas realizaciones, el cebador específico de la diana inicial y el último cebador específico de la diana pueden denominarse cebador específico de la diana "primero" y "segundo", respectivamente.

Debería apreciarse que, en algunas realizaciones, el uso de los términos "primero", "segundo", "tercero", etc. pueden usarse relativamente, de modo que estos términos pueden referirse a diferentes clases de cebadores según el contexto de la técnica que se describe. Por ejemplo, en algunas realizaciones, se utiliza un cebador de transcriptasa inversa específico de la diana con una molécula de ARNm para generar un ADNc, que se somete además a reacciones de PCR utilizando cebadores específicos de diana adicionales. En tales realizaciones, el cebador de transcriptasa inversa específico de la diana puede denominarse "primer cebador específico de la diana" y los cebadores de PCR posteriores se unen a una secuencia diana conocida denominada "segundo cebador específico de la diana" "tercer cebador específico de la diana", etc. En algunas realizaciones, se utiliza una pluralidad de cebadores aleatorios de transcriptasa inversa con una molécula de ARNm para generar un ADNc, que luego se somete a reacciones de PCR utilizando cebadores específicos de la diana. En tales realizaciones, la pluralidad de cebadores aleatorios no se denominan "específicos de la diana"; por consiguiente, si las reacciones de PCR posteriores utilizan cebadores específicos del objetivo, las expresiones "primer cebador específico de la diana", "segundo cebador específico de la diana", etc. pueden usarse de acuerdo con distintas reacciones (por ejemplo, en un método de preparación de ácidos nucleicos para la secuenciación).

En algunos aspectos, la divulgación proporciona métodos que pueden incluir la realización de una reacción de síntesis de la primera cadena utilizando un primer cebador específico de la diana (por ejemplo, un cebador modificado con el resto de captura). En algunas realizaciones, se lleva a cabo una primera ronda de amplificación utilizando un segundo cebador específico de la diana (por ejemplo, un cebador específico de la diana) y un primer cebador adaptador.

En algunas realizaciones, un "cebador específico de la diana" es un oligonucleótido que comprende una secuencia de ácido nucleico que puede hibridarse específicamente, en condiciones de apareamiento adecuadas, a una secuencia de nucleótidos diana de una molécula de ácido nucleico (por ejemplo, un ácido nucleico molde). En algunas realizaciones, los términos ordinales (por ejemplo, primero, segundo, tercero) se pueden utilizar para distinguir un cebador específico de diana de otro utilizado en diferentes etapas de un método de varias etapas. Por ejemplo, en algunas realizaciones, un segundo cebador específico de la diana es un cebador específico de la diana para su uso en una reacción de amplificación en un proceso que comprende una síntesis previa de la primera cadena que implica el uso de un primer cebador específico de la diana. En tales realizaciones, durante la amplificación, el segundo cebador específico de la diana genera una cadena que es complementaria a su molde y esta cadena complementaria es capaz de hibridarse con un primer cebador adaptador.

Como se utiliza en el presente documento, un "cebador adaptador" es un oligonucleótido que comprende una secuencia de ácido nucleico que puede aparearse específicamente, en condiciones de apareamiento adecuadas, a una secuencia complementaria de un ácido nucleico adaptador. En algunas realizaciones, un cebador adaptador (por ejemplo, un primer cebador adaptador) es idéntico a al menos una parte del adaptador, y se aparea con la cadena

complementaria generada por un cebador específico de la diana (por ejemplo, un segundo cebador específico de la diana) para permitir que la amplificación proceda.

En algunas realizaciones, en el primer ciclo de amplificación por PCR de la primera etapa de amplificación, un segundo cebador específico de la diana puede aparearse específicamente con una cadena molde de un ácido nucleico que comprende una secuencia de nucleótidos diana. En algunas realizaciones, dependiendo de la orientación con la cual se diseñó el primer cebador específico de diana, una secuencia aguas arriba o aguas abajo de la secuencia de nucleótidos diana se sintetizará como una cadena complementaria a la cadena molde. En algunas realizaciones, si, durante la fase de extensión de la PCR, el extremo 5' de una cadena molde termina en un adaptador ligado, el extremo 3' de la cadena complementaria recién sintetizada comprenderá una secuencia capaz de hibridarse con un primer cebador adaptador. En ciclos de amplificación por PCR posteriores, tanto el segundo cebador específico de la diana como el primer cebador adaptador serán capaces de aparearse de forma específica con las hebras apropiadas de la secuencia de ácido nucleico diana y pueden amplificarse la secuencia entre la secuencia de ácido nucleico diana conocida y el adaptador. En algunas realizaciones, un segundo cebador específico de la diana comprende una parte de la cola en 5' que no se aparea específicamente con la secuencia de nucleótidos diana. Por ejemplo, en algunas realizaciones, una porción de la cola en 5' puede comprender una región que proporciona un sitio de unión al cebador para reacciones de extensión posteriores (por ejemplo, durante la amplificación). En algunas realizaciones, se lleva a cabo una segunda ronda de amplificación utilizando un cebador de cola y un segundo cebador adaptador.

Como se utiliza en el presente documento, un "cebador de cola" es un oligonucleótido que comprende una secuencia de ácido nucleico que comprende una porción en 3' que puede aparearse específicamente, en condiciones de apareamiento adecuadas, a una secuencia complementaria de la porción de la cola en 5' de un cebador específico de la diana (por ejemplo, segundo cebador específico de la diana) compuesto por el amplicón resultante de un paso de amplificación anterior. En algunas realizaciones, un cebador de cola comprende una porción en 5' que no se aparea específicamente con la secuencia complementaria de la porción de cola en 5' del cebador específico de la diana. En algunas realizaciones, la porción en 5' del cebador de cola comprende al menos una de una región de índice de muestra, una región de unión al cebador de PCR, una región de código de barras molecular y una región de sitio de cebador de secuenciación. Aunque un cebador de cola, como se utiliza en el presente documento, generalmente se refiere a un cebador que se usa en una segunda ronda de PCR, debe tenerse en cuenta que el término puede usarse para referirse a cualquier cebador que se hibrida con una secuencia que es complementaria a una porción de la cola en 5' de un cebador usado en una reacción anterior.

En algunas realizaciones, un cebador adaptador (por ejemplo, un segundo cebador de adaptador) es idéntico a al menos una parte del adaptador, y se hibrida con la cadena complementaria generada por el cebador de cola para permitir que continúe la amplificación.

En algunas realizaciones, un segundo cebador adaptador está anidado en relación con un primer cebador adaptador. En algunas realizaciones, el uso de los cebadores adaptadores anidados elimina la posibilidad de producir amplicones finales que sean amplificables (por ejemplo, durante una PCR puente o PCR en emulsión) pero que no se puedan secuenciar, una situación que puede surgir durante los métodos semianidados. En otras situaciones, los planteamientos semianidados que utilizan un cebador idéntico a un cebador de secuenciación pueden dar como resultado el paso de los productos de amplificación no deseados de la primera etapa de PCR a la segunda etapa de PCR y producirían por último lecturas de secuenciación artificiales. En algunas realizaciones, un segundo cebador adaptador está anidado con respecto a un primer cebador adaptador en al menos 1 nucleótido, por ejemplo, en 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 o más nucleótidos. En algunas realizaciones, un segundo cebador adaptador está anidado con respecto a un primer cebador adaptador en aproximadamente 5 nucleótidos a aproximadamente 10 nucleótidos, de aproximadamente 10 nucleótidos a aproximadamente 15 nucleótidos, de aproximadamente 15 nucleótidos a aproximadamente 20 nucleótidos o de aproximadamente 20 nucleótidos o más.

Entre otros aspectos, las técnicas descritas en el presente documento pueden implicar el uso de uno o más cebadores anidados. En algunas realizaciones, el uso de cebadores anidados puede reducir la unión no específica en productos de PCR debido a la amplificación de sitios de unión de cebadores inesperados. Como se utiliza en el presente documento, el término "anidado" se usa para describir una relación posicional entre el sitio de hibridación de un cebador de un par de cebadores y el sitio de hibridación de otro cebador de otro par de cebadores. Por ejemplo, en algunas realizaciones, un segundo cebador está anidado por 1, 2, 3 o más nucleótidos en relación con un primer cebador, lo que significa que se une a un sitio en la cadena de la plantilla que está desplazada en el marco de 1, 2, 3 o más nucleótidos.

En algunas realizaciones, un cebador específico de la diana (por ejemplo, un segundo cebador específico de la diana) comprende una porción en 3' que se aparea específicamente con una secuencia de nucleótidos diana y una cola en 5' que no se aparea con la secuencia de nucleótidos diana. En algunas realizaciones, la cola en 5' comprende una secuencia de ácido nucleico que es idéntica a la porción en 3' de un cebador de cola. En algunas realizaciones, varios cebadores (por ejemplo, uno o más cebadores específicos de la diana y/o uno o más cebadores adaptadores) presentes en una reacción pueden comprender porciones de secuencia de cola en 5' idénticas.

En algunas realizaciones, una cola en 5' puede ser una secuencia rica en GC. En algunas realizaciones, una secuencia

de cola en 5' puede comprender al menos un 50 % de contenido de GC, al menos un 55 % de contenido de GC, al menos un 60 % de contenido de GC, al menos un 65 % de contenido de GC, al menos un 70 % de contenido de GC, al menos un 75 % de contenido de GC, al menos un 80 % de contenido de GC o contenido de GC superior. En algunas realizaciones, una secuencia de cola en 5' puede comprender al menos un 60 % de contenido de GC. En algunas realizaciones, una secuencia de cola en 5' puede comprender al menos un 65 % de contenido de GC.

En algunas realizaciones, una primera ronda de amplificación incluye un segundo cebador específico de la diana que comprende una cola en 5', un primer cebador adaptador y un cebador adicional. En algunas realizaciones, el cebador adicional comprende una porción en 3' que es idéntica a la cola en 5' del segundo cebador específico de la diana. En algunas realizaciones, el cebador adicional puede comprender secuencias adicionales en 5' a la secuencia de hibridación que pueden incluir código de barras, índice, secuencias adaptadoras o sitios de cebadores de secuenciación. En algunas realizaciones, el cebador adicional es un cebador índice/adaptador de secuenciación genérico.

En algunas realizaciones, dos cebadores específicos de la diana (por ejemplo, primer y segundo cebadores específicos de la diana) son sustancialmente complementarios a la misma cadena del ácido nucleico diana. En algunas realizaciones, las porciones de los primeros y segundos cebadores específicos a diana que se aparean de forma específica con la secuencia diana conocida pueden comprender un total de al menos 20 bases únicas de la secuencia nucleotídica diana conocida, por ejemplo, 20 o más bases únicas, 25 o más bases únicas, 30 o más bases únicas, 35 o más bases únicas, 40 o más bases únicas o 50 o más bases únicas. En algunas realizaciones, las porciones de los primeros y segundos cebadores específicos a diana que se aparean de forma específica con la secuencia diana conocida pueden comprender un total de al menos 30 bases únicas de la secuencia nucleotídica diana conocida.

En algunas realizaciones, el primer cebador adaptador puede comprender una secuencia de ácido nucleico idéntica a aproximadamente las 20 bases más 5' de la cadena de amplificación del adaptador y el segundo cebador adaptador puede comprender una secuencia de ácido nucleico idéntica a aproximadamente 30 bases de la cadena de amplificación del adaptador, con una base en 5' que es al menos 1 nucleótido en 3' del extremo 5' de la cadena de amplificación.

En algunas realizaciones, un ácido nucleico ligado al adaptador (por ejemplo, un producto de ligadura) es mínimo. En tales realizaciones, se puede utilizar un primer cebador adaptador que contenga una parte de la secuencia nucleica adaptadora en su extremo 3' y luego información adicional importante para el secuenciador en su extremo 5'. En tales realizaciones, se puede usar un segundo cebador adaptador que contiene, en su extremo 3', el extremo 5' del primer cebador adaptador. En tales realizaciones, el segundo cebador adaptador también puede tener una secuencia de nucleótidos que permita la secuenciación en su extremo 5'. En tales realizaciones, es posible producir, usando PCR, una biblioteca que sea compatible con el secuenciador.

Cebadores

En general, un cebador que comprende una secuencia que es complementaria a una secuencia de interés (por ejemplo, una secuencia diana o una secuencia adaptadora) puede consistir solo en una secuencia complementaria o también puede incluir una secuencia adicional que no es complementaria a la secuencia de interés (por ejemplo, una secuencia de cola, una secuencia adaptadora, una secuencia índice, etc.). En algunas realizaciones, un cebador también puede incluir restos no nucleótidos (por ejemplo, restos de captura, etc.).

En algunas realizaciones, los cebadores (por ejemplo, primer y segundo cebadores específicos de la diana, cebadores adaptadores primero y segundo, cebadores de cola, cebadores modificados con el resto de captura) se diseñan de modo que se aparearán de forma específica con sus secuencias complementarias a una temperatura de apareamiento de aproximadamente 61 a 72 °C, por ejemplo, de aproximadamente 61 a 69 °C, de aproximadamente 63 a 69 °C, de aproximadamente 63 a 67 °C, de aproximadamente 64 a 66 °C. En algunas realizaciones, los cebadores se diseñan de modo que se aparearán de forma específica con sus secuencias complementarias a una temperatura de apareamiento de menos de 72 °C. En algunas realizaciones, los cebadores se diseñan de modo que se aparearán de forma específica con sus secuencias complementarias a una temperatura de apareamiento de menos de 70 °C. En algunas realizaciones, los cebadores se diseñan de modo que se aparearán de forma específica con sus secuencias complementarias a una temperatura de apareamiento de menos de 68 °C. En algunas realizaciones, los cebadores se diseñan de modo que se aparearán de forma específica con sus secuencias complementarias a una temperatura de apareamiento de menos de aproximadamente 65 °C. En algunas realizaciones, los sistemas proporcionados en el presente documento están configurados para alterar la temperatura del recipiente (por ejemplo, ciclando entre diferentes intervalos de temperatura) para facilitar el apareamiento del cebador.

En algunas realizaciones, las porciones de los cebadores específicos a diana que se aparean de forma específica con la secuencia nucleotídica diana (por ejemplo, secuencia nucleotídica diana conocida) se aparearán de forma específica a una temperatura de aproximadamente 61 a 72 °C, por ejemplo, de aproximadamente 61 a 69 °C, de aproximadamente 63 a 69 °C, de aproximadamente 63 a 67 °C, de aproximadamente 64 a 66 °C. En algunas realizaciones, las porciones de los cebadores específicos a diana que se aparearán de forma específica con la secuencia nucleotídica diana conocida se aparearán de forma específica a una temperatura de aproximadamente

65 °C en un tampón de PCR.

En algunas realizaciones, los cebadores (por ejemplo, cebadores aleatorios, cebadores específicos de la diana) descritos en el presente documento comprenden cebadores de transcriptasa inversa. En algunas realizaciones, los cebadores de transcriptasa inversa se aparean específicamente con una molécula de ARNm a una temperatura de aproximadamente 50 a 52 °C, de aproximadamente 51 a 53 °C, de aproximadamente 52 a 54 °C, de aproximadamente 53 a 55 °C, de aproximadamente 54 a 56 °C, de aproximadamente 55 a 57 °C, de aproximadamente 56-58 °C, de aproximadamente 57 a 59 °C, de aproximadamente 58 a 60 °C. Por ejemplo, en algunas realizaciones, los cebadores de transcriptasa inversa tienen una temperatura de apareamiento de aproximadamente 53 °C, aproximadamente 53,5 °C, aproximadamente 54 °C, aproximadamente 54,5 °C, aproximadamente 56 °C. En algunas realizaciones, los cebadores de transcriptasa inversa comprenden uno o más restos de captura (por ejemplo, como se describe en el presente documento). En algunas realizaciones, el uno o más restos de captura pueden unirse a un cebador de transcriptasa inversa en el extremo 5' del ácido nucleico cebador. En algunas realizaciones, los cebadores de transcriptasa inversa comprenden un enlace de fosforotioato que une la base más en 5' con la penúltima base en 5' adyacente.

Extensión de ácido nucleico, Amplificación y PCR

En algunas realizaciones, los métodos descritos en el presente documento comprenden un régimen o etapa de extensión. En tales realizaciones, la extensión puede proceder de uno o más cebadores aleatorios hibridados, utilizando las moléculas de ácido nucleico con las que se hibridan los cebadores como moldes. Las etapas de extensión se describen en el presente documento. En algunas realizaciones, uno o más cebadores aleatorios pueden hibridar sustancialmente con todos los ácidos nucleicos en una muestra, muchos de los cuales pueden no comprender una secuencia de nucleótidos diana. En consecuencia, en algunas realizaciones, la extensión de cebadores aleatorios puede ocurrir debido a la hibridación con moldes que no comprenden una secuencia de nucleótidos diana.

En algunas realizaciones, los métodos descritos en el presente documento pueden implicar un régimen de amplificación de reacción en cadena de la polimerasa (PCR), implicando uno o más ciclos de amplificación. Las etapas de amplificación de los métodos descritos en el presente documento pueden comprender cada uno un régimen de amplificación por PCR, es decir, un conjunto de ciclos de amplificación por la reacción en cadena de la polimerasa (PCR). Como se utiliza en el presente documento, la expresión "régimen de amplificación" se refiere a un proceso de amplificación específica (aumento de la abundancia de) un ácido nucleico de interés. En algunas realizaciones, la amplificación exponencial que se da cuando los productos de una extensión por polimerasa previa sirven como moldes para las sucesivas rondas de extensión. En algunas realizaciones, un régimen de amplificación por PCR según los métodos desvelados en el presente documento puede comprender al menos uno y, en algunos casos, al menos 5 o más ciclos iterativos. En algunas realizaciones, cada ciclo iterativo comprende las etapas de: 1) separación de cadena (por ejemplo, desnaturalización térmica); 2) apareamiento del cebador oligonucleotídico con moléculas molde; y 3) extensión por polimerasa de ácido nucleico de los cebadores apareados. Debe apreciarse que se pueden utilizar cualesquiera condiciones y tiempos adecuados implicados en cada una de estas etapas. En algunas realizaciones, las condiciones y los tiempos seleccionados pueden depender de la duración, el contenido de la secuencia, los valores de temperatura de fusión, las características estructurales secundarias u otros factores relacionados con la plantilla de ácido nucleico y/o los cebadores utilizados en la reacción. En algunas realizaciones, un régimen de amplificación de acuerdo con los métodos descritos en el presente documento preferentemente se realiza en un termociclador, muchos de los cuales están disponibles en el mercado. En algunas realizaciones, los métodos descritos en el presente documento pueden comprender amplificación lineal. Por ejemplo, en algunas realizaciones, las etapas de amplificación realizadas usando cebadores anidados pueden realizarse usando amplificación lineal. En algunas realizaciones, la amplificación se puede realizar usando amplificación basada en secuencias de ácidos nucleicos (NASBA). Por ejemplo, en algunas realizaciones, la amplificación comprende una reacción NASBA mediada por T7.

En algunas realizaciones, una reacción de extensión de ácido nucleico implica el uso de una polimerasa de ácido nucleico. Como se utiliza en el presente documento, la expresión "polimerasa de ácido nucleico" se refiere a una enzima que cataliza la polimerización dependiente de molde de nucleósido trifosfatos para formar productos de extensión de cebador que son complementarios a la secuencia del ácido nucleico molde. Una enzima polimerasa de ácido nucleico inicia la síntesis en el extremo 3' de un cebador apareado y avanza en dirección hacia el extremo 5' del molde. Numerosas polimerasas de ácido nucleico son conocidas en la técnica y están disponibles en el mercado. Un grupo de polimerasas de ácido nucleico son termoestables, es decir, conservan la función después de ser sometidas a temperaturas suficientes para desnaturalizar las cadenas apareadas de los ácidos nucleicos complementarios, por ejemplo, a 94 °C, o a veces más. Un ejemplo no limitante de un protocolo de amplificación implica el uso de una polimerasa (por ejemplo, Phoenix Taq, VeraSeq) en las siguientes condiciones: 98 °C durante 30 segundos, seguido de 14-22 ciclos que comprenden fusión a 98 °C durante 10 segundos, seguido de apareamiento a 68 °C durante 30 segundos, seguido de extensión a 72 °C durante 3 minutos, seguido de mantenimiento de la reacción a 4 °C. Sin embargo, se pueden usar otras condiciones de reacción apropiadas. En algunas realizaciones, las temperaturas de apareamiento/extensión pueden ajustarse para tener en cuenta las diferencias en la concentración de sal (por ejemplo, 3 °C más para concentraciones de sal más altas). En algunas realizaciones, ralentizar la velocidad de rampa (por ejemplo, 1 °C/s, 0,5 °C/s, 0,28 °C/s, 0,1 °C/s o más lenta), por ejemplo, de 98 °C a 65 °C, mejora el rendimiento de la imprimación y la uniformidad de la cobertura en muestras altamente multiplexadas. En algunas realizaciones, los

sistemas proporcionados en el presente documento están configurados para alterar la temperatura del recipiente (por ejemplo, ciclando entre diferentes intervalos de temperatura, teniendo velocidades de aumento o disminución controladas) para facilitar la amplificación.

- 5 En algunas realizaciones, se utiliza una polimerasa de ácido nucleico en condiciones en las que la enzima realiza una extensión dependiente del molde. En algunas realizaciones, la polimerasa de ácido nucleico es ADN polimerasa I, polimerasa Taq, polimerasa Phoenix Taq, polimerasa Phusion, polimerasa T4, polimerasa T7, fragmento Klenow, Klenow exo, polimerasa phi29, transcriptasa inversa del VMA, transcriptasa inversa M-MuLV, transcriptasa inversa del VIH-1, polimerasa VeraSeq ULtra, polimerasa VeraSeq HF 2.0, EnzScript u otra polimerasa adecuada. En algunas
- 10 realizaciones, una polimerasa de ácido nucleico no es una transcriptasa inversa. En algunas realizaciones, una polimerasa de ácido nucleico actúa sobre un molde de ADN. En algunas realizaciones, la polimerasa de ácido nucleico actúa sobre un molde de ARN. En algunas realizaciones, una reacción de extensión implica la transcripción inversa realizada en un ARN para producir una molécula de ADN complementaria (actividad de polimerasa de ADN dependiente de ARN). En algunas realizaciones, una transcriptasa inversa es una polimerasa del virus de la leucemia
- 15 murina de moloney de ratón (M-MLV), transcriptasa inversa del VMA, transcriptasa inversa del VSR, transcriptasa inversa del VIH-1, transcriptasa inversa del VIH-2 u otra transcriptasa inversa apropiada.

