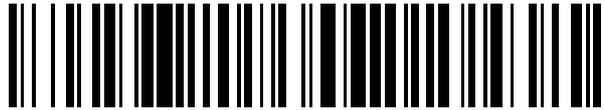


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 565 809**

51 Int. Cl.:

C12Q 1/68 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **06.03.2013 E 13714865 (6)**

97 Fecha y número de publicación de la concesión europea: **02.03.2016 EP 2823058**

54 Título: **Métodos mejorados de secuenciación de ácidos nucleicos**

30 Prioridad:

06.03.2012 US 201261607418 P

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

07.04.2016

73 Titular/es:

**ILLUMINA CAMBRIDGE LIMITED (100.0%)
Chesterford Research Park
Little Chesterford Saffron Walden Essex CB10
1XL, GB**

72 Inventor/es:

**GORMLEY, NIALL ANTHONY;
FRASER, LOUISE y
KOKKO-GONZALES, PAULA**

74 Agente/Representante:

LAZCANO GAINZA, Jesús

ES 2 565 809 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

DESCRIPCIÓN

Métodos mejorados de secuenciación de ácidos nucleicos

Antecedentes

5 Se han desarrollado técnicas de secuenciación para secuenciar ácidos nucleicos que incluyen ARN. Las técnicas de secuenciación incluyen, por ejemplo, la secuenciación mediante síntesis. La secuenciación mediante síntesis o la secuenciación en ciclo se pueden lograr mediante la adición en etapas de nucleótidos que contienen, por ejemplo, una etiqueta de tinte separable o fotolixiviable como se describió, por ejemplo, en la Patente U.S. 7, 427, 673; la Patente U.S. 7, 414, 116; la WO 04/018497; la WO 91/06678; la WO 07/123744; y la Patente U.S. No. 7, 057, 026.

10 De manera alternativa, se pueden emplear técnicas de pirosecuenciamiento. El pirosecuenciamiento detecta la liberación de pirofosfato inorgánico (PPi) en la medida en que se incorporan nucleótidos particulares en la cadena insipiente (Ronaghi et al., (1996) "Real-time DNA sequencing using detection of pyrophosphate release." *Analytical Biochemistry* 242 (1), 84-9; Ronaghi, M. (200 1) "Pyrosequencing sheds light on DNA sequencing." *Genome Res.* 11(1), 3-11; Ronaghi, M., Uhlen, M. and Nyren, P. (1998) "A sequencing method based on real-time pyrophosphate." *Science* 281(5375), 363; Patente de los Estados Unidos No. 6, 210, 891; Patente de los Estados Unidos No. 6,258,568; and Patente de los Estados Unidos No. 6,274,320. El pirosecuenciamiento, se puede detectar el PPi liberado al ser inmediatamente convertido a adenosín trifosfato (ATP) mediante ATP sulfurilasa, y el nivel del ATP generado se detecta por vía de los fotones producidos por la luciferasa.

15 Las técnicas de secuenciación también incluyen secuenciación mediante técnicas de ligado. Tales técnicas utilizan ADN ligasa para incorporar oligonucleótidos e identificar la incorporación de tales oligonucleótidos y se describen en la Patente U.S. No 6, 969,488; La Patente U.S. No. 6, 172, 218; y la Patente U.S. No. 6, 306, 597.

20 Otras técnicas de secuenciación incluyen, por ejemplo, secuenciación in situ fluorescente, (FISSEQ), y Secuenciación Masivamente Paralela de la Firma (MPSS).

25 La preparación de muestras de ADN para secuenciación pueden ser relativamente directas e incluir el uso de reacciones de transposición para fragmentar y agregar secuencias del adaptador a los fragmentos de ADN, lo que simplifica el proceso de preparación de la muestra. Ver, por ejemplo, Publicación Internacional No. WO 2010/048605, que describe un método de transposición in vitro que ambos ADN de etiquetas y fragmentos, pero no describe el etiquetado del transposón de un blanco inmovilizado. En contraste, los protocolos habituales para secuenciación de muestras de ARN emplean un método de preparación de muestra que convierte el ARN en la muestra en un formato de cADN de doble cadena antes de la secuenciación. Así, la preparación de las muestras de ARN para secuenciación es un trabajo más intensivo. Además, los protocolos corrientes son menos que óptimos por su capacidad para preservar la información específica de cadena. Más específicamente, la mayoría de los métodos no pueden preservar la información de la cadena en relación con la dirección de la molécula de ARN de cadena simple original después de ser convertida a una cADN de cadena doble. Preservar la información específica de la cadena es importante para la anotación de nuevos genes y para determinar los niveles de expresión del gen. Algunos métodos intentan preservar la información específica de cadena mediante los adaptadores de ligado en los extremos de las moléculas de ARN de cadena única. Los adaptadores pueden tener secuencias que suministran información distinguible para ambos extremos del cADN de cadena doble generado por las moléculas de ARN. Sin embargo, este método tiene desventajas. Por ejemplo, si se fragmentan las moléculas de ARN, después de que la fragmentación de las partes internas de las moléculas pierden su información direccional (es decir, específica de la cadena).

30 Resúmen

35 Se suministra aquí un método para etiquetar dúplex de ácido nucleico. El método incluye las etapas de suministrar una transposasa y una composición de transposón, suministrando uno o más dúplex de ácido nucleico inmovilizado sobre un soporte, y poner en contacto la transposasa y la composición de transposón con uno o más dúplex de ácido nucleico bajo condiciones en donde el uno o más dúplex de ácido nucleico y la composición de transposón sufre una reacción de transposición para producir uno o más dúplex de ácido nucleico etiquetado, en donde la composición de transposón comprende una molécula de ácido nucleico de cadena doble que comprende una cadena transferida y una cadena no transferida. El método también se puede efectuar al etiquetar los dúplex de ADN: ADN o ADN: ARN que son inmovilizados sobre un soporte sólido.

40 Los detalles de una o más realizaciones se establecen en los dibujos que la acompañan y en la descripción de adelante. Otras características, objetos, y ventajas serán evidentes de la descripción y los dibujos, y de las reivindicaciones.

Descripción de los dibujos

45 La Figura 1 es un esquema que muestra un método de ejemplo suministrado aquí. El mRNA con una cola de poliA se captura sobre un soporte por vía de hibridación a una sonda de captura de poliT ADN o al cebador acoplado a la superficie de un soporte. La cadena poliT es luego extendida con una polimerasa de transcriptasa inversa para hacer una molécula de cadena doble que comprende un dúplex de ADN: ARN. Luego, el complejo del transposoma (por ejemplo, una transposasa Tn5 unida con una secuencia de extremo de mosaico (ME) y secuencias complementarias a los cebadores de amplificación de superficie) se agregan al soporte, que sufre una reacción de transposición y

fragmenta el dúplex, ligando un oligo adaptador de ADN al extremo 5' de la cadena de ARN. Una polimerasa que desplaza la cadena (por ejemplo, la Bst polimerasa) se puede entonces utilizar para extender el extremo 3' de la cadena de ADN, desplazando la cadena no transferida del complejo de transposoma y copiando la cadena de ARN a su extremo quimérico de ADN 5'. La molécula de cadena doble se puede entonces amplificar (por ejemplo, amplificación de un conjunto) y secuenciada con el cebador de secuenciación. El cebador comprende parcialmente la secuencia ME y la secuencia de adaptador corriente arriba. De manera alternativa, en otro extremo de la molécula (el extremo poli T) se puede secuenciar con un cebador que hibrida corriente arriba la secuencia poliT y se extiende con un nucleótido dATP natural antes de comenzar los ciclos de secuencia mediante la química de síntesis (SBS).

La Figura 2 es un esquema que muestra un método de ejemplo suministrado aquí. El ARN se fragmenta y trata con una fosfatasa. Una molécula adaptadora de cadena única se liga al extremo 3' de cada fragmento de ARN que comprende el complemento de un cebador unido a la superficie. Los fragmentos se agregan entonces a un soporte y se capturan por vía de hibridación. Las moléculas de ARN hibridadas se convierten a un dúplex de ADN: ARN con una polimerasa de transcriptasa inversa. El complejo de transposoma o la composición que comprende una transposasa y un dúplex adaptador (es decir, el transposón) de un ME con P5 se utiliza para tagmentar el dúplex. Luego de la extensión de la cadena de ADN al extremo con una polimerasa que desplaza la cadena, las moléculas se pueden amplificar (por ejemplo, amplificación de conjunto) y secuenciar.

La Figura 3 es un esquema que muestra un método de ejemplo suministrado aquí. Se utiliza un soporte que contiene dos cebadores injertados de superficie: un cebador de injerto estándar (por ejemplo P5) y un cebador de injerto modificado (por ejemplo P7) que tiene una secuencia de captura específica de blanco en su lado corriente abajo (3'). Un ejemplo de una secuencia específica blanco es una secuencia oligo complementaria a la transcriptasa inversa retroviral (por ejemplo la VIH polimerasa). El ARN viral purificado se agrega al soporte, capturado por vía de hibridación, copiado con transcriptasa inversa y tagmentado. La secuenciación se puede lograr con un cebador hibridado al adaptador tagmentado o al otro extremo de la sonda de captura.

La Figura 4 es un esquema que muestra un método de ejemplo suministrado aquí. Los transcritos de ARN se generan de un plásmido que contiene la secuencia de transcripción de la Proteína Fluorescente Verde (GFP) y la secuencia complementaria al cebador de unión de superficie (por ejemplo, la secuencia P7'). Los transcritos se hibridan a un soporte que comprende los cebadores que comprenden, por ejemplo, una secuencia P7. Las moléculas de ARN hibridadas se convierten a un dúplex de ADN: ARN con una polimerasa de transcriptasa inversa. Un complejo de transposoma se utiliza para tagmentar el dúplex. Luego de la extensión de la cadena de ADN al extremo con una polimerasa que desplaza la cadena y la remoción de la cadena de ARN, las moléculas se pueden amplificar (por ejemplo amplificación de conjunto y secuenciar).

La Figura 5 es una fotografía de un gel que muestra los transcritos de ARN generados de un plásmido que contiene Proteína Fluorescente Verde (GFP) y, opcionalmente, tratada con DNasa para remover el ADN (es decir, plásmido). Ningún ADN residual (es decir, plásmido) fue visible luego del tratamiento con DNasa del transcrito de ARN.

La Figura 6 muestra fotografías de conjuntos tenidos con verde SYBR. Los carriles 1-4 contenían ADN PhiX y los carriles 5- 8 contenían ARN GFP. Los carriles 5 y 6 contenían ARN que fue pretratado con DNasa para remover el ADN. Los carriles 7 y 8 contenían ARN que fue pretratado con DNasa y tratado con RNasa como un control adicional. La primera extensión se llevó a cabo utilizando la Transcriptasa Inversa del Virus de la Mieloblastosis de aves (AMV - RT) (carriles 2, 4, 6 y 8) o la ADN polimerasa de Fusión (carriles 1, 3, 5, y 7). Los carriles 3- 8 fueron tagmentados con un adaptador P5. La amplificación del conjunto isotérmico se llevó a cabo de manera estándar y los conjuntos se tiñeron con un verde SYBR.

Las Figuras 7A y 7B muestran gráficas de lotes de cubrimiento de los datos de secuenciación alineados provenientes de la secuenciación de los carriles 1- 8 tal como se describió para la Figura 6.

La Figura 8 es un esquema que muestra un método de ejemplo suministrado aquí para el etiquetado del dúplex de ADN: ADN. El ADN de cadena única (ssADN) se fragmenta y los fragmentos se marcan con polinucleótidos mediante la desoxinucleotidil transferasa terminal. Los fragmentos son luego agregados a un soporte y capturados por vía de hibridación de la cola poliT con su complemento inmovilizado sobre el soporte sólido. Las moléculas de ssADN hibridadas se convierten a un dúplex de ADN: ADN con una ADN polimerasa. El complejo o composición de transposoma que comprende una transposasa y un dúplex adaptador (es decir, transposón) de una secuencia P5-seq se utiliza para tagmentar el dúplex. Luego de la extensión de la cadena de ADN al extremo con una polimerasa que desplaza la cadena, las moléculas se pueden amplificar (por ejemplo amplificación de conjunto) y secuenciar.

La Figura 9 muestra fotografías de conjuntos teñidos con Verde SYBR de un experimento diseñado para generar datos que utilizan la muestra de transcrito de mRNA completa. El carril 1 es una biblioteca estándar PhiX donde no existe tagmentación. El carril 2 es un control negativo donde no se efectuó la adición de cola y el carril 8 es un control negativo del dsADN. Los carriles 3 y 4 son el mismo experimento, exceptuando que los cebadores fueron diluidos para el carril 4. Los carriles 5 y 6 son el mismo experimento, excepto que los cebadores se diluyeron para el carril 6. El carril 7 utiliza hexámeros de ARN aleatorios en lugar de hexámeros de ADN aleatorios.

La Figura 10 muestra una fotografía del cubrimiento de lecturas de secuenciación alineadas para el GAPDH que sigue el método de la Figura 1. La captura del 3' superior demuestra la captura, fragmentación, conjunto y alineada de un

control de una muestra de mARN con cola poliA. El transcripto completo de fondo demuestra que el alineamiento del mARN de una muestra de mARN que fue enriquecida de un complejo de muestra de UHR ARN total, fragmentado enzimáticamente y poliadenilado para demostrar el cubrimiento del transcripto de mARN completo utilizando los métodos de la presente descripción.

5 Descripción detallada

Los protocolos habituales para secuenciación de las muestras de ARN emplean toda una preparación de muestra que convierte el ARN en la muestra en un formato de cADN de doble cadena antes de la secuenciación. Se suministran aquí métodos para secuenciar muestras de ARN que evitan una preparación de la fase de solución de un cADN de cadena doble intermedio. Los métodos suministrados también dan como resultado la preservación de la información encadenada durante la secuenciación. Sin embargo, los métodos descritos aquí se podrían utilizar para marcación y secuenciación para marcar y secuenciar ADN.

Se suministra aquí un método para etiquetar dúplex de ADN: ARN. El método incluye las etapas de suministrar una composición de transposasa y transposón, suministrando uno o más dúplex de ADN: ARN inmovilizados en un soporte, y poniendo en contacto la composición de transposasa y transposón con el uno o más dúplex de ADN: ARN bajo condiciones en donde uno o más dúplex de ADN: ARN y la composición de transposón sufren de una reacción de transposición para producir uno o más dúplex de ADN: ARN etiquetados. La composición de transposón comprende una molécula de ácido nucleico de doble cadena que comprende una cadena transferida y una cadena no transferida. Aunque se pueden ejemplificar los siguientes ejemplos utilizando dúplex de ADN: ARN, ellos también podrían ser responsables por los dúplex de ADN:ARN donde sea apropiado (ver Figura 8).

Opcionalmente, los uno o más dúplex de ADN: ARN se etiquetan en el extremo 5' de la cadena de ARN. Opcionalmente, la cadena transferida comprende una etiqueta para preservar la información de la cadena. La reacción de transposición da como resultado una cadena de ARN etiquetada en 5' que comprende la cadena transferida y la composición de transposón y el espacio entre el extremo 3' de la cadena de ADN y la cadena no transferida de la composición de transposón. Opcionalmente, el método comprende además poner en contacto los uno o más dúplex de ADN: ARN etiquetados con la enzima que modifica el ácido nucleico bajo condiciones para extender el extremo 3' de la cadena de ADN para copiar la cadena de ARN a su extremo 5'. La enzima que modifica el ácido nucleico puede desplazar la cadena no transferida de la composición de transposón.

La Figura 1 es un esquema que muestra un método de ejemplo suministrado aquí. En resumen, el mARN con una cola de poliA es capturado sobre un soporte (por ejemplo "flowcell") por vía de hibridación a una sonda de captura de ADN poliT (o cebador) acoplada a la superficie del soporte. La cadena poliT se extiende luego polimerasa de transcriptasa inversa para hacer una molécula de cadena doble que comprende un dúplex de ADN: ARN. Luego, una secuencia de complejo de transposoma (por ejemplo Tn5 unido con un transposón (por ejemplo, extremo de mosaico (ME)) y las secuencias complementarias a los cebadores de amplificación de superficie (se agregan al soporte, los cuales "tagmentan" el dúplex, ligando el oligo adaptador de ADN al extremo 5' de la cadena de ARN). Un polimerasa que desplaza la cadena (por ejemplo una polimerasa Bst) se puede utilizar para extender el extremo 3' de la cadena de ADN, desplazando la "cadena no transferida" del transposoma y copiando la cadena de ARN a su extremo quimérico de ADN 5'. La molécula de cadena doble se puede entonces amplificar (por ejemplo en conjunto) y secuenciar con un cebador de secuenciación que comprende parcialmente la secuencia ME y la secuencia adaptadora corriente arriba. De manera alternativa, el otro extremo de la molécula (el extremo poliT) se puede secuenciar con un cebador que hibridiza corriente arriba la secuencia de poliT y se extiende con un nucleótido dATP natural antes de comenzar los ciclos de la química SBS. La secuenciación del extremo pareado también se posibilita mediante este método.

