

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 706 531**

51 Int. Cl.:

**C12Q 1/68** (2008.01)

**C12N 15/10** (2006.01)

**C12N 15/63** (2006.01)

**C12N 15/90** (2006.01)

12

## TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **10.11.2015 PCT/US2015/059959**

87 Fecha y número de publicación internacional: **19.05.2016 WO16077350**

96 Fecha de presentación y número de la solicitud europea: **10.11.2015 E 15797569 (9)**

97 Fecha y número de publicación de la concesión europea: **31.10.2018 EP 3218513**

54 Título: **Amplificación de polinucleótidos empleando sistemas CRISPR-Cas**

30 Prioridad:

**11.11.2014 US 201462078355 P**

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

**29.03.2019**

73 Titular/es:

**ILLUMINA, INC. (100.0%)  
5200 Illumina Way  
San Diego, CA 92122, US**

72 Inventor/es:

**MANDELL, JEFFREY G.**

74 Agente/Representante:

**ELZABURU, S.L.P**

**ES 2 706 531 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

Amplificación de polinucleótidos empleando sistemas CRISPR-Cas

La presente invención se refiere generalmente a métodos para amplificar polinucleótidos, y más específicamente a métodos para amplificar polinucleótidos y a aplicaciones de los mismos empleando sistemas CRISPR-Cas.

5 **Antecedentes**

La amplificación de ácidos nucleicos es una etapa clave de muchos ácidos nucleicos basada en procesos tales como la secuenciación de ácidos nucleicos. Los métodos de amplificación de ácidos nucleicos más utilizados actualmente, *por ejemplo*, aquellos empleados en la creación de un grupo en la secuenciación de última generación requieren tanto de un ciclo de temperatura como de un intercambio de fluidos. Por otro lado, la amplificación isotérmica puede ser eficaz tanto en tiempo como en energía mediante la eliminación de la rampa de temperatura y de los tiempos de equilibrio. Se han desarrollado varios métodos de amplificación isotérmica, *por ejemplo*, la amplificación con polimerasa y recombinasa (RPA, de sus siglas en inglés) basados en la amplificación isotérmica. Estos sistemas de amplificación isotérmica normalmente carecen de la velocidad y la eficacia deseadas adecuadas de una manera ideal para algunas aplicaciones. Además, algunos sistemas requieren de enzimas y reactivos adicionales que incluyen ATP. Por tanto, existe una necesidad en la técnica de métodos isotérmicos de amplificación de ácidos nucleicos que sean fáciles, rápidos, y eficaces. La presente divulgación satisface esta necesidad proporcionando métodos para amplificar ácidos nucleicos empleando sistemas CRISPR-Cas. También se proporcionan ventajas relacionadas.

Las Repeticiones Palindrómicas Cortas Agrupadas y Regularmente interespaciadas (CRISPRs, de sus siglas en inglés) están implicadas en una ruta de interferencia que protege a las células de los bacteriófagos y los plásmidos conjugativos en muchas bacterias y arqueas (Marraffini y Sontheimer, 2010, *Nat Rev Genet.* 11(3):181-190). Las CRISPR consisten en matrices de secuencias de repetición cortas interespaciadas mediante secuencias de ADN variables únicas de tamaño similar llamadas espaciadores, que a menudo se originan a partir de ADN de un fago o plásmido (Barrangou *et al.*, 2007, *Science* 315:1709-12; Bolotin *et al.*, 2005, *Microbiology* 151:2551-61; Mojica *et al.*, 2005, *J Mol Evol* 60:174-82). Por tanto, las secuencias CRISPR proporcionan un registro heredable de infecciones pasadas y se pueden transcribir en los ARNs CRISPR (ARNcrs) – ARNs pequeños que se dirigen a ácidos nucleicos invasivos (Marraffini y Sontheimer, 2010, *Nat Rev Genet.* 11(3):181-190). Las CRISPRs a menudo se asocian a genes asociados a CRISPR (Cas) que codifican para proteínas relacionadas a CRISPRs. Las proteínas Cas pueden proporcionar mecanismos para destruir ácidos nucleicos extraños invasores dirigidos por ARNcrs. Las CRISPR junto con los genes Cas (CRISPR-asociado) comprenden un sistema inmunológico adaptativo que proporciona resistencia adquirida contra ácidos nucleicos extraños invasores en bacterias y arqueas (Barrangou *et al.*, 2007, *Science* 315:1709-12).

**Compendio**

La presente divulgación proporciona métodos para amplificar polinucleótidos, y más específicamente métodos y aplicaciones de los mismos para amplificar una secuencia de ADN diana empleando sistemas CRISPR-Cas.

En un aspecto, en la presente memoria se proporciona un método para amplificar un ácido nucleico bicatenario diana que incluye: (a) proporcionar un sistema que tiene: un ARN (ARNcr) con repeticiones palindrómicas cortas agrupadas y regularmente interespaciadas (CRISPRs) o un derivado del mismo, y una proteína asociada a CRISPR (Cas) o una variante de la misma, en donde el ARNcr o el derivado del mismo contiene una región de nucleótidos específica de la diana complementaria a una región en una primera cadena del ácido nucleico bicatenario diana; (b) poner en contacto el ácido nucleico bicatenario diana con el sistema para formar un complejo; (c) hibridar un cebador a una segunda cadena del ácido nucleico bicatenario diana, el cebador contiene una secuencia complementaria para una región de la segunda cadena del ácido nucleico bicatenario diana, y (d) prolongar un ácido nucleico complementario para la segunda cadena del ácido nucleico bicatenario diana a partir del cebador empleando una polimerasa.

En algunas realizaciones, el método proporcionado en la presente memoria incluye además repetir de la etapa (c) a la etapa (d) durante una o dos veces, *por ejemplo*, hasta que se alcanza un grado de amplificación deseado. En algunas realizaciones, el método proporcionado en la presente memoria incluye además repetir de la etapa (a) a la etapa (d) hasta que se alcanza un grado de amplificación deseado.

En algunas realizaciones, el ácido nucleico diana proporcionado en la presente memoria es un ADN bicatenario (ADNs). En algunas realizaciones, el ácido nucleico diana es un ARN bicatenario (ARNds).

En algunas realizaciones, el sistema es un sistema CRISPR-Cas de Tipo I o un derivado del mismo. En algunas realizaciones, el sistema es un sistema CRISPR-Cas Tipo II o un derivado del mismo. En algunas realizaciones, el sistema es un sistema CRISPR-Cas Tipo III o un derivado del mismo.

En algunas realizaciones el sistema comprende además un ARNcr trans-activador (ARNtracr) o un derivado del mismo. En algunas realizaciones, el ARNcr o el derivado del mismo es un polinucleótido que comprende un polinucleótido ARNcr fusionado a un polinucleótido ARNtracr.

5 En algunas realizaciones, la primera cadena del ácido nucleico bicatenario diana contiene una secuencia complementaria a un motivo adyacente de protoespaciador (PAM) 5'-NGG.

En algunas realizaciones, la primera cadena del ácido nucleico bicatenario diana contiene una secuencia universal, y en donde el ARNcr o el derivado del mismo contiene una secuencia complementaria para una región de la secuencia universal. En algunas realizaciones, el cebador contiene una secuencia de una región de la secuencia universal.

10 En algunas realizaciones, la secuencia universal tiene una secuencia de SEQ ID No.3. En algunas realizaciones, el ARNcr contiene una secuencia de SEQ ID No.7. En algunas realizaciones, la secuencia universal tiene una secuencia de SEQ ID No.5.

15 En algunas realizaciones, la secuencia universal tiene una secuencia de SEQ ID No.4. En algunas realizaciones, el ARNcr contiene una secuencia de SEQ ID No.8. En algunas realizaciones, la secuencia universal tiene una secuencia de SEQ ID No.6.

20 En algunas realizaciones, la proteína Cas o la variante de la misma es una proteína Cas9 o una variante de la misma. En algunas realizaciones, la proteína Cas9 contiene dos dominios nucleasa inactivados. En algunas realizaciones, los dos dominios de nucleasa inactivados comprenden una primera mutación en el dominio que separa la cadena complementaria para el ARNcr y una segunda mutación en el dominio que separa la cadena no complementaria para el ARNcr. En algunas realizaciones, la primera mutación es D10A y la segunda mutación es H840A. En algunas realizaciones, la proteína Cas o la variante de la misma es una proteína en Cascada o una variante de la misma. En algunas realizaciones, la proteína Cas o la variante de la misma es una proteína Cas3 o una variante de la misma.

25 En algunas realizaciones, la polimerasa es una polimerasa que desplaza la cadena. En algunas realizaciones, la polimerasa se selecciona de un grupo que consiste en Bst, Bsu, y Phi29.

30 En algunas realizaciones, el método proporcionado en la presente memoria comprende además: aplicar al menos una transposasa y al menos una composición de extremo del transposón que contiene una cadena transferida a una muestra que contiene un ácido nucleico diana bajo condiciones donde el ácido nucleico diana y la composición de extremo del transposón sufren una reacción de transposición para generar una mezcla, en donde el ácido nucleico diana se fragmenta para generar una pluralidad de fragmentos de ácido nucleico diana, e incorporan una secuencia cebadora universal en cada uno de la pluralidad de fragmentos del ácido nucleico diana, en donde el ARNcr o el derivado del mismo contiene una región de nucleótidos específica a la diana complementaria para una región del cebador universal.

35 En algunas realizaciones, el cebador universal se incorpora en una pluralidad de fragmentos de ácido nucleico diana mediante una reacción PCR. En algunas realizaciones, el cebador universal tiene una secuencia complementaria a un motivo adyacente de protoespaciador (PAM) 5'-NGG.

En algunas realizaciones, la secuencia universal tiene una secuencia de SEQ ID No.3. En algunas realizaciones, el ARNcr contiene una secuencia de SEQ ID No.7. En algunas realizaciones, la secuencia universal tiene una secuencia de SEQ ID No.5.

40 En algunas realizaciones, el cebador universal tiene una secuencia de SEQ ID No.4. En algunas realizaciones, el ARNcr contiene una secuencia de SEQ ID No.8. En algunas realizaciones, la secuencia universal tiene una secuencia de SEQ ID No.6.

45 En algunas realizaciones, se incorporan dos cebadores universales a los dos extremos de cada uno de la pluralidad de fragmentos de ácido nucleico diana. En algunas realizaciones, los dos cebadores universales tienen secuencias de SEQ ID No.3 y SEQ ID No.4. En algunas realizaciones, los dos cebadores universales tienen secuencias de SEQ ID No.5 y SEQ ID No.6.

En algunas realizaciones, el ácido nucleico bicatenario diana se amplifica de manera lineal. En algunas realizaciones, el ácido nucleico bicatenario diana se amplifica exponencialmente.

50 En algunas realizaciones, en la presente memoria se proporciona un método para amplificar un ácido nucleico bicatenario diana que comprende: (a) proporcionar un primer sistema que tiene: un primer ARN (ARNcr) con repeticiones palindrómicas cortas agrupadas y regularmente interespaciadas (CRISPS) o un derivado del mismo, y una primera proteína asociada a CRISPR (Cas) o una variante de la misma, en donde el primer ARNcr o el derivado del mismo contiene una región de nucleótidos específica de la diana complementaria para una región de una primera cadena del ácido nucleico bicatenario diana; (b) proporcionar un segundo sistema que tiene: un segundo ARN (ARNcr) con repeticiones palindrómicas cortas agrupadas y regularmente interespaciadas (CRISPS) o un

55

derivado del mismo, y una segunda proteína asociada a CRISPR (Cas) o una variante de la misma, en donde el segundo ARNcr o el derivado del mismo contiene una región de nucleótidos específica de la diana complementaria para una región de una segunda cadena del ácido nucleico bicatenario diana; (c) poner en contacto el ácido nucleico bicatenario diana con el primer sistema y el segundo sistema; (d) hibridar el primer sistema a una segunda cadena del ácido nucleico bicatenario diana, el primer cebador que contiene una secuencia complementaria para una región de la segunda cadena del ácido nucleico bicatenario diana, e hibridar un segundo cebador a una primera cadena del ácido nucleico bicatenario diana, el segundo cebador que contiene una secuencia complementaria para una región de la primera cadena del ácido nucleico bicatenario diana, y (e) prolongar el extremo 3' del primer cebador y el segundo cebador con una o más polimerasas para generar un primer y un segundo ácido nucleico diana bicatenario.

5 En algunas realizaciones, el método proporcionado en la presente memoria incluye además repetir de la etapa (a) a la etapa (e) durante una o dos veces, *por ejemplo*, hasta que se alcance el grado de amplificación deseado.

10 En algunas realizaciones, el ácido nucleico diana es un ADN bicatenario (ADNs). En algunas realizaciones, el ácido nucleico diana es un ARN bicatenario (ARNs).

15 En algunas realizaciones, el primer sistema o el segundo sistema es un sistema CRISPR-Cas de Tipo I o un derivado del mismo. En algunas realizaciones, el primer sistema o el segundo sistema es un sistema CRISPR-Cas Tipo II o un derivado del mismo. En algunas realizaciones, el primer sistema o el segundo sistema es un sistema CRISPR-Cas Tipo III o un derivado del mismo.

20 En algunas realizaciones el primer sistema o el segundo sistema comprende además un ARNcr trans-activador (ARNtracr) o un derivado del mismo. En algunas realizaciones, el ARNcr o el derivado del mismo del primer sistema o del segundo sistema, es un polinucleótido que comprende un polinucleótido ARNcr fusionado a un polinucleótido ARNtracr. En algunas realizaciones, la primera cadena y la segunda cadena del ácido nucleico bicatenario diana contiene una secuencia complementaria para un motivo adyacente de protoespaciador (PAM) 5'-NGG.

25 En algunas realizaciones, la primera cadena del ácido nucleico bicatenario diana contiene una primera secuencia universal, y en donde el ARNcr o el derivado del mismo del primer sistema contiene una secuencia complementaria para una región de la primera secuencia universal, y la segunda cadena del ácido nucleico bicatenario diana contiene una segunda secuencia universal, y en donde el crATN o el derivado del mismo del segundo sistema contiene una secuencia complementaria para una región de la segunda secuencia universal.

30 En algunas realizaciones, el primer cebador contiene una secuencia de una región de la primera secuencia universal, y el segundo cebador contiene una secuencia de una región de la segunda secuencia universal. En algunas realizaciones, la primera secuencia universal (que contiene un primer cebador) tiene una secuencia de SEQ ID No.3, el ARNcr o el derivado del mismo del primer sistema contiene una secuencia de SEQ ID No.7, y el primer cebador contiene una secuencia de SEQ ID No.5, y la segunda secuencia universal (que contiene un segundo cebador) tiene una secuencia de SEQ ID No.4, el ARNcr o el derivado del mismo del segundo sistema contiene una secuencia de SEQ ID No.8, y el segundo cebador contiene una secuencia de SEQ ID No.6.

35 En algunas realizaciones, la proteína Cas o la variante de la misma del primer sistema o el segundo sistema es una proteína Cas9 o una variante de la misma. En algunas realizaciones, la proteína Cas9 contiene dos dominios nucleasa inactivados. En algunas realizaciones, los dos dominios nucleasa inactivados comprenden una primera mutación en el dominio que separa la cadena complementaria del ARNcr y una segunda mutación en el dominio que separa la cadena no complementaria del ARNcr. En algunas realizaciones, la primera mutación es DA10A y la segunda mutación es H840A. En algunas realizaciones, la proteína Cas o la variante de la misma del primer sistema o del segundo sistema es una proteína en Cascada o una variante de la misma. En algunas realizaciones, la proteína Cas o la variante de la misma del primer sistema o del segundo sistema es una proteína Cas3 o una variante de la misma.

40

45 En algunas realizaciones, la polimerasa es una polimerasa que desplaza la cadena. En algunas realizaciones, la polimerasa se selecciona de un grupo que consiste en Bst, Bsu, u Phi29.

En algunas realizaciones, el ácido nucleico diana es ADN genómico. En algunas realizaciones, el ácido nucleico diana contiene ADN cromosómico o un fragmento del mismo. En algunas realizaciones, el ácido nucleico diana comprende un genoma o un genoma parcial.

50 En algunas realizaciones, el método proporcionado en la presente memoria incluye además secuenciar el ácido nucleico diana o los fragmentos de ácido nucleico diana. En algunas realizaciones, la secuenciación comprende el uso de una o más secuenciaciones mediante síntesis, PCR de puente, secuenciación por terminación en cadena, secuenciación por hibridación, secuenciación por nanoporos, y secuenciación por ligación.

### Breve descripción de los dibujos

55 La Figura 1A muestra los cebadores diseñados según los métodos actuales para amplificar fragmentos de ADN que contienen todos o partes de los adaptadores P5 y P7 del cebador de secuenciación universal de Illumina. Los cebadores se pueden añadir empleando el método de preparación de biblioteca la Nextera (Illumina, Inc). La Figura 1B ilustra un recorrido de amplificación lineal mediado por Cas9 empleando un ARNcr dirigido a un cebador P5

modificado. 2B(I) representa el ADN a amplificar que posee secuencias cebadoras P5 y P7 adecuadas. El ARN guía que se dirige a P5, unido a Cas9, también se muestra como P5'. 1B(II) muestra el bucle R creado por Cas9 después de que el ARN se haya hibridado a la primera cadena. 1B(III) ilustra el cebador P5 inmovilizado que hibrida a la segunda cadena desplazada, seguido de la extensión por la polimerasa. 1B(IV) muestra el cebador P5 extendido resultante. El amplicón resultante se puede redirigir mediante Cas9+ARNcr como se muestra en la etapa 1B(II). La Figura 1C ilustra un recorrido de la amplificación lineal mediada por Cas9 empleando un ARNcr dirigido a un cebador P7 modificado. 1C(I) representa el ADN a amplificar que posee las secuencias adecuadas del cebador P5 y P7. El ARN guía que se dirige a P7, unido a Cas9, se muestra también como P7'. 1C(II) muestra el bucle R creado por Cas9 después de que el ARN se haya hibridado a la primera cadena. 1C(III) ilustra el cebador P7 inmovilizado que hibrida a la segunda cadena desplazada, seguido de la extensión por la polimerasa. 1C(IV) muestra el cebador P7 extendido resultante. El amplicón resultante se puede redirigir mediante Cas9+ARNcr como se muestra en la etapa 1C(II). Sometiendo a la amplificación a ambos lados del amplicón (Figuras 1B y 1C), se puede conseguir la amplificación exponencial.