- En algunas realizaciones, una reacción de amplificación de ácido nucleico implica ciclos que incluyen una etapa de separación de cadenas que generalmente implica el calentamiento de la mezcla de reacción. Como se utiliza en el
- 20 presente documento, la expresión "separación de cadena" o "separar las cadenas" significa tratamiento de una muestra de ácido nucleico de modo que las moléculas bicatenarias complementarias se separen en dos hebras sencillas disponibles para el apareamiento con un cebador de oligonucleótido. En algunas realizaciones, la separación de cadena de acuerdo con los métodos descritos en el presente documento. se consigue calentando la muestra de ácido nucleico por encima de su temperatura de fusión (T_m). En algunas realizaciones, para una muestra que contiene
- 25 moléculas de ácido nucleico en una preparación de reacción adecuada para una polimerasa de ácido nucleico, el calentamiento a 94 °C es suficiente para conseguir la separación de cadenas. En algunas realizaciones, una preparación de reacción adecuada contiene una o más sales (por ejemplo, KCl de 1 a 100 mM, $MgCl_2$ de 0,1 a 10 mM), al menos un agente tampón (por ejemplo, Tris-HCl de 1 a 20 mM) y un vehículo (por ejemplo, BSA del 0,01 al 0,5 %). Un ejemplo no limitativo de un tampón adecuado comprende KCl 50 mM, Tris-HCl 10 mM (pH 8,8 a 25 °C), $MgCl_2$ de
- 30 0,5 a 3 mM y BSA al 0,1 %. Otro ejemplo no limitante de un tampón adecuado comprende KCl 50 mM, Tris-HCl 10 mM (pH 8,8 a 25 °C), $MgCl_2$ de 0,5 a 5 mM (por ejemplo, aproximadamente 0,5 mM, aproximadamente 1 mM, aproximadamente 2 mM, aproximadamente 3 mM, aproximadamente 4 mM, aproximadamente 5 mM) y BSA al 0,1 %.

- En algunas realizaciones, una amplificación de ácido nucleico implica aparear cebadores con moldes de ácido nucleico que tienen cadenas características de un ácido nucleico diana. En algunas realizaciones, una cadena de un ácido nucleico diana puede servir como ácido nucleico molde. Como se utiliza en el presente documento, el término "aparear" se refiere a la formación de uno o más pares de bases complementarias entre dos ácidos nucleicos. En algunas
- 35 realizaciones, el apareamiento implica la hibridación de dos cadenas de ácido nucleico complementarias o sustancialmente complementarias. En algunas realizaciones, en el contexto de una reacción de extensión, el apareamiento implica la hibridación del cebador con una plantilla de modo que se forme un sustrato de extensión del cebador para una enzima polimerasa dependiente de la plantilla. En algunas realizaciones, las condiciones para el apareamiento (por ejemplo, entre un cebador y el molde de ácido nucleico) pueden variar en función de la longitud y la secuencia de un cebador. En algunas realizaciones, las condiciones para el apareamiento se basan en una T_f (por
- 40 ejemplo, una T_f calculada) de un cebador. En algunas realizaciones, una etapa de apareamiento de un régimen de extensión implica reducir la temperatura después de una etapa de separación de cadenas a una temperatura basada en la T_f (por ejemplo, una T_f calculada) para un cebador, durante un tiempo suficiente para permitir tal apareamiento. En algunas realizaciones, se puede determinar una T_f utilizando cualquiera de una serie de algoritmos (por ejemplo, Software de diseño de cebadores OLIGO™ (Molecular Biology Insights Inc. Colorado) y software y programas de
- 45 diseño de cebadores VENTRO NTI™ (Invitrogen, Inc. California) disponibles en Internet, que incluyen Cebador3, Oligo Calculator y NetPrimer (Premier Biosoft; Palo Alto, CA; y disponible gratuitamente en Internet (por ejemplo, en premierbiosoft.com/netprimer/netprlaunch/Help/xnetprlaunch.html)). En algunas realizaciones, la T_f de un cebador puede calcularse utilizando la siguiente fórmula, que es utilizada por el software NetPrimer y se describe con más detalle en Frieir, *et al.* PNAS 1986 83:9373-9377.

55
$$T_f = \Delta H / (\Delta S + R \cdot \ln(C/4)) + 16,6 \log ([K^+]/(1 + 0,7 [K^+])) - 273,15$$

- en donde: ΔH es entalpía para la formación de la hélice; ΔS es entropía para la formación de la hélice; R es la constante molar de un gas (1,987 cal/°C * mol); C es la concentración del ácido nucleico; y $[K^+]$ es la concentración de sal. Para la mayoría de los regímenes de amplificación, la temperatura de apareamiento se selecciona para ser de
- 60 aproximadamente 5 °C por debajo de la T_m prevista, aunque pueden usarse temperaturas más cercanas a y por encima de la T_f (por ejemplo, entre 1 °C y 5 °C por debajo de la T_f prevista o entre 1 °C y 5 °C por encima de la T_f prevista), así como, por ejemplo, temperaturas de más de 5 °C por debajo de la T_f prevista (por ejemplo, 6 °C por debajo, 8 °C por debajo, 10 °C por debajo o menos). En algunas realizaciones, cuanto más cercana es la temperatura de apareamiento a la T_f , más específico es el apareamiento. En algunas realizaciones, el tiempo utilizado para el
- 65 apareamiento del cebador durante una reacción de extensión (por ejemplo, en el contexto de un régimen de amplificación por PCR) se determina en función de, al menos en parte, el volumen de la reacción (por ejemplo, por lo

que mayores volúmenes requieren mayores tiempos). En algunas realizaciones, el tiempo utilizado para el apareamiento del cebador durante una reacción de extensión (por ejemplo, dentro del contexto de un régimen de amplificación por PCR) se determina en función de, al menos en parte, las concentraciones de cebador y molde (por ejemplo, por lo que concentraciones relativamente superiores de cebador a molde implican menos tiempo que las concentraciones relativamente inferiores. En algunas realizaciones, dependiendo del volumen y la concentración relativa de cebador/molde, las etapas de apareamiento del cebador en una reacción de extensión (por ejemplo, dentro del contexto de un régimen de amplificación) pueden estar en el intervalo de 1 segundo a 5 minutos, de 10 segundos a 2 minutos o de 30 segundos a 2 minutos. Como se utiliza en el presente documento, "aparear sustancialmente" se refiere a la medida en que se forman pares de bases complementarias entre dos ácidos nucleicos que, cuando se utiliza en el contexto de un régimen de amplificación por PCR, es suficiente para producir un nivel detectable de un producto específicamente amplificado.

Como se utiliza en el presente documento, la expresión "extensión por polimerasa" se refiere a la adición dependiente del molde de al menos un nucleótido complementario, mediante una polimerasa de ácido nucleico, al extremo 3' de un cebador que se hibrida con un molde de ácido nucleico. En algunas realizaciones, la extensión por polimerasa añade más de un nucleótido, por ejemplo, hasta e incluidos nucleótidos que corresponden a la longitud completa del molde. En algunas realizaciones, las condiciones para la extensión de la polimerasa se basan, al menos en parte, en la identidad de la polimerasa utilizada. En algunas realizaciones, la temperatura utilizada para la extensión por polimerasa se basa en las propiedades de actividad conocidas de la enzima. En algunas realizaciones, en las que las temperaturas de apareamiento están por debajo de las temperaturas óptimas para la enzima, puede ser aceptable utilizar una temperatura de extensión inferior. En algunas realizaciones, las enzimas pueden retener al menos una actividad parcial por debajo de sus temperaturas de extensión óptimas. En algunas realizaciones, una extensión de polimerasa (por ejemplo, realizada con polimerasas termoestables tales como polimerasa Taq y variantes de la misma) se realiza a una temperatura de 65 °C a 75 °C o de 68 °C a 72 °C. En algunas realizaciones, los métodos proporcionados en el presente documento implican la extensión con polimerasa de cebadores que se hibridan con moldes de ácido nucleico en cada ciclo de un régimen de amplificación por PCR. En algunas realizaciones, una extensión de polimerasa se realiza utilizando una polimerasa que tiene una actividad de desplazamiento de cadena relativamente fuerte. En algunas realizaciones, las polimerasas que tienen un fuerte desplazamiento de cadena son útiles para preparar ácidos nucleicos con el fin de detectar fusiones (por ejemplo, fusiones en 5'). En algunas realizaciones, las polimerasas que tienen actividad exonucleasa 5'→3' (por ejemplo, polimerasa Taq) son útiles para producir fragmentos de biblioteca largos.

En algunas realizaciones, la extensión del cebador se realiza en condiciones que permiten la extensión de los cebadores oligonucleotídicos apareados. Como se utiliza en el presente documento, la expresión "condiciones que permiten la extensión de un oligonucleótido apareado de modo que se generen productos de extensión" se refiere al conjunto de condiciones (por ejemplo, temperatura, las concentraciones de sal y el cofactor, pH y la concentración de enzima) en las cuales una polimerasa de ácido nucleico cataliza la extensión del cebador. En algunas realizaciones, tales condiciones se basan, al menos en parte, en la polimerasa de ácido nucleico que se utiliza. En algunas realizaciones, una polimerasa puede realizar una reacción de extensión del cebador en una preparación de reacción adecuada.

En algunas realizaciones, una preparación de reacción adecuada contiene una o más sales (por ejemplo, KCl de 1 a 100 mM, MgCl₂ de 0,1 a 10 mM), al menos un agente tampón (por ejemplo, Tris-HCl de 1 a 20 mM), un vehículo (por ejemplo, del 0,01 al 0,5 % de BSA) y uno o más NTP (por ejemplo, de 10 a 200 μM de cada uno de dATP, dTTP, dCTP y dGTP). Un conjunto no limitante de condiciones es KCl 50 mM, Tris-HCl 10 mM (pH 8,8 a 25 °C), MgCl₂ 0,5 a 3 mM, 200 μM de cada dNTP y BSA al 0,1 % a 72 °C, en las cuales una polimerasa (por ejemplo, polimerasa Taq) cataliza la extensión del cebador.

En algunas realizaciones, una preparación de reacción adecuada contiene una o más sales (por ejemplo, KCl de 1 a 100 mM, MgCl₂ de 0,5 a 5 mM), al menos un agente tampón (por ejemplo, Tris-HCl de 1 a 20 mM), un vehículo (por ejemplo, del 0,01 al 0,5 % de BSA) y uno o más NTP (por ejemplo, de 50 a 350 μM de cada uno de dATP, dTTP, dCTP y dGTP). Un conjunto no limitante de condiciones es KCl 50 mM, Tris-HCl 10 mM (pH 8,8 a 25 °C), MgCl₂ 3 mM, 200 μM de cada dNTP y BSA al 0,1 % a 72 °C, en las cuales una polimerasa (por ejemplo, polimerasa Taq) cataliza la extensión del cebador. Otro conjunto no limitante de condiciones es KCl 50 mM, Tris-HCl 10 mM (pH 8,8 a 25 °C), MgCl₂ 3 mM, dATP 266 μM, dCTP 200 μM, dGTP 133 μM, dTTP 200 μM y BSA al 0,1 % a 72 °C, en las cuales una polimerasa (por ejemplo, polimerasa Taq) cataliza la extensión del cebador.

En algunas realizaciones, las condiciones para la iniciación y la extensión pueden incluir la presencia de uno, dos, tres o cuatro trifosfatos de desoxirribonucleósido diferentes (por ejemplo, seleccionado de entre dATP, dTTP, dCTP y dGTP) y un agente inductor de la polimerización como la ADN polimerasa o la transcriptasa inversa, en un tampón adecuado. En algunas realizaciones, un "tampón" puede incluir disolventes (por ejemplo, disolventes acuosos) más cofactores y reactivos apropiados que afectan al pH, la fuerza iónica, etc. En algunas realizaciones, los dos, tres o cuatro trifosfatos de desoxirribonucleósido diferentes están presentes a concentraciones equimolares, o aproximadamente equimolares. En algunas realizaciones, los dos, tres o cuatro trifosfatos de desoxirribonucleósido diferentes están presentes en diferentes concentraciones, que se ha determinado experimentalmente que son adecuadas para una implementación particular de la tecnología.

En algunas realizaciones, la amplificación de ácidos nucleicos implica hasta 5, hasta 10, hasta 20, hasta 30, hasta 40 o más rondas (ciclos) de amplificación. En algunas realizaciones, la amplificación de ácido nucleico puede comprender un conjunto de ciclos de un régimen de amplificación por PCR de 5 ciclos a 20 ciclos de largo. En algunas realizaciones, una etapa de amplificación puede comprender un conjunto de ciclos de un régimen de amplificación por PCR de 10 ciclos a 20 ciclos de largo. En algunas realizaciones, cada etapa de amplificación puede comprender un conjunto de ciclos de un régimen de amplificación por PCR de 12 ciclos a 16 ciclos de duración. En algunas realizaciones, una temperatura de apareamiento puede ser inferior a 70 °C. En algunas realizaciones, una temperatura de apareamiento puede ser inferior a 72 °C. En algunas realizaciones, una temperatura de apareamiento puede ser de aproximadamente 65 °C. En algunas realizaciones, una temperatura de apareamiento puede ser de aproximadamente 61 a aproximadamente 72 °C.

En diversas realizaciones, los métodos y composiciones descritos en el presente documento se refieren a la realización de un régimen de amplificación por PCR con uno o más de los tipos de cebadores descritos en el presente documento. Como se utiliza en el presente documento, "cebador" se refiere a un oligonucleótido capaz de hibridarse específicamente con un molde de ácido nucleico y proporcionar un extremo 3' que sirve como sustrato para una polimerasa dependiente del molde para producir un producto de extensión que es complementario al molde. En algunas realizaciones, un cebador es monocatenario, de modo que el cebador y su complementario puedan aparearse para formar dos cadenas. Los cebadores de acuerdo con los métodos y composiciones descritos en el presente documento pueden comprender una secuencia de hibridación (por ejemplo, una secuencia que se hibrida con una plantilla de ácido nucleico) que tiene una longitud inferior o igual a 300 nucleótidos, por ejemplo, menos de o igual a 300 o 250, o 200, o 150, o 100, o 90, o 80, o 70, o 60, o 50, o 40, o 30 o inferior, o 20 o inferior, o 15 o inferior, pero al menos 6 nucleótidos de longitud. En algunas realizaciones, una secuencia de apareamiento de un cebador puede tener de 6 a 50 nucleótidos de longitud, de 6 a 35 nucleótidos de longitud, de 6 a 20 nucleótidos de longitud, de 10 a 25 nucleótidos de longitud.

Puede usarse cualquier método adecuado para sintetizar oligonucleótidos y cebadores. En algunas realizaciones, las fuentes comerciales ofrecen servicios de síntesis de oligonucleótidos adecuados para proporcionar cebadores para usar en métodos y composiciones descritos en el presente documento (por ejemplo, INVITROGEN™ Custom DNA Oligos (Life Technologies, Grand Island, NY) u oligos de ADN personalizados de Integrated DNA Technologies (Coralville, IA)).

Ácido nucleico diana

Como se utiliza en el presente documento, las expresiones "ácido nucleico diana", "molécula de ácido nucleico que comprende una secuencia de nucleótidos diana", y "ácido nucleico que comprende una secuencia de nucleótidos diana" se refiere a una molécula de ácido nucleico de interés (por ejemplo, un ácido nucleico que se va a preparar para el análisis). En algunas realizaciones, un ácido nucleico diana comprende tanto una secuencia de nucleótidos diana (por ejemplo, una secuencia de nucleótidos conocida o predeterminada) como una secuencia de nucleótidos adyacente que se va a determinar (que puede denominarse secuencia desconocida). Un ácido nucleico diana puede tener cualquier longitud apropiada. En algunas realizaciones, un ácido nucleico diana es bicatenario. En algunas realizaciones, un ácido nucleico diana es ADN. En algunas realizaciones, un ácido nucleico diana comprende ADN genómico o cromosómico (ADNg). En algunas realizaciones, un ácido nucleico diana comprende ADN complementario (ADNc). En algunas realizaciones, un ácido nucleico diana es monocatenario. En algunas realizaciones, un ácido nucleico diana comprende ARN (por ejemplo, ARNm, ARNr, ARNt, ADNsc, ARNcf, ARN largo no codificante, microARN).

Muchos de los métodos de secuenciación adecuados para su uso en los métodos descritos en el presente documento proporcionan carreras de secuenciación con longitudes óptimas de lectura de decenas a cientos de bases nucleotídicas (por ejemplo, la tecnología Ion Torrent puede producir longitudes de lectura de 200-400 pb). Los ácidos nucleicos diana comprendidos, por ejemplo, por ADN genómico o ARNm, pueden comprender moléculas de ácido nucleico que son sustancialmente mayores que esta longitud óptima de lectura. Para que la porción de ácido nucleico amplificado resultante de la segunda etapa de amplificación tenga una longitud adecuada (por ejemplo, hasta 100 pb, 200 pb, 300 pb, 400 pb, 500 pb, 1 kb, 2 kb) para su uso en una tecnología de secuenciación particular, la distancia promedio entre la secuencia nucleotídica diana conocida y un extremo del ácido nucleico diana al que se puede ligar el adaptador debe ser tan próxima a la longitud óptima de lectura de la tecnología seleccionada como sea posible. Por ejemplo, si la longitud óptima de lectura de una tecnología de secuenciación dada es de 200 pb, entonces las moléculas de ácido nucleico amplificadas de acuerdo con los métodos descritos en el presente documento deben tener una longitud media de aproximadamente 400 pb o menos. Sin embargo, debería apreciarse que, en algunas realizaciones, las técnicas descritas en el presente documento pueden implementarse cuando las moléculas de ácido nucleico superan los 400 pb de longitud. Por ejemplo, en algunas realizaciones, los fragmentos de ácido nucleico pueden tener aproximadamente 400 o más nucleótidos, 500 o más nucleótidos, 600 o más nucleótidos, 700 o más nucleótidos, 800 o más nucleótidos, 900 o más nucleótidos, 1000 o más nucleótidos, 1500 o más nucleótidos, 2000 o más nucleótidos, 2500 o más nucleótidos, 3000 o más nucleótidos, 4000 o más nucleótidos, 5000 o más nucleótidos, 10000 o más nucleótidos.

Los ácidos nucleicos diana comprendidos por, por ejemplo, ADN genómico o ARNm, se pueden romper, *por ejemplo*, cortado mecánica o enzimáticamente, para generar fragmentos de cualquier tamaño deseado. Ejemplos no limitantes de los procesos de rotura mecánica incluyen sonicación, nebulización y la tecnología de ruptura AFA™ disponible en Covaris (Woburn, MA). En algunas realizaciones, un ácido nucleico diana comprendido por ADN genómico se puede romper de forma mecánica mediante sonicación.

En algunas realizaciones, cuando el ácido nucleico diana está comprendido por ARN, la muestra se puede someter a un régimen de transcriptasa inversa para generar un molde de ADN. En algunas realizaciones, el molde de ADN puede luego ser cortado. En algunas realizaciones, el molde de ADN no se corta. Por ejemplo, en algunas realizaciones, la concentración de cebadores utilizados durante un régimen de transcriptasa inversa se puede ajustar de manera que el ADNc producto tenga una longitud "fragmentada" apropiada. En algunas realizaciones, el ARN diana se puede romper antes de la realización del régimen de transcriptasa inversa. En algunas realizaciones, una muestra que comprende el ARN diana puede utilizarse en los métodos descritos en el presente documento utilizando ácidos nucleicos totales extraídos de especímenes frescos o degradados; sin la necesidad de eliminación de ADN genómico para la secuenciación de ADNc; sin la necesidad de agotamiento de ARN ribosómico para la secuenciación de ADNc; sin la necesidad de rotura mecánica o enzimática en cualquiera de las etapas; sometiendo el ARN a síntesis de ADNc bicatenario utilizando hexámeros aleatorios; y sometiendo al ácido nucleico a reparación del extremo, fosforilación y adenilación.

En algunas realizaciones, una secuencia nucleotídica diana puede estar comprendida por un reordenamiento génico. Los métodos descritos en el presente documento son adecuados para determinar la presencia y/o la identidad de un reordenamiento de genes ya que previamente se debe conocer la identidad de solo una mitad del reordenamiento génico (es decir, la mitad del reordenamiento génico que es identificada por los cebadores específicos del gen). En algunas realizaciones, el reordenamiento génico puede comprender un oncogén. En algunas realizaciones, el reordenamiento génico puede comprender un oncogén de fusión. En algunas realizaciones, el reordenamiento génico puede comprender un producto de recombinación V(D)J.

Como se utiliza en el presente documento, la expresión "secuencia nucleotídica diana conocida" se refiere a una porción de un ácido nucleico diana para la cual se conoce la secuencia (por ejemplo, la identidad y el orden de las bases nucleotídicas del ácido nucleico). Por ejemplo, en algunas realizaciones, una secuencia de nucleótidos diana conocida es una secuencia de nucleótidos de un ácido nucleico que se conoce o que se ha determinado antes de una interrogación de una secuencia adyacente desconocida del ácido nucleico. Una secuencia de nucleótidos diana conocida puede tener cualquier longitud apropiada.

En algunas realizaciones, una secuencia de nucleótidos diana (por ejemplo, una secuencia de nucleótidos diana conocida) tiene una longitud de 10 o más nucleótidos, 30 o más nucleótidos, 40 o más nucleótidos, 50 o más nucleótidos, 100 o más nucleótidos, 200 o más nucleótidos, 300 o más nucleótidos, 400 o más nucleótidos, 500 o más nucleótidos, 600 o más nucleótidos, 700 o más nucleótidos, 800 o más nucleótidos, 900 o más nucleótidos, 1000 o más nucleótidos, 1500 o más nucleótidos, 2000 o más nucleótidos, 2500 o más nucleótidos, 3000 o más nucleótidos, 4000 o más nucleótidos, 5000 o más nucleótidos, 10000 o más nucleótidos. En algunas realizaciones, una secuencia de nucleótidos diana (por ejemplo, una secuencia de nucleótidos diana conocida) tiene una longitud en el intervalo de 10 a 100 nucleótidos, de 10 a 500 nucleótidos, de 10 a 1000 nucleótidos, de 100 a 500 nucleótidos, de 100 a 1000 nucleótidos, de 500 a 1000 nucleótidos, de 500 a 5000 nucleótidos.

En algunas realizaciones, en el presente documento se proporcionan métodos para determinar secuencias de porciones contiguas (o adyacentes) de un ácido nucleico. Como se utiliza en el presente documento, la expresión "secuencia de nucleótidos contigua a" se refiere a una secuencia de nucleótidos de una molécula de ácido nucleico (por ejemplo, un ácido nucleico diana) que está inmediatamente aguas arriba o aguas abajo de otra secuencia de nucleótidos (por ejemplo, una secuencia de nucleótidos conocida). En algunas realizaciones, una secuencia de nucleótidos contigua a una secuencia de nucleótidos diana conocida puede tener cualquier longitud apropiada. En algunas realizaciones, una secuencia nucleotídica contigua a una secuencia nucleotídica diana conocida comprende 1 kb o menos de la secuencia nucleotídica, por ejemplo, 1 kb o menos de la secuencia nucleotídica, 750 pb o menos de la secuencia nucleotídica, 500 pb o menos de la secuencia nucleotídica, 400 pb o menos de la secuencia nucleotídica, 300 pb o menos de la secuencia nucleotídica, 200 pb o menos de la secuencia nucleotídica, 100 pb o menos de la secuencia nucleotídica. En algunas realizaciones, en la que una muestra comprende diferentes ácidos nucleicos diana que comprenden una secuencia nucleotídica diana conocida (por ejemplo, una célula en la que una secuencia nucleotídica diana conocida se da múltiples veces en su genoma, o por separado, cromosomas no idénticos), puede haber múltiples secuencias que comprendan una "secuencia nucleotídica contigua a" la secuencia nucleotídica diana conocida. Como se utiliza en el presente documento, la expresión "determinar una (o la) secuencia nucleotídica", se refiere a determinar la identidad y las posiciones relativas de las bases nucleotídicas de un ácido nucleico.