Cuando se suministra ssADN para secuenciación se puede utilizar una aproximación similar. Por ejemplo, el extremo 3' de los poli nucleótidos de ADN de cadena única se pueden agregar a los nucleótidos al utilizar una desoxinucleotidil transferasa (TdT) terminal y cualquier dNTP tal como el dATP o el dTTP. Se podría utilizar cualquier método para agregar una cadena de nucleótidos al extremo de una molécula de ssADN. La Figura 8 es un ejemplo donde el poliA que contiene sondas de captura se inmoviliza sobre la superficie de soporte y se capturan las moléculas con una adición de cola de ssADN- poliT. Cualquier secuencia de captura, que incluye aquella del extremo ssADN, se podría utilizar en tanto que las secuencias complementarias se suministran mediante la sonda de captura sobre el soporte y los nucleótidos del ssADN de tal manera que podría ocurrir la hibridación. La extensión de la sonda de captura para crear el dsADN mediante una polimerasa de ADN para crear el dúplex de ADN: ADN, la ligadura transposicional de los adaptadores oligos y la amplificación del desplazamiento de la cadena como se describió previamente se podrían efectuar al suministrar moléculas de doble cadena para la formación del conjunto. La molécula de doble cadena se podría amplificar (por ejemplo amplificación de conjunto) y secuenciar.

Por vía de otro ejemplo (Figura 2) el ARN (total o enriquecido con poliA) se fragmenta, tratado con una fosfatasa, luego la molécula adaptadora de cadena simple se liga al extremo 3' de cada fragmento que comprende el complemento del cebador unido a la superficie P7. Los fragmentos son entonces agregados a un soporte (por ejemplo, "flowcell") se capturan por vía de hibridación. Las moléculas de aire ARN hibridadas se convierten a un dúplex de ADN: ARN con una polimerasa de transcriptasa inversa. El complejo de transposoma que comprende una transposasa y un dúplex adaptador (transposón) de una secuencia ME con una secuencia cebadora P5 se puede utilizar para tagmentar el

dúplex. Luego de la extensión de la cadena de ADN al extremo con una polimerasa que desplaza la cadena, las moléculas se pueden amplificar y secuenciar.

Por vía de un ejemplo adicional (Figura 3), se utiliza un soporte especial (por ejemplo "flowcell") que contiene dos cebadores injertados de superficie: un cebador de injerto estándar (por ejemplo P5) y un cebador de injerto modificado (por ejemplo P7) que tiene una sonda de captura específica de blanco a su lado (3') corriente abajo. Un ejemplo de una sonda específica de blanco es una secuencia oligo complementaria a una transcriptasa inversa retroviral (por ejemplo la VIH polimerasa). Se agrega el ARN viral purificado a un soporte, capturado por vía de hibridación, copiado con transcriptasa inversa y tagmentado. Se puede lograr la secuenciación con un cebador hibridado al adaptador tagmentado o en el otro extremo de la sonda de captura. Opcionalmente, el soporte especial contiene múltiples diferentes sondas de captura específicas de blanco para posibilitar la captura simultánea para muchos diferentes blancos de ARN.

El uso de una reacción de transposición *in vitro* para etiquetar los dúplex de ADN: ADN o de ADN: ARN para generar los dúplex de ADN: ADN o de ADN: ARN etiquetados involucra una transposasa, una composición de secuencia de transposón, y unas condiciones de reacción adecuadas.

Como se utiliza en todo el documento, el término transposón se refiere a un ADN de cadena doble que contiene las secuencias de nucleótido que son necesarias para formar el complejo con la transposasa o la enzima integrasa que es funcional en una reacción de transposición *in vitro*. Un transposón forma un complejo o un complejo sináptico o un complejo de transposoma. El transposón también puede formar una composición de transposoma con una transposasa o integrasa que reconoce y se une a la secuencia de transposón, y cuyo complejo es capaz de insertar o transponer el transposón en el ADN blanco con el cual este se incubaba en una reacción de transposición *in vitro*. Un transposón exhibe dos secuencias complementarias que consisten de una secuencia de transposón transferida o una cadena transferida o una secuencia de transposón no transferida, o una cadena no transferida. Por ejemplo, un transposón que forma un complejo con una transposasa de Tn5 hiperactiva (por ejemplo, la Transposasa EZ- Tn5TM, EPICENTRE Biotechnologies, Madison, WI, USA) que está activa en una reacción de transposición *in vitro* comprende la cadena transferida que exhibe una secuencia de transposón transferida de 5' AGATGTGTATAAGAGACAG 3', (SEQ ID NO: 1) y una cadena no transferida que exhibe una secuencia de transposón no transferida de 5'

CTGTCTCTTATACACATCT 3'. (SEQ ID NO: 2). El extremo 3' de una cadena transferida se une o se transfiere a el ácido nucleico blanco en una reacción de transposición *in vitro*. La cadena no transferida, que exhibe la secuencia extrema de transposón que es complementaria a la secuencia extrema de transposón transferida, no se une o se transfiere al ácido nucleico blanco en una reacción de transposición *in vitro*. La composición de transposón, como se utiliza aquí, se refiere a una composición que comprende un transposón (es decir, el segmento de ADN de doble cadena mínimo que es capaz de actuar con una transposasa para sufrir una reacción de transposición), que incluye opcionalmente secuencias adicionales. Por ejemplo, la composición de transposón comprende dos oligonucleótidos de transposón que contienen el oligonucleótido de transposón transferido con la cadena transferida y el oligonucleótido de cadena no transferida o la cadena no transferida, la cual, en combinación, exhiben las secuencias del transposón. Una o ambas cadenas pueden comprender la secuencia adicional. El transposón puede incluir oligonucleótidos de ocurrencia natural y/o de ocurrencia no natural y enlaces de estructura natural o no natural. Opcionalmente, el transposón puede también incluir una o más porciones unidas a uno o más nucleótidos que conforman el transposón. Por ejemplo, una o ambas cadenas del transposón pueden ser biotiniladas o pueden contener una etiqueta, por ejemplo una etiqueta fluorescente.

Los términos oligonucleótido de transposón transferido y cadena transferida se utilizan de manera intercambiable y se refieren a la porción transferida de tanto los transposones como las composiciones de transposón, es decir, sin importar si el extremo de transposón se une a una etiqueta o a otra secuencia o porción. De manera similar, los términos oligonucleótidos de transposón no transferido y cadena no transferida se utilizan intercambiamente y se refieren a la porción no transferida de tanto los transposones como las composiciones de transposón.

En algunas realizaciones, la composición de transposón comprende o consiste de al menos un transposón con una o más de otras secuencias de nucleótido además de las secuencias de transposón. Así, en algunas realizaciones, la composición de transposón comprende una cadena transferida con una o más de otras secuencias 5' de nucleótido de la secuencia de transposón transferida, por ejemplo, una secuencia de etiqueta. Además de la secuencia de transposón transferida, la etiqueta puede tener una o más porciones de etiqueta o dominios de etiqueta.

Como se utiliza aquí, una "etiqueta" se refiere a un componente de ácido nucleico, generalmente ADN, que suministra unos medios para identificar o dirigir un fragmento de ácido nucleico al cual éste se une. Por ejemplo, una etiqueta comprende una secuencia de nucleótido que permite la identificación, reconocimiento, y/o manipulación molecular o bioquímica del ADN al cual se une la etiqueta (por ejemplo, al suministrar un sitio para hibridar un oligonucleótido, tal como un cebador para la extensión del ADN polimerasa, al suministrar un oligonucleótido para la captura o para una reacción de ligado, o al suministrar la identificación del ácido nucleico como originario de una fuente particular, y similar). El proceso de unir la etiqueta a una molécula de ácido nucleico se denomina algunas veces aquí como "etiquetado" y los ácidos nucleicos que sufren el etiquetado o que contienen la etiqueta se denominan como "etiquetado" (por ejemplo "ARN etiquetado")

Como se utiliza en toda la especificación, el término encadenamiento o información específica de cadena se refiere a la preservación del conocimiento en la dirección de la molécula de cadena única original. Este se preserva en los métodos suministrados ya que se sabe que la cadena de ADN es complementaria a la cadena de ARN en los dúplex de ADN: ARN. Así, cuando se secuencia la cadena de ADN, la secuencia será la secuencia de la cadena de ARN que preserva la información específica de la cadena y permite la identificación correcta de la molécula de ARN y/o su nivel de expresión. Los métodos para preservar la información específica de cadena también se describen en la WO 2011/003630.

Sin embargo, el método descrito en la WO 2011/003630 requiere a una conversión de las moléculas de ARN en moléculas de cADN de cadena doble, el cual, como se describen aquí, no es tan eficiente como los métodos suministrados en la presente solicitud. Además, el método descrito en la WO 2011/003630 requiere una etiqueta con el fin de preservar la información de cadena. En los métodos suministrados aquí, no se requiere una etiqueta para preservar la información específica de cadena o el encadenamiento. En realizaciones donde la cadena de ADN (es decir, la primera cadena de ADN) de los dúplex de ADN: ARN se amplifica para producir las primeras u segundas cadenas de ADN amplificadas. Se mantiene el encadenamiento mediante el conocimiento de que la primera cadena de ADN es complementaria a la cadena de ARN original y la segunda cadena de ADN es la misma secuencia que la cadena de ARN original (con la excepción del Ts en la secuencia en lugar del Us). Así, aunque una etiqueta (por ejemplo, una secuencia de etiqueta se puede incluir en la cadena transferida del transposón) se puede utilizar para preservar el encadenamiento, esta no se requiere.

Como se utiliza aquí, una porción de etiqueta o un dominio de etiqueta significa una porción o dominio de una etiqueta que exhibe una secuencia para un propósito o aplicación pretendida deseada. Una porción de etiqueta o un dominio de etiqueta es el dominio de transposón, cuya porción de etiqueta o dominio de etiqueta exhibe la secuencia de transposón transferida. En algunas realizaciones en donde la cadena transferida también exhibe una o más secuencias de nucleótido, la etiqueta también tiene uno o más de otros dominios de etiqueta, cada uno de cuyos dominios de etiqueta se suministra para cualquier propósito deseado. Por ejemplo, una composición de transposón puede comprender (i) una cadena transferida que exhibe una o más secuencias adicionales (además de la secuencia de transposón) que puede comprender un dominio de etiqueta seleccionado de entre una o más de un dominio de etiqueta de sitio de restricción, un dominio de etiqueta de captura, un dominio de etiqueta de secuenciación, un dominio de etiqueta de amplificación, un dominio de etiqueta de detección, un dominio de etiqueta de dirección, y un dominio promotor de transcripción; y (ii) una cadena no transferida que exhibe la secuencia de transposón no transferida.

Si se utiliza una descripción para un dominio de etiqueta, los nombres y las descripciones de los diferentes dominios de etiqueta son por conveniencia, tal como para hacer más fácil el entendimiento y discutir los propósitos pretendidos y las aplicaciones de las diferentes porciones o dominios de la etiqueta en las diferentes realizaciones. Sin embargo, estos nombres y descripciones no pretenden limitar el uso o las aplicaciones de la etiqueta o de ninguno de sus dominios de etiqueta de ninguna manera. Así, cualquier etiqueta particular o dominio de etiqueta se puede utilizar para cualquier propósito además de, o en lugar del propósito o la aplicación primaria pretendida. También, un dominio de etiqueta puede comprender dos o más de otros dominios de etiqueta (por ejemplo, un dominio de etiqueta de secuenciación puede comprender tanto el dominio de etiqueta de captura como un dominio de etiqueta de amplificación) o un dominio de etiqueta puede suministrar las funciones o los propósitos o aplicaciones de dos o más diferentes dominios de etiqueta (por ejemplo, un dominio de etiqueta de captura también puede suministrar la función o propósito de secuenciar el dominio de etiqueta y/o un dominio de etiqueta de amplificación para una aplicación particular). Aun adicionalmente, la etiqueta no requiere ser descrita en términos de uno o más diferentes dominios con el fin de ser utilizada para cualquier propósito o aplicación o función particular.

Como se utiliza en toda la especificación, el término transposasa se refiere a una enzima que es capaz de formar un complejo funcional con una composición que contiene transposón (por ejemplo transposones, composiciones de transposón) y la inserción catalizadora o la transposición de la composición que contiene transposón en el ácido nucleico blanco de doble cadena con el cual se incubaba en una reacción de transposición *in vitro*. Una transposasa de los métodos suministrados también incluye integrasas provenientes de retrotransposones y retrovirus. Transposasas de ejemplo que se pueden utilizar en los métodos suministrados incluyen el tipo silvestre o las formas mutantes de la transposasa Tn5 y la transposasa MuA.

Una "reacción de transposición" es una reacción en donde uno o más de los transposones se insertan en los ácidos nucleicos blanco en sitios aleatorios o casi en sitios aleatorios. Los componentes esenciales en la reacción de transposición son una transposasa y los oligonucleótidos de ADN que exhiben las secuencias de nucleótido de un transposón, incluyendo la secuencia de transposón transferida y su complemento (es decir, el transposón y la secuencia no transferida) así como también otros componentes necesarios para formar una transposición funcional o un complejo de transposoma. El método de esta invención se ejemplifica al emplear un complejo de transposición formado por una transposasa Tn5 hiperactiva y un extremo del transposón tipo Tn5 o por medio de una transposasa MuA o HYPERMu y un extremo de transposón Mu que comprende las secuencias de extremo R1 y R2 (See e.g., Goryshin, I. and Reznikoff, W. S., J. Biol. Chem., 273: 7367, 1998; and Mizuuchi, K., Cell, 35: 785, 1983; Savilahti, H, et al., EMBO J., 14: 4893, 1995; las cuales se incorporan aquí como referencia en su totalidad). Sin embargo, cualquier Sistema de transposición que sea capaz de insertar un extremo del transposón de una manera aleatoria o casi aleatoria con suficiente eficiencia para etiquetar los ácidos nucleicos blancos para su propósito pretendido se pueden utilizar en los métodos suministrados. Otros ejemplos de los sistemas de transposición conocidos que se podrían utilizar en los métodos

suministrados incluyen pero no están limitados a *Staphylococcus aureus* Tn552, Tyl, el Transposon Tn7, Tn/O y IS10, la transposasa Maruner, Tcl, el Elemento P, Tn3, las secuencias de inserción bacteriana, los retrovirus, y el retrotransposón de levadura (See, e.g., Colegio O R et al., J. Bacteriol., 183: 2384-8, 2001; Kirby C et al., Mol.

5 Microbiol., 43: 173-86, 2002; Devine S E, and Boeke J D., Nucleic Acids Res., 22: 3765-72, 1994; International Patent Application No. WO 95/23875; Craig, N L, Science. 27 1:1512, 1996; Craig, N L, Review in: Curr Top Microbiol Immunol., 204: 27-48, 1996;

Kleckner N, et al., Curr Top Microbiol Immunol., 204: 49-82, 1996; Lampe D J, et al., EMBO J., 15: 5470-9, 1996; Plasterk R H, Curr Top Microbiol Immunol, 204: 125-43,

10 1996; Gloor, G B, Methods Mol. Biol., 260: 97-1 14, 2004; Ichikawa H, and Ohtsubo E., J Biol. Chem. 265: 18829-32, 1990; Ohtsubo, F and Sekine, Y, Curr. Top. Microbiol. Immunol. 204: 1-26, 1996; Brown P O, et al., Proc Natl Acad Sci USA, 86: 2525-9,

1989; Boeke J D and Corces V G, Annu Rev Microbiol. 43: 403-34, 1989;

15 El método para insertar un transposón en la secuencia blanco se puede llevar a cabo *in vitro* utilizando cualquier sistema de transposón adecuado para lo cual está disponible un sistema de transposición *in vitro* adecuado o se puede desarrollar con base en el conocimiento de la técnica. En general, un sistema de transposición *in vitro* adecuado para uso en los métodos de la presente invención requiere como mínimo una enzima transposasa de pureza suficiente, concentración suficiente, actividad de transposición *in vitro* suficiente un transposón con el cual la transposasa forme un complejo funcional con la respectiva transposasa que sea capaz de catalizar la reacción de transposición. La secuencia del transposón de transposasa adecuadas que se pueden utilizar en la invención incluye pero no están limitadas a la tipo silvestre, las secuencias de transposón derivadas o mutantes que forman un complejo con la transposasa seleccionada de entre el tipo silvestre, la forma derivada mutante de la transposasa.

20 En los métodos suministrados, los dúplex de ADN: ARN se pueden suministrar en una variedad de maneras. Por vía de ejemplo, el soporte puede comprender una pluralidad de cebadores y los dúplex de ADN: ARN se suministran al hibridar una o más moléculas de ARN a los cebadores inmovilizados sobre el soporte y extendiendo los cebadores hibridados a las moléculas de ARN utilizando las moléculas de ARN como plantillas para producir los uno o más dúplex de ADN: ARN. Opcionalmente, se suministra una pluralidad de dúplex de ADN: ARN al hibridar una pluralidad de moléculas de ARN a los cebadores inmovilizados sobre el soporte y extendiendo los cebadores hibridados a las moléculas de ARN utilizando las moléculas de ARN como plantilla para producir la pluralidad de dúplex de ADN: ARN.

