

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 645 418**

51 Int. Cl.:

**C12P 19/34** (2006.01)

**C12Q 1/68** (2006.01)

**C12Q 1/70** (2006.01)

**C12N 15/10** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **21.12.2012 PCT/US2012/071303**

87 Fecha y número de publicación internacional: **27.06.2013 WO13096798**

96 Fecha de presentación y número de la solicitud europea: **21.12.2012 E 12858832 (4)**

97 Fecha y número de publicación de la concesión europea: **06.09.2017 EP 2794904**

54 Título: **Amplificación de una secuencia de un ácido ribonucleico**

30 Prioridad:

**22.12.2011 US 201161579512 P**

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

**05.12.2017**

73 Titular/es:

**IBIS BIOSCIENCES, INC. (100.0%)  
2251 Faraday Avenue Suite 150  
Carlsbad, CA 92008, US**

72 Inventor/es:

**MOTLEY, STANLEY y  
ESHOO, MARK W.**

74 Agente/Representante:

**IZQUIERDO BLANCO, María Alicia**

ES 2 645 418 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).

**Amplificación de una secuencia de un ácido ribonucleico****Descripción**5 **CAMPO**

La presente invención se refiere a procedimientos y composiciones para marcar, amplificar, purificar y o caracterizar ácido ribonucleico (ARN) en una muestra. En particular, se proporcionan procedimientos para la preparación de ARN viral de secuencia desconocida a partir de una muestra para su análisis posterior.

10

**ANTECEDENTES**

Las tecnologías de secuenciación de próxima generación están revolucionando el proceso para la identificación de patógenos en una muestra. Sin embargo, estas tecnologías están típicamente limitadas en cuanto a que requieren una cantidad relativamente grande de ADN de entrada para secuenciar, mientras que las muestras clínicas/forenses/ambientales importantes de interés pueden contener niveles bajos de ácido nucleico del patógeno viral importante. Para agravar aún más este problema está el hecho de que muchos virus patógenos importantes tienen genomas que están compuestos por ARN.

15

El documento US 5.545.522 describe procesos en los que se sintetiza ácido desoxirribonucleico complementario (ADNc) a partir de una secuencia de ácido ribonucleico (ARN) usando un cebador complementario unido a una región promotora de la ARN polimerasa complementaria y, después, se transcribe el ARN antisentido (ARNa) a partir del ADNc introduciendo una ARN polimerasa capaz de unirse a la región promotora.

20

El documento WO 99/29907 describe procedimientos para fabricar ácidos nucleicos que implican añadir una secuencia de nucleótidos conocida al extremo 3' del primer ARN que tiene una secuencia conocida en el extremo 5' para formar un segundo ARN y transcribir de forma inversa el segundo ARN para formar un ADNc.

25

**SUMARIO**

30

La presente invención proporciona procedimientos de amplificación de ARN viral de secuencia desconocida de una muestra, que comprende: (a) añadir una poli-A-polimerasa a una muestra de ARN viral no poliadenilada para unir un tramo de poli(A) sobre el extremo 3' del ARN viral; (b) hibridar un oligonucleótido al tramo de poli(A), en el que dicho oligonucleótido comprende un dominio de hibridación que comprende un tramo de poli(T) y un dominio funcional que comprende una región promotora de ARN polimerasa; (c) sintetizar ADNc de doble cadena a partir del ARN viral y el oligonucleótido, en el que la síntesis de ADNc bicatenario a partir del ARN viral y el oligonucleótido comprende la transcripción inversa de una primera hebra de dicho ADNc y sintetizar una segunda hebra de dicho ADN; (d) amplificar el ADNc bicatenario con una ARN polimerasa para producir ARN antisentido amplificado; y (e) convertir el ARN antisentido amplificado en ADNc amplificado. En algunas realizaciones, dicho dominio de hibridación está en el extremo 3' de dicho oligonucleótido y dicho dominio funcional está en el extremo 5' de dicho oligonucleótido. En algunas realizaciones, dicha región promotora de la ARN polimerasa comprende una región promotora de T7. En algunas realizaciones, dicha segunda hebra de dicho ADN. se sintetiza usando ARNasa H y ADN polimerasa. En algunas realizaciones, el ARN antisentido amplificado es una amplificación de al menos 50 veces de dicho ADNc bicatenario. En algunas realizaciones, dicho ARN antisentido amplificado es una amplificación de al menos 100 veces de dicho ADNc bicatenario. En algunas realizaciones, dicho ARN antisentido amplificado es una amplificación de al menos 500 veces de dicho ADNc bicatenario. En algunas realizaciones, el ARN antisentido amplificado es una amplificación de al menos 1000 veces de dicho ADNc bicatenario. En algunas realizaciones, convertir el ARN antisentido amplificado en ADNc amplificado comprende la transcripción inversa. En algunas realizaciones, dicha muestra comprende una muestra biológica, ambiental y / o forense.

35

40

45

50

La presente invención también proporciona procedimientos de secuenciación de ARN viral de secuencia desconocida de una muestra, que comprende: (1) amplificar dicho ARN viral mediante: (a) añadir una poli-A-polimerasa a una muestra de ARN viral no poliadenilada para unir un tramo de poli(A) sobre el extremo 3' del ARN viral; (b) hibridar un oligonucleótido al tramo de poli(A), en el que dicho oligonucleótido comprende un dominio de hibridación que comprende un tramo de poli(T) y un dominio funcional que comprende una región promotora de ARN polimerasa; (c) sintetizar ADNc de doble cadena a partir del ARN viral y el oligonucleótido, en el que la síntesis de ADNc bicatenario a partir del ARN viral y el oligonucleótido comprende la transcripción inversa de una primera hebra de dicho ADNc y sintetizar una segunda hebra de dicho ADN; (d) amplificar el ADNc bicatenario con una ARN polimerasa para producir ARN antisentido amplificado; y (e) convertir el ARN antisentido amplificado en ADNc amplificado; y (2) secuenciar dicho ARN viral usando el ADNc amplificado. En algunas realizaciones, la secuenciación es mediante una técnica de secuenciación de la próxima generación.

55

60

También se divulga en el presente documento kits para realizar la secuenciación (por ejemplo, secuenciación de siguiente generación) y/o amplificación de ARN (por ejemplo, poliadenilación, transcripción inversa, síntesis de ARN y transcripción inversa) mediante los métodos de la invención.

65

**DESCRIPCIÓN DETALLADA**

La presente invención se refiere a procedimientos y composiciones para marcar, amplificar, purificar y o caracterizar ácido ribonucleico (ARN) en una muestra. La presente invención se refiere a métodos para marcar un ARN diana con un marcador de ácido nucleico y usar dicho marcador para amplificar el ARN diana para el análisis aguas abajo. La presente invención se refiere al uso de poli-A-polimerasa para unir una cola de poli(A) a ARN (por ejemplo, un material genómico viral etc.) dentro de una muestra (por ejemplo, biológica, clínica, forense, ambiental, de investigación etc.). En la invención, el ARN en una muestra está marcado con poli(A). En la invención, el ARN marcado (marcado con poli(A) se amplifica (por ejemplo, mediante ARN polimerasa de T7 usando el procedimiento de Eberwine, etc.). En algunas realizaciones, los procedimientos de la invención generan una gran cantidad de ácido nucleico a partir de niveles traza de ARN

(por ejemplo, material genómico viral) presente en una muestra (por ejemplo, biológica, clínica, forense, ambiental, investigación, etc.). La presente invención implica la conversión de ARN (amplificado a partir de una muestra) en ADN (ADNc). En algunas realizaciones, la presente invención proporciona métodos para la secuenciación (por ejemplo, secuenciación de próxima generación) de ácidos nucleicos (por ejemplo, ADN, ADNc). En algunas realizaciones, la presente invención proporciona la identificación y / o caracterización de virus de ARN patogénicos en una muestra (por ejemplo, mediante marcado con poli(A), amplificación de ARN, conversión de ARN a ADNc y secuenciación de ADN). En algunas realizaciones, los procedimientos y kits descritos en el presente documento se usan para amplificar ARN de una muestra (por ejemplo, una muestra que contiene traza de ARN de una muestra).