En algunas realizaciones, un ácido nucleico diana conocido puede contener una secuencia de fusión resultante de un reordenamiento génico. En algunas realizaciones, los métodos descritos en el presente documento son adecuados para determinar la presencia y/o identidad de un reordenamiento génico. En algunas realizaciones, la identidad de una parte de un reordenamiento génico se conoce previamente (por ejemplo, la parte de un reordenamiento génico que

será el objetivo de los cebadores específicos de genes) y la secuencia de la otra parte puede determinarse utilizando los métodos descritos en el presente documento. En algunas realizaciones, un reordenamiento génico puede implicar un oncogén. En algunas realizaciones, un reordenamiento génico puede comprender un oncogén de fusión.

5 *Códigos de barras moleculares y secuencias índice*

En algunas realizaciones, los cebadores y/o adaptadores pueden contener secuencias adicionales, tal como una secuencia identificadora (por ejemplo, un código de barras, un índice), cebador de secuenciación de secuencias de hibridación (por ejemplo, Rd1), y secuencias adaptadoras. En algunas realizaciones, las secuencias adaptadoras son secuencias utilizadas con un sistema de secuenciación de próxima generación. En algunas realizaciones, las secuencias adaptadoras son secuencias P5 y P7 para la tecnología de secuenciación basada en Illumina. En algunas realizaciones, la secuencia adaptadora son P1 y A compatibles con la tecnología de secuenciación Ion Torrent.

En algunas realizaciones, como se utiliza en el presente documento, "código de barras", "código de barras molecular", y "etiqueta de código de barras molecular" pueden usarse indistintamente, y generalmente se refieren a una región de un ácido nucleico adaptador que es útil como identificador del ácido nucleico específico al que está ligado. En algunas realizaciones, un código de barras molecular comprende una secuencia de ácido nucleico aleatoria que proporciona un identificador único para el ácido nucleico al que está ligado. En algunas realizaciones, se puede usar un código de barras molecular para identificar fragmentos únicos y "desduplicar" las lecturas de secuenciación de una muestra. En algunas realizaciones, se puede utilizar un código de barras molecular para identificar y eliminar duplicados de PCR. En algunas realizaciones, un código de barras molecular puede tener de 2 a 25 nucleótidos de longitud, de 2 a 15 nucleótidos de longitud, de 2 a 10 nucleótidos de longitud, de 2 a 6 nucleótidos de longitud. En algunas realizaciones, un código de barras molecular comprende al menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 o al menos 25 nucleótidos. En algunas realizaciones, un código de barras molecular comprende 8 nucleótidos.

En algunas realizaciones, como se utiliza en el presente documento, "índice", "secuencia índice", "región índice", e "índice de muestra" pueden usarse indistintamente, y generalmente se refieren a una región de un ácido nucleico adaptador que es útil como identificador de la población a la que pertenece el ácido nucleico ligado. En algunas realizaciones, un índice comprende una secuencia de ácido nucleico fija que puede usarse para identificar una colección de secuencias que pertenecen a una biblioteca común. Por ejemplo, se puede usar un índice para identificar una muestra que corresponde a un ácido nucleico. En algunas realizaciones, se puede usar un índice, por ejemplo, como identificador de fuente, identificador de ubicación, identificador de fecha u hora (por ejemplo, fecha u hora de muestreo o procesamiento), u otro identificador de un ácido nucleico relacionado con una propiedad compartida o común (*por ejemplo*, común entre otros ácidos nucleicos de una biblioteca). En algunas realizaciones, dichas secuencias índice son útiles para identificar diferentes aspectos de un ácido nucleico que están presentes en una población de ácidos nucleicos. En algunas realizaciones, las secuencias de índice pueden proporcionar un identificador de fuente o ubicación para un ácido nucleico diana. Por ejemplo, una secuencia índice puede servir para identificar a un paciente del que se obtiene un ácido nucleico. En algunas realizaciones, las secuencias índice permiten la secuenciación de múltiples muestras diferentes en una sola reacción (por ejemplo, realizado en una sola celda de flujo). En algunas realizaciones, se puede usar una secuencia índice para orientar un generador de imágenes de secuencia con el fin de detectar reacciones de secuenciación individuales. En algunas realizaciones, una secuencia índice puede tener de 2 a 25 nucleótidos de longitud, de 2 a 15 nucleótidos de longitud, de 2 a 10 nucleótidos de longitud, de 2 a 6 nucleótidos de longitud. En algunas realizaciones, un índice comprende al menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, o al menos 25 nucleótidos.

En algunas realizaciones, cuando se usa una población de cebadores aleatorios con cola de acuerdo con los métodos descritos en el presente documento, pueden estar presentes múltiples productos de amplificación distinguibles después de la amplificación. En algunas realizaciones, porque los cebadores aleatorios con cola se hibridan en varias posiciones a lo largo de las moléculas de ácido nucleico de una muestra, un conjunto de cebadores específicos del objetivo pueden hibridar (y amplificar) los productos de extensión creados por más de 1 evento de hibridación, por ejemplo, un cebador aleatorio con cola puede hibridar a una primera distancia (por ejemplo, 100 nucleótidos) desde un sitio de hibridación de cebador específico de la diana, y otro cebador aleatorio con cola puede hibridar a una segunda distancia (por ejemplo, 200 nucleótidos) desde un sitio de hibridación de cebador específico de la diana, dando como resultado dos productos de amplificación (por ejemplo, un primer producto de amplificación que comprende aproximadamente 100 pb y un segundo producto de amplificación que comprende aproximadamente 200 pb). En algunas realizaciones, cada uno de estos múltiples productos de amplificación puede secuenciarse utilizando tecnología de secuenciación de próxima generación. En algunas realizaciones, la secuenciación de estos múltiples productos de amplificación es ventajosa porque proporciona múltiples lecturas de secuencia superpuestas que pueden compararse entre sí para detectar errores de secuencia introducidos durante los procesos de amplificación o secuenciación. En algunas realizaciones, los productos de amplificación individuales (por ejemplo, derivados de una sola molécula) pueden alinearse y difieren en la secuencia presente en una base particular, puede haber un artefacto o error de PCR y/o secuenciación.

65 *Corte/fragmentación de ADN*

Las moléculas de ácido nucleico descritas en el presente documento se pueden cortar (por ejemplo, cortado mecánica o enzimáticamente, cortado mediante nebulizador) para generar fragmentos de cualquier tamaño deseado. Ejemplos no limitantes de los procesos de rotura mecánica incluyen sonicación, nebulización y la tecnología de ruptura AFA™ disponible en Covaris (Woburn, MA). En algunas realizaciones, un ácido nucleico se puede cortar mecánicamente mediante ultrasonidos. En algunas realizaciones, un ácido nucleico diana no se corta ni se digiere. En algunas realizaciones, los productos de ácido nucleico de etapas preparativas (por ejemplo, productos de extensión, productos de amplificación) no se cortan ni se digieren enzimáticamente.

En algunas realizaciones, cuando una secuencia de nucleótidos diana comprende ARN, la muestra puede someterse a un régimen de transcriptasa inversa para generar molde de ADN y, a continuación, el molde de ADN puede romperse. En algunas realizaciones, el ARN diana se puede romper antes de la realización de un régimen de transcriptasa inversa. En algunas realizaciones, una muestra que comprende el ARN diana puede utilizarse en los métodos descritos en el presente documento utilizando ácidos nucleicos totales extraídos de especímenes frescos o degradados; sin la necesidad de eliminación de ADN genómico para la secuenciación de ADNc; sin la necesidad de agotamiento de ARN ribosómico para la secuenciación de ADNc; sin la necesidad de rotura mecánica o enzimática en cualquiera de las etapas; sometiendo el ARN a síntesis de ADNc bicatenario utilizando hexámeros aleatorios.

Secuenciación

En algunos aspectos, la tecnología descrita en el presente documento se refiere a métodos para enriquecer muestras de ácido nucleico para la secuenciación de oligonucleótidos. En algunas realizaciones, se puede secuenciar mediante un método de secuenciación de última generación. Como se utiliza en el presente documento, "secuenciación de última generación" se refiere a las tecnologías de secuenciación de oligonucleótidos que tienen la capacidad de secuenciar oligonucleótidos a velocidades por encima de aquellas posibles con los métodos de secuenciación convencionales (por ejemplo, secuenciación Sanger), debido a la realización y lectura de miles a millones de reacciones de secuenciación en paralelo. Ejemplos no limitantes de métodos/plataformas de secuenciación de última generación incluyen secuenciación de firma masivamente paralela (Lynx Therapeutics); pirosecuenciación 454 (454 Life Sciences/Roche Diagnostics); fase sólida, secuenciación de terminador de colorante reversible (Solexa/Illumina); tecnología SOLiD (Applied Biosystems); secuenciación por semiconductores iónicos (Ion Torrent); secuenciación de nanoesferas de ADN (Complete Genomics); y tecnologías disponibles en Pacific biosciences, Intelligen Biosystems y Oxford Nanopore Technologies. En algunas realizaciones, los cebadores de secuenciación pueden comprender porciones compatibles con el método de secuenciación de última generación seleccionado. Las tecnologías de secuenciación de última generación y las restricciones y los parámetros de diseño de cebadores de secuenciación asociados son bien conocidos en la técnica (véase, por ejemplo, Shendure, *et al.*, "Next-generation DNA sequencing", *Nature*, 2008, vol. 26, n.º 10, 1135-1145; Mardis, "The impact of next-generation sequencing technology on genetics", *Trends in Genetics*, 2007, vol. 24, n.º 3, pág. 133-141; Su, *et al.*, "Next-generation sequencing and its applications in molecular diagnostics" *Expert. Rev. Mol. Diagn.*, 2011, 11(3):333-43; Zhang *et al.*, "The impact of next-generation sequencing on genomics", *J. Genet. Genomics*, 2011, 38(3):95-109; (Nyren, P. *et al.* *Anal Biochem* 208: 17175 (1993); Bentley, D. R. *Curr. Opin. Genet. Dev.* 16:545-52 (2006); Strausberg, R. L., *et al.* *Drug Disc Today* 13:569-77 (2008); patente de Estados Unidos n.º 7.282.337; patente de Estados Unidos n.º 7.279.563; patente de Estados Unidos n.º 7.226.720; patente de Estados Unidos n.º 7.220.549; patente de Estados Unidos n.º 7.169.560; patente de Estados Unidos n.º 6.818.395; patente de Estados Unidos n.º 6.911.345; las publicaciones de Estados Unidos n.º 2006/0252077; 2007/0070349 y 20070070349).

En algunas realizaciones, la etapa de secuenciación depende de la utilización de un primer y un segundo cebador de secuenciación. En algunas realizaciones, el primer y segundo cebador de secuenciación se selecciona por ser compatible con un método de secuenciación de última generación como se describe en el presente documento.

Los métodos para alinear las lecturas de secuenciación con las bases de datos de secuencia conocidas de secuencias de ADNc y/o genómico son bien conocidos en la técnica y el programa informático para este proceso está disponible en el mercado. En algunas realizaciones, las lecturas (menos la secuenciación de la secuencia nucleotídica del adaptador y/o cebador) que no se mapean, en su totalidad, en las bases de datos de secuencia tipo silvestre pueden ser reordenamientos genómicos o mutaciones indel grandes. En algunas realizaciones, las lecturas (menos la secuenciación de la secuencia nucleotídica de adaptador y/o cebador) que comprenden las secuencias que se mapean en para localizaciones múltiples en el genoma pueden ser reordenamientos genómicos. En algunas realizaciones, un ensamblaje *de novo* de lecturas superpuestas en secuencias contiguas, o "contigs", puede construirse y utilizarse en la alineación de lecturas de secuenciación. En algunas realizaciones, se puede utilizar una referencia de punto caliente que no dependa de una base de datos genómica de acceso público.

Muestras

En algunas realizaciones, un ácido nucleico (por ejemplo, ácido nucleico diana, ácido nucleico que comprende una secuencia de nucleótidos diana) está presente o se obtiene de una muestra apropiada (por ejemplo, una muestra de comida, muestra ambiental, muestra biológica, por ejemplo, muestra de sangre, etc.). En algunas realizaciones, el ácido nucleico diana es una muestra biológica obtenida de un sujeto. En algunas realizaciones una muestra puede ser una muestra diagnóstica obtenida de un sujeto. En algunas realizaciones, una muestra además puede comprender

proteínas, células, líquidos, fluidos biológicos, conservantes y/u otras sustancias. A modo de ejemplo no limitante, una muestra puede ser un frotis de mejilla, sangre, suero, plasma, esputo, líquido cefalorraquídeo, orina, lágrimas, aislados alveolares, líquido pleural, líquido pericárdico, líquido quístico, tejido tumoral, tejido, una biopsia, saliva, una aspiración o combinaciones de los mismos. En algunas realizaciones, una muestra puede obtenerse por extirpación o biopsia.

5 En algunas realizaciones, la muestra puede obtenerse de un sujeto que necesita tratamiento para una enfermedad asociada a una alteración genética, por ejemplo, cáncer o una enfermedad hereditaria. En algunas realizaciones, una secuencia diana conocida está presente en un gen asociado a una enfermedad.

10 En algunas realizaciones, una muestra de tumor obtenida de un sujeto que necesita tratamiento para el cáncer. En algunas realizaciones, la muestra comprende una población de células tumorales, por ejemplo, al menos una célula tumoral. En algunas realizaciones, la muestra comprende una biopsia de tumor, incluyendo pero sin limitación, tejido de biopsia no tratado o tejido de biopsia tratado (por ejemplo, tejido de biopsia fijado en formalina y/o incluido en parafina).

15 En algunas realizaciones, la muestra está obtenida de forma reciente. En algunas realizaciones, la muestra se almacene antes de utilizarse en los métodos y composiciones descritos en el presente documento. En algunas realizaciones, la muestra es una muestra no tratada. Como se utiliza en el presente documento, "muestra no tratada" se refiere a una muestra biológica que no ha tenido ningún pretratamiento de muestra previo excepto la dilución y/o suspensión en una solución. En algunas realizaciones, una muestra se obtiene un sujeto y conservarse o procesarse antes de ser utilizada en los métodos y composiciones descritos en el presente documento. A modo de ejemplo no limitante, una muestra puede ser incrustada en cera de parafina, refrigerada o congelada. Una muestra congelada puede descongelarse antes de determinar la presencia de un ácido nucleico de acuerdo con los métodos y composiciones descritos en el presente documento. En algunas realizaciones, la muestra puede ser una muestra procesada o tratada. Métodos ilustrativos para tratar o procesar una muestra incluyen, aunque sin limitación, centrifugación, filtración, sonicación, homogeneización, calentamiento, congelación y descongelación, contacto con un conservante (por ejemplo, anticoagulante o inhibidor de nucleasa) y cualquier combinación de los mismos. En algunas realizaciones, una muestra puede tratarse con un reactivo químico y/o biológico. Los reactivos químicos y/o biológicos pueden emplearse para proteger y/o conservar la estabilidad de la muestra o el ácido nucleico comprendida por la muestra durante el procesamiento y/o almacenamiento. Además, o como alternativa, los reactivos químicos y/o biológicos pueden emplearse para liberar los ácidos nucleicos de otros componentes de la muestra. A modo de ejemplo no limitante, una muestra sanguínea puede tratarse con un anticoagulante antes de utilizarse en los métodos y composiciones descritos en el presente documento.

35 En el método desvelado en el presente documento se pueden usar métodos y procesos adecuados para el procesamiento, la conservación o el tratamiento de muestras para el análisis de ácidos nucleicos. En algunas realizaciones, una muestra puede ser una muestra fluida aclarada. En algunas realizaciones, una muestra se puede clarificar por centrifugación a baja velocidad (por ejemplo, 3000 x g o menos) y la recolección del sobrenadante que comprende la muestra de fluido clarificado.

40 En algunas realizaciones, un ácido nucleico presente en una muestra puede aislarse, enriquecerse o purificarse antes de utilizarse en los métodos y composiciones descritos en el presente documento. Se pueden usar métodos adecuados para aislar, enriquecer o purificar ácidos nucleicos de una muestra. Por ejemplo, los kits para el aislamiento de ADN genómico de varios tipos de muestras están disponibles comercialmente (por ejemplo, números de catálogo 51104, 51304, 56504 y 56404; Qiagen; Germantown, MD).

50 En algunas realizaciones, los métodos descritos en el presente documento se refieren a métodos de enriquecimiento de ácidos nucleicos diana, por ejemplo, antes de una secuenciación de los ácidos nucleicos diana. En algunas realizaciones, no se conoce una secuencia de un extremo del ácido nucleico diana a enriquecer antes de la secuenciación. En algunas realizaciones, los métodos descritos en el presente documento se refieren a métodos de enriquecimiento de secuencias nucleotídicas específicas antes de determinar la secuencia nucleotídica utilizando una tecnología de secuenciación de última generación. En algunas realizaciones, los métodos de enriquecimiento de secuencias nucleotídicas específicas no comprenden enriquecimiento por hibridación.

55 *Genes diana y aplicaciones terapéuticas*

Los aspectos de la divulgación pueden ser útiles en el análisis genético de un sistema inmunitario. Sin embargo, debe apreciarse que las técnicas descritas en el presente documento pueden aplicarse a cualquier gen diana o ácido nucleico de interés. En algunas realizaciones de técnicas descritas en el presente documento, una determinación de la secuencia contigua a una secuencia diana de oligonucleótido conocida puede proporcionar información relevante para el tratamiento de la enfermedad. Por lo tanto, en algunas realizaciones, los métodos desvelados en el presente documento pueden usarse para ayudar en el tratamiento de enfermedades. En algunas realizaciones, una muestra puede ser de un sujeto que necesita tratamiento para una enfermedad asociada a una alteración genética. En algunas realizaciones, una secuencia diana conocida es una secuencia de un gen asociado a una enfermedad, por ejemplo, un oncogén. En algunas realizaciones, una secuencia contigua a una secuencia diana oligonucleotídica conocida y/o la secuencia diana oligonucleotídica conocida pueden comprender una mutación o anomalía genética que está

asociada a la enfermedad, por ejemplo, un SNP, una inserción, una delección y/o un reordenamiento génico. En algunas realizaciones, una secuencia contigua a una secuencia diana conocida y/o una secuencia diana conocida presente en una secuencia compuesta de muestra de un producto de reordenamiento génico. En algunas realizaciones, un reordenamiento génico puede ser un oncogén, por ejemplo, un oncogén de fusión.

Determinados tratamientos contra el cáncer son en particular efectivos contra los tumores que comprenden determinados oncogenes, por ejemplo, un agente de tratamiento que se dirige a la acción o la expresión de un oncogén de fusión dado puede ser efectivo contra tumores que comprenden ese oncogén de fusión pero no contra los tumores que carecen del oncogén de fusión. Los métodos descritos en el presente documento pueden facilitar la determinación de secuencias específicas que revelan el estado del oncogén (por ejemplo, mutaciones, SNP y/o reordenamientos). En algunas realizaciones, los métodos descritos en el presente documento además pueden permitir la determinación de secuencias específicas cuando se conoce solo la secuencia de una región flanqueante, por ejemplo, los métodos descritos en el presente documento pueden determinar la presencia e identidad de reordenamientos génicos que implican genes conocidos (por ejemplo, oncogenes) en los que la ubicación precisa y/o la pareja de reordenamiento no se conocen antes de realizar los métodos descritos en el presente documento.

En algunas realizaciones, un sujeto necesita tratamiento para el cáncer de pulmón (por ejemplo, con EGFR-TKI, una terapia dirigida contra el cáncer). En algunas realizaciones, por ejemplo, cuando la muestra es obtenida de un sujeto que necesita tratamiento para el cáncer de pulmón, la secuencia diana conocida puede comprender una secuencia de un gen seleccionado entre el grupo de ALK, ROS1 y RET. En consecuencia, en algunas realizaciones, los reordenamientos génicos dan como resultado fusiones que implican ALK, ROS1 o RET. Ejemplos no limitantes de conjuntos de genes que implican ALK, ROS1 o RET se describen en, por ejemplo, Soda *et al.* Nature 2007 448:561-6; Rikova *et al.* Cell 2007 131:1190-1203; Kohno *et al.* Nature Medicine 2012 18:375-7; Takouchi *et al.* Nature Medicine 2012 18:378-81. Sin embargo, debe tenerse en cuenta que la ubicación precisa de un reordenamiento génico y la identidad del segundo gen implicado en el reordenamiento pueden no conocerse de antemano. En consecuencia, en los métodos descritos en el presente documento, la presencia y la identidad de tales reordenamientos pueden detectarse sin tener que conocer la localización del reordenamiento o la identidad del segundo gen implicado en el reordenamiento de genes.

En algunas realizaciones, la secuencia diana conocida puede comprender la secuencia de un gen seleccionado entre el grupo de: ALK, ROS1 y RET.

En algunas realizaciones, la presencia de un reordenamiento génico de ALK en una muestra obtenida de un tumor en un sujeto puede indicar que el tumor es susceptible a tratamiento con un tratamiento seleccionado del grupo que consiste en: un inhibidor de ALK; EGFR; crizotinib (PF-02341066); AP26113; LDK378; 3-39; AF802; IPI-504; ASP3026; AP-26113; X-396; GSK-1838705A; CH5424802; inhibidores de diamino y aminopirimidina de la actividad cinasa ALK tal como NVP-TAE684 y PF-02341066 (véase, por ejemplo, Galkin *et al.*, Proc Natl Acad Sci USA, 2007, 104:270-275; Zou *et al.*, Cancer Res, 2007, 67:4408-4417; Hallberg y Palmer *F1000 Med Reports* 2011 3:21; Sakamoto *et al.*, Cancer Cell 2011 19:679-690; y las moléculas desveladas en el documento WO 04/079326). Un inhibidor de ALK puede incluir cualquier agente que reduzca la expresión y/o la actividad cinasa de ALK o una porción del mismo, incluyendo, por ejemplo, oligonucleótidos, moléculas pequeñas y/o péptidos que reduzcan la expresión y/o la actividad de ALK o una porción de los mismos. Como se utiliza en el presente documento, "cinasa de linfoma anaplásico" o "ALK" se refiere a una tirosina cinasa transmembrana normalmente implicada en regulación neuronal en la forma tipo silvestre. La secuencia nucleotídica del gen ALK y el ARNm es conocida para un número de especies, incluyendo seres humanos (por ejemplo, como se indica en NCBI Gene ID: 238).

En algunas realizaciones, la presencia de un reordenamiento génico de ROS1 en una muestra obtenida de un tumor en un sujeto puede indicar que el tumor es susceptible a tratamiento con un tratamiento seleccionado del grupo que consiste en: un inhibidor de ROS1 y un inhibidor de ALK como se ha descrito anteriormente en el presente documento (por ejemplo, crizotinib). Un inhibidor de ROS1 puede incluir cualquier agente que reduzca la expresión y/o la actividad cinasa de ROS1 o una porción del mismo, incluyendo, por ejemplo, oligonucleótidos, moléculas pequeñas y/o péptidos que reduzcan la expresión y/o la actividad de ROS1 o una porción de los mismos. Como se utiliza en el presente documento, "oncogén 1 c-ros" o "ROS1" (también denominado en la técnica ros-1) se refiere a una tirosina cinasa transmembrana de la subfamilia sevenless y que interactúa con PTPN6. Las secuencias nucleotídicas del gen ROS1 y el ARNm es conocida para una serie de especies, incluyendo seres humanos (por ejemplo, como se indica en NCBI Gene ID: 6098).