25 Como se estableció anteriormente, los métodos pueden comprender suministrar un soporte con una pluralidad de cebadores; los cebadores o un subconjunto de los mismos comprenden una secuencia capaz de unir a una o más moléculas de ARN. Por ejemplo, los cebadores inmovilizados pueden incluir una secuencia de poliT y el ARN puede incluir una secuencia de poliA capaz de hibridar la secuencia de poliT. Alternativa o adicionalmente, la pluralidad de cebadores inmovilizados pueden incluir cebadores específicos blanco capaces de hibridar a una o más de las moléculas de ARN en la pluralidad de moléculas de ARN. Así, la cadena de ARN de los uno o más dúplex de ADN: ARN comprende una secuencia complementaria a al menos una porción de uno o más de los cebadores inmovilizados. Opcionalmente, la pluralidad de cebadores inmovilizados comprende un primer subconjunto de cebadores de una primera secuencia y un segundo subconjunto de cebadores de una segunda secuencia. El primer o segundo subconjunto de cebadores puede comprender la secuencia poliT.

30 Opcionalmente, se puede agregar un adaptador 3' a la pluralidad de moléculas de ARN, el adaptador 3' comprende una secuencia complementaria a la pluralidad de cebadores inmovilizados o un subconjunto de los mismos. Tales moléculas de ARN ligadas al adaptador 3' se pueden entonces hibridar a los cebadores inmovilizados.

35 Así, los cebaderos inmovilizados o un subconjunto de los mismos pueden comprender una secuencia poliT, una secuencia específica blanco de ARN o una secuencia complementaria a un adaptador ligado a la molécula de ARN. Opcionalmente, la pluralidad de cebadores comprende al menos dos subconjuntos de cebadores, el primer subconjunto comprende una secuencia poliT, una secuencia específica blanco de ARN o una secuencia complementaria a un adaptador ligado a la molécula de ARN, y el segundo subconjunto de cebadores comprende una secuencia que es capaz de unirse a una secuencia sobre la cadena de ADN de los dúplex de ADN: ARN. Tal secuencia puede ser, por ejemplo, la misma secuencia que la secuencia de la cadena transferida del transposón. Como se describió en toda la especificación, después de la transposición, habrá un bache entre el extremo de la cadena de ADN y la cadena no transferida del transposón. La cadena de ADN se puede extender para copiar la cadena de ARN. La copia incluirá copiar las secuencias de la cadena transferida del transposón. La cadena de ADN incluirá entonces las secuencias complementarias a las secuencias de la cadena transferida del transposón y, así, los cebadores o el subconjunto de los mismos sobre la superficie de soporte. En otras palabras, si uno o más de los cebadores comprende una secuencia igual o similar a la cadena transferida del transposón, la cadena de ADN en los dúplex de ADN: ARN será capaz de
40
45
50
55 hibridar a los cebadores ya que la cadena de ADN contiene una secuencia complementaria al cebador.

Se conocen las enzimas que modifican el ácido nucleico adecuado capaces de extender el extremo 3' de las cadenas de ADN para copiar las cadenas de ARN a su extremo 5' y desplazar la cadena no transferida del transposón. En resumen, algunas ADN polimerasas son capaces de desplazar la cadena complementaria a la cadena de plantilla como

una nueva cadena de ADN y sintetizada mediante la polimerasa. Este proceso se denomina desplazamiento de cadena y las ADN polimerasa que tienen esta actividad se denominan aquí como polimerasas de ADN que desplazan la cadena. En general, la polimerasa de ADN específica de la plantilla de ADN utilizada para los métodos suministrados sintetiza eficientemente el ADN de una longitud adecuada para el propósito pretendido sin desacoplarse de la plantilla (o terminar la síntesis del ADN), lo cual se denomina como la procesividad de la enzima. La capacidad de una ADN polimerasa para desplazar la cadena se puede determinar fácilmente utilizando la polimerasa en un ensayo de replicación de círculo rodante como se describió por Fire and Xu (*Proc. Natl. Acad. Sci. USA* 92: 4641-4645, 1995).

El desplazamiento de la cadena y la procesividad de la ADN polimerasa se puede ensayar utilizando los métodos descritos en Kong et al. (*J. Biol. Chem.* 268: 1965-1975, 1993).

La transferasa terminal también se define como una ADN polimerasa aquí, cuya ADN polimerasa se utiliza como una composición en algunas realizaciones de los métodos suministrados. La transferasa terminal se puede utilizar porque esta cataliza la adición independiente de la plantilla de los dNTP a los 3'-hidroxil terminales de ADN.

En los métodos suministrados aquí, el método puede además comprender secuenciar al menos una porción de las cadenas de ADN y/o amplificar al menos una porción de las cadenas de ADN. Opcionalmente, las cadenas de ARN provenientes de los dúplex de ADN: ARN se pueden remover antes de secuenciación y/o amplificación. A manera de ejemplo, el método además comprende retirar las cadenas de ARN de los dúplex de ADN: ARN y secuenciar al menos una porción de las cadenas de ADN (es decir las primeras cadenas de ADN). El método también puede incluir copiar al menos una porción de las cadenas de ADN para producir una segunda cadena de ADN complementaria a la cadena de ADN (es decir, la primera cadena de ADN) de los dúplex de ADN: ARN. La segunda cadena de ADN complementaria se puede secuenciar, si se desea. Opcionalmente, la primera cadena de ADN de los dúplex de ADN: ARN se puede retirar antes de secuenciar la segunda cadena de ADN complementaria.

En los métodos suministrados, opcionalmente, después de retirar la cadena de ARN de los dúplex de ADN: ARN, las cadenas de ADN se pueden amplificar para producir una pluralidad de moléculas de ADN de cadena doble que comprenden primeras y segundas cadenas amplificadas. Opcionalmente, la amplificación produce un conjunto, descrito con más detalle adelante.

En algunas realizaciones, cuando las cadenas de ADN han sido amplificadas para producir una pluralidad de moléculas de ADN de cadena doble, una o ambas cadenas se pueden secuenciar. Por vía de ejemplo, los métodos pueden incluir retirar las primeras cadenas amplificadas seguidas al secuenciar al menos una porción de las segundas cadenas amplificadas. Opcionalmente, las primeras cadenas amplificadas se pueden generar al copiar al menos una porción de las segundas cadenas amplificadas. Las segundas cadenas amplificadas se pueden entonces retirar con el fin de secuenciar al menos una porción de las primeras cadenas amplificadas. Opcionalmente, las lecturas de secuencia de una porción o todo de uno o ambos de las primeras y segundas cadenas amplificadas se puede efectuar sin retirar todo o una porción de cualquier cadena.

Se pueden utilizar varios protocolos para generar ácidos nucleicos amplificadas, por ejemplo, ácidos nucleicos amplificadas sobre un soporte. Por ejemplo, se pueden amplificar ácidos nucleicos mediante PCR de emulsión, o PCR de puente (Mitra & Church *Nucleic Acids Res.* 27, e34 (1999); Dressman et al. *Proc. Natl. Acad. Sci USA* 100, 88 17-8822 (2003); Adessi, C. et al. *Nucleic Acids Res.* 28, e87 (2000); Fedurco et al. *Nucleic Acids Res.* 34, e22 (2006).

En las realizaciones que utilizan PCR de emulsión, los ácidos nucleicos pueden ser amplificadas mediante PCR en una emulsión de agua en aceite. En una realización, se utiliza un par cebador simple. Uno de los cebadores de PCR está atado a la superficie (unido a 5') de un soporte (por ejemplo glóbulos a micro escala) y el otro cebador está en solución. Opcionalmente, el soporte comprende cebadores de más de una secuencia, los cebadores son cebadores específicos blanco capaces de hibridar a una o más moléculas de ARN blanco y el cebador en solución es la misma secuencia (por ejemplo una secuencia complementaria a la secuencia agregada a la cadena de ADN al copiar la cadena de ARN etiquetada a su extremo 5'). En general, una concentración de plantilla baja da como resultado en la mayoría de los compartimentos que contienen glóbulos que tienen cero o una molécula de plantilla presente. En los compartimentos de emulsión productiva (donde tanto el glóbulo como la molécula de plantilla está presente), las moléculas de ARN se pueden capturar y/o el correspondiente complemento del ADN de la molécula de ARN amplificada a la superficie del glóbulo. Después de descomponer la emulsión, los productos de amplificación que llevan glóbulos se pueden enriquecer selectivamente. Cada glóbulo amplificado clonalmente llevará sobre su superficie los productos de PCR que corresponden a la amplificación de una molécula única proveniente de una biblioteca de plantilla. Varias realizaciones de los métodos PSR de emulsión que son útiles se establecen en las Publicaciones de las Solicitudes de Patente U.S. 2005/0042648 A1; 2005/0079510 A 1 y 2005/0130173 A 1, y WO 05/010145.

En las realizaciones que utilizan PCR puente, también conocidas como formación de conjunto, los ácidos nucleicos provenientes de la biblioteca de plantillas se pueden amplificar utilizando los cebadores recubiertos sobre la superficie de un soporte. Los cebadores se pueden unir a sus extremos 5' mediante un ligador flexible. Los productos de amplificación que se originan de cualquier miembro dado de la biblioteca de plantilla permanecen localmente atados cerca al punto de origen. A la conclusión del PCR, cada conjunto clonal contiene varias copias de un miembro único de la biblioteca de plantilla. Los métodos suministrados, cada dúplex de ADN: ARN forma el origen de un conjunto clonal. Luego de la remoción de la cadena de ARN la cadena de ADN se puede copiar utilizando los cebadores unidos al soporte para generar copias amplificadas de la cadena de ADN para producir los conjuntos clonales. Varias

realizaciones de los métodos PCR puente que son útiles se establecen en las Publicaciones de las Solicitudes de Patente U.S. No. 2007/0128624 A1, WO 07/010251, Patente U.S. No. 6, 090,592 y la Patente U.S. No.5, 641,658. Los métodos para llevar a cabo la amplificación se describen en la Publicación U.S. No. 2009/0226975; WO 98/44151; WO 00/18957; WO 02/46456; WO 06/064 199; y WO 07/0 10251.

5 Los métodos establecidos aquí pueden hacer o utilizar arreglos que tengan características de una cualquiera de una variedad de densidades que incluyen, por ejemplo, al menos aproximadamente 10 características /cm², 100 características/cm², 500 características /cm², 1.000 características/cm², 5.000 características/cm², 10.000 características/cm², 50.000características/cm², 100.000 características/cm², 1.000.000 características/cm², 5.000.000 características/cm², o más.

10 Como se utiliza aquí, el término “ácido nucleico” se puede utilizar para referirse a al menos dos monómeros de análogo de nucleótido ligados. Un ácido nucleico puede contener enlaces de fosfodiéster, sin embargo, en algunas realizaciones, el ácido nucleico puede ser un análogo que tenga otros tipos de estructuras, que comprenden, por ejemplo, fosforamida, fosforotioato, fosforoditioato, estructuras de péptido de ácido nucleico y enlaces, estructuras positivas, o estructuras no iónicas. Un ácido nucleico puede incluir una porción de pentosa tal como ribosa (actualmente en ARN de ocurrencia natural), desoxirribosa (presente en ADN de ocurrencia natural) o didesoxirribosa . En algunas realizaciones un ácido nucleico puede tener una porción no pentosa o un azúcar carbocíclico en lugar de una ribosa o una porción desoxirribosa. Un ácido nucleico puede tener una o más diferentes porciones base que incluyen, pero no están limitadas a, adenina, (A), guanina (G), timina (T), uracilo (U), citosina (C), inosina, xantantina, hipoxantantina, isocitosina, isoguanina, nitropirrol (que incluye 3- nitropirrol) y/o nitroindol (que incluye 5-nitroindol). Un ácido nucleico utilizado aquí puede incluir bases nativas o no nativas. Así, un ácido nucleico puede incluir nucleótidos de ocurrencia natural y/o de ocurrencia no natural y enlaces de estructuras naturales y no naturales. Los ácidos nucleicos pueden ser de cadena simple o de cadena doble, tal como se especificó, o contener porciones de secuencias tanto de cadena doble como de cadena sencilla. El ácido nucleico puede ser ADN (por ejemplo ADN genómico o cADN), ARN o un híbrido.

25 Como se utiliza aquí, el término “arreglo” significa una población de diferentes moléculas que están unidas a uno o más soportes de tal manera que se pueden diferenciar diferentes moléculas una de la otra de acuerdo a su ubicación relativa. Un arreglo puede incluir diferentes moléculas que se ubican cada una en diferentes ubicaciones dirigibles (por ejemplo una característica) sobre un soporte. De manera alternativa, un arreglo puede incluir soportes separados que lleva cada uno una molécula diferente, en donde las moléculas de sonda diferentes se pueden identificar de acuerdo a las ubicaciones de los soportes sobre una superficie a la cual los soportes se unen o de acuerdo a las ubicaciones de los soportes en líquido tal como una corriente de fluido. Las moléculas del arreglo pueden ser, por ejemplo, cebadores de ácido nucleico, sondas de ácido nucleico, plantillas de ácido nucleico o enzimas de ácido nucleico tal como polimerasas. Por ejemplo, en realizaciones particulares los ácidos nucleicos blanco se pueden unir a una superficie de un detector o a una capa (por ejemplo una capa de acrilamida) que está presente en la superficie del soporte. Los hidrogeles son particularmente útiles tal como aquellos establecidos en la Publicación de Patente US No. 2011/0059865 A 1.

35 Como se utiliza aquí el término “arreglo de ácido nucleico” significa un soporte sólido que tenga una pluralidad de ácidos nucleicos espacialmente distinguibles dispuestos sobre éste o en éste. Los ácidos nucleicos se pueden disponer en un patrón de características ordenado o aleatorio. Una característica individual puede ser, por ejemplo, una molécula de ácido nucleico espacialmente aislada, o un ensamble de moléculas de ácido nucleico tal como un conjunto. Un arreglo puede ser un arreglo compuesto que comprenda una pluralidad de arreglos individuales configurados para permitir el procesamiento de múltiples muestras. Los arreglos individuales denominados aquí como “subarreglos”, incluyen grupos de características de ácido nucleico. Los subarreglos aparecen en distintas regiones con un arreglo mayor. Los subarreglos mismos se pueden ordenar o no ordenar. Tales subarreglos pueden ser dirigibles espacialmente de manera opcional. Los subarreglos pueden incluir conjuntos de ácidos nucleicos idénticos. Un ejemplo de un arreglo compuesto en subarreglos individuales es una placa de microtítulo que tiene pozos en cada placa como un todo es un arreglo de ácidos nucleicos (o un arreglo de compuesto) en el cual cada pozo individual representa un subarreglo dentro del arreglo de compuesto mayor.

40 Como se utiliza aquí el término “soporte” se refiere a un sustrato para inmovilizar un arreglo de ácidos nucleicos. Un “soporte” es un material que tiene una superficie rígida o semirrígida a la cual el arreglo de ácido nucleico se puede unir o a la cual los ácidos nucleicos se pueden sintetizar y/o modificar. Los soportes pueden incluir cualquier resina, microglóbulo, vidrio, vidrio poroso controlado (CPG), soporte de polímero, membrana, papel, plástico, tubo o tableta plástica, glóbulo plástico, glóbulo de vidrio, lámina, cerámica, chip de silicio, placa multipozo, membrana de nylon, fibra óptica, y membrana de PVDF.

55 Un soporte puede incluir cualquier sustrato similar a galleta plana y los sustratos planos que tengan pozos, tales como las placas de microtítulo, que incluyen placas de 96 pozos. Los sustratos planos de ejemplo incluyen chips, láminas, sustratos grabados, placas de microtítulo, y reactores de celda de flujo, que incluyen reactores de celda de flujo multicarril que tienen múltiples canales de microfluído, tales como la celda de flujo de 8 canales utilizada en la estación de trabajo de secuenciación cBot (Illumina, Inc., San Diego, CA). Las celdas de flujo de ejemplo que se pueden utilizar también se describen en la WO 2007/123744.

Un soporte también puede incluir glóbulos, que incluyen glóbulos magnéticos, glóbulos huecos, y glóbulos sólidos. Los glóbulos se pueden utilizar en conjunto con soportes planos, tales soportes planos opcionalmente también contienen pozos. Los glóbulos o alternativamente las microesferas, se refieren de manera general a un cuerpo pequeño hecho de un material rígido o semirrígido. El cuerpo puede tener una forma caracterizada, por ejemplo, una esfera, oval, microesfera, u otra forma de partícula reconocida sea que tenga dimensiones regulares o irregulares. Los tamaños de los glóbulos, en particular, incluyen, sin limitación, aproximadamente un 1 μm , aproximadamente 2 μm , aproximadamente 3 μm , aproximadamente 5 μm , aproximadamente 10 μm , aproximadamente 20 μm , aproximadamente 30 μm , aproximadamente 40 μm , aproximadamente 60 μm , aproximadamente 100 μm , aproximadamente 150 μm , o aproximadamente 200 μm de diámetro. Se pueden utilizar otras partículas de manera similares a aquellas descritas aquí para los glóbulos y las microesferas.

