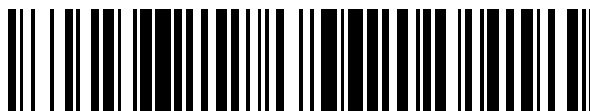


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 763 822**

51 Int. Cl.:

C12Q 1/6813 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **09.12.2016 PCT/IB2016/057499**

87 Fecha y número de publicación internacional: **15.06.2017 WO17098468**

96 Fecha de presentación y número de la solicitud europea: **09.12.2016 E 16820351 (1)**

97 Fecha y número de publicación de la concesión europea: **02.10.2019 EP 3387150**

54 Título: **Análisis exento de marcas de la eficacia de la adición de caperuzas al arn utilizando rnasa h, sondas y cromatografía líquida/espectrometría de masas**

30 Prioridad:

09.12.2015 US 201562265099 P

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

01.06.2020

73 Titular/es:

**NOVARTIS AG (100.0%)
Lichtstrasse 35
4056 Basel, CH**

72 Inventor/es:

BEVERLY, MICHAEL

74 Agente/Representante:

CARVAJAL Y URQUIJO, Isabel

ES 2 763 822 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

Análisis exento de marcas de la eficacia de la adición de caperuzas al arn utilizando rnasa h, sondas y cromatografía líquida/espectrometría de masas

Campo de la invención

La invención se refiere en general a formas modificadas de ARN que no se utilizan en tecnología recombinante y a procesos de medición que involucran enzimas, con la liberación de un marcador enlazado.

Antecedentes de la invención

El interés cada vez mayor en el desarrollo de ARNm con fines terapéuticos requerirá una caracterización exhaustiva del ARNm, incluida la caperuza 5'. Thess A *et al.*, *Mol. Ther.* 23(9): 1456-1464 (2015); Weissman D *et al.*, *Mol. Ther.* 23(9): 1416-1417 (2015). Tal desarrollo necesitará métodos que sean robustos, precisos y capaces de procesar muchas muestras rápidamente.

Los métodos previos para caracterizar la caperuza 5' han utilizado la digestión enzimática del ARNm para obtener mononucleótidos, seguida de la electroforesis en gel de poliacrilamida (PAGE) o la cromatografía líquida de alto rendimiento (HPLC) de intercambio aniónico de los productos de escisión, junto con la detección de radiomarcas de P³² incorporado como el fosfato terminal 5'. La escisión del ARNm con RNasa T2 permitió una determinación de la eficacia de la adición de la caperuza (*capping*). El uso de la digestión con RNasa T2 seguida de la pirofosfatasa ácida del tabaco ha proporcionado información sobre la orientación de la caperuza. Pasquinelli AE *et al.*, *RNA* 1(9): 957-967 (1995); Peng ZH *et al.*, *Org. Lett.* 4(2):161-164 (2002); Stepinski J *et al.*, *RNA* 7 (10):1486-1495 (2001); Jemielity J *et al.*, *RNA* 9(9): 1108-1122 (2003); Jemielity J *et al.*, *Nucleosides Nucleotides Nucleic Acids* 22(5-8): 691-694 (2003); Grudzien E *et al.*, *RNA* 10(9): 1479-1487 (2004). Esta combinación de enzimas proporcionó pGp (si la caperuza está en la orientación directa o correcta) o pm7Gp (si está en la orientación inversa). A continuación, estas dos especies se separaron por HPLC y se cuantificaron midiendo el P³².

La eficacia de la reacción de adición de la caperuza también se ha determinado utilizando sondas de RNasa H específicas de la secuencia que escinden pequeñas secciones del extremo 5' del ARNm y resolviendo el trifosfato sin caperuza procedente del ARNm con caperuza de m7GppG utilizando PAGE con detección radiomarcada. Peng ZH *et al.*, *Org. Lett.* 4(2):161-164 (2002); Tcherepanova IY *et al.*, *BMC Mol. Biol.* 9: 90 (2008).

Con respecto a otros métodos, la solicitud de PCT publicada WO 2012/135805 (Moderna Therapeutics) divulgó la detección de ARNm con caperuza mediante LC-MS, correspondiente a la eficacia de la reacción de adición de la caperuza. La solicitud de PCT publicada WO 2014/152659 (Shire Human Genetic Therapies) divulga y reivindica un método para cuantificar la eficacia de la adición de la caperuza en ARNm mediante cromatografía. El documento WO2015101416 (Curevac GmbH) divulga y reivindica un método para analizar ARN que tiene un sitio de escisión para una molécula de ácido nucleico catalítica, en el que la cantidad de ARN con caperuza se puede medir mediante espectroscopía de masas cuantitativa. Lapham y Crothers, *RNA* 2(3): 289-296 (1996) divulgaron un método para escindir secuencias cortas del extremo 5' del ARNm en una ubicación específica.

Si bien los métodos de la técnica anterior proporcionaron información sobre la estructura y orientación de la caperuza, estos están limitados al ARNm radiomarcado. Estos métodos también se limitan a la identificación de diferentes especies de caperuza en función de la resolución cromatográfica y el tiempo de retención de los patrones. A medida que aumenta la investigación sobre el uso de ARNm con fines terapéuticos, existe la necesidad en el campo biotecnológico de caracterizar el ARNm sintetizado transcrito *in vitro* (IVT). Sahin U *et al.*, *Nat. Rev. Drug Discov.* 13(10):759-780 (2014); Geall AJ *et al.*, *Semin. Immunol.* 25(2):152-159 (2013).

Compendio de la invención

En una primera realización, la invención proporciona un método exento de radiomarcas para identificar una caperuza del extremo 5' en un ácido ribonucleico (ARN) diana, que comprende los pasos de: (a) hibridar una sonda etiquetada no radiomarcada con el ARN diana, donde la secuencia de nucleótidos de la sonda etiquetada no radiomarcada es complementaria respecto al extremo 5' del ARN diana, con lo que se forma así un polinucleótido dúplex; (b) tratar el polinucleótido dúplex con RNasa H, con lo que se escinde así el extremo 5' del ARN diana y se forma un polinucleótido dúplex que contiene el extremo 5' del ARN diana; (c) aislar el polinucleótido dúplex, utilizando un sustrato de superficie recubierta que está recubierto con un reactivo que se une a la sonda etiquetada no radiomarcada; (d) eliminar el polinucleótido dúplex del sustrato de superficie recubierta; (e) desnaturalizar el polinucleótido dúplex, con lo que se produce así un fragmento monocatenario del extremo 5' del ARN diana y la sonda etiquetada no radiomarcada; (f) aislar el fragmento monocatenario del extremo 5' del ARN diana; (g) analizar el fragmento monocatenario del extremo 5' del ARN diana mediante cromatografía líquida/espectrometría de masas (LC-MS); y (h) identificar la caperuza del extremo 5'. En una realización, el ARN diana es un ARN mensajero (ARNm). En una segunda realización, el ARN diana es un ARN mensajero eucariota.

El método de la invención proporciona a los expertos en el campo biotecnológico una forma de identificar la caperuza 5' sin la necesidad de radiomarcas, utilizando cromatografía líquida acoplada a espectrometría de masas por electronebulización (LC-MS) y detectando sus diferencias en masa y tiempo de retención. El uso de la espectrometría de masas como detector permite la identificación basada en una propiedad intrínseca de la especie que se puede medir con precisión y con alta resolución. La identificación inequívoca de los productos de escisión de la caperuza 5' mediante espectrometría de masas también permite la identificación de una variedad de especies sin caperuza.

En una tercera realización, el sustrato de superficie recubierta son microesferas magnéticas. En una cuarta realización, el reactivo en el sustrato de superficie recubierta que se une a la sonda etiquetada no radiomarcada es (a) estreptavidina, donde la etiqueta no radiomarcada es biotina; (b) avidina, donde la etiqueta no radiomarcada es biotina; (c) un anticuerpo anti-biotina, donde la etiqueta no radiomarcada es biotina; (d) un anticuerpo anti-digoxigenina, donde la etiqueta no radiomarcada es digoxigenina; o (e) un anticuerpo anti-péptido, donde la etiqueta no radiomarcada es un péptido. En una quinta realización, el ARN diana se sintetiza *in vitro*.

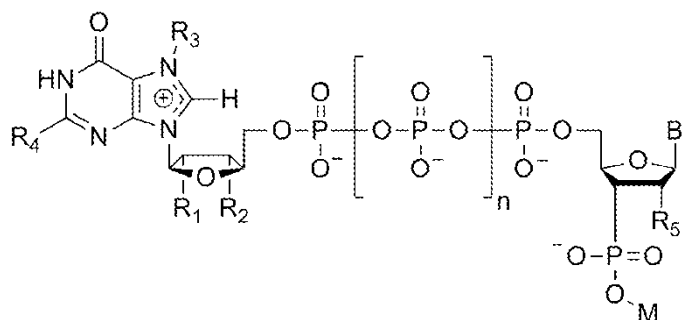
En una sexta realización, el método de identificación comprende además determinar información cuantitativa y cualitativa sobre la caperuza 5' a partir de la masa única del fragmento monocatenario del extremo 5' del ARN diana. El método de la invención puede determinar cualitativa y cuantitativamente la adición de la caperuza en ARNm, la eficacia de la adición de la caperuza 5' y la identidad de la caperuza 5' en muestras de ARNm.

En una séptima realización, también se proporciona un método exento de radiomarcas para determinar la orientación del extremo 5' en ARN diana, que comprende los pasos de: (a) hibridar una sonda etiquetada no radiomarcada con el ARN diana para formar un polinucleótido parcialmente dúplex, donde la secuencia de nucleótidos de la sonda etiquetada no radiomarcada es complementaria respecto al extremo 5' del ARN diana; (b) tratar el polinucleótido dúplex con RNasa H, para escindir el extremo 5' del ARN y formar un polinucleótido dúplex; (c) aislar el polinucleótido dúplex que contiene el extremo 5' del ARN diana, utilizando microesferas magnéticas recubiertas con un reactivo que se une a la sonda etiquetada no radiomarcada; (d) tratar el polinucleótido dúplex con 5' ARN pirofosfohidrolasa (RppH); (e) eliminar el polinucleótido dúplex de las microesferas magnéticas; (f) desnaturalizar el polinucleótido dúplex, para producir un fragmento monocatenario del extremo 5' del ARN diana y la sonda etiquetada no radiomarcada; (g) aislar el fragmento monocatenario del extremo 5' del ARN diana; (h) analizar el fragmento monocatenario del extremo 5' del ARN mediante cromatografía líquida/espectrometría de masas (LC-MS); e (i) detectar la diferencia de masa en el producto para determinar la orientación de la caperuza 5'.

El método de la invención se utilizó para analizar de forma sencilla la orientación de la caperuza. La combinación de la sonda de escisión de RNasa H con una 5' ARN pirofosfohidrolasa (RppH) permitió la determinación de la orientación de la caperuza, lo que puede ser útil al trabajar con métodos que introducen la caperuza de forma cotranscripcional en el ARNm.

En una octava realización, el ARN diana es un ARN mensajero (ARNm). En una novena realización, el ARN diana es un ARN mensajero eucariota.

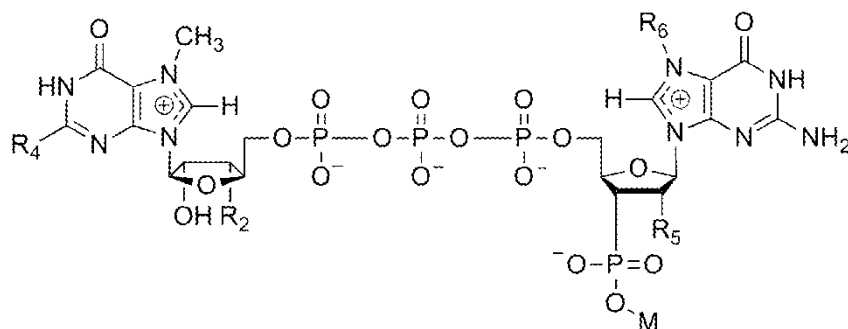
En una décima realización, la caperuza del extremo 5' tiene una estructura de fórmula I:



(I), donde B es una base nitrogenada; R₁ se selecciona entre un halógeno, OH y OCH₃; R₂ se selecciona entre H, OH y OCH₃; R₃ es CH₃, CH₂CH₃, CH₂CH₂CH₃ o no existente; R₄ es NH₂;

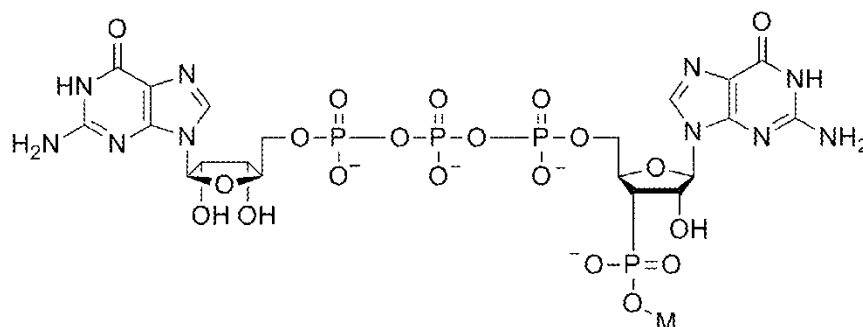
R₅ se selecciona entre OH, OCH₃ o un halógeno; n es 1, 2 o 3, o un fosforotioato; y M es un nucleótido del ARNm. En una undécima realización más específica, la base nitrogenada es guanina.

En una duodécima realización, la caperuza del extremo 5' es una caperuza m⁷G con una estructura de fórmula II:



(II) donde R_2 es H o CH_3 ; R_4 es NH_2 ; R_6 es OH o OCH_3 ; y M es un nucleótido del ARN; o (en una decimotercera realización) una caperuza no metilada con una estructura de fórmula III

5



(III), donde M es una base nitrogenada del ARN.