En algunas realizaciones, los procedimientos de la presente invención comprenden una o más de las etapas de: (1) poliadenilación del ARN diana en una muestra mediante la adición de polimerasa poli(A) (o un equivalente funcional) a una muestra para producir una cola de poli (A) en ARN no poliadenilado (por ejemplo, ARN viral) dentro de la muestra; (2) adición de un cebador de transcripción inversa (cebador de TI), que comprende un segmento 3'-poli(T) y una región promotora en 5' (por ejemplo, promotor de ARN polimerasa (por ejemplo, promotor de ARN polimerasa T7), al ARN en una muestra; (3) hibridación del segmento 3'-poli(T) de un cebador de TI al segmento de poli(A) de un ARN diana; (4) transcripción inversa del ARN diana del cebador de TI; (5) síntesis de la segunda hebra (por ejemplo, usando ARNasa H y ADN polimerasa I), para producir ADNc bicatenario del ARN diana; (6) terminación roma del ADNc (por ejemplo, con ADN polimerasa T4); (7) transcripción de ADNc por una ARN polimerasa (por ejemplo, ARN polimerasa de T7) en ARN antisentido amplificado (ARNa), (8) conversión (mediante transcripción inversa de ARNm amplificado en ADNc y (9) análisis o manipulación adicional de ADNc (por ejemplo, secuenciación). En algunas realizaciones, los procedimientos comprenden cada una de las etapas anteriores en orden. En algunas realizaciones, se realizan una o más de las etapas anteriores. En algunas realizaciones, el orden de las etapas anteriores se modifica. En algunas realizaciones, las etapas se llevan a cabo usando diferentes técnicas, componentes y / o reactivos que los descritos anteriormente. En algunas realizaciones, modificaciones, etapas adicionales, reordenamiento de etapas, etc. están dentro del alcance de las realizaciones de la presente invención.

En algunas realizaciones, en los procedimientos en el presente documento se usa una muestra que contiene ARN. En algunas realizaciones, una muestra puede ser cualquier muestra que comprenda una molécula de ARN diana. La muestra de ARN puede obtenerse de una célula, cultivo celular, un fluido corporal, un tejido, un órgano, el medio ambiente (por ejemplo, muestras de aceite, muestras de agua y muestras de aire), etc. En ciertas realizaciones, las muestras comprenden una mezcla de uno o más de ARN eucariótico, ARN bacteriano y / o ARN viral. En algunas realizaciones, la muestra de ARN comprende uno o más tipos de ARN viral.

En algunas realizaciones de la divulgación, cualquier ácido nucleico adecuado (por ejemplo, ADN, ARN, etc.) proporciona un material de partida y / o un molde inicial para los procedimientos del presente documento. En algunas realizaciones de la divulgación, el material de partida y / o el molde inicial comprenden ARN (por ejemplo, ARN viral, ARN genómico viral, ARN bacteriano, ARN humano, etc.). En la invención, el material de partida y / o el molde inicial comprenden ARN viral y / o ARN genómico viral. En algunas realizaciones, el ARN viral puede ser monocatenario o bicatenario. El ARN de cualquier virus adecuado (conocido o desconocido) encuentra uso como diana en realizaciones de la presente invención. En algunas realizaciones, el ARN objetivo comprende ARN de un virus ARNdc, virus ARN de cadena positiva, virus de ARNmc de sentido positivo, virus de ARNmc de sentido negativo o cualquier virus de familias de los mismos. En algunas realizaciones de la divulgación, el material de partida y / o el molde inicial comprenden ARN bacteriano y/o ARN de transcriptoma bacteriano. Un material de partida inicial puede convertirse a partir de un primer tipo de ácido nucleico (por ejemplo, ADN o ARN) en un segundo tipo de ácido nucleico (por ejemplo, ARN o ADN). En la invención, el ácido nucleico del molde inicial comprende ARN viral no poliadenilado.

En algunas realizaciones, se usa un molde de ácido nucleico adecuado para una poli (A) polimerasa (por ejemplo, ARN, ARN genómico viral, etc.). En algunas realizaciones, se usa una muestra que comprende un molde adecuado para una poli (A) polimerasa (por ejemplo, ARN, ARN genómico viral, etc.). En algunas realizaciones, se usa una muestra (por ejemplo, clínica, biológica, ambiental, forense, etc.) que puede contener, o se sospecha que contiene, uno o más moldes para una poli (A) polimerasa (por ejemplo, ARN genómico viral). En algunas realizaciones, una muestra contiene ARN de uno o más genomas virales (por ejemplo, genomas completos, genomas parciales, etc.). En algunas realizaciones, el ARN genómico viral carece de una cola de poli (A).

En algunas realizaciones, se agrega un marcador de ácido nucleico a un ARN diana. En algunas realizaciones, se agrega un marcador de ácido nucleico al extremo 3' de un ARN diana. En algunas realizaciones, se añade químicamente un marcador de ácido nucleico. En algunas realizaciones, un marcador de ácido nucleico comprende un sitio de hibridación para uno o más oligonucleótidos (por ejemplo, para las siguientes etapas del presente documento). En las realizaciones descritas en el presente documento, el marcador de ácido nucleico comprende, consiste en, o consiste esencialmente en un tracto de poli (A) (por ejemplo, 5-500 nucleótidos A, 20-250 nucleótidos A, etc.). Debe observarse que las realizaciones que se describen en el presente documento junto con un marcador poli (A) también pueden implementarse con marcadores de ácido nucleico de otras secuencias. El alcance de la divulgación en el presente documento no debe estar limitado por el tipo de marcador (por ejemplo, poli (A)) descrito junto con cualquier realización descrita en el presente documento.

En algunas realizaciones, se usan cebadores y otros oligonucleótidos adecuados para llevar a cabo las etapas de los presentes procedimientos. En algunas realizaciones, los oligonucleótidos hibridan con una porción (por ejemplo, cola de poli (A)) de ácidos nucleicos diana. En algunas realizaciones, al menos una parte (por ejemplo, el extremo 3') de un oligonucleótido se hibrida con al menos una porción (por ejemplo, cola de poli (A)) de un ácido nucleico diana. En algunas realizaciones, al menos una parte (por ejemplo, el extremo 5') de un oligonucleótido no es complementaria del objetivo. En algunas realizaciones, se usan oligonucleótidos que comprenden un segmento poli (T). En algunas realizaciones, la parte en 3' de un oligonucleótido comprende un segmento poli (T). En algunas realizaciones, se usan oligonucleótidos que comprenden una región promotora de la ARN polimerasa. En algunas realizaciones, la parte en 5' de un oligonucleótido comprende una región promotora de ARN polimerasa. En algunas realizaciones, se usan oligonucleótidos que comprenden un segmento poli (T) en el extremo 3' y una región promotora de ARN polimerasa en o cerca del extremo 5'.