La composición de un soporte puede variar, dependiendo por ejemplo, del formato, la química y/o el método de unión y/o el método de síntesis del ácido nucleico. Los materiales de soporte que se pueden utilizar de acuerdo con la presente descripción incluyen, pero no están limitados a, polipropileno, polietileno, polibutileno, poliuretanos, nylon, metales, y otros materiales adecuados. Las composiciones de ejemplo incluyen soportes, y funcionalidades químicas impartidas a estos, utilizadas en el polipéptido, polinucleótido y/o las síntesis de porciones orgánicas. Tales composiciones incluyen, por ejemplo, plásticos, cerámicas, vidrio, poliestireno, melamina, metilostireno, polímeros acrílicos, materiales paramagnéticos, sol toria, grafito de carbono, dióxido de titanio, látex, o dextranos reticulados tales como cefaroseTM, celulosa, nylon, micelas reticuladas y teflónTM, así como también otros materiales que se pueden encontrar descritos en, por ejemplo, "Microsphere Detection Guide" de Bangs Laboratories, Fishers IN.

Una partícula de soporte se puede hacer de un almidón reticulado, dextranos, celulosa, proteínas, polímeros orgánicos que incluyen polímeros de estireno que incluyen poliestireno y metilostireno así como también otros copolímeros de estireno, plásticos, vidrio, cerámicas, polímeros acrílicos, materiales de respuesta magnética, coloides, toreasol, grafito de carbono, dióxido de titanio, nylon, látex o TEFLON® "Microsphere Detection Guide" de Bangs Laboratories, Fishers IN. Es una guía de ayuda. Soportes de ejemplo adicionales dentro del alcance de la presente descripción incluyen, por ejemplo, aquellos descritos en la Publicación de la Solicitud US No.2012/0102578 y de la Patente U.S. No. 6, 429, 027.

La unión de un ácido nucleico a un soporte, sea rígida o semirrígida, puede ocurrir por vía covalente o de enlaces no covalentes. Los enlaces de ejemplo se establecen en la patente U.S. Nos. 6, 737, 236; 7,259, 258; 7, 375, 234 y 7, 427, 678; y la Publicación de Patente US No. 2011/0059865 A1. En algunas realizaciones, un ácido nucleico u otro componente de reacción se puede unir a un gel u otro soporte semisólido que a su vez esté unido o adherido a un soporte de fase sólida. En tales realizaciones, el ácido nucleico u otro componente de la reacción se entenderá como la fase sólida.

Opcionalmente, el soporte es un glóbulo o una pluralidad de glóbulos. Opcionalmente, el soporte es un soporte plano. Opcionalmente, se suministran una pluralidad de glóbulos, cada glóbulo comprende uno o más dúplex de ADN: ARN. Si un glóbulo comprende más de un dúplex de ADN: ARN, los dúplex pueden ser de la misma secuencia o de diferente secuencia. Opcionalmente, se suministra una pluralidad de glóbulos cada glóbulo comprendiendo un dúplex de ADN: ARN. Los glóbulos en la pluralidad de glóbulo pueden comprender el mismo o diferentes dúplex de ADN: ARN. Por ejemplo, un primer subconjunto de glóbulos en la pluralidad de glóbulos puede comprender un dúplex de ADN: ARN de una primera secuencia mientras que el segundo subconjunto de glóbulos en la pluralidad de glóbulos puede comprender un dúplex de ADN: ARN de una segunda secuencia.

Cualquiera de una variedad de protocolos de secuenciación y los respectivos reactivos se pueden utilizar en cualquier método o dispositivo establecido aquí. Las técnicas mediante síntesis de secuenciación (SBS) generalmente involucra la extensión enzimática de una cadena de ácido nucleico naciente a través de la adición iterativa de nucleótidos contra la cadena de la plantilla. El SBS puede utilizar monómero de nucleótido que tengan una porción terminadora o aquellos que les falta las porciones terminadoras.

Los métodos que utilizan monómeros que tienen terminadores incluyen, por ejemplo, aquellos descritos en la WO 04/018497, US 7,057,026, WO 91/106678, WO 07/ 123744, US 2007/0166705, US 2006/0188901, US 2006/0240439, US 2006/0281109, WO 05/065814, US 2005/0100900, WO 06/064199 o WO 070 1025 1.

También son útiles los métodos SBS que están comercialmente disponibles de Illumina, Inc., San Diego CA.

Las técnicas SBS pueden utilizar monómeros de nucleótido que tengan una porción de rótulo. De acuerdo con esto, los eventos de incorporación se pueden detectar con base en las características del rótulo, tal como la fluorescencia del rótulo; una característica del monómero del nucleótido tal como el peso molecular o la carga; un subproducto de incorporación del nucleótido, tal como la liberación del pirofosfato o los protones; o similares. Los diferentes nucleótidos pueden ser distinguibles uno del otro, o alternativamente, los dos o más diferentes rótulos pueden ser indistinguibles bajo las técnicas de detección que sean utilizadas. Por ejemplo, los diferentes nucleótidos presentes en un reactivo de secuenciación pueden tener diferentes rótulos y ellos se pueden distinguir utilizando ópticas apropiadas como las ejemplificadas por medio de los métodos de secuenciación desarrollados por Solexa (ahora Illumina, Inc.). Sin embargo, también es posible utilizar el mismo rótulo para los dos o más diferentes nucleótidos presentes en un reactivo de secuenciación o para utilizar las ópticas de detección que no distinguen necesariamente los rótulos diferentes.

Los métodos que utilizan los monómeros de nucleótidos a los que le faltan los terminadores son también útiles incluyendo, por ejemplo, pirosecuenciación. La pirosecuenciación detecta la liberación del pirofosfato inorgánico (PPi) en la medida en que nucleótidos particulares se incorporan en la cadena nascente. (Ronaghi, M., Karamohamed, S., Pettersson, B., Uhlen, M. and Nyren, P. (1996) "Real-time DNA sequencing using detection of pyrophosphate release." Analytical Biochemistry 242 (1), 84-9; Ronaghi, M. (2001) "Pyrosequencing sheds light on DNA sequencing." Genome Res. 11 (1), 3-11; Ronaghi, M., Uhlen, M. and Nyren, P. (1998) "A sequencing method based on real-time pyrophosphate." Science 281(5375), 363; Patente US No. 6, 210, 891; Patente US No. 6, 258,568 y Patente US No. 6, 274,320.

En la pirosecuenciación, el PPi liberado se puede detectar al ser convertido a adenosina trifosfato (ATP) mediante ATP sulfúrilasa, y el nivel del ATP generado se detecta por vía de los fotones producidos por luciferasa.

Se pueden utilizar algunas realizaciones secuenciando mediante técnicas de ligado. Tales técnicas utilizan ADN ligasa para incorporar los oligonucleótidos. Los sistemas SBS de ejemplo y los métodos que se pueden utilizar con los métodos y sistemas descritos aquí se describen en la Patente U.S. No. 6, 969, 488, la Patente US No. 6, 172, 218, y la Patente U.S. No. 6, 306, 597.

Algunas realizaciones pueden utilizar métodos que involucran la vigilancia en tiempo real de la actividad de la ADN polimerasa. Las incorporaciones del nucleótido se pueden detectar a través de interacciones de transferencia de energía con resonancia de fluorescencia (FRET) entre la polimerasa que lleva el fluoróforo y los nucleótidos rotulados con γ -fosfato como se describió, por ejemplo, en la Patente U.S. No. 7, 329, 492 y la Patente U.S. No. 7, 211, 414, o las incorporaciones de nucleótido se pueden detectar con guías onda de modo cero como se describió, por ejemplo, en la Patente U.S. No. 7, 315, 019; y utilizando análogos de nucleótido fluorescente y polimerasas desarrolladas con ingeniería como se describió, por ejemplo, en la Patente U.S. No. 7, 405, 281 y la Publicación de la Solicitud de Patente No. 2008/0108082.

La iluminación se puede restringir a un volumen de escala de zeplolitro alrededor de polimerasas atadas a la superficie de tal manera que la incorporación de los nucleótidos rotulados fluorescentemente se pueden observar con bajo trasfondo (Levene, M. 1. et al. "Zero-mode waveguides for single-molecule analysis at high concentrations." Science 299, 682-686 (2003);

Lundquist, P.M. et al. "Parallel confocal detection of single molecules in real time." Opt. Lett. 33, 1026-1028 (2008); Korlach, J. et al. "Selective aluminum passivation for targeted immobilization of single DNA polymerase molecules in zero-mode waveguide nanostructures." Proc. Nat'l. Acad. Sci. USA 105, 1176-1181 (2008).

Se describieron materiales, composiciones y componentes que pueden ser utilizados para, se puede utilizar en conjunto, o se pueden utilizar en la preparación para, o son productos de los métodos y composiciones descritos. Estos y otros materiales se describen aquí, y se entiende que cuando combinaciones, subconjuntos, interacciones, grupos, etc. de éstos y otros materiales se describen aquí, y se entiende que cuando se describen las combinaciones, subconjuntos, interacciones, grupos, etc. de estos materiales que aunque no se describa de manera explícita una referencia específica de cada una de las varias combinaciones individuales y colectivas y permutaciones, cada una se contempla específicamente y se describe aquí. Por ejemplo, si un método se describe y se discute y se puede hacer un número de modificaciones a las etapas del método que se discuten, cada una de las combinaciones y permutaciones de las etapas del método y las modificaciones que son posibles se contemplan específicamente a menos que específicamente se indique lo contrario. De manera similar, cualquier subconjunto o combinación de estos es también contemplada y descrita específicamente. Este concepto aplica a todos los aspectos de esta descripción. Así, si existe una variedad de etapas adicionales que se puedan efectuar se entiende que cada una de estas etapas adicionales se puede efectuar con alguna de las etapas del método específicas o combinaciones de las etapas de los métodos descritos, y que cada una de tales combinaciones o subconjunto de combinaciones se contempla específicamente y se debe considerar descrita.

Se han descrito un número de realizaciones. Sin embargo, se entenderá que se pueden hacer varias modificaciones.

EJEMPLOS

Ejemplo 1. Secuenciación de ARN que emplea reacción de tagmentación de un dúplex de ADN: ARN.

Se efectúa un experimento de ejemplo con un transcripto de ARN adaptado a P7', cuyo esquema se esboza en la Figura 4. Los transcriptos de ADN se generaron de un plásmido que contiene Proteína Fluorescente Verde (GFP) que utiliza el kit del sistema de transcripción *In vitro* Riboprobe® de Promega (Madison, WI) siguiendo el protocolo del fabricante.

La secuencia del cartucho de expresión GFP se muestra adelante (SEQ ID NO: 3).
aatgatacggcgaccaccgagatctacactcttccctacacgacgctctccgatctfaatagactcactataggaatttaa
 ctttactaaggagaattcaccatgaaacatcaccatcaccacacGACTACAAAGACGATGACGACAAGg
 cgatcgtgagcaagggcgaggagctgtcaccgggtggtgccatcctggcagctggagcggcgacgtaaacggccaca
 agttcagcgtgtccggcgggcgaggcgatgccacctacggcaagctgacctgaagttcatctgcaccaccggcaagctg
 cccgtgccccggccccctcgtgaccaccctgacctacggcgtgagctgctcagccgctaccggaccacatgaagcagcac
 gacttctcaagtccgcatgcccgaaggctacgtccaggagcgcaccatcttcaaggacgacggcaactacaagaccgg
 cgccgaggtgaagttcagggcggaaccctggtgaaccgcatcgagctgaaggcctcagctcaaggagggacggcaacatc
 ctggggcacaagctggagtacaactacaacagccacaacgtctatatcatggccgacaagcagaagaacgcatcaaggtgaa
 cttcaagatccgccacaacatcaggagcggcagcgtgcagctgccgaccactaccagcagaacccccatcggcgacgg
 ccccgctgctgcccgacaaccactacctgagccccagtcgcccctgagcaagaccccacgagaagcgcgatcacatgg
 tctgctggagttcgtgaccgccgggatcactctcggcatggagcagctgtacaagtaactgctgccaccgctgagaataa
ctagcataacccttggggcctctaaacgggtcttgggggtttttgagatcggagagcgggttcagcaggaatgccgagac
 cgatcTCGTATGCCGCTTCTGCTTG (SEQ ID NO:3)

5 El plásmido basado en pMA-T contiene P5 (subrayado), un promotor de polimerasa T7 (resaltado), un codón de inicio (resaltado y en itálica), y una etiqueta His (en itálica y subrayada), una etiqueta FLAG (en letras mayúsculas) y una secuencia GFP, un codón de parada TAA (subrayado y resaltado), un terminador T7 (resaltado, en itálica y subrayado), y P7' (en letras mayúsculas y resaltado).

10 El transcrito de ARN se debe extender desde la secuencia promotora a la secuencia de terminación T7. Sin embargo, el terminador T7 no detiene la transcripción completamente y así algunos de los transcritos de ADN resultantes son His_FLAG_GFP_P7'. El transcrito de ARN se trató con DNasa para remover el ADN que de otra manera formaría conjuntos. Con el fin de revisar que el tratamiento de DNasa fuera efectivo, se efectuó una reacción y se analizó sobre un gel para sondear que el tratamiento de DNasa fuera efectivo al retirar el ADN. No fue visible ADN residual (es decir, plásmido) luego del tratamiento con DNasa del transcrito de ARN. (Figura 5).

15 Una biblioteca de ADN PhiX y los transcritos de ARN GFP- P7' tratada con DNasa se hibridaron sobre diferentes carriles de un "flowcell" siguiendo el protocolo de conjunto estándar para la hibridación de plantilla. Los carriles 1- 4 contenían el ADN PhiX y los carriles 5 – 8 contenían el ARN GFP. Los carriles 5 y 6 contenían ARN que fue pretratado con DNasa para retirar el ADN. Los carriles 7 y 8 contenían ARN que fue pretratado con DNasa y tratado con RNasa sobre el "flowcell" como un control adicional. La biblioteca de ADN PhiX se puede hibridar por vía de P5 o P7 en la medida en que ambas secuencias y sus complementos estén presentes en la plantilla. En contraste, las plantillas de ARN GFP-P7' hibridan sobre los cebadores de superficie P7 solo en razón de su "encadenamiento" y la falta de una secuencia P5.

20 La primera extensión se llevó a cabo utilizando transcriptasa Inversa del Virus de Mieloblastosis Avian (AMV - RT) (carriles 2, 4, 6, Y 8) o ADN polimerasa de fusión (carriles 1, 3, 5 y 7). El AMV- RT puede generar una cadena de cADN de cualquier plantilla de ARN o ADN, mientras que el "Phusion" solo puede generar una cadena de ADN de una plantilla de ADN.

25 Algunos carriles fueron transpuestos utilizando un complejo de transposoma que contiene la secuencia adaptadora P5 de la secuencia del transposón (carriles 3 - 8). Los baches dejados en la secuencia de ADN después del evento de transposición fueron llenados al utilizar una reacción de extensión de desplazamiento de cadena que contiene la ADN polimerasa Bst. El evento de transposición se requiere en los carriles que contienen ARN GFP_P7' para agregar el adaptador P5 para generar una plantilla que pueda hacer conjuntos. La amplificación del conjunto isotérmico se llevó a cabo de manera estándar y los conjuntos se tiñeron con Verde SYBR. Las fotografías de los conjuntos se muestran en la Figura 6.

30 El carril 1 se controló para generación de conjuntos en la medida en que éste contiene una muestra de ADN con formato estándar extendida con ADN polimerasa PHUSION. La generación exitosa de conjuntos resultó como se muestra en la Figura 6.

35 El carril 2 demostró que las plantillas de ADN podían ser extendidas de manera exitosa mediante una transcriptasa inversa (que genera un dúplex de ADN: ADN) y hace los conjuntos bajo condiciones estándar (Figura 6).

Los carriles 3 y 4 demostraron que los dúplex de ADN: ADN (extendidos con una ADN polimerasa PHUSION o un AMV - RT) se pueden tagmentar con un adaptador Tn5 y generar conjuntos (Figura 6).

El carril 5 no se esperaba que generará conjuntos por que la ADN polimerasa PHUSION ha sido previamente reportada por no extenderse opuesta a una cadena de ARN. El número pequeño de conjuntos observado puede ser debido a la plantilla de ADN residual utilizada para generar el ARN a pesar del tratamiento de DNasa, o algún grado de extensión mediante la ADN polimerasa PHUSION del ARN opuesto al ADN (Figura 6).