10 En una decimocuarta realización, el ARN diana se sintetiza *in vitro*. En una decimoquinta realización más específica, el ARN sintetizado *in vitro* contiene uno o más nucleótidos seleccionados entre Ψ (pseudouridina); m^5C (5-metilcitidina); m^5U (5-metiluridina); m^6A (N^6 -metiladenosina); s^2U (2-tiouridina); Um (2'-O-metil-U; 2'-O-metiluridina); m^1A (1-metiladenosina); m^2A (2-metiladenosina); Am (2'-O-metiladenosina); $ms^2 m^6A$ (2-metiltio- N^6 -metiladenosina); i^6A (N^6 -isopenteniladenosina); $ms^2 i^6A$ (2-metiltio- N^6 -isopenteniladenosina); io^6A (N^6 -(cis-hidroxiisopentenil)adenosina); $ms^2 i^6A$ (2-metiltio- N^6 -(cis-hidroxiisopentenil)adenosina); g^6A (N^6 -glicinilcarbamoiladenosina); t^6A (N^6 -treonilcarbamoiladenosina); $ms^2 t^6A$ (2-metiltio- N^6 -treonilcarbamoiladenosina); $m^6 t^6A$ (N^6 -metil- N^6 -treonilcarbamoiladenosina); hn^6A (N^6 -hidroxinorvalilcarbamoiladenosina); $ms^2 hn^6A$ (2-metiltio- N^6 -hidroxinorvalilcarbamoiladenosina); $Ar(p)$ (2'-O-ribosiladenosina (fosfato)); I (inosina); m^1I (1-metilinosina); m^1Im (1,2'-O-dimetilinosina); m^3C (3-metilcitidina); Cm (2'-O-metilcitidina); s^2C (2-tiocitidina); ac^4C (N^4 -acetilcitidina); f^5C (5-formilcitidina); $m^5 Cm$ (5,2'-O-dimetilcitidina); ac^4Cm (N^4 -acetil-2'-O-metilcitidina); k^2C (lisidina); m^1G (1-metilguanosa); m^2G (N^2 -metilguanosa); m^7G (7-metilguanosa); Gm (2'-O-metilguanosa); $m^2 _2G$ (N^2, N^2 -dimetilguanosa); m^2Gm ($N^2, 2'$ -O-dimetilguanosa); $m^2 _2Gm$ ($N^2, N^2, 2'$ -O-trimetilguanosa); $Gr(p)$ (2'-O-ribosilguanosa (fosfato)); yW (wibutosina); o_2yW (peroxiwibutosina); $OHyW$ (hidroxiwibutosina); $OHyW^*$ (hidroxiwibutosina submodificada); imG (wiosina); $mimG$ (metilwiosina); Q (queuosina); oQ (epoxiqueuosina); $galQ$ (galactosilqueuosina); $manQ$ (manosilqueuosina); $preQ_0$ (7-ciano-7-desazaguanosina); $preQ_1$ (7-aminometil-7-desazaguanosina); G^+ (arcaeosina); D (dihidrouridina); m^5Um (5,2'-O-dimetiluridina); s^4U (4-tiouridina); m^5s^2U (5-metil-2-tiouridina); s^2Um (2-tio-2'-O-metiluridina); acp^3U (3-(3-amino-3-carboxipropil)uridina); ho^5U (5-hidroxiuridina); mo^5U (5-metoxiuridina); cmo^5U (ácido uridino-5-oxiacético); $mcmo^5U$ (éster metílico del ácido uridino-5-oxiacético); chm^5U (5-(carboxihidroximetil)uridina); $mchm^5U$ (éster metílico de la 5-(carboxihidroximetil)uridina); mcm^5U (5-metoxicarbonilmetiluridina); mcm^5Um (5-metoxicarbonilmetil-2'-O-metiluridina); mcm^5s^2U (5-metoxicarbonilmetil-2-tiouridina); nm^5s^2U (5-aminometil-2-tiouridina); mnm^5U (5-metilaminometiluridina); mnm^5s^2U (5-metilaminometil-2-tiouridina); mnm^5se^2U (5-metilaminometil-2-selenouridina); ncm^5U (5-carbamoilmetiluridina); ncm^5Um (5-carbamoilmetil-2'-O-metiluridina); $cmnm^5U$ (5-carboximetilaminometiluridina); $cmnm^5Um$ (5-carboximetilaminometil-2'-O-metiluridina); $cmnm^5s^2U$ (5-carboximetilaminometil-2-tiouridina); $m^6 _2A$ (N^6, N^6 -dimetiladenosina); Im (2'-O-metilinosina); m^4C (N^4 -metilcitidina); $m^4 Cm$ ($N^4, 2'$ -O-dimetilcitidina); hm^5C (5-hidroximetilcitidina); m^3U (3-metiluridina); cm^5U (5-carboximetiluridina); m^6Am ($N^6, 2'$ -O-dimetiladenosina); $m^6 _2Am$ (N^6, N^6, O -2'-trimetiladenosina); $m^2 _7G$ ($N^2, 7$ -dimetilguanosa); $m^{2,2,7}G$ ($N^2, N^2, 7$ -trimetilguanosa); m^3Um (3,2'-O-dimetiluridina); m^5D (5-metildihidrouridina); f^5Cm (5-formil-2'-O-metilcitidina); m^1Gm (1,2'-O-dimetilguanosa); m^1Am (1,2'-O-dimetiladenosina); tm^5U (5-aurinometiluridina); tm^5s^2U (5-aurinometil-2-tiouridina); $imG-14$ (4-desmetilwiosina); $imG2$ (isowiosina); y ac^6A (N^6 -acetiladenosina). En una decimosexta realización específica, el ARN sintetizado *in vitro* contiene uno o más nucleótidos seleccionados entre Ψ (pseudouridina), m^5C (5-metilcitidina) o ambos.

En una decimoséptima realización, el sustrato de superficie recubierta son microesferas magnéticas. En otra realización, el reactivo en el sustrato de superficie recubierta que se une a la sonda etiquetada no radiomarcada es: (a) estreptavidina, donde la etiqueta no radiomarcada es biotina; (b) avidina, donde la etiqueta no radiomarcada es biotina; (c) un anticuerpo anti-biotina, donde la etiqueta no radiomarcada es biotina; (d) un anticuerpo anti-digoxigenina, donde la etiqueta no radiomarcada es digoxigenina; o (e) un anticuerpo anti-péptido, donde la etiqueta no radiomarcada es un péptido.

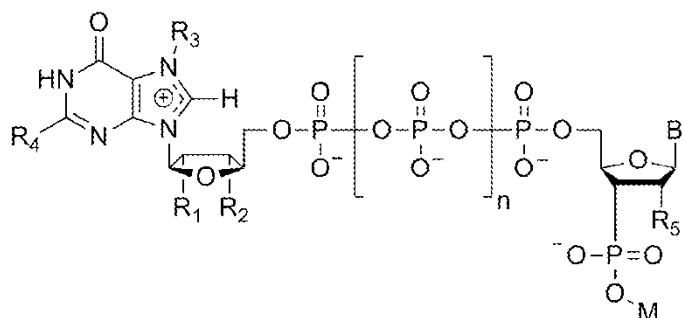
En una decimoctava realización, la invención proporciona además un método exento de radiomarcas para determinar la eficacia de la adición de la caperuza, que comprende los pasos de: (a) proporcionar una muestra de ARN que comprende ARN con caperuza y ARN sin caperuza; (b) hibridar una sonda etiquetada no radiomarcada con el ARN diana para formar un polinucleótido dúplex, donde la secuencia de nucleótidos de la sonda etiquetada no radiomarcada es complementaria respecto al extremo 5' del ARN diana; (c) tratar el polinucleótido dúplex con RNasa H, para escindir el extremo 5' del ARN y formar un polinucleótido dúplex; (d) aislar el polinucleótido dúplex que contiene el extremo 5' del ARN diana, utilizando microesferas magnéticas recubiertas con un reactivo que se une a la sonda etiquetada no radiomarcada; (e) eliminar el polinucleótido dúplex de las microesferas magnéticas; (f) desnaturalizar el polinucleótido dúplex, para producir un fragmento monocatenario del extremo 5' del ARN diana y la sonda etiquetada no radiomarcada; (g) aislar el fragmento monocatenario del extremo 5' del ARN diana; (h) analizar el fragmento monocatenario del extremo 5' del ARN mediante cromatografía líquida/espectrometría de masas (LC-MS); e (i) determinar la cantidad relativa de los fragmentos monocatenarios con caperuza y sin caperuza, con lo que se cuantifica así la eficacia de la adición de la caperuza en el ARNm.

En una decimonovena realización, el método de la invención proporciona una forma de diferenciar el ARN con caperuza del ARN sin caperuza sin la necesidad de radiomarcas, utilizando cromatografía líquida acoplada a espectrometría de masas por electronebulización (LC-MS) y detectando sus diferencias en masa y tiempo de retención. En una vigésima realización, la estrategia basada en LC-MS descrita anteriormente está automatizada para que se lleve a cabo en un formato de alto rendimiento con el fin de analizar números mayores de muestras.

El método de la invención se puede aplicar favorablemente para determinar la eficacia de la adición de la caperuza en ARNm sintetizados por transcripción *in vitro* (IVT). En una vigésima primera realización, el método de la invención se utiliza para analizar ARNm con una longitud de 2,2 K y 9 K nucleótidos que contenía los nucleótidos modificados pseudouridina (Ψ) y 5-metilcitidina (m^5C). El resultado exitoso del uso de esta realización demostró la utilidad del método de la invención para estudiar la adición de la caperuza en ARNm y la funcionalidad en un intervalo de longitudes de ARNm y con tipos de modificaciones habituales. La respuesta del ensayo para el ARNm sin caperuza fue lineal, lo que permitió la determinación cuantitativa de las especies sin caperuza. El ARNm con trifosfato sin caperuza de 2,2 K bases se pudo detectar en presencia de 100 picomoles de ARNm con caperuza a lo largo del intervalo evaluado de un 0,5 a un 25%. El análisis de varios lotes de ARNm con caperuza reveló eficacias de la adición de la caperuza comprendidas en el intervalo de 88-98%, lo que parece ser coherente con el sistema enzimático de adición de la caperuza en *Vaccinia*. Se observó que la adición de la caperuza en ARNm utilizando el sistema ARCA tenía una eficacia de adición de la caperuza de un 70%, que está de acuerdo con la descrita por Grudzien E *et al.*, *RNA* 10(9): 1479-1487 (2004).

En una vigésima segunda realización, el método de la invención se utiliza para solucionar problemas del proceso de adición de la caperuza IVT. En el uso de esta realización, la presencia de difosfato en el material con caperuza indicó que la trifosfatasa inicial que elimina un fosfato del ARNm con trifosfato sin caperuza estaba funcionando bien, pero que la guaniltransferasa que añade GTP al difosfato necesitaba mejorar. De manera similar, la presencia de ARNm con caperuza G no metilada indicó que la N7-metiltransferasa necesitaba un ajuste.

En una vigésima tercera realización, la caperuza del extremo 5' tiene una estructura de fórmula I:

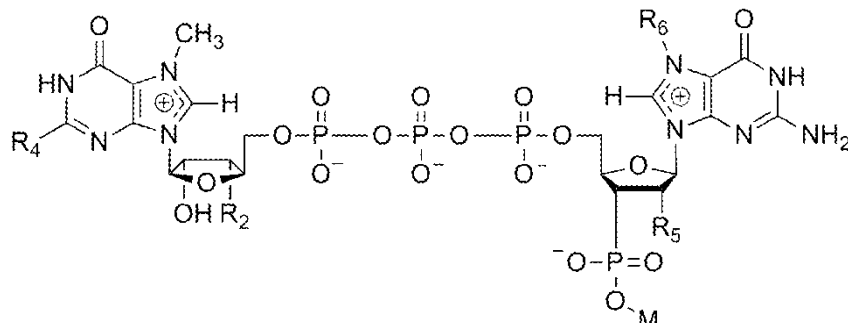


(I), donde B es una base nitrogenada; R_1 se selecciona entre un halógeno, OH y OCH_3 ; R_2 se selecciona entre H, OH y OCH_3 ; R_3 es CH_3 , CH_2CH_3 , $CH_2CH_2CH_3$ o no existente; R_4 es NH_2 ; R_5 se selecciona entre OH, OCH_3 o un halógeno;

n es 1, 2 o 3, o un fosforotioato; y M es un nucleótido del ARNm. En una vigésima cuarta realización más específica, la base nitrogenada es guanina.

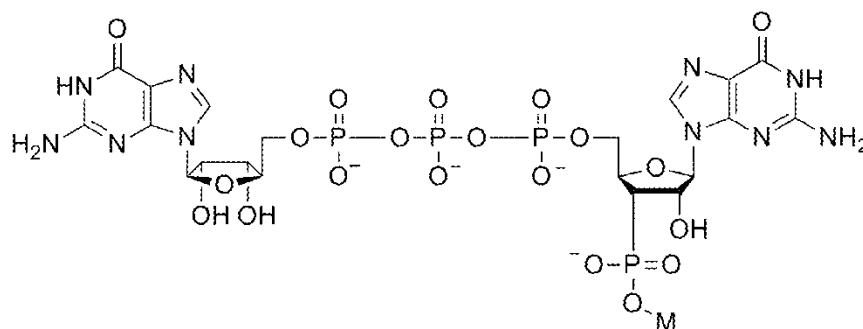
En una vigésima quinta realización, la caperuza es una caperuza m⁷G con una estructura de fórmula II:

5



(II) donde R₂ es H o CH₃; R₄ es NH₂; R₆ es OH o OCH₃; y M es un nucleótido del ARN; o (en una vigésima sexta realización) una caperuza no metilada con una estructura de fórmula III

10



(III), donde M es una base nitrogenada del ARN.

15

En una vigésima séptima realización, la muestra de ARN es una mezcla de reacción de transcripción *in vitro* (IVT). En una vigésima octava realización más específica, la reacción de IVT es un preparado de ARNm con caperuza de *Vaccinia*.

20

En una vigésima novena realización, el ARN sintetizado *in vitro* contiene uno o más nucleósidos seleccionados entre Ψ (pseudouridina); m⁵C (5-metilcitidina); m⁵U (5-metiluridina); m⁶A (N⁶-metiladenosina); s²U (2-tiouridina); Um (2'-O-metil-U; 2'-O-metiluridina); m¹A (1-metiladenosina); m²A (2-metiladenosina); Am (2'-O-metiladenosina); ms² m⁶A (2-metil-N⁶-metiladenosina); i⁶A (N⁶-isopenteniladenosina); ms²i⁶A (2-metil-N⁶-isopenteniladenosina); io⁶A (N⁶-(cis-hidroxiisopentenil)adenosina); ms²io⁶A (2-metil-N⁶-(cis-hidroxiisopentenil)adenosina); g⁶A (N⁶-glicinilcarbamoiladenosina); t⁶A (N⁶-treonilcarbamoiladenosina); ms²t⁶A (2-metil-N⁶-treonilcarbamoiladenosina); m⁶t⁶A (N⁶-metil-N⁶-treonilcarbamoiladenosina); hn⁶A (N⁶-hidroxinorvalilcarbamoiladenosina); ms²hn⁶A (2-metil-N⁶-hidroxinorvalilcarbamoiladenosina); Ar(p) (2'-O-ribosiladenosina (fosfato)); I (inosina); m¹I (1-metilinosina); m¹Im (1,2'-O-dimetilinosina); m³C (3-metilcitidina); Cm (2'-O-metilcitidina); s²C (2-tiocitidina); ac⁴C (N⁴-acetilcitidina); f⁶C (5-formilcitidina); m⁵Cm (5,2'-O-dimetilcitidina); ac⁴Cm (N⁴-acetil-2'-O-metilcitidina); k²C (lisidina); m¹G (1-metilguanina); m²G (N²-metilguanina); m⁷G (7-metilguanina); Gm (2'-O-metilguanina); m² 2G (N²,N²-dimetilguanina); m²Gm (N²,2'-O-dimetilguanina); m² 2Gm (N²,N²,2'-O-trimetilguanina); Gr(p) (2'-O-ribosilguanina (fosfato)); yW (wibutosina); o₂yW (peroxiwibutosina); OHyW (hidroxiwibutosina); OHyW* (hidroxiwibutosina submodificada); imG (wiosina); mimG (metilwiosina); Q (queuosina); oQ (epoxiqueuosina); galQ (galactosilqueuosina); manQ (manosilqueuosina); preQ₀ (7-ciano-7-desazaguanosina); preQ₁ (7-aminometil-7-desazaguanosina); G⁺ (arcaeosina); D (dihidrouridina); m⁵Um (5,2'-O-dimetiluridina); s⁴U (4-tiouridina); m⁵s²U (5-metil-2-tiouridina); s²Um (2-tio-2'-O-metiluridina); acp³U (3-(3-amino-3-carboxipropil)uridina); ho⁵U (5-hidroxiuridina); mo⁵U (5-metoxiuridina); cmo⁵U (ácido uridino-5-oxiacético); mcmo⁵U (éster metílico del ácido uridino-5-oxiacético); chm⁵U (5-(carboxihidroximetil)uridina); mchm⁵U (éster metílico de la 5-(carboxihidroximetil)uridina); mcm⁵U (5-metoxicarbonilmetiluridina); mcm⁵Um (5-metoxicarbonilmetil-2'-O-metiluridina); mcm⁵s²U (5-metoxicarbonilmetil-2-tiouridina); nm⁵s²U (5-aminometil-2-tiouridina); mnm⁵U (5-metilaminometiluridina); mnm⁵s²U (5-metilaminometil-2-tiouridina); mnm⁵se²U (5-metilaminometil-2-selenouridina); ncm⁵U (5-carbamoilmetiluridina); ncm⁵Um (5-carbamoilmetil-2'-O-metiluridina); cmnm⁵U (5-carboximetilaminometiluridina); cmnm⁵Um (5-carboximetilaminometil-2'-O-metiluridina); cmnm⁵s²U (5-carboximetilaminometil-2-tiouridina); m⁶ 2A (N⁶,N⁶-dimetiladenosina); Im (2'-O-metilinosina); m⁴C (N⁴-metilcitidina); m⁴ Cm (N⁴,2'-O-dimetilcitidina); hm⁵C (5-hidroximetilcitidina); m³U (3-metiluridina); cm⁵U (5-carboximetiluridina); m⁶Am (N⁶,2'-O-dimetiladenosina); m⁶ 2Am (N⁶,N⁶,O-2'-trimetiladenosina); m^{2,7}G (N²,7-dimetilguanina); m^{2,2,7}G (N²,N²,7-

25

30

35

40

trimetilguanosina); m³Um (3,2'-O-dimetiluridina); m⁵D (5-metildihidrouridina); f⁵Cm (5-formil-2'-O-metilcitidina); m¹Gm (1,2'-O-dimetilguanosina); m¹Am (1,2'-O-dimetiladenosina); τ m⁵U (5-taurinometiluridina); τ m⁵s²U (5-taurinometil-2-tiouridina); imG-14 (4-desmetilwiosina); imG2 (isowiosina); y ac⁶A (N⁶-acetiladenosina). En una trigésima realización, el ARN sintetizado *in vitro* contiene uno o más nucleótidos seleccionados entre Ψ (pseudouridina), m⁵C (5-metilcitidina) o ambos.

En una trigésima primera realización, el sustrato de superficie recubierta comprende microesferas magnéticas.