En algunas realizaciones, la presencia de un reordenamiento génico de RET en una muestra obtenida de un tumor en un sujeto puede indicar que el tumor es susceptible a tratamiento con un tratamiento seleccionado del grupo que consiste en: un inhibidor de RET; DP-2490, DP-3636, SU5416; BAY 43-9006, BAY 73-4506 (regorafenib), ZD6474, NVP-AST487, sorafenib, RPI-1, XL184, vandetanib, sunitinib, imatinib, pazopanib, axitinib, motesanib, gefitinib y witaferina A (véase, por ejemplo, Samadi *et al.*, Surgery 2010 148:1228-36; Cuccuru *et al.*, JNCI 2004 13:1006-1014; Akeno-Stuart *et al.*, Cancer Research 2007 67:6956; Grazma *et al.*, J Clin Oncol 2010 28:15s 5559; Mologni *et al.*, J. Mol. Endocrinol 2006 37:199-212; Calmomagno *et al.*, Journal NCI 2006 98:326-334; Mologni, Curr Med Chem 2011 18:162-175; y los compuestos desvelados en el documento WO 06/034833; la publicación de patente de Estados Unidos 2011/0201598 y la patente de Estados Unidos n.º 8.067.434. Un inhibidor de RET puede incluir cualquier

agente que reduzca la expresión y/o la actividad cinasa de RET o una porción del mismo, incluyendo, por ejemplo, oligonucleótidos, moléculas pequeñas y/o péptidos que reduzcan la expresión y/o la actividad de RET o una porción de los mismos. Como se utiliza en el presente documento, "reordenado durante la transfección" o "RET" se refiere a un receptor tirosina cinasa de la superfamilia de las cadherinas que está implicado en el desarrollo de la cresta neural y reconoce las moléculas de señalización de la familia del factor neurotrófico derivado de la línea de células gliales. La secuencia nucleotídica del gen RET y el ARNm es conocida para una serie de especies, incluyendo seres humanos (por ejemplo, como se indica en NCBI Gene ID: 5979).

En algunas realizaciones, la secuencia diana conocida puede comprender un gen seleccionado de la Tabla 2.

Tabla 2. Secuencias diana conocidas

GEN	TRANSCRITO de secuencias de referencia NCBI (RefSeq)	EXONES	DIRECCIÓN	TIPO
AKT3	NM_005465	1,2,3	5'	Fusión
ALK	NM_004304	19, (intrón19), 20, 21, 22	5'	Fusión
ARHGAP26	NM_015071	2, 10, 11, 12	5'	Fusión
AXL	NM_021913	19, 20	3'	Fusión
BRAF	NM_004333	7, 8	3'	Fusión
BRAF	NM_004333	7, 8, 9, 10, 11, 12	5'	Fusión
BRAF	NM_004333	15	5'	Fusión
BRAF	NM_004333	V600E	n/d	Mutación
BRD3	NM_007371	9, 10, 11, 12	3'	Fusión
BRD4	NM_014299	10, 11	3'	Fusión
EGFR	NM_005228	7,9, 16,20	5'	Fusión
EGFR	NM_005228	8 (evento de omisión de exón 2-7)	n/d	Mutación
EGFR	NM_005228	24, 25	3'	Fusión
ERG	NM_004449	2, 3, 4, 5, 6, 7, 8, 9, 10, 11	5'	Fusión
ESR1	NM_001122742	3, 4, 5, 6	3'	Fusión
ETV1	NM_004956	3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13	5'	Fusión
ETV4	NM_001986	2, 4, 5, 6, 7, 8, 9, 10	5'	Fusión
ETV5	NM_004454	2, 3, 7, 8, 9	5'	Fusión
ETV6	NM_001987	1, 2, 3, 4, 5, 6	3'	Fusión
ETV6	NM_001987	2, 3, 5, 6, 7	5'	Fusión
EWSR1	NM_005243	4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14	3'	Fusión
FGFR1	NM_015850	2, 8, 9, 10, 17	5'	Fusión
FGFR2	NM_000141	2, 8, 9, 10	5'	Fusión
FGFR2	NM_000141	17	3'	Fusión
FGFR3	NM_000142	17, Intrón 17	3'	Fusión
FGFR3	NM_000142	8, 9, 10	5'	Fusión
FGR	NM_005248	2	5'	Fusión
INSR	NM_000208	20, 21, 22	3'	Fusión

(continuación)

GEN	TRANSCRITO de secuencias de referencia NCBI (RefSeq)	EXONES	DIRECCIÓN	TIPO
INSR	NM_000208	12, 13, 14, 15, 16, 17, 18, 19	5'	Fusión
MAML2	NM_032427	2,3	5'	Fusión
MAST1	NM_014975	7, 8, 9, 18, 19, 20, 21	5'	Fusión
MAST2	NM_015112	2, 3, 5, 6	5'	Fusión
MET	NM_000245	13	3'	Fusión
MET	NM_000245	13, 15 (evento de omisión del exón 14)	n/d	Mutación
MSMB	NM_002443	2,3,4	3'	Fusión
MUSK	NM_005592	7, 8, 9, 11, 12, 13, 14	5'	Fusión
MYB	NM_001130173	7, 8, 9, 11, 12, 13, 14, 15, 16	3'	Fusión
NOTCH1	NM_017617	2, 4, 29, 30, 31	3'	Fusión
NOTCH1	NM_017617	26, 27, 28, 29 (delección interna del exón 3-27)	5'	Fusión
NOTCH2	NM_024408	5, 6, 7	3'	Fusión
NOTCH2	NM_024408	26, 27, 28	5'	Fusión
NRG1	NM_004495	1, 2, 3, 6	5'	Fusión
NTRK1	NM_002529	8, 10, 11, 12, 13	5'	Fusión
NTRK2	NM_006180	11, 12, 13, 14, 15, 16, 17	5'	Fusión
NTRK3	NM_002530	13, 14, 15, 16	5'	Fusión
NTRK3	NM_001007156	15	5'	Fusión
NUMBL	NM_004756	3	5'	Fusión
NUTM1	NM_175741	3	5'	Fusión
PDGFRA	NM_006206	7 (delección del exón 8)	n/d	Mutación
PDGFRA	NM_006206	10, 11, 12, 13, 14,	5'	Fusión
PDGFRA	NM_006206	T674I, D842V	n/d	Mutación
PDGFRB	NM_002609	8, 9, 10, 11, 12, 13, 14	5'	Fusión
PIK3CA	NM_006218	2	5'	Fusión
PKN1	NM_002741	10, 11, 12, 13	5'	Fusión
PPARG	NM_015869	1,2,3	5'	Fusión
PRKCA	NM_002737	4, 5, 6	5'	Fusión
PRKCB	NM_002738	3	5'	Fusión
RAF1	NM_002880	4, 5, 6, 7, 9	3'	Fusión
RAF1	NM_002880	4, 5, 6, 7, 9, 10, 11, 12	5'	Fusión
RELA	NM_021975	3,4	5'	Fusión
RET	NM_020630	8, 9, 10, 11, 12, 13	5'	Fusión
ROS1	NM_002944	31, 32, 33, 34, 35, 36, 37	5'	Fusión

(continuación)

GEN	TRANSCRITO de secuencias de referencia NCBI (RefSeq)	EXONES	DIRECCIÓN	TIPO
RSPO2	NM_178565	1,2	5'	Fusión
RSPO3	NM_032784	2	5'	Fusión
TERT	NM_198253	2	5'	Fusión
TFE3	NM_006521	2, 3, 4, 5, 6	3'	Fusión
TFE3	NM_006521	2, 3, 4, 5, 6, 7, 8	5'	Fusión
TFEB	NM_007162	1,2	5'	Fusión
THADA	NM_022065	28	3'	Fusión
TMPRSS2	NM_005656	1, 2, 3, 4, 5, 6	3'	Fusión
TMPRSS2	NM_001135099	1	3'	Fusión

Ejemplos no limitantes adicionales de aplicaciones de los métodos descritos en el presente documento incluyen detección de marcadores de malignidad hematológica y paneles de los mismos (por ejemplo, incluidos aquellos para detectar los reordenamientos genómicos en linfomas y leucemias), detección de los reordenamientos genómicos relacionados con sarcoma y paneles de los mismos; y detección de los reordenamientos genómicos de IGH/TCR y paneles de los mismos para el ensayo del linfoma.

En algunas realizaciones, los métodos descritos en el presente documento se refieren a tratar a un sujeto que tiene o ha sido diagnosticado como que tiene, por ejemplo, cáncer con un tratamiento para el cáncer. Los sujetos que tienen cáncer pueden ser identificados por un médico utilizando métodos actuales de diagnóstico de cáncer. Por ejemplo, los síntomas y/o complicaciones del cáncer de pulmón que caracterizan estas condiciones y ayudan en el diagnóstico son bien conocidos en la técnica e incluyen, pero sin limitación, respiración débil, ganglios linfáticos inflamados por encima de la clavícula, sonidos anómalos en los pulmones, matidez cuando se da golpecitos en el pecho y dolor torácico. Las pruebas que pueden ayudar al diagnóstico de, por ejemplo, el cáncer de pulmón incluye, aunque sin limitación, rayos x, análisis sanguíneo para detectar niveles elevados de determinadas sustancias (por ejemplo, calcio), escaneos por TC y biopsia tumoral. Un historial familiar de cáncer de pulmón, o la exposición a factores de riesgo para el cáncer de pulmón (por ejemplo, fumar o exposición al humo y/o a la contaminación del aire) pueden ayudar también en la determinación de si el sujeto es probable que tenga cáncer de pulmón o en la realización de un diagnóstico de cáncer de pulmón.

El cáncer puede incluir, pero sin limitación, carcinoma, incluido adenocarcinoma, linfoma, blastoma, melanoma, sarcoma, leucemia, cáncer de células escamosas, cáncer de pulmón microcítico, cáncer de pulmón no microcítico, cáncer gastrointestinal, linfoma de Hodgkin y no Hodgkin, cáncer de páncreas, glioblastoma, carcinoma de células basales, cáncer de las vías biliares, cáncer de vejiga, cáncer cerebral, incluidos glioblastomas y meduloblastomas; cáncer de mama, cáncer cervicouterino, coriocarcinoma; cáncer de colon, cáncer colorrectal, carcinoma de endometrio, cáncer endometrial; cáncer de esófago, cáncer gástrico; diversos tipos de cáncer de cabeza y cuello, neoplasias intraepiteliales, incluida la enfermedad de Bowen y la enfermedad de Paget; neoplasias hematológicas, incluida la leucemia linfocítica y mielógena aguda; sarcoma de Kaposi, tricoleucemia; leucemia mielógena crónica, leucemias relacionadas al SIDA y linfoma o leucemia de linfocitos T del adulto; cáncer renal tal como cáncer de células renales, leucemia/linfoma linfoblástica aguda de linfocitos T, linfomas, incluida la enfermedad de Hodgkin y linfomas linfocíticos; cáncer de hígado tal como carcinoma hepático y hepatoma, carcinoma de células de Merkel, melanoma, mieloma múltiple; neuroblastomas; cáncer oral, incluyendo carcinoma de células escamosas; cáncer de ovario, incluyendo aquellos que surgen de células epiteliales, sarcomas, incluidos leiomiomasarcoma, rabdomiosarcoma, liposarcoma, fibrosarcoma y osteosarcoma; cáncer pancreático; cáncer de piel, incluido el melanoma, células estromales, células germinales y células mesenquimatosas; cáncer de próstata, cáncer de recto; cáncer de vulva, cáncer renal, incluido el adenocarcinoma; cáncer testicular, incluidos tumores germinales tales como seminoma, no seminoma (teratomas, coriocarcinomas), tumores estromales y tumores de células germinales; cáncer de tiroides, incluyendo el adenocarcinoma de tiroides y el carcinoma medular; cáncer de esófago, carcinoma de las glándulas salivales y tumores de Wilms. En algunas realizaciones, el cáncer puede ser cáncer de pulmón.

Métodos multiplex

Los métodos descritos en el presente documento se pueden emplear en un formato multiplex. En realizaciones de los métodos descritos en el presente documento, las múltiples aplicaciones pueden incluir determinar la secuencia nucleotídica contigua a una o más secuencias nucleotídicas diana conocidas. Como se utiliza en el presente documento, "amplificación múltiple" se refiere a un proceso que implica la amplificación simultánea de más de un ácido

nucleico diana en uno o más recipientes de reacción. En algunas realizaciones, los métodos implican la determinación posterior de la secuencia de los productos de amplificación multiplex utilizando uno o más conjuntos de cebadores. Múltiple puede referirse a la detección de entre aproximadamente 2-1.000 secuencias diana diferentes en una única reacción. En algunas realizaciones, sin embargo, múltiple puede referirse a la detección de entre aproximadamente 1.000-10.000 secuencias diana diferentes en una única reacción. En algunas realizaciones, múltiple puede referirse a la detección de entre aproximadamente 10.000-100.000 secuencias diana diferentes en una única reacción. Como se utiliza en el presente documento, múltiple se refiere a la detección de cualquier intervalo entre 2-1.000, por ejemplo, entre 5-500, 25-1.000 o 10-100 secuencias diana diferentes en una reacción única, etc. El término "múltiple" como se aplica a la PCR implica que hay cebadores específicos para al menos dos secuencias diana diferentes en la misma reacción de PCR.

En algunas realizaciones, los ácidos nucleicos diana en una muestra o distintas porciones de una muestra, se puede amplificar con una pluralidad de cebadores (por ejemplo, una pluralidad de cebadores específicos de diana primero y segundo). En algunas realizaciones, la pluralidad de cebadores (por ejemplo, una pluralidad de primer y segundo cebadores específicos de diana) pueden estar presentes en una sola mezcla de reacción, por ejemplo, pueden producirse múltiples productos de amplificación en la misma mezcla de reacción. En algunas realizaciones, la pluralidad de cebadores (por ejemplo, una pluralidad de conjuntos de primeros y segundos cebadores específicos a diana puede aparearse de forma específica con secuencias diana conocidas comprendidas por distintos genes. En algunas realizaciones, al menos dos juegos de cebadores (por ejemplo, al menos dos conjuntos de primeros y segundos cebadores específicos a diana pueden aparearse de forma específica con diferentes porciones de una secuencia diana conocida. En algunas realizaciones, al menos dos conjuntos de cebadores (por ejemplo, al menos dos conjuntos de primeros y segundos cebadores específicos de diana) pueden aparearse de forma específica con diferentes porciones de una secuencia diana conocida comprendida por un único gen. En algunas realizaciones, al menos dos juegos de cebadores (por ejemplo, al menos dos conjuntos de primeros y segundos cebadores específicos a diana pueden aparearse de forma específica con diferentes exones de un gen que comprende una secuencia diana conocida. En algunas realizaciones, la pluralidad de cebadores (por ejemplo, primeros cebadores específicos de diana) puede comprender idénticas porciones de la secuencia del marcador en 5'.

En realizaciones de los métodos descritos en el presente documento, las múltiples aplicaciones pueden incluir determinar la secuencia nucleotídica contigua a una o más secuencias nucleotídicas diana conocidas en múltiples muestras en una reacción de secuenciación o carrera de secuenciación. En algunas realizaciones, múltiples muestras pueden ser de diferentes orígenes, por ejemplo, de diferentes tejidos y/o diferentes sujetos. En tales realizaciones, los cebadores (por ejemplo, cebadores aleatorios con cola) pueden comprender además una porción de código de barras. En algunas realizaciones, se puede añadir un cebador (por ejemplo, un cebador aleatorio con cola) con una parte de código de barras única a cada muestra y ligarlo a los ácidos nucleicos en la misma; las muestras pueden agruparse posteriormente. En tales realizaciones, cada lectura de secuenciación resultante de un producto de amplificación comprenderá un código de barras que identifica la muestra que contiene el ácido nucleico molde del que deriva el producto de amplificación.

Ejemplos

Los siguientes ejemplos pretenden ilustrar ciertas realizaciones descritas en el presente documento, incluyendo determinados aspectos de la presente invención, pero no ejemplifican el alcance completo de la invención.

Ejemplo 1: Preparación de muestras de ácido nucleico

En la figura 1 se muestra un ejemplo de un protocolo que ilustra un método de preparación de una muestra de ácido nucleico para análisis.

Un cebador específico del receptor biotinilado se aparea con la muestra de ARN. Un termociclador se calienta a 65 °C y el ácido nucleico total purificado o el ARN (20-250 ng) se combina con agua libre de nucleasas mientras se encuentra en hielo antes de combinarse con el cebador específico del receptor. Luego, la muestra se transfiere al termociclador y se incuba con una tapa calentada (≥ 100 °C) a 65 °C durante 5 minutos y luego se mantiene a 4 °C. Una vez terminado, la muestra se coloca en hielo durante al menos 2 minutos.

Después del apareamiento, la primera cadena de ADNc se sintetiza por extensión del cebador específico del receptor por la acción de una enzima transcriptasa inversa para generar un híbrido de ADN/ARN. La muestra se incuba en un termociclador con tapa calentada (≥ 100 °C) a 50 °C durante 30 minutos, seguido de 20 minutos a 80 °C y luego mantenido a 4 °C.

La cadena de ARN del híbrido ADN/ARN resultante se degrada parcialmente mediante la acción de una ribonucleasa, lo que deja fragmentos de ARN hibridados con la primera cadena que sirven como cebadores para una reacción de síntesis de la segunda cadena. La segunda cadena de ADNc se sintetiza después de la incubación con ADN Pol en un termociclador con tapa calentada (≥ 100 °C) a 16 °C durante 60 minutos, seguido de 20 minutos a 75 °C y luego mantenido a 4 °C.

La muestra de ADNc bicatenario se somete a reparación de extremos para hacer romos los extremos del ADNc y fosforilar los extremos 5'. En esta etapa, se añade un exceso de ADN polimerasa T4 y polinucleótido cinasa T4 a la muestra junto con suficientes dNTP y se deja incubar durante 30 minutos a 25 °C en un termociclador (sin tapa calentada). El ADN se somete a un paso de limpieza utilizando perlas AMPure® XP (2,5x). Las perlas se vuelven a suspender por completo mediante agitación vorticial y se agregan a cada reacción con mezclado para asegurar una mezcla homogénea. A continuación, la reacción se incuba durante 5 minutos a temperatura ambiente (de 20 a 25 °C). Luego, los tubos se centrifugan y se colocan en un imán durante 4 minutos para garantizar que las perlas se sedimenten por completo contra la pared del tubo. El sobrenadante se descarta sin perturbar el sedimento de perlas y se usa más aumento según sea necesario para volver a peletizar las perlas. Las perlas se lavan con etanol al 70 % durante 30 segundos mientras aún están en el imán antes de desechar el sobrenadante. El lavado se repite dos veces. Tras el lavado final, el residuo sobrenadante visible se elimina por completo y las perlas se secan durante 5 minutos a temperatura ambiente con las tapas abiertas. Las perlas no deben secarse demasiado, ya que esto reduce significativamente el rendimiento general de ácido nucleico. El ADN se eluye volviendo a suspender las perlas en Tris-HCl 10 mM, a pH 8,0. Luego, los tubos se vuelven a colocar en el imán durante 2 minutos.

En una primera etapa de ligadura, el ADN purificado se somete después a una reacción de cola dA que incorpora dAMP en los extremos 3' de las cadenas de ADN durante la incubación en un termociclador con tapa calentada (≥ 100 °C) durante 15 minutos a 37 °C y luego se mantiene a 4 °C. A continuación, la reacción se centrifuga y se coloca en hielo.

Siguiendo la cola A, las muestras se limpian utilizando perlas AMPure® XP (2,5x) siguiendo el mismo procedimiento descrito anteriormente. El ADN se eluye usando agua libre de nucleasas.

En una segunda etapa de ligadura, las secuencias de nucleótidos únicas o códigos de barras moleculares (MBC) se ligan al ADN a través de la acción de la ADN ligasa después de la incubación en un termociclador (sin tapa calentada) durante 15 minutos a 25 °C y luego se mantienen a 4 °C. A continuación, las muestras se purifican utilizando perlas recubiertas de estreptavidina. Las muestras se colocan en un imán durante 1 minuto o hasta que se sedimenten las perlas. El sobrenadante se elimina con una pipeta sin perturbar el sedimento de perlas. Luego, las perlas se vuelven a suspender en tampón de limpieza de ligadura (NaCl 1 M, EDTA 1 mM, 0,1 % de Tween, Tris 10 mM a pH 8). El producto de ADN ligado (50 μ l) se mezcla con perlas de limpieza de ligadura (50 μ l para un total de 100 μ l) y la reacción se incuba a temperatura ambiente durante 5 minutos, seguido de mezclado y otros 5 minutos de incubación. Las muestras se centrifugan y se colocan en un imán durante 1 minuto para garantizar que las perlas se sedimenten por completo contra la pared del tubo. El sobrenadante se desecha sin perturbar las perlas y las perlas se lavan con tampón de ligadura y se colocan en un imán durante 1 minuto. Una vez que la suspensión se haya aclarado, el sobrenadante se descarta. Luego, las perlas se lavan dos veces con tampón y luego una vez con agua libre de nucleasas. Luego, las perlas unidas al adaptador MBC se transfieren a una mezcla separada de componentes para una primera etapa de PCR.

Se realiza una primera ronda de PCR utilizando un primer cebador específico de gen y un primer cebador adaptador. La reacción se mantiene en hielo antes de realizar la primera PCR en un termociclador con tapa calentada (≥ 100 °C) utilizando el programa descrito en la Tabla 3. Una vez que la reacción ha alcanzado los 4 °C, las reacciones se centrifugan brevemente y se colocan en hielo.

Tabla 3: Condiciones de PCR para la primera reacción de PCR.

Etapa	Temperatura (°C)	Tiempo (min)	Ciclos
1	95	3 min	1
2	95	30 s	24
3	65	3 min (velocidad de rampa del 100 %)	
4	72	3 min	1
5	4	Retención 1	

Luego, la reacción de PCR se limpia con perlas AMPure® XP (1,2x) y se incuba durante 5 minutos a temperatura ambiente (20 a 25 °C). Luego, los tubos se centrifugan brevemente y se colocan en un imán durante 4 minutos para garantizar que las perlas se sedimenten por completo contra la pared del tubo. El sobrenadante se desecha sin perturbar el sedimento de perlas. Luego, los tubos se lavan durante 30 segundos con etanol al 70 % mientras permanecen en el imán. El lavado se repite dos veces. Tras el lavado final, se elimina el sobrenadante con una pipeta y se secan los tubos durante 3 minutos a temperatura ambiente. Las perlas no deben secarse demasiado, ya que esto reduce significativamente el rendimiento general de ácido nucleico. El ADN se eluye volviendo a suspender las perlas en Tris-HCl 10 mM, a pH 8,0. Luego, los tubos se colocan en el imán durante 2 minutos antes de que el sobrenadante se transfiera a una mezcla separada de componentes para una segunda etapa de PCR.

Se realiza una segunda ronda de PCR utilizando un segundo cebador específico de gen y un segundo cebador

adaptador. La segunda etapa de PCR incorpora una cola P7, que se incorpora como una región de cola en 5' del segundo cebador específico del gen, como se muestra en la Figura 1. Las secuencias para las etiquetas de Índice 1 (P7) se muestran en la Tabla 4.