5 Los carriles 7 y 8 no se esperarían que exhibieran ninguna formación de conjunto por que las plantillas han sido tratadas con RNasa y DNasa. Como se vio en el carril 5, el número pequeño de conjuntos observado se puede deber a la plantilla de ADN residual utilizada para generar el ARN a pesar del tratamiento con DNasa (Figura 6).

10 El carril 6 de la Figura 6 demuestra la extensión de la cadena de ADN contra la plantilla de ARN tal como se esperó. No se esperara que se formaran estas plantillas extendidas de los conjuntos ya que ellos no poseen una secuencia P5. Sin embargo, luego de la tagmentación con el adaptador P5 (*q. e. d.* carril 6), ellos forman conjuntos. El número pequeño de conjuntos en los carriles 5, 7 y 8 sugiere que existe un bajo nivel de contaminación de ADN en la muestra de ARN, pero muestra que la mayoría de los conjuntos en el carril 6 se generaron del ARN.

Los conjuntos del “flowcell” fueron entonces secuenciados. La Tabla 1 muestra los resultados de la secuenciación.

Tabla 1. Resumen de secuenciación

	Carril	Rendimiento del carril	Conjuntos (bruto)	Conjuntos (PF)	1mer Ciclo Int PF)	% de Intensidad después de 20 ciclos (PF)	%PF conjunto	% de alineación (PF)	Calificación de alineación (PF)	% de tasa de error (PF)
PhiX	1	116	177455 +/- 11971	161186 +/- 10237	289+/-9	86080+/-0.97	90388 +/-1.64	98.28+/-0.22	166.43+/-0.00166.03 +/-0.11	0.04+/-0.00
PhiX	2	78	116338 +/-3547	108649 +/-3484	282+/-8	86.4+/-0.75	93.39+/-0.37	98.11+/-0.07	166.03+/-0.11	0.05+/-0.01
PhiX+Tn	3	40	63725+/-1557	55752+/-1477	289+/-4	87.78+/-1.13	87.49+/-0.34	58.26+/-0.25	96.55+0.39	0.28+/-0.00
PhiX+Tn	4	23	42441+/-1497	32075+/-1307	267+/-16	88.48+/-1.77	75.56+/-0.76	35.54+/-0.63	57.81+/-0.99	0.49+/-0.02
RNA+Tn	5	3	13608+/-731	4332+/-510	295+/-54	121.18+/-17.78	32.00+/-4.62	55.95+/-15.91	49.00+/-24.03	2.76+/-0.37
RNA+Tn	6	4	20916+/-1519	5701+/-1642	179+/-35	219.02+/-82.12	26.98+/-6.47	25.45+/-8.27	19.48+/-13.01	2.77+/-0.70
RNA+Tn+RN Asa	7	1	22597+/-4274	1627+/-317	190+/-65	104.24+/-35.64	7.43+/-1.99	12.12+/-6.14	3.69+/-2.40	3.45+/-0.83
RNA+Tn+RN Asa	8	1	45904+/-10022	1033+/-156	178+/-22	108.42+/-12.14	2.26+/-0.35	41.32+/-9.19	46.18+/-19.94	1.47+/-0.72

15

Carriles 1, 3, 5 y 7 se amplificaron con Phusion, no se asumió para amplificar ARN.

Los carriles 2, 4, 6, y 8 se amplificaron con AMV-RT, el cual amplificó el ADN y el ARN.

La secuenciación para los carriles 1 y 2 no traspuestos con SBS3+T, para los carriles traspuestos 3-8 con cebador Nx-R1

20 La matriz y la fase ajustada, carriles 5 – 8 alineados a GFP.

25 Como se esperó, más del 90% de los conjuntos pasaron los filtros de castidad para el carril 1 y 2 y de estos más del 98% se alinearon con PhiX tal como se esperó (Tabla 1). Los carriles 3 y 4 que contenían los dúplex de ADN: ADN tagmentado exhibieron 10 – 20% de reducción en los conjuntos que pasaron el filtro, de los cuales 75 – 87% de los conjuntos se alinearon a PhiX. Dado que esa tagmentación puede reducir la longitud de una plantilla, en algunos casos a una longitud demasiado corta para alinearse efectivamente, no se espera una reducción en los conjuntos que pasen los filtros y alineación. Los conjuntos en los carriles 7 y 8 no deben secuenciarse bien ya que no debe haber ninguna plantilla presente (con la excepción de las plantillas de ADN contaminantes o tocones de ARN no digeridos). Como se esperó muy pocos conjuntos pasaron los filtros. Menos del 7% de los conjuntos pasaron los filtros de los cuales solamente el 12% se alineó para las plantillas extendidas de ADN polimerasa PHUSION el 41% se alinea para las plantillas extendidas AMV-RT. Cuando no se hizo un tratamiento con RNasa, solamente la DNasa y la ARN extendida con Phusion 32% de los conjuntos pasaron los filtros de los cuales 56 +/-16% se alinearon (carril 5, Tabla 1). Esto se puede deber a la combinación de las plantillas de ADN residual y alguna extensión del ARN opuesto al ADN mediante la ADN polimerasa PHUSION. En contraste, aproximadamente el 50% más de los conjuntos de observaron en el carril,

30

donde la plantilla de ARN tratada con DNasa se extendió con el AMV – RT y de los cuales un porcentaje similar pasó el filtro (~ 27%) al carril 6 con el 25% de alineación.

5 Los datos alineados se utilizaron para generar gráficas de cubrimiento (Figura 7). Los carriles 1 – 4 dieron el cubrimiento del genoma completo del PhiX tal como se esperó. Los carriles que contenían ADN tagmentado (carriles 3 y 4) dieron más cubrimiento no homogéneo. Los carriles 5 – 8, que contenían muestras de ARN tagmentado mostraron todos cubrimiento parcial del GFP, que indica que la tagmentación de la plantilla ha generado conjuntos. Dado que algunos de estos se pueden derivar de la plantilla de ADN residual, el carril 6 muestra el cubrimiento más amplio de la plantilla GFP, que indica que las moléculas de ARN extendidas con AMV – RT han sido tagmentadas exitosamente.

Ejemplo 2. Secuenciación de ARN que emplea una reacción de tagmentación de muestras humanas.

10 Un “flowcell” con 8 carriles se preparó comprendiendo cebadores capaces de hibridar a las moléculas de ARN que comprenden una cola de poliA como sigue. El carril 1 fue injertado con una mezcla oligo estándar que comprende solamente los oligos P5 y P7 y los carriles 2 – 8 fueron injertados con una mezcla estándar (oligos P5 y P7) más el oligo de captura (es decir el cebador que comprende una secuencia poliT para unirse a las moléculas de ARN que comprenden una cola PoliA). Después del injerto del cebador, el “flowcell” se almacenó a 4°C hasta que se utilizó.

15 Para los carriles 1 y 2, 5 pM de las muestras de biblioteca de control PhiX se prepararon y agregaron al “flowcell” para hibridación. Para cada carril 3 – 8, se prepararon 400 ng de una muestra de ARN y se agregaron al “flowcell” para hibridación. Los carriles 3 y 4 contenían ARN humano de Clontech (Mountain View, CA) Los carriles 5 y 6 contenían ARN humano del cerebro. Los carriles 7 y 8 contenían ARN de referencia humana universal (UHR). Después de la hibridación de la plantilla, se administró el amortiguador a través del “flowcell”, para retirar la plantilla no hibridada. Las
20 plantillas hibridadas se extendieron utilizando AMV – RT (NEB, Ipswich, MA) en todos los carriles, que produjeron dúplex de ADN: ARN en los carriles 3 – 8.

Mientras que los carriles 3- 8 fueron conectados con los complejos de transposoma, los carriles 1 y 2 fueron puestos en contacto con un volumen equivalente de amortiguador de lavado. Se prepararon complejos de transposoma mezclas de
25 dos diferentes concentraciones. Las mezclas para los carriles 3, 5 y 7 se prepararon con 1.25 µl del complejo de transposoma, 100 µl de amortiguador y 400 µl de agua. La mezcla para los carriles 4, 6 y 8 se preparó con 0.625 µl de complejo de transposoma, 100 µl de amortiguador y 400 µl de agua. 95 µl de mezclas de complejo de transposoma se agregaron a los carriles 3– 8 del “flowcell” para tagmentación. Para retirar la transposasa después de la tagmentación, se agregó amortiguador caotrópico a los carriles 3- 8 del “flowcell” y se incubó durante 2 minutos. Los carriles del
30 “flowcell” fueron entonces lavados dos veces. Después de lavado, se utilizó la enzima Bst para extensión de desplazamiento de cadena de los dúplex de ADN: ARN tagmentados para remover la cadena no transferida del transposón y hacer la cadena de ADN de los dúplex de ADN: ARN de longitud completa para formar conjuntos. Las cadenas de ARN fueron removidas y los conjuntos fueron entonces generados utilizando amplificación isotérmica. Los conjuntos fueron entonces secuenciados. La Tabla 2 muestra los resultados de la secuenciación.

Tabla 2. Resumen de la secuenciación

Carril	Muestra	Conjuntos (bruto)	Conjuntos (PF)	1 Ciclo Int (PF)	% int después de 20 ciclos (PF)	% PF conjuntos	% de alineación (PF)	% de tasa de error (PF)
1	ADN PhiX Cebador 2	73569+/-8007	68422+/-7981	284+/-13	87.51+/-2.89	92.94+/-0.77	97.99+/-0.23	0.06+/-0.00
2	ADN PhiX Cebador 3	18553+/-2932	11200+/-2451	206+/-10	99.83+/-6.39	60.00+/-4.34	1.55+/-1.36	8.78+/-1.60
3	Clontech 1x Tn5	187046+/-29545	164971+/-26054	209+/-13	85.35+/-2.94	88.23+/-1.30	73.27+/-0.46	0.30+/-0.00
4	Clontech 0.5x Tn5	109889+/-13109	99558+/-13108	211+/-10	87.58+/-4.32	90.49+/-1.19	73.91+/-0.56	0.27+/-0.02
5	Cerebro 0.5 xTn5	226164+/-31941	198031+/-28192	218+/-6	84.55+/-2.36	87.55+/-0.97	75.27+/-0.36	0.36+/-0.20
6	Cerebro 0.5x Tn5	125939+/-21818	113273+/-18279	212+/-12	86.85+/-30.9	90.06+/-1.20	75.91+/-0.13	0.24+/-0.00
7	UHR 1X Tn5	310276+/-21976	269047+/-17778	195+/-7	96.77+/-2.14	86.75+/-1.68	67.70+/-0.38	0.27+/-0.04

8	UHR 0.5X Tn5	195323+/- 16530	172838+/- 15327	211+/-16	86.70+/- 1.77	86.47+/- 0.67	68.14+/- 0.52	0.36+/- 0.30
---	-----------------	--------------------	--------------------	----------	------------------	------------------	------------------	-----------------

5 Los resultados de la secuenciación se compararon con los resultados obtenidos para la secuenciación de ARN estándar del ARN de cerebro humano y el ARN de referencia humana universal, que se llevó a cabo de acuerdo con los métodos de secuenciación Illumina estándar utilizando los reactivos de secuenciación Illumina estándar. Tales métodos se describen en la guía de reparación de muestra de ARN TruSeq y la Guía de Usuario HiSeq2000. Las guías y los reactivos están disponibles de Illumina, Inc. (San Diego, CA). Los resultados se muestran en la Tabla 3.

Tabla 3. Comparación del método de tagmentación con el método de secuenciación de ARN estándar

	ARN Clontech			ARN UHR			ARN de cerebro			Secuenciación de ARN Estándar			
	Lectura 1		Lectura 2	Lectura 1		Lectura 2	Lectura 1		Lectura 2	UHR	UHR	Cerebro	Cerebro
Conjuntos Totales	16.000	.000		16.000	.000		16.000	.000					
PFClusters	14.113	100	100	14.008	100	100	13.957	100	100	78.895	100	80.670.	100
usableClusters	9.005.	63.	62.6	8.756.	62.	60.9	8.956.	64.	61.1	59.010	74.	59.821.	74.2
noMatch	622.29	4.4	2.2	600.13	4.3	2.7	484.15	3.5	3.1	1.352.	1.7	1.503.3	1.9%
repeatMasked	4.484.	31.	35.1	4.651.	33.	36.4	4.517.	32.	35.7	18.525	23.	19.342.	24%
spliceUsable	9.058	0.1	4.9	6.427	0%	3.3	22.719	0.2	7.0	9.357.	11.	6.959.6	8.6%
genomeUsable	8.996.	63.	57.8	8.750.	62.	57.6	8.933.	64	54.1	49.652	62.	52.861.	65.5
chrM.fa	2.045.	14.	14.1	2.570.	18.	17.9	737.70	5.3	5.3	5.710.	7.2	10.414.	12.9
humRibosomal.fa	9.409	0.1	0.1	7.248	0.1	0.1	8.839	0.1	0.1	1.789.	2.3	2.349.2	2.9%

10 Los resultados muestran una distribución de alineamiento normal para las muestras de ARN secuenciadas utilizando el método de tagmentación suministrado aquí. Los resultados muestran mayores conjuntos enmascarados repetidos probablemente debido a los mayores números de secuencia de poliA y más repetidos en las regiones UTR3' de las muestras de ARN analizadas mediante el método de tagmetación. Las lecturas utilizables fueron aproximadamente 10% inferiores que el protocolo de secuenciación de ARN estándar de nuevo probablemente debido a más repeticiones en el ARN que fue analizado. La cantidad de ARN ribosómico es baja como se esperaba ya que el mARN se aisló y secuenció en el método de tagmentación suministrado aquí. El ARN mitocondrial está dentro de los límites normales.

Ejemplo 3: Secuenciación de ARN que emplea una reacción de tagmentación y un lisado de célula.

20 Este ejemplo demuestra que las plantillas de ácido nucleico se pueden capturar, tagmentar y secuenciar sobre un soporte sólido utilizando un lisado de célula cruda. En resumen, células de ratón fueron lisadas utilizando una solución de Tritón- X y Proteinasa K. El lisado se aplicó a un "flowcell", el mARN se capturó y tagmentó, y los conjuntos fueron creados y secuenciados. Como un control y para comparación, también se capturó, tagmentó, se puso en conjunto y se secuenció ARN total de Referencia Humana Universal (UHR). La Tabla 4 hace un resumen de los resultados de las lecturas duplicadas para cada tipo de muestra.

25 Tabla 4. Comparación de los datos de secuenciación entre UHR y mARN de lisado.

	UHR		Lisado de célula de ratón	
	R1	R2	R1	R2
Lecturas totales	4.726.081		1.905.434	
%PF	86.89%	84.59%	83.94%	84.79%

ES 2 565 809 T3

	UHR		Lisado de célula de ratón	
	R1	R2	R1	R2
Alineado (de las lecturas % PF)	61.14%	73.97%	49.17%	68.16%
No alineado (de las lecturas %PF)	28.42%	12.00%	34.01%	8.45%
Abundante (de las lecturas %PF)	10.44%	14.03%	16.82%	23.39%
Alineamientos empalmados (% de bases alineadas)	0.80%	11.14%	0.62%	10.73%
Alineamientos empalmados (de las lecturas %PF)	0.4893%	8.2435	0.3054	7.3138
Ribosoma humano	0.21%	0.04%	3.98%	4.88%
5Sr humano	0.01%	0.01%	0.02%	0.17%
Inserto mediano	135		129	
Inserto SD	67.99%		66.97	30
duplicados	41.66%		45.54	

5 La Tabla 4 demuestra que la secuencia se obtuvo directamente del mRNA capturado de un lisado crudo de célula de ratón. El porcentaje de las lecturas alineadas cayó solo aproximadamente el 10% cuando se capturó el mRNA directamente de los lisados crudos de célula comparados con la muestra de ARN UHR (alineados de las lecturas % PF).
 5 Datos de secuenciación adicionales que comparan el control UHR con el lisado de ratón derivado del mRNA reportaron que se capturó la cadena correcta y que se alineo al > 97% tanto para el UHR como para el mRNA de lisado de ratón. Además, el cubrimiento fue comparable entre el control de UHR y el mRNA del lisado; aproximadamente el 65% de la región no traducida (UTR), aproximadamente el 16% de la región codificante, aproximadamente el 13% de la región intergénica, y pequeños porcentajes de lectura sintrónicos. Como tal, los presentes métodos pueden ser utilizados para
 10 capturar, tagmentar, hacer conjuntos y secuenciar mRNA de los listados de crudo.

Ejemplo 4. Secuenciación de ARN que emplea una reacción de tagmentación del transcrito de mRNA completo.