En una trigésima segunda realización, el reactivo en el sustrato de superficie recubierta que se une a la sonda etiquetada no radiomarcada es: (a) estreptavidina, donde la etiqueta no radiomarcada es biotina; (b) avidina, donde la etiqueta no radiomarcada es biotina; (c) un anticuerpo anti-biotina, donde la etiqueta no radiomarcada es biotina; (d) un anticuerpo anti-digoxigenina, donde la etiqueta no radiomarcada es digoxigenina; y (e) un anticuerpo anti-péptido, donde la etiqueta no radiomarcada es un péptido.

En una trigésima tercera realización, el ARNm con trifosfato sin caperuza se puede detectar en un intervalo evaluado de un 0,1 a un 90% con una respuesta lineal. En una trigésima cuarta realización, el ARNm con trifosfato sin caperuza se puede detectar en un intervalo evaluado de un 0,5 a un 25% con una respuesta lineal.

En una trigésima quinta realización, la invención proporciona un método exento de radiomarcas para detectar una impureza de la reacción de adición de la caperuza en un preparado de ARN, que comprende los pasos de: (a) hibridar una sonda etiquetada no radiomarcada con el ARN diana, donde la sonda etiquetada no radiomarcada es complementaria respecto al extremo 5' del ADN diana; (b) tratar el ARN hibridado con RNasa H para escindir el extremo 5' del ARN; (c) aislar la secuencia del extremo 5' escindida, utilizando microesferas magnéticas recubiertas con un reactivo que se une a la sonda etiquetada no radiomarcada; (e) analizar la secuencia del extremo 5' escindida mediante LC-MS, y (f) detectar una impureza de la reacción de adición de la caperuza en un preparado de ARN. En una trigésima sexta realización, el método incluye además el paso de: (f) identificar una impureza de la reacción de adición de la caperuza en un preparado de ARN.

El método de la invención mostró buena sensibilidad para detectar impurezas de la reacción de adición de la caperuza. El ARNm con trifosfato sin caperuza con el que se enriquecieron 100 picomoles de ARNm con caperuza pudo ser detectado a lo largo del intervalo evaluado de un 0,5 a un 25% con una respuesta lineal. Se determinó que la eficacia de la adición de la caperuza para varios preparados de ARNm con caperuza de *Vaccinia* estuvo entre un 88 y un 98% dependiendo del tipo de modificación y la longitud del ARNm.

En una trigésima séptima realización, el ARN diana se sintetiza *in vitro*. En una realización más específica, el ARN sintetizado *in vitro* contiene un nucleótido seleccionado del grupo constituido por Ψ (pseudouridina); m⁵C (5-metilcitidina); m⁵U (5-metiluridina); m⁶A (N⁶-metiladenosina); s²U (2-tiouridina); Um (2'-O-metil-U; 2'-O-metiluridina); m¹A (1-metiladenosina); m²A (2-metiladenosina); Am (2'-O-metiladenosina); ms² m⁶A (2-metiltio-N⁶-metiladenosina); i⁶A (N⁶-isopenteniladenosina); ms²i⁶A (2-metiltio-N⁶-isopenteniladenosina); io⁶A (N⁶-(cis-hidroxiisopentenil)adenosina); ms²i⁶A (2-metiltio-N⁶-(cis-hidroxiisopentenil)adenosina); g⁶A (N⁶-glicinilcarbamoiladenosina); t⁶A (N⁶-treonilcarbamoiladenosina); ms²t⁶A (2-metiltio-N⁶-treonilcarbamoiladenosina); m⁶t⁶A (N⁶-metil-N⁶-treonilcarbamoiladenosina); hn⁶A (N⁶-hidroxinorvalilcarbamoiladenosina); ms²hn⁶A (2-metiltio-N⁶-hidroxinorvalilcarbamoiladenosina); Ar(p) (2'-O-ribosiladenosina (fosfato)); I (inosina); m¹I (1-metilinosina); m¹Im (1,2'-O-dimetilinosina); m³C (3-metilcitidina); Cm (2'-O-metilcitidina); s²C (2-tiocitidina); ac⁴C (N⁴-acetilcitidina); f⁵C (5-formilcitidina); m⁵ Cm (5,2'-O-dimetilcitidina); ac⁴Cm (N⁴-acetil-2'-O-metilcitidina); k²C (lisidina); m¹G (1-metilguanosina); m²G (N²-metilguanosina); m⁷G (7-metilguanosina); Gm (2'-O-metilguanosina); m² ₂G (N²,N²-dimetilguanosina); m²Gm (N²,2'-O-dimetilguanosina); m²₂Gm (N²,N²,2'-O-trimetilguanosina); Gr(p) (2'-O-ribosilguanosina (fosfato)); yW (wibutosina); o₂yW (peroxiwibutosina); OHyW (hidroxiwibutosina); OHyW* (hidroxiwibutosina submodificada); imG (wiosina); mimG (metilwiosina); Q (queuosina); oQ (epoxiqueuosina); galQ (galactosilqueuosina); manQ (manosilqueuosina); preQ₀ (7-ciano-7-desazaguanosina); preQ₁ (7-aminometil-7-desazaguanosina); G⁺ (arcaeosina); D (dihidrouridina); m⁵Um (5,2'-O-dimetiluridina); s⁴U (4-tiouridina); m⁵s²U (5-metil-2-tiouridina); s²Um (2-tio-2'-O-metiluridina); acp³U (3-(3-amino-3-carboxipropil)uridina); ho⁵U (5-hidroxiuridina); mo⁵U (5-metoxiuridina); cmo⁵U (ácido uridino-5-oxiacético); mcmo⁵U (éster metílico del ácido uridino-5-oxiacético); chm⁵U (5-(carboxihidroximetil)uridina); mchm⁵U (éster metílico de la 5-(carboxihidroximetil)uridina); mcm⁵U (5-metoxicarbonilmetiluridina); mcm⁵Um (5-metoxicarbonilmetil-2'-O-metiluridina); mcm⁵s²U (5-metoxicarbonilmetil-2-tiouridina); nm⁵s²U (5-aminometil-2-tiouridina); mnm⁵U (5-metilaminometiluridina); mnm⁵s²U (5-metilaminometil-2-tiouridina); mnm⁵se²U (5-metilaminometil-2-selenouridina); ncm⁵U (5-carbamoilmetiluridina); ncm⁵Um (5-carbamoilmetil-2'-O-metiluridina); cmnm⁵U (5-carboximetilaminometiluridina); cmnm⁵Um (5-carboximetilaminometil-2'-O-metiluridina); cmnm⁵s²U (5-carboximetilaminometil-2-tiouridina); m⁶ ₂A (N⁶,N⁶-dimetiladenosina); Im (2'-O-metilinosina); m⁴C (N⁴-metilcitidina); m⁴ Cm (N⁴,2'-O-dimetilcitidina); hm⁵C (5-hidroximetilcitidina); m³U (3-metiluridina); cm⁵U (5-carboximetiluridina); m⁶Am (N⁶,2'-O-dimetiladenosina); m⁶ ₂Am (N⁶,N⁶,O-2'-trimetiladenosina); m^{2,7}G (N²,7-dimetilguanosina); m^{2,2,7}G (N²,N²,7-trimetilguanosina); m³Um (3,2'-O-dimetiluridina); m⁵D (5-metildihidrouridina); f⁵Cm (5-formil-2'-O-metilcitidina); m¹Gm (1,2'-O-dimetilguanosina); m¹Am (1,2'-O-dimetiladenosina); τ m⁵U (5-taurinometiluridina); τ m⁵s²U (5-taurinometil-2-tiouridina); imG-14 (4-desmetilwiosina); imG2 (isowiosina); y ac⁶A (N⁶-acetiladenosina). En una trigésima octava realización específica, el ARN sintetizado *in vitro* comprende Ψ

(pseudouridina), m⁵C (5-metilcitidina) o ambos nucleósidos. En una trigésima novena realización, el sustrato del sustrato de superficie recubierta son microesferas magnéticas.

5 En una cuadragésima realización, el reactivo en el sustrato de superficie recubierta que se une a la sonda etiquetada no radiomarcada es: (a) estreptavidina, donde la etiqueta no radiomarcada es biotina; (b) avidina, donde la etiqueta no radiomarcada es biotina; (c) un anticuerpo anti-biotina, donde la etiqueta no radiomarcada es biotina; (d) un anticuerpo anti-digoxigenina, donde la etiqueta no radiomarcada es digoxigenina; o (e) un anticuerpo anti-péptido, donde la etiqueta no radiomarcada es un péptido.

10 Breve descripción de los dibujos

La FIG. 1 muestra la vía enzimática involucrada en la adición de la caperuza en ARNm que da como resultado Cap0 y Cap1 (adoptada de Ghosh y Lima, *Wiley Interdiscip. Rev. RNA* 1(1): 152-172 (2010)).

15 La FIG. 2 es una representación esquemática del procedimiento para la escisión por parte de la RNasa H del ARNm y el aislamiento del fragmento escindido utilizando microesferas magnéticas. Las flechas indican los dos sitios de escisión observados por LC-MS.

20 La FIG. 3 muestra el flujo de trabajo del análisis de la orientación de la caperuza que muestra los productos de reacción procedentes del tratamiento con la enzima RppH añadidos a la secuencia de escisión de la RNasa H. La orientación inversa de la caperuza p7mG ... tiene una masa diferente a la del monofosfato.

Descripción detallada de la invención

25 Definiciones

Para que la invención pueda comprenderse más fácilmente, se definen en la presente algunos términos. A lo largo de la memoria descriptiva se pueden exponer definiciones adicionales.

30 Cuando «3'» se utiliza en una posición de un nucleótido, se refiere a una región o posición en un polinucleótido u oligonucleótido 3' (es decir, en dirección de 5' a 3') desde otra región o posición en el mismo polinucleótido u oligonucleótido. Las expresiones «extremo 3'» y «extremo terminal 3'», tal como se utilizan en la presente en referencia a una molécula de ácido nucleico, se refieren al extremo del ácido nucleico que contiene un grupo hidroxilo libre unido al carbono 3' del azúcar de tipo pentosa terminal.

35 Cuando «5'» se utiliza en una posición de un nucleótido, se refiere a una región o posición en un polinucleótido u oligonucleótido 5' (es decir, en dirección de 3' a 5') desde otra región o posición en el mismo polinucleótido u oligonucleótido. Las expresiones «extremo 5'» y «extremo terminal 5'», tal como se utilizan en la presente en referencia a una molécula de ácido nucleico, se refieren al extremo de la molécula de ácido nucleico que contiene un grupo hidroxilo o fosfato libre unido al carbono 5' del azúcar de tipo pentosa terminal. En algunas realizaciones de la invención, los cebadores oligonucleotídicos comprenden tramos de poliadenosina en sus extremos 5'.

40 La «afinidad», tal como se utiliza en la técnica, es una medida de la firmeza con la que se une un ligando particular a (por ejemplo, se asocia de forma no covalente con) y/o la velocidad o frecuencia con la que se disocia de, su pareja. 45 El experto sabrá que se han utilizado y se pueden utilizar varios métodos para determinar la afinidad. La afinidad es una medida de la unión específica.

Los términos «anelar», «hibridación», «hibridar» y sus equivalentes gramaticales significan la formación de complejos (también denominados «dúplex» o «híbridos») entre secuencias de nucleótidos que son suficientemente complementarias para formar complejos mediante el apareamiento de bases de Watson-Crick o el apareamiento de bases no canónico. Las secuencias que se anelan o hibridan no necesitan tener una complementariedad perfecta para proporcionar híbridos estables. Normalmente se pueden formar híbridos estables cuando menos de aproximadamente un 10% de las bases presentan apareamientos erróneos. Por consiguiente, el término «complementario», tal como se utiliza en la presente, se refiere a una molécula de ácido nucleico que forma un dúplex estable con su complemento 55 en condiciones particulares, por lo general donde existe una homología > 90% o superior (por ejemplo, una homología de aproximadamente un 95% o superior, aproximadamente un 98% o superior, o aproximadamente un 99% o superior). El experto sabe cómo estimar y ajustar la rigurosidad de las condiciones de hibridación de tal manera que las secuencias que tienen al menos un nivel deseado de complementariedad se hibridarán de manera estable, mientras que las que tienen una complementariedad más baja no lo harán. Para consultar ejemplos de condiciones de hibridación, remítase a Green & Sambrook, *Molecular Cloning: A Laboratory Manual*, cuarta edición (Cold Spring Harbor Press, Plainview, NY, 2012) y Ausubel, *Current Protocols in Molecular Biology* (John Wiley & Sons, Somerset NJ EE. UU., 2016). La complementariedad entre dos moléculas de ácido nucleico y la formación del dúplex es «completa», «total» o «perfecta» si todas las bases de los ácidos nucleicos coinciden, y de lo contrario es «parcial». 60

«ARCA» son análogos de caperuza anti-inversos, que se pueden adquirir del proveedor comercial TriLink, San Diego CA 92121 EE. UU. En las células, el ribosoma traduce ARNm en proteínas. El ARNm en las células eucariotas tiene una caperuza 5' [m⁷G(5')ppp(5')G]. Se puede incorporar una caperuza diferente en la transcripción del ARNm (por ejemplo, mediante una transcripción *in vitro*) incluyendo una mezcla de un análogo de la caperuza y GTP. De este modo, aproximadamente un 80% del ARNm sintetizado poseerá una caperuza 5', mientras que el 20% restante tendrá un trifosfato 5'. El primer análogo de caperuza que se introdujo fue m⁷G(5')ppp(5')G. El 50% de las veces, esta caperuza se inserta en la orientación correcta para mejorar la traducción. El otro 50% de las moléculas no son sustratos para una traducción eficaz, lo que reduce la actividad específica del transcrito de ARN. Más recientemente, se ha introducido el análogo de caperuza anti-inverso (ARCA) (Trilink, parte # N-7003) [3' O-Me-m⁷G(5')ppp(5')G] Stepinski J *et al.*, *RNA* 7(10):1486-95 (octubre de 2001). El ARCA solo se puede insertar en la orientación adecuada, lo que da como resultado unos ARNm con caperuza que se traducen con el doble de eficacia en comparación con los iniciados con m⁷G(5')ppp(5')G. La traducción eficaz del ARNm en una proteína también requiere una cola de poli(A). Esta se puede introducir incluyendo un tramo de poli(dT) al final de la plantilla de transcripción. Esto se consigue a menudo con un paso de PCR que utiliza un cebador que contiene el tramo de poli(dT)

Una «etiqueta de biotina» es una biotina unida químicamente a un compuesto de interés. Por lo tanto, una «sonda etiquetada con biotina» es una sonda polinucleotídica que tiene una etiqueta de biotina unida covalentemente a la sonda.

«Cap0» es una caperuza m⁷GpppG. Remítase a la FIG. 1 y la SEQ ID NO.: 1. Las caperuzas 5' terminales se pueden adquirir de proveedores comerciales, por ejemplo, TriLink BioTechnologies, Inc., San Diego CA EE. UU.

«Cap1» es una caperuza m⁷GpppmG, donde la Cap0 está metilada además en el OH de la posición 2' de la penúltima guanina. Remítase a la FIG. 1 y la SEQ ID NO.: 6. Las caperuzas 5' terminales se pueden adquirir de proveedores comerciales, por ejemplo, TriLink BioTechnologies, Inc., San Diego CA EE. UU.

La «cromatografía» es una técnica para separar mezclas. La mezcla se disuelve habitualmente en un fluido denominado «fase móvil», que la transporta a través de una estructura que contiene otro material denominado «fase estacionaria». Algunos ejemplos incluyen LC y HPLC.

Un «compuesto» es cualquier molécula de origen natural o no natural (es decir, sintética o recombinante) tal como una macromolécula biológica (por ejemplo, ácido nucleico, polipéptido o proteína), molécula orgánica o inorgánica, o un extracto obtenido a partir de materiales biológicos tales como bacterias, plantas, hongos, o células o tejidos animales (en particular de mamífero, incluido el ser humano). El compuesto puede ser una sola molécula o una mezcla o complejo de al menos dos moléculas.

El término «control» tiene el significado que se sobreentiende en la técnica de ser un patrón con el que se comparan los resultados. En algunas realizaciones, un control es una reacción o ensayo que se realiza simultáneamente con una reacción o ensayo de prueba para proporcionar un comparador. En un experimento, se aplica la «prueba» (es decir, la variable que se está evaluando). En el segundo experimento, el «control», la variable que se está evaluando no se aplica. En algunas realizaciones, un control es un control histórico (es decir, de una prueba o ensayo que se ha realizado previamente, o una cantidad o resultado que se conoce con anterioridad). En algunas realizaciones, un control es o comprende un registro impreso o guardado de otro modo. Un control puede ser un control positivo o un control negativo.