5

Tabla 4: Índice 1 (P7) Tabla de secuencias.

Número de muestra	Índice Illumina 1 P7/i7 Secuencia
1	TAAGGCGA
2	CGTACTAG
3	AGGCAGAA
4	TCCTGAGC
5	GGACTCCT
6	TAGGCATG
7	CTCTCTAC
8	CAGAGAGG

El ADN de biblioteca purificado de la primera PCR se mezcla con el segundo cebador específico del gen y el segundo cebador adaptador y los componentes de la PCR, y la segunda PCR se realiza en el termociclador con tapa calentada ($\geq 100^\circ\text{C}$) utilizando el programa descrito en Tabla 5. Una vez que la reacción ha alcanzado los 4°C , las reacciones se centrifugan brevemente y se colocan en hielo.

10

Tabla 5: Condiciones de PCR para la segunda reacción de PCR.

Etapas	Temperatura ($^\circ\text{C}$)	Tiempo	Ciclos
1	95	3 min	1
2	95	30 s	6
3	65	3 min (velocidad de rampa del 100 %)	
4	72	3 min	1
5	4	Retención	1

A continuación, la reacción de PCR se limpia con perlas AMPure® XP (1,2x) siguiendo el mismo procedimiento descrito en la primera reacción de PCR. El ADN se eluye volviendo a suspender las perlas en Tris-HCl 10 mM pH 8,0 e incubando en un imán durante 2 minutos. Luego, el ADN marcado con la biblioteca se transfiere a un nuevo tubo de PCR para su almacenamiento, cuantificación o normalización y secuenciación.

15

Ejemplo 2: Diversidad y reproducibilidad

20

La información de secuenciación se obtuvo de muestras de cantidades de entrada variadas (25, 100, 200, 400, 800 y 1600 ng) en experimentos repetidos utilizando los métodos descritos anteriormente. Tal como se muestra en la Figura 4, la superposición de clonotipos entre experimentos repetidos aumenta con la cantidad entrante. Se realizó un análisis pareado de muestras replicadas utilizando datos de 256 000 lecturas. Los resultados, que demostraron la naturaleza altamente reproducible del ensayo, se muestran para muestras de entrada de 800 y 1600 ng en la Figura 5. Las Figuras 6A y 6B representan el análisis completo lado a lado de todos los clones (FIG. 6A) frente a los clones superpuestos (FIG. 6B). La Figura 7 muestra un gráfico de diversidad versus tamaño de muestra (arriba) y un gráfico del índice de diversidad de Shannon por tamaño de entrada (abajo), que muestran que la cantidad entrante impulsa la complejidad y la diversidad de la observación.

25

30

Como se muestra en las Figuras 8-10, se utilizó una serie de diluciones de Jurkat para probar la reproducibilidad intra e interlaboratorios. La línea celular Jurkat que expresa el receptor $\text{TCR}\alpha:\beta$ se añadió al ARN de linfocitos de sangre periférica (PBL) de donantes sanos para determinar los límites de detección de la cadena beta del receptor de linfocitos T (TRB) y para evaluar la variación del ensayo entre laboratorios. Se realizó una dilución en serie de ARN total de Jurkat en ARN de PBL que varió de dilución 1:10 a 1:10.000 [ng Jurkat / (ng Jurkat + ng PBL)] por duplicado.

35

La dilución en serie se dividió en dos alícuotas por dilución y las bibliotecas se prepararon y secuenciaron. Los resultados en la figura 11 muestran una fuerte correlación entre ambos laboratorios.

40

Todas las bibliotecas se normalizaron a 600.000, se desduplicaron y se corrigieron los errores. Las frecuencias

esperadas se determinaron multiplicando las frecuencias de clonotipo de las diluciones 1:10 por 10 y dividiendo el número resultante por los respectivos factores de dilución (por ejemplo, factor = 0,5 (dilución 1:10) x 10 = 5. 5/factor de dilución = frecuencia experimental para un factor de dilución dado).

5 Ejemplo 3: Cebadores

La Figura 12 representa una estrategia optimizada por FFPE para el análisis del repertorio de TCR (β/γ). Tal como se muestra, el sitio de unión del cebador de la transcriptasa inversa (RT) corresponde a una región en 5' del exón del dominio constante. Los primeros cebadores específicos de genes (GSP) de PCR están diseñados para unirse a una región que corresponde a una región en 3' del segmento J o una región que abarca la intersección del segmento J: exón C. La distancia promedio desde el extremo 5' del cebador RT hasta el extremo 5' de CDR3 es inferior a 100 pares de bases. Se diseñaron paneles de cebadores para la secuenciación del repertorio inmunitario de TCR (β/γ) y TCR (α/δ), enumerados en la Tabla 6 y en la Tabla 7, respectivamente.

15 Tabla 6: Panel de cebadores de receptores de linfocitos T (β/γ)

Cebadores de transcriptasa inversa	
Nombre del cebador	Secuencia (secuencia con nucleótidos modificados debajo de cada uno)
TRAC_RT_19_60.7 _BIOTINA_rediseño	CACTGGATTTAGAGTCTCTCAGC (SEQ ID NO: 1) /52-Bio/C*ACTGGATTTAGAGTCTCTCAGC (SEQ ID NO: 2)
TRBC1_RT_2_59.8 _BIOTINA	GAACACCTTGTTTCAGGTCCT (SEQ ID NO: 3) /52-Bio/G*AACACCTTGTTTCAGGTCCT (SEQ ID NO: 4)
TRBC2_RT_1_59.8 _BIOTINA	GAACACGTTTTTCAGGTCCTC (SEQ ID NO: 5) /52-Bio/G*AACACGTTTTTCAGGTCCTC (SEQ ID NO: 6)
TRDC_RT_37_59_ _BIOTIN	TCTTATATCCTTGGGGTAGAATTCC (SEQ ID NO: 7) /52-Bio/T*CTTATATCCTTGGGGTAGAATTCC (SEQ ID NO: 8)
TRGC_RT_24_univ _60.6_BIOTIN	GGGAAACATCTGCATCAAGTTG (SEQ ID NO: 9) /52-Bio/G*GGGAAACATCTGCATCAAGTTG (SEQ ID NO: 10)
Primeros cebadores específicos de genes de PCR	
Nombre del cebador	Secuencia (secuencia con cola debajo de cada una)
TRBJ1-1*01_12_69.7	CAGGTCCTCTACAACGTGTGAGTCTGG (SEQ ID NO: 11) AGACGTGTGCTCTTCCGATCTCAGGTCCTCTACAACGTGTGAGTCTGG (SEQ ID NO: 12)
TRBJ1-2*01_3_70.1	GGTCCTCTACAACGGTTAACCTGGTC (SEQ ID NO: 13) AGACGTGTGCTCTTCCGATCTGGTCCTCTACAACGGTTAACCTGGTC (SEQ ID NO: 14)
TRBJ1-3*01_10_70.1	GGTCCTCTACAACAGTGAGCCAACTT (SEQ ID NO: 15) AGACGTGTGCTCTTCCGATCTGGTCCTCTACAACAGTGAGCCAACTT (SEQ ID NO: 16)
TRBJ1-4*01_11_70.1	GGTCCTCCAAGACAGAGAGCTGG (SEQ ID NO: 17) AGACGTGTGCTCTTCCGATCTGGTCCTCCAAGACAGAGAGCTGG (SEQ ID NO: 18)
TRBJ1-5*01_9_70.4	GGTCCTCTAGGATGGAGAGTCGAGTC (SEQ ID NO: 19) AGACGTGTGCTCTTCCGATCTGGTCCTCTAGGATGGAGAGTCGAGTC (SEQ ID NO: 20)
TRBJ1-6_7_70.5	GGTCCTCTGTACAGTGAGCCTG (SEQ ID NO: 21) AGACGTGTGCTCTTCCGATCTGGTCCTCTGTACAGTGAGCCTG (SEQ ID NO: 22)
TRBJ2-1*01_1_70.1	TCTAGCACGGTGAGCCGTGT (SEQ ID NO: 23) AGACGTGTGCTCTTCCGATCTTCTAGCACGGTGAGCCGTGT (SEQ ID NO: 24)

(continuación)

Primeros cebadores específicos de genes de PCR

Nombre del cebador	Secuencia (secuencia con cola debajo de cada una)
TRBJ2-2*01_6_70.0	CAGTACGGTCAGCCTAGAGCCTTC (SEQ ID NO: 25) AGACGTGTGCTCTTCCGATCTCAGTACGGTCAGCCTAGAGCCTTC (SEQ ID NO: 26)
TRBJ2-3*01_2_70.0	TTCAGGTCCTCGAGCACTGTCAG (SEQ ID NO: 27) AGACGTGTGCTCTTCCGATCTTTCAGGTCCTCGAGCACTGTCAG (SEQ ID NO: 28)
TRBJ2-4*01_2_69.5	TTCAGGTCCTCCAGCACTGAGAG (SEQ ID NO: 29) AGACGTGTGCTCTTCCGATCTTTCAGGTCCTCCAGCACTGAGAG (SEQ ID NO: 30)
TRBJ2-5*01_1_69.9	CAGGTCCTCGAGCACCAGGA (SEQ ID NO: 31) AGACGTGTGCTCTTCCGATCTCAGGTCCTCGAGCACCAGGA (SEQ ID NO: 32)
TRBJ2-6*01_1_69.5	CAGCACGGTCAGCCTGCT (SEQ ID NO: 33) AGACGTGTGCTCTTCCGATCTCAGCACGGTCAGCCTGCT (SEQ ID NO: 34)
TRBJ2-7*01_2_69.7	TTCAGGTCCTCTGTGACCGTGAG (SEQ ID NO: 35) AGACGTGTGCTCTTCCGATCTTTCAGGTCCTCTGTGACCGTGAG (SEQ ID NO: 36)
TRGJ1_2_C_Ex1_2_0_69.5	ACATCTGCATCAAGTTGTTTATCTGTGACAAC (SEQ ID NO: 37) AGACGTGTGCTCTTCCGATCTACATCTGCATCAAGTTGTTTATCTGTGACAAC (SEQ ID NO: 38)
TRGJP*01_C_Ex1_1_69.6	TGTTTATCTGTAATGATAAGCTTTGTTCCGGGA (SEQ ID NO: 39) AGACGTGTGCTCTTCCGATCTTGTTTATCTGTAATGATAAGCTTTGTTCCGGGA (SEQ ID NO: 40)
TRGJP1*01_C_Ex1_8_69.8	TCAGGTGAAGTTACTATGAGCTTAGTCCCT (SEQ ID NO: 41) AGACGTGTGCTCTTCCGATCTTCAGGTGAAGTTACTATGAGCTTAGTCCCT (SEQ ID NO: 42)
TRGJP2*01_C_Ex1_5_69.3	GCGAAGTTACTATGAGCCTAGTCCCTT (SEQ ID NO: 43) AGACGTGTGCTCTTCCGATCTGCGAAGTTACTATGAGCCTAGTCCCTT (SEQ ID NO: 44)

Tabla 7: Panel de cebadores de receptores de linfocitos T (α/δ)**Cebadores de transcriptasa inversa**

Nombre del cebador	Secuencia (secuencia con nucleótidos modificados debajo de cada uno)
TRAC_RT_19_60.7 _BIOTINA_rediseño	CACTGGATTTAGAGTCTCTCAGC(SEQ ID NO: 1) /52-Bio/C*ACTGGATTTAGAGTCTCTCAGC(SEQ ID NO: 2)
TRBC1_RT_2_59.8 _BIOTINA	GAACACCTTGTTTCAGGTCCT(SEQ ID NO: 3) /52-Bio/G*AACACCTTGTTTCAGGTCCT(SEQ ID NO: 4)
TRBC2_RT_1_59.8 _BIOTINA	GAACACGTTTTTCAGGTCCTC(SEQ ID NO: 5) /52-Bio/G*AACACGTTTTTCAGGTCCTC(SEQ ID NO: 6)
TRDC_RT_37_59 _BIOTIN	TCTTATATCCTTGGGGTAGAATTCC(SEQ ID NO: 7) /52-Bio/T*CTTATATCCTTGGGGTAGAATTCC(SEQ ID NO: 8)
TRGC RT 24 univ _60.6 _BIOTIN	GGGAAACATCTGCATCAAGTTG(SEQ ID NO: 9) /52-Bio/G*GGAAACATCTGCATCAAGTTG(SEQ ID NO: 10)

(continuación)

Primeros cebadores específicos de genes de PCR

Nombre del cebador	Secuencia (secuencia con cola debajo de cada una)
TRAC_PCR_9_69	TCTCTCAGCTGGTACACGGCA(SEQ ID NO: 45) AGACGTGTGCTCTTCCGATCTTCTCTCAGCTGGTACACGGCA(SEQ ID NO: 46)
TRDC_PCR_21_69	TCACCAGACAAGCGACATTTGTTCC(SEQ ID NO: 47) AGACGTGTGCTCTTCCGATCTTACCAGACAAGCGACATTTGTTCC(SEQ ID NO: 48)

El panel de cebadores de BCR IgH tiene algunos requisitos adicionales además de la identificación del clonotipo (CDR3), en comparación con los paneles de TCR, ya que el panel de IgH determina el isotipo (A, D, E, G, M) y selecciona las subclases. También determina si el clon ha sufrido una hipermutación somática (análisis del segmento V). Las ubicaciones de los cebadores de IgH se muestran en la Figura 13. El cebador RT y el primer cebador específico del gen PCR se unen a las regiones correspondientes al exón CH1 de la región constante del isotipo respectivo para distinguir las subclases de isotipo seleccionadas. Se diseñaron paneles de cebadores para la secuenciación del repertorio inmunitario de BCR (IgH) y BCR (IgK/IgL), enumerados en la Tabla 8 y en la Tabla 9, respectivamente

Tabla 8: Panel de cebadores de receptores de linfocitos B (IgH)

Cebadores de transcriptasa inversa

Nombre del cebador	Secuencia (secuencia con nucleótidos modificados debajo de cada uno)
IGHA_RT_4_59_BIO TINA	GGCTCCTGGGGGAAGA(SEQ ID NO: 49) /52-Bio/G*GCTCCTGGGGGAAGA(SEQ ID NO: 50)
IGHD_RT_40_59_BI OTINA	GTACCCAGTTATCAAGCATGC(SEQ ID NO: 51) /52-Bio/G*TACCCAGTTATCAAGCATGC(SEQ ID NO: 52)
IGHE_RT_33_60_BI OTINA	AGTCACGGAGGTGGCAT(SEQ ID NO: 53) /52-Bio/A*GTCACGGAGGTGGCAT(SEQ ID NO: 54)
IGHG_RT_18_59_BI OTINA	GACACCGTCACCGGTTC(SEQ ID NO: 55) /52-Bio/G*ACACCGTCACCGGTTC(SEQ ID NO: 56)
IGM_RT_27_58_BIO TINA	GAAGGAAGTCCTGTGCGA(SEQ ID NO: 57) /52-Bio/G*AAGGAAGTCCTGTGCGA(SEQ ID NO: 58)
IG LC_RT_72_mayor _60_BIOTIN	TTGACGGGGCTGCTATCT(SEQ ID NO: 59) /52-Bio/T*TGACGGGGCTGCTATCT(SEQ ID NO: 60)
IG LC_RT_72_minor _59_BIOTIN	ACGGGGCTGCCATCT(SEQ ID NO: 61) /52-Bio/A*CGGGGCTGCCATCT(SEQ ID NO: 62)
IGKC_RT_10_59_BI OTINA	CAGATTTCAACTGCTCATCAGA(SEQ ID NO: 63) /52-Bio/C*AGATTTCAACTGCTCATCAGA(SEQ ID NO: 64)

Cebadores de transcriptasa inversa (diseño alternativo 1)

Nombre del cebador	Secuencia (secuencia con nucleótidos modificados debajo de cada uno)
IGHA_RT_4_59_BIO TINA	GGCTCCTGGGGGAAGA(SEQ ID NO: 49) /52-Bio/G*GCTCCTGGGGGAAGA(SEQ ID NO: 50)
IGHD_RT_40_59_BI OTINA	GTACCCAGTTATCAAGCATGC(SEQ ID NO: 51) /52-Bio/G*TACCCAGTTATCAAGCATGC(SEQ ID NO: 52)
IGHE_RT_33_60_BI OTINA	AGTCACGGAGGTGGCAT(SEQ ID NO: 53) /52-Bio/A*GTCACGGAGGTGGCAT(SEQ ID NO: 54)
IGHG_PCR_4_58_ BIOTINA	GGAAGTAGTCCTTGACCA(SEQ ID NO: 65) /52-Bio/G*GGAAGTAGTCCTTGACCA(SEQ ID NO: 66)

(continuación)

Cebadores de transcriptasa inversa (diseño alternativo 1)

Nombre del cebador	Secuencia (secuencia con nucleótidos modificados debajo de cada uno)
IGM_RT_27_58_BIO TINA	GAAGGAAGTCCTGTGCGA(SEQ ID NO: 57) /52-Bio/G*AAGGAAGTCCTGTGCGA(SEQ ID NO: 58)
_60_BIOTINA	IGLC_RT_72_major TTGACGGGGCTGCTATCT(SEQ ID NO: 59) /52-Bio/T*TGACGGGGCTGCTATCT(SEQ ID NO: 60)
IGLC_RT_72_minor _59_BIOTIN	ACGGGGCTGCCATCT(SEQ ID NO: 61) /52-Bio/A*CGGGGCTGCCATCT(SEQ ID NO: 62)
IGKC_RT_10_59_BI OTINA	CAGATTCAACTGCTCATCAGA(SEQ ID NO: 63) /52-Bio/C*AGATTCAACTGCTCATCAGA(SEQ ID NO: 64)

Primeros cebadores específicos de genes de PCR

Nombre del cebador	Secuencia (secuencia con cola debajo de cada una)
IGHA_PCR_4_69	AGGCTCAGCGGGAAGACCT(SEQ ID NO: 67) AGACGTGTGCTCTTCCGATCTAGGCTCAGCGGGAAGACCT(SEQ ID NO: 68)
IGHD_PCR_27_69	CAGGGCTGTTATCCTTTGGGTGTC(SEQ ID NO: 69) AGACGTGTGCTCTTCCGATCTCAGGGCTGTTATCCTTTGGGTGTC(SEQ ID NO: 70)
IGHE_PCR_24_68	GAGGTGGCATTGGAGGGAATGT(SEQ ID NO: 71) AGACGTGTGCTCTTCCGATCTGAGGTGGCATTGGAGGGAATGT(SEQ ID NO: 72)
IGHG_PCR_4_69	TTCGGGGAAGTAGTCCTTGACCA(SEQ ID NO: 73) AGACGTGTGCTCTTCCGATCTTTCGGGGAAGTAGTCCTTGACCA(SEQ ID NO: 74)
IGHG_PCR_4_min or_69	GGTCTGGGAAGTAGTCCTTGACCA(SEQ ID NO: 75) AGACGTGTGCTCTTCCGATCTGGTCTGGGAAGTAGTCCTTGACCA(SEQ ID NO: 76)
IGM_PCR_17_69	TCGTATCCGACGGGAATTCTCAC(SEQ ID NO: 77) AGACGTGTGCTCTTCCGATCTTCGTATCCGACGGGAATTCTCAC(SEQ ID NO: 78)
IGKC_PCR_22_70	TGCTCATCAGATGGCGGGAAGAT(SEQ ID NO: 79) AGACGTGTGCTCTTCCGATCTTGCTCATCAGATGGCGGGAAGAT(SEQ ID NO: 80)
IGLC_PCR_17_princi pal_69	CCTTGTGGCTTGAAGCTCCTCA(SEQ ID NO: 81) AGACGTGTGCTCTTCCGATCTCCTTGTGGCTTGAAGCTCCTCA(SEQ ID NO: 82)
IGLC_PCR_17_minor itario_68	CTTGTGGCTTGGAGCTCCTCA(SEQ ID NO: 83) AGACGTGTGCTCTTCCGATCTCTTGTGGCTTGGAGCTCCTCA(SEQ ID NO: 84)

Primeros cebadores específicos de genes de PCR (diseño alternativo 1)

Nombre del cebador	Secuencia (secuencia con cola debajo de cada una)
IGHA_PCR_4_69	AGGCTCAGCGGGAAGACCT(SEQ ID NO: 67) AGACGTGTGCTCTTCCGATCTAGGCTCAGCGGGAAGACCT(SEQ ID NO: 68)
IGHD_PCR_27_69	CAGGGCTGTTATCCTTTGGGTGTC(SEQ ID NO: 69) AGACGTGTGCTCTTCCGATCTCAGGGCTGTTATCCTTTGGGTGTC(SEQ ID NO: 70)
IGHE_PCR_24_68	GAGGTGGCATTGGAGGGAATGT(SEQ ID NO: 71) AGACGTGTGCTCTTCCGATCTGAGGTGGCATTGGAGGGAATGT(SEQ ID NO: 72)
IGHG1_p38_70.2	CCCAGAGGTGCTCTTGAGGAG(SEQ ID NO: 85) AGACGTGTGCTCTTCCGATCTCCCAGAGGTGCTCTTGAGGAG(SEQ ID NO: 86)

(continuación)

Primeros cebadores específicos de genes de PCR (diseño alternativo 1)

Nombre del cebador	Secuencia (secuencia con cola debajo de cada una)
IGHG2_4_p48_71.1	GCTGTGCTCTCGGAGGTGCT (SEQ ID NO: 87) AGACGTGTGCTCTTCCGATCTGCTGTGCTCTCGGAGGTGCT (SEQ ID NO: 88)
IGHG3_p40_70.6	CGGAGGTGCTCCTGGAGCA (SEQ ID NO: 89) AGACGTGTGCTCTTCCGATCTCGGAGGTGCTCCTGGAGCA (SEQ ID NO: 90)
IGM_PCR_17_69	TCGTATCCGACGGGGAATTCTCAC (SEQ ID NO: 77) AGACGTGTGCTCTTCCGATCTTCGTATCCGACGGGGAATTCTCAC (SEQ ID NO: 78)

Primeros cebadores específicos de genes de PCR (diseño alternativo 2)

Nombre del cebador	Secuencia (secuencia con cola debajo de cada una)
IGNHA_universal	GCGA/ideoxil/GACCACGTTCCCATCT (SEQ ID NO: 91) AGACGTGTGCTCTTCCGATCTGCGA/ideoxyl/GACCACGTTCCCATCT (SEQ ID NO: 92)
IGHA1_p54_69.0	GCGATGACCACGTTCCCATCT (SEQ ID NO: 93) AGACGTGTGCTCTTCCGATCTGCGATGACCACGTTCCCATCT (SEQ ID NO: 94)
IGHA1_p54_69.7	GCGATGACCACGTTCCCATCTG (SEQ ID NO: 95) AGACGTGTGCTCTTCCGATCTGCGATGACCACGTTCCCATCTG (SEQ ID NO: 96)

Primeros cebadores específicos de genes de PCR (diseño alternativo 2)