### Descripción detallada

15 La presente divulgación proporciona métodos que emplean sistemas CRISPR-Cas para una amplificación rápida y eficaz del ácido nucleico diana.

La amplificación de ácidos nucleicos es una etapa de muchos procesos basados en ácidos nucleicos, tal como la secuenciación de última generación. Actualmente la reacción en cadena de la polimerasa (PCR) es el método más ampliamente utilizado para la amplificación del ADN, *por ejemplo*, para la detección e identificación de enfermedades infecciosas, trastornos genéticos y otros propósitos de investigación. Una reacción de PCR normalmente emplea dos cebadores de oligonucleótidos, que se hibridan a los extremos 3' de una secuencia de un ácido nucleico diana doble, y una ADN polimerasa, que puede extender los cebadores re-asociados mediante la adición de desoxirribonucleósidos trifosfato (dNTPs) para generar productos de ácidos nucleicos bicatenarios. Gill y Ghaemi, *Nucleosides, Nucleotides, and Nucleic Acids*, 2008, 27:224-243. Sin embargo, una reacción PCR requiere de un ciclo térmico para separar las dos hebras de ADN. De manera similar, muchos métodos de amplificación de ácidos nucleicos empleados actualmente, *por ejemplo*, aquellos empleados en la producción de un grupo en la secuenciación de la próxima generación requieren tanto de temperatura de ciclado como de intercambio de fluidos.

Se han desarrollado varios métodos de amplificación isotérmicos para eliminar la rampa de temperatura y los tiempos de equilibrado, tal como la amplificación de la recombinasa polimerasa (RPA) basada en la amplificación isotérmica, la amplificación mediada por transcripción, la amplificación basada en la secuencia de ácido nucleico, amplificación mediada por señal de ARN, amplificación por desplazamiento de la hebra, amplificación en círculo rodante, amplificación de ADN mediada por bucle, amplificación de desplazamiento múltiple isotérmico, amplificación dependiente de helicasa, amplificación isotérmica de un único cebador, y amplificación dependiente de la helicasa circular, como se describen en Gill y Ghaemi, *Nucleosides, Nucleotides, and Nucleic Acids*, 2008, 27:224-243.

Por ejemplo, en la amplificación mediada por transcripción (TMA), para producir ARN se emplea una ARN polimerasa de un promotor de ingeniería genética en la región del cebador, y después una transcriptasa inversa sintetiza ADNc a partir del cebador. Se puede emplear una tercera enzima, *por ejemplo*, la RNasa H, para degradar el ARN diana a partir de un ADNc sin la etapa de desnaturalización por calor. Esta técnica de amplificación es muy similar a la Replicación de Secuencia Auto-sostenida (3SR) y a la Amplificación Basada en Secuencias de Ácidos Nucleicos (NASBA), pero varía en las enzimas empleadas. *Id.* Otro ejemplo, la amplificación dependiente de helicasa (HDA) emplea una helicasa termoestable (Tte-UvrD) en lugar del calor para desenrollar los ADNds para crear cadenas simples que después están disponibles para la hibridación y la extensión de los cebadores mediante la polimerasa. Se ha demostrado que los tiempos de reacción son superiores a 1 hora para amplificar productos de 70-120 pares de bases de longitud. *Id.* Otro ejemplo, la amplificación mediada por bucle (LAMP) emplea una polimerasa termoestable con capacidad de desplazar la hebra y un grupo de cuatro o más cebadores diseñados específicamente. Cada cebador se diseña para que tenga extremos en horquilla, que una vez desplazado, se rompa en una horquilla para facilitar el autocebado y otra extensión de la polimerasa. En la reacción LAMP, aunque la reacción procede bajo condiciones isotérmicas, se requiere de una etapa inicial de desnaturalización por calor para dianas de doble hebra. Además, la amplificación produce un patrón en escalera de varios productos de longitud. *Id.* Otro ejemplo, la amplificación por desplazamiento de la hebra (SDA) combina la capacidad de una endonucleasa de restricción para cortar la hebra sin modificar de su ADN diana y una polimerasa de ADN deficiente de exonucleasa para prolongar el extremo 3' y cortar y desplazar la hebra de ADN más adelante. *Id.* Otros ejemplos de métodos de amplificación isotérmico incluyen, pero no se limitan a, los descritos en Craw y Balachandran, *Lab Chip*, 2012, 12:2469-2486.

55 Sin embargo, estos sistemas de amplificación isotérmicos desarrollados en la actualidad carecen normalmente de la velocidad deseada y de la eficacia adecuada idealmente para algunas aplicaciones. Además, algunos sistemas requieren de enzimas adicionales y reactivos que incluyen ATP. Por tanto, existe una necesidad en la técnica de métodos de amplificación isotérmicos de ácidos nucleicos que sean cómodos, rápidos, y eficaces. La presente divulgación satisface esta necesidad proporcionando métodos para amplificar ácidos nucleicos empleando sistemas CRISPR-Cas. Por ejemplo, una ventaja proporcionada por los presentes métodos es que la proteína Cas reconoce

el ácido nucleico diana sin consumir ATP o invertir energía, y por tanto, los métodos de la presente memoria proporcionan métodos de amplificación isotérmica eficaces en coste y en tiempo.

### Definiciones

- 5 Como se emplea en la presente memoria, los términos “incluye”, “que incluye”, “contienen”, “que contiene”, “tiene”, “que tiene”, y cualquiera de las variaciones de los mismos, se destinan a cubrir una inclusión no exclusiva, tal como un proceso, método, producto por procedimiento, o composición de materia que incluye, incluyen, o contienen un elemento o lista de elementos que no incluye sólo a esos elementos sino que pueden incluir otros elementos no enumerados expresamente o inherentes para tal proceso, método, producto por procedimiento, o composición de materia.
- 10 Como se emplea en la presente memoria, las formas singulares “un”, “uno” y “el”, “la” incluyen los referentes en plural a menos que el contexto indique claramente lo contrario. Por tanto, por ejemplo, la referencia a “una proteína” incluye una mezcla de dos o más proteínas, y similares.
- Como se emplea en la presente memoria, el término “aproximadamente” o “aproximadamente” significa dentro del 5% de un intervalo o valor dado.
- 15 Como se emplea en la presente memoria, el término “ácido nucleico” significa polímeros de cadena simple y de cadena doble de monómeros de nucleótidos, que incluyen 2'-desoxirribonucleótidos (ADN) y ribonucleótidos (ARN) unidos mediante enlaces internucleótido fosfodiéster, o análogos de internucleótidos, y contraiones asociados, *por ejemplo*, H<sup>+</sup>, NH<sup>4+</sup>, trialkilamonio, tetraalkilamonio, Mg<sup>2+</sup>, Na<sup>+</sup>, y similares. Un ácido nucleico puede ser un polinucleótido o un oligonucleótido. Un ácido nucleico se puede componer completamente de desoxirribonucleótidos, completamente de ribonucleótidos, o de mezclas quiméricas de los mismos. Las unidades de monómero del nucleótido pueden comprender cualquiera de los nucleótidos descritos en la presente memoria, que incluyen, pero no se limitan a, nucleótidos que aparecen de manera natural y a análogos de nucleótidos. Normalmente el ácido nucleico oscila de un tamaño de unas pocas unidades monoméricas, *por ejemplo*, 5-40, a varios miles de unidades de nucleótidos momoméricos. Los ácidos nucleicos incluyen, pero no se limitan a, ADN genómico, ADNc, ARNhn, ARNm, ARNr, ARNt, ácido nucleico fragmentado, ácido nucleico obtenido a partir de orgánulos sub-celulares, tales como mitocondrias o cloroplastos, y ácido nucleico obtenido a partir de microorganismos o virus de ADN o ARN que pueden presentarse o estar en una muestra biológica.
- 20 Como se emplea en la presente memoria, el término “ácido nucleico diana” se destina a indicar un ácido nucleico que es objeto de un análisis o acción. El análisis o acción incluye someter al ácido nucleico a la copia, amplificación, secuenciación y/u otro procedimiento para la interrogación del ácido nucleico. Un ácido nucleico diana puede incluir secuencias de nucleótidos adicionales a la secuencia diana a analizar. Por ejemplo, un ácido nucleico diana puede incluir uno o más adaptadores, que incluyen un adaptador que funciona como un cebador de unión al sitio, que flanquea una secuencia de ácido nucleico diana a analizar. Un ácido nucleico diana hibridado a un oligonucleótido de captura o a un cebador de captura puede contener nucleótidos que se extienden más allá de los extremos 3' y 5' del oligonucleótido de captura, de tal modo que no todo el ácido nucleico diana es susceptible de extensión.
- 25 Como se emplea en la presente memoria, el término “diana específica” cuando se emplea en referencia a un ARN guía, un ARNcr o un derivado de los mismos, u otro nucleótido, se destina a significar un polinucleótido que incluye una secuencia de nucleótidos específica para una secuencia de polinucleótidos diana, especialmente una secuencia de nucleótidos capaz de re-asociarse selectivamente a una región de identificación de un polinucleótido diana, *por ejemplo*, un ADN diana. El nucleótido específico diana puede tener especies únicas de oligonucleótidos, o puede incluir dos o más especies con secuencias diferentes. Por tanto, el nucleótido específico diana puede ser dos o más secuencias, que incluyen, 3, 4, 5, 6, 7, 8, 9, ó 10 o más secuencias diferentes. En una realización, un ARNcr o el derivado del mismo, contiene una región de nucleótidos específica complementaria para una región de una secuencia de ADN diana. En una realización, un ARNcr o el derivado del mismo puede contener otras secuencias de nucleótidos además de una región de nucleótidos específica de la diana. En una realización, las otras secuencias de nucleótidos pueden ser a partir de una secuencia de ARNtracr.
- 30 Como se emplea en la presente memoria, el término “complementario” cuando se emplea en referencia a un polinucleótido se destina a significar un polinucleótido que incluye una secuencia de nucleótidos capaz de re-asociarse selectivamente a una región de identificación de un polinucleótido diana bajo determinadas condiciones.
- 35 Como se emplea en la presente memoria, el término “sustancialmente complementario” y equivalentes gramaticales se destinan a significar un polinucleótido que incluye una secuencia de nucleótidos capaz de re-asociarse específicamente a una región de identificación de un polinucleótido diana bajo determinadas condiciones. Re-asociación se refiere a la interacción del emparejamiento de bases de nucleótidos de un ácido nucleico con otro ácido nucleico que da como resultado la formación de una estructura doble, triple, u otra estructura de mayor orden.
- 40 La interacción primaria es normalmente específica de la base, *por ejemplo*, A:T, A:U, y G:C, mediante uniones de hidrógeno tipo Watson-Crick y Hoogsteen. En determinadas realizaciones, las interacciones de apilamiento de bases y las hidrofóbicas pueden contribuir también a la estabilidad del dúplex. Las condiciones bajo las que un polinucleótido se re-asocia a regiones complementarias o sustancialmente complementarias de ácidos nucleicos diana son bien conocidos en la técnica, *por ejemplo*, como se describe en Nucleic Acid Hybridization, A Practical
- 45 Como se emplea en la presente memoria, el término “complementario” cuando se emplea en referencia a un polinucleótido se destina a significar un polinucleótido que incluye una secuencia de nucleótidos capaz de re-asociarse selectivamente a una región de identificación de un polinucleótido diana bajo determinadas condiciones.
- 50 Como se emplea en la presente memoria, el término “sustancialmente complementario” y equivalentes gramaticales se destinan a significar un polinucleótido que incluye una secuencia de nucleótidos capaz de re-asociarse específicamente a una región de identificación de un polinucleótido diana bajo determinadas condiciones. Re-asociación se refiere a la interacción del emparejamiento de bases de nucleótidos de un ácido nucleico con otro ácido nucleico que da como resultado la formación de una estructura doble, triple, u otra estructura de mayor orden.
- 55 La interacción primaria es normalmente específica de la base, *por ejemplo*, A:T, A:U, y G:C, mediante uniones de hidrógeno tipo Watson-Crick y Hoogsteen. En determinadas realizaciones, las interacciones de apilamiento de bases y las hidrofóbicas pueden contribuir también a la estabilidad del dúplex. Las condiciones bajo las que un polinucleótido se re-asocia a regiones complementarias o sustancialmente complementarias de ácidos nucleicos diana son bien conocidos en la técnica, *por ejemplo*, como se describe en Nucleic Acid Hybridization, A Practical

Approach, Hames y Higgins, eds., IRL Press, Washington, D.C. (1985) y Wetmur y Davidson, Mol. Biol. 31:349 (1968). Las condiciones de re-asociación dependerán de la aplicación en particular, y se pueden determinar de manera rutinaria por los expertos en la técnica, sin excesiva experimentación.

5 Como se emplea en la presente memoria, el término "hibridación" se refiere al proceso en el que dos polinucleótidos de hebra simple se unen no covalentemente para formar un polinucleótido de doble hebra. Un polinucleótido de doble hebra resultante es un "híbrido" o "dúplex". Las condiciones de hibridación normalmente incluyen concentraciones de sales de menos de aproximadamente 1 M, más normalmente de menos de aproximadamente 500 mM y pueden ser menores de aproximadamente 200 mM. Un tampón de hibridación incluye una disolución de sal tamponada, tal como SSPE al 5%, u otros tampones conocidos en la técnica. Las temperaturas de hibridación  
10 pueden ser tan bajas como 5°C, pero normalmente son superiores a 22°C, y más normalmente superiores a aproximadamente 30°C, y normalmente de más de 37°C. Normalmente las hibridaciones se realizan bajo estrictas condiciones, es decir, condiciones bajo las que una sonda se hibridará a su subsecuencia diana pero que no se hibridará a la otra, secuencias sin complementariedad. Las condiciones estrictas son dependientes de la secuencia y son diferentes en las distintas circunstancias, y se pueden determinar de manera rutinaria por los expertos en la  
15 técnica.

En el contexto de "polinucleótidos", los términos "variante" y "derivado", como se emplean en la presente memoria se refieren a un polinucleótido que comprende una secuencia de nucleótidos de un polinucleótido o un fragmento de un polinucleótido, que se ha alterado mediante la introducción de sustituciones, deleciones o adiciones de nucleótidos. Una variante o derivado de un polinucleótido puede ser un polinucleótido de fusión que contiene parte de una  
20 secuencia de nucleótidos de un polinucleótido. El término "variante" o "derivado" como se emplea en la presente memoria se refiere también a un polinucleótido o a un fragmento del mismo, que se ha modificado químicamente, por ejemplo, mediante la unión covalente de cualquier tipo de molécula al polinucleótido. Por ejemplo, pero no a modo de limitación, un polinucleótido o un fragmento del mismo se puede modificar químicamente, por ejemplo, mediante acetilación, fosforilación, metilación, etc. Las variantes o derivados de nucleótidos o polinucleótidos se modifican de una manera que es diferente a la que aparece o se inicia de manera natural, bien en el tipo o bien en la localización de las moléculas unidas. Las variantes o derivados incluyen además la deleción de uno o más grupos  
25 químicos que se presentan de manera natural en el nucleótido o polinucleótido. Una variante o un derivado de un polinucleótido o un fragmento de un polinucleótido se puede modificar químicamente mediante modificaciones químicas empleando técnicas conocidas por los expertos en la técnica, que incluyen, pero no se limitan a, la separación química específica, acetilación, formulación, etc. Además, una variante o derivado de un polinucleótido o un fragmento de un polinucleótido puede contener uno o más dNTPs o análogos de nucleótidos. Una variante o derivado de polinucleótido puede poseer unas funciones similares o idénticas al polinucleótido o fragmento de un polinucleótido descrito en la presente memoria. Una variante o derivado de un polinucleótido puede poseer una función diferente o adicional en comparación con un polinucleótido o un fragmento de un polinucleótido descrito en la  
30 presente memoria.

Como se emplea en la presente memoria, el término "dNTP" se refiere a desoxirribonucleósidos trifosfato. NTP se refiere a ribonucleótidos trifosfato tal como los empleados para sintetizar ARNcr o ARNtracr. Las bases de purina (Pu) incluyen adenina (A), guanina (G) y derivados y análogos de las mismas. Las bases de pirimidina (Py) incluyen citosina (C), timidina (T), uracilo (U) y derivados y análogos de las mismas. Ejemplos de tales derivados o análogos,  
40 a modo de ilustración y no de limitación, son aquellos que se modifican con un grupo reportero, biotinilado, amino modificado, radiomarcado, alquilado, y similares, y también los que incluyen fosforotioato, fosfito, derivados modificados en el átomo del anillo, y similares. El grupo reportero puede ser un grupo fluorescente, tal como fluoresceína, un grupo quimioluminiscente, tal como luminol, un quelante de terbio tal como ácido N-(hidroxietil) etilenediaminetriacético que es capaz de la detección mediante fluorescencia retardada, y similares.

45 Como se emplea en la presente memoria, el término "análogos de nucleótidos" se refiere a análogos sintéticos que tienen porciones de bases de nucleótidos modificados, porciones de pentosa modificadas, y/o porciones de fosfato modificadas, y, en el caso de polinucleótidos, uniones de internucleósidos modificadas, como se describe generalmente en otras fuentes (por ejemplo, *Nucleotide Analogs*, John Wiley, Nueva York, 1980; *Englisch, Angew. Chem. Int. Ed. Engl.* 30:613-29, 1991; Agarwal, *Protocols for Polynucleotides and Analogs*, Humana Press, 1994; y S. Verma y F. Eckstein, *Ann. Rev. Biochem.* 67:99-134, 1998). Ejemplos de análogos de fosfato incluyen, pero no se limitan a, fosforotioato, fosforoditioato, fosforoselenoato, fosforodiselenoato, fosforoanilotoato, fosforaniladato, fosforoamidato, boranofosfatos, incluyendo contraiones asociados, por ejemplo, H<sup>+</sup>, NH<sub>4</sub><sup>+</sup>, Na<sup>+</sup>, si se presentan tales  
50 contraiones. Ejemplos de porciones de bases de nucleótido modificadas incluyen, pero no se limitan a, 5-metilcitosina (5mC); análogos de C-5-propinilo, que incluyen pero no se limitan a, C-5 propinil-C y C-5 propinil-U; 2,6-diaminopurina, también conocida como 2-amino adenina o 2-amino-dA); hipoxantina, pseudouridina, 2-tiopirimidina, isocitosina (isoC), 5-metil isoC, e isoguanina (isoG; véase, por ejemplo, la Patente U.S. Nº 5.432.272). Ejemplos de porciones modificadas de pentosa incluyen, pero no se limitan a, análogos de ácido nucleico bloqueado (LNA) que incluyen sin limitación Bz-A-LNA, 5-Me-Bz-C-LNA, dmf-G-LNA, y T-LNA (véase, por ejemplo, The Glen Report, 16(2):5, 2003; Koshkin *et al.*, *Tetrahedron* 54:3607-30, 1998), y modificaciones 2' o 3' en las cuales la  
55 posición 2' o 3' es hidrógeno, hidroxil, alcoxi (por ejemplo, metoxi, etoxi, aliloxi, isopropoxi, butoxi, isobutoxi y fenoxi), azido, amino, alquilamino, fluoro, cloro, o bromo. Enlaces internucleótidos modificados incluyen análogos de fosfato, análogos que tienen enlaces aquirales y enlaces no cargados de intersubunidad (por ejemplo, Sterchak, E. P. *et al.*, *Organic Chem.*, 52:4202, 1987), y polímeros a base de morfolino no cargados que tienen enlaces aquirales

de intersubunidad (véase, *por ejemplo*, la Patente U.S. N° 5.034.506). Algunos análogos de enlace internucleótido incluyen morfolidato, acetal y heterociclos enlazados a poliamida.