El término «homología» significa la relación general entre moléculas de ácido nucleico (por ejemplo, moléculas de ADN o moléculas de ARN) o entre moléculas de polipéptido. El término «homólogo» se refiere necesariamente a una comparación entre al menos dos secuencias (secuencias de polinucleótidos o polipéptidos).

«HPLC» es una cromatografía líquida de alta resolución, conocida previamente como cromatografía líquida de alta presión, una forma de cromatografía en columna que bombea una mezcla de muestra o analito en un disolvente (conocido como la fase móvil) a presión elevada a través de una columna con material de relleno cromatográfico (fase estacionaria). En el campo biotecnológico se conocen muchas técnicas de HPLC. Para obtener más información, remítase a *HPLC Primer* (2015) que se puede adquirir de Waters Corporation, Milford MA EE. UU.

El término «IVT» se refiere a la transcripción *in vitro* de ácido ribonucleico (ARN) a partir de un modelo de ácido desoxirribonucleico (ADN). En el campo biotecnológico se conocen muchas técnicas de IVT. Para obtener información, remítase a *The Basics: In Vitro Transcription* (2015), que se puede adquirir de Thermo Fisher Scientific Inc., Waltham MA EE. UU.

El término «LC» se refiere a cromatografía líquida, una técnica utilizada para separar una muestra en sus partes individuales. Esta separación se produce en función de las interacciones de la muestra con las fases móvil y estacionaria. En el campo biotecnológico se conocen muchas técnicas de LC. Para obtener más información, remítase a *Beginners Guide to UPLC* (2015) y *HPLC Primer* (2015), que se pueden adquirir ambos de Waters Corporation, Milford MA EE. UU.

El término «modificado» se refiere a una estructura o estado cambiado de una molécula de la invención. Un ARNm «modificado» contiene ribonucleósidos que engloban modificaciones respecto a los nucleósidos estándar de guanina (G), adenina (A), citidina (C) y uridina (U). Los nucleósidos no estándar pueden ser de origen natural o no natural. El ARN puede modificarse de muchas maneras, que incluyen química, estructural y funcionalmente, mediante métodos conocidos por los expertos en el campo biotecnológico. Tales modificaciones del ARN pueden incluir, por ejemplo, modificaciones que se introducen normalmente después de la transcripción en ARNm de células de mamífero. Además, las moléculas de ARNm pueden modificarse mediante la introducción durante la transcripción de nucleósidos o nucleótidos de origen natural y no natural, como se describe en la Pat. de EE. UU. N.º 8 278 036 (Karikó *et al.*); Sol. de Pat. de EE. UU. N.º 2013/0102034 (Schrum); Sol. de Pat. de EE. UU. N.º 2013/0115272 (deFougerolles *et al.*) y Sol. de Pat. de EE. UU. N.º 2013/0123481 (deFougerolles *et al.*).

«MS» es la espectrometría de masas, una técnica de química analítica que ayuda a identificar la cantidad y el tipo de productos químicos presentes en una muestra midiendo la relación masa-carga y la abundancia de iones en fase gaseosa. Un espectro de masas (espectros en plural) es una gráfica de la señal iónica en función de la relación masa-carga. En el campo biotecnológico se conocen muchas técnicas de MS. Para obtener más información, remítase a *MS Primer* (2015) que se puede adquirir de Waters Corporation, Milford MA EE. UU. Remítase también a Basiri *et al.*, *Bioanalysis* 6(11): 1525-1542 (2014).

El «ARNm» es ARN mensajero, incluido el ARN mensajero eucariota. Los ARNm transmiten información genética del ADN al ribosoma, donde especifican la secuencia de aminoácidos de los productos proteicos de la expresión génica mediante un proceso conocido como transcripción. Estructural e informativamente, el ARNm codifica la información para una proteína en una región codificante. El ARNm eucariota puede comenzar en el extremo 5' con una caperuza de ARNm que se sintetiza enzimáticamente después de que el ARNm ha sido transcrito por una ARN polimerasa *in vitro*. La caperuza del ARNm facilita el inicio de la traducción a la vez que evita el reconocimiento del ARNm como exógeno y protege el ARNm de la degradación mediada por la exonucleasa 5'. La caperuza 5' puede ser un nucleótido de guanina modificado que está unido al extremo 5' de una molécula de ARN utilizando un conector de tipo trifosfato 5'-5'. La caperuza 5' también puede ser un análogo de caperuza 5' tal como una caperuza de diguanosina 5', análogos de caperuzas de tipo tetrafosfato que tienen un resto de metilenobis(fosfonato) (remítase, por ejemplo, a Rydzik AM *et al.*, *Org. Biomol. Chem.* 7(22):4763-76) (2009), análogos de caperuzas de tipo dinucleótido que tienen una modificación de fosforotioato (remítase, por ejemplo, a Kowalska J *et al.*, *RNA* 14(6):1119-1131 (2008)), análogos de caperuzas que tienen una sustitución de azufre por un oxígeno que no actúa como puente (remítase, por ejemplo, a Grudzien-Nogalska E *et al.*, *RNA* 13(10): 1745-1755 (2007)), análogos de tetrafosfato de dinucleósidos N7-bencilados (remítase, por ejemplo, a Grudzien E *et al.*, *RNA* 10(9):1479-1487 (2004)) o análogos de caperuzas anti-inversos (remítase, por ejemplo, a Jemielity J *et al.*, (2003) *RNA* 9(9): 1108-1122 y Stepinski J *et al.*, *RNA* 7(10):1486-1495 (2001)). El análogo de caperuza 5' puede ser una caperuza de diguanosina 5'. Remítase también a los métodos para la adición de la caperuza de *New England Biolabs* (Beverly MA EE. UU.).

Una «etiqueta no radiomarcada» es uno o más marcadores, señales o restos no radioactivos que están unidos, incorporados o asociados con otra entidad (tal como un polinucleótido, como el ARN) que se detecta fácilmente mediante métodos conocidos en la técnica, que incluyen fluorescencia, quimioluminiscencia, actividad enzimática, absorbancia y similares. Las marcas detectables incluyen fluoróforos, cromóforos, enzimas, colorantes, iones metálicos, ligandos tales como biotina, avidina, estrepavidina, digoxigenina y haptenos, puntos cuánticos y similares. Las marcas detectables pueden ubicarse en cualquier posición en los ARN descritos en la presente. Una «sonda etiquetada no radiomarcada» es una sonda polinucleotídica que tiene una etiqueta no radiomarcada unida covalentemente.

Un «nucleósido» o una «base nitrogenada» se refieren a una base (adenina (A), guanina (G), citosina (C), uracilo (U), timina (T) y sus análogos) unida a un carbohidrato, por ejemplo, D-ribosa (en el ARN) o 2'-desoxi-D-ribosa (en el ADN), a través de un enlace N-glucosídico entre el carbono anomérico del carbohidrato y la base nitrogenada. Cuando la base nitrogenada es purina (por ejemplo, A o G), el azúcar de tipo ribosa se une generalmente a la posición N9 del anillo heterocíclico de la purina. Cuando la base nitrogenada es pirimidina (por ejemplo, C, T o U), el azúcar se une normalmente a la posición N1 del anillo heterocíclico. El carbohidrato puede estar sustituido o no sustituido. Los azúcares de tipo ribosa sustituidos incluyen, sin carácter limitante, aquellos en los que uno o más de los átomos de carbono, por ejemplo, el átomo de carbono 2', están sustituidos con uno o más de los siguientes grupos iguales o diferentes: -Cl, -F, -R, -OR, -NR₂ o halógeno, donde cada R es independientemente H, alquilo C₁-C₆ o arilo C₅-C₁₄. Los ejemplos de ribosas incluyen ribosa, 2'-desoxirribosa, 2', 3'-didesoxirribosa, 2'-haloribosa, 2'-fluororibosa, 2'-clororibosa y 2'-alquirribosa, por ejemplo, nucleótidos 2'-O-metilo, 4'-alfa-anoméricos, modificaciones de azúcares bicíclicos con enlaces 2*-4* y 3*-4* y otras modificaciones de azúcares bicíclicos «bloqueados» o de tipo «LNA». Remítase al documento WO 98/22489 (Takeshi Imanishi), WO 98/39352 (Exiqon A/S, Santaris Pharma A/S) y WO 99/14226 (Exiqon A/S).

Un «nucleótido» es un nucleósido en forma fosforilada (un éster de tipo fosfato de un nucleósido), como unidad monomérica o dentro de un polímero polinucleotídico. Un «nucleótido 5'-trifosfato» es un nucleótido con un grupo éster de tipo trifosfato en la posición 5', en ocasiones denominado «NTP» o «dNTP» y «ddNTP» para señalar

particularmente las características estructurales del azúcar de tipo ribosa. El grupo éster de tipo trifosfato puede incluir sustituciones de azufre para los diferentes restos de oxígeno, por ejemplo, α -tio-nucleótido 5'-trifosfatos. Los nucleótidos pueden existir en las formas mono, di o trifosforiladas. Los átomos de carbono de la ribosa presente en los nucleótidos se designan con un carácter de prima (') para distinguirlos de la numeración del esqueleto en las bases. Para consultar un artículo de revisión sobre la química de polinucleótidos y ácidos nucleicos, remítase a Shabarova Z & Bogdanov A, *Advanced Organic Chemistry of Nucleic Acids* (VCH, Nueva York, 1994).

Las expresiones «ácido nucleico», «molécula de ácido nucleico», «polinucleótido» y «oligonucleótido» se refieren indistintamente a polímeros de monómeros nucleotídicos o análogos de estos, tales como el ácido desoxirribonucleico (ADN) y el ácido ribonucleico (ARN) y combinaciones de estos. Los nucleótidos pueden ser de origen genómico, sintético o semisintético. A menos que se indique lo contrario, los términos abarcan estructuras de tipo ácido nucleico con esqueletos sintéticos, así como productos de amplificación. La longitud de estos polímeros (es decir, el número de nucleótidos que contiene) puede variar ampliamente, a menudo dependiendo de su función o uso previsto. Los polinucleótidos pueden ser moléculas lineales, lineales ramificadas o circulares. Los polinucleótidos también tienen contraiones asociados, tales como H^+ , NH_4^+ , trialquilamonio, Mg^+ , Na^+ y similares. Un polinucleótido puede estar constituido en su totalidad por desoxirribonucleótidos, en su totalidad por ribonucleótidos, o mezclas químicas de estos. Los polinucleótidos pueden estar constituidos por análogos de azúcares y bases nitrogenadas internucleotídicos.

Un «oligonucleótido» es un polinucleótido que comprende entre aproximadamente 5 y aproximadamente 150 nucleótidos, por ejemplo, entre aproximadamente 10 y aproximadamente 100 nucleótidos, entre aproximadamente 15 y aproximadamente 75 nucleótidos o entre aproximadamente 15 y aproximadamente 50 nucleótidos. A lo largo de la memoria descriptiva, siempre que un oligonucleótido esté representado por una secuencia de letras (seleccionadas, por ejemplo, entre las cuatro letras de las bases: A (adenosina), C (citidina), G (guanosa) y T (timidina), los nucleótidos se presentan en el orden de 5' a 3' de izquierda a derecha. Una «secuencia de polinucleótido» se refiere a la secuencia de monómeros nucleotídicos a lo largo del polímero. A menos que se indique lo contrario, siempre que se representa una secuencia de polinucleótido, los nucleótidos están en la orientación de 5' a 3' de izquierda a derecha. Los ácidos nucleicos, polinucleótidos y oligonucleótidos pueden comprender bases nucleotídicas estándar o pueden estar sustituidos con análogos de isoformas de nucleótidos, que incluyen, sin carácter limitante, bases de tipo iso-C e iso-G, que pueden hibridarse de forma más o menos permisible que las bases estándar, y que se hibridan preferentemente con bases análogas de isoformas complementarias. Muchas de estas bases de isoformas se describen, por ejemplo, en Benner *et al.*, *Cold Spring Harb. Symp. Quant. Biol.* 52: 53-63 (1987). Los análogos de monómeros nucleotídicos de origen natural incluyen, por ejemplo, 7-desazaadenina, 7-desazaguanina, 7-desaza-8-azaguanina, 7-desaza-8-azaadenina, 7-metilguanina, inosina, nebularina, nitropirrol, nitroindol, 2-aminopurina, 2-amino-6-cloropurina, 2,6-diaminopurina, hipoxantina, pseudouridina (Ψ), pseudocitosina, pseudoisocitosina, 5-propinilcitosina, isocitosina, isoguanina, 7-desazaguanina, 2-azapurina, 2-tiopirimidina, 6-tioguanina, 4-tiotimina, 4-tiouracilo, O-6-metilguanina, N6-metiladenina, O-4-metiltimina, 5,6-dihidrotimina, 5,6-dihidrouracilo, 4-metilindol, pirazolo[3,4-D]pirimidinas, «PPG» (Patentes de EE. UU. N.ºs 6 143 877 (Meyer)) y etenoadenina (Fasman, en *Practical Handbook of Biochemistry and Molecular Biology*, págs. 385-394 (CRC Press, Boca Ratón FL EE. UU., 1989)). Remítase también a los nucleósidos descritos en la Patente de EE. UU. N.º 8 278 036 (Karikó *et al.*); WO 2015/095351 (Novartis AG); Karikó K *et al. Curr. Opin. Drug Disc. Devel.* 10(5): 523-532 (2007); Karikó K *et al. Mol. Therap.* 16(11): 1833-1840 (2008) y Anderson BR *et al., Nucleic Acids Res.* 38(17): 5884-5892 (2010).

«PAGE» es la electroforesis en gel de poli(acrilamida), una técnica utilizada de forma generalizada en el campo biotecnológico para separar macromoléculas biológicas, por ejemplo, ácidos nucleicos, de acuerdo con su movilidad electroforética. En el campo biotecnológico se conocen muchas técnicas de PAGE. Para consultar más información, remítase a Green & Sambrook, *Molecular Cloning: A Laboratory Manual, cuarta edición* (Cold Spring Harbor Press, Plainview, NY, 2012).

Los «cebadores» son secuencias cortas de ácido nucleico. Los cebadores de la reacción en cadena de la polimerasa (PCR) son habitualmente oligonucleótidos de longitud corta (por ejemplo, 8-30 nucleótidos) que se utilizan en las reacciones en cadena de la polimerasa. Los cebadores de PCR y las sondas de hibridación pueden ser desarrollados y producidos con facilidad por los expertos en la técnica, utilizando información sobre la secuencia de la secuencia diana. Remítase a Green & Sambrook, *Molecular Cloning: A Laboratory Manual, cuarta edición* (Cold Spring Harbor Press, Plainview, NY, 2012).

Una «sonda», tal como se utiliza en la presente, es una sonda de oligonucleótidos, una molécula de ácido nucleico cuyo tamaño varía habitualmente desde aproximadamente 50-100 nucleótidos hasta varios cientos de nucleótidos hasta varios miles de nucleótidos de longitud, en incrementos de números enteros. Una sonda puede tener cualquier longitud adecuada para su uso en el método de la invención que se describe en la presente. Tal molécula se utiliza habitualmente para identificar una secuencia específica de ácido nucleico en una muestra hibridándose con la secuencia específica de ácido nucleico en condiciones de hibridación rigurosas. Las condiciones de hibridación son conocidas en el campo biotecnológico. Remítase, por ejemplo, a Green & Sambrook, *Molecular Cloning: A Laboratory Manual, cuarta edición* (Cold Spring Harbor Press, Plainview, NY, 2012).

Un «reactivo que se une a una sonda etiquetada no radiomarcada» o un «reactivo que se une a una etiqueta no radiomarcada» es un reactivo que se une a una etiqueta no radiomarcada que puede unirse a la sonda de oligonucleótidos utilizada en el método de la invención. Los ejemplos de un reactivo que se une a una sonda etiquetada no radiomarcada incluyen (a) estreptavidina, donde la etiqueta no radiomarcada es biotina; (b) avidina, donde la etiqueta no radiomarcada es biotina; (c) un anticuerpo anti-biotina, donde la etiqueta no radiomarcada es biotina; (d) un anticuerpo anti-digoxigenina, donde la etiqueta no radiomarcada es digoxigenina; y (e) un anticuerpo anti-péptido, donde la etiqueta no radiomarcada es un péptido. Los expertos en el campo biotecnológico estarán familiarizados con otros pares de etiquetas no radiomarcadas y reactivos que se unen a la sonda etiquetada no radiomarcada. Remítase, por ejemplo, a Green & Sambrook, *Molecular Cloning: A Laboratory Manual, cuarta edición* (Cold Spring Harbor Press, Plainview, NY, 2012).