En algunas realizaciones, los oligonucleótidos y / o ácidos nucleicos producidos por las etapas descritas en el presente documento comprenden al menos un dominio de amplificación. Como se usa en el presente documento, un dominio de amplificación será principalmente una secuencia que puede soportar la amplificación de un ácido nucleico que comprende dicha secuencia (por ejemplo, mediante PCR, mediante ARN polimerasa, etc.). El uso de secuencias de ácido nucleico en reacciones de amplificación es bien conocido en la técnica, y en el presente documento se describen ejemplos no limitantes. En algunas realizaciones, un dominio de amplificación comprenderá una secuencia que puede soportar la unión y extensión del cebador. Se aplican reglas estándar para el diseño de cebadores (Sambrook, 1994). En algunas realizaciones, un dominio de amplificación comprenderá preferiblemente un sitio de unión al cebador para la amplificación por PCR. Los parámetros para el diseño de cebadores para PCR son bien conocidos en la técnica (véase, por ejemplo, Beasley et al, 1999). También se contemplan sitios de unión al cebador para otros tipos de procedimientos de amplificación. En algunas realizaciones, un dominio de amplificación comprende una secuencia promotora (por ejemplo, para una ARN polimerasa (por ejemplo, ARN polimerasa de T7)). En algunas realizaciones, una secuencia promotora (un "dominio de transcripción") es una secuencia de ácido nucleico a la que se une una ARN polimerasa para iniciar la transcripción. En algunas realizaciones, un dominio de amplificación en un ácido nucleico monocatenario comprende una única hebra de la región promotora (por ejemplo, para una ARN polimerasa (por ejemplo, ARN polimerasa de T7)). En algunas realizaciones, se usa un oligonucleótido que comprende la segunda hebra de una región promotora (por ejemplo, para una ARN polimerasa (por ejemplo, ARN polimerasa de T7)). En algunas realizaciones, un oligonucleótido también comprende otros dominios funcionales (por ejemplo, dominios de marcaje, dominios de hibridación (por ejemplo, poli (T), etc.). En algunas realizaciones, un oligonucleótido que comprende una cadena de una región promotora se hibrida con un ácido nucleico que comprende la segunda cadena de una región promotora para producir una región promotora bicatenaria (por ejemplo, para una ARN polimerasa (por ejemplo, ARN polimerasa de T7)).

En ciertas realizaciones de la invención, los procedimientos comprenden reducir el ADN de una muestra (por ejemplo, antes del marcaje con poli (A) del ARN objetivo, antes de la TI del objetivo, antes de la amplificación del ARN objetivo, etc.). Los expertos en la técnica conocen bien los procedimientos para reducir el ADN o separar el ADN del ARN. Un abordaje habitual para reducir el ADN es incubar la muestra con ADNasa. Otro procedimiento para separar el ADN del ARN es la precipitación en cloruro de litio. Los sistemas de aislamiento de ARN basados en filtros también son conocidos en la técnica.

En algunas realizaciones de la invención, los procedimientos comprenden reducir el ARNm poliadenilado de una muestra antes de marcar con poli(A) el ARN objetivo. Los expertos en la técnica conocen bien procedimientos para aislar específicamente ARNm poliadenilado de una muestra. Por ejemplo, un procedimiento común para aislar ARNm poliadenilado comprende hibridar el ARNm poliadenilado con un oligonucleótido poli (T). Típicamente, el oligonucleótido poli (T) está unido a una superficie, tal como una columna o una esfera. Después de que el ARNm poliadenilado se hibrida con el oligonucleótido de poli (T), se puede separar de la muestra. Por ejemplo, si el ARNm poliadenilado se hibrida con el oligonucleótido de poli (T) inmovilizado sobre una esfera magnética. A continuación, las esferas pueden separar de la muestra usando un imán.

En algunas realizaciones, los procedimientos comprenden agotar el ARNr de la muestra. Dependiendo de la composición de la muestra, puede ser deseable agotar el ARNr eucariótico, el ARNr bacteriano o ambos. El ARNr puede hibridar con uno o más oligonucleótidos complementarios al menos a una parte de uno o más del ARNr 17S, ARNr 18S o ARNr 28S o eucariotas o al menos una parte de uno o más del ARNr 16S o ARNr 23S o procariotas.

Los complejos de hibridación se eliminan de la muestra con un sistema de captura apropiado.

La poliadenilación es la adición de una cola de poli (A) a una molécula de ARN. La poliadenilación se lleva a cabo mediante una poli-A polimerasa. La polimerasa Poli-A es una enzima comercialmente disponible que cataliza la adición de adenosina al extremo 3' del ARN de una manera independiente de la secuencia. En algunas realizaciones, cualquier enzima adecuada para añadir una cola de poli (A) al extremo 3' de un ARN, de una manera independiente de la secuencia encuentra uso en la presente invención. La poli-A polimerasa también se conoce como: polinucleótido adenililtransferasa, NTP polimerasa, enzima de adenilación de ARN, AMP polinucleotidilxotransferasa, ATP-polinucleótido adenililtransferasa, ATP-polinucleotidilxotransferasa, poli(A) sintetasa, poliadenilato nucleotidiltransferasa, poliadenilato polimerasa, poliadenilato sintetasa, ácido poliadenílico polimerasa, polimerasa poliadenilica, riboadenilato terminal transferasa, factores de formación de poli(A) hidrolasa ARN, PF1 o adenosina trifosfato; ácido ribonucleico adenililtransferasa. En algunas realizaciones, cualquier poli (A) polimerasa adecuada para la adición de adenosinas (por ejemplo, > 5, > 10, > 20, > 50, etc.) al extremo 3' de un molde de ARN inicial (por ejemplo, ARN genómico viral) encuentra uso con realizaciones de la presente invención. En algunas realizaciones, cualquier poli (A) polimerasa adecuada para producir un sitio de unión para un promotor de ARN polimerasa (por ejemplo, ARN polimerasa de T7) encuentra uso en realizaciones de la presente invención. En algunas realizaciones, cualquier polimerasa adecuada para la síntesis de ARN independiente de molde encuentra uso con las realizaciones descritas en el presente documento. En algunas realizaciones, cualquier polimerasa adecuada para la síntesis de poli(A) independiente de molde encuentra uso con las realizaciones descritas en el presente documento. Por ejemplo, se pueden usar poli (A) polimerasas tales como poli (A) polimerasa I de *E. coli* o poli (A) polimerasa de levadura. También se pueden usar poli (A) polimerasas de origen de mamífero o viral. Las enzimas poli (A) polimerasas recombinantes pueden usarse en los procedimientos y kits de la presente divulgación. Las enzimas poli (A) polimerasas se describen en, por ejemplo, Yehudai-Resheff, S. et al. (2000); Mohanty y Kushner (1999); Mohanty y Kushner (2000); Cao y Sarkar (1997); Raynal et al. (1996).

En algunas realizaciones, los procedimientos del presente documento implican la amplificación de ARN usando el procedimiento de Eberwine (Van Gelder et al. Proc Natl Acad Sci USA. 1990 Mar; 87(5):1663–7; Eberwine et al. Proc Natl Acad Sci USA. 1992; 89:3010–3014), procedimientos de Eberwine modificados y porciones de los mismos. El procedimiento de Eberwine aprovecha el hecho de que el ARNm eucariótico tiene una cola de poli (A) de origen natural. Este procedimiento usa la cola de poli (A) para unir un promotor de ARN polimerasa de T7 al ARNm, seguido de amplificación de ARN extensa con ARN polimerasa de T7.