15 Este experimento se efectuó para demostrar que la muestra de mRNA que presenta un transcrito completo se puede capturar y tagmentar sobre un soporte sólido para suministrar información de secuencia siguiendo los métodos descritos aquí. En resumen, el enriquecimiento del ARN poliA se efectuó de 50ug de ARN total UHR (Agilent) que utiliza el kit Puris PoliA (Ambion). La fragmentación del ARN y el mRNA enriquecido con poliA fue hecho en 25 ul de amortiguador 1X T4 PNK (Epicentre) con 100ng de ARN poliA, donde la muestra se calentó a 95°C durante 5 min y se enfrió sobre hielo. El ARN fragmentado se fosforiló con PNK T4 y los fragmentos fueron puestos en cola con poliA utilizando 4
 20 unidades de polimerasa poliA de E. coli en 50 ul de amortiguador de polimerasa PoliA 2X que contiene 2mM de ATP de (Epicentre). El mRNA fragmentado poliadenilado se purificó utilizando un kit de limpieza y concentración de ARN (Zymo Research). Los controles incluyeron un control PhiX para validar el desempeño de la química de secuenciación y una muestra poliadenilada derivada del mRNA no total que se capturó y tagmentó para comparar con el transcrito completo del mRNA capturado de la agrupación UHR de ARN total complejo.

Los datos de secuenciación del replicado 1 (R1) de 2 se reporta en la Tabla 5 para el control (ctrl) pHiX, la muestra mRNA de control (captura 3') y el mRNA derivado de la muestra de ARN total (transcrito completo).

25 Tabla 5. Resumen de la secuencia de replicado para el transcrito mRNA completo

R1	Rendimiento de muestra (Mb)	Conjuntos (bruto)	Conjuntos (PF)	Primer ciclo Int (PF)	% de intensidad después de 20 ciclos (PF)	% de conjuntos (PF)	% de alineación (PF)	Calificación de alineamiento (PF)	% tasa de no coincidencia (PF)	% >= bases Q30 (PF)	SCore Calidad Mediana (PF)
Control	317	7.079.810	6.343.692	268	86.78	89.6	97.89	251.82	015	98.35	39.12
Captura 3'	1.077	28.753.789	21.543.617	290	85.17	74.92	57.15	75.08	0.74	92.34	36.58

R1	Rendimiento de muestra (Mb)	Conjuntos (bruto)	Conjuntos (PF)	Primer ciclo Int (PF)	% de intensidad después de 20 ciclos (PF)	% de conjuntos (PF)	% de alineación (PF)	Calificación de alineamiento (PF)	% tasa de no coincidencia (PF)	% >= bases Q30 (PF)	Score Calidad Mediana (PF)
Transcripto completo.	752	19.474.750	15.040.807	323	83.03	77.23	49.85	70.27	1.63	91.89	36.51

5 La Tabla 5 reporta que el porcentaje de lecturas alineadas fue comparable sin importar la fuente del mRNA (% de alineación (PF)) con una generación de conjunto alta. Adicionalmente, los datos de secuencia mostraron que el cubrimiento del transcripto del mRNA de control (captura 3') fue de aproximadamente 70%UTR, 19% de la región codificante seguida por el cubrimiento de la región intergénica e intrónica. El cubrimiento del transcripto del mRNA derivado del ARN total del complejo fue de aproximadamente 43% UTR, 37% codificante y relativamente similar para las regiones intergénica e intrónica. La Figura 10 demuestra el cubrimiento del transcripto alineado para un gen representativo, GAPDH; el mRNA de control (captura 3') muestra el cubrimiento principalmente en la región 3' del gen tal como se esperaba, mientras que el cubrimiento del mRNA derivado del ARN total (transcripto completo) mostró más cubrimiento completo de ambas regiones la exónica como la UTR. Como tal, aunque las secuencias mRNA de control alineadas a aquellas regiones asociadas con el extremo 3' del transcripto (región de cola poliA), el cubrimiento del mRNA derivado del ARN total demostró ser más completo, lecturas del transcripto total, demostrando de esta manera la utilidad de los métodos para obtener la información del transcripto completo de una muestra.

15 Un flujo de trabajo alternativo se efectuó para enriquecer el mRNA de la muestra de ARN total UHR para la secuenciación del transcripto completo. El cADN de cadena doble se preparó de 500ng del ARN total UHR y 50 ng de hexámeros de ADN aleatorio. Los cebadores en exceso fueron degradados al agregar 20 unidades de Exonucleasa 1 (Epicentro), incubando a 37°C durante 30 min seguidos por inactivación con calor de la enzima. El ARN se removió y la mezcla de enzima de un U de RNasa I/10 U Hibridaza (RNasa H, Epicentro) a 55°C durante 10 min. La reacción se purificó utilizando volúmenes iguales de glóbulos AMPure (Agencourt) y el ADN se eluyó en 10 mM de amortiguador Tris HCL (pH 8.0). EL cADN se le colocó una cola poliA utilizando 20 U de Transferasa Terminal (New England Biolabs), 1mM de ATP y 1X de amortiguador de Transferasa, incubando a 37°C durante 10 min. Seguido por inactivación con calor. Para algunas de las muestras, se utilizó una dilución 1:10 de los hexámeros de ADN aleatorios. Además, para algunas de las muestras se omitió la etapa de la Exonucleasa I. Las muestras fueron entonces aplicadas a un "flowcell" capturadas, tagmentadas, puestas en conjunto y secuenciadas. Los controles incluyeron el control (ctrl) PhiX , el cADN sin cola (ctrl sin cola), un control negativo dsADN, una muestra de mRNA purificada que siguió el mismo método como se describió anteriormente excepto utilizando los exámenes de ARN aleatorio y omitiendo la etapa de la Exonucleasa I.

La Tabla 6 resume las primeras dos corridas de secuenciación de réplica

Tabla 6. Resumen de secuencia de réplica para el transcripto de mRNA completo alternativo.

R1	Rendimiento (Mb)	Conjunto (crudo)	Conjunto (PF)	1er ciclo Int (PF)	% de intensidad después de 20 ciclos (PF)	% de conjuntos (PF)	% de alineamiento (PF)	Calificación de alineamiento (PF)	% de índice de no coincidencia (PF)	% >= Q30 bases (PF)	Score de calidad mediana (PF)
Carril 1-ctrl PhiX	227	7.040.939	6.480.317	317	89.58	92.04	98.16	165.52	0.08	98.25	38.88
Carril 2-ctrl sin cola	3	773.384	76.281	246	84.82	9.86	0.15	0.05	5.04	42.11	20.13
Carril 3-ARNas a, AMP, cola	62	3.504.536	1.783.482	420	75.32	50.89	20.14	13.65	1.29	85.96	34.57
Carril 4-1:10	15	1.353.520	429.352	339	91.53	31.72	8.37	6.38	1	62.02	27.88

R1	Rendimiento (Mb)	Conjunto (crudo)	Conjunto (PF)	1er ciclo Int (PF)	% intensidad después de 20 ciclos (PF)	% de conjuntos (PF)	% de alineamiento (PF)	Calificación de alineamiento (PF)	% de índice de no coincidencia (PF)	% >= Q30 bases (PF)	Score de calidad media (PF)
CARRIL 5-Exo, ARNasa, AMP, cola	97	4.922.187	2.774.812	373	83.4	56.37	40.74	27.62	1.47	86.68	34.8
Carril 6-1:10	7	844.731	198.052	317	79.07	23.45	13.09	6.81	1.38	42.76	21.15
Carril 7-cebadores de ARN	29	2.164.633	840.373	377	92.34	38.82	17.28	13.89	0.7	79.14	32.88
Carril 8-ads cADN	8	852.450	219.356	311	97.61	25.73	1.3	0.32	2.61	48.15	23.6

5 La Tabla 6 demuestra que el método de preparación de tratar una muestra con nucleasas, Exo 1 y RNasa H e I, seguido por purificación de glóbulo y colocación de cola poliA (Exco, ARNasa, AMP, cola) se puede utilizar para suministrar la muestra de transcripto de mRNA completa para secuenciación. La Figura 9 muestra fotografías de puestas en conjunto sobre el "flowcell" con respecto a las diferentes condiciones identificadas en la Tabla 6. Los carriles 1-8 en la Tabla 1 corresponden a los carriles 1- 8 en la Figura 9. El control positivo PhiX muestra un gran número de conjuntos que corresponden al más alto rendimiento del conteo del conjunto en los datos de secuenciación. Los carriles 2 y 8 de control negativo, que muestran menor número de conjuntos también corresponden a dos de los tres rendimientos más bajos y los conteos de conjuntos en los datos de secuenciación. Diluir los cebadores de examen o de ADN aleatorio 1: 10, sin importar la digestión de la exonucleasa en el método de preparación, no fue óptimo para la secuenciación, lo que muestra un bajo conteo de conjunto en la Tabla 6 soportado por más pocos conjuntos vistos en la Figura 9. El método que utiliza la RNsa H e I con Exonulceasa 1 durante la preparación de la muestra de los transcriptos del mRNA completo dieron como resultado el más grande número de conjuntos generados así como también el porcentaje de alineación después del control positivo PhiX, seguido por el método d preparación donde no se practicó la digestión de la Exonucleasa 1. Además, el porcentaje de lecturas alineadas (% de alineación (PF)) son las más altas para el carril 5 y el carril 3, respectivamente, entre los carriles de prueba. Los resultados demuestran que el método alternativo descrito en este ejemplo para generar cADN se puede utilizar para suministrar la información del transcripto de mRNA completa al secuenciar utilizando los métodos de captura y tagmentación descritos en esta solicitud.

20 Las opciones adicionales para la preparación de la muestra incluyen, pero no están limitadas a, utilizar los métodos descritos para secuenciar ARN de especies que no tengan ARN poliadenilado, tal como mRNA bacteriano. En este caso, el ARN ribosómico se puede primero remover y el mRNA que permanece se puede fragmentar y poner la cola poliA como se describió previamente. El mRNA podría entonces ser capturado, tagmentado, puesto en conjunto amplificado y secuenciado como se describió anteriormente.

25 Estos resultados muestran que una variedad de los diferentes tipos de muestras de ARN y muestras de ADN derivadas de las muestras de ARN se pueden secuenciar utilizando los métodos suministrados aquí en que los métodos suministrados aquí suministran aproximadamente resultados de secuencia equivalentes a los protocolos de secuenciación de ARN estándar corriente.

LISTA DE SECUENCIAS

<110> Illumina Cambridge Limited
 <120> Métodos Mejorados de Secuenciamiento de Ácido Nucleico
 <130> ILC130301PCT
 5 <150> US61/607418
 <151> 2012-03-06
 <160> 3
 <170> BiSSAP 1.0
 <210> 1
 10 <211> 19
 <212> ADN
 <213> secuencias artificiales
 <220>
 <221> fuente
 15 <222> 1 .. 19
 <223> /mol_type="ADN"
 /nota="secuencia de transposón transferida"
 /organismo="secuencias artificiales"
 <400> 1 19
 20 aqatqtqtat aaqaqacaq
 <210> 2
 <211> 19
 <212> ADN
 <213> secuencias artificiales
 25 <220>
 <221> secuencias
 <222> 1 .. 19
 <223> /mol_type="ADN"
 /note="secuencia de transposón no transferida"
 30 /organismo="secuencias artificiales"
 <400> 2 19
 ctqtctctta tacacatct
 <210> 3
 <211> 1003
 35 <212> DNA
 <213> secuencias artificiales
 <220>

ES 2 565 809 T3

<221> fuente

<222> 1. .1003

<223> /mol_type="DNA"

/note="cassette de expresión GFP"

5 / organismo="secuencias artificiales"

<400> 3

```

aatgatacgg cgaccaccga gatctacact cttccctac acgacgctct tccgatctta      60
atagactca ctataggcaa tttaacttt actaaggaga attcaccatg aacatcacc      120

atcaccacac gactacaaag acgatgacga caaggcgatc gtgagcaagg gcgaggagct      180
gttcaccggg gtggtgccca tcctggtcga gctggacggc gacgtaaacg gccacaagtt      240
cagcgtgtcc ggcggggcca gggcgatgcc acctacggca agctgaccct gaagtcatc      300
tgaccaccg gcaagctgcc cgtgccctgg ccaccctcg tgaccaccct gacctacggc      360
gtgcagtgtc tcagccgcta cccgaccaca tgaagcagca cgacttcttc aagtccgcca      420
tgcccgaagg ctacgtccag gagcgcacca tcttcttcaa ggacgacggc aactacaaga      480
cccgcgccga ggtgaagttc gagggcgaac cctggtgaac cgcacgagc tgaagggcat      540
cgacttcaag gaggacggca acatcctggg gcacaagctg gagtacaact acaacagcca      600
caacgtctat atcatggccg acaagcagaa gaacgcatca aggtgaactt caagatccgc      660
cacaacatcg aggacggcag cgtgcagctc gccgaccact accagcagaa cacccccatc      720
ggcgacggcc ccgtgctgct gcccgacaac cactacctga gcccagtc ccctgagca      780
aagaccccaa cgagaagcgc gatcacatgg tcctgctgga gttegtgacc gccgccggga      840
tactctcgg catggacgag ctgtacaagt aactgctgcc accgctgaga ataactagca      900
taacccttg gggcctctaa acgggtcttg aggggttttt tgagatcgga agagcggttc      960
agcaggaatg ccgagaccga tctcgtatgc cgtcttctgc ttg      1003