El «ARN» es el ácido ribonucleico, un polímero de ribonucleósidos. Cada nucleótido en una molécula de ARN contiene un azúcar de tipo ribosa, con carbonos numerados del 1' al 5'. Hay una base unida a la posición 1'. En general, las bases son adenina (A), citosina (C), guanina (G) o uracilo (U), aunque los expertos en la técnica estarán familiarizados con muchas modificaciones. Por ejemplo, tal como se describe en la presente, un ARN puede contener una o más bases de tipo pseudouracilo (Ψ), de modo que los nucleótidos de pseudouridina sustituyen a los nucleótidos de uridina. Los expertos en la técnica estarán familiarizados con muchas otras modificaciones del ARN, tal como se describe en la presente. El experto en la técnica, tal como un científico de laboratorio, estará familiarizado con los procedimientos para aislar y producir ARN. Remítase a Green & Sambrook, *Molecular Cloning: A Laboratory Manual, cuarta edición* (Cold Spring Harbor Press, Plainview, NY, 2012).

La «RNasa H» es la ribonucleasa H, una familia de endonucleasas no específicas de la secuencia que catalizan la escisión del ARN mediante un mecanismo hidrolítico. La actividad ribonucleasa de la RNasa H escinde el enlace O-P 3' del ARN en un sustrato dúplex de ADN/ARN para producir productos terminados en 3'-hidroxilo y 5'-fosfato.

La «RppH» es la enzima 5' ARN pirofosfohidrolasa.

El «ARNnp» es el ARN nuclear pequeño. Los ARNnp contienen caperuzas 5' únicas. Los ARNnp de clase Sm se encuentran con caperuzas de 5'-trimetilguanosina, mientras que los ARNnp de clase LSM se encuentran con caperuzas de 5'-monometilfosfato. Remítase a Matera, A *et al.*, *Nature Rev. Mol. Cell Biol.* 8(3): 209-220 (marzo de 2007)

Una «sustitución» es una mutación que intercambia una base por otra (es decir, un cambio en una sola «letra química», tal como cambiar una A por una G). Tal sustitución podría (a) cambiar un codón por uno que codifique un aminoácido diferente y provocar un pequeño cambio en la proteína producida; (b) cambiar un codón por uno que codifique el mismo aminoácido y no provocar ningún cambio en la proteína producida («mutaciones silenciosas»); o (c) cambiar un codón codificador de aminoácidos por un único codón de «parada» y generar una proteína incompleta.

Un «sustrato de superficie recubierta» es un sustrato que está recubierto con un reactivo que se une a una sonda etiquetada no radiomarcada. En una realización, el sustrato del sustrato de superficie recubierta son microesferas magnéticas.

El término «sintético» significa producido, preparado o fabricado mediante intervención humana. La síntesis de polinucleótidos o polipéptidos u otras moléculas de la invención puede ser química o enzimática.

El término «diana» se refiere a una molécula de interés. El «ARN diana» es un ARN de interés, que puede ser analizado mediante el método de la invención.

Números de identificación de las secuencias

TABLA 1 Secuencias de nucleótidos		
Polinucleótido (con caperuza 5')	Secuencia	SEQ ID NO.
Cap0	m7GpppGGGAGACGCGUGUUAAAUAACA	1
Trifosfato sin caperuza	pppGGGAGACGCGUGUUAAAUAACA	2
Difosfato sin caperuza	ppGGGAGACGCGUGUUAAAUAACA	3
Monofosfato sin caperuza (el pico esperado para el producto enzimático de RppH con la orientación correcta)	pGGGAGACGCGUGUUAAAUAACA (gggagacgcguguuaaauaaca)	4

la posición 2' de la penúltima guanina para producir m⁷GpppG (Cap1). Remítase a la FIG. 1. Esta estrategia se conoce como adición de la caperuza «postranscripcional». Tcherepanova IY *et al.*, *BMC Mol. Biol.* 9: 90 (2008).

En la segunda manera, se añade la caperuza al ARNm «de forma cotranscripcional» a medida que se transcribe, mediante la adición del análogo de caperuza de dinucleótido m⁷GpppG a la reacción de IVT. La ARN polimerasa incorpora m⁷GpppG en el extremo 5' terminal del ARNm en lugar de GTP. El ARNm producido de esta manera no presenta una adición de la caperuza de un 100% debido a la competencia entre GTP y el análogo de caperuza de dinucleótido durante la síntesis. Los análogos de caperuzas de dinucleótidos se pueden incorporar en la orientación inversa (GpppGm⁷) de modo que m⁷G ya no es la base terminal. Hasta un 50% del ARNm en el que se ha introducido la caperuza «de forma cotranscripcional» puede contener caperuzas inversas. Pasquinelli AE *et al.*, *RNA* 1(9): 957-967 (1995).

Debido a que la orientación de la caperuza es crítica para la traducción del ARNm, se han desarrollado análogos de caperuzas «antiinversos» (ARCA) para garantizar que las caperuzas se incorporan en la orientación correcta. Peng ZH *et al.*, *Org. Lett.* 4 (2):161-164 (2002); Stepinski J *et al.*, *RNA* 7(10): 1486-1495 (2001). Con los dinucleótidos ARCA, el OH de la ribosa en posición 3' en la N7-metilguanosa se metila en los dinucleótidos ARCA para evitar la adición de la caperuza en esa cara del polinucleótido. Aunque el método de adición de la caperuza ARCA garantiza la orientación correcta de la caperuza 5', el ARNm producido «de forma cotranscripcional» produce inherentemente ARNm con un trifosfato en el extremo 5', debido a la competencia entre GTP y el motivo de caperuza durante la transcripción. El trifosfato 5' es inmunogénico. Su presencia debe ser mitigada para que el ARNm no sea citotóxico. Hornung V *et al.*, *Science* 314(5801): 994-997 (2006). Debido a la importancia de la caperuza 5', el hecho de conocer el grado de adición de la caperuza y la orientación de esa caperuza es vital para comprender la eficacia de la traducción del ARNm.

Un método exento de radiomarcas para identificar la caperuza del extremo 5' en un ARN diana

El ARNm de longitud completa es demasiado grande para ser analizado directamente por LC-MS, por lo que el método de la invención utiliza en primer lugar una técnica descrita por primera vez por Lapham & Crothers, *ARN* 2(3): 289-296 (1996), para escindir secuencias cortas del extremo 5' del ARNm en una ubicación específica. Esta técnica utiliza una sonda de escisión de RNasa H etiquetada con biotina que es complementaria respecto al extremo 5' del ARNm (remítase, por ejemplo, a la SEQ ID NO.: 15, que es complementaria respecto al extremo 5' de un ARNm de 2,2 K y un ARNm de 9 K, ambos descritos más adelante). La FIG. 2 presenta una representación esquemática de este flujo de trabajo.

Materiales y métodos En esta cuadragésima primera realización, la sonda de escisión tiene una secuencia de ARN que contiene modificaciones 2' de O-metilo, excepto en el sitio de escisión del extremo 3', que contiene 4-6 nucleótidos de ADN (remítase, por ejemplo, a la SEQ ID NO.: 15). El tramo terminal del ADN dirige a la RNasa H para que escinda el ARNm en una ubicación específica, porque el ARN modificado en la posición 2' con O-metilo no es reconocido por la RNasa H como un sitio de escisión. Inoue H *et al.*, *FEBS Lett.* 215(2): 327-330 (1987).

En un ejemplo, evaluamos el método de la invención utilizando dos ARN diana, que eran ARNm de dos longitudes diferentes, uno de aproximadamente 2200 nucleótidos (ARNm de luciferasa de *Photinus pyralis* (luciérnaga), SEQ ID NO.: 17) y el otro de aproximadamente 9000 nucleótidos. Los primeros 50 nucleótidos del extremo 5' de cada ARNm tenían la misma secuencia, por lo que utilizamos una sonda de escisión de RNasa H para ambos. El ARNm se preparó utilizando procedimientos de transcripción *in vitro* estándar (IVT) y a continuación se añadió la caperuza utilizando el sistema de adición de caperuzas de *Vaccinia* de New England Biolabs, Ipswich MA EE. UU. (Parte # M2080S). Tras la adición de la caperuza, se hizo precipitar el ARNm con LiCl y a continuación se recaptó en agua.

Para el ARNm que contenía Ψ (pseudouridina) o m⁵C (5-metilcitidina), se utilizaron los NTP modificados correspondientes en lugar de U (uridina) o C (citidina) en las reacciones usuales de síntesis de ARNm.

Para consultar ejemplos sobre la incorporación de Ψ (pseudouridina) o m⁵C (5-metilcitidina) en ARNm, remítase a la Patente de EE. UU. N.º 8 278 036 (Karikó *et al.*); WO 2015/095351 (Novartis AG); Karikó K *et al. Curr. Opin. Drug Disc. Devel.* 10(5): 523-532 (2007); Karikó K *et al. Mol. Therap.* 16(11): 1833-1840 (2008) y Anderson BR *et al., Nucleic Acids Res.* 38(17): 5884-5892 (2010). La Pat. de EE. UU. N.º 8 278 036 (Karikó *et al.*) también describe la incorporación de otros nucleótidos en el ARNm, que incluyen m⁵U (5-metiluridina); m⁶A (N⁶-metiladenosina); s²U (2-tiouridina); Um (2'-O-metil-U; 2'-O-metiluridina); m¹A (1-metiladenosina); m²A (2-metiladenosina); Am (2'-O-metiladenosina); ms²m⁶A (2-metiltio-N⁶-metiladenosina); i⁶A (N⁶-isopenteniladenosina); ms²i⁶A (2-metiltio-N⁶-isopenteniladenosina); io⁶A (N⁶-*(cis-hidroxiisopentenil)*adenosina); ms²i⁶A (2-metiltio-N⁶-*(cis-hidroxiisopentenil)*adenosina); g⁶A (N⁶-glicinilcarbamoiladenosina); t⁶A (N⁶-treonilcarbamoiladenosina); ms²t⁶A (2-metiltio-N⁶-treonilcarbamoiladenosina); m⁶t⁶A (N⁶-metil-N⁶-treonilcarbamoiladenosina); hn⁶A (N⁶-hidroxinorvalilcarbamoiladenosina); ms²hn⁶A (2-metiltio-N⁶-hidroxinorvalilcarbamoiladenosina); Ar(p) (2'-O-ribosiladenosina (fosfato)); I (inosina); m¹I (1-metilinosina); m¹Im (1,2'-O-dimetilinosina); m³C (3-metilcitidina); Cm (2'-O-metilcitidina); s²C (2-tiocitidina); ac⁴C (N⁴-acetilcitidina); f⁵C (5-formilcitidina); m⁵ Cm (5,2'-O-dimetilcitidina); ac⁴Cm (N⁴-acetil-2'-O-metilcitidina); k²C (lisidina); m¹G (1-metilguanosa); m²G (N²-metilguanosa); m⁷G (7-metilguanosa); Gm (2'-O-metilguanosa); m² γ G (N²,N²-

dimetilguanosa); m²Gm (N²,2'-O-dimetilguanosa); m²₂Gm (N²,N²,2'-O-trimetilguanosa); Gr(p) (2'-O-ribosilguanosa (fosfato)); yW (wibutosina); o₂yW (peroxiwibutosina); OHyW (hidroxiwibutosina); OHyW* (hidroxiwibutosina submodificada); imG (wiosina); mimG (metilwiosina); Q (queuosina); oQ (epoxiqueuosina); galQ (galactosilqueuosina); manQ (manosilqueuosina); preQ₀ (7-ciano-7-desazaguanosina); preQ₁ (7-aminometil-7-desazaguanosina); G⁺ (arcaeosina); D (dihidrouridina); m⁵Um (5,2'-O-dimetiluridina); s⁴U (4-tiouridina); m⁵s²U (5-metil-2-tiouridina); s²Um (2-tio-2'-O-metiluridina); acp³U (3-(3-amino-3-carboxipropil)uridina); ho⁵U (5-hidroxiuridina); mo⁵U (5-metoxiuridina); cmo⁵U (ácido uridino-5-oxiacético); mcmo⁵U (éster metílico del ácido uridino-5-oxiacético); chm⁵U (5-(carboxihidroximetil)uridina); mchm⁵U (éster metílico de la 5-(carboxihidroximetil)uridina); mcm⁵U (5-metoxycarbonilmetiluridina); mcm⁵Um (5-metoxycarbonilmetil-2'-O-metiluridina); mcm⁵s²U (5-metoxycarbonilmetil-2-tiouridina); nm⁵s²U (5-aminometil-2-tiouridina); mnm⁵U (5-metilaminometiluridina); mnm⁵s²U (5-metilaminometil-2-tiouridina); mnm⁵se²U (5-metilaminometil-2-selenouridina); ncm⁵U (5-carbamoilmetiluridina); ncm⁵Um (5-carbamoilmetil-2'-O-metiluridina); cmnm⁵U (5-carboximetilaminometiluridina); cmnm⁵Um (5-carboximetilaminometil-2'-O-metiluridina); cmnm⁵s²U (5-carboximetilaminometil-2-tiouridina); m⁶ ₂A (N⁶,N⁶-dimetiladenosina); Im (2'-O-metilinosina); m⁴C (N⁴-metilcitidina); m⁴ Cm (N⁴,2'-O-dimetilcitidina); hm⁵C (5-hidroximetilcitidina); m³U (3-metiluridina); cm⁵U (5-carboximetiluridina); m⁶Am (N⁶,2'-O-dimetiladenosina); m⁶ ₂Am (N⁶,N⁶,O-2'-trimetiladenosina); m^{2,7}G (N²,7-dimetilguanosa); m^{2,2,7}G (N²,N²,7-trimetilguanosa); m³Um (3,2'-O-dimetiluridina); m⁵D (5-metildihidrouridina); f⁵Cm (5-formil-2'-O-metilcitidina); m¹Gm (1,2'-O-dimetilguanosa); m¹Am (1,2'-O-dimetiladenosina); tm⁵U (5-taurinometiluridina); tm⁵s²U (5-taurinometil-2-tiouridina); imG-14 (4-desmetilwiosina); imG2 (isowiosina); o ac⁶A (N⁶-acetiladenosina). Cada posibilidad representa una realización separada del método de la invención.

La longitud e integridad tanto del ARNm de 2 K como del ARNm de 9 K se confirmó mediante electroforesis en gel en un analizador de ARN BioRad Experion (Hercules CA EE. UU.).

Anelación del ARNm y de la sonda de escisión: La anelación de la sonda de escisión y el ARNm se realizó en 100-120 µL del tampón de reacción de RNasa H. Una mezcla típica consistía en una sonda de RNasa H 500 pmol (5 µL), tampón de reacción de RNasa H 10X (12 µL) y ARNm 100 pmol (103 µL) para asegurar que el tampón de reacción de RNasa H estuviera en la concentración recomendada de 1X. La sonda de escisión de RNasa H se adquirió de Integrated DNA Technologies (Coralville IA EE. UU.) y se utilizó directamente. La sonda de escisión tenía la siguiente secuencia, de 5' a 3': TTTGTTmAmUmUmAmAmCmAmCmGmCmGmUmCmUmCmCmC/3BioTEG/ (SEQ ID NO.: 15). Las bases con una m que las precede son ARN modificado con O-metilribosa en la posición 2'; las bases sin prefijo en minúscula son ADN.

Para asegurar que todo el ARNm se uniera a la sonda de escisión, se utilizó un exceso de sonda 5:1. La cantidad de 500 pmol es la cantidad máxima de oligo etiquetado con biotina que puede unirse a la microesfera según el fabricante y nuestras propias pruebas han respaldado este número. La mezcla de la sonda de escisión y el ARNm se aneló en un termociclador a 95 °C durante 5 minutos, a continuación se redujo en intervalos de 2 minutos a 65 °C, 55 °C, 40 °C y finalmente a 22 °C.

Unión a microesferas y escisión por parte de RNasa H: Se obtuvieron microesferas magnéticas recubiertas con estreptavidina (Dynabeads MyOne Streptavidin C1, 10 mg/mL, parte # 65002) de Invitrogen (Life Technologies, Grand Island NY, EE. UU.) y se prepararon para su uso de acuerdo con los protocolos del fabricante. En resumen, utilizando un imán para condensar las microesferas, se retiró la solución de almacenamiento y las microesferas se mezclaron con un volumen equivalente de NaOH 0,1 M + NaCl 0,05 M. Se retiró esta solución y las microesferas se suspendieron de nuevo a continuación en un volumen equivalente de NaCl 0,1 M y se guardaron hasta que fueron necesarias.