En algunas realizaciones, se describe el marcaje (por ejemplo, marcaje en el extremo 3') del ARN objetivo con un marcador de ácido nucleico (por ejemplo, el tramo de poli (A)) (por ejemplo, químicamente, enzimáticamente). En algunas realizaciones, la poliadenilación en 3' se realiza mediante una poli (A) polimerasa (o un equivalente funcional) a una muestra para producir una cola de poli (A) sobre ARN no poliadenilado (por ejemplo, ARN viral). Los expertos en la técnica conocen procedimientos para añadir secuencias de ácidos nucleicos a moléculas de ARN. Por ejemplo, en algunas realizaciones, se usa una poli (A) polimerasa, tal como poli (A) polimerasa I de *E. coli* o poli (A) polimerasa de levadura, para añadir una secuencia de poli (A) al extremo 3' de una ARN.

En algunas realizaciones, el ADN complementario (ADNc) se sintetiza a partir de un molde de ARN (por ejemplo, ARN objetivo, ARN objetivo poliadenilado, ARN objetivo marcado, etc.) en una reacción catalizada por una enzima transcriptasa inversa y una enzima ADN polimerasa. En algunas realizaciones, se añaden a la muestra un cebador de transcripción inversa (cebador TI), transcriptasa inversa y desoxinucleótido trifosfatos (A, T, G, C). En algunas realizaciones, un cebador de TI comprende un segmento de hibridación en 3' (por ejemplo, para hibridar a un ARN objetivo marcado) y una región promotora en 5' (por ejemplo, promotor de ARN polimerasa (por ejemplo, promotor de ARN polimerasa de T7)). En algunas realizaciones, un cebador de TI comprende un segmento 3'-poli(T) y una región promotora en 5' (por ejemplo, promotor de ARN polimerasa (por ejemplo, promotor de ARN polimerasa de T7)). En algunas realizaciones se sintetiza una cadena de ADN que es complementaria al ARN objetivo. En algunas realizaciones, se produce un ADN / ARN híbrido (una hebra de cada uno). En algunas realizaciones, para sintetizar una cadena de ADN adicional (para producir un ADNc dúplex), el ARN objetivo de la cadena híbrida se digiere (por ejemplo, usando una enzima como ARNasa H). Después de la digestión del ARN, queda un ADN monocatenario (ADNmc). En algunas realizaciones, debido a que los ácidos nucleicos monocatenarios son hidrófobos, tiende a enrollarse alrededor de sí mismo y formar un lazo en horquilla en 3'. En algunas realizaciones, a partir del bucle de la horquilla, una ADN polimerasa puede usarlo como un cebador para transcribir una secuencia complementaria del ADNc mc, dando como resultado un ADNc bicatenario con secuencia idéntica al ARN objetivo. Cualquier otro procedimiento adecuado para la síntesis de ADNc y / o la síntesis de las segundas hebras está dentro del alcance de las realizaciones de la presente invención (Véase, por ejemplo, Nature Methods 2, 151–152 (2005); D'Alessio & Gerard. Nucleic Acids Res. 1988 March 25; 16(5 Pt B): 1999–2014). En algunas realizaciones descritas en el presente documento, después de la síntesis de una primera cadena de ADNc a partir de un molde de ARN, la síntesis de la segunda cadena se lleva a cabo por cualquier procedimiento adecuado (por ejemplo, usando ARNasa H y ADN polimerasa I) para producir un ADNc de doble cadena.

Cualquier transcriptasa inversa adecuada para llevar a cabo los procedimientos descritos en el presente documento encuentra uso con la presente invención. En algunas realizaciones, la transcriptasa inversa es la transcriptasa inversa del virus de la leucemia murina Moloney (MMLV) o la transcriptasa inversa del virus de la

mieloblastosis aviar (AMV). La transcriptasa inversa puede ser una transcriptasa inversa mutante, siempre que los mutantes conserven la actividad de síntesis de ADNc. Los ejemplos de mutantes de transcriptasa inversa incluyen aquellos con actividad de ARNasa H reducida o ausente (por ejemplo, Superscript™ II, Superscript® III y ThermoScript® (Invitrogen)) y aquellos con actividad potenciada a temperaturas más altas (Superscript III y ThermoScript). (Invitrogen).

En algunas realizaciones, se realizan etapas para eliminar los extremos salientes y / o adhesivos de ácidos nucleicos (por ejemplo, ADNc) a través de un procedimiento de terminación roma. En algunas realizaciones, se llevan a cabo etapas para eliminar los extremos salientes y / o adhesivos del ADNc producido por los procedimientos descritos en el presente documento mediante técnicas comprendidas en la materia. En algunas realizaciones se producen extremos romos en el ADNc. Los extremos romos se pueden realizar mediante diversas enzimas y / o combinaciones de enzimas. Por ejemplo, la ADN ligasa de T4 formará los extremos romos en el ADN rellenando los salientes de 5' usando su actividad de polimerasa 5'-3', o recortando los salientes en 3' con su actividad de exonucleasa 3'-5'. De manera similar, la nucleasa de frijol mungo cortará los salientes 5' o 3' como una exonucleasa. Cualesquiera otras técnicas apropiadas de formación de extremos romos conocidas en la materia encuentran uso en el presente documento.

En algunas realizaciones descritas en el presente documento, los ADNc producidos por los procedimientos descritos en el presente documento se transcriben mediante una ARN polimerasa (por ejemplo, ARN polimerasa de T7) en ARN (por ejemplo, ARN antisentido amplificado (ARNa)). En algunas realizaciones descritas en el presente documento, pueden producirse copias múltiples de ARNa a partir de un único ADNc a través de la transcripción de ARN, amplificando así la secuencia. En algunas realizaciones, una secuencia de ácido nucleico se amplifica mediante síntesis procesiva de múltiples moléculas de ARN a partir de un único molde de ADNc, lo que da como resultado una ARN antisentido amplificado (ARNa). Los procedimientos para la síntesis de ARNa se describen en las patentes de Estados Unidos n.º 5.545.522, 5.716.785 y 5.891.636. En algunas realizaciones descritas en el presente documento, estos procedimientos implican la incorporación de un promotor de ARN polimerasa en una molécula de ADNc cebando la síntesis de ADNc con un complejo de cebador que comprende un oligonucleótido sintético que contiene el promotor. Después de la síntesis de ADNc de doble cadena, se agrega una transcriptasa inversa y el ARN antisentido se transcribe a partir del molde de ADNc. La amplificación, que típicamente será de al menos aproximadamente 500 veces, pero puede ser de al menos aproximadamente 1.000, 5.000, 10.000, 15.000 o 20.000 veces o más, se puede lograr a partir de cantidades de nanogramos o menos de ADNc. Una ventaja de la síntesis procesiva de ARNa sobre la PCR es que solo se necesita conocer una región de secuencia compartida para sintetizar ARNa; la PCR generalmente requiere que las secuencias compartidas sean conocidas tanto en 5' como en 3' a la región de interés, y que estas regiones flanqueantes estén suficientemente cerca para permitir una amplificación eficiente.

En algunas realizaciones, se produce ADNc que contiene un promotor para la ARN polimerasa de SP6, T3 o T7. La ARN polimerasa utilizada para la transcripción debe ser capaz de unirse operativamente a la región promotora particular empleada en el complejo promotor-cebador. Una ARN polimerasa preferente es la que se encuentra en los bacteriófagos, en particular los fagos SP6, T3 y T7.

En algunas realizaciones, el ARNa amplificado creado por los procedimientos descritos en el presente documento se transcribe de forma inversa en ADNc. Las transcriptasas inversas adecuadas se describen en el presente documento y las condiciones adecuadas para dicha transcripción inversa están dentro de la experiencia de los expertos en el campo.