Los términos “reacción en cadena de polimerasa”, o “PCR”, como se emplean en la presente memoria, se refieren a un procedimiento donde cantidades pequeñas de un ácido nucleico, *por ejemplo*, ARN y/o ADN, se amplifican como se describe, *por ejemplo*, en la Patente U.S. N° 4.683.195 por Mullis. Generalmente, se necesita tener información de la secuencia a partir de los extremos de la región de interés, o se necesita tenerla disponible de tal forma que se puedan diseñar cebadores de oligonucleótidos; estos cebadores serán idénticos o similares en la secuencia para hebras opuestas del molde a amplificar. Los nucleótidos en el extremo 5' de dos cebadores pueden coincidir con los extremos del material amplificado. El PCR se puede emplear para amplificar secuencias de ARN específicas, secuencias de ADN específicas a partir del ADN genómico total, y ADNc transcrito a partir de ARN celular total, secuencias de bacteriófagos o plásmidos, etc. Véase *generalmente*, Mullis *et al.*, Cold Spring Harbor Symp. Quant. Biol., 51:263 (1987); Erlich, ed., PCR Technology, (Stockton Press, NY, 1989).

Como se emplea en la presente memoria, los términos “ligación”, “que liga”, y los equivalentes de los mismos se destinan a significar que forman un enlace o unión covalente entre el extremo de dos o más ácidos nucleicos, *por ejemplo*, oligonucleótidos y/o polinucleótidos, normalmente en una reacción dirigida por el molde. La naturaleza del enlace o de la unión puede variar ampliamente y la ligación se puede llevar a cabo enzimática o químicamente. Como se emplea en la presente memoria, las ligaciones se llevan a cabo normalmente de manera enzimática para formar un enlace fosfodiéster entre el carbono 5' del nucleótido terminal de un oligonucleótido con el carbono 3' de otro nucleótido. Las reacciones de ligación dirigidas por el molde se describen en las siguientes referencias: Patentes U.S. N°s 4.883.750; 5.476.930; 5.593.826, y 5.871.921. El término “ligación” engloba también la formación no enzimática de enlaces fosfodiéster, así como la formación de enlaces covalentes no fosfodiéster entre los extremos de oligonucleótidos, tal como enlaces fosforotioato, enlaces disulfuro, y similares.

Como se emplea en la presente memoria, el término “adaptador” es una molécula de ácido nucleico de cadena simple o de cadena doble que se puede unir al extremo de otros ácidos nucleicos. En una realización, un adaptador es una molécula corta, sintetizada químicamente, de doble cadena que se puede emplear para unir los extremos de otras dos moléculas de ácido nucleico. En una realización, un adaptador es un ácido nucleico de doble cadena (*por ejemplo*, oligonucleótidos) que comprende salientes en los extremos 5' y/o 3'. En algunas realizaciones, los salientes de cadena simple son 1, 2 o más nucleótidos. En algunas realizaciones, los adaptadores comprenden una secuencia de ácido nucleico adicional para clonar o analizar “insertos”. En algunas realizaciones, los adaptadores comprenden marcadores o marcadores de afinidad para el análisis o la purificación de “insertos”. El término “inserto” se refiere a una secuencia de ácido nucleico de interés. En algunas realizaciones, los insertos son ADNs bicatenarios que comprenden salientes de nucleótidos de una sola cadena en los extremos 5' y/o 3'. En algunas realizaciones, los salientes de la hebra simple son 1, 2, o más nucleótidos.

Como se emplea en la presente memoria, “sistema CRISPR-Cas” se refiere a un sistema enzimático que incluye una secuencia de ARN guía que contiene una secuencia de nucleótidos complementaria o sustancialmente complementaria para una región de un polinucleótido diana, y una proteína con actividad nucleasa. Los sistemas CRISPR-Cas incluyen el sistema CRISPR-Cas Tipo I, el sistema CRISPR-Cas Tipo II, el sistema CRISPR-Cas Tipo III, y derivados de los mismos. Los sistemas CRISPR-Cas incluyen sistemas de nucleasa modificados genéticamente y/o de sistemas nucleasa programados derivados de los sistemas CRISPR-Cas que aparecen de manera natural. Los sistemas CRISPR-Cas pueden contener proteínas Cas modificadas genéticamente y/o proteínas Cas mutadas. Los sistemas CRISPR-Cas pueden contener ARN guía modificado genéticamente y/o programado.

Como se emplea en la presente memoria, el término “ARN guía” se refiere a un ARN que contiene una secuencia complementaria o sustancialmente complementaria para una región de una secuencia de ADN diana. Un ARN guía puede contener secuencias de nucleótidos en otra región distinta a la región complementaria o sustancialmente complementaria para una región de una secuencia de ADN diana. Un ARN guía puede ser un ARNcr o un derivado del mismo, *por ejemplo*, un ARNcr: quimera ARNtracr.

Como se emplea en la presente memoria, el término “nucleasa” se refiere a una enzima capaz de separar los enlaces fosfodiéster entre subunidades de nucleótidos de ácidos nucleicos; el término “endonucleasa” se refiere a una enzima capaz de separar el enlace fosfodiéster dentro de una cadena de polinucleótido; y el término “nickasa” se refiere a una endonucleasa que separa sólo una hebra simple de un ADN dúplex. El término “nickasa Cas9” se refiere a una nickasa derivada de una proteína Cas9, normalmente mediante la inactivación de un dominio nucleasa de la proteína Cas9.

En el contexto de un polipéptido, los términos “variante” y “derivado” como se emplean en la presente memoria, se refieren a un polipéptido que comprende una secuencia de aminoácido de un polipéptido o un fragmento de un polipéptido, que se ha alterado mediante la introducción de sustituciones, deleciones o adiciones de restos de aminoácido. Una variante o un derivado de un polipéptido puede ser una proteína de fusión que contiene parte de la secuencia de aminoácido de un polipéptido. El término “variante” o “derivado” como se emplea en la presente memoria se refiere también a un polipéptido o a un fragmento de un polipéptido, que se ha modificado químicamente, *por ejemplo*, mediante la unión covalente de cualquier tipo de molécula al polipéptido. *Por ejemplo*,

pero no a modo de limitación, un polipéptido o un fragmento de un polipéptido se puede modificar químicamente, *por ejemplo*, mediante glicosilación, acetilación, pegilación, fosforilación, amidación, derivación, mediante grupos protectores/bloqueantes, separación proteolítica, unión a un ligando celular o a otra proteína, etc. Las variantes o derivados se modifican de una manera que es diferente a la que ocurre o comienza en péptidos o polipéptidos de manera natural, bien en el tipo o bien en la localización de las moléculas unidas. Las variantes o derivados incluyen además la delección de uno o más grupos químicos que se presentan de manera natural en el péptido o polipéptido. Una variante o un derivado de un polipéptido o un fragmento de un polipéptido se puede modificar mediante modificaciones químicas empleando técnicas conocidas por los expertos en la técnica, que incluyen, pero no se limitan a, separación química específica, acetilación, formulación, síntesis metabólica de tunicamicina, etc. Además, una variante o derivado de un polipéptido o un fragmento de un polipéptido puede contener uno o más aminoácidos no clásicos. Una variante o derivado de polipéptido puede poseer una función idéntica o similar al de un polipéptido o fragmento de un polipéptido descrito en la presente memoria. Una variante o derivado del polipéptido puede poseer una función adicional o diferente en comparación con un polipéptido o un fragmento de un polipéptido descrito en la presente memoria.

Como se emplea en la presente memoria, el término “detectar” una molécula de ácido nucleico o fragmento del mismo se refiere a determinar la presencia de la molécula de ácido nucleico, normalmente cuando la molécula de ácido nucleico o fragmento del mismo se ha separado completa o parcialmente de otros componentes de una muestra o composición, y puede incluir también determinar la relación de carga sobre masa, la masa, la cantidad, la absorbancia, la fluorescencia, u otra propiedad de la molécula de ácido nucleico o fragmento del mismo.

Como se emplea en la presente memoria, el término “cebador” se refiere a un cebador oligonucleótido, ya sea natural o sintético, que es capaz de actuar como un punto de inicio de la síntesis de ácido nucleico cuando se sitúa bajo condiciones en las que el cebador inicia la extensión (no limitado a un número de bases extendidas). Un cebador puede ser un oligodesoxirribonucleótido monocatenario. La longitud de un cebador puede oscilar de aproximadamente 10 a aproximadamente 50 nucleótidos, por ejemplo, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, o aproximadamente 50 nucleótidos. Un cebador no necesita reflejar la secuencia exacta del molde pero debe ser lo suficientemente complementario como para hibridar con un molde para que ocurra la elongación del cebador. Un cebador se puede marcar, si se desea, mediante la incorporación de un marcador que es detectable mediante, por ejemplo, espectroscopía, o por medio fotoquímico, bioquímico, o inmunológico. Ejemplos de marcadores incluyen, pero no se limitan a, biotina, radiomarcadores (*por ejemplo*, 32P), tintes fluorescentes, reactivos electron-densos, enzimas (como se emplean normalmente en los ELISAS), o biotina.

Como se emplea en la presente memoria, el término “amplificación lineal” se refiere a un proceso de amplificación que emplea ciclos múltiples de reacciones de extensión del cebador para amplificar un ácido nucleico diana. Con la amplificación lineal, la abundancia de un transcrito incrementa proporcionalmente con el número de ciclos y de escalas de manera lineal. Un ejemplo de un procedimiento de amplificación lineal es el LCR (reacción en cadena de ligasa, de sus siglas en inglés), el método ARNa de Phillips y Eberwine, citado anteriormente, y el método de amplificación lineal descrito en la presente memoria. A diferencia de la amplificación exponencial, la cantidad de productos de amplificación no crece de manera exponencial. Por ejemplo, en una reacción ideal de amplificación lineal de 4 horas cuya tasa de replicación es de 2000 copias por minuto, 2000 copias del ADN molde rendirán 960.000.000 de copias.

Como se emplea en la presente memoria, el término “amplificación exponencial” se refiere a un procedimiento de amplificación donde el producto (es decir, amplicón) se dobla con cada ciclo de reacción. La “amplificación exponencial” es una amplificación no lineal que da como resultado un crecimiento exponencial en el número de copias de ácido nucleico presentes. Por ejemplo, la amplificación exponencial puede ocurrir cuando la extensión del cebador se inicia a partir de ambos extremos de un amplicón en un ciclo de amplificación. Por ejemplo, el PCR es un procedimiento de amplificación exponencial. Por ejemplo, una reacción PCR ideal con 30 ciclos, dos copias del ADN molde rendirán  $2^{30}$  ó 1.073.741.824 copias.

Como se emplea en la presente memoria, el término “polimerasa” se refiere a una proteína que es capaz de catalizar la incorporación específica de nucleótidos para prolongar el hidroxilo 3' terminal de una molécula cebadora, tal como, por ejemplo, el molde de oligonucleótidos, frente a una secuencia diana de ácido nucleico. La polimerasa puede ser, por ejemplo, termófila para que se active a una temperatura de reacción elevada. Puede tener también, por ejemplo, capacidades de desplazamiento.

#### **Métodos para la amplificación de polinucleótidos**

En un aspecto, la presente divulgación proporciona un método para la amplificación mediada por un sistema CRISPR-Cas. Los métodos proporcionados en la presente memoria se basan parcialmente en que la unión de un ARN guía a un ácido nucleico bicatenario altera la interacción entre las dos cadenas del ácido nucleico diana, y crea de esta manera una estructura en bucle (también llamada “bucle-R”) que expone la cadena no complementaria al ARN guía. Esta hebra al descubierto se puede someter a la hibridación con el cebador para la extensión mediante una polimerasa adecuada, *por ejemplo*, en un proceso de amplificación de ácido nucleico. Como se ilustra en el Ejemplo 1, la estructura en bucle creada mediante los sistemas CRISPR-Cas, *por ejemplo*, sistemas que contienen

proteínas Cas9 o en Cascada, pueden ser accesibles para otras enzimas. Por tanto, la estructura en bucle se puede utilizar además como un molde para iniciar la hibridación del cebador e interactuar con las enzimas polimerasa para la amplificación.

5 En algunas realizaciones, la presente divulgación proporciona un método para amplificar un ácido nucleico bicatenario diana que incluye proporcionar un sistema que tiene repeticiones palindrómicas cortas agrupadas y regularmente interespaciadas (CRISPR) de ARN (ARN) (ARNcr) o derivados de las mismas, y una proteína CRISPR-asociada (Cas) o una variante de la misma, en donde el ARNcr o el derivado del mismo contiene una región de nucleótidos específica de la diana complementaria para una región de la primera hebra del ácido nucleico bicatenario diana; poniendo en contacto el ácido nucleico bicatenario diana con el sistema para formar un complejo;  
10 hibridando un cebador a una segunda hebra del ácido nucleico bicatenario diana, el cebador que contiene una secuencia complementaria para una región de la segunda hebra del ácido nucleico bicatenario diana, y extendiendo un ácido nucleico complementario a la segunda hebra del ácido nucleico bicatenario diana a partir del cebador empleando una polimerasa.

15 Los métodos proporcionados en la presente memoria se pueden emplear en varios métodos de amplificación, que incluyen, pero no se limitan a, amplificación de ácido nucleico lineal, y amplificación de ácido nucleico exponencial.

En algunas realizaciones, el ácido nucleico diana se amplifica de manera lineal según los métodos proporcionados en la presente memoria. Por ejemplo, en algunas realizaciones, el método proporcionado en la presente memoria incluye además repetir la hibridación del cebador para la segunda hebra del ácido nucleico bicatenario diana, y prolongar un ácido nucleico complementario a la segunda hebra del ácido nucleico bicatenario diana a partir del  
20 cebador empleando una polimerasa durante una o más veces, *por ejemplo*, hasta que se consigue la cantidad de amplificación deseada.

En otras realizaciones, el ácido nucleico diana se amplifica de manera exponencial. En un ejemplo de amplificación exponencial de ácido nucleico, el producto (es decir, amplicón) se dobla con cada ciclo. La “amplificación exponencial” es una amplificación no lineal que da como resultado un crecimiento exponencial del número de copias de ácido nucleico presente. Normalmente, en una amplificación exponencial, la extensión (o copia) del cebador ocurre a partir de ambos extremos de un amplicón. Para asegurar que la hebra recién creada tiene sitios de unión al cebador en ambos extremos, en una amplificación exponencial, el extremo 3' del nuevo ácido nucleico sintetizado contiene el complemento inverso de un cebador, y los extremos del molde se copian normalmente en cada ciclo de  
25 amplificación.

30 El número de reacciones del ciclo según los métodos actuales depende de las aplicaciones y de la cantidad deseada de los productos de amplificación. En algunas realizaciones, el número de ciclos es de 5 a 100. En algunas realizaciones, el número de ciclos es de 10 a 90. En algunas realizaciones, el número de ciclos es de 20 a 80. En algunas realizaciones, el número de ciclos es de 30 a 70. En algunas realizaciones, el número de ciclos es de 40 a 60. Ejemplos de números de ciclo incluyen 10, 15, 20, 25, 30, 35, 40, 45, y 50. Los ciclos no necesitan estar  
35 sincronizados entre los amplicones, como están en una reacción de PCR cuando comienza cada ciclo se controla mediante el cambio de temperatura. Ciclos, como se emplean en la presente memoria se refiere, por tanto, al número medio de series de amplificaciones para un amplicón seleccionado.

En algunas realizaciones, el inicio de la etapa incluye proporcionar sistemas CRISPR-Cas para el ácido nucleico diana para abrir la estructura del ácido nucleico bicatenario en dos o más secuencias diana y formar dos o más  
40 estructuras en bucle R. La etapa de iniciación no requiere calentar la reacción, como se requiere en la reacción por PCR, ya que la etapa de iniciación se dirige enzimáticamente. La próxima etapa según los presentes métodos implica proporcionar y re-asociar los cebadores diana para las regiones en bucle R para formar la interacción relativamente estable del ácido nucleico-ácido nucleico, *por ejemplo*, híbridos de ADN-ADN. Normalmente, los híbridos de ácido nucleico-ácido nucleico estables se forman cuando la secuencia cebadora tiene  
45 complementariedad sustancial con la secuencia molde. Después, una o más polimerasas se unen al híbrido cebador-molde y comienza la síntesis del ácido nucleico en una etapa de extensión/elongación. En algunas realizaciones, la temperatura en esta etapa de extensión/elongación depende de la polimerasa empleada. En esta etapa la polimerasa sintetiza una nueva hebra de ácido nucleico complementaria a la hebra molde mediante la adición de dNTPs que son complementarios al molde. En algunas realizaciones, cuando se emplea la ADN  
50 polimerasa, la reacción condensa el grupo 5'-fosfato de los dNTPs con el grupo 3'-hidroxilo en el extremo de la hebra de ADN naciente (que se extiende). El tiempo de extensión depende tanto de la polimerasa empleada como de la longitud del fragmento del ácido nucleico a amplificar. Una o más de estas etapas se pueden repetir durante una o dos veces, *por ejemplo*, hasta que se alcance la cantidad de amplificación deseada.