En un ejemplo, la inmovilización de la sonda de escisión se realizó con 100 µL de las microesferas, con una concentración de 10 mg/mL, que se habían preparado según el procedimiento de Invitrogen. Antes de su uso, el tampón de almacenamiento de NaCl 0,1 M se retiró de las microesferas con un imán o centrifugación a 15 000 x g durante 5 minutos. El ARNm anelado y la solución de la sonda (120 µL) se añadieron a las microesferas de estreptavidina y se incubaron durante 30 minutos a temperatura ambiente con una rotación suave del tubo para asegurar que se uniera toda la sonda marcada con biotina.

A continuación, se añadieron 10 µL de RNasa H (50 unidades), se mezclaron con una pipeta y se incubaron durante 3 horas a 37 °C. La RNasa H y el tampón de reacción de RNasa H se obtuvieron de New England Biolabs, Ipswich MA EE. UU. (Parte # M0297L).

Después de que el extremo 5' del ARNm se escinde, la sonda de escisión etiquetada no radiomarcada (remítase, por ejemplo, a la SEQ ID NO.: 15, que es una sonda etiquetada con biotina) sigue unida al fragmento escindido como un dúplex. Este dúplex que contiene el fragmento 5' se captura en microesferas magnéticas de estreptavidina, para aislar el fragmento escindido del ARNm restante y el tampón de reacción.

Después de tres horas, las microesferas se condensaron magnéticamente y el sobrenadante se retiró y se desechó sin tocar las microesferas. A continuación, las microesferas se lavaron con 100 µL de tampón de lavado (Tris-HCl 5 mM, (pH 7,5), EDTA 0,5 mM, NaCl 1 M) pipeteando hacia arriba y hacia abajo. Después de mezclar, las muestras se colocaron en un imán durante dos minutos y se retiró el sobrenadante. Esta etapa de lavado se repitió tres veces con

tampón de lavado para eliminar la enzima, el ARNm y el tampón de reacción y a continuación se realizaron tres lavados con agua destilada para eliminar el exceso de sal.

Una vez eliminados los componentes de reacción, el dúplex que contiene el fragmento 5' se libera de las microesferas y se analiza con LC-MS. El acoplamiento de la sonda de escisión de RNasa H a una microesfera magnética proporciona una plataforma sencilla para aislar y purificar el producto de escisión de los tampones no compatibles con LC-MS.

Elución : El agua destilada se eliminó de las microesferas y el ARNm unido se liberó de la sonda añadiendo 100 µL de MeOH al 75% que se había calentado hasta 80 °C. Esta mezcla se calentó hasta 80 °C en una placa caliente durante 3 minutos, a continuación se colocó en un imán durante 3 minutos. El sobrenadante que contenía el producto de escisión y la sonda liberada se recolectó a continuación y se secó hasta 10 µL con una centrifuga evaporativa a TA durante ~ 45 min. La muestra se volvió a suspender a continuación en 50 µL de EDTA 100 µM/MeOH al 1% para el análisis de LC-MS.

Análisis de LC-MS: El análisis del fragmento de ARNm 5' escindido se realizó con un Acuity UPLC (Waters, Milford MA EE. UU.) conectado a un QExactive Orbitrap (Thermo Scientific, Grand Island NY EE. UU.). La fase móvil A consistió en hexafluoroisopropanol 200 mM + trietilamina 8,15 mM, pH 7,9 y la fase móvil B era metanol al 100%. Para todos los análisis se utilizó una columna Waters Acuity C18, 2,1 X 50 mm calentada hasta 75 °C con un caudal de 300 µL/min. El perfil del gradiente para la elución comenzó con un 5% de B durante 1 minuto seguido de una rampa lineal hasta un 25% durante 12 minutos. A los doce minutos, se inició un lavado de un minuto con un 90% de B, seguido de un retorno a un 5% de B a los 13 minutos.

Los espectros de masas se obtuvieron en el modo de iones negativos, en un intervalo de barrido de 600-3000 *m/z* con una resolución de 70 000 o 140 000. La temperatura capilar y de la fuente se estableció a 400 °C. Los espectros se analizaron utilizando el software Promass (Novatia, Newtown PA EE. UU.) en el modo de alta resolución.

Resultados para la identificación de los productos de escisión: La masa única de los fragmentos de escisión que tienen una caperuza m7GpppG en comparación con los que contienen una caperuza G difosfato, trifosfato o no metilada (GpppG) permite la identificación de estos fragmentos y una estimación de la eficacia de la adición de la caperuza, como se muestra en la TABLA 2.

TABLA 2

iones de escisión observados y teóricos para diversos productos de escisión de ARNm

	Productos de escisión de la sonda de 5' a 3'	Mono iso p mol	Estado de carga -3*	observado
Cap0	m7GpppGGGAGACG CGUGUUAAAU AACA	7645,0256	2546,9980	2546,9993
trifos caperuza	sin pppGGGAGACG CGUGUUAAAU AACA	7364,9205	2453,9656	2453,9678
difos	ppGGGAGACG CGUGUUAAAU AACA	7284,9547	2427,3103	2427,3105
monofos	pGGGAGACG CGUGUUAAAU AACA	7204,9879	2400,6547	2400,6563
Cap inv	pm7GGGAGACG CGUGUUAAAU AACA	7220,0035	2405,3240	2405,3271
Cap1	m7GpppmGGGAGACG CGUGUUAAAU AACA	7659,0412	2551,6699	2551,6707
Cap1 trifos	pppmGGGAGACG CGUGUUAAAU AACA	7378,9361	2458,6375	no observado
Cap1 difos	ppmGGGAGACG CGUGUUAAAU AACA	7298,9702	2431,9822	no observado
Cap1 monofos	pmGGGAGACG CGUGUUAAAU AACA	7219,0039	2405,3267	2405,3271
Cap G	GpppGGGAGACG CGUGUUAAAU AACA	7630,0021	2542,3261	2542,3219
5mC Cap0	m7GpppGGGAGAmCG AAmCA	mCGUGUUAAAU 7687,0731	2561,0138	2561,0137

5mC trifos sin cap	pppGGGAGAmCG mCGUGUUAAAU AAmCA	7406,9685	2467,9816	2467,9686
5mC difos sin cap	ppGGGAGAmCG mCGUGUUAAAU AAmCA	7327,0022	2441,3262	2441,327
5mC cap G	GpppGGGAGAmCG mCGUGUUAAAU AAmCA	7672,0496	2556,3420	2556,3487
* m7G tiene una carga positiva, por lo que se debe eliminar un hidrógeno adicional para producir una carga negativa.				

Remítase a la TABLA1 para consultar los SEQ ID Nos. correspondientes.

Un método exento de radiomarcas para determinar la orientación del extremo 5' en ARN diana

5 Análisis de la orientación de la caperuza: En esta cuadragésima segunda realización, una vez separada del ARNm, la caperuza puede identificarse según las diferencias en los tiempos de retención de HPLC. Dado que la orientación de la caperuza es importante para la eficacia de la traducción, queríamos comprobar si nuestro sistema de escisión y análisis del extremo 5' del ARNm podría utilizarse también para determinar la orientación de la caperuza.

10 Tal como se ha descrito anteriormente, la síntesis de ARNm con caperuza mediante la adición de un dinucleótido (m7GpppG) en la mezcla de trifosfatos puede dar como resultado una caperuza con dos orientaciones en el ARNm: m7GpppG...ARNm o Gpppm7G...ARNm. El desarrollo de análogos de caperuzas anti-inversas (ARCA) para permitir la orientación correcta de la caperuza se basó en varios métodos para evaluar la orientación de la caperuza. Los métodos para estudiar la orientación de la caperuza han utilizado 5' ARN pirofosfohidrolasa (RppH, New England Biolabs, Ipswich MA EE. UU.), pirofosfatasa ácida del tabaco u otras pirofosfatasas para escindir selectivamente la caperuza. Pasquinelli AE *et al.*, *RNA* 1(9):957-967 (1995); Peng ZH *et al.*, *Org. Lett.* 4(2): 161-164 (2002); Stepinski J *et al.*, *RNA* 7(10): 1486-1495 (2001); Jemielity J *et al.*, *RNA* 9(9): 1108-1122 (2003); Grudzien E *et al.*, *RNA* 10(9): 1479-1487 (2003); y Grudzien-Nogalska E *et al.*, *Methods Enzymol.* 431: 203-227 (2007).

20 La estrategia enzimática consistía en añadir la enzima RppH al fragmento Cap0 escindido por la RNasa H y a continuación detectar la diferencia de masa en el producto para determinar la orientación de la caperuza. Remítase a la FIG. 3. En esta realización, la reacción de la RNasa H se llevó a cabo como se ha descrito originariamente y las microesferas se lavaron para eliminar el tampón de RNasa H, la enzima y el ARNm no unido, pero no se eluyeron. A continuación, se llevó a cabo la reacción de RppH sobre el producto de escisión Cap0 unido.

25 En un ejemplo de una cuadragésima tercera realización de la invención, la RNasa H, el tampón de reacción de RNasa H y la ARN 5'-pirofosfohidrolasa (RppH) con tampón NE 2 se obtuvieron de New England Biolabs (partes # M0297L y M0356S, Ipswich MA EE. UU.). Para el análisis de la orientación de la caperuza, se añadió una etapa al método de la invención, mediante la cual la enzima RppH (10 unidades) junto con su tampón de reacción se añadió a las microesferas lavadas y se hizo reaccionar durante una hora a temperatura ambiente. Después de esta reacción, las microesferas se lavaron tres veces con agua desionizada antes de eluir.

35 Los resultados del análisis de LC-MS para los productos de reacción muestran el fragmento de escisión de la RNasa H esperado para la sonda junto con picos más pequeños correspondientes a los productos de reacción de RppH (remítase a la TABLA 2). Si bien la enzima RppH no escindió completamente todo el sustrato del fragmento de RNasa H disponible como se observa por el pico grande a m/z 2548, se produjo suficiente producto para determinar la orientación de la caperuza. El pico observado a m/z 2401 es el pico esperado para el producto enzimático de RppH con la orientación correcta (pGGGAGACGCGUGUUAAAUACA) (SEQ. ID. NO.: 4) y no se observó ningún producto para la orientación inversa a m/z 2405.

40 Si bien no cabía esperar que se observara la especie de caperuza inversa con el procedimiento de adición de la caperuza «postranscripcional» que utilizamos para producir el ARNm, este resultado demuestra la viabilidad del uso del proceso de escisión con RNasa para examinar la orientación.

45 Un método exento de radiomarcas para determinar la eficacia de la adición de la caperuza

50 Ensayos de enriquecimiento sin caperuza. En una cuadragésima cuarta realización, para determinar la cantidad de ARNm sin caperuza (trifosfato 5') que el ensayo era capaz de detectar, 100 pmol de ARNm con Cap0 se enriquecieron con 0,5, 1, 5, 10 y 25 pmol de ARNm sin caperuza. Se registraron las áreas de los picos para los iones de Cap0 y trifosfato. El área de fondo del trifosfato presente en la muestra de Cap0 sola se restó de las muestras enriquecidas y a continuación las áreas se representaron gráficamente frente al porcentaje de enriquecimiento.

Medición de la eficacia de la adición de la caperuza. Para cuantificar las diferentes especies del extremo 5' y calcular la cantidad de adición de la caperuza en una muestra de ARNm dada, supusimos que la eficacia de la ionización de las diferentes especies de escisión era la misma. Las pruebas con cantidades iguales de muestra con y sin caperuza dieron respuestas similares según el UV (260 nm) y los recuentos de iones, por lo que nuestra suposición fue razonable.

Para evaluar la linealidad y la reproducibilidad del procedimiento, 100 picomoles de ARNm de 2,2 K con caperuza se enriquecieron con porcentajes variables de ARNm sin caperuza. Las muestras enriquecidas se procesaron mediante el procedimiento como muestras separadas, tres para cada nivel porcentual. Se midieron las áreas de los picos del ion de estado de carga más intenso, -8, para las especies con y sin caperuza y se utilizaron para determinar la cantidad de muestra sin caperuza presente. El área del pico de la especie de ARNm con trifosfato sin caperuza presente como fondo en la muestra de Cap0 se restó de las muestras enriquecidas a la hora de determinar la cantidad de ARNm sin caperuza presente.

Resultados de la mezcla de fragmentos de escisión 5' de ARNm de 2,2 K sin caperuza y con Cap1.

La cromatografía de fase inversa separó fácilmente el producto de escisión de tipo trifosfato o sin caperuza del producto de escisión con Cap1. Los espectros de masas por electronebulización del ARN con y sin caperuza consistieron en estados de carga de -9 a -3, siendo el -8 el más intenso. Los iones con estado de carga -3 de la muestra con Cap1 y sin caperuza (trifosfato) fueron aproximadamente equivalentes. Ambos mostraron una pequeña cantidad de aductos de sodio y potasio procedentes del tampón de reacción. Además de los fragmentos de escisión esperados, también había iones correspondientes a la sonda de escisión de biotina e iones que identificaban un segundo sitio de escisión de la RNasa H, como se muestra en la FIG. 2 con la flecha pequeña.

La presencia de la sonda de escisión de biotina fue el resultado de unas condiciones de elución duras que se utilizan para liberar el fragmento de escisión 5' unido. Parte del procedimiento implica secar y suspender de nuevo las muestras, lo que otros han descrito como problemático para los oligonucleótidos según Zhang G *et al.*, *Anal. Chem.* 79(9): 3416-3424 (2007).

El ion difosfato se observó tanto en las muestras con caperuza como en las que no tenían caperuza. Al parecer no se crea debido a las condiciones de LC-MS. La alteración de las temperaturas de la fuente, las temperaturas de la columna y los voltajes de la lente no afectó al ion difosfato observado.

Los resultados mostraron el porcentaje observado de ARNm sin caperuza frente al porcentaje conocido sin caperuza para 100 pmol de ARNm de 2,2 K enriquecido con ARNm con trifosfato sin caperuza. Para los datos de la TABLA 3, se enriquecieron tres muestras individuales con cada porcentaje y se analizaron. El ARNm con Cap0 contenía 0,36% de trifosfato sin caperuza y este nivel de fondo se restó de las muestras enriquecidas.

TABLA 3

Detección de ARNm con trifosfato sin caperuza con el que se enriqueció el ARNm con Cap0

% de enriquecimiento añadido	% de enriquecimiento detectado (Trifos de fondo restado)	% de enriquecimiento	Promedio	Desv. est.	% de CV
25	22,68	25	26,26	3,36	12,79
	26,76	10	10,63	0,13	1,24
	29,33	5	4,89	0,22	4,59
10	10,49	1	0,92	0,03	3,01
	10,67	0,5	0,43	0,04	9,42
	10,74	0	0	0	0
5	4,63				
	4,98				
	5,06				
1	0,91				
	0,95				
	0,90				
0.5	0,46				
	0,39				
	0,46				
0	0,01				
	-0,01				
	0,00				

5 LA TABLA 3 muestra la respuesta lineal y los valores calculados del ensayo desde un 25% hasta un 0,5% de ARNm sin caperuza en 100 pmoles de ARNm con caperuza. El coeficiente de varianza porcentual fue superior a 10 para la serie de enriquecimiento de un 25%.

10 Para las pruebas que se muestran en la TABLA 4, 100 pmol de ARNm de 2,2 K se enriquecieron con varias cantidades de ARNm con trifosfato sin caperuza y se analizaron mediante LC-MS. Se enumeran las principales especies de escisión identificadas en las muestras analizadas, junto con sus porcentajes. Los porcentajes se calcularon a partir de las áreas de los picos para los iones con estado de carga -8 para cada especie individual.

TABLA 4

15 % calculado de especies con caperuza determinado utilizando áreas de picos de iones seleccionadas para especies con y sin caperuza observadas en ARNm sintetizado por IVT

% de especies con caperuza

Información sobre el ARNm	% con caperuza = m7GpppN	% de cap G = GpppN	% de difosfato = ppN	% de trifosfato = pppN	% total sin caperuza
ARNm de 2,2 K-Pseudo U Cap0	95,3	2,59	1,77	0,29	4,7
ARNm de 2,2 K -Pseudo U Cap0	94,9	2,91	1,93	0,22	5,1
ARNm de 2,2 K sin Cap0	98,0	1,32	0,56	0,14	2,0
ARNm de 2,2 K sin Cap0	97,4	1,60	0,77	0,23	2,6
ARNm de 2,2 K <i>psU 5MeC</i> Cap0	87,8	7,44	4,15	0,62	12,2
ARNm de 2,2 K Pseudo U-Cap1	97,5	1,38	0,91	0,25	2,5
ARNm de 9 K sin Cap0	96,5	2,19	1,05	0,25	3,5
ARNm de 9 K sin Cap0	95,4	3,71	0,87	0,00	4,6
ARNm de 9 K sin Cap0	91,7	5,16	2,78	0,36	8,3
ARNm de 9 K sin Cap0	95,1	3,94	0,82	0,12	4,9
ARNm de 9 K sin Cap0	95,9	3,37	0,60	0,17	4,1
	% con caperuza = m7GpppN	tras la adición de la caperuza, la desfosfatasa reaccionó		5' OH = N	
ARNm de 9 K ARCA Cap 0	71,9	0,0	0,0	28,12	28,1

* No se observó Cap0 en la muestra de Cap1.