Se dispone de una serie de procesos dependientes de molde para amplificar las secuencias presentes en una muestra dada. Un ejemplo no limitante es la reacción en cadena de la polimerasa (denominada PCR) que se describe con detalle en las patentes de Estados Unidos números 4.683.195, 4.683.202 y 4.800.159, y en Innis et al, 1988. Otros procedimientos no limitantes para la amplificación de secuencias de ácido nucleico objetivo que se pueden usar en la práctica de la presente invención se divulgan en las patentes de Estados Unidos números 5.843.650, 5.846.709, 5.846.783, 5.849.546, 5.849.497, 5.849.547, 5.858.652, 5.866.366, 5.916.776, 5.922.574, 5.928.905, 5.928.906, 5.932.451, 5.935.825, 5.939.291 y 5.942.391, solicitud de GB N° 2 202 328, y en la solicitud de PCT N°. PCT/US89/01025.

En algunas realizaciones, se puede realizar un procedimiento de amplificación por PCR con transcriptasa inversa para amplificar poblaciones de ARN. Los procedimientos de transcripción inversa del ARN en ADNc son bien conocidos (véase Sambrook 1989). Procedimientos alternativos para la transcripción inversa usan ADN polimerasas termoestables. Estos procedimientos se describen en el documento WO 90/07641. Adicionalmente, se describen procedimientos representativos de RT-PCR en la patente de Estados Unidos N.º 5.882.864.

Otros procedimientos de amplificación de ácido nucleico no limitantes incluyen sistemas de amplificación basados en la transcripción (TAS), incluida la amplificación basada en la secuencia de ácido nucleico (NASBA) y 3SR (Kwoh y col., 1989; Gingeras y col., solicitud de PCT WO 88/10315). La solicitud europea N.º 329 822 divulga un procedimiento de amplificación de ácido nucleico que implica sintetizar cíclicamente el ARN monocatenario ("ARNmc"), ADNmc y ADN bicatenario (ADNbc), que se puede usar de acuerdo con la presente invención.

La amplificación basada en secuencia de ácido nucleico (NASBA) (Guatelli, 1990; Compton, 1991) hace uso de tres enzimas, transcriptasa inversa del virus de la mieloblastosis aviar (AMV-RT), ARNasa H de *E. coli* y ARN polimerasa de T7 para inducir ciclos repetidos de transcripción inversa y transcripción de ARN. La reacción de NASBA comienza con el cebado de la síntesis de ADNc de la primera cadena con un oligonucleótido específico de gen (cebador 1) que comprende un promotor de ARN polimerasa de T7. La ARNasa H digiere el ARN en el dúplex de ADN:ARN resultante que proporciona acceso de un cebador (o cebadores) específico del objetivo aguas arriba (cebador 2) a la copia de ADNc del objetivo u objetivos de ARN específicos. AMV-RT extiende el segundo cebador, produciendo un segmento de ADNc de doble cadena (ADNdc) con un promotor de polimerasa de T7 en un extremo. El ADNc sirve como molde para la ARN polimerasa de T7 que sintetizará muchas copias de ARN en la primera fase de la reacción NASBA cíclica. El ARN sirve como moldes para una segunda ronda de transcripción inversa con el segundo cebador específico del gen, produciendo finalmente más moldes de ADN que respaldan la transcripción adicional. En algunas realizaciones, NASBA puede adaptarse a la presente invención para proporcionar amplificación de las secuencias objetivo.

En ciertas realizaciones, se usa amplificación por desplazamiento de cadena (SDA). La SDA es un esquema de amplificación isotérmica que incluye cinco etapas: unión de cebadores de amplificación a una secuencia objetivo, extensión de los cebadores mediante una polimerasa de exonucleasa deficiente que incorpora un trifosfato de desoxinucleósido alfa-tio, mellado del ácido nucleico de doble cadena hemifosforotioato en un sitio de restricción, disociación de la enzima de restricción del sitio de la muesca, y extensión desde el extremo 3' de la muesca mediante una polimerasa deficiente en exonucleasa con desplazamiento de la cadena no molde aguas abajo. La formación de mellas, la polimerización y el desplazamiento se producen de forma concurrente y continua a una temperatura constante, ya que la extensión desde la mella regenera otro sitio de restricción de hemifosforotioato. En realizaciones en las que se usan cebadores para ambas cadenas de una secuencia objetivo bicatenaria, la amplificación es exponencial, ya que las cadenas sentido y antisentido sirven como moldes para el cebador opuesto en rondas de amplificación posteriores. En algunas realizaciones, SDA puede adaptarse a la presente invención para proporcionar amplificación de las secuencias objetivo.

En ciertas realizaciones, se usa amplificación por transcripción de ARN. Las moléculas de ADN con promotores pueden ser moldes para una cualquiera de varias ARN polimerasas (Sambrook 1989). Una reacción de transcripción *in vitro* eficiente puede convertir un único molde de ADN en cientos e incluso miles de transcritos de ARN. Si bien este nivel de amplificación es órdenes de magnitud menor que la que se logra mediante otros procedimientos de amplificación, es suficiente para la amplificación de ácidos nucleicos, y proporciona síntesis de ácidos nucleicos sin la necesidad de cebadores opuestos en ninguno de los extremos de una secuencia deseada.

En algunas realizaciones, los procedimientos y kits divulgados en el presente documento encuentran uso con técnicas de biología molecular adicionales para el análisis de ácidos nucleicos o que implican el mismo. En algunas realizaciones, las secuencias de ácido nucleico (por ejemplo, polimorfismos de un solo nucleótido, mutaciones, secuencia completa, etc.) se detectan y / o se determinan por cualquier procedimiento adecuado (por ejemplo, ensayo de detección de ácido nucleico, secuenciación de ácido nucleico, hibridación de ácido nucleico, etc.). El alcance de la presente invención no está limitado por la aplicación, los procedimientos o las técnicas con los que encuentran uso (por ejemplo, detección de ácidos nucleicos, secuenciación de ácidos nucleicos, identificación de ácidos nucleicos, etc.).

En algunas realizaciones, las composiciones y procedimientos descritos en el presente documento encuentran uso con la secuenciación de ácidos nucleicos (por ejemplo, producción de material de partida para la secuenciación a partir de pequeñas cantidades de ARN viral). En algunas realizaciones, los procedimientos proporcionados en el presente documento se usan para preparar un molde de ácido nucleico a partir de una muestra para su uso en la secuenciación de ácido nucleico. En algunas realizaciones, se detecta ácido nucleico y / o se determina la secuencia de un ácido nucleico mediante secuenciación de ácido nucleico. En algunas realizaciones, el ácido nucleico se secuenciará usando cualquier tipo de tecnología de secuenciación adecuada. La presente invención no está limitada por el tipo de método de secuenciación usado. A continuación se describen ejemplos de procedimientos de secuenciación. Entre los ejemplos no limitantes ilustrativos de las técnicas de secuenciación de ácidos nucleicos se incluyen, pero no se limitan a, secuenciación de terminación de cadena (Sanger), secuenciación de terminación de colorante y métodos de secuenciación de la próxima generación.