Nombre del cebador	Secuencia (secuencia con cola debajo de cada una)
IGHA2_p54_71.1	GCGACGACCACGTTCCCATCT (SEQ ID NO: 97) AGACGTGTGCTCTTCCGATCTGCGACGACCACGTTCCCATCT (SEQ ID NO: 98)
IGHA2_p54_70	GCGACGACCACGTTCCCATC (SEQ ID NO: 99) AGACGTGTGCTCTTCCGATCTGCGACGACCACGTTCCCATC (SEQ ID NO: 100)
IGHA1_p43_70.1	TTCCCATCTGGCTGGGTGCT (SEQ ID NO: 101) AGACGTGTGCTCTTCCGATCTTTCCCATCTGGCTGGGTGCT (SEQ ID NO: 102)
IGHA2_p43_70.2	TTCCCATCTTGGGGGGTGCT (SEQ ID NO: 103) AGACGTGTGCTCTTCCGATCTTTCCCATCTTGGGGGGTGCT (SEQ ID NO: 104)

Primeros cebadores específicos de genes de PCR (diseño alternativo 3)

Nombre del cebador	Secuencia (secuencia con cola debajo de cada una)
IGHD_PCR_27_69	CAGGGCTGTTATCCTTTGGGTGTC (SEQ ID NO: 69) AGACGTGTGCTCTTCCGATCTCAGGGCTGTTATCCTTTGGGTGTC (SEQ ID NO: 70)
IGHE_PCR_24_68	GAGGTGGCATTGGAGGGAATGT (SEQ ID NO: 71) AGACGTGTGCTCTTCCGATCTGAGGTGGCATTGGAGGGAATGT (SEQ ID NO: 72)
IGHG1_p38_70.2	CCCAGAGGTGCTCTTGAGGAG (SEQ ID NO: 85) AGACGTGTGCTCTTCCGATCTCCCAGAGGTGCTCTTGAGGAG (SEQ ID NO: 86)
IGHG2_4_p48_71.1	GCTGTGCTCTCGGAGGTGCT (SEQ ID NO: 87) AGACGTGTGCTCTTCCGATCTGCTGTGCTCTCGGAGGTGCT (SEQ ID NO: 88)
IGHG3_p40_70.6	CGGAGGTGCTCCTGGAGCA (SEQ ID NO: 89) AGACGTGTGCTCTTCCGATCTCGGAGGTGCTCCTGGAGCA (SEQ ID NO: 90)
IGM_PCR_17_69	TCGTATCCGACGGGGAATTCTCAC (SEQ ID NO: 77) AGACGTGTGCTCTTCCGATCTTCGTATCCGACGGGGAATTCTCAC (SEQ ID NO: 78)

Tabla 9: Panel de cebadores de receptores de linfocitos B (IgK/IgL)

Cebadores de transcriptasa inversa	
Nombre del cebador	Secuencia (secuencia con nucleótidos modificados debajo de cada uno)
IGHA_RT_4_59_BIOTINA	GGCTCCTGGGGGAAGA(SEQ ID NO: 49) /52-Bio/G*GCTCCTGGGGGAAGA(SEQ ID NO: 50)
IGHD_RT_40_59_BIOTINA	GTACCCAGTTATCAAGCATGC(SEQ ID NO: 51) /52-Bio/G*TACCCAGTTATCAAGCATGC(SEQ ID NO: 52)
IGHE_RT_33_60_BIOTINA	AGTCACGGAGGTGGCAT(SEQ ID NO: 53) /52-Bio/A*GTCACGGAGGTGGCAT(SEQ ID NO: 54)
IGHG_RT_18_59_BIOTINA	GACACCGTCACCGGTTTC(SEQ ID NO: 55) /52-Bio/G*ACACCGTCACCGGTTTC(SEQ ID NO: 56)
IGM_RT_27_58_BIOTINA	GAAGGAAGTCCTGTGCGA(SEQ ID NO: 57) /52-Bio/G*AAGGAAGTCCTGTGCGA(SEQ ID NO: 58)
IGLC_RT_72_mayor_60_BIOTINA	TTGACGGGGCTGCTATCT(SEQ ID NO: 59) /52-Bio/T*TGACGGGGCTGCTATCT(SEQ ID NO: 60)
IGLC_RT_72_minor_59_BIOTINA	ACGGGGCTGCCATCT(SEQ ID NO: 61) /52-Bio/A*CGGGGCTGCCATCT(SEQ ID NO: 62)
IGKC_RT_10_59_BIOTINA	CAGATTTCAACTGCTCATCAGA(SEQ ID NO: 63) /52-Bio/C*AGATTTCAACTGCTCATCAGA(SEQ ID NO: 64)
Cebadores de transcriptasa inversa (diseño alternativo 1)	
Nombre del cebador	Secuencia (secuencia con nucleótidos modificados debajo de cada uno)
IGHA_RT_4_59_BIOTINA	GGCTCCTGGGGGAAGA(SEQ ID NO: 49) /52-Bio/G*GCTCCTGGGGGAAGA(SEQ ID NO: 50)
IGHD_RT_40_59_BIOTINA	GTACCCAGTTATCAAGCATGC(SEQ ID NO: 51) /52-Bio/G*TACCCAGTTATCAAGCATGC(SEQ ID NO: 52)
IGHE_RT_33_60_BIOTINA	AGTCACGGAGGTGGCAT(SEQ ID NO: 53) /52-Bio/A*GTCACGGAGGTGGCAT(SEQ ID NO: 54)
IGHG_PCR_4_58_BIOTINA	GGGAAGTAGTCCTTGACCA(SEQ ID NO: 65) /52-Bio/G*GGAAGTAGTCCTTGACCA(SEQ ID NO: 66)
IGM_RT_27_58_BIOTINA	GAAGGAAGTCCTGTGCGA(SEQ ID NO: 57) /52-Bio/G*AAGGAAGTCCTGTGCGA(SEQ ID NO: 58)
_60_BIOTINA	IGLC_RT_72_mayoritaria TTGACGGGGCTGCTATCT(SEQ ID NO: 59) /52-Bio/T*TGACGGGGCTGCTATCT(SEQ ID NO: 60)
IGLC_RT_72_minoritaria ACGGGGCTGCCATCT(SEQ ID NO: 61) _59_BIOTIN	/52-Bio/A*CGGGGCTGCCATCT(SEQ ID NO: 62)
IGKC_RT_10_59_BIOTINA	CAGATTTCAACTGCTCATCAGA(SEQ ID NO: 63) /52-Bio/C*AGATTTCAACTGCTCATCAGA(SEQ ID NO: 64)

(continuación)

Primeros cebadores específicos de genes de PCR

Nombre del cebador	Secuencia (secuencia con cola debajo de cada una)
IGKC_PCR_22_70	TGCTCATCAGATGGCGGGAAGAT(SEQ ID NO: 79) AGACGTGTGCTCTTCCGATCTTGCTCATCAGATGGCGGGAAGAT(SEQ ID NO: 80)
IGLC_PCR_17_principal_69	CCTTGTTGGCTTGAAGCTCCTCA(SEQ ID NO: 81) AGACGTGTGCTCTTCCGATCTCCTTGTTGGCTTGAAGCTCCTCA(SEQ ID NO: 82)
IGLC_PCR_17_minoritario_68	CTTGTTGGCTTGGAGCTCCTCA(SEQ ID NO: 83) AGACGTGTGCTCTTCCGATCTCTTGTTGGCTTGGAGCTCCTCA(SEQ ID NO: 84)

Las secuencias de nucleótidos de los cebadores de RT en las Tablas 6-9 se enumeran alternativamente con las modificaciones que se muestran con notación de acuerdo con la nomenclatura de Integrated DNA Technologies (IDT), donde "*" denota un enlace de fosforotioato entre el nucleótido que precede a "*" y el nucleótido que sigue a "*" en la secuencia, y "/52-Bio/" indica un resto de biotina dual en 5'. La Figura 14 muestra un ejemplo de un cebador de RT que tiene un resto de biotina dual en 5' unido a la base más 5' ("A"), con un enlace fosforotioato entre el 5' más base y el penúltimo 5' más base ("C"). El resto de biotina dual permite la selección de productos de RT y asegura una mayor eficiencia de captura. Sin embargo, es posible utilizar una sola biotina. El enlace fosforotioato evita la degradación por exonucleasa para garantizar que no se elimine la base que contiene biotina.

REIVINDICACIONES

1. Un método para preparar ácidos nucleicos para análisis, comprendiendo el método:

- 5 (a) poner en contacto una molécula de ácido nucleico que comprende una secuencia de nucleótidos diana con un cebador modificado con un resto de captura que se aparea específicamente con la secuencia de nucleótidos diana en condiciones de hibridación;
- (b) llevar a cabo una reacción de síntesis de la primera cadena que está cebada por un cebador modificado con un resto de captura hibridado y que usa la molécula de ácido nucleico como molde;
- 10 (c) llevar a cabo una reacción de síntesis de la segunda cadena que usa un producto de la reacción de síntesis de la primera cadena como molde para generar un ácido nucleico bicatenario que comprende un resto de captura;
- (d) ligar un ácido nucleico adaptador al ácido nucleico bicatenario para producir un producto de ligadura que comprende el resto de captura;
- 15 (e) capturar el producto de ligadura poniendo en contacto el producto de ligadura con un compañero de unión inmovilizado del resto de captura; y
- (f) amplificar el producto de ligadura capturado mediante la reacción en cadena de la polimerasa utilizando un cebador específico de la diana que comprende una porción en 3' que se aparea específicamente con la secuencia de nucleótidos diana y un primer cebador adaptador que se aparea específicamente con una secuencia complementaria del ácido nucleico adaptador, en donde el cebador específico de la diana comprende una parte de la cola en 5' que no se aparea específicamente con la secuencia de nucleótidos diana.
- 20

2. El método de la reivindicación 1, que comprende además:

- 25 (g) amplificar un producto de amplificación de la etapa (f) mediante reacción en cadena de la polimerasa utilizando un cebador de cola que comprende una porción en 3' que se aparea específicamente con una secuencia complementaria de la porción de cola en 5' del cebador específico de la diana y un segundo cebador adaptador que se aparea específicamente con una secuencia complementaria del ácido nucleico adaptador, en donde el cebador de cola comprende una porción en 5' que no se aparea específicamente con una secuencia complementaria del cebador específico de la diana, opcionalmente en donde la porción en 5' del cebador de cola comprende al menos una de una región de índice de muestra, una región de código de barras molecular y una región de sitio de cebador de secuenciación;
- 30

comprendiendo opcionalmente el método:

- 35 (h) inmovilizar un producto de amplificación de la etapa (g) sobre un sustrato paramagnético;
- (i) lavar el producto de amplificación inmovilizado; y
- (j) liberar el producto de amplificación inmovilizado lavado del sustrato paramagnético.

3. El método de las reivindicaciones 1 o 2, en el que la etapa (d) comprende:

- 40 i) combinar el ácido nucleico adaptador, el ácido nucleico bicatenario y una ligasa en condiciones en las que la ligasa liga el ácido nucleico adaptador al ácido nucleico bicatenario, en donde el ácido nucleico adaptador que se combina con el ácido nucleico bicatenario comprende una porción dúplex y una secuencia saliente, en donde la secuencia saliente comprende una secuencia de nucleótidos que es complementaria a una secuencia saliente en el extremo 3' del ácido nucleico bicatenario
- 45 o
- ii) combinar el ácido nucleico adaptador, el ácido nucleico bicatenario y una ligasa en condiciones en las que la ligasa liga el ácido nucleico adaptador al ácido nucleico bicatenario, en donde el ácido nucleico adaptador que se combina con el ácido nucleico bicatenario es monocatenario;
- 50 y/o
- iii) ligar el ácido nucleico bicatenario con el ácido nucleico adaptador en presencia de un agente de agregación; y/o en donde el método comprende además una etapa de lavado después de la etapa (d) y antes de la etapa (e).

4. El método de una cualquiera de las reivindicaciones anteriores, en el que el cebador modificado con el resto de captura comprende un primer grupo de acoplamiento químico configurado para unirse a un segundo grupo de acoplamiento químico unido a un resto de captura.

5. El método de una cualquiera de las reivindicaciones anteriores, en el que el resto de captura es un resto de biotina, opcionalmente en donde el resto de biotina comprende biotina-trietilenglicol, bis-biotina, biotina fotoescindible, destiobiotina, destiobiotina-trietilenglicol o azida de biotina; y/o en donde el compañero de unión inmovilizado es estreptavidina unida a un sustrato, opcionalmente en donde el sustrato comprende una superficie sólida que opcionalmente comprende una perla paramagnética.

6. El método de una cualquiera de las reivindicaciones anteriores que además comprende, después de la etapa (e) y antes de la etapa (f):

- i) inmovilizar el ácido nucleico bicatenario, que comprende el resto de captura, sobre un sustrato paramagnético;
y
- ii) lavar el ácido nucleico bicatenario inmovilizado; y, opcionalmente
- iii) liberar el ácido nucleico bicatenario inmovilizado lavado del sustrato paramagnético.

7. El método de una cualquiera de las reivindicaciones anteriores que además comprende, después de la etapa (c) y antes de la etapa (d):

- i) fosforilar en 5' el ácido nucleico bicatenario; y/o
- ii) reparar los extremos del ácido nucleico bicatenario para producir un ácido nucleico bicatenario de extremos romos.

8. El método de la reivindicación 7 (ii), que comprende además la adición de uno o más nucleótidos a un extremo 3' del ácido nucleico bicatenario de extremos romos, opcionalmente en donde:

- i) el uno o más nucleótidos comprenden nucleótidos de desoxiadenosina; o
- ii) el ácido nucleico adaptador comprende una secuencia de nucleótidos en un extremo 3' que comprende uno o más nucleótidos complementarios al uno o más nucleótidos añadidos al extremo 3' del ácido nucleico bicatenario de extremos romos,

opcionalmente en donde la secuencia de nucleótidos en el extremo 3' del ácido nucleico adaptador comprende uno o más nucleótidos de desoxitimidina, o en donde el ácido nucleico adaptador comprende además una cadena de bloqueo apareada con una cadena de amplificación que comprende la secuencia de nucleótidos en el extremo 3' y en donde la secuencia de nucleótidos en el extremo 3' no está emparejada de manera que forma una secuencia saliente.

9. El método de la reivindicación 7 (ii), que comprende adicionalmente, después de la reparación de los extremos:

- i) inmovilizar el ácido nucleico bicatenario, que comprende el resto de captura, sobre un sustrato paramagnético;
- ii) lavar el ácido nucleico bicatenario inmovilizado; y
- iii) liberar el ácido nucleico bicatenario inmovilizado lavado del sustrato paramagnético;
- o
- i) inmovilizar el ácido nucleico bicatenario, que comprende el resto de captura, sobre un sustrato paramagnético;
- y
- ii) lavar el ácido nucleico bicatenario inmovilizado.

10. El método de una cualquiera de las reivindicaciones anteriores, en el que la molécula de ácido nucleico se obtiene de una muestra que comprende un linfocito T, un linfocito B o una mezcla de los mismos, opcionalmente en donde la muestra se obtiene de un sujeto que tiene, o se sospecha que tiene, una neoplasia maligna de linfocitos T o una neoplasia maligna de linfocitos B.

11. El método de una cualquiera de las reivindicaciones anteriores, en el que la secuencia de nucleótidos diana comprende una secuencia de nucleótidos correspondiente a una porción de un gen del receptor de linfocitos T (TCR) o un gen del receptor de linfocitos B (BCR); y/o en donde el cebador modificado con el resto de captura comprende una secuencia de nucleótidos que es complementaria a un gen de receptor inmunitario o un gen de inmunoglobulina.

12. El método de la reivindicación 11, en el que el cebador modificado con el resto de captura se aparea específicamente con una región constante que está corriente abajo de una región determinante de complementariedad 3 (CDR3), opcionalmente en donde el cebador específico de la diana se aparea específicamente con una región constante o un segmento J que está aguas abajo de una CDR3.

13. Un método para determinar un repertorio inmunitario en una muestra, comprendiendo el método:

- (a) obtener una muestra que comprende una molécula de ácido nucleico que codifica al menos uno de un receptor inmunitario y una inmunoglobulina;
- (b) poner en contacto la molécula de ácido nucleico de la muestra con un primer cebador específico de la diana que se aparea específicamente con una secuencia de nucleótidos diana de la molécula de ácido nucleico en condiciones de hibridación, en donde el primer cebador específico de la diana comprende un resto de captura;
- (c) llevar a cabo una reacción de síntesis de la primera cadena que está cebada por un primer cebador específico de la diana hibridado y que usa la molécula de ácido nucleico como molde;
- (d) llevar a cabo una reacción de síntesis de la segunda cadena que usa un producto de la reacción de síntesis de la primera cadena como molde para generar un ácido nucleico bicatenario que comprende el resto de captura;
- (e) ligar un ácido nucleico adaptador al ácido nucleico bicatenario para producir un producto de ligadura;
- (f) capturar el producto de ligadura poniendo en contacto el producto de ligadura con un compañero de unión inmovilizado del resto de captura;
- (g) amplificar el producto de ligadura capturado mediante la reacción en cadena de la polimerasa utilizando un segundo cebador específico de la diana que comprende una porción en 3' que se aparea específicamente con la

secuencia de nucleótidos diana y un primer cebador adaptador que se aparee específicamente con una secuencia complementaria del ácido nucleico adaptador, en donde el segundo cebador específico de la diana comprende una parte de la cola en 5' que no se aparee específicamente con la secuencia de nucleótidos diana;

- 5 (h) amplificar un producto de amplificación de la etapa (g) mediante reacción en cadena de la polimerasa utilizando un cebador de cola que comprende una porción en 3' que se aparee específicamente con una secuencia complementaria de la porción de cola en 5' del segundo cebador específico de la diana y un segundo cebador adaptador que se aparee específicamente con una secuencia complementaria del ácido nucleico adaptador, en donde el cebador de cola comprende una porción en 5' que no se aparee específicamente con una secuencia complementaria del segundo cebador específico de la diana; y
- 10 (i) secuenciar un producto de amplificación de la etapa (h) utilizando un primer y un segundo cebador de secuenciación.

14. El método de la reivindicación 13, en el que el receptor inmunitario comprende un TCR y en donde la inmunoglobulina comprende un BCR; y/o
- 15 en donde la secuencia de nucleótidos diana corresponde a una secuencia genéticamente recombinada.

15. El método de la reivindicación 13, en el que la muestra comprende un linfocito T, un linfocito B o una mezcla de los mismos, opcionalmente en donde la muestra se obtiene de un sujeto humano.