Para permitir el crecimiento exponencial de los productos amplificados, es beneficioso que el producto del ácido nucleico recién creado en cada ciclo contenga sitios de unión al cebador en ambos extremos. En algunas realizaciones, el extremo 3' de la molécula recién sintetizada es el complemento inverso de un cebador. En algunas realizaciones, en una amplificación exponencial, se pueden emplear dos cebadores, cada uno dirigido a un extremo de un ácido nucleico diana. En algunas realizaciones, es también beneficioso que el cebador se diseñe de tal forma que se pueda dirigir mediante los sistemas CRISPR-Cas proporcionados en la presente memoria – el cebador se  
60 puede dirigir mediante el ARN guía, *por ejemplo*, ARNcr, de los sistemas CRISPR-Cas para que los sistemas

CRISPR-Cas se puedan emplear repetidamente para unir el ácido nucleico diana para iniciar una nueva ronda de amplificación.

Por tanto, en algunas realizaciones, se pueden emplear dos o más sistemas CRISPR-Cas para iniciar la unión del cebador a ambos extremos del ácido nucleico diana. En algunas realizaciones, en la presente memoria se proporciona un método para amplificar un ácido nucleico bicatenario diana que incluye: (a) proporcionar un primer sistema que tiene: un primer grupo de repeticiones palindrómicas cortas agrupadas y regularmente interespaciadas (CRISPR) de ARN (ARNcr) o un derivado de las mismas, y una primera proteína CRISPR-asociada (Cas) o una variante de la misma, en donde el primer ARNcr o el derivado del mismo contiene una región de nucleótidos específica de la diana complementaria para una región de una primera hebra del ácido nucleico bicatenario diana; (b) proporcionar un segundo sistema que tiene: un segundo grupo de repeticiones palindrómicas cortas agrupadas y regularmente interespaciadas (CRISPR) de ARN (ARNcr) o un derivado de las mismas, y una segunda proteína CRISPR-asociada (Cas) o una variante de la misma, en donde el segundo ARNcr o el derivado del mismo contiene una región de nucleótidos específica de la diana complementaria para una región de una segunda hebra del ácido nucleico bicatenario diana; (c) poner en contacto el ácido nucleico bicatenario diana con el primer sistema y el segundo sistema; (d) hibridar un primer cebador a una segunda hebra del ácido nucleico bicatenario diana, el primer cebador que contiene una secuencia complementaria para una región de la segunda hebra del ácido nucleico bicatenario diana, e hibridar un segundo cebador a una primera hebra del ácido nucleico bicatenario diana, el segundo cebador que contiene una secuencia complementaria para una región de la primera hebra del ácido nucleico bicatenario diana, y (e) prolongar el extremo 3' del primer cebador y del segundo cebador con una o más polimerasas para generar un primer y un segundo ácido nucleico diana bicatenario. En algunas realizaciones, la hibridación del cebador y la extensión mediante las etapas de la polimerasa se repiten durante una o más veces, *por ejemplo*, hasta que se consiga el grado de amplificación deseado.

En otras realizaciones, el método proporcionado en la presente memoria se emplea para amplificación multiplex. Como se emplea en la presente memoria, "amplificación multiplex" se refiere a la amplificación de más de un ácido nucleico de interés, *por ejemplo*, una amplificación de secuencias múltiples a partir de la misma muestra. El término "amplificación multiplex" se refiere también a la amplificación de una o más secuencias presentes en múltiples muestras bien simultáneamente o de manera gradual. Por tanto, en algunas realizaciones, dos o más secuencias de ácido nucleico diana que se están amplificando en una reacción de amplificación, y la reacción de amplificación, comprenden los moldes y enzimas adecuados para amplificar al menos dos secuencias del ácido nucleico diana. Una aplicación de la amplificación multiplex proporcionada en la presente memoria es detectar dos o más secuencias diana en una muestra ya que la amplificación multiplex es capaz de amplificar dos o más secuencias diana. Cuando sólo está realmente presente una de las secuencias diana en la muestra que se está ensayando, el resultado de la amplificación multiplex puede ser la amplificación de sólo una de las secuencias que está presente. La amplificación multiplex puede emplear el mismo par de cebadores para amplificar uno o más de las secuencias del amplicón de intervención. Alternativamente, la amplificación multiplex puede emplear uno o más pares de cebadores.

En algunas realizaciones, el ácido nucleico bicatenario proporcionado en la presente memoria es un ADN bicatenario. En otras realizaciones, el ácido nucleico bicatenario proporcionado en la presente memoria es un ARN bicatenario. En algunas realizaciones, el ácido nucleico diana es ADN genómico. En otras realizaciones, el ácido nucleico diana contiene ADN cromosómico o un fragmento del mismo. En otras realizaciones, el ácido nucleico diana comprende un genoma o un genoma parcial.

En algunas realizaciones, los sistemas proporcionados en la presente memoria son derivados de los sistemas CRISPR-Cas. Los sistemas CRISPR-Cas se pueden clasificar generalmente en tres tipos principales (Tipo I-III), que se subdividen además en diez subtipos, en base al contenido del elemento central y de las secuencias (Makarova *et al.*, *Nat Rev Microbiol* 9:467-77). Los dos elementos clave de estos sistemas CRISPR-Cas son las proteínas Cas y el ARN CRISPR (ARNcr). El ARNcr consiste en secuencias repetidas cortas con secuencias espaciadoras derivadas del ADN invasor. Las proteínas Cas tienen diferentes actividades, *por ejemplo*, actividad nucleasa. Por tanto, los sistemas CRISPR-Cas proporcionan mecanismos para dirigirse a una secuencia específica, así como determinadas actividades enzimáticas en la secuencia.

Un sistema CRISPR-Cas de Tipo I típico contiene proteína Cas3 con actividades helicasa y DNasa separadas. Por ejemplo, en el sistema Tipo 1-E, los ARNcrs se incorporan al complejo efector multisubunidad llamado Cascada (complejo CRISPR-asociado para defensa antiviral) (Brouns *et al.*, 2008, *Science* 321:960-4), que se une al ADN diana y desencadena la degradación mediante la proteína Cas3 (Sinkunas *et al.*, 2011, *EMBO J* 30:1335-1342; Beloglazova *et al.*, 2011, *EMBO J* 30:616-627).

Los sistemas CRISPR-Cas Tipo II incluyen la característica proteína Cas9, una proteína en solitario (aproximadamente 160KDa), capaz de generar ARNcr y separar el ADN diana. La proteína Cas9 normalmente contiene dos dominios de nucleasa, un dominio de nucleasa similar a RuvC cerca del extremo amino y el dominio de nucleasa HNH (o similar a McrA) cerca de la mitad de la proteína. Cada dominio nucleasa de la proteína Cas9 se especializa en cortar una hebra de la doble hélice (Jinek *et al.*, 2012, *Science* 337 (6096):816-821).

Los sistemas CRISPR-Cas Tipo III contienen polimerasa y componentes RAMP. Los sistemas Tipo III se pueden dividir además en los sub-tipos III-A y III-B. Los sistemas CRISPR-Cas Tipo III-A han demostrado dirigirse a plásmidos, y las proteínas similares a la polimerasa de los sistemas de Tipo III-A están implicadas en la separación del ADN diana (Marraffini y Sontheimer, 2008, *Science* 322:1843-1845). Los sistemas CRISPR-Cas Tipo III-B también han demostrado dirigirse al ARN diana (Hale *et al.*, 2009, *Cell* 139:945-956).

Por tanto, en algunas realizaciones, el sistema es un sistema CRISPR-Cas Tipo I o un derivado del mismo. En otras realizaciones, el sistema es un sistema CRISPR-Cas Tipo II o un derivado del mismo. En otras realizaciones, el sistema es un sistema CRISPR-Cas Tipo III o un derivado del mismo.

Los elementos clave de un sistema CRISPR-Cas incluyen un ARN guía, *por ejemplo*, un ARNcr, y una proteína Cas. El ARNcr o derivado del mismo contiene una región de nucleótidos específica de la diana complementaria o sustancialmente complementaria a una región del ácido nucleico diana. En algunas realizaciones, el ARNcr o el derivado del mismo contiene una secuencia de ARN que el usuario puede seleccionar que permite la dirección específica de la enzima a un ADN bicatenario complementario. En algunas realizaciones, la secuencia de ARN seleccionable por el usuario contiene 20-50 nucleótidos complementarios o sustancialmente complementarios para una región de la secuencia de ADN diana. En algunas realizaciones, la secuencia de ARN seleccionable por el usuario contiene menos de 20 nucleótidos complementarios o sustancialmente complementarios para una región de la secuencia de ADN diana. Ejemplos de secuencia de ARN seleccionable por el usuario contienen 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, ó 19 nucleótidos complementarios o sustancialmente complementarios para una región de la secuencia de ADN diana. En algunas realizaciones, la región de nucleótidos específicos de la diana del ARNcr tienen el 100% de pares de bases equivalentes con la región del ácido nucleico diana. En algunas realizaciones, la región de nucleótidos específicos de la diana del ARNcr tienen el 90%-100%, 80%-100%, ó 70-100% de pares de bases equivalentes con la región del ácido nucleico diana. En algunas realizaciones, hay un par de bases desapareado entre la región de nucleótidos específica de la diana del ARNcr y la región del ácido nucleico diana. En algunas realizaciones, hay dos pares de bases desapareadas entre la región de nucleótidos específica de la diana del ARNcr y la región del ácido nucleico diana. En algunas realizaciones, hay tres pares de bases desapareadas entre la región de nucleótidos específica de la diana del ARNcr y la región del ácido nucleico diana. En algunas realizaciones, hay cuatro pares de bases desapareadas entre la región de nucleótidos específica de la diana del ARNcr y la región del ácido nucleico diana. En algunas realizaciones, hay cinco pares de bases desapareadas entre la región de nucleótidos específica de la diana del ARNcr y la región del ácido nucleico diana.

En algunas realizaciones, el sistema proporcionado en la presente memoria incluye además un ARNcr transactivador (ARNtracr) o un derivado del mismo.

Los sistemas CRISPR-Cas proporcionados en la presente memoria incluyen sistemas nucleasa modificados genéticamente y/o programados a partir de sistemas CRISPR-Cas que aparecen de manera natural. Los sistemas CRISPR-Cas pueden incluir proteínas Cas modificadas genéticamente y/o mutadas. Los sistemas CRISPR-Cas pueden contener también ARN guía modificado genéticamente y/o programado. En algunas realizaciones, el ARNcr o el derivado del mismo proporcionado en la presente memoria es un polinucleótido que tiene un polinucleótido ARNcr fusionado a un polinucleótido ARNtracr. Un ARN guía único (sgARN) quimérico se describe en Jinek *et al.*, 2012, *Science* 337, 816-821. En una realización, la proteína Cas o la variante de la misma proporcionado en la presente memoria se puede dirigir por un sgARN quimérico hacia cualquier locus genómico seguido por un motivo adyacente de protoespaciador (PAM) 5'-NGG. Por ejemplo, en algunas realizaciones, el ARNcr y el ARNtracr se sintetizan mediante transcripción *in vitro*, empleando un molde de ADN bicatenario sintético que contienen el promotor T7. El ARNtracr tiene una secuencia fija, mientras que la secuencia diana ordena parte de la secuencia de ARNcr. Se mezclan molaridades iguales de ARNcr y ARNtracr y se calientan a 55°C durante 30 segundos. El Cas9 se añade a la misma molaridad a 37°C y se incuba durante 10 minutos con la mezcla de ARN. Después, se añade un exceso de 10-20 veces del complejo de Cas9 al ADN diana. La reacción de unión puede ocurrir dentro de 15 minutos.

En algunas realizaciones, la proteína Cas o la variante de la misma es una proteína Cas9 o una variante de la misma. El complejo Cas9-ARNcr aislado a partir del sistema CRISPR-Cas de *S. thermophilus*, así como el complejo ensamblado *in vitro* a partir de componentes separados demuestran que se une tanto al oligodesoxinucleótido sintético como al ADN plasmídico que tiene una secuencia de nucleótidos complementaria al ARNcr. Se ha demostrado que la Cas9 tiene dos dominios nucleasa- dominios nucleasa/sitios activos para RuvC- y HNH-, y estos dos dominios nucleasa son responsables de la separación de cadenas de ADN opuestas. En algunas realizaciones, la proteína Cas9 se deriva de la proteína Cas9 del sistema CRISPR-Cas de *S. thermophilus*. En algunas realizaciones, la proteína Cas9 es una proteína multi-dominio que tiene aproximadamente 1.409 restos de aminoácidos.

En algunas realizaciones, la proteína Cas9 o la variante de la misma es una variante de nucleasa nula, en la que están mutados ambos dominios de nucleasa/sitios activos para RuvC- y HNH-. Una variante de nucleasa nula de la proteína Cas9 se une a una cadena bicatenaria, pero no separa el ADN, y por tanto, se puede emplear también para el enriquecimiento de ADN específico. En algunas realizaciones, la proteína Cas9 tiene dos dominios nucleasa inactivados con una primera mutación en el dominio que separa la hebra complementaria del ARNcr y una segunda

mutación en el dominio que separa la hebra no complementaria del ARNcr. En algunas realizaciones, la proteína Cas9 tiene una primera mutación D10A y una segunda mutación H840A.

5 En algunas realizaciones, la proteína Cas o la variante de la misma es una proteína de Cascada o una variante de la misma. El complejo en Cascada en *E. coli* reconoce de una manera específica las dianas del ADN bicatenario (dsADN) en una secuencia. El complejo en Cascada de *E. coli* es un complejo de 405-kDa que comprende cinco proteínas asociadas a CRISPR (Cas) esenciales de manera funcional (CasA1B2C6D1E1, también llamada proteína en Cascada) y un ARNcr de 61 nucleótidos. El complejo en Cascada guía al ARNcr hacia las secuencias diana de ADNds para formar pares de bases con la cadena de ADN complementario mientras que la hebra no complementaria se desplaza para formar un bucle en R. La cascada reconoce el ADN diana sin consumir ATP, lo que sugiere que el control del ADN invasor continuo tiene lugar sin inversión de energía. Mattheijs *et al.*, Nature Structural & Molecular Biology, 2011, 18, 529-536.

15 En algunas realizaciones, la proteína Cas o la variante de la misma es una proteína Cas3 o una variante de la misma. La proteína Cas3 de *E. coli* puede catalizar la re-asociación de ARN de manera independiente a ATP con el ADN formando bucles en R, e hibrida pares de bases de ARN en el dúplex de ADN. La proteína Cas3 puede emplear un gARN que es más largo que para Cas9. Howard *et al.*, Biochem J., 2011, 439(1):85-95. Tal ARN más largo puede permitir a otros elementos un acceso más fácil al ADN diana, *por ejemplo*, el acceso de un cebador a prolongar por la polimerasa. Otra ventaja proporcionada por la proteína Cas3 es que la proteína Cas3 no requiere una secuencia PAM como para Cas9, por tanto, proporciona más flexibilidad para dirigir la secuencia deseada. La formación del bucle en R por la Cas3 puede requerir magnesio como co-factor. Howard *et al.*, Biochem J., 2011, 20 439(1):85-95. Por tanto, en algunas realizaciones, el sistema proporcionado en la presente memoria comprende magnesio.

25 Se debe tener en cuenta que en los presentes métodos se puede emplear cualquiera de los sistemas CRISPR-Cas capaces de romper la doble cadena de ácido nucleico y crear una estructura en bucle. Por ejemplo, las proteínas Cas proporcionadas en la presente memoria pueden incluir, pero no se limitan a, las proteínas Cas descritas en Haft *et al.*, PLoS Comput Biol., 2005, 1(6): e60, y Zhang *et al.*, Nucl. Acids Res., 2013, 10.1093/nar/gkt1262. Algunos de estos sistemas CRISPR-Cas requieren que se presente una secuencia específica para estos sistemas CRISPR-Cas para reconocer y unirse a la secuencia diana. Por ejemplo, la Cas9 requiere la presencia de un motivo adyacente de protoespaciador (PAM) 5'-NGG. Por tanto, en algunas realizaciones, una secuencia PAM o una secuencia complementaria a una PAM se modifica genéticamente en el ácido nucleico diana para iniciar la unión de los sistemas CRISPR-Cas al ácido nucleico diana.

35 En algunas realizaciones, el cebador proporcionado en la presente memoria es un oligodesoxirribonucleótido monocatenario. En algunas realizaciones, la longitud de un cebador oscila de aproximadamente 10 a aproximadamente 50 nucleótidos. Ejemplos de longitudes del cebador proporcionadas en la presente memoria son 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, ó 50 nucleótidos.

40 En algunas realizaciones, el cebador proporcionado en la presente memoria tiene el 100% de pares de bases emparejadas con una región del ácido nucleico diana. En algunas realizaciones, el cebador proporcionado en la presente memoria tiene 90%-100%, 80%-100%, o 70%-100% de pares de bases emparejadas con la región del ácido nucleico diana. En algunas realizaciones, hay un par de bases desapareadas entre el cebador y la región del ácido nucleico diana. En algunas realizaciones, hay dos pares de bases desapareadas entre el cebador y la región del ácido nucleico diana. En algunas realizaciones, hay tres pares de bases desapareadas entre el cebador y la región del ácido nucleico diana. En algunas realizaciones, hay cuatro pares de bases desapareadas entre el cebador y la región del ácido nucleico diana. En algunas realizaciones, hay cinco pares de bases desapareadas entre el cebador y la región del ácido nucleico diana.

45 En algunas realizaciones, el cebador proporcionado en la presente memoria está marcado, *por ejemplo*, para el enriquecimiento o la detección de los productos amplificados. En algunas realizaciones, el cebador proporcionado en la presente memoria se marca para la detección mediante espectroscopía, o por medio fotoquímico, bioquímico, o inmunológico. En algunas realizaciones, el cebador se marca con biotina, amina, radiomarcadores (*por ejemplo*, 32P), tintes fluorescentes, reactivos electron-densos, enzimas (como se emplean normalmente en los ELISAS), o biotina.