5 Cada muestra representa una preparación sintética de IVT por separado. El porcentaje total sin caperuza se determina calculando el (área del pico de todas las especies sin caperuza/área del pico de las especies con y sin caperuza)*100. El ARNm al cual se le añadió la caperuza con el procedimiento ARCA se trató con difosfatasa y, por lo tanto, solo tiene 5' OH como la especie sin caperuza detectada.

10 La TABLA 5 muestra los datos sin procesar para la TABLA 4.

TABLA 5

15 Áreas de los picos de cromatogramas iónicos seleccionados para especies con y sin caperuza detectadas en reacciones de IVT

Información sobre el ARNm	G sin metilar de				
	Área del pico de escisión Cap 0 = m7GpppN	Área del pico de escisión de m7G = GpppN (sin caperuza)	Área de difosfato sin caperuza = ppN	Área de trifosfato sin caperuza = pppN	Área total sin caperuza
ARNm de 2,2 K-Pseudo U Cap0	143 866 658	3 913 867	2 663 938	438 836	7 016 641
ARNm de 2,2 K-Pseudo U Cap0	127 157 089	3 893 792	2 590 580	294 181	6 778 553
ARNm de 2,2 K sin Cap0	135 857 464	1 835 325	770 040	187 661	2 793 026
ARNm de 2,2 K sin Cap0	115 067 975	1 889 422	915 586	269 918	3 074 926
ARNm de 2,2 K <i>psU 5MeC</i> Cap0	122 150 875	10 350 263	5 778 955	867 862	16 997 080
ARNm de 9 K sin Cap0	10 042 471	227 905	108 761	26 249	362 915
ARNm de 9 K sin Cap0	6 592 062	256 475	59 804	0	316 279
ARNm de 9 K sin Cap0	14 609 106	822 621	443 386	57 059	1 323 066
ARNm de 9 K sin Cap0	6 206 318	256 904	53 268	7 853	318 025
ARNm de 9 K sin Cap0	7 166 897	252 037	44 698	12 831	309 566
	Área del pico de escisión Cap 1 = m7GpppN	Área de G sin metilar de m7G = GpppN (sin caperuza)	Área de difosfato sin caperuza = ppN	Área de trifosfato sin caperuza = pppN	Área total sin caperuza*
ARNm de 2,2 K Pseudo U-Cap1	151 688 119	2 141 731	1 411 031	382 468	3 935 230
	% con caperuza = m7GpppN	tras la adición de la caperuza, la desfosfatasa reaccionó		5' OH = N	
ARNm de 9 K ARCA Cap 0	27 070 588	0	0	10 592 109	10 592 109

* No se observó Cap0 en la muestra de Cap1.

5 El análisis reveló que todos los ARNm estudiados, independientemente del tamaño, tipo de caperuza o base modificada, contenían la caperuza G no metilada, di- y trifosfatos como las principales especies sin caperuza. Las reacciones de síntesis de IVT individuales fueron en gran medida uniformes dentro de un conjunto dado de condiciones de reacción. Por ejemplo, todos los lotes de ARNm de 2,2 K de origen natural (wt) presentaron perfiles de adición de la caperuza similares, así como todos los lotes de ARNm de 2,2 K de pseudouridina y ARNm de origen natural de 9 K. La eficacia de la adición de la caperuza se calculó dividiendo el área del pico de todas las especies sin caperuza identificadas por el área del pico de las especies sin caperuza más las especies con caperuza. Las eficiencias de la adición de la caperuza calculadas del ARNm de origen natural fueron de un 98 y un 95% para el ARNm de 2,2 K y 9 K, respectivamente.

15 Estos resultados son coherentes con la bibliografía del producto suministrada con el sistema de adición de la caperuza de *Vaccinia* (New England Biolabs, Ipswich MA EE. UU.), que establece que cabe esperar que las eficacias de la adición de la caperuza sean cercanas a un 100%. Sin embargo, no somos conscientes de que se hayan descrito determinaciones de esto en la bibliografía científica. El ARNm al cual se le añadió la caperuza utilizando el sistema ARCA y a continuación se trató con fosfatasa para eliminar todos los fosfatos libres en el extremo 5' presentó una eficacia de la adición de la caperuza significativamente menor. Este resultado concuerda con una eficacia de la adición de la caperuza de un 70% con ARCA observada por Grudzien E *et al.*, *RNA* 10(9): 1479-1487 (2004).

20 En una cuadragésima sexta realización, la TABLA 6 proporciona los datos para producir una curva de calibración. Las áreas de los picos son para el estado de carga -8.

25 **TABLA 6**

Muestra	estado de carga -8 área del pico de escisión de Cap0	Cap0, media	Cap0, desv. est.	Cap0, % de CV	área del pico de trifosfato sin caperuza	sin caperuza, media	sin caperuza, desv. est.	sin caperuza, % de CV	fondo restado
25	218 407 220	197 861 847	26 628 795	13	50 284 860	52 514 453	8 198 941	16	49 528 288
25	167 777 334				45 660 917				44 904 345
25	207 400 988				61 597 582				60 841 010
10	228 597 114	222 773 390	5 796 090	3	24 709 479	24 421 647	459,911	2	23 971 366
10	217 005 336				23 891 231				23 153 118
10	222 717 720				24 664 231				23 926 118
5	205 191 057	200 598 928	5 250 336	3	10 247 095	10 541 001	358 625	3	9 508 982
5	194 874 993				10 435 314				9 697 201
5	201 730 733				10 940 593				10 202 480
1	205 326 924	182 532 252	33 491 166	18	2 616 537	2 427 770	338 856	14	1 878 424
1	198 189 683				2 630 200				1 892 087
1	144 080 149				2 036 572				1 298 459
0,5	223 407 722	209 412 972	27 028 179	13	1 763 775	1 642 331	109 989	7	1 025 662
0,5	226 574 088				1 613 796				875 683
0,5	178 257 106				1 549 421				811 308
0	212 808 157	212 014 762	5 951 111	3	756 572	738 113	16 052	2	18 459
0	205 706 752				727 430				-10 683
0	217529376				730 336				-7 777
sin caperuza, desv. est.					212 052 194	194 739 256	15 144 950	8	
sin caperuza, desv. est.					183 945 935				
sin caperuza, desv. est.					188 219 638				

30 Si bien la curva de calibración no se extendió más allá de un 25%, no hay razón para esperar que la respuesta no permanezca lineal. El tamaño de las áreas de los picos de una mezcla al 50% (como se muestra en la FIG. 2) respalda este razonamiento.

35 Análisis de la síntesis de ARNm por IVT que incorpora nucleósidos modificados. En una cuadragésima séptima realización, el ensayo de la invención se utilizó para analizar ARNm de diferentes longitudes (2,2 K y 9 K), así como ARNm sintetizado con Cap1, Cap0 y con el nucleótido modificado 5-metilcitosina y pseudouridina. Se evaluaron estas dos modificaciones. Otros han descrito que la incorporación de pseudouridina y 5-metilcitosina en ARNm da como resultado una menor inmunogenicidad y estabilidad del ARNm terapéutico, pero la eficacia de la adición de la caperuza de estas incorporaciones no ha sido descrita por Kormann MS *et al.*, *Nat. Biotechnol.* 29(2): 154-157 (2011); Warren L *et al.*, *Cell Stem Cell* 7(5): 618-630 (2010); Karikó K *et al.*, *Immunity* 23 (2):165-175 (2005); o Karikó K *et al.*, *Mol. Ther.* 16(11): 1833-1840 (2008).

En general, parece ser que la incorporación de pseudouridina y/o 5-metilcitidina en el ARNm aumenta la cantidad de ARNm sin caperuza presente, aunque los resultados no son claros para el ARNm de 2,2 K. En comparación con el origen natural, la incorporación de pseudouridina en el ARNm con Cap0 de 2,2 K aumentó la cantidad de especies sin caperuza en una pequeña cantidad; aproximadamente un 2,5%. Sin embargo, la incorporación de pseudouridina en ARNm de 2,2 K que además se metiló en Cap1 no mostró ningún aumento en las especies sin caperuza. Por el contrario, la incorporación de pseudouridina y 5-metilcitidina en el ARNm de 9 K dio como resultado un aumento significativo en el ARNm sin caperuza respecto al caso de origen natural. Se ha demostrado que estos dos nucleósidos aumentan la T_m de los oligonucleótidos, por lo que es posible que su incorporación altere la estructura del ARNm haciendo que el extremo 5' sea menos accesible para la adición de la caperuza. Davis DR *et al.*, *J. Biomol. Struct. Dyn.* 15(6): 1121-1132 (1998).

Eficacia de escisión de la RNasa H, especificidad y recuperación. La examinación de las etapas enzimáticas involucradas en la adición de la caperuza en ARNm de *Vaccinia* proporciona información sobre cómo se crean las especies sin caperuza observadas. Ghosh & Lima, *Wiley Interdiscip. Rev. RNA* 1(1): 152-172 (2010). Remítase también a la FIG. 1.

Utilizando los porcentajes o las áreas de los picos como guía para la cantidad de cada especie presente, la enzima inicial en el proceso (trifosfatasa) funciona bien, juzgando por la presencia de difosfato en el extremo 5' de la secuencia escindida. La siguiente etapa en el proceso, la adición de GTP al difosfato por parte de la guaniltransferasa, también funciona, según la observación de la secuencia G con caperuza, pero la presencia de difosfato remanente es una señal de que la guaniltransferasa podría ser más eficaz. De forma similar, la observación de especies de G con caperuza no metiladas indica que la metilación de la guanina terminal mediante N7-metiltransferasa es ineficaz, con lo que deja especies de caperuza G no metiladas. Las observaciones en la muestra de Cap1 son interesantes; se observó caperuza G no metilada, pero no Cap0. Por consiguiente, la 2' O-ribosa metiltransferasa necesaria para producir Cap1 parece ser más eficaz en comparación con la N7-metiltransferasa.

El sustrato de ARNm de 48 mer utilizado para determinar la eficacia de la reacción y la recuperación tenía la siguiente secuencia de ARN de 5' a 3':
 rGrGrGrArGrArCrGrCrGrUrGrUrUrArArArUrArArCrArArArUrCrUrCrArArCrArCrArCrArUrArUrArCrArArArCrA
 (SEQ ID NO.: 16). La sonda utilizada en estas pruebas contenía seis nucleótidos de ADN (remítase a la SEQ ID NO.: 15). La sonda provocó la escisión en dos posiciones adyacentes para formar dos productos que diferían en una adenosina, tal como se muestra con las flechas en la FIG. 2.

Comparando las áreas de los picos de los iones, el producto de escisión más largo (CAA) representó de un 7 a un 8% del producto de escisión más corto (CA).

En una cuadragésima octava realización, la eficacia de la reacción de escisión de la RNasa H y la recuperación del producto de escisión a partir de la microesfera se evaluaron con un ARN de 48 mer sintetizado con la misma secuencia que el extremo 5' del ARNm. El ARN de 48 mer se sometió en una cantidad de 200 pmol al procedimiento de escisión y aislamiento sin añadir RNasa H y el área del pico se comparó con la preextracción de 48 mer sola. Se observó que la recuperación era de un 94%, lo cual demostró que se perdía una pequeña cantidad del oligo una vez que se unía a la microesfera. A continuación, el procedimiento se llevó a cabo con la enzima RNasa H y las áreas de los picos de escisión se compararon entre el ARN de 48 mer y el ARNm de 2,2 kb (ambos en una cantidad de 200 pmol). En este caso, el pico de escisión del ARNm fue de un 58% de 48 mer, lo cual sugiere que es o bien la unión inicial de la sonda de escisión al ARNm y/o la reacción de escisión de la RNasa H la causa de la recuperación más baja en comparación con el 48 mer.

Este resultado no es sorprendente, teniendo en cuenta el tamaño del ARNm y los elementos estructurales que podrían interferir con la unión del ARNm a la sonda de escisión de biotina.

Un método exento de radiomarcas para detectar impurezas en la reacción de adición de la caperuza en un preparado de ARN

En una cuadragésima novena realización, el método de la invención mostró buena sensibilidad para detectar impurezas en la reacción de adición de la caperuza. El ARNm con trifosfato sin caperuza con el que se enriquecieron 100 picomoles de ARNm con caperuza pudo ser detectado a lo largo del intervalo evaluado de un 0,5 a un 25% con una respuesta lineal. Se determinó que la eficacia de la adición de la caperuza para varios preparados de ARNm con caperuza de *Vaccinia* estuvo entre un 88 y un 98% dependiendo del tipo de modificación y la longitud del ARNm.

La descripción detallada proporcionada en la presente es para ilustrar la invención, pero no para limitar su alcance. Otras variantes de la invención serán muy obvias para un experto en el campo biotecnológico y están incluidas en las reivindicaciones adjuntas.

REIVINDICACIONES

1. Un método exento de radiomarcas para identificar una caperuza del extremo 5' en un ácido ribonucleico (ARN) diana, que comprende los pasos de:
- 5 (a) hibridar una sonda etiquetada no radiomarcada con el ARN diana, donde la secuencia de nucleótidos de la sonda etiquetada no radiomarcada es complementaria respecto al extremo 5' del ARN diana, con lo que se forma así un polinucleótido dúplex;
- 10 (b) tratar el polinucleótido dúplex con RNasa H, con lo que se escinde así el extremo 5' del ARN diana y se forma un polinucleótido dúplex que contiene el extremo 5' del ARN diana;
- (c) aislar el polinucleótido dúplex, utilizando un sustrato de superficie recubierta que está recubierto con un reactivo que se une a la sonda etiquetada no radiomarcada;
- 15 (d) eliminar el polinucleótido dúplex del sustrato de superficie recubierta;
- (e) desnaturalizar el polinucleótido dúplex, con lo que se produce así un fragmento monocatenario del extremo 5' del ARN diana y la sonda etiquetada no radiomarcada;
- 20 (f) aislar el fragmento monocatenario del extremo 5' del ARN diana;
- (g) analizar el fragmento monocatenario del extremo 5' del ARN diana mediante cromatografía líquida/espectrometría de masas (LC-MS); y
- 25 (h) identificar la caperuza del extremo 5'.
2. El método de la reivindicación 1, donde el ARN diana es un ARN mensajero eucariota.
- 30 3. El método de la reivindicación 1, donde el sustrato de superficie recubierta comprende microesferas magnéticas.
4. El método de la reivindicación 1, donde el reactivo en el sustrato de superficie recubierta que se une a la sonda etiquetada no radiomarcada se selecciona del grupo constituido por:
- 35 (a) estreptavidina, donde la etiqueta no radiomarcada es biotina;
- (b) avidina, donde la etiqueta no radiomarcada es biotina;
- 40 (c) un anticuerpo anti-biotina, donde la etiqueta no radiomarcada es biotina;
- (d) un anticuerpo anti-digoxigenina, donde la etiqueta no radiomarcada es digoxigenina; y
- (e) un anticuerpo anti-péptido, donde la etiqueta no radiomarcada es un péptido.
- 45 5. Un método exento de radiomarcas para determinar la orientación del extremo 5' en ARN diana, que comprende los pasos de:
- (a) hibridar una sonda etiquetada no radiomarcada con el ARN diana para formar un polinucleótido dúplex, donde la secuencia de nucleótidos de la sonda etiquetada no radiomarcada es complementaria respecto al extremo 5' del ARN diana;
- 50 (b) tratar el polinucleótido dúplex con RNasa H, para escindir el extremo 5' del ARN y formar un polinucleótido dúplex;
- (c) aislar el polinucleótido dúplex que contiene el extremo 5' del ARN diana, utilizando microesferas magnéticas recubiertas con un reactivo que se une a la sonda etiquetada no radiomarcada;
- 55 (d) tratar el polinucleótido dúplex con 5' ARN pirofosfohidrolasa (RppH);
- (e) eliminar el polinucleótido dúplex de las microesferas magnéticas;
- 60 (f) desnaturalizar el polinucleótido dúplex, para producir un fragmento monocatenario del extremo 5' del ARN diana y la sonda etiquetada no radiomarcada;
- (g) aislar el fragmento monocatenario del extremo 5' del ARN diana;
- 65

(h) analizar el fragmento monocatenario del extremo 5' del ARN mediante cromatografía líquida/espectrometría de masas (LC-MS); y

(i) detectar la diferencia de masa en el producto para determinar la orientación de la caperuza 5'.