La secuenciación de terminación de cadena utiliza la terminación específica de secuencia de una reacción de síntesis de ADN usando sustratos de nucleótidos modificados. La extensión se inicia en un sitio específico en el ADN molde utilizando un cebador oligonucleotídico radioactivo corto, u otro marcado, complementario del molde en esa región. El cebador oligonucleotídico se extiende utilizando una ADN polimerasa, cuatro bases de desoxinucleótidos estándar y una concentración baja de un nucleótido de terminación de cadena, más habitualmente un didesoxinucleótido. Esta reacción se repite en cuatro tubos separados, turnándose cada una de las bases como didesoxinucleótido. La incorporación limitada del nucleótido de terminación de cadena por la ADN polimerasa da como resultado una serie de fragmentos de ADN relacionados que se terminan solamente en las posiciones en las que se utiliza dicho didesoxinucleótido particular. Para cada tubo de reacción, los fragmentos se separan por tamaño mediante electroforesis en un gel de poliacrilamida plano o un tubo capilar relleno con un polímero viscoso. La secuencia se determina leyendo qué calle produce una marca visualizada desde el cebador marcado mientras se

escanea desde la parte superior del gel hasta la parte inferior.

La secuenciación del terminador de colorante marca, alternativamente los terminadores. La secuenciación completa se puede llevar a cabo en una sola reacción marcando cada uno de los terminadores de cadena de los didesoxinucleótidos con un colorante fluorescente separado, que brilla a una longitud de onda diferente.

Un conjunto de métodos conocidos como técnicas de "secuenciación de próxima generación" han surgido como alternativas a los métodos de secuenciación de Sanger y por terminador colorante (Voelkerding et al., *Clinical Chem.*, 55: 641–658, 2009; MacLean et al., *Nature Rev. Microbiol.*, 7: 287–296). La tecnología de secuenciación de próxima generación permite la secuenciación *de novo* de genomas completos para determinar la secuencia de ácidos nucleicos primaria de un organismo. La tecnología de secuenciación de próxima generación también proporciona resecuenciación específica (secuenciación profunda), que permite la detección de mutaciones sensibles dentro de una población de secuencias de tipo salvaje. Algunos ejemplos incluyen trabajos recientes que describen la identificación de variantes de VIH resistentes a fármacos, así como mutaciones del EGFR para determinar la respuesta a fármacos terapéuticos anti-TK. Las publicaciones que describen la secuenciación de próxima generación permiten la secuenciación simultánea de múltiples muestras durante un ciclo de secuenciación típica, incluyendo, por ejemplo: Margulies, M. et al. "Genome Sequencing in Microfabricated High-Density Picolitre Reactors", *Nature*, 437, 376–80 (2005); Mikkelsen, T. et al. "Genome-Wide Maps of Chromatin State in Pluripotent and Lineage-Committed Cells", *Nature*, 448, 553–60 (2007); McLaughlin, S. et al. "Whole-Genome Resequencing with Short Reads: Accurate Mutation Discovery with Mate Pairs and Quality Values", *ASHG Annual Meeting* (2007); Shendure J. et al. "Accurate Multiplex Polony Sequencing of an Evolved Bacterial Genome", *Science*, 309, 1728–32 (2005); Harris, T. et al. "Single-Molecule DNA Sequencing of a Viral Genome", *Science*, 320, 106–9 (2008); Simen, B. et al. "Prevalence of Low Abundance Drug Resistant Variants by Ultra Deep Sequencing in Chronically HIV-infected Antiretroviral (ARV) Naive Patients and the Impact on Virologic Outcomes", 16th International HIV Drug Resistance Workshop, Barbados (2007); Thomas, R. et al. "Sensitive Mutation Detection in Heterogeneous Cancer Specimens by Massively Parallel Picoliter Reactor Sequencing", *Nature Med.*, 12, 852–855 (2006); Mitsuya, Y. et al. "Minority Human Immunodeficiency Virus Type 1 Variants in Antiretroviral-Naive Persons with Reverse Transcriptase Codon 215 Revertant Mutations", *J. Vir.*, 82, 10747–10755 (2008); Binladen, J. et al. "The Use of Coded PCR Primers Enables High-Throughput Sequencing of Multiple Homolog Amplification Products by 454 Parallel Sequencing", *PLoS ONE*, 2, e197 (2007); y Hoffmann, C. et al. "DNA Bar Coding and Pyrosequencing to Identify Rare HIV Drug Resistance Mutations", *Nuc. Acids Res.*, 35, e91 (2007).

En comparación con la secuenciación tradicional de Sanger, la tecnología de secuenciación de próxima generación produce grandes cantidades de puntos de datos de secuenciación. Un ciclo típico puede generar fácilmente de decenas a cientos de megabases por ciclo, con una potencia salida diaria que llega al rango de gigabases. Esto se traduce en varios órdenes de magnitud mayor que una placa estándar de 96 pocillos, que puede generar varios cientos de puntos de datos en ciclo múltiplex típico. Los amplicones diana que difieren en tan poco como un nucleótido pueden distinguirse fácilmente, incluso cuando están presentes múltiples dianas de especies u organismos relacionados. Esto mejora en gran medida la capacidad de realizar un genotipado exacto. Los programas de software de alineación de secuencia de próxima generación usados para producir secuencias consenso pueden identificar fácilmente nuevas mutaciones puntuales, lo que podría dar lugar a nuevas cepas con resistencias a fármacos asociadas. El uso de la codificación de barras de cebadores también permite la multiplexación de diferentes muestras de pacientes dentro de un solo ciclo de secuenciación.

Los métodos de secuenciación de próxima generación (NGS) comparten la característica común de las estrategias masivas paralelas de alto rendimiento, con el objetivo de reducir los costes en comparación con los métodos de secuenciación más antiguos. Los métodos de NGS pueden dividirse en líneas generales en aquellos que requieren amplificación del molde y los que no lo hacen. Los métodos que requieren amplificación incluyen la pirosecuenciación comercializada por Roche como las plataformas tecnológicas 454 (por ejemplo, GS 20 y GS FLX), la plataforma Solexa comercializada por Illumina y la plataforma de Supported Oligonucleotide Ligation and Detection (SOLiD) comercializada por Applied Biosystems. Los enfoques de no amplificación, también conocidos como secuenciación de una sola molécula, se ilustran mediante la plataforma HeliScope comercializada por Helicos BioSciences, y plataformas emergentes comercializadas por VisiGen y Pacific Biosciences, respectivamente.

En la pirosecuenciación (Voelkerding et al., *Clinical Chem.*, 55: 641–658, 2009; MacLean et al., *Nature Rev. Microbiol.*, 7: 287–296; patente de Estados Unidos n.º 6.210.891; patente de Estados Unidos n.º 6.258.568), el ADN molde se fragmenta, se reparan los extremos, se liga a los adaptadores y se amplifica clonalmente *in situ* capturando moléculas molde sencillas con perlas que llevan oligonucleótidos complementarios a los adaptadores. Cada perla portadora de un solo tipo de molde se compartimenta en una microvesícula de agua en aceite y el molde se amplifica clonalmente utilizando una técnica denominada PCR en emulsión. La emulsión se rompe después de la amplificación y las perlas se depositan en pocillos individuales de una placa de picotitulación que funciona como una celda de flujo durante las reacciones de secuenciación. La introducción iterativa ordenada de cada uno de los cuatro reactivos de dNTP se produce en la celda de flujo en presencia de enzimas de secuenciación e indicador luminiscente, tal como luciferasa. En el caso de que se añada un dNTP apropiado al extremo 3' del cebador de secuenciación, la producción resultante de ATP provoca una explosión de luminiscencia dentro del pocillo, que se registra utilizando una cámara CCD. Es posible conseguir longitudes de lectura mayores o iguales a 400 bases 400



bases y se pueden conseguir  $1 \times 10^6$  lecturas de secuencias, dando como resultado hasta 500 millones de pares de bases (Mb) de secuencia.