55 En algunas realizaciones, la polimerasa proporcionada en la presente memoria cataliza la incorporación de nucleótidos para prolongar un extremo 3' hidroxilo de una molécula cebadora frente a una secuencia diana de ácido nucleico. En algunas realizaciones, la reacción por la polimerasa proporcionada en la presente memoria se realiza en condiciones isotérmicas. En algunas realizaciones, la polimerasa es una polimerasa que desplaza la cadena. Ejemplos de polimerasa incluyen, pero no se limita a, ADN polimerasa Bst, ADN polimerasa 9<sup>o</sup>Nm, ADN polimerasa Phi29, ADN polimerasa I (*E. coli*) ADN polimerasa I (grande)m, fragmento (de Klenow), fragmento de Klenow (3'-5' exo-), ADN polimerasa T4, ADN polimerasa T7, ADN polimerasa Deep VentR<sup>TM</sup> (exo-), ADN polimerasa Deep VentR<sup>TM</sup>, DyNAzyme<sup>TM</sup> EXT ADN, ADN Polimerasa DyNAzyme<sup>TM</sup> II Hot Start, ADN Polimerasa Phusion<sup>TM</sup> High Fidelity, ADN Polimerasa Therminator<sup>TM</sup>, ADN Polimerasa Therminator<sup>TM</sup> II, ADN Polimerasa VentR<sup>®</sup>, ADN Polimerasa VentR<sup>®</sup> (exo-), ADN Polimerasa RepliPHI<sup>TM</sup> Phi29, ADN Polimerasa rBst, ADN Polimerasa rBst

(grande), Fragmento (ADN Polimerasa IsoTherm™), MasterAmp™ AmpliThem™, ADN Polimerasa, ADN Polimerasa Taq, ADN polimerasa Tth, ADN polimerasa Tfl, ADN polimerasa Tgo, ADN polimerasa SP6, ADN polimerasa Tbr, ADN polimerasa Beta, y ADN polimerasa ThermoPhi. En algunas realizaciones, la polimerasa se selecciona del grupo que consiste en Bst, Bsu, y Phi29. Como la polimerasa extiende la cadena hibridada, puede ser beneficioso incluir un proteína de unión a cadena sencilla (SSB). La SSB puede estabilizar la hebra desplazada (no molde). Por tanto, en algunas realizaciones, el método proporcionado en la presente memoria puede incluir además una proteína SSB.

Como se indicó, una ventaja proporcionada por las amplificaciones lineales o exponenciales proporcionadas en la presente memoria es que las amplificaciones se pueden realizar bajo temperatura constante o condiciones isotérmicas. Como se emplea en la presente memoria, el término “temperatura constante”, “condiciones isotérmicas”, o “isotérmicamente” se refiere a un conjunto de condiciones de reacción donde la temperatura de la reacción se mantiene esencialmente constante durante el transcurso de la reacción de amplificación. Sin embargo, no es necesario que la temperatura se mantenga a una temperatura exacta. Si el equipamiento empleado para mantener una temperatura elevada permite que la temperatura de la mezcla de reacción varíe unos pocos grados, esto no es perjudicial para la reacción de amplificación, y se puede considerar aún que es una reacción isotérmica. En algunas realizaciones, la reacción de amplificación se realiza a una temperatura constante de entre 20°C a 70°C. En algunas realizaciones, la reacción de amplificación se realiza a una temperatura constante de entre 25°C a 60°C. En algunas realizaciones, la reacción de amplificación se realiza a una temperatura constante de entre 40°C a 55°C. En algunas realizaciones, la reacción de amplificación se realiza a una temperatura constante de entre 30°C a 40°C. Ejemplos de temperaturas de reacción de amplificación incluyen 30°C, 31°C, 32°C, 33°C, 34°C, 35°C, 36°C, 37°C, 38°C, 39°C, 40°C, 41°C, 42°C, 43°C, 44°C, 45°C, 46°C, 47°C, 48°C, 49°C, 50°C, 51°C, 52°C, 53°C, 54°C, 55°C, 56°C, 57°C, 58°C, 59°C, y 60°C. En una realización, la temperatura de la reacción de amplificación es de aproximadamente 37°C.

En algunas realizaciones, el tiempo a la que se realiza la reacción de amplificación puede variar desde, por ejemplo, 1 minuto a varias horas hasta que se alcanza la cantidad de amplificación deseada. En algunas realizaciones, el tiempo de la reacción de amplificación es de 5 minutos a 1 hora. Ejemplos de tiempos de la reacción de amplificación incluyen 5 minutos, 10 minutos, 15 minutos, 20 minutos, 30 minutos, 40 minutos, 50 minutos, y 60 minutos.

En algunas realizaciones, se pueden emplear los presentes métodos en determinadas aplicaciones en donde se desea la detección y/o cuantificación de los productos del ácido nucleico amplificado. La secuencia diana amplificada se puede detectar mediante cualquier método conocido por un experto habitual en la técnica.

En algunas realizaciones, para detectar o cuantificar los productos amplificados se pueden emplear tintes que tiñen específicamente el ADN bicatenario. En algunas realizaciones, los tintes intercalantes muestran mejor fluorescencia en la unión al ADN o ARN. En algunas realizaciones, los tintes pueden ser, por ejemplo, fluoróforos que se intercalan en el ADN o el ARN. Ejemplos de fluoróforos que se intercalan en el ADN o el ARN incluyen, pero no se limitan a, naranja de acridina, bromuro de etidio, tintes Hoechst, picoGreen, yoduro de propidio, SYBR I (un tinte de cianina asimétrico), SYBR II, TOTO (un dímero de naranja de tiazol) y YOYO (un dímero de amarillo de oxazol).

En algunas realizaciones, se pueden detectar los productos amplificados con tamaños específicos mediante electroforesis en gel. En algunas realizaciones, los nucleótidos empleados en la reacción de amplificación se pueden marcar, *por ejemplo*, con biotina. Las secuencias amplificadas marcadas con biotina se pueden capturar empleando avidina unida a una señal que genera enzimas, *por ejemplo*, peroxidasa.

En algunas realizaciones, los nucleótidos marcados se pueden incorporar directamente en la secuencia diana o en los cebadores que contienen secuencias complementarias para la diana de interés. Tales marcadores pueden ser radiactivos y/o fluorescentes en la naturaleza y se pueden resolver de cualquiera de las maneras indicadas en la presente memoria.

Los métodos para detectar y/o controlar de manera continua la amplificación de productos de ácido nucleico son bien conocidos por los expertos en la técnica. *Por ejemplo*, en algunas realizaciones, la producción o la presencia de ácidos nucleicos diana y de secuencias de ácidos nucleicos se puede detectar y controlar mediante Balizas Moleculares. Las Balizas Moleculares son oligonucleótidos con forma de horquilla que contienen un fluoróforo en uno de los extremos y un tinte de extinción en el extremo opuesto. El bucle de la horquilla contiene una secuencia sonda que es complementaria a una secuencia diana y el tallo se forma por re-asociación de las secuencias en el brazo complementario localizado en cada lado de la secuencia sonda. Cuando una baliza molecular se enfrenta a una molécula diana, ocurre la hibridación; la estructura en bucle se convierte en una conformación más rígida que causa la separación del fluoróforo y las moléculas silenciadoras conducen la fluorescencia. Tyagi *et al.*, Nature Biotechnology, 1996, 303-308. Por tanto, la generación de fluorescencia indica la síntesis del producto amplificado previsto. Para otro ejemplo, en algunas realizaciones, la producción o presencia de secuencias de ácidos nucleicos y ácidos nucleicos diana se puede detectar o controlar mediante transferencia de energía de resonancia fluorescente (FRET). La FRET es una herramienta útil para cuantificar interacciones dinámicas moleculares, *por ejemplo*, en interacciones de ADN-ADN. Para controlar la producción de un producto específico se puede marcar una sonda con una molécula donadora en un extremo y una molécula aceptora en el otro. La hibridación de la sonda-diana produce

- un cambio en la distancia u orientación del donador y del aceptor y se observa un cambio de FRET. Lakowicz, "Principles of Fluorescence Spectroscopy", Plenum Publishing Corporation, 2ª edición (1 de Julio, 1999). Otros ejemplos de métodos para detectar y/o controlar de manera continua la amplificación de los productos de ácidos nucleicos incluyen, pero no se limitan a, la Espectroscopía de Masas, electroforesis capilar en gel, secuenciación, y
- 5 varios métodos de captura superficial como se conocen por el experto en la técnica. En algunas realizaciones, llevar a cabo la reacción de amplificación con cantidades asimétricas de los cebadores, *es decir*, con un cebador presente en mayor concentración que el otro, permitiría la amplificación con preferencia hacia un extremo. Un exceso de una cadena simple puede facilitar la detección mediante tales modalidades como las Balizas Moleculares, que detectan ssADN.
- 10 Los métodos de amplificación proporcionados en la presente memoria se pueden emplear en varias aplicaciones. Por ejemplo, en algunas realizaciones, el método proporcionado en la presente memoria se puede emplear para aislar fragmentos de ADN a partir de ADN genómico mediante amplificación selectiva de una región específica de ADN.
- 15 Como otro ejemplo, en algunas realizaciones, los métodos de amplificación proporcionados en la presente memoria se pueden emplear para diagnosticar varias enfermedades en base a la producción o presencia de un ácido nucleico diana vinculado a una enfermedad específica. En algunas realizaciones, los métodos proporcionados en la presente memoria se pueden emplear para el diagnóstico temprano de enfermedades malignas, *por ejemplo*, leucemia y linfomas. En otras realizaciones, los métodos proporcionados en la presente memoria se pueden emplear directamente sobre muestras de ADN genómico para amplificar y detectar la translocación específica de células malignas.
- 20 En otras realizaciones, los métodos proporcionados en la presente memoria se pueden emplear para el diagnóstico de enfermedades infecciosas, que incluyen aquellas causadas por bacterias o virus. Por ejemplo, los presentes métodos se pueden emplear para detectar agentes infecciosos y distinguir cepas no patogénicas de las cepas patogénicas en virtud de secuencias de ácidos nucleicos específicas. En algunas realizaciones, los métodos de amplificación proporcionados en la presente memoria pueden amplificar e identificar microorganismos no cultivables o de crecimiento lento, *por ejemplo*, micobacterias, bacterias anaerobias, o virus a partir de ensayos de cultivo tisular y modelos animales. En otras realizaciones, los métodos de amplificación proporcionados en la presente memoria pueden amplificar e identificar ADN viral en una muestra a partir de un individuo.
- 25 En otras realizaciones, los presentes métodos de amplificación se pueden emplear para generar materiales de ácido nucleico para aumentar otros procesos. Por ejemplo, en algunas realizaciones, los presentes métodos de amplificación se pueden emplear para generar sondas de hibridación por hibridación Southern o Northern y clonación de ADN, que requieren cantidades relativamente grandes de ADN, que representan una región de ADN específica. Como otro ejemplo, los métodos proporcionados en la presente memoria se pueden emplear para la secuenciación de ácido nucleico.
- 30 En una realización específica, los métodos de amplificación de ácido nucleico mediado por CRISPR-Cas proporcionados en la presente memoria se pueden emplear para amplificar una biblioteca de fragmentos de ácido nucleico, *por ejemplo*, generada utilizando métodos de preparación de biblioteca y/o kits disponibles en Illumina, Inc. (San Diego, CA).
- 35 En algunas realizaciones, los métodos proporcionados en la presente memoria se pueden emplear para amplificar una biblioteca de fragmentos de ácido nucleico generado a partir de ADN genómico. En una realización específica, la biblioteca de fragmentos de ácido nucleico se genera empleando tagmentación. En una realización específica, la biblioteca de fragmentos de ácido nucleico se genera a partir de ADN genómico empleando métodos de preparación de biblioteca y kits de Illumina Nextera (disponibles en Illumina, Inc., San Diego, CA).
- 40 Por tanto, en algunas realizaciones, el método proporcionado en la presente memoria incluye además aplicar al menos una transposasa y al menos una composición de extremo del transposón que contiene una hebra transferida a una muestra que contiene ácido nucleico diana bajo condiciones donde el ácido nucleico diana y la composición de extremo del transposón experimentan una reacción de transposición para generar una mezcla, en donde el ácido nucleico diana se fragmenta para generar una pluralidad de fragmentos de ácido nucleico diana, y por tanto, incorpora una secuencia cebadora universal en cada uno de la pluralidad de fragmentos de ácido nucleico diana, en donde el ARNcr o el derivado del mismo contiene una región de nucleótidos específica de la diana complementaria para una región del cebador universal.
- 45 En algunas realizaciones, el ácido nucleico diana se somete a una transposasa mediada por fragmentación que da como resultado la fragmentación del ácido nucleico diana y la unión de los adaptadores al extremo 5' de ambas hebras de los fragmentos de ADN bicatenario. Opcionalmente, el ácido nucleico se puede fragmentar y se pueden añadir adaptadores a los extremos 5' y 3' empleando tagmentación o transposición como se describe en la Publicación U.S. N° 2010/0120098. En resumen, una reacción de transposición es una reacción en donde se insertan uno o más transposones en los ácidos nucleicos en sitios aleatorios. Componentes esenciales en una reacción de transposición son una transposasa y oligonucleótidos de ADN que muestran secuencias de nucleótidos de un transposón, incluyendo la secuencia del transposón y su complemento (la secuencia de extremo del
- 55

transposón no transferido) así como otros componentes necesarios para formar una transposición o complejo del transposoma. Los oligonucleótidos de ADN pueden incluir además secuencias adicionales (*por ejemplo*, un adaptador o secuencias cebadoras) como se necesite o desee. Ejemplos de complejos de transposición, adecuados para el empleo en los métodos proporcionados en la presente memoria, incluyen, pero no se limitan a, aquellos formados por una transposasa Tn5 hiperactiva y un extremo de transposón de tipo Tn5 o mediante una transposasa MuA y un extremo de transposón Mu que comprende secuencias de extremo R1 y R2 (*véase, por ejemplo*, Goryshin y Reznikoff, *J. Biol. Chem.* 273:7367, 1998; y Mizuuchi, *Cell* 35:785, 1983; Savilahti *et al.*, *EMBO J.* 14:4893, 1995. Sin embargo, en los métodos proporcionados se puede emplear cualquier sistema que sea capaz de insertar un extremo de transposón con la suficiente eficacia como para marcar ácidos nucleicos diana para su propósito destinado. Otros ejemplos de sistemas de transposición conocidos que se podrían emplear en los métodos proporcionados incluyen, pero no se limitan a, Tn552 de *Staphylococcus aureus*, Tyl. Transposón Tn7, Tn/O y IS10, transposasa Marina, Tel, Elemento P, Tn3, secuencias de inserción bacteriana, retrovirus, y el retrotransposón de levadura (*véase, por ejemplo*, Colegio *et al.*, 2001, *J. Bacteriol.* 183:2384-8; Kirby *et al.*, 2002, *Mol. Microbiol.* 43:173-86; Devine y Boeke, 1994, *Nucleic Acids Res.*, 22:3765-72; Solicitud de Patente Internacional N° WO 95/23875; Craig, 1996, *Science* 271:1512; Craig, 1996, Review in: *Curr Top Microbiol Immunol.* 204:27-48; Kleckner *et al.*, 1996, *Curr Top Microbiol Immunol.* 204:49-82; Lampe *et al.*, 1996, *EMBO J.* 15:5470-9; Plasterk, 1996, *Curr Top Microbiol Immunol.* 204:125-43; Gloor, 2004, *Methods Mol. Biol.* 260:97-114; Ichikawa y Ohtsubo, 1990, *J Biol.Chem.* 265:18829-32, Ohtsubo y Sekine, 1996, *Curr. Top. Microbiol. Immunol.* 204:1-26, Brown *et al.*, 1989, *Proc Natl Acad Sci EE.UU.* 86:2525-9; Boeke y Corces, 1989, *Annu Rev Microbiol.* 43:403-34. En algunas realizaciones, el método de la presente divulgación comprende además eliminar la enzima transposasa y añadir mediante PCR fragmentos a los extremos de ADN adaptado.

El término “tagmentación”, “tagmento” o “tagmentar” como se emplea en la presente memoria, se refiere a transformar un ácido nucleico, *por ejemplo*, un ADN, en moldes modificados por un adaptador en disolución lista para la formación de grupos y secuenciación mediante el empleo de transposasa mediada por fragmentación y marcaje. Este proceso a menudo implica la modificación del ácido nucleico mediante un complejo de transposoma que comprende una enzima transposasa complejada con adaptadores que comprenden una secuencia de extremo del transposón. La tagmentación da como resultado la fragmentación simultánea del ácido nucleico y la unión de los adaptadores a los extremos 5' de ambas hebras de los fragmentos dobles. Después de una etapa de purificación para eliminar la enzima transposasa, se añaden mediante PCR secuencias adicionales a los extremos de los fragmentos adaptados.

El término “complejo de transposoma”, como se emplea en la presente memoria, se refiere a una enzima transposasa no unida de manera covalente a un ácido nucleico bicatenario. Por ejemplo, el complejo puede ser una enzima transposasa preincubada con ADN de un transposón bicatenario bajo condiciones que soporten la formación no covalente del complejo. El ADN del transposón bicatenario puede incluir, sin limitación, ADN Tn5, una porción de ADN Tn5, una composición con el extremo del transposón, una mezcla de composiciones del transposón final u otros ADNs bicatenarios capaces de interactuar con una transposasa, tal como la Tn5 transposasa hiperactiva.

Una “transposasa” significa una enzima que es capaz de formar un complejo funcional con una composición que contiene un extremo de transposón (*por ejemplo*, transposones, extremos de transposón, composiciones de extremo de transposón) y catalizar la inserción o transposición de la composición que contiene el extremo de transposón en un ácido nucleico diana bicatenario con el que se incubaba, por ejemplo, en una reacción de transposición *in vitro*. Una transposasa como se presenta en la presente memoria puede incluir también integrasas a partir de retrotransposones y retrovirus. Las transposasas, transposomas y complejos de transposomas generalmente son bien conocidos por los expertos en la técnica, como se ejemplifica mediante la divulgación de la Patente EE.UU. 2010/0120098. Aunque muchas realizaciones descritas en la presente memoria se refieren a una transposasa Tn5 y/o una transposasa Tn5 hiperactiva, se apreciará que en la presente invención se puede emplear cualquier sistema de transposición que sea capaz de insertar un extremo de transposón con la suficiente eficacia para marcar y fragmentar un ácido nucleico 5' con el propósito destinado. En realizaciones particulares, un sistema de transposición preferido es capaz de insertar el extremo de transposón de una manera aleatoria o de una manera casi aleatoria para marcar y fragmentar el ácido nucleico diana 5'.