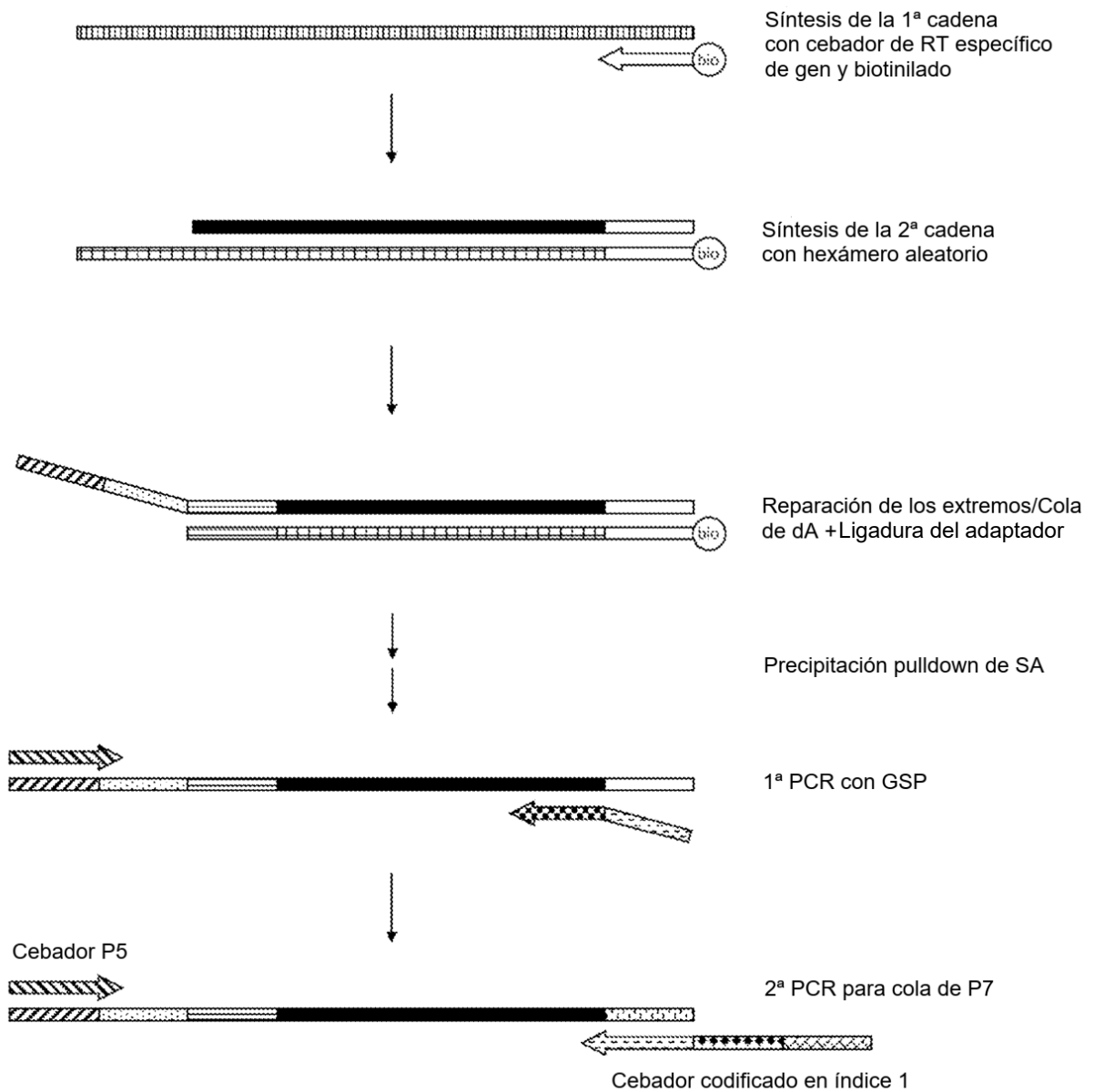


FIG. 1

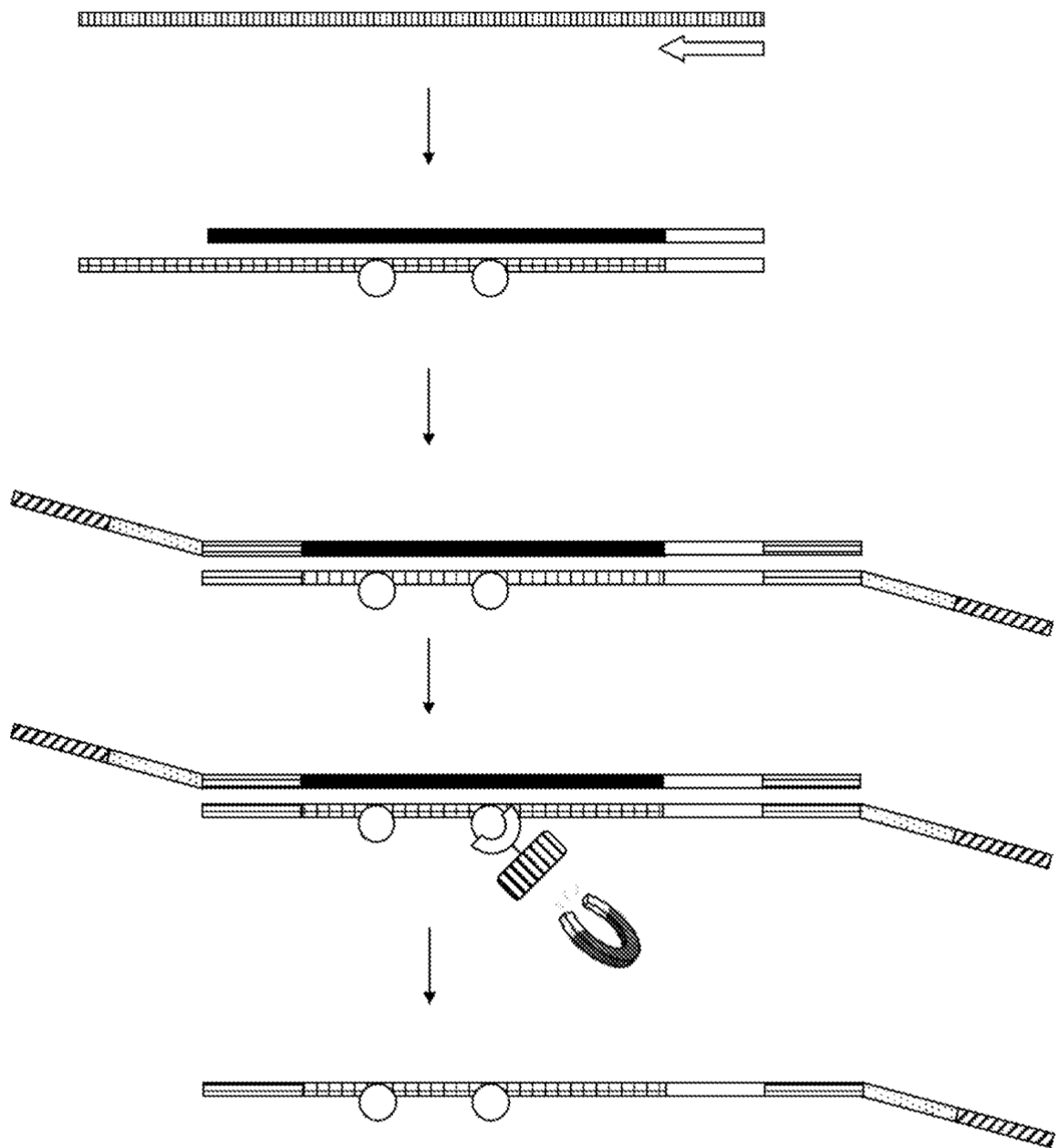


FIG. 2

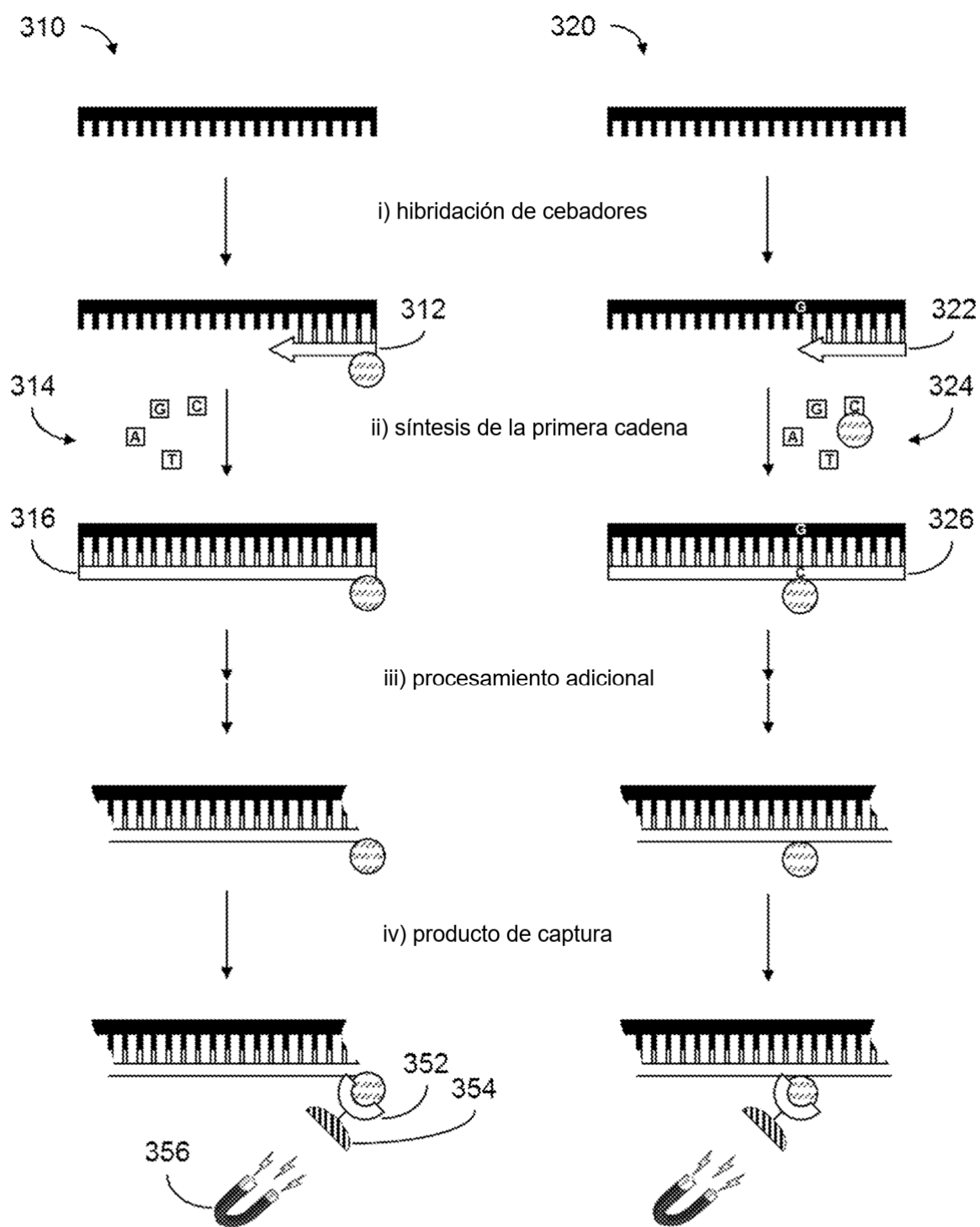


FIG. 3

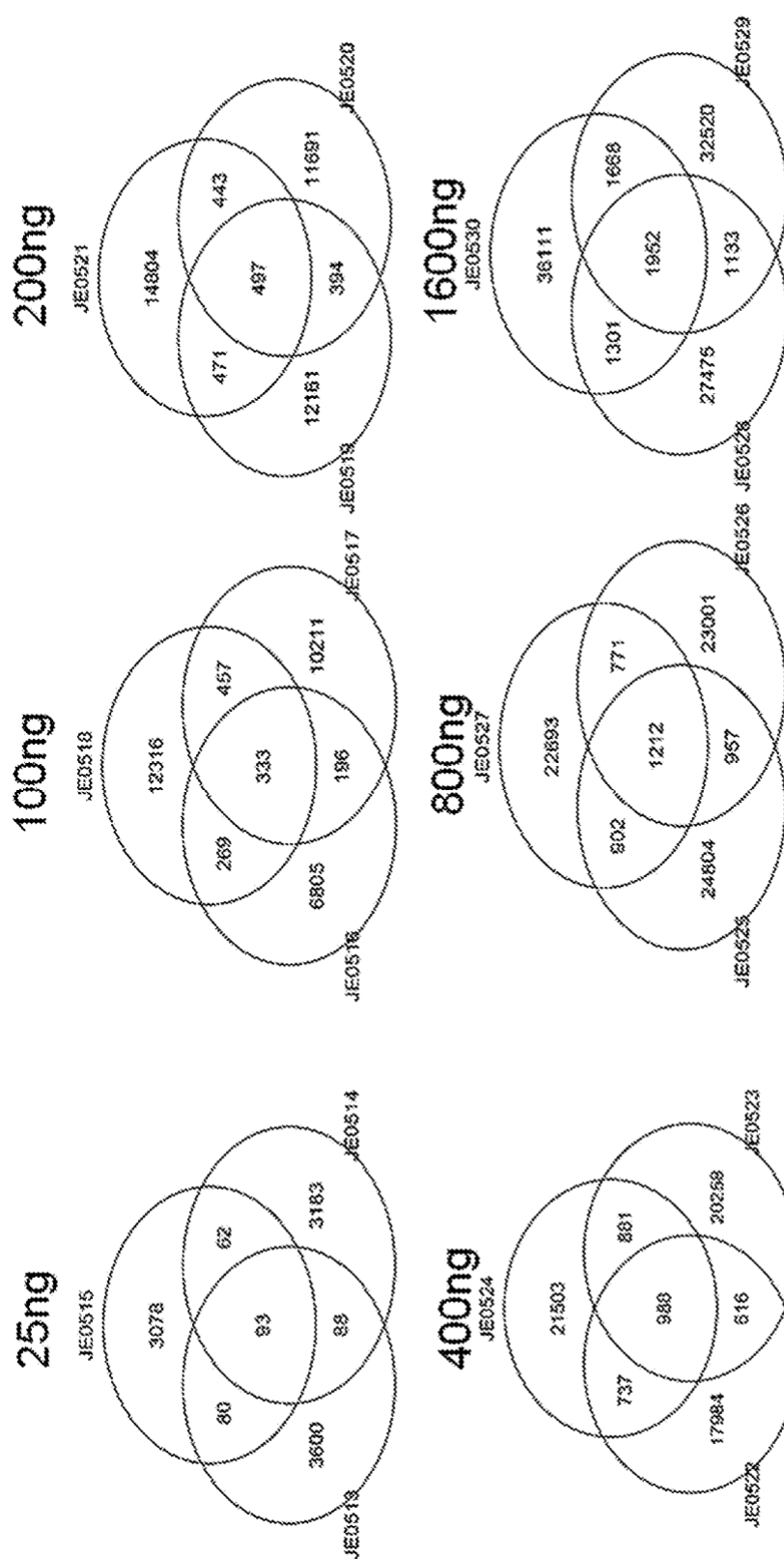


FIG. 4

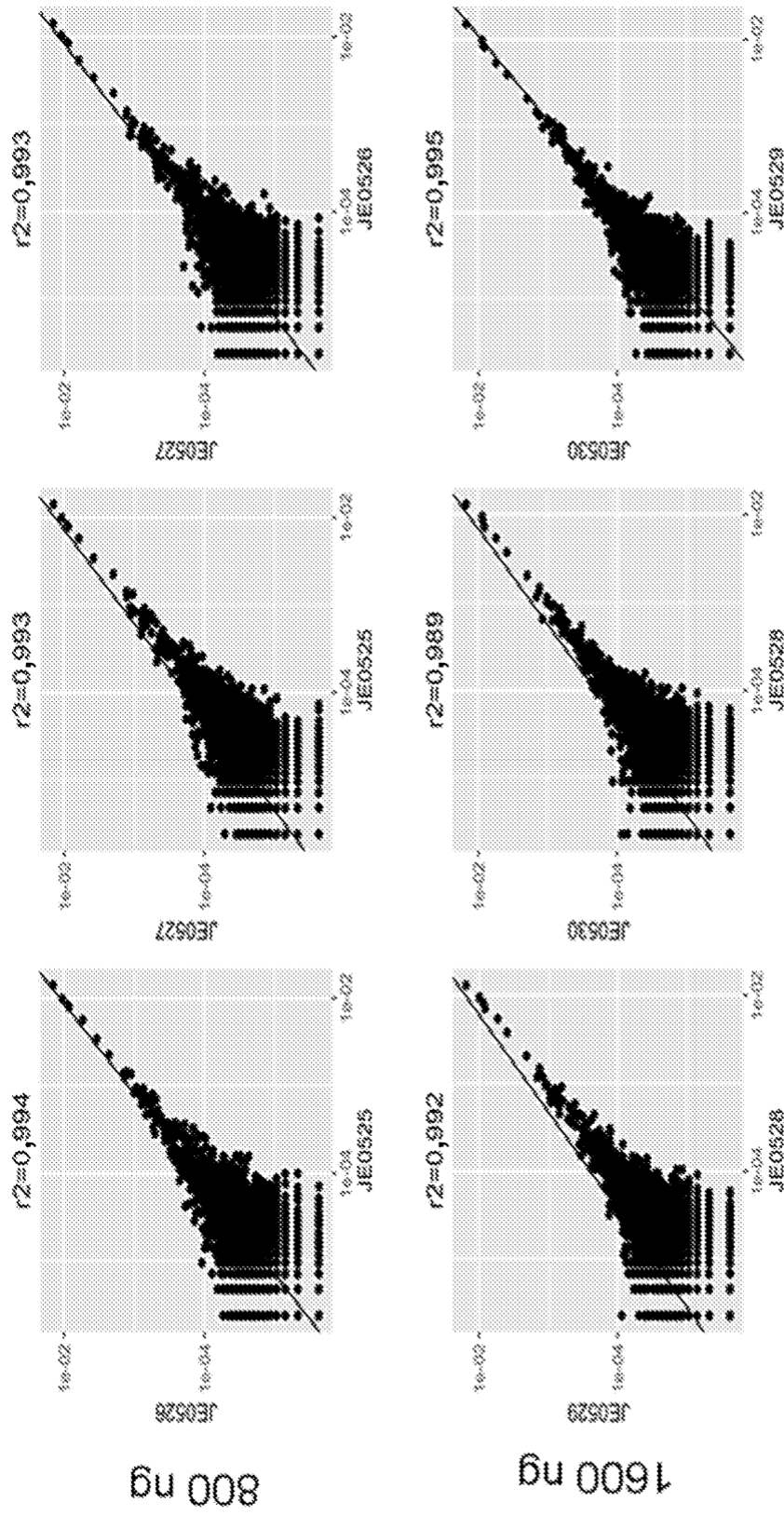


FIG. 5

Todos los clones (unión de ambas muestras)

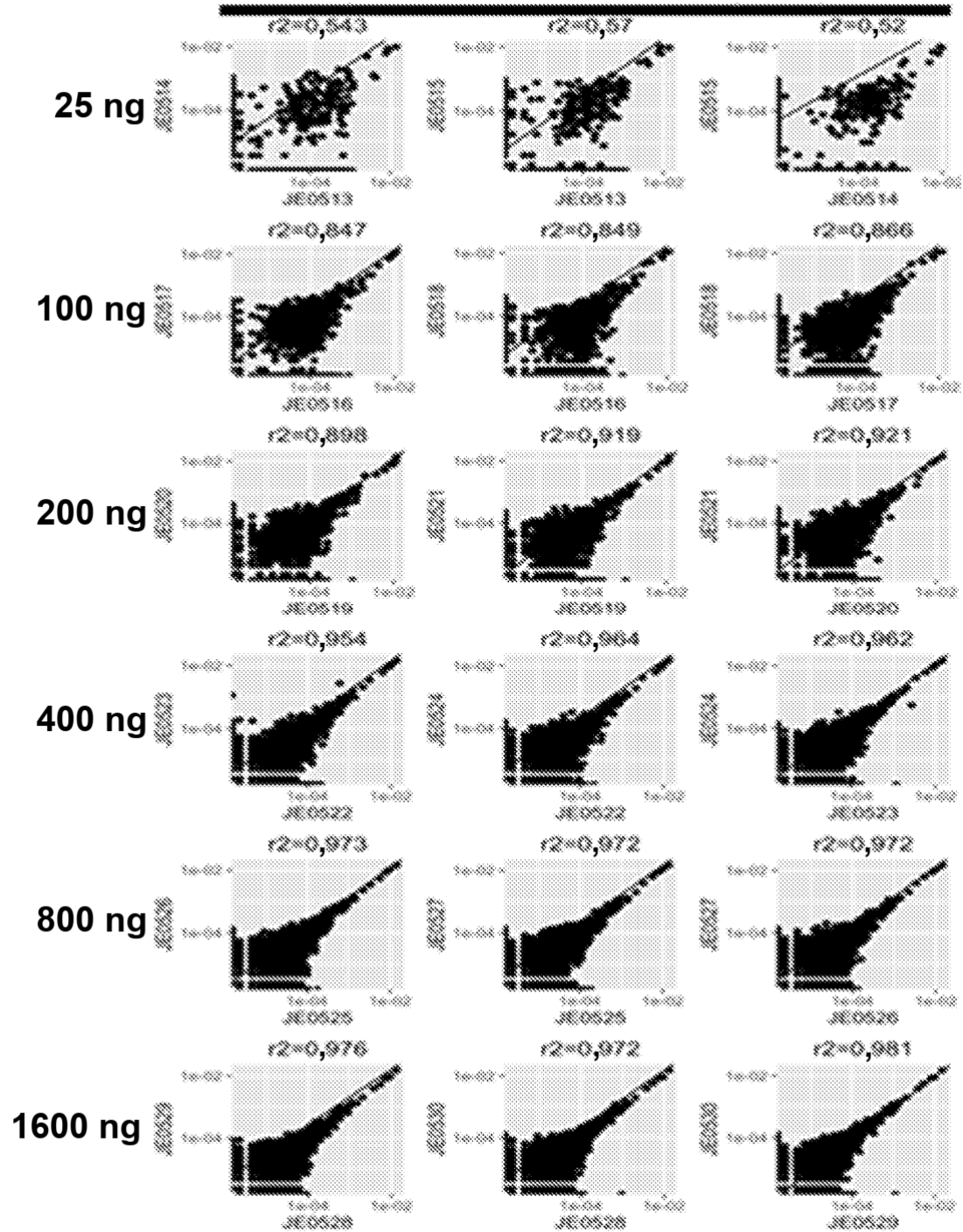


FIG. 6A

Clones solapantes (intersección de ambas muestras)

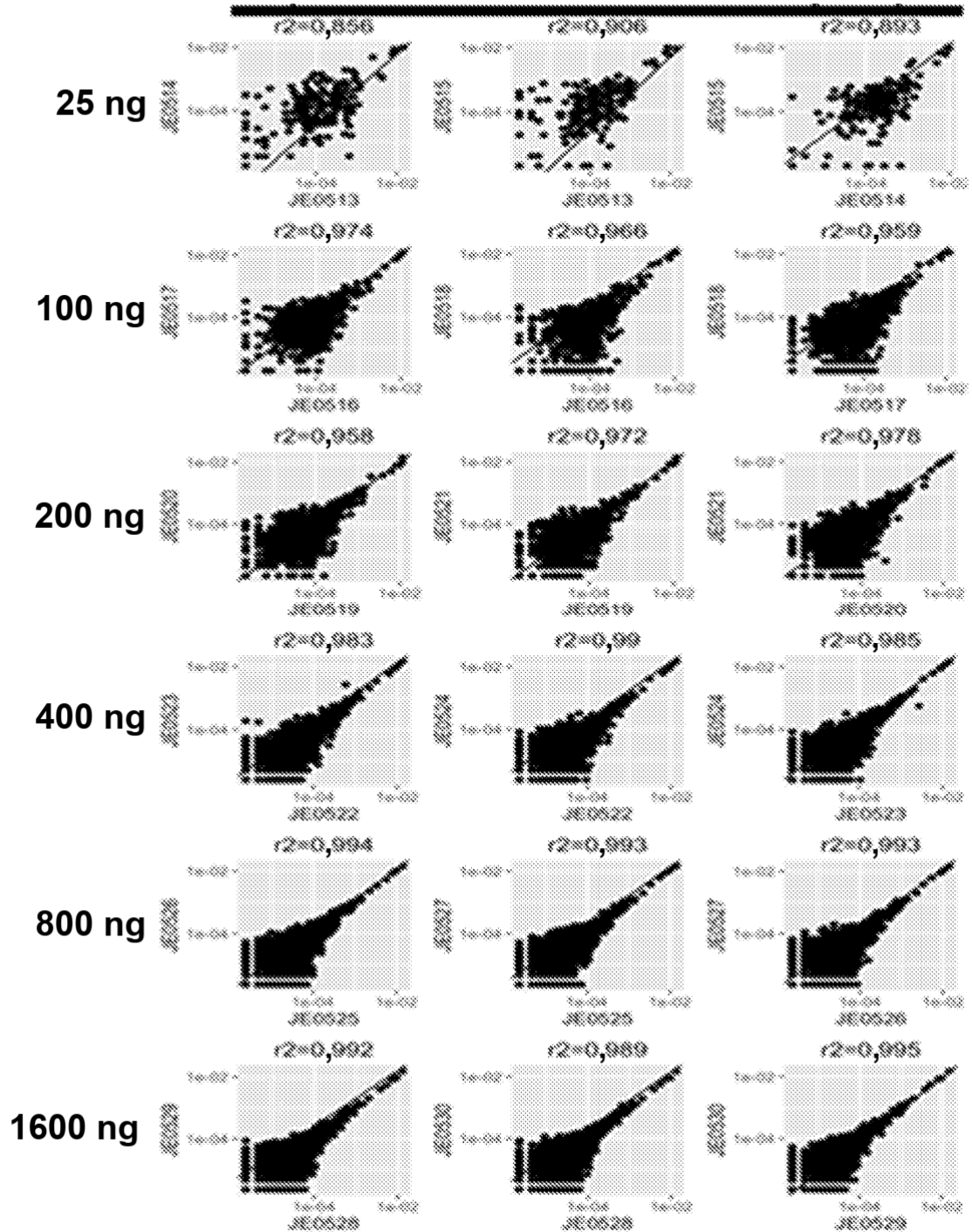


FIG. 6B

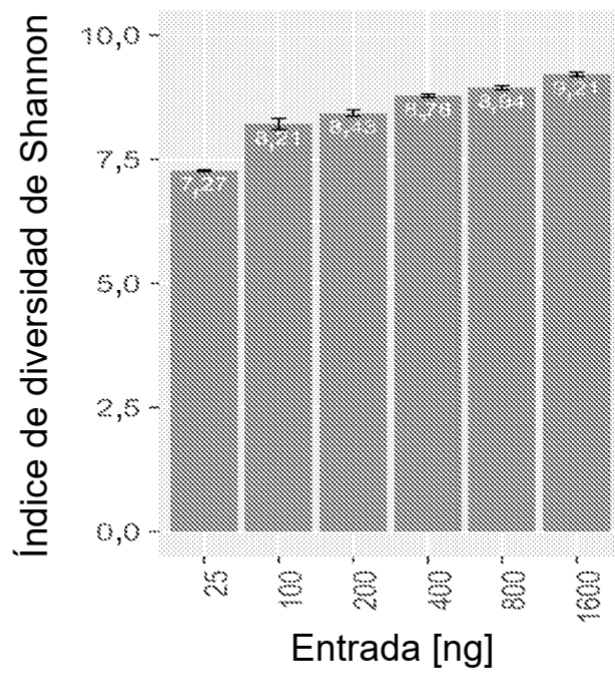
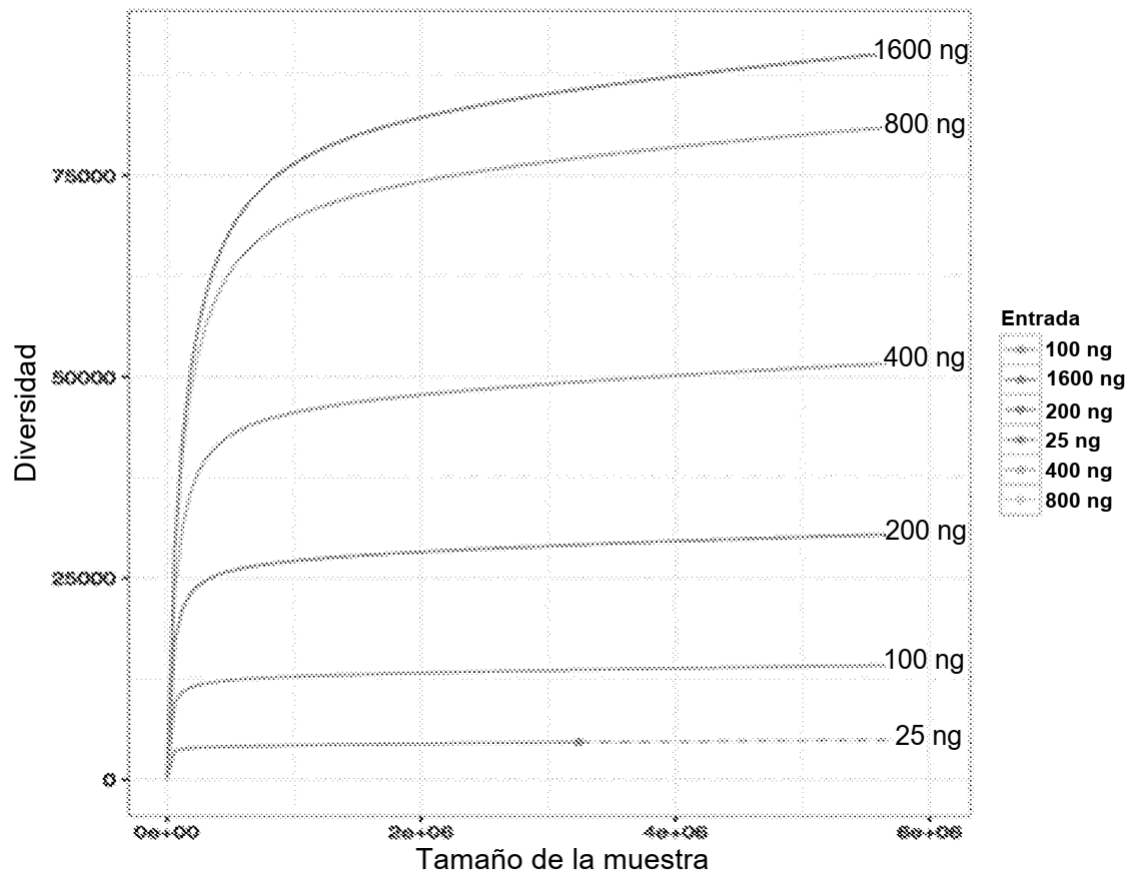