En la plataforma Solexa / Illumina (Voelkerding et al., *Clinical Chem.*, 55: 641–658, 2009; MacLean et al., *Nature Rev. Microbiol.*, 7: 287–296; patente de Estados Unidos n.º 6.833.246; patente de Estados Unidos n.º 7.115.400; patente de Estados Unidos n.º 6.969.488), los datos de secuenciación se producen en forma de lecturas de longitud más corta. En este método, el ADN fragmentado monocatenario se repara en los extremos para generar extremos romos fosforilados en 3', seguido de la adición mediada por el fragmento Klenow de una única base A al extremo 3' de los fragmentos. La adición de A facilita la adición de oligonucleótidos adaptadores sobresalientes de T, que se utilizan posteriormente para capturar las moléculas adaptadoras del molde sobre la superficie de una celda de flujo que está tachonada con anclajes de oligonucleótidos. El anclaje se utiliza como cebador de PCR, pero debido a la longitud del molde y su proximidad a otros oligonucleótidos de anclaje cercanos, la extensión por PCR da como resultado el "arqueamiento" de la molécula para hibridar con un oligonucleótido de anclaje adyacente para formar una estructura de puente sobre la superficie de la celda de flujo. Estos bucles de ADN son desnaturizados y escindidos. A continuación, las hebras delanteras se secuencian con terminadores colorantes reversibles. La secuencia de los nucleótidos incorporados se determina mediante detección de fluorescencia posterior a la incorporación, con cada fluorescencia y bloque eliminados antes del siguiente ciclo de adición de dNTP. La longitud de la lectura de la secuencia varía desde 36 nucleótidos a más de 50 nucleótidos, con una producción total superior a mil millones de pares de nucleótidos por análisis.

La secuenciación de moléculas de ácido nucleico usando tecnología SOLID (Voelkerding et al., *Clinical Chem.*, 55: 641–658, 2009; MacLean et al., *Nature Rev. Microbiol.*, 7: 287–296; patente de Estados Unidos n.º 5.912.148; patente de Estados Unidos n.º 6.130.073) también implica la fragmentación del molde la unión a los adaptadores de oligonucleótidos, la unión a perlas y la amplificación clonal por PCR en emulsión. Después de esto, las perlas portadoras de molde se inmovilizan sobre una superficie derivatizada de una celda de flujo de vidrio y se hibrida un cebador complementario al oligonucleótido adaptador. Sin embargo, en lugar de utilizar este cebador para la extensión en 3', en lugar de ello se utiliza para proporcionar un grupo fosfato en 5' para la unión sondas identificadoras que contienen dos bases específicas de sonda seguidas de 6 bases degeneradas y uno de cuatro marcadores fluorescentes. En el sistema SOLID, las sondas identificadoras tienen 16 posibles combinaciones de las dos bases en el extremo 3' de cada sonda y uno de cuatro flúor en el extremo 5'. El color fluorado y, por lo tanto, la identidad de cada sonda corresponde a esquemas de codificación de espacio de color especificados. A varias rondas (normalmente 7) de hibridación de la sonda, unión y detección de flúor les sigue desnaturización y, después, una segunda ronda de secuenciación usando un cebador que está compensado por una base con respecto al cebador inicial. De esta manera, la secuencia molde puede reconstruirse por ordenador y las bases del molde se interrogan dos veces, lo que da como resultado una mayor precisión. La longitud de lectura de la secuencia es de 35 nucleótidos y la producción total supera los 4 mil millones de bases por ciclo de secuenciación.

En ciertas realizaciones, se usa secuenciación en nanoporos (véase, por ejemplo, Astier et al., *J Am Chem Soc.* 2006 Feb. 8; 128(5):1705–10). La teoría detrás de la secuenciación en nanoporos tiene que ver con lo que se produce cuando el nanoporo se sumerge en un fluido conductor y se aplica un potencial (voltaje) a su través: en estas condiciones se puede observar una ligera corriente eléctrica debido a la conducción de iones a través del nanoporo y la cantidad de corriente es extremadamente sensible al tamaño del nanoporo. Si las moléculas de ADN pasan (o parte de la molécula de ADN pasa) a través del nanoporo, se puede crear un cambio en la magnitud de la corriente a través del nanoporo, lo que permite determinar las secuencias de la molécula de ADN.

HeliScope de Helicos BioSciences (Voelkerding et al., *Clinical Chem.*, 55: 641–658, 2009; MacLean et al., *Nature Rev. Microbiol.*, 7: 287–296; patente de Estados Unidos n.º 7.169.560; patente de Estados Unidos n.º 7.282.337; patente de Estados Unidos n.º 7.482.120; patente de Estados Unidos n.º 7.501.245; patente de Estados Unidos n.º 6.818.395; patente de Estados Unidos n.º 6.911.345; patente de Estados Unidos n.º 7.501.245) no requiere amplificación clonal. El ADN molde se fragmenta y poliadenila en el extremo 3', portando la adenosina final un marcador fluorescente. Los fragmentos del molde poliadenilados desnaturizados se ligan a oligonucleótidos poli(dT) sobre la superficie de una celda de flujo. Las localizaciones físicas iniciales de las moléculas molde capturadas son grabadas por una cámara CCD y, después, se escinde el marcador y se elimina mediante lavado. La secuenciación se consigue mediante la adición de polimerasa y la adición en serie de reactivos de dNTP marcados con fluorescencia. Los acontecimientos de incorporación producen una señal de flúor correspondiente al dNTP y la señal es capturada por una cámara CCD antes de cada ronda de adición de dNTP. La longitud de la lectura de la secuencia varía desde 25 nucleótidos a 50 nucleótidos, con una producción total superior a mil millones de pares de nucleótidos por análisis.

Otros métodos emergentes de secuenciación de moléculas individuales incluyen la secuenciación en tiempo real mediante la síntesis utilizando una plataforma VisiGen (Voelkerding et al., *Clinical Chem.*, 55: 641–658, 2009; patente de Estados Unidos n.º 7.329.492; solicitud de patente de Estados Unidos n.º de serie 11/671.956; solicitud de patente de Estados Unidos n.º de serie 11/781.166) en la que el molde de ADN cebado inmovilizado se somete a extensión de cadena usando una polimerasa modificada por fluorescencia y moléculas aceptoras fluorescentes, lo que da como resultado una transferencia de energía de resonancia de fluorescencia detectable (FRET) después de la adición de nucleótidos.

Otro sistema de secuenciación de moléculas individuales en tiempo real desarrollado por Pacific Biosciences (Voelkerding et al., *Clinical Chem.*, 55: 641–658, 2009; MacLean et al., *Nature Rev. Microbiol.*, 7: 287–296; patente de Estados Unidos n.º 7.170.050; patente de Estados Unidos n.º 7.302.146; patente de Estados Unidos n.º 7.313.308; patente de Estados Unidos n.º 7.476.503) utiliza pocillos de reacción de 50-100 nm de diámetro y que abarca un volumen de reacción de aproximadamente 20 zeptolitros ( $10 \times 10^{-21}$  l). Las reacciones de secuenciación se llevan a cabo usando un molde inmovilizado, ADN polimerasa de phi29 modificada y concentraciones locales altas de dNTP marcados con fluorescencia. Las concentraciones locales altas y las condiciones de reacción continuas permiten capturar acontecimientos de incorporación en tiempo real mediante la detección de la señal fluorescente mediante excitación por láser, una guía de ondas óptica y una cámara CCD.