Como se emplea en la presente memoria, el término “reacción de transposición” se refiere a una reacción en donde uno o más transposones se insertan en ácidos nucleicos diana, *por ejemplo*, en sitios aleatorios o en sitios casi aleatorios. Componentes esenciales en una reacción de transposición son una transposasa y oligonucleótidos de ADN que muestran las secuencias de nucleótidos de un transposón, que incluyen la secuencia del transposón transferido y su complemento (secuencia del extremo del transposón no transferido) así como otros componentes necesarios para formar una transposición funcional o un complejo de transposoma. Los oligonucleótidos de ADN pueden comprender además secuencias adicionales (*por ejemplo*, secuencias adaptadoras o cebadoras) como se desee o necesite. En algunas realizaciones, el método proporcionado en la presente memoria se ejemplifica mediante el empleo de un complejo de transposición formado por una transposasa Tn5 hiperactiva y un extremo de transposón tipo Tn5 (Goryshin y Reznikoff, 1998, *J. Biol. Chem.*, 273:7367) o mediante una transposasa MuA y un extremo de transposón Mu que comprende secuencias finales R1 y R2 (Mizuuchi, *Cell*, 35:785; Savilahti *et al.*, *EMBO J.*, 14:4893). Sin embargo, en la presente invención se puede emplear cualquier sistema de transposición que sea capaz de insertar un extremo de transposón de una manera aleatoria o casi aleatoria con la suficiente eficacia como para marcar y fragmentar un ADN 5' con el propósito destinado. Ejemplos de sistemas de

transposición conocidos en la técnica que se pueden emplear para los presentes métodos incluyen, pero no se limitan a, *Staphylococcus aureus* Tn552 (Colegio *et al.*, 2001, *J. Bacteriol.* 183:2384-8; Kirby *et al.*, 2002, *Mol. Microbiol.* 43:173-86), Tyl (Devine y Boeke, 1994, *Nucleic Acids Res.*, 22:3765-72; Solicitud de Patente Internacional N° WO 95/23875), Transposón Tn7 (Craig, 1996, *Science.* 271:1512; Craig, 1996, Review in: *Curr Top Microbiol Immunol.*, 204:27-48), Tn10 y IS10 (Kleckner *et al.*, 1996, *Curr Top Microbiol Immunol.*, 204:49-82), transposasa Marina (Lampe *et al.*, 1996, *EMBO J.*, 15:5470-9), Tci (Plasterk, *Curr Top Microbiol Immunol.*, 204:125-43), Elemento P (Gloor, 2004, *Methods Mol. Biol.*, 260:97-114), TnJ (Ichikawa y Ohstsubo, 1990, *J Biol.Chem.* 265:18829-32), secuencias de inserción bacteriana (Oshtsubo y Sekine, 1996, *Curr. Top. Microbiol. Immunol.* 204:1-26), retrovirus (Brown *et al.*, 1989, *Proc Natl Acad Sci EE.UU.* 86:2525-9), y retrotransposón de levadura (Boeke y Corces, 1989, *Annu Rev Microbiol.* 43:403-34). El método para insertar un extremo del transposón en una secuencia diana se puede llevar a cabo *in vitro* empleando cualquier sistema de transposón adecuado para el que está disponible un sistema de transposición *in vitro* adecuado o que se puede desarrollar en base al conocimiento en la técnica. En general, un sistema de transposición *in vitro* adecuado para emplear en los métodos proporcionados en la presente memoria requiere, como mínimo, una enzima transposasa de la pureza suficiente, suficiente concentración, y suficiente actividad de transposición *in vitro* y un extremo de transposón con el que la transposasa forma un complejo funcional con la transposasa respectiva que es capaz de catalizar la reacción de transposición. Transposasas y secuencias de extremo de transposón adecuadas que se pueden emplear en la invención incluye, pero no se limitan a, secuencias de extremo de transposón de tipo salvaje, derivado o mutante que forman un complejo con una transposasa elegida de entre una transposasa de tipo salvaje, derivada o mutante de la transposasa.

El término "extremo de transposón" (TE) se refiere a un ácido nucleico bicatenario, *por ejemplo*, un ADN bicatenario, que muestra sólo las secuencias de nucleótidos (las "secuencias del extremo de transposón") que son necesarias para formar el complejo con la transposasa o la enzima integrasa que es funcional en una reacción de transposición *in vitro*. En algunas realizaciones, un extremo de transposón es capaz de formar un complejo funcional con la transposasa en una reacción de transposición. Como ejemplos no limitantes, extremos de transposones pueden incluir el transposón de extremo 19-bp externo ("OE"), el transposón de extremo interno ("IE"), o el transposón de "extremo en mosaico" ("ME") reconocido por una transposasa Tn5 de tipo salvaje o mutante, o el extremo de transposón R1 y R2 como se establece en la divulgación de EE:UU. 2010/0120098.

Los extremos del transposón pueden incluir cualquier ácido nucleico o análogo de ácido nucleico adecuado para formar un complejo funcional con la transposasa o la enzima integrasa en una reacción de transposición *in vitro*. Por ejemplo, el transposón de extremo puede incluir ADN, ARN, bases modificadas, bases no naturales, estructuras modificadas, y pueden incluir cortes en una o en ambas hebras. Aunque el término "ADN" se usa a veces en la presente divulgación en relación con la composición de los transposones de extremo, se entenderá que en un transposón de extremo se puede utilizar cualquier ácido nucleico o análogo de ácido nucleico.

A través de una reacción de transposición *in vitro*, se marcan los fragmentos de ácido nucleico diana en el extremo 5'. En algunas realizaciones, el método proporcionado en la presente memoria incluye además las etapas para incorporar un marcador en el extremo 3' de los fragmentos de ácido nucleico marcados en 5', para producir una biblioteca de fragmentos de ácido nucleico di-marcados. Añadir marcadores en el extremo 3' se puede realizar a través de varios métodos, *por ejemplo*, mediante el empleo de ADN polimerasa, transferasa terminal, y/o ligasa como se describe en WO 2010/048605.

En algunas realizaciones, los fragmentos de ácido nucleico di-marcados se generan mediante el empleo de una polimerasa, *por ejemplo*, una ADN polimerasa, con actividad de desplazamiento de la hebra o 5' nucleasa. En algunas realizaciones, el método proporcionado en la presente memoria incluye incubar la población de fragmentos de ácido nucleico 5'-marcados re-asociados con una ADN polimerasa que tiene actividad de desplazamiento de la hebra o 5'-nucleasa bajo condiciones sin termociclado, y en donde los fragmentos de ácido nucleico 5'-marcados no están desnaturalizados, en donde la ADN polimerasa extiende el extremo 3' de cada una de las hebras de los fragmentos de ácido nucleico 5'-marcados re-asociados empleando la hebra complementaria como un molde y que desplaza o digiere la hebra no transferida, generando de este modo la biblioteca de fragmentos de ADN bicatenario di-marcado.

En otras realizaciones, los fragmentos de ácido nucleico 5'-marcados se incuban con una ADN polimerasa que consiste en una transferasa terminal y al menos un sustrato para la transferasa terminal bajo condiciones y durante el tiempo suficiente en donde la transferasa terminal une el segundo marcador al extremo 3' de los fragmentos de ácido nucleico 5'-marcados, generando de este modo una biblioteca de fragmentos de ácido nucleico di-marcados. En algunas realizaciones, el extremo 3' del transposón no transferido que compone la composición del extremo de transposón, se bloquea (*por ejemplo*, mediante el empleo de un transposón de extremo no transferido que tiene un dideoxinucleótido o un 3'-O-metil-nucleótido como el nucleótido terminal 3').

En otras realizaciones, los fragmentos de ácido nucleico di-marcados se generan mediante el empleo de una ligasa dependiente del molde y un oligonucleótido de marcaje de ligación. En algunas realizaciones, los fragmentos de ácido nucleico 5'-marcados se incuban con una ligasa de ADN dependiente del molde y un oligodesoxinucleótido de marcaje de ligación que tienen una porción 3' y una porción 5', en donde la porción 3' muestra un segundo marcador que muestra cualquier secuencia que se desee unir al extremo 3' de los fragmentos de ADN 5'-marcados y la porción 5' tienen un grupo 5'-monofosfato y muestra una secuencia aleatoria, bajo condiciones y durante el tiempo

suficiente en donde el segundo marcador se une a los fragmentos de ADN 5'-marcados re-asociados, generando de este modo una biblioteca de fragmentos de ADN que comprende fragmentos de ADN di-marcados re-asociados.

5 En algunas realizaciones, los fragmentos de ácido nucleico generados contienen secuencias universales en los dos extremos de los fragmentos de ácido nucleico. Las secuencias universales en los dos extremos de los fragmentos de ácido nucleico, *por ejemplo*, en una biblioteca, se pueden marcar mediante el sistema CRISPR-Cas según los métodos proporcionados en la presente memoria, y como tales fragmentos se pueden amplificar, *por ejemplo*, en una reacción de formación de grupos.

10 Tales secuencias universales se pueden introducir en una biblioteca en los dos extremos de los fragmentos de ácido nucleico mediante PCR o una reacción de transposón Nextera. En algunas realizaciones, después de generar una biblioteca de fragmentos de ácido nucleico marcados, los fragmentos de ácido nucleico marcados se pueden amplificar, *por ejemplo*, empleando la reacción en cadena de polimerasa de ciclo limitado (PCR), para introducir otras secuencias de extremo o adaptadores, *por ejemplo*, índices, cebadores universales y otras secuencias requeridas para la formación y la secuenciación del grupo. En una realización específica, se realiza una amplificación PCR de ciclo limitado para añadir un índice 1 (P7) y un índice 2 (P5) (disponible en Illumina, Inc, San Diego, CA) a los dos extremos de los fragmentos de ácido nucleico.

15 En algunas realizaciones, tal amplificación se realiza para una biblioteca de fragmentos de ácido nucleico 5'-marcado. En algunas realizaciones, tal amplificación se realiza para una biblioteca de fragmentos de ácido nucleico di-marcado. Ejemplos de métodos de amplificación incluyen la reacción en cadena de la polimerasa (PCR), reacción de amplificación por desplazamiento de la hebra, reacción de amplificación en ciclo rodante, reacción en cadena de la ligasa, reacción de amplificación mediada por transcripción, y reacción de amplificación mediada por bucle.

20 En algunas realizaciones, el método proporcionado en la presente memoria incluye amplificar la biblioteca de fragmentos de ácido nucleico di-marcados empleando un PCR. En algunas realizaciones, la presente memoria proporcionada incluye amplificar la biblioteca de fragmentos de ácido nucleico di-marcado empleando el método de amplificación mediado por Cas9 proporcionado en la presente memoria. En algunas realizaciones, el método proporcionado en la presente memoria emplea una amplificación PCR de cebador único de una biblioteca de fragmentos de ADN di-marcados. En algunas realizaciones, la etapa de amplificar los fragmentos de ADN di-marcados incluye emplear una ADN polimerasa y al menos un cebador que es complementario para el segundo marcador. En algunas realizaciones, la etapa de amplificar la biblioteca de fragmentos de ADN di-marcados incluye amplificar la biblioteca de fragmentos de ADN marcados mediante PCR empleando sólo un oligodesoxirribonucleótido que muestra la secuencia de al menos una porción de la hebra transferida como un cebador PCR y los fragmentos de ADN di-marcados como moldes. En algunas realizaciones, el cebador contiene una porción 5' que contiene una secuencia adicional, *por ejemplo*, una secuencia adaptadora.

25 En algunas realizaciones, se emplean dos cebadores PCR diferentes, cada uno de los cebadores PCR muestra la secuencia de al menos una porción del extremo de transposón transferido que compone la composición de extremo del transposón. En algunas realizaciones, cada cebador PCR incluye una porción 3' y una porción 5', en donde la porción 3' muestra la secuencia de extremo del respectivo transposón transferido y la porción 5' muestra la secuencia del dominio marcador respectivo o un adaptador para un propósito en particular (*por ejemplo*, un dominio/adaptador de marcaje de secuenciación o un dominio/adaptador marcador de amplificación, y opcionalmente un dominio/adaptador marcador de dirección para la secuenciación o amplificación de la siguiente generación). Por ejemplo, cuando se usa una composición única del extremo de transposón en la reacción de transposición *in vitro* para generar la biblioteca de fragmentos de ADN di-marcados empleando una ADN polimerasa que tiene una actividad de desplazamiento de la hebra o 5' nucleasa, los fragmentos de ADN di-marcados se pueden amplificar mediante PCR empleando dos cebadores PCR diferentes. Cada cebador PCR contiene una porción 3' y una porción 5', en donde la porción 3' muestra la secuencia de extremo del respectivo transposón transferido y la porción 5' muestra la secuencia de un dominio/adaptador marcador respectivo para un propósito en particular (*por ejemplo*, un dominio/adaptador de marcaje de secuenciación o un dominio/adaptador marcador de amplificación, y opcionalmente un dominio/adaptador marcador de dirección para la secuenciación o amplificación de la siguiente generación). En algunas realizaciones, la porción 5' de cada cebador PCR es diferente para cada cebador, y tales secuencias de ambos extremos del producto PCR son diferentes. Por ejemplo, un extremo contiene un índice y/o secuencia cebadora universal, y el otro extremo contienen un índice y/o secuencia cebadora universal diferente.

35 En algunas realizaciones, los dos extremos de los fragmentos de ácido nucleico di-marcados se originan de dos secuencias de hebra transferida diferentes. Por ejemplo, en algunas realizaciones, se pueden usar dos transposones diferentes en la reacción de transposición *in vitro*, y cada uno de los dos transposomas contiene la misma transposasa pero diferente composición de extremo de transposón. En algunas realizaciones, se usan dos transposasas diferentes, y las dos transposasas diferentes contiene cada una la misma transposasa y las composiciones de extremo de transposón contienen diferentes hebras transferidas. En algunas realizaciones, se usan dos transposomas diferentes, y cada uno de los dos transposomas incluyen enzimas transposasas diferentes y composiciones de extremo de transposón diferentes, cada una de las cuales forma un complejo funcional con la transposasa respectiva. En algunas realizaciones, en donde se emplean dos composiciones de extremo de transposón diferentes en la reacción de transposición *in vitro*, y la biblioteca de fragmentos de ácido nucleico

monocatenario di-marcados se genera empleando una ADN polimerasa que tiene actividad de desplazamiento de la hebra o 5' nucleasa, el primer marcador muestra la secuencia de la hebra transferida de una composición de extremo de transposón y el segundo marcador muestra la secuencia de la hebra no transferida de la otra composición de extremo de trasposón.

5 En las realizaciones anteriormente mencionadas y en otras realizaciones en donde dos hebras transferidas se unen al extremo 5' de cada hebra opuesta del ácido nucleico bicatenario, el método proporcionado en la presente memoria puede incluir además la etapa de amplificar los fragmentos de ácido nucleico di-marcados mediante PCR empleando dos cebadores PCR diferentes. Uno de los cebadores PCR muestra la secuencia de al menos una porción de una hebra transferida que compone una composición de extremo de transposón, y el otro cebador PCR muestra la secuencia de al menos una porción de la otra hebra transferida que compone la otra composición de extremo de transposón.

10 En algunas realizaciones en donde se emplean dos cebadores, cada cebador PCR contiene una porción 3' y una porción 5', en donde la porción 3' muestra la secuencia de extremo del respectivo transposón transferido y la porción 5' muestra la secuencia de un dominio/adaptador marcador respectivo para un propósito en particular (*por ejemplo*, un dominio marcador de secuenciación o un dominio marcador de amplificación, y opcionalmente un dominio marcador de dirección para la secuenciación o amplificación de la siguiente generación). En algunas realizaciones, la porción 5' de cada cebador PCR es diferente para cada cebador, y tales secuencias de ambos extremos del producto PCR son diferentes. En algunas realizaciones, la porción del primer cebador PCR o la porción 5' del segundo cebador PCR, o las porciones 5' de tanto el primer como el segundo cebador PCR contiene el primer o el segundo marcador/adaptador, respectivamente, para la generación de moldes para la secuenciación de la siguiente generación para una plataforma de secuenciación particular (*por ejemplo*, marcadores de secuenciación para una plataforma de secuenciación Illumina Nextera). En algunas realizaciones, la porción 5' del primer cebador PCR o la porción 5' del segundo cebador PCR contiene adicionalmente un dominio/adaptador marcador de dirección u otro dominio/adaptador marcador para un propósito particular.

25 Están disponibles una amplia variedad de enzimas y kits para realizar la reacción de amplificación mediante PCR como se conoce por los expertos en la técnica. Por ejemplo, en algunas realizaciones, la amplificación PCR se realiza empleando bien el Sistema FAILSAFE™ PCR o bien el Sistema MASTERAMP™ Extra-Long de EPICENTRE Biotechnologies, Madison, WI, como se describe por el fabricante. Sin embargo, la presente divulgación no se limita al empleo de estos productos o condiciones para la reacción de amplificación y se puede emplear cualquier ADN polimerasa termoestable y mezcla de reacción adecuadas que permitan la amplificación de la secuencia entre el cebador que re-asocia la secuencia diana y el cebador que re-asocia el transposón.

35 El método proporcionado en la presente memoria no se limita al empleo del PCR para amplificar la biblioteca de fragmentos de ácido nucleico marcado. En realizaciones de la presente invención se puede emplear cualquier método de amplificación adecuado (*por ejemplo*, amplificación en círculo rodante, amplificación del ribocebador (*por ejemplo*, la Patente U.S. Nº 7.413.857), ICAN (amplificación de ácidos nucleicos iniciada por cebadores químicos e isotérmicos), UCAN, ribospia, marcado terminal (Solicitud de Patente U.S. Nº 20050153333), una amplificación de aARN de tipo Eberwine o amplificación de desplazamiento de cadena) que amplifica la misma secuencia, y genera una composición apropiada y una cantidad de producto de amplificación para el fin pretendido. Por ejemplo, algunos métodos de desplazamiento de cadena que se pueden usar se describen en la Publicación de Patente PCT Nºs WO 02/16639; WO 00/56877; y AU 00/29742; de Takara Shuzo Company, Kioto, Japón; Patentes U.S. Nºs 5.523.204; 5.536.649; 5.624.825; 5.631.147; 5.648.211; 5.733.752; 5.744.311; 5.756.702; y 5.916.779 de Becton Dickinson and Company; Patentes U.S. Nºs 6.238.868; 6.309.833; y 6.326.173 de Nanogen/Becton Dickinson Partnership; Patentes U.S. Nºs 5.849.547; 5.874.260; y 6.218.151 de Bio Merieux; Patentes U.S. Nºs 5.786.183; 6.087.133; y 6.214.587 de Gen-Probe, Inc.; Patente U.S. Nº 6.063.604 de Wick *et al.*; Patente U.S. Nº 6.251.639 de Kurn; Patente U.S. Nº 6.410.278; y la publicación PCT Nº WO 00/28082 de Eiken Kagaku Kabushiki Kaishi, Tokio, Japón; Patentes U.S. Nºs 5.591.609; 5.614.389; 5.773.733; 5.834.202; y 6.448.017 de Auerbach; y Patentes U.S. Nºs 6.124.120; y 6.280.949 de Lizardi. En algunas realizaciones, la amplificación mediada por Cas proporcionada en la presente memoria se emplea para amplificar la biblioteca de los fragmentos de ácido nucleico diana.