5

6. El método de la reivindicación 5, donde el ARN diana se sintetiza *in vitro*.

7. El método de la reivindicación 6, donde el ARN sintetizado *in vitro* contiene un nucleótido seleccionado del grupo constituido por Ψ (pseudouridina); m⁵C (5-metilcitidina); m⁵U (5-metiluridina); m⁶A (N⁶-metiladenosina); s²U (2-tiouridina); Um (2'-O-metil-U; 2'-O-metiluridina); m¹A (1-metiladenosina); m²A (2-metiladenosina); Am (2'-O-metiladenosina); ms² m⁶A (2-metiltio-N⁶-metiladenosina); i⁶A (N⁶-isopenteniladenosina); ms²i⁶A (2-metiltio-N⁶-isopenteniladenosina); io⁶A (N⁶-(cis-hidroxiisopentenil)adenosina); ms²i⁶A (2-metiltio-N⁶-(cis-hidroxiisopentenil)adenosina); g⁶A (N⁶-glicinilcarbamoiladenosina); t⁶A (N⁶-treonilcarbamoiladenosina); ms²t⁶A (2-metiltio-N⁶-treonilcarbamoiladenosina); m⁶t⁶A (N⁶-metil-N⁶-treonilcarbamoiladenosina); hn⁶A (N⁶-hidroxinorvalilcarbamoiladenosina); ms²hn⁶A (2-metiltio-N⁶-hidroxinorvalilcarbamoiladenosina); Ar(p) (2'-O-ribosiladenosina (fosfato)); I (inosina); m¹I (1-metilinosina); m¹Im (1,2'-O-dimetilinosina); m³C (3-metilcitidina); Cm (2'-O-metilcitidina); s²C (2-tiocitidina); ac⁴C (N⁴-acetilcitidina); f⁶C (5-formilcitidina); m⁵ Cm (5,2'-O-dimetilcitidina); ac⁴Cm (N⁴-acetil-2'-O-metilcitidina); k²C (lisidina); m¹G (1-metilguanosina); m²G (N²-metilguanosina); m⁷G (7-metilguanosina); Gm (2'-O-metilguanosina); m² zG (N²,N²-dimetilguanosina); m²Gm (N²,2'-O-dimetilguanosina); m²zGm (N²,N²,2'-O-trimetilguanosina); Gr(p) (2'-O-ribosilguanosina (fosfato)); yW (wibutosina); o₂yW (peroxiwibutosina); OHyW (hidroxiwibutosina); OHyW* (hidroxiwibutosina submodificada); imG (wiosina); mimG (metilwiosina); Q (queuosina); oQ (epoxiqueuosina); galQ (galactosilqueuosina); manQ (manosilqueuosina); preQ₀ (7-ciano-7-desazaguanosina); preQ₁ (7-aminometil-7-desazaguanosina); G⁺ (arcaeosina); D (dihidrouridina); m⁵Um (5,2'-O-dimetiluridina); s⁴U (4-tiouridina); m⁵s²U (5-metil-2-tiouridina); s²Um (2-tio-2'-O-metiluridina); acp³U (3-(3-amino-3-carboxipropil)uridina); ho⁵U (5-hidroxiuridina); mo⁵U (5-metoxiuridina); cmo⁵U (ácido uridino-5-oxiacético); mcmo⁵U (éster metílico del ácido uridino-5-oxiacético); chm⁵U (5-(carboxihidroximetil)uridina); mchm⁵U (éster metílico de la 5-(carboxihidroximetil)uridina); mcm⁵U (5-metoxicarbonilmetiluridina); mcm⁵Um (5-metoxicarbonilmetil-2'-O-metiluridina); mcm⁵s²U (5-metoxicarbonilmetil-2-tiouridina); nm⁵s²U (5-aminometil-2-tiouridina); mnm⁵U (5-metilaminometiluridina); mnm⁵s²U (5-metilaminometil-2-tiouridina); mnm⁵se²U (5-metilaminometil-2-selenouridina); ncm⁵U (5-carbamoilmetiluridina); ncm⁵Um (5-carbamoilmetil-2'-O-metiluridina); cmnm⁵U (5-carboximetilaminometiluridina); cmnm⁵Um (5-carboximetilaminometil-2'-O-metiluridina); cmnm⁵s²U (5-carboximetilaminometil-2-tiouridina); m⁶ zA (N⁶,N⁶-dimetiladenosina); Im (2'-O-metilinosina); m⁴C (N⁴-metilcitidina); m⁴ Cm (N⁴,2'-O-dimetilcitidina); hm⁵C (5-hidroximetilcitidina); m³U (3-metiluridina); cm⁵U (5-carboximetiluridina); m⁶Am (N⁶,2'-O-dimetiladenosina); m⁶ zAm (N⁶,N⁶,O-2'-trimetiladenosina); m^{2,7}G (N²,7-dimetilguanosina); m^{2,2,7}G (N²,N²,7-trimetilguanosina); m³Um (3,2'-O-dimetiluridina); m⁵D (5-metildihidrouridina); f⁶Cm (5-formil-2'-O-metilcitidina); m¹Gm (1,2'-O-dimetilguanosina); m¹Am (1,2'-O-dimetiladenosina); tm⁵U (5-taurinometiluridina); tm⁵s²U (5-taurinometil-2-tiouridina); imG-14 (4-desmetilwiosina); imG2 (isowiosina); y ac⁶A (N⁶-acetiladenosina).

8. El método de la reivindicación 5, donde el ARN sintetizado *in vitro* comprende un nucleótido seleccionado del grupo constituido por Ψ (pseudouridina) y m⁵C (5-metilcitidina).

9. Un método exento de radiomarcas para determinar la eficacia de la adición de la caperuza, que comprende los pasos de:

(a) proporcionar una muestra de ARN que comprende ARN con caperuza y ARN sin caperuza;

(b) hibridar una sonda etiquetada no radiomarcada con el ARN diana para formar un polinucleótido dúplex, donde la secuencia de nucleótidos de la sonda etiquetada no radiomarcada es complementaria respecto al extremo 5' del ARN diana;

(c) tratar el polinucleótido dúplex con RNasa H, para escindir el extremo 5' del ARN y formar un polinucleótido dúplex;

(d) aislar el polinucleótido dúplex que contiene el extremo 5' del ARN diana, utilizando microesferas magnéticas recubiertas con un reactivo que se une a la sonda etiquetada no radiomarcada;

(e) eliminar el polinucleótido dúplex de las microesferas magnéticas;

(f) desnaturalizar el polinucleótido dúplex, para producir un fragmento monocatenario del extremo 5' del ARN diana y la sonda etiquetada no radiomarcada;

(g) aislar el fragmento monocatenario del extremo 5' del ARN diana;

(h) analizar el fragmento monocatenario del extremo 5' del ARN mediante cromatografía líquida/espectrometría de masas (LC-MS); y

(i) determinar la cantidad relativa de los fragmentos monocatenarios con caperuza y sin caperuza, con lo que se cuantifica así la eficacia de la adición de la caperuza en el ARNm.

- 5 10. El método de la reivindicación 9, donde el ARN diana es una mezcla de reacción de transcripción *in vitro* (IVT).
11. El método de la reivindicación 10, donde la reacción de IVT es un preparado de ARNm con caperuza de *Vaccinia*.
- 10 12. El método de la reivindicación 9, donde el ARNm con trifosfato sin caperuza se puede detectar en un intervalo evaluado de un 0,1 a un 90% con una respuesta lineal.
13. El método de la reivindicación 9, donde el ARNm con trifosfato sin caperuza se puede detectar en un intervalo evaluado de un 0,5 a un 25% con una respuesta lineal.
- 15 14. Un método exento de radiomarcas para detectar impurezas en la reacción de adición de la caperuza en un preparado de ARN, que comprende los pasos de:
- (a) hibridar una sonda etiquetada no radiomarcada con el ARN diana, donde la sonda etiquetada no radiomarcada es complementaria respecto al extremo 5' del ARN diana;
- 20 (b) tratar el ARN hibridado con RNasa H para escindir el extremo 5' del ARN;
- (c) aislar la secuencia del extremo 5' escindida, utilizando microesferas magnéticas recubiertas con un reactivo que se une a la sonda etiquetada no radiomarcada;
- 25 (d) analizar la secuencia del extremo 5' escindida mediante LC-MS, y
- (e) detectar una impureza de la reacción de adición de la caperuza en un preparado de ARN.
15. El método de la reivindicación 14, donde el ARN diana se sintetiza *in vitro*.

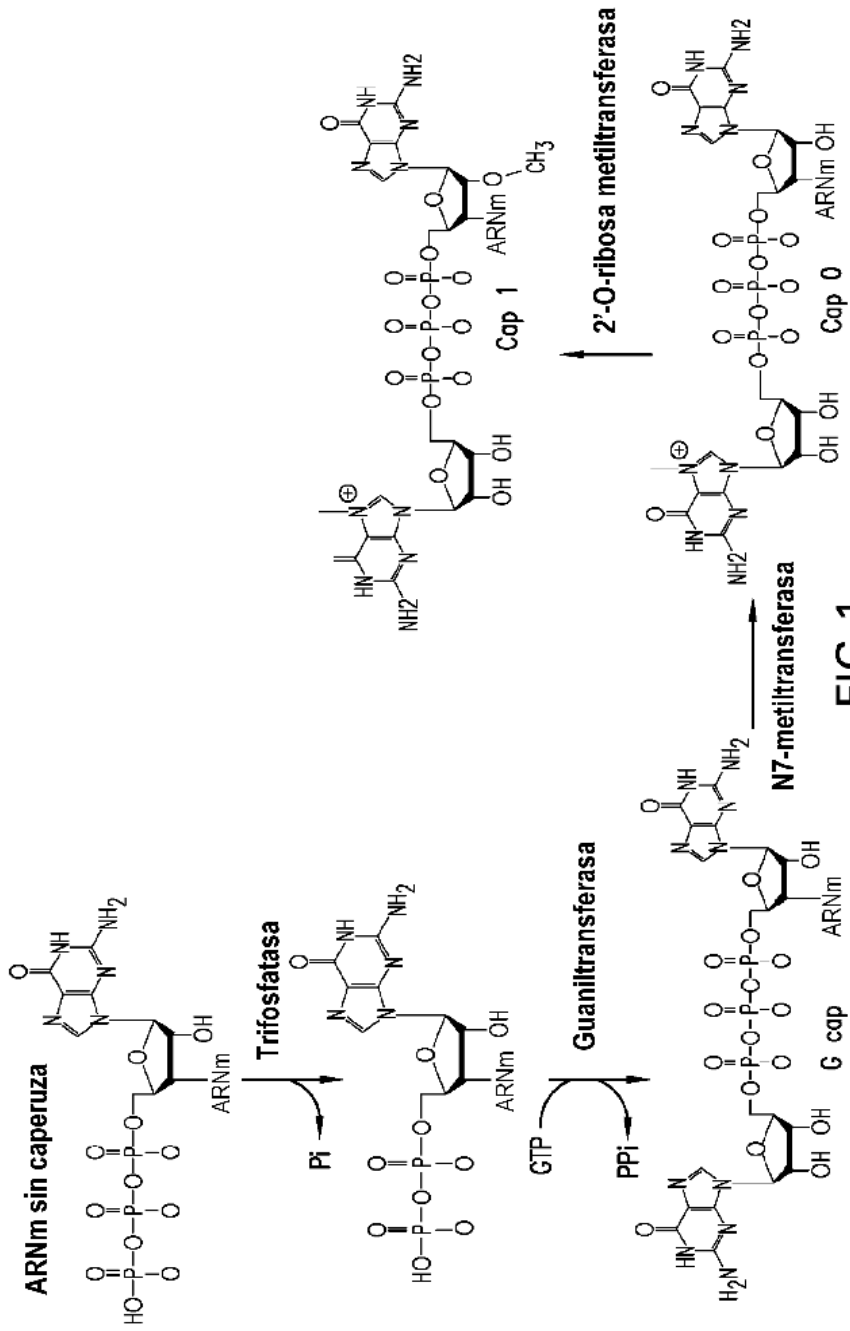


FIG.1

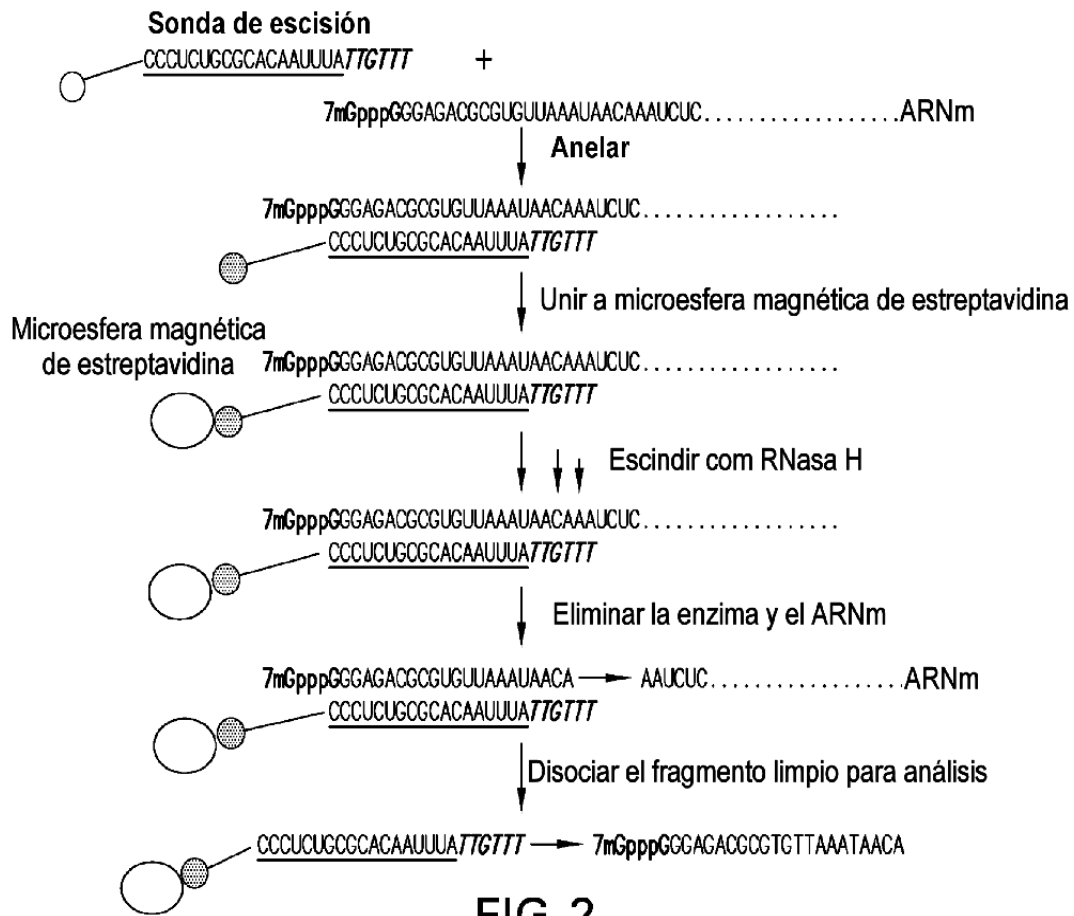


FIG.2

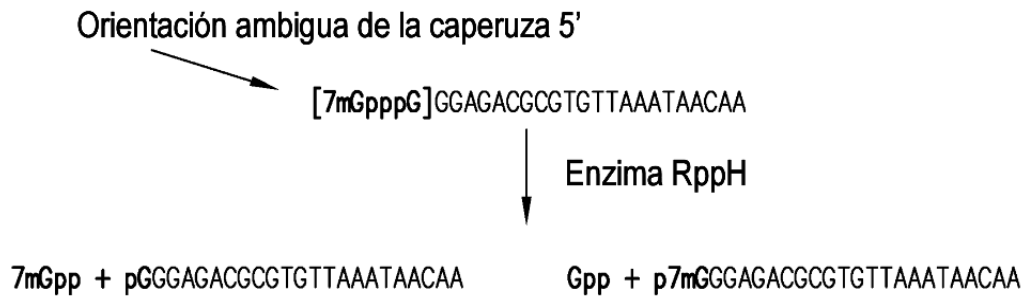


FIG.3