```

Reivindicaciones

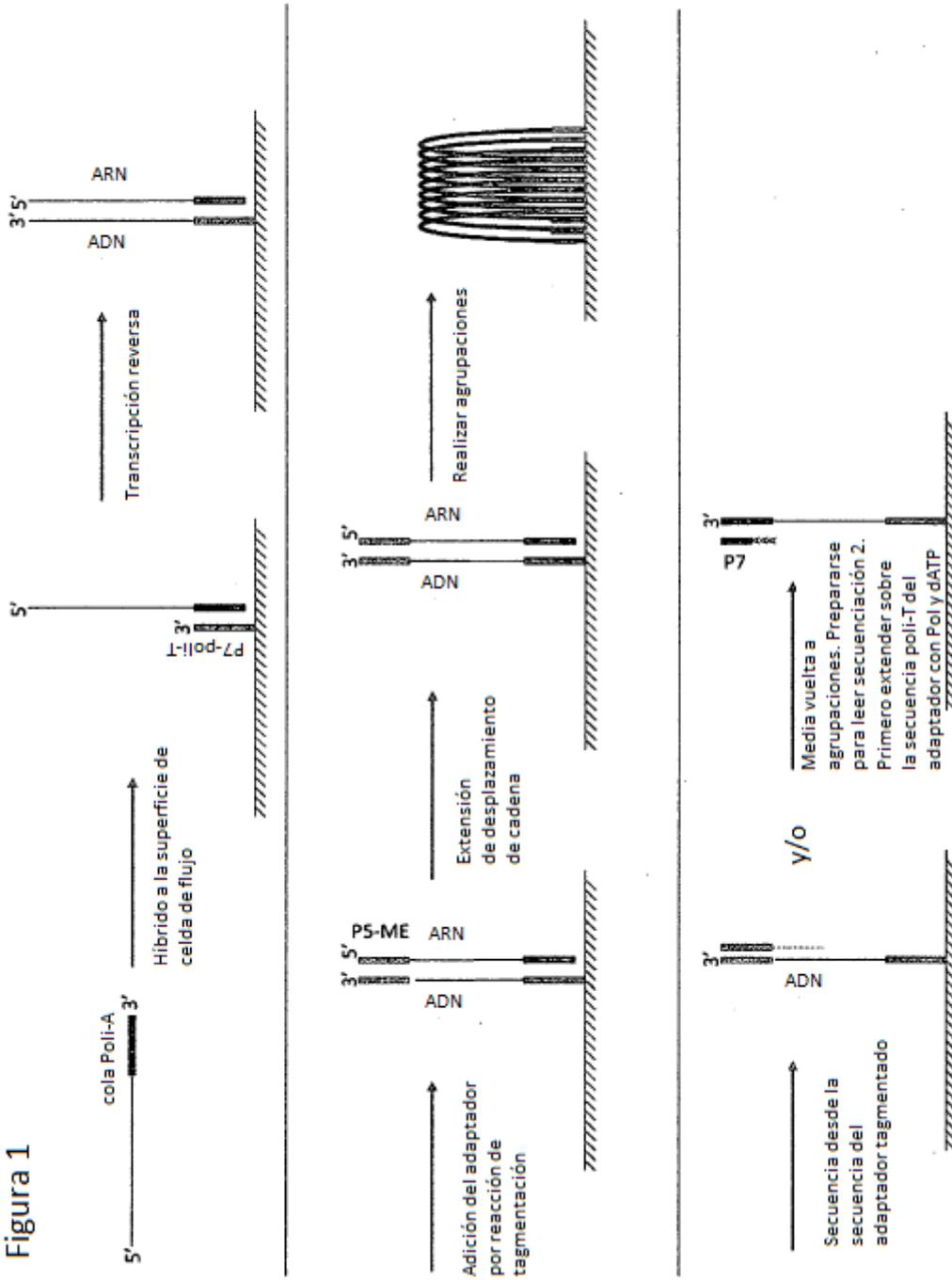
1. Un método de etiquetar dúplex de ácido nucleico que comprende las etapas de:
 - (a) suministrar una transposasa y una composición de transposón;
 - (b) suministrar uno o más dúplex de ácido nucleico y movilizado sobre un soporte;
 - (c) poner en contacto la transposasa y la composición de transposón con los uno o más dúplex de ácido nucleico bajo condiciones en donde los uno o más dúplex de ácido nucleico y la composición de transposón sufren una reacción de transposición para producir uno o más dúplex de ácido nucleico etiquetados, en donde la composición de transposón comprende una molécula de ácido nucleico de doble cadena que comprende una cadena transferida y una cadena no transferida.
2. El método de la reivindicación 1, en donde los uno o más son los dúplex de ADN: ARN y en donde los dúplex de ADN: ARN son etiquetados sobre el extremo 5' de la cadena de ARN, o, en donde los dúplex de ácido nucleico son los dúplex de ADN: ADN y en donde una de las cadenas del dúplex de ADN: ADN está etiquetada en el extremo 5' de la cadena de ADN.
3. El método de las reivindicaciones 1 o 2, en donde el soporte comprende una pluralidad de cebadores inmovilizados, y opcionalmente, en donde la cadena de ARN de los uno o más dúplex de ADN: ARN o de la cadena de ADN de los dúplex de ADN: ADN comprende la secuencia complementaria a al menos una porción de uno o más de los cebadores inmovilizados.
4. El método de la reivindicación 3, en donde los dúplex de ADN: ARN o los dúplex de ADN: ADN se suministran al hibridar una o más moléculas de ARN o moléculas de ADN de cadena simple a los cebadores inmovilizados sobre el soporte y extendiendo los cebadores hibridados a las moléculas de ARN o a las moléculas de ADN de cadena sencilla utilizando las moléculas de ARN o las moléculas de ADN de cadena sencilla como plantilla para producir los uno más dúplex de ADN: ARN o los dúplex de ADN: ADN.
5. El método de la reivindicación 3, en donde la pluralidad de los dúplex de ADN: ARN o de los dúplex de ADN: ADN se suministran al hibridar la pluralidad de las moléculas de ARN o las moléculas de ADN de cadena sencilla a los cebadores inmovilizados sobre el soporte y extendiendo los cebadores civilizados a las moléculas de ARN o a las moléculas de ADN de cadena sencilla utilizando las moléculas de ARN o las moléculas de ADN de cadena sencilla como plantilla para producir la pluralidad de dúplex de ADN: ARN o moléculas de ADN: ADN.
6. El método de la reivindicación 3, en donde la pluralidad de cebadores inmovilizados comprende un primer subconjunto de cebadores en una primera secuencia y un segundo subconjunto de cebadores de una segunda secuencia, y opcionalmente, en donde el primer subconjunto de cebadores comprende una secuencia poliT.
7. El método de la reivindicación 3, que comprende además agregar un adaptador 3' a la pluralidad de moléculas de ARN o las moléculas de ADN de cadena sencilla, el adaptador 3' que comprende una secuencia complementaria a la pluralidad de cebadores inmovilizados o un subconjunto de los mismos; e hibridar las moléculas de ARN ligadas al adaptador 3' o a las moléculas de ADN de cadena sencilla a los cebadores inmovilizados.
8. El método de una cualquiera de las reivindicaciones 1- 7, en donde la reacción de transposición da como resultado una cadena de ARN etiquetada 5' o una cadena de ADN etiquetada 5' que comprende la cadena transferida de la composición del transposón y un bache entre el extremo 3' de la cadena de ADN complementaria y la cadena no transferida de la composición de transposón, el método opcionalmente comprende además poner en contacto los uno o más dúplex de ADN: ARN etiquetado o los dúplex de ADN: ADN con una enzima que modifica el ácido nucleico bajo condiciones para extender el extremo 3' de las cadenas de ADN para copiar las cadenas de ARN a su extremo 5' o copiar las cadenas de ADN de cadena sencilla a sus extremos 5'.
9. El método de la reivindicación 8, que comprende además secuenciar al menos una porción de las cadenas de ADN, y opcionalmente además comprende copiar al menos una porción de las cadenas de ADN para producir una segunda cadena de ADN complementaria a la cadena de ADN de los dúplex de ADN: ARN, y opcionalmente además que comprende secuenciar la segunda cadena de ADN complementaria, en donde la cadena de ADN de los dúplex de ADN: ARN se remueve opcionalmente antes de secuenciar la segunda cadena de ADN complementaria.
10. El método de la reivindicación 8, que comprende además amplificar las cadenas de ADN para producir una pluralidad de moléculas de ADN de doble cadena que comprende primeras y segundas cadenas amplificadas, opcionalmente además que comprende remover las primeras cadenas amplificadas, y opcionalmente además que comprende secuenciar al menos una porción de las segundas cadenas amplificadas.
11. El método de la reivindicación 10, que comprende además copiar al menos una porción de las segundas cadenas amplificadas para regenerar las primeras cadenas amplificadas, opcionalmente además que comprende remover las segundas cadenas amplificadas, y opcionalmente, además que comprende secuenciar al menos una porción de las primeras cadenas amplificadas.
12. El método de una cualquiera de las reivindicaciones precedentes, en donde el soporte es un glóbulo o una pluralidad de glóbulos, o en donde el soporte es un soporte plano.

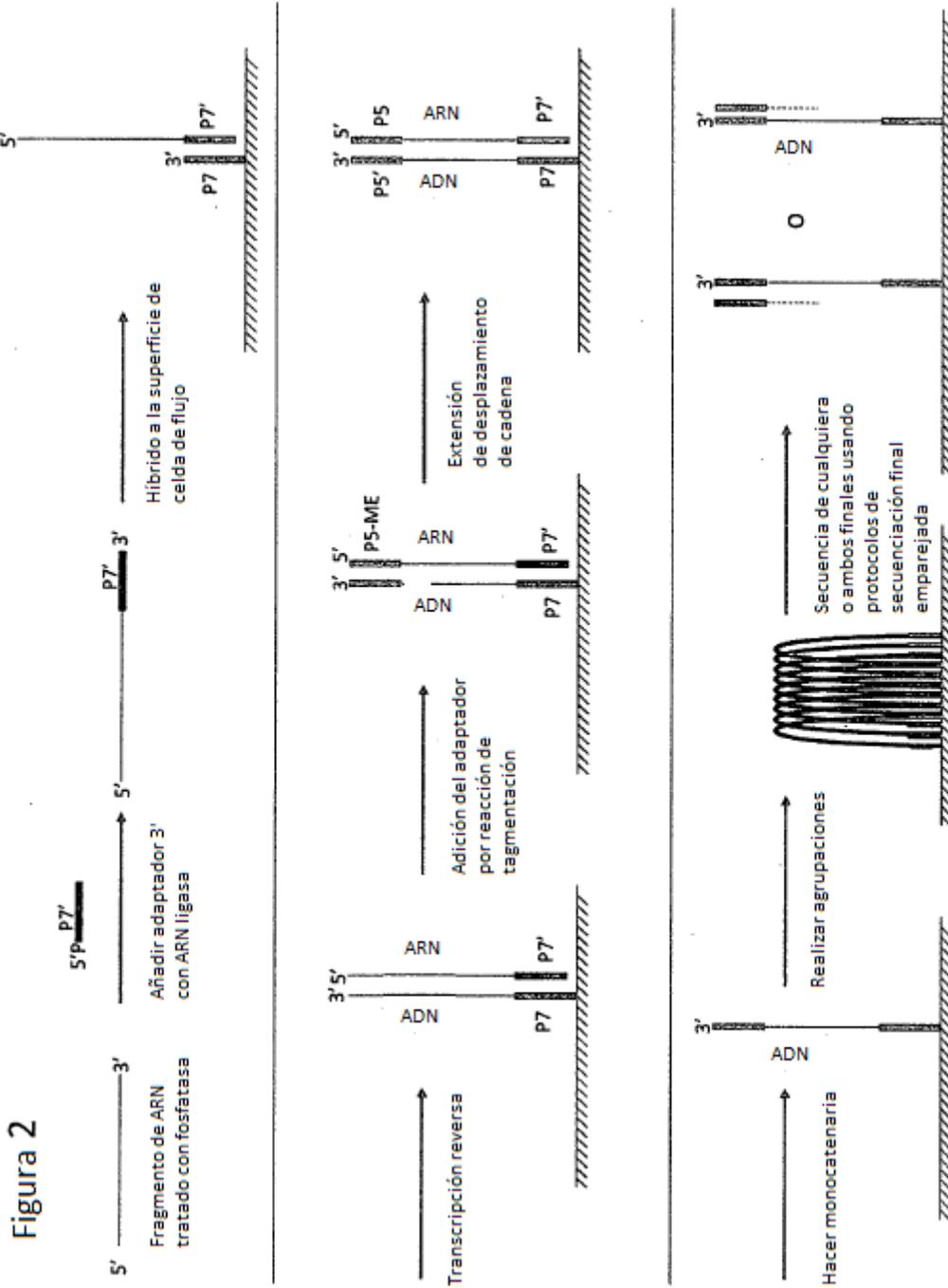
13. El método de la reivindicación 12, en donde se suministra una pluralidad de dúplex, cada dúplex ubicado sobre una cadena sobre un glóbulo simple.

14. El método de la reivindicación 10, en donde la amplificación produce un conjunto de amplicones

5 15. El método de una cualquiera de las reivindicaciones precedentes, en donde la cadena transferida comprende una etiqueta para preservar la información de cadena.

Figura 1





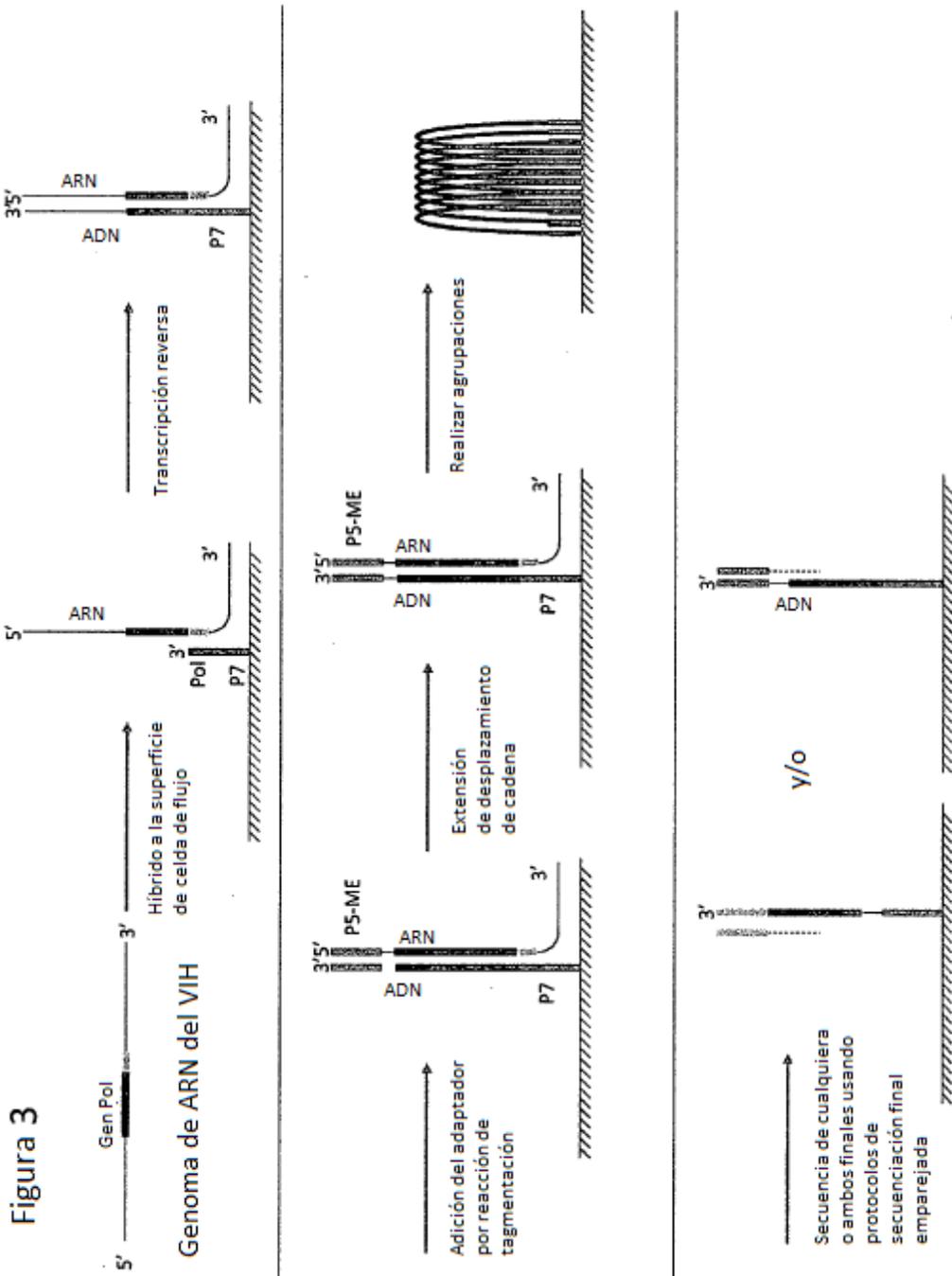


Figura 4

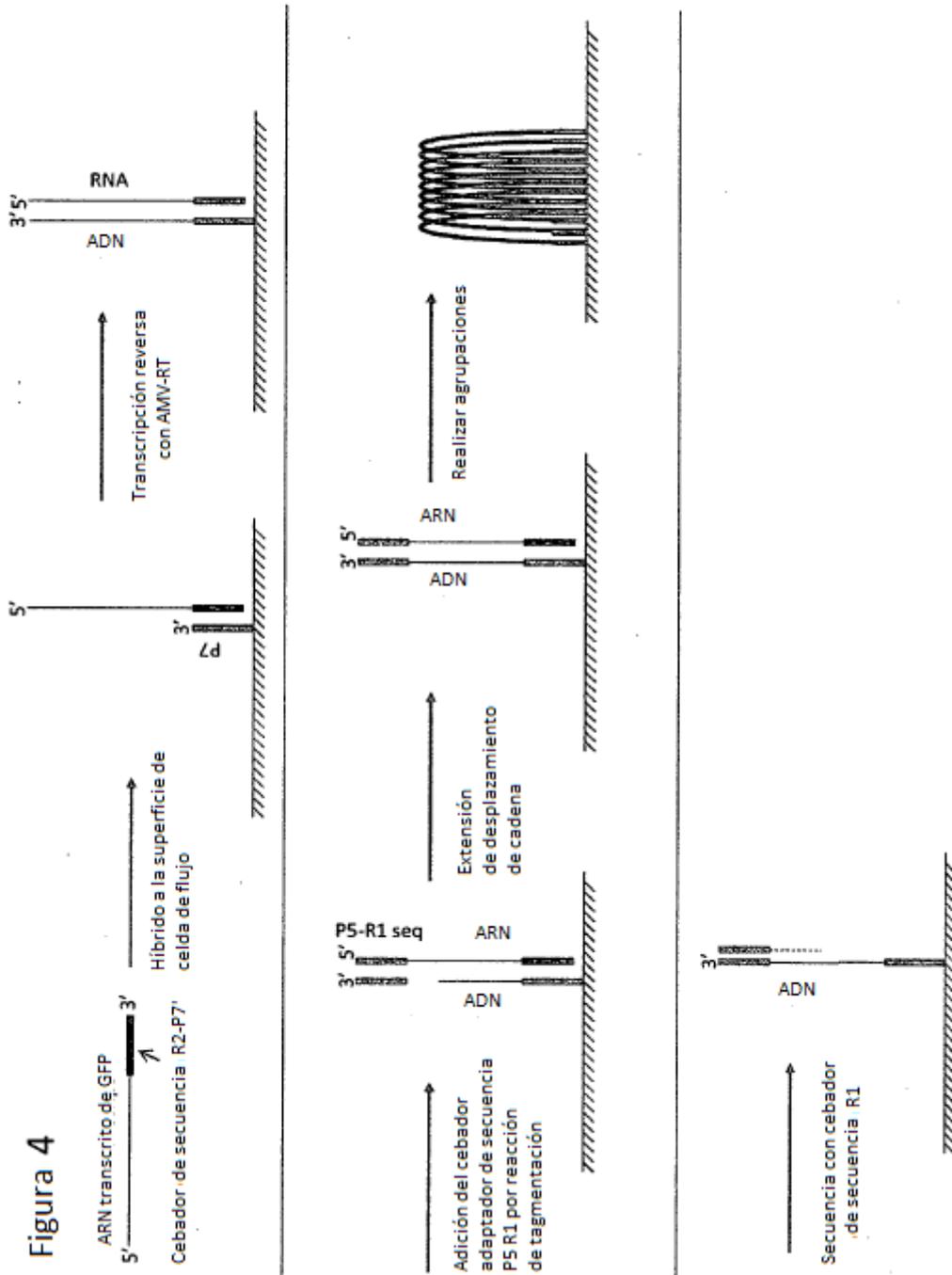
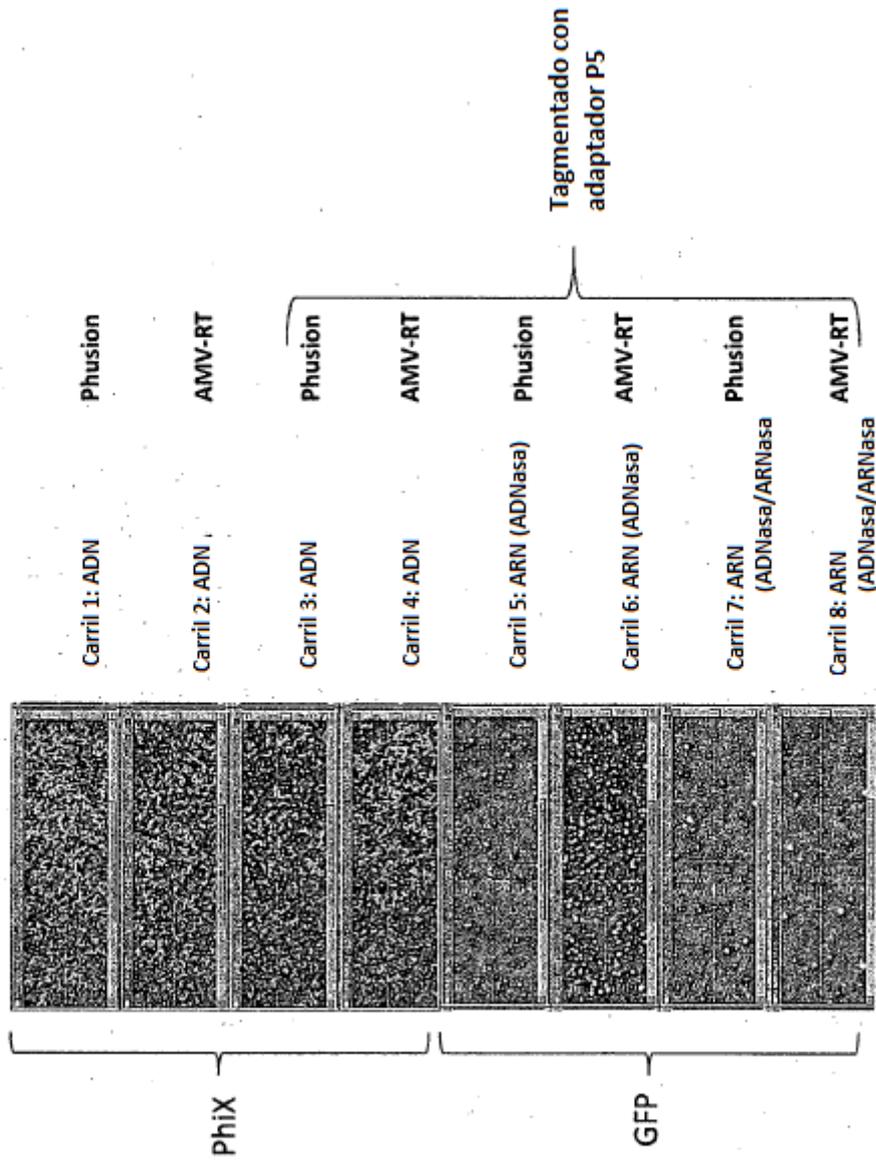