50 En algunas realizaciones, cuando el ácido nucleico diana tiene una secuencia cebadora universal, *por ejemplo*, como se genera empleando los métodos proporcionados anteriormente, un ARN guía, *por ejemplo*, un ARNcr, puede dirigir la secuencia cebadora universal de los fragmentos de ácido nucleico en la biblioteca para amplificar la biblioteca de fragmentos de ácido nucleico de manera isotérmica.

55 La Figura 1 ilustra un método para amplificar fragmentos de ácido nucleico generados mediante la preparación de la biblioteca Nextera (disponible en Illumina Inc., San Diego, CA) según la presente divulgación. En algunas realizaciones, las proteínas Cas proporcionadas en la presente memoria requieren una secuencia PAM para reconocer una diana. Por ejemplo, la proteína Cas9 requiere una secuencia del motivo NGG que esté adyacente a la secuencia diana. Se puede incorporar una secuencia complementaria a la secuencia PAM en los cebadores de flujo (*por ejemplo*, P5 y P7) que se añaden a los insertos de biblioteca para dar como resultado cebadores P5 y P7 modificados en PAM capaces de dirigirse a las secuencias que contienen PAM. En la Figura 1A (SEQ ID No.1 y SEQ ID No.2) se muestran las secuencias de los cebadores universales estándar P5 y P7. Las secuencias de los cebadores P5 y P7 modificados en PAM se muestran también en la Figura 1A (SEQ ID No.3 y SEQ ID No.4). Por

tanto, en algunas realizaciones, los cebadores P5 y P7 modificados en PAM se añaden a los fragmentos de ácidos nucleicos en la biblioteca. En una realización específica, los cebadores P5 y P7 modificados en PAM se añaden en los fragmentos de ácido nucleico empleando un PCR de ciclo limitado.

5 Como se muestra en la Figura 1B-C, los fragmentos de ADN (*por ejemplo*, amplicones Nextera en disolución) contienen secuencias del cebador P5 y P7 (SEQ ID No.1 y SEQ ID No.2) en los extremos del amplicón. Se añade un sistema CRISPR-Cas9 que contiene un ARN guía de SEQ ID NO:7 que se dirige a una región de la secuencia del cebador P5 (véase la Figura 1B), y un sistema CRISPR-Cas9 que contiene un ARN guía de SEQ ID No.8 que se dirige a una región de la secuencia del cebador P7 (véase la Figura 1C). El sistema CRISPR-Cas9 abre algunas regiones del ADN bicatenario para crear estructuras en bucle R cerca de los dos extremos del fragmento de ADN que se unen al ARNcr de las secuencias cebadoras. Después, se pueden emplear los cebadores P5 y P7 con PAM modificado truncado (SEQ ID No.5 y SEQ ID No.6) para amplificar el fragmento de ADN.

10 Por tanto, en algunas realizaciones, la primera hebra del ácido nucleico bicatenario diana contienen una secuencia universal, y en donde el ARNcr o el derivado del mismo contienen una secuencia complementaria a una región de la secuencia universal. En algunas realizaciones, el cebador contiene una secuencia de una región de la secuencia universal.

15 En algunas realizaciones, el cebador universal tiene una secuencia de SEQ ID No.3. En algunas realizaciones, el ARNcr contiene una secuencia de SEQ ID No.7. En algunas realizaciones, el cebador contiene una secuencia de SEQ ID No.5.

20 En algunas realizaciones, la secuencia del cebador universal tiene una secuencia de SEQ ID No.4. En algunas realizaciones, el ARNcr contiene una secuencia de SEQ ID No.8. En algunas realizaciones, el cebador contiene una secuencia de SEQ ID No.6.

25 En algunas realizaciones, la primera hebra del ácido nucleico bicatenario diana contiene una primera secuencia universal, y en donde el ARNcr o el derivado del mismo del primer sistema contiene una secuencia complementaria para una región de la primera secuencia universal, y la segunda hebra del ácido nucleico bicatenario diana contiene una segunda secuencia universal, y en donde el ARNcr o el derivado del mismo del segundo sistema contiene una secuencia complementaria para una región de la segunda secuencia universal. En algunas realizaciones, el primer cebador contiene una secuencia de una región de la primera secuencia universal, y el segundo cebador contiene una secuencia de una región de la segunda secuencia universal. En una realización específica, la primera secuencia universal tiene una secuencia de SEQ ID No.3, el ARNcr o el derivado del mismo del primer sistema contiene una secuencia de SEQ ID No.7, y el primer cebador contiene una secuencia de SEQ ID No.5, y la segunda secuencia universal tiene una secuencia de SEQ ID No.4, el ARNcr o el derivado del mismo del segundo sistema contiene una secuencia de SEQ ID No.8, y el segundo cebador contiene una secuencia de SEQ ID No.6.

30 Los métodos proporcionados en la presente memoria se pueden emplear en amplificación isotérmica para secuenciar, *por ejemplo*, en una amplificación en grupo desarrollada por Illumina, Inc. (San Diego, CA). En algunas realizaciones, el ácido nucleico diana de los presentes métodos se puede inmovilizar sobre una superficie para la amplificación. Por ejemplo, en algunas realizaciones, los fragmentos de ácido nucleico inmovilizados se amplifican empleando metodologías de amplificación en grupo como se ejemplifica mediante las divulgaciones de las Patentes EE.UU. N<sup>os</sup> 7.985.565 y 7.115.400. Los materiales de las Patentes EE.UU.N<sup>os</sup> 7.985.565 y 7.115.400 describen métodos de amplificación de ácido nucleico en fase sólida que permiten la amplificación de productos a inmovilizar sobre una superficie sólida para formar matrices comprendidas en grupos o "colonias" de moléculas de ácido nucleico inmovilizado. Cada grupo o colonia sobre tal matriz se forma a partir de una pluralidad de hebras de polinucleótidos idénticos inmovilizados y una pluralidad de hebras de polinucleótidos complementarios inmovilizados idénticas. Las matrices así formadas se refieren generalmente en la presente memoria como "matrices agrupadas". Los productos de reacciones de amplificación en fase sólida tal como las descritas en las Patentes EE.UU.N<sup>os</sup> 35 7.985.565 y 7.115.400 se denominan estructuras "puenteadas" formadas mediante re-asociación de pares de hebras de polinucleótidos inmovilizados y hebras complementarias inmovilizadas, estando ambas hebras inmovilizadas sobre el soporte sólido en el extremo 5', preferiblemente a través de un enlace covalente. Las metodologías de amplificación en grupo son ejemplos de métodos en donde se emplea un molde de ácido nucleico para producir amplicones inmovilizados. Se pueden emplear también otras metodologías adecuadas para producir amplicones inmovilizados a partir de fragmentos de ácido nucleico inmovilizado producido según los métodos proporcionados en la presente memoria. Por ejemplo, se pueden formar uno o más grupos o colonias a través de PCR en fase sólida si uno o ambos de cada par de cebadores de amplificación están inmovilizados.

40 Como se emplea en la presente memoria, los términos "superficie sólida", "soporte sólido" y otros equivalentes gramaticales en la presente memoria se refieren a cualquier material que sea apropiado o se pueda modificar para ser apropiado para la unión de un polinucleótido. Posibles sustratos incluyen, pero no se limitan a, vidrio o vidrio modificado o funcionalizado, plásticos (incluyendo acrílicos, poliestireno y copolímeros de estireno y otros materiales, polipropileno, polietileno, polibutileno, poliuretanos, Teflon<sup>TM</sup>, etc.), polisacáridos, nilón o nitrocelulosa, cerámicas, resinas, sílica o materiales basados en sílica que incluyen silicona y silicona modificada, carbono, metales, vidrios inorgánicos, plásticos, haces de fibra óptica, y una variedad de otros polímeros. En algunas realizaciones, los soportes sólidos y las superficies sólidas se colocan dentro de un aparato celular de flujo. En

- algunas realizaciones el soporte sólido comprende una superficie diseñada adecuada para la inmovilización de moléculas en un patrón ordenado. Una "superficie diseñada" se refiere a una disposición de diferentes regiones en o sobre una capa expuesta de un soporte sólido. En algunas realizaciones, el soporte sólido comprende una matriz de pocillos o depresiones en una superficie. La composición y geometría del soporte sólido puede variar con su uso. En algunas realizaciones, el soporte sólido es una estructura plana tal como un portaobjetos, chip, microchip y/o matriz. Como tal, la superficie de un sustrato puede estar en forma de una capa plana. En algunas realizaciones, el soporte sólido comprende una o más superficies de una célula de flujo. El término "célula de flujo" como se emplea en la presente memoria se refiere a una cámara que comprende una superficie sólida a través de la cual pueden fluir uno o más reactivos de fluido. Ejemplos de células de flujo y sistemas fluidicos relacionados y plataformas de detección que se pueden emplear fácilmente en los métodos de la presente divulgación se describen, por ejemplo, en Bentley *et al.*, *Nature* 456:53-59 (2008), WO 04/018497; EE.UU. 7.057.026; WO 91/06678; WO 07/123744; EE.UU. 7.329.492; EE.UU. 7.211.414; EE.UU. 7.315.019; EE.UU. 7.405.281, y EE.UU. 2008/0108082. En algunas realizaciones, el soporte sólido o su superficie no es plana, tal como la superficie interna o externa de un tubo o vaso. En algunas realizaciones, el soporte sólido comprende microesferas o perlas. "Microesferas", "perlas", "partículas", o equivalentes gramaticales en la presente memoria están destinados a significar partículas discretas hechas de varios materiales que incluyen, pero no se limitan a, plásticos, cerámicas, vidrio, y poliestireno. En determinadas realizaciones, las microesferas son microesferas o perlas magnéticas. Alternativa o adicionalmente, las perlas pueden ser porosas. Los tamaños de la perla oscilan de nanómetros, *por ejemplo*, 100 nm, a milímetros, *por ejemplo*, 1 mm.
- En otras realizaciones, los fragmentos de ácido nucleico inmovilizado se amplifican en disolución. Por ejemplo, en algunas realizaciones, los fragmentos de ácido nucleico inmovilizado se separan o liberan de otra manera del soporte sólido y los cebadores de amplificación se hibridan después en disolución para liberar moléculas. En otras realizaciones, los cebadores de amplificación se hibridan a los fragmentos de ácido nucleico inmovilizados en una o más etapas de amplificación inicial, seguido de las posteriores etapas de amplificación. Por tanto, en algunas realizaciones, se puede emplear un molde de ácido nucleico inmovilizado para producir amplicones en fase de disolución. Se apreciará que para amplificar los fragmentos de ácido nucleico inmovilizados se puede emplear cualquiera de las metodologías de amplificación descritas en la presente memoria con los cebadores universales o específicos de diana.
- El ácido nucleico amplificado según el método proporcionado en la presente memoria se puede secuenciar según cualquier metodología de secuenciación adecuada, tal como la secuenciación directa, incluyendo la secuenciación mediante síntesis, secuenciación mediante ligación, secuenciación mediante hibridación, secuenciación por nanoporos, y similares. En algunas realizaciones, los fragmentos de ADN inmovilizados se secuencian sobre un soporte sólido. En algunas realizaciones, el soporte sólido para la secuenciación es el mismo soporte sólido en el que ocurre la amplificación.
- En algunas realizaciones, la metodología de secuenciación empleada en el método proporcionado en la presente memoria es la secuenciación mediante síntesis (SBS, de sus siglas en inglés). En la SBS, la extensión del cebador del ácido nucleico a lo largo de un molde de ácido nucleico (*por ejemplo*, un ácido nucleico diana o amplicón del mismo) se controla para determinar la secuencia de nucleótidos en el molde. El proceso químico subyacente puede ser la polimerización (por ejemplo, como el catalizado mediante una enzima polimerasa). En una realización de SBS particular basada en la polimerasa, los nucleótidos marcados fluorescentemente se añaden a un cebador (extendiendo de esta manera al cebador) en un molde de una manera dependiente tal que la detección del orden y del tipo de nucleótidos añadidos al cebador se puede emplear para determinar la secuencia del molde.
- Se pueden emplear otros procesos de secuenciación que utilizan reacciones cíclicas, tal como la pirosecuenciación. La pirosecuenciación detecta la liberación de pirofosfato inorgánico (PPi) como nucleótidos particulares que se incorporan en una hebra de ácido nucleico naciente (Ronaghi, *et al.*, 1996), *Analytical Biochemistry* 242(1), 84-9; Ronaghi, 2001, *Genome Res.* 11(1), 3-11; Ronaghi *et al.*, 1998, *Science* 281 (5475), 363; EE.UU. 6.210.891; EE.UU. 6.258.568 y EE.UU. 6.274.320. En la pirosecuenciación, el PPi liberado se puede detectar siendo convertido directamente en trifosfato adenosina (ATP) mediante la ATP sulfurilasa, y el nivel de ATP generado se puede detectar a través de fotones producidos por luciferasa. Por tanto, la reacción de secuenciación se puede controlar a través de un sistema de detección de luminiscencia. Las fuentes de radiación de excitación empleadas para los sistemas de detección basados en fluorescencia no son necesarias para procesos de pirosecuenciación. Detectores y procedimientos útiles para los sistemas fluidicos que se pueden adaptar para la aplicación de la pirosecuenciación de amplicones producidos según la presente divulgación se describen, por ejemplo, en la Solicitud de Patente WIPO N° PCT/US11/57111, EE.UU. 2005/0191698 A1, EE.UU. 7.595.883, y EE.UU. 7.244.559.
- Algunas realizaciones pueden utilizar métodos que implican controlar a tiempo real la actividad de la ADN polimerasa. Por ejemplo, se pueden detectar las incorporaciones de nucleótidos a través de las interacciones de transferencia de energía de resonancia fluorescente (FRET) entre una polimerasa que tiene un fluoróforo y nucleótidos marcados con  $\gamma$ -fosfato, o con guías de onda de modo cero (ZMWs, de sus siglas en inglés). Técnicas y reactivos para la secuenciación basada en FRET se describen, por ejemplo, en Levene *et al.*, 2003, *Science* 299, 682-686; Lundquist *et al.*, 2008, *Opt. Lett.* 33, 1026-1028; Korfach *et al.*, 2008, *Proc. Natl. Acad. Sci.* EE.UU. 105, 1176-1181.

- Algunas realizaciones de SBS incluyen la detección de un protón liberado en la incorporación de un nucleótido en un producto de extensión. Por ejemplo, en la secuenciación basada en la detección de protones liberados se puede usar un detector eléctrico y técnicas asociadas que están disponibles comercialmente en Ion Torrent (Guilford, CT, Life Technologies subsidiary) o métodos y sistemas de secuenciación descritos en EE.UU. 2009/0026082 A1; EE.UU. 2009/0127589 A1; EE.UU. 2010/0137143 A1; o EE.UU. 2010/0282617 A1. Los métodos establecidos en la presente memoria para amplificar ácidos nucleicos diana empleando la exclusión cinética se pueden aplicar fácilmente a los sustratos empleados para detectar protones. Más específicamente, los métodos establecidos en la presente memoria se pueden emplear para producir poblaciones clónicas de amplicones que se usan para detectar protones.
- 5
- 10 Otra técnica de secuenciación útil es la secuenciación por nanoporos (véase, por ejemplo, Deamer *et al.*, 2000, *Trends Biotechnol.*, 18, 147-151; Deamer *et al.*, 2002, *Acc. Chem. Res.* 35:817-825; Li *et al.*, 2003, *Nat. Mater.* 2:611-615).
- En algunas realizaciones por nanoporo, el ácido nucleico diana o los nucleótidos individuales eliminados a partir del ácido nucleico diana pasan a través de un nanoporo. Como el ácido nucleico o los nucleótidos pasan a través del nanoporo, se puede identificar cada tipo de nucleótido mediante la medición de las fluctuaciones en la conductancia eléctrica del poro. (Patente U.S. Nº 7.001.792; Soni *et al.*, 2007, *Clin. Chem.*, 53, 1996-200; Healy, 2007, *Nanomed.* 2, 459-481; Cockroft *et al.*, 2008, *J. Am. Chem. Soc.*, 130, 818-820).
- 15
- En otras realizaciones de secuenciación por nanoporos, los fragmentos de ADN a secuenciar crean una única corriente característica de nanoporos empleando nucleótidos modificados con gamma fosfato.
- 20 La recitación de un listado de elementos en cualquier definición de una variable en la presente memoria incluye las definiciones de la variable como cualquier elemento único o en combinación (o sub-combinación) de los elementos listados. La recitación de una realización en la presente memoria incluye la realización como cualquier realización única o en combinación con otras realizaciones o porciones de las mismas.