En algunas realizaciones, las composiciones y procedimientos descritos en el presente documento encuentran uso con detección de ácido nucleico, caracterización y/o identificación. En algunas realizaciones, los procedimientos proporcionados en el presente documento se usan para preparar un molde de ácido nucleico a partir de una muestra para su uso en el análisis posterior. En algunas realizaciones, el ácido nucleico se detecta, caracteriza y/o identifica usando cualquier tecnología adecuada. La presente invención no está limitada por el tipo de métodos de análisis de ácido nucleico usados. Ejemplos de procedimientos se describen a continuación.

En algunas realizaciones, el ácido nucleico se detecta mediante hibridación con una sonda oligonucleotídica. Se dispone de diversos ensayos de hibridación que usan varias tecnologías de hibridación y detección. Por ejemplo, en algunas realizaciones se usa el ensayo TaqMan (PE Biosystems, Foster City, Calif.; véanse, por ejemplo, las patentes de Estados Unidos n.º 5.962.233 y 5.538.848) El ensayo se realiza durante una reacción de PCR. El ensayo TaqMan explota la actividad de exonucleasa 5'-3' de la ADN polimerasa AMPLITAQ GOLD. Una sonda compuesta por un oligonucleótido con un pigmento indicador en 5' (por ejemplo, un pigmento fluorescente) y un pigmento inactivador en 3' se incluye en la reacción de PCR. Durante la PCR, si la sonda se une a su diana, la actividad nucleolítica 5'-3' de la polimerasa AMPLITAQ GOLD escinde la sonda entre el pigmento indicador y el inactivador. La separación del pigmento indicador del pigmento inactivador tiene como resultado un incremento de la fluorescencia. La señal se acumula con cada ciclo de PCR y se puede monitorizar con un fluorímetro.

En algunas realizaciones, el ácido nucleico se detecta mediante análisis de transferencia Northern (por ejemplo, tras amplificación mediante los procedimientos descritos en el presente documento). El análisis de transferencia Northern implica la separación de ácido nucleico y la hibridación de una sonda marcada complementaria.

En algunas realizaciones, el ácido nucleico se detecta utilizando un ensayo de detección (por ejemplo, después de la amplificación por los procedimientos descritos en el presente documento) que incluye, pero no se limita a los mismos, procedimientos de escisión por apareamiento erróneo enzimático (por ejemplo, Variagenics, Patentes de Estados Unidos n.º 6.110.684, 5.958.692, 5.851.770); reacción en cadena de la polimerasa; procedimientos de hibridación ramificada (por ejemplo, Chiron, Patentes de Estados Unidos n.º 5.849.481, 5.710.264, 5.124.246, y 5.624.802); replicación por círculo rodante (por ejemplo, Patentes de Estados Unidos n.º 6.210.884, 6.183.960 y 6.235.502); NASBA (por ejemplo, Patente de Estados Unidos n.º 5.409.818); tecnología de balizas moleculares (por ejemplo, Patente de Estados Unidos n.º 6.150.097); tecnología de sensor E (Motorola, Patentes de Estados Unidos n.º 6.248.229, 6.221.583, 6.013.170 y 6.063.573); tecnología de sonda cíclica (por ejemplo, Patentes de Estados Unidos n.º 5.403.711, 5.011.769 y 5.660.988); procedimientos de amplificación de señal de Dade Behring (por ejemplo, Patentes de Estados Unidos n.º 6.121.001, 6.110.677, 5.914.230, 5.882.867, and 5.792.614); reacción en cadena de la ligasa (Barnay Proc. Natl. Acad. Sci USA 88, 189–93 (1991)); ensayos de FULL-VELOCITY; y procedimientos de hibridación en sándwich (por ejemplo, Patente de Estados Unidos n.º 5.288.609). En otras realizaciones, el ensayo de detección empleado es el ensayo INVADER (Third Wave Technologies) que se describe en las patentes de EE.UU. N.º 5.846.717, 5.985.557, 5.994.069, 6.001.567, y 6.090.543, WO 97/27214, WO 98/42873, Lyamichev et al., *Nat. Biotech.*, 17: 292 (1999), y Hall et al., *PNAS, USA*, 97: 8272 (2000).

Cualquiera de las composiciones descritas en el presente documento puede proporcionarse en un kit. En un ejemplo no limitativo, el kit, en un medio contenedor adecuado, comprende, uno o más de: una poli (A) polimerasa, transcriptasa inversa, ARN polimerasa, nucleótidos; tampón; ARN de control, oligonucleótido, etc. Cualquiera de los otros reactivos tratados en el presente documento o implicados por los procedimientos descritos también puede incluirse en un kit.

**Reivindicaciones**

1. Un procedimientos de amplificación de ARN viral de secuencia desconocida de una muestra, que comprende:
- 5 a) añadir una poli-A-polimerasa a una muestra de ARN viral no poliadenilado para unir un tramo de poli (A) al extremo 3' del ARN viral;
- (b) hibridar un oligonucleótido con el tramo de poli (A), en el que dicho oligonucleótido comprende un dominio de hibridación que comprende un tramo de poli (T) y un dominio funcional que comprende una región promotora de ARN polimerasa;
- 10 (c) sintetizar ADNc bicatenario a partir del ARN viral y oligonucleótido, en el que la síntesis del ADNc bicatenario a partir del ARN viral y el oligonucleótido comprende la transcripción inversa de una primera cadena de dicho ADNc y la síntesis de una segunda cadena de dicho ADN;
- (d) amplificar el ADNc bicatenario con una ARN polimerasa para producir ARN antisentido amplificado; y
- 15 (e) convertir el ARN antisentido amplificado en ADNc amplificado.
2. El procedimiento de la reivindicación 1, en el que dicho dominio de hibridación está en el extremo 3' de dicho oligonucleótido y dicho dominio funcional está en el extremo 5' de dicho oligonucleótido.
3. El procedimiento de la reivindicación 1, en el que dicha región promotora de ARN polimerasa comprende una
- 20 región promotora de T7.
4. El procedimiento de la reivindicación 1, en el que dicha segunda cadena de dicho ADN se sintetiza usando ARNasa H y ADN polimerasa.
- 25 5. El procedimiento de la reivindicación 1, en el que dicho ARN antisentido amplificado:
- (i) es al menos una amplificación de 50 veces de dicho ADNc de doble hebra;
- (ii) es al menos una amplificación de 100 veces de dicho ADNc de doble hebra;
- (iii) es al menos una amplificación de 500 veces de dicho ADNc de doble hebra; y/o
- 30 (iv) es al menos una amplificación de 1000 veces de dicho ADNc de doble hebra.
6. El procedimiento de la reivindicación 1, en el que la conversión del ARN antisentido amplificado en ADNc amplificado comprende la transcripción inversa.
- 35 7. El procedimiento de la reivindicación 1, en el que dicha muestra comprende una muestra biológica, ambiental y / o forense.
8. El procedimiento de la reivindicación 1, que comprende además la formación de extremos romos en el ADNc.
- 40 9. El procedimiento de la reivindicación 1, que comprende además deplecionar el ADN de la muestra.
10. El procedimiento de la reivindicación 1, que comprende además reducir el ARNm poliadenilado de la muestra antes del ARN objetivo de marcado con poli (A).
- 45 11. El procedimiento de la reivindicación 1, en el que la poli-A-polimerasa es poli (A) polimerasa I de E. coli o poli (A) polimerasa de levadura.
12. Un procedimientos de secuenciación de ARN viral de secuencia desconocida de una muestra, que comprende:
- 50 (1) amplificar dicho ARN viral por el procedimiento de la reivindicación 1; y
- (2) secuenciar dicho ARN viral usando el ADNc amplificado.
13. El procedimiento de la reivindicación 12, en el que dicha secuenciación es mediante una técnica de
- 55 secuenciación de próxima generación.

60

65