Figura 6



ARN pol T7
ADNasa

+

-

+

-

+

-



Figura 5

Figura 7A

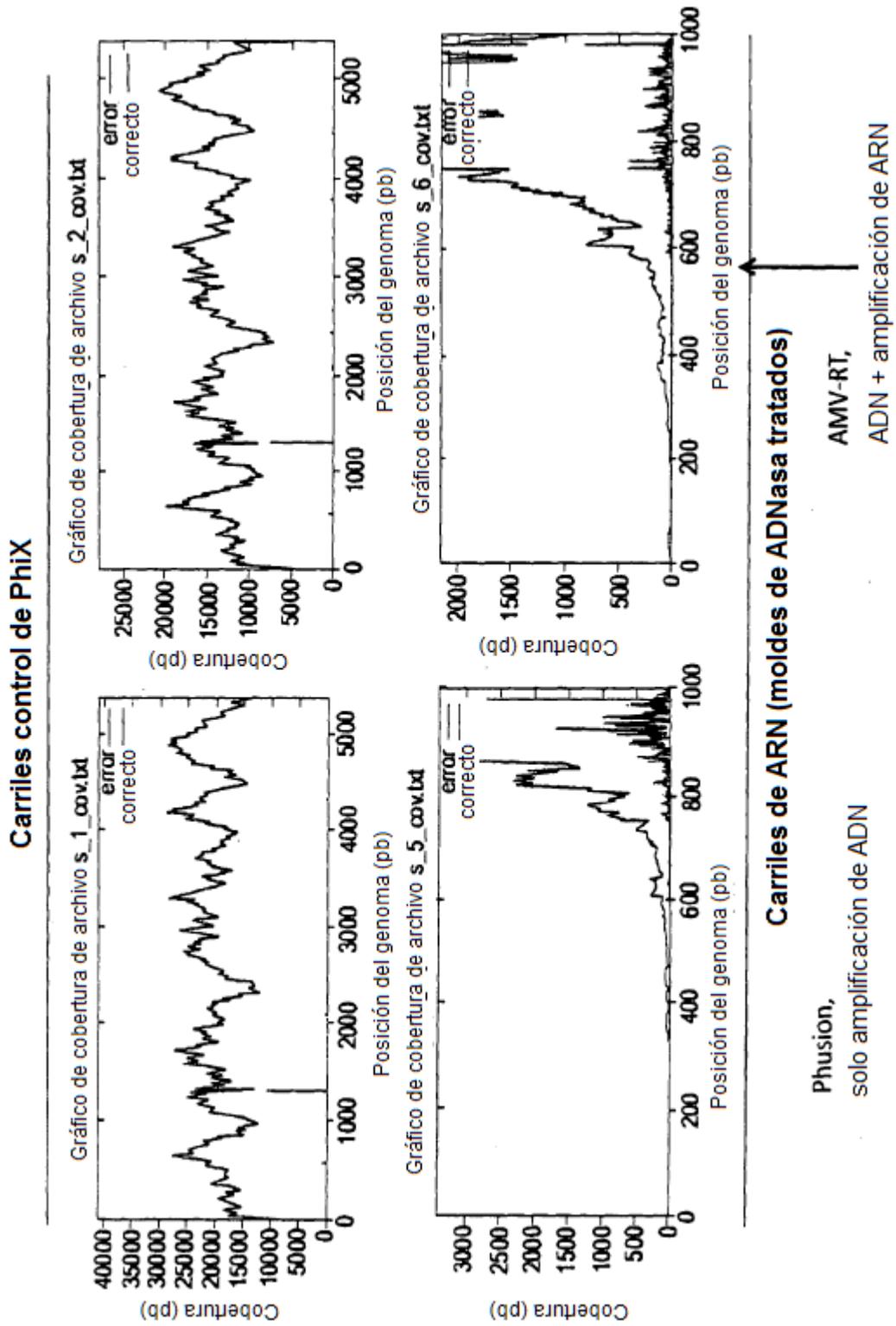
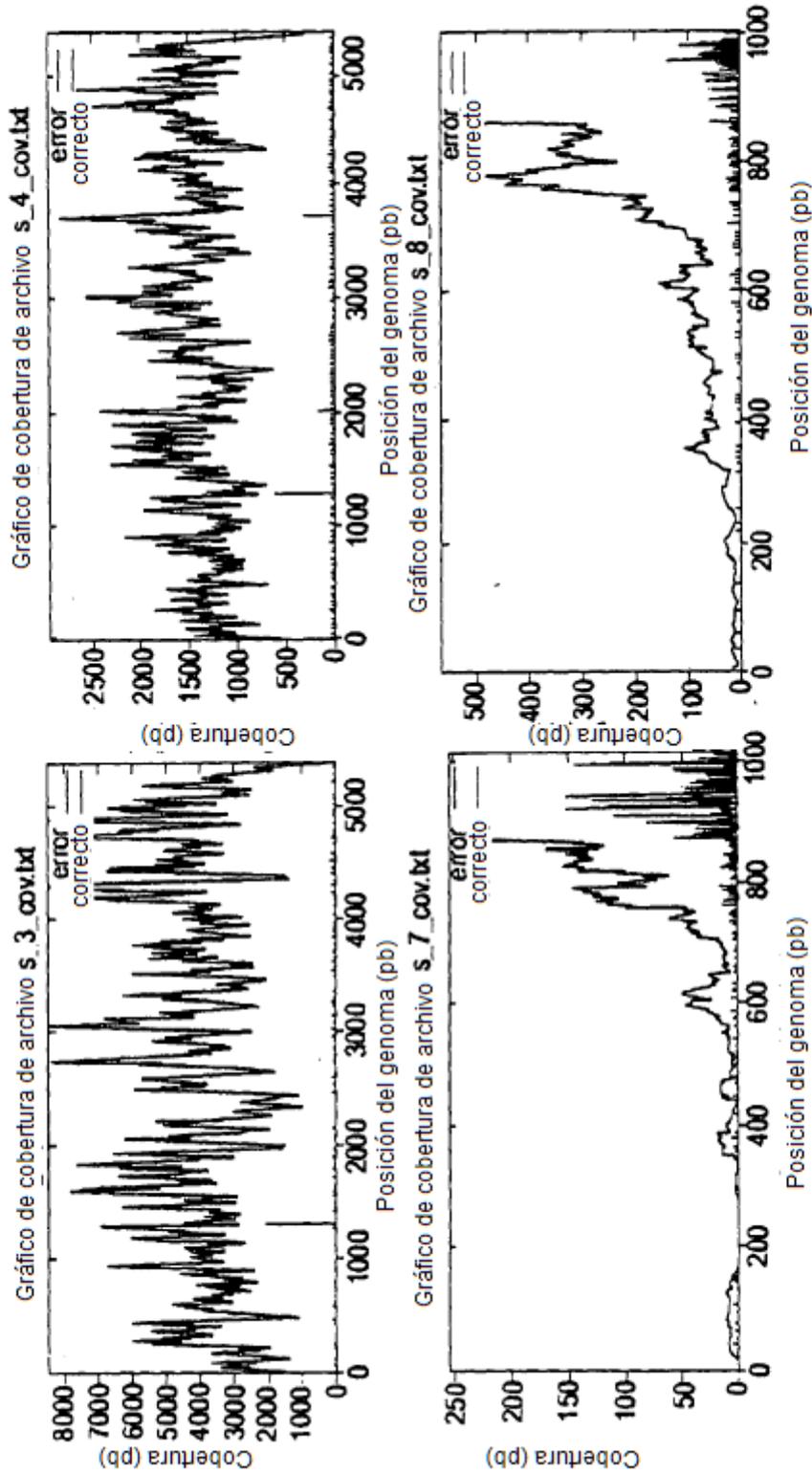


Figura 7B

PhiX + transposón



Carriles de ARN (moldes de ADNasa/ARNasa tratados)

Phusion resto único de amplificación de ADN AMV-RT

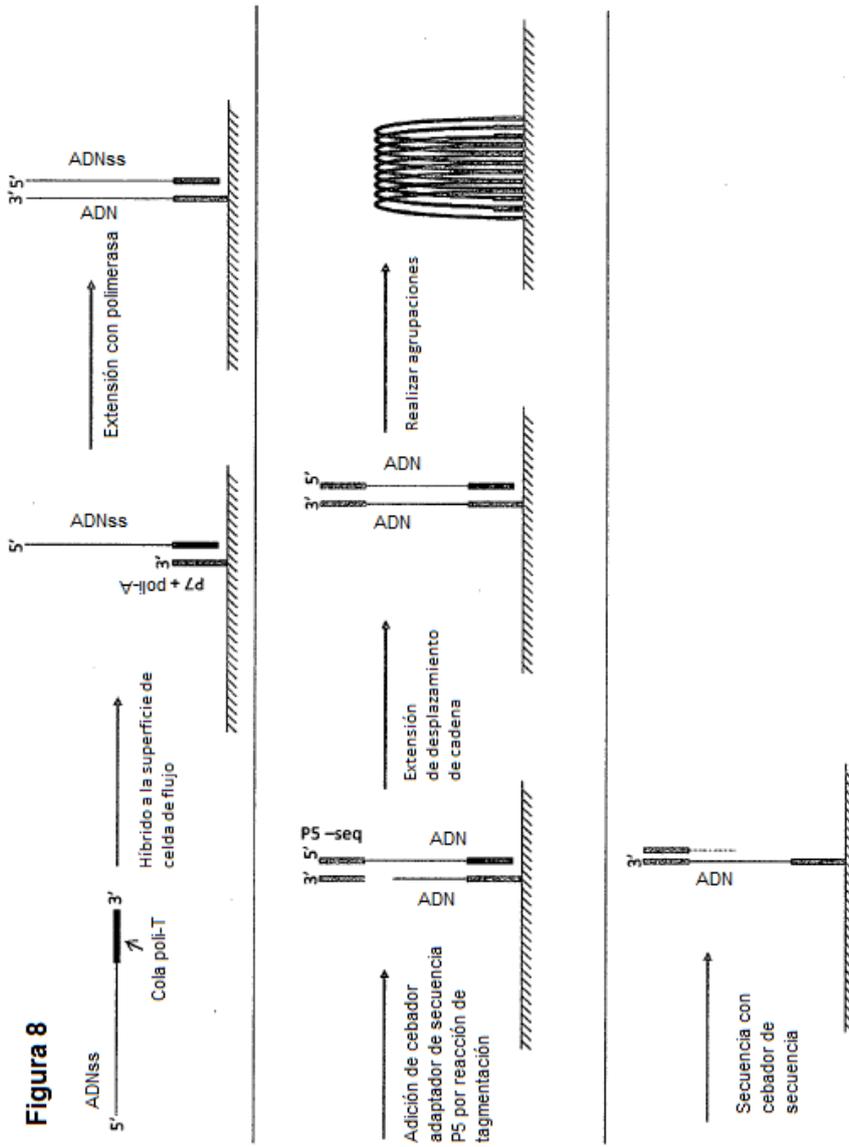


Figura 8

Figura 9

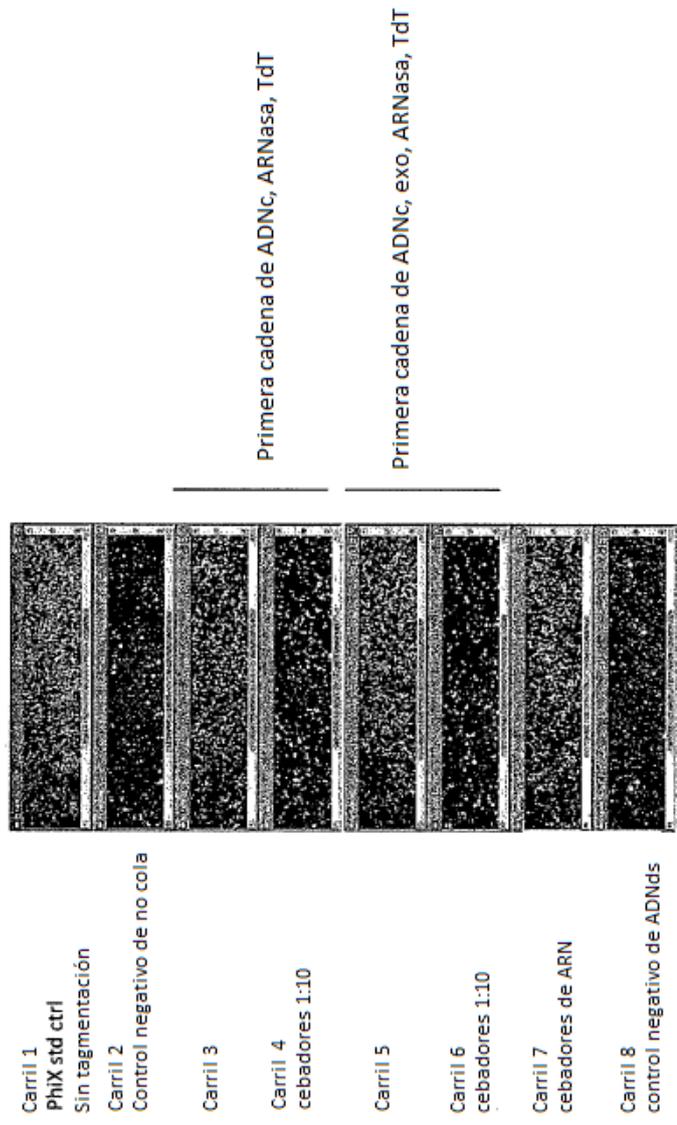


Figura 10