**Lista de secuencias**

- <110> Illumina, Inc.
- 5 <120> Amplificación de polinucleótido que usa sistemas CRISPR-CAS  
 <130> 12957-170-228
- 10 <140> Para asignar  
 <141> 10-11-2015  
 <150> 62/078,355  
 <151> 11-11-2014
- 15 <160> 8  
 <170> PatentIn versión 3.5
- 20 <210> 1  
 <211> 29  
 <212> DNA  
 <213> Secuencia Artificial
- 25 <220>  
 <223> Secuencia P5 estándar  
 <400> 1  
 aatgatacgg cgaccaccga gatctacac 29
- 30 <210> 2  
 <211> 24  
 <212> DNA  
 <213> Secuencia Artificial
- 35 <220>  
 <223> Secuencia P7 estándar  
 <400> 2  
 caagcagaag acggcatacg agat 24
- 40 <210> 3  
 <211> 32  
 <212> DNA  
 <213> Secuencia Artificial
- 45 <220>  
 <223> Secuencia P5 modificada con PAM
- 50 <220>  
 <221> característica miscelánea  
 <222> (3)..(3)  
 <223> n es a, c, g, o t
- 55 <400> 3  
 ccnaatgata cggcgaccac cgagatctac ac 32
- 60 <210> 4  
 <211> 27  
 <212> DNA  
 <213> Secuencia Artificial
- <220>  
 <223> Secuencia P7 modificada con PAM

<220>  
 <221> característica miscelánea  
 <222> (3)..(3)  
 <223> n es a, c, g, o t  
 5  
 <400> 4  
 ccncaagcag aagacggcat acgagat 27  
 <210> 5  
 10 <211> 21  
 <212> DNA  
 <213> Secuencia Artificial  
 <220>  
 15 <223> Secuencia P5 con PAM modificado truncado  
 <220>  
 <221> característica miscelánea  
 <222> (3)..(3)  
 20 <223> n es a, c, g, o t  
 <400> 5  
 ccnaatgata cggcgaccac c 21  
 25 <210> 6  
 <211> 21  
 <212> DNA  
 <213> Secuencia Artificial  
 30 <220>  
 <223> Secuencia P7 con PAM modificado truncado  
 <220>  
 <221> característica miscelánea  
 35 <222> (3)..(3)  
 <223> n es a, c, g, o t  
 <400> 6  
 40 ccncaagcag aagacggcat a 21  
 <210> 7  
 <211> 20  
 <212> RNA  
 <213> Secuencia Artificial  
 45 <220>  
 <223> ARN guía dirigido a P5  
 <400> 7  
 50 ucgguggucg ccguaucuu 20  
 <210> 8  
 <211> 20  
 <212> RNA  
 55 <213> Secuencia Artificial  
 <220>  
 <223> ARN guía dirigido a P7  
 60 <400> 8  
 cguaugccgu cuucugcuug 20  
 65

**REIVINDICACIONES**

1. Un método para amplificar un ácido nucleico bicatenario diana que comprende:
  - a) proporcionar un sistema que tiene:
    - 5 un ARN (ARNcr) con repeticiones palindrómicas cortas agrupadas y regularmente interespaciadas (CRISPR), y una proteína asociada a CRISPR (Cas), en donde el ARNcr contiene una región de nucleótidos específica de la diana complementaria para una región de una primera cadena de ácido nucleico bicatenario diana;
    - b) poner en contacto el ácido nucleico bicatenario diana con el sistema para formar un complejo;
    - c) hibridar un cebador a una segunda cadena del ácido nucleico bicatenario diana, el cebador contiene una secuencia complementaria para una región de la segunda cadena del ácido nucleico bicatenario diana, y
    - 10 d) prolongar un ácido nucleico complementario a la segunda cadena de ácido nucleico bicatenario diana a partir del cebador empleando una polimerasa.
2. El método de la reivindicación 1,
  - i) en donde el ácido nucleico bicatenario diana se amplifica de manera lineal; o
  - ii) en donde el ácido nucleico bicatenario diana se amplifica de manera exponencial, o
  - 15 iii) en donde el ácido nucleico diana es un ADN bicatenario (ADNs), o
  - iv) en donde el ácido nucleico diana es un ARN bicatenario (ARNs), o
  - v) en donde el sistema es un sistema CRISPR-Cas de Tipo I, o
  - vi) en donde el sistema es un sistema CRISPR-Cas de Tipo II, o
  - vii) en donde el sistema es un sistema CRISPR-Cas de Tipo III, o
  - 20 viii) en donde el sistema comprende además un ARNcr trans-activador (ARNtracr), o
  - ix) en donde el ARNcr es un polinucleótido que comprende un polinucleótido ARNcr fusionado a un polinucleótido ARNtracr, o
  - x) en donde la primera hebra del ácido nucleico bicatenario diana contiene una secuencia complementaria a un motivo adyacente de protoespaciador (PAM) 5'-NGG, o
  - 25 xi) en donde la primera hebra del ácido nucleico bicatenario diana contiene una secuencia universal, y en donde el ARNcr contiene una secuencia complementaria a una región de la secuencia universal, o
  - xii) que comprende además repetir de la etapa (a) a la etapa (d) durante una o más veces.
3. El método de la reivindicación 2 en donde la primera hebra del ácido nucleico bicatenario diana contiene una secuencia universal, y en donde el ARNcr contiene una secuencia complementaria a una región de la secuencia universal, en donde
  - i) el cebador contiene una secuencia de una región de la secuencia universal, o
  - ii) la secuencia universal tiene una secuencia de SEQ ID No.3, en donde opcionalmente el ARNcr contiene una secuencia de SEQ ID No.7, en donde opcionalmente el cebador contiene una secuencia de SEQ ID No.5, o
  - 35 iii) en donde la secuencia universal tiene una secuencia de SEQ ID No.4, en donde opcionalmente el ARNcr contiene una secuencia de SEQ ID No.8, en donde opcionalmente el cebador contiene una secuencia de SEQ ID No.6.
4. El método de la reivindicación 1, en donde la proteína Cas es una proteína Cas9 en donde opcionalmente la proteína Cas9 contiene dos dominios nucleasa inactivados, en donde los dos dominios nucleasa inactivados comprenden una primera mutación en el dominio que separa la hebra complementaria para el ARNcr y una segunda mutación en el dominio que separa la cadena no complementaria para el ARNcr, en donde opcionalmente la primera mutación es D10A y la segunda mutación es H840A.
5. El método de la reivindicación 1, en donde
  - i) la proteína Cas es una proteína en Cascada, o

- ii) la proteína Cas es una proteína Cas3, o
- iii) la polimerasa es una polimerasa que desplaza la cadena, en donde opcionalmente la polimerasa se selecciona del grupo que consiste en Bst, Bsu, y Phi29.

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

- 5 aplicar al menos una transposasa y al menos una composición de extremo de transposón que contiene una cadena transferida a una muestra que contiene un ácido nucleico diana bajo condiciones donde el ácido nucleico diana y la composición extremo del transposón sufren una reacción de transposición para generar una mezcla, en donde el ácido nucleico diana se fragmenta para generar una pluralidad de fragmentos de ácido nucleico diana, e
- 10 incorporar una secuencia cebadora universal en cada uno de la pluralidad de fragmentos del ácido nucleico diana, en donde el ARNcr contiene una región de nucleótidos específica a la diana complementaria para una región del cebador universal, en donde opcionalmente el cebador universal se incorpora en una pluralidad de fragmentos de ácido nucleico diana mediante una reacción PCR, en donde opcionalmente
  - a) el cebador universal tiene una secuencia complementaria a un motivo adyacente de protoespaciador (PAM) 5'-NGG o
  - 15 b) la secuencia universal tiene una secuencia de SEQ ID No.3, en donde opcionalmente el ARNcr contiene una secuencia de SEQ ID No.7, en donde opcionalmente el cebador contiene una secuencia de SEQ ID No.5, o
  - c) el cebador universal tiene una secuencia de SEQ ID No.4, en donde opcionalmente el ARNcr contiene una secuencia de SEQ ID No.8, en donde opcionalmente el cebador contiene una secuencia de SEQ ID No.6.

7. El método de la reivindicación 6, en donde los cebadores universales se incorporan en los dos extremos de cada uno de la pluralidad de fragmentos de ácido nucleico diana, en donde opcionalmente los dos cebadores universales tienen secuencias de SEQ ID No.3 y SEQ ID No.4.

8. Un método para amplificar un ácido nucleico bicatenario diana que comprende:

- 25 a) proporcionar un primer sistema que tiene: un primer ARN (ARNcr) con repeticiones palindrómicas cortas agrupadas y regularmente interespaciadas (CRISPR), y una primera proteína asociada a CRISPR (Cas), en donde el primer ARNcr contiene una región de nucleótidos específica de la diana complementaria para una región de una primera cadena del ácido nucleico bicatenario diana;
- 30 b) proporcionar un segundo sistema que tiene: un segundo ARN (ARNcr) con repeticiones palindrómicas cortas agrupadas y regularmente interespaciadas (CRISPR), y una segunda proteína asociada a CRISPR (Cas), en donde el segundo ARNcr contiene una región de nucleótidos específica de la diana complementaria para una región de una segunda cadena del ácido nucleico bicatenario diana;
- c) poner en contacto el ácido nucleico bicatenario diana con el primer sistema y el segundo sistema;
- 35 d) hibridar el primer sistema a una segunda cadena del ácido nucleico bicatenario diana, el primer cebador que contiene una secuencia complementaria para una región de la segunda cadena del ácido nucleico bicatenario diana, e hibridar un segundo cebador a una primera cadena del ácido nucleico bicatenario diana, el segundo cebador que contiene una secuencia complementaria para una región de la primera cadena del ácido nucleico bicatenario diana, y
- e) prolongar el extremo 3' del primer cebador y el segundo cebador con una o más polimerasas para generar un primer y un segundo ácido nucleico diana bicatenario.

9. El método de la reivindicación 8 en donde

- 40 a) el ácido nucleico diana es un ADN bicatenario (ADNs), o
- b) el ácido nucleico diana es un ARN bicatenario (ARNds), o
- c) el primer sistema o el segundo sistema es un sistema CRISPR-Cas Tipo I, o
- d) el primer sistema o el segundo sistema es un sistema CRISPR-Cas Tipo II, o
- e) el primer sistema o el segundo sistema es un sistema CRISPR-Cas Tipo III, o
- 45 f) el primer sistema o el segundo sistema comprende además un ARNcr trans-activador (ARNtracr), o
- g) el ARNcr del primer sistema o del segundo sistema es un polinucleótido que comprende un polinucleótido ARNcr fusionado a un polinucleótido ARNtracr, o

- h) la primera cadena y la segunda cadena del ácido nucleico bicatenario diana contienen una secuencia complementaria a un motivo adyacente de protoespaciador (PAM) 5'-NGG, o
- i) el método comprende además repetir de la etapa (a) a la etapa (e) durante una o más veces.

10. El método de la reivindicación 8, en donde:

5 la primera hebra del ácido nucleico bicatenario diana contiene una primera secuencia universal, y en donde el ARNcr del primer sistema contiene una secuencia complementaria para una región de la primera secuencia universal, y la segunda hebra del ácido nucleico bicatenario diana contiene una segunda secuencia universal, y en donde el ARNcr del segundo sistema contiene una secuencia complementaria para una región de la segunda secuencia universal, en donde opcionalmente el primer cebador contiene una secuencia de una región de la primera secuencia universal, y el segundo cebador contiene una secuencia de una región de la segunda secuencia universal, en donde opcionalmente:

10 la primera secuencia universal tiene una secuencia de SEQ ID No.3, el ARNcr del primer sistema contiene una secuencia de SEQ ID No.7, y el primer cebador contiene una secuencia de SEQ ID No.5, y la segunda secuencia universal tiene una secuencia de SEQ ID No.4, el ARNcr del segundo sistema contiene una secuencia de SEQ ID No.8, y el segundo cebador contienen una secuencia de SEQ ID No.6.

11. El método de la reivindicación 8, en donde la proteína Cas del primer sistema o del segundo sistema es una proteína Cas9, en donde opcionalmente la proteína Cas9 contiene dos dominios nucleasa inactivados, en donde opcionalmente los dos dominios nucleasa inactivados comprenden una primera mutación en el dominio que separa la cadena complementaria para el ARNcr y una segunda mutación en el dominio que separa la cadena no complementaria para el ARNcr, en donde opcionalmente la primera mutación es D10A y la segunda mutación es H840A.

12. El método de la reivindicación 8, en donde

- a) la proteína Cas del primer sistema o del segundo sistema es una proteína en Cascada, o
- b) la proteína Cas del primer sistema o del segundo sistema es una proteína Cas3, o
- 25 c) la polimerasa es una polimerasa que desplaza la cadena, en donde opcionalmente la polimerasa se selecciona del grupo que consiste en Bst, Bsu, y Phi29.

13. El método de cualquiera de las reivindicaciones 1-12, en donde el ácido nucleico diana es ADN genómico.

14. El método de cualquiera de las reivindicaciones 1-12, en donde el ácido nucleico diana contiene ADN cromosómico.

30 15. El método de cualquiera de las reivindicaciones 1-12, en donde el ácido nucleico diana comprende un genoma o un genoma parcial, que comprende además opcionalmente secuenciar el ácido nucleico diana o los fragmentos de ácido nucleico diana, en donde opcionalmente la secuenciación comprende el uso de una o más secuenciaciones mediante síntesis, PCR de puente, secuenciación por terminación en cadena, secuenciación por hibridación, secuenciación por nanoporos, y secuenciación por ligación.

35

**P5 y P7 estándar:**

P5 seq: 5' AATGATACGGCGACCACCGAGATCTACAC '3 (SEQ ID No. 1)

P7 seq: 5' CAAGCAGAAAGACGGCATACGAGAT '3 (SEQ ID No. 2)

**P5 y P7 modificado en PAM:**

P5 seq: 5' CCN AATGATACGGCGACCACCGAGATCTACAC '3 (SEQ ID No. 3)

P7 seq: 5' CCN CAAGCAGAAAGACGGCATACGAGAT '3 (SEQ ID No. 4)

**P5 y P7 modificado en PAM, truncado:**

P5 seq: 5' CCN AATGATACGGCGACCACCGAGATCTAGAG '3 (SEQ ID No. 5)

P7 seq: 5' CCN CAAGCAGAAAGACGGCATAGGAGAT '3 (SEQ ID No. 6)

**ARNs guía:**

P5 diana : 5' UCGGUGGUCGCCGUAUCAUU 3' (SEQ ID No. 7)

P7 diana : 5' CGUAUGCCGUCUCUGCUUG 3' (SEQ ID No. 8)

Figura 1A

Amplicón Nextera en disolución:

**P5** 3' GGA TTACTATGCCCGCTGGCTCTAGATGTG----TAGAGCATACGGCAGAGACGGAAC TCC 5'  
 5' CCT AATGATACGGGACCAACCGAGATCTACAC----ATCTCGTATGCCGTCTTCTGCTTG AGG 3'

+

3' UUACUAUGCCCGUGGUGGU 5'

Cas9+ARN guía (P5')



**II** 3' GGA TTACTATGCCCGCTGGCTCTAGATGTG----TAGAGCATACGGCAGAGACGGAAC TCC 5'  
 5' CCT AATGATACGGGACCAACCGAGATCTACAC----ATCTCGTATGCCGTCTTCTGCTTG AGG 3'



**III** 3' GGA TTACTATGCCCGCTGGCTCTAGATGTG----TAGAGCATACGGCAGAGACGGAAC TCC 5'  
 5' CCT AATGATACGGGACCAACCGAGATCTACAC----ATCTCGTATGCCGTCTTCTGCTTG AGG 3'



**IV** 3' GGA TTACTATGCCCGCTGGCTCTAGATGTG----TAGAGCATACGGCAGAGACGGAAC TCC 5'  
 5' CCT AATGATACGGGACCAACCGAGATCTACAC----ATCTCGTATGCCGTCTTCTGCTTG AGG 3'

Figura 1B

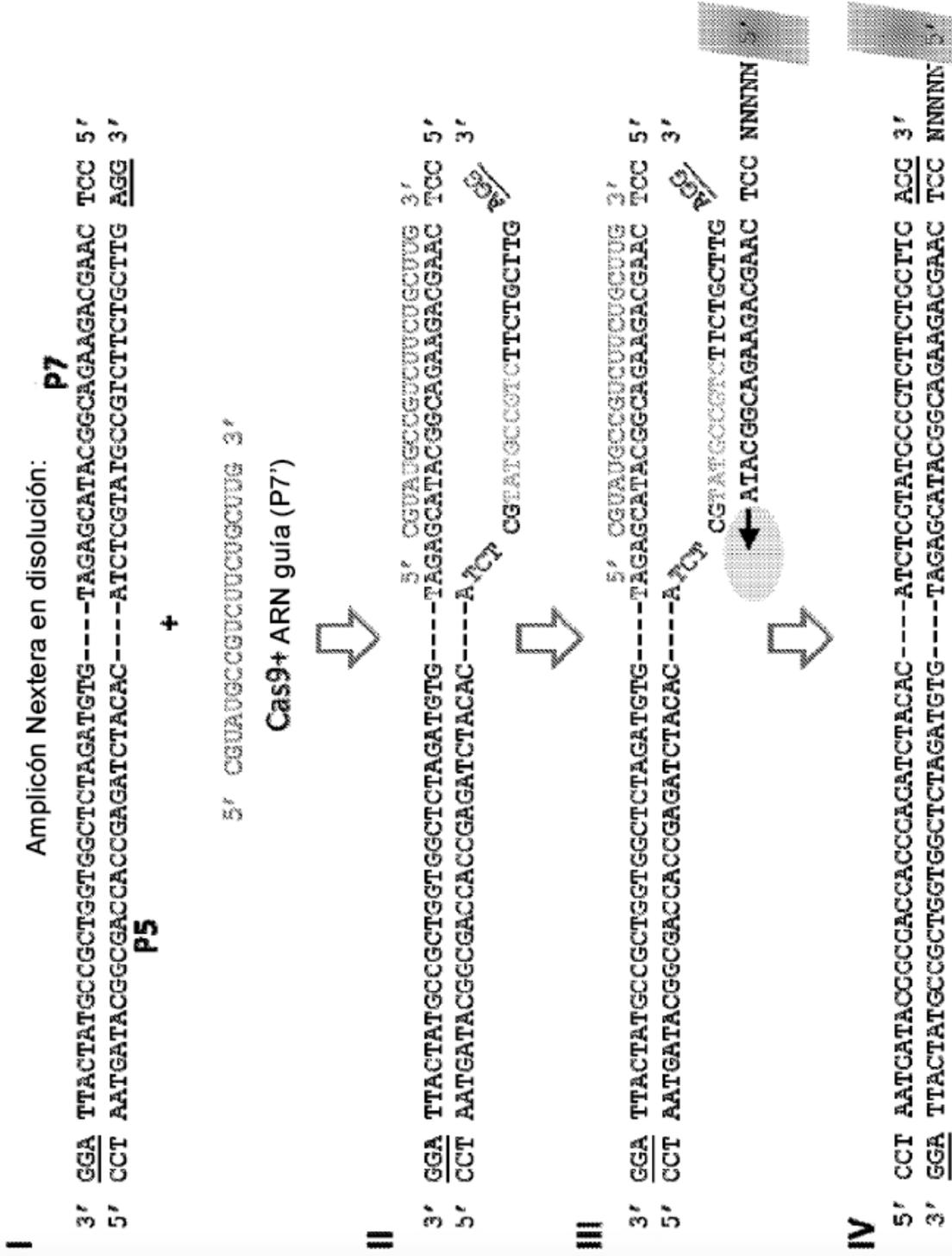


Figura 1C