FIG. 7

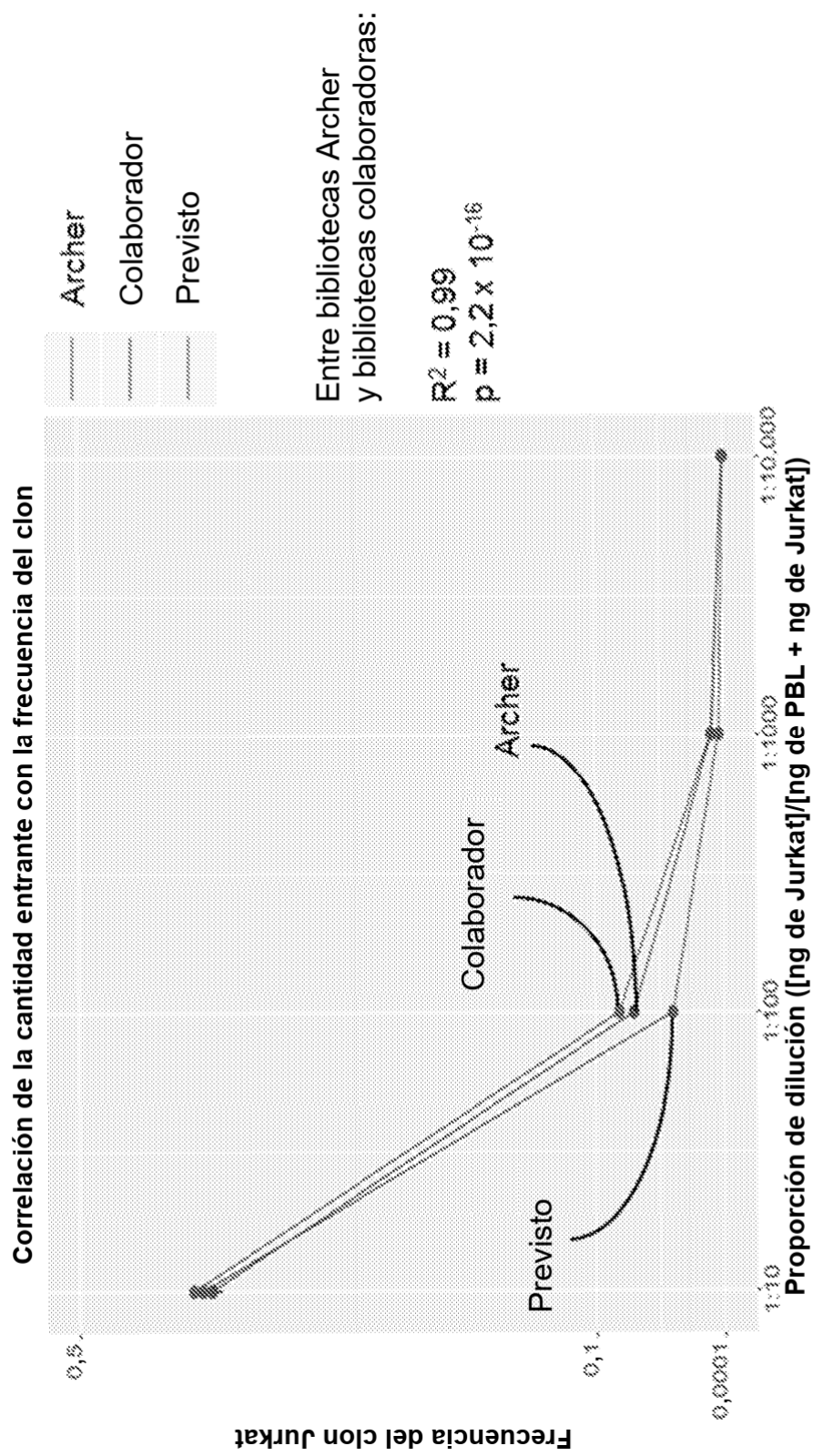


FIG. 8

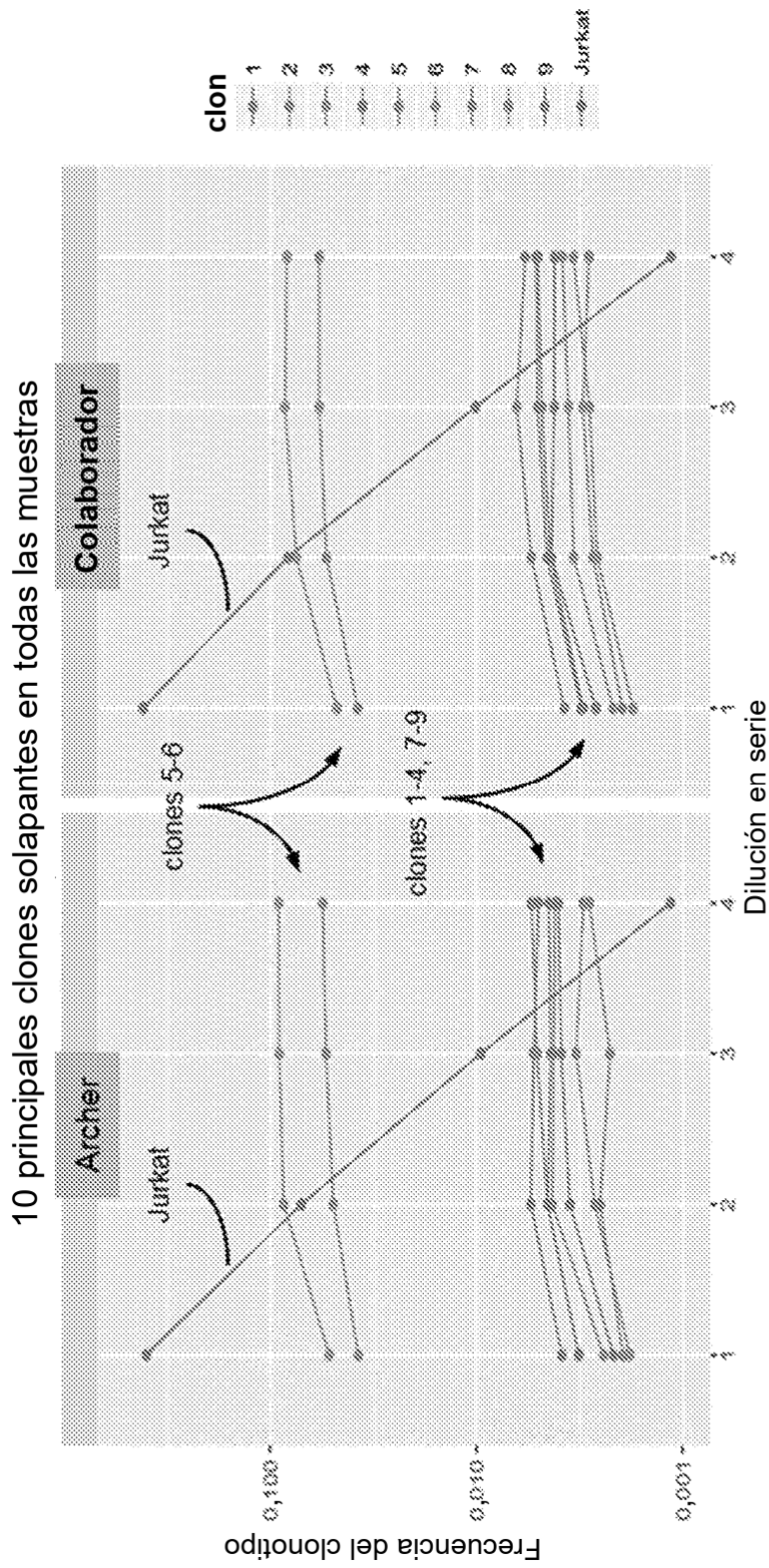


FIG. 9

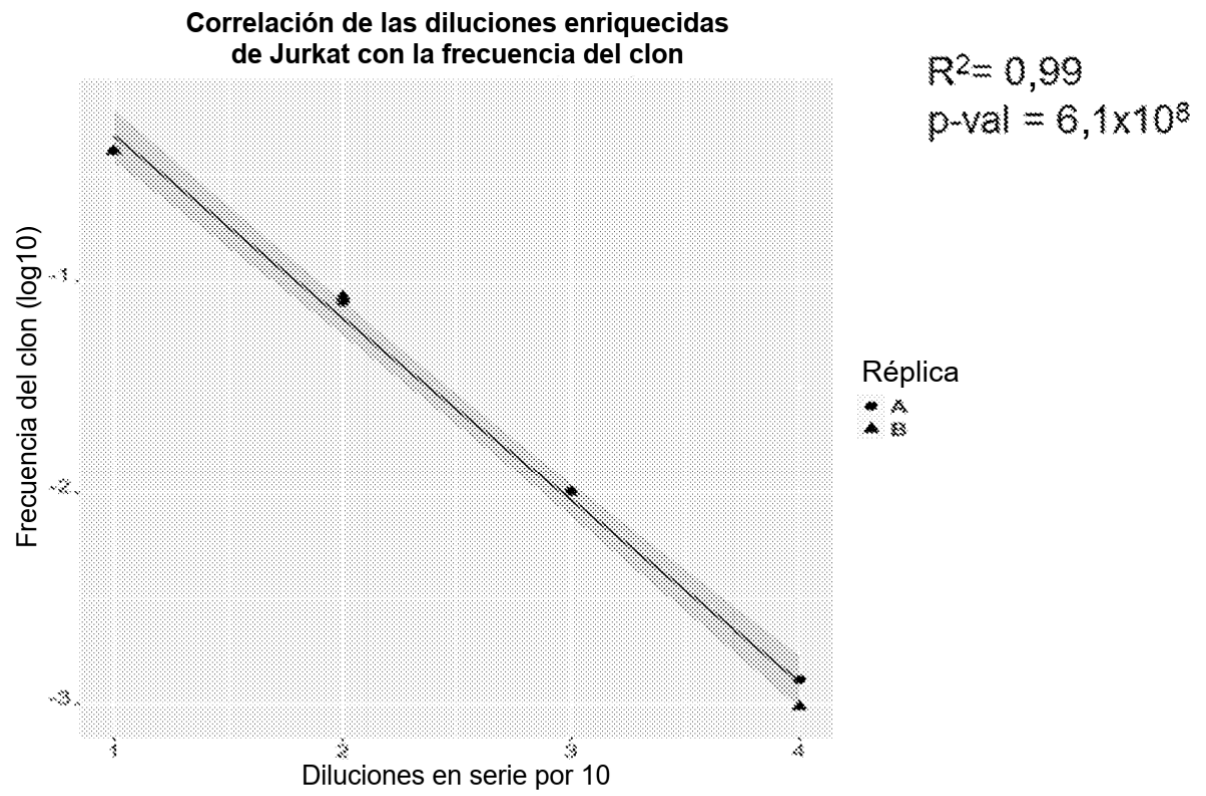


FIG. 10

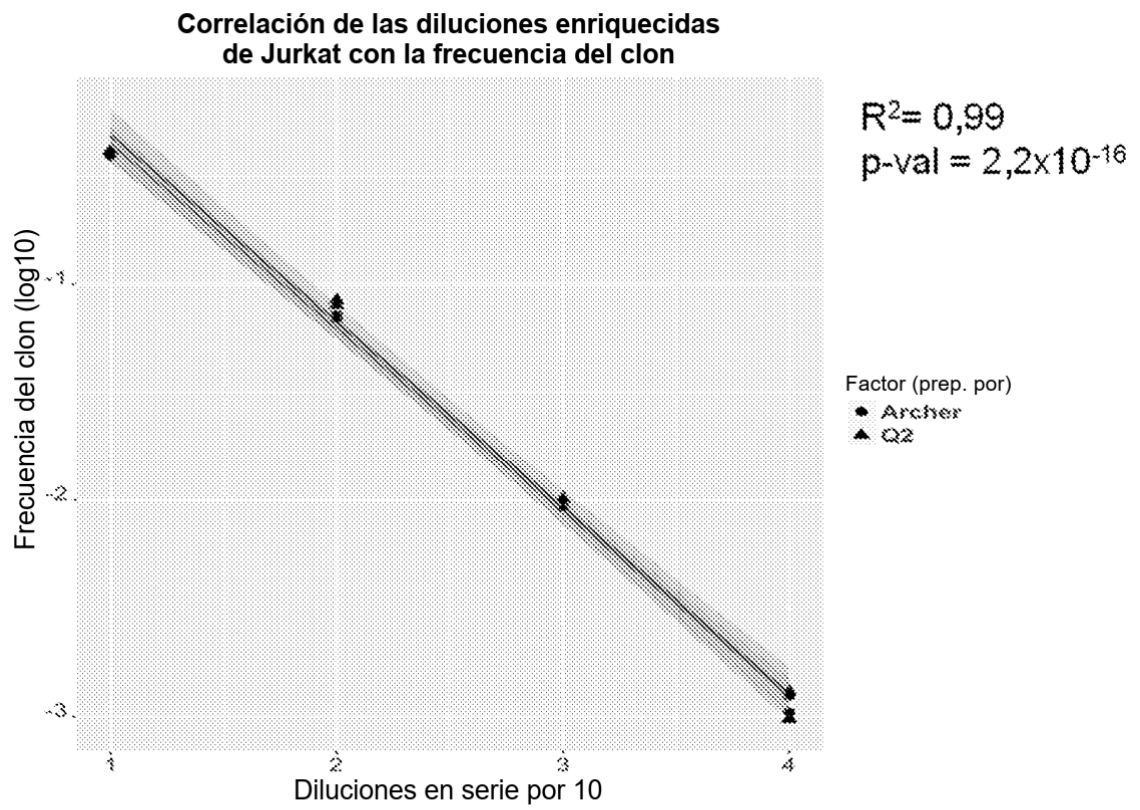


FIG. 11

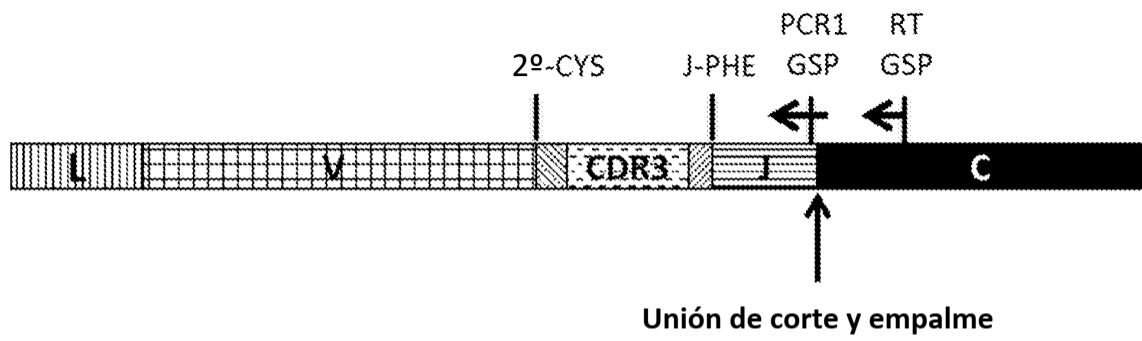


FIG. 12

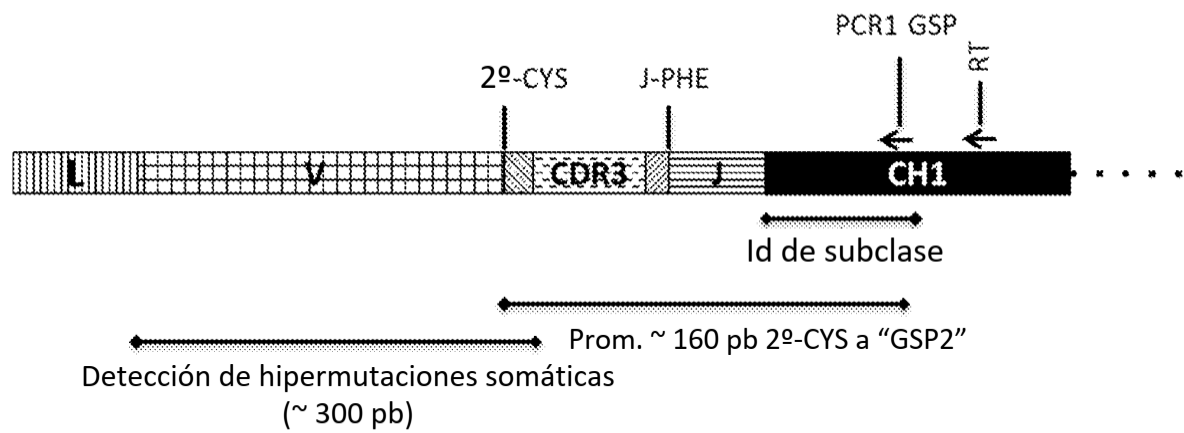


FIG. 13

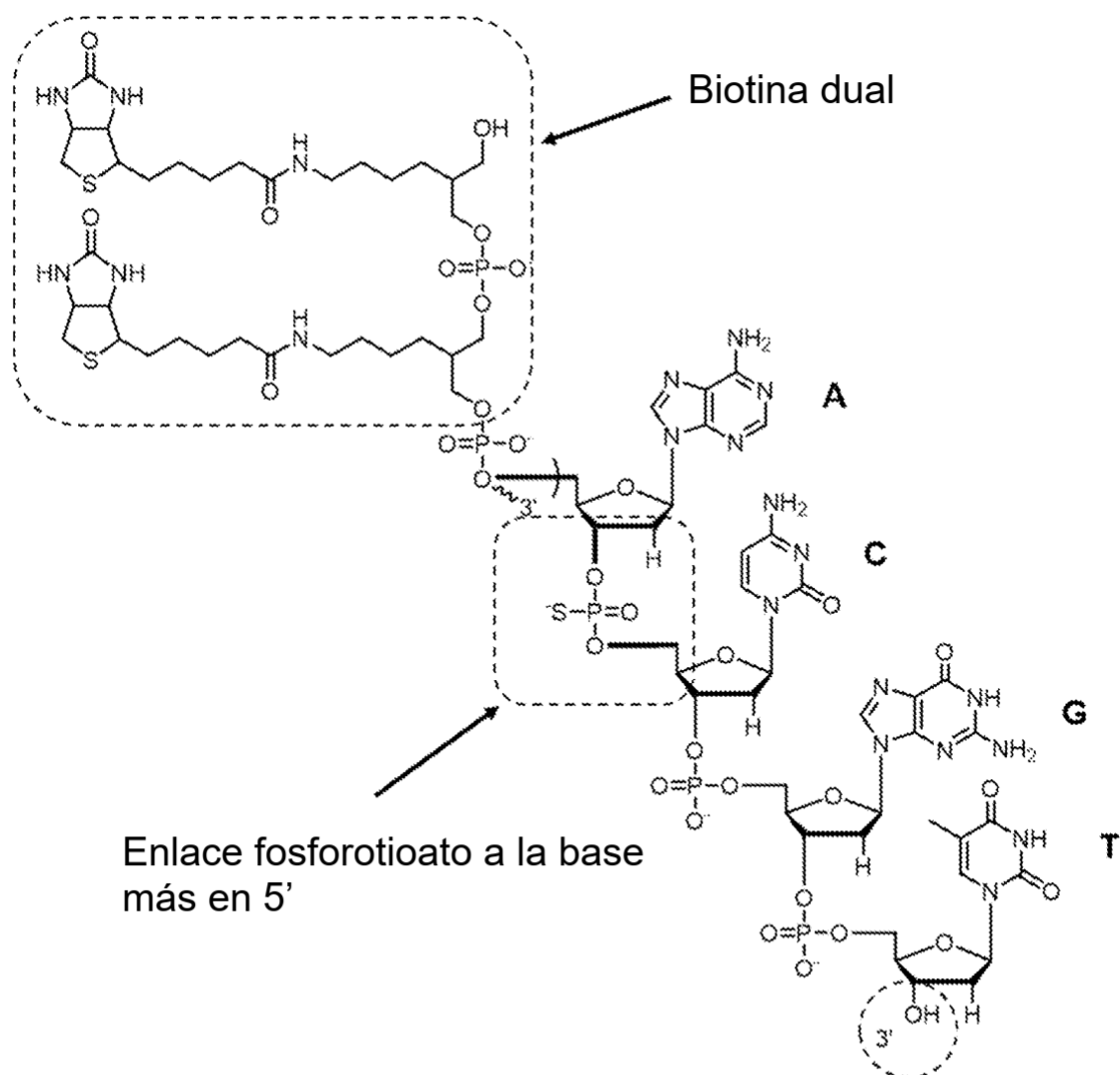


FIG